



**HAL**  
open science

# Étude phonétique de la production du Human Beatbox : approche articulatoire, aérodynamique et acoustique

Alexis Dehais Underdown

► **To cite this version:**

Alexis Dehais Underdown. Étude phonétique de la production du Human Beatbox : approche articulatoire, aérodynamique et acoustique. Linguistique. Université de la Sorbonne nouvelle - Paris III, 2023. Français. NNT : 2023PA030069 . tel-04585701

**HAL Id: tel-04585701**

**<https://theses.hal.science/tel-04585701v1>**

Submitted on 23 May 2024

**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 destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Étude phonétique de la production du Human  
Beatbox :  
Approche articulatoire, aérodynamique et  
acoustique**

*Alexis Dehais-Underdown*

**Université Sorbonne Nouvelle  
ED 622 Sciences du langage  
Laboratoire de Phonétique et Phonologie (CNRS / USN)**

*Thèse présentée pour l'obtention d'un doctorat en Phonétique, Phonologie et Sciences  
de la Parole*

**SOUS LA DIRECTION DU PROFESSEUR DIDIER DEMOLIN ET  
DU DOCTEUR LISE CREVIER-BUCHMAN**

*Thèse soutenue le 8 septembre 2023  
Composition du jury :*

Co-Directeur	Didier Demolin (Professeur Émérite, Université Sorbonne Nouvelle)
Co-Directrice	Lise Crevier-Buchman (Chargée de Recherche CNRS, Hôpital Foch, Université Sorbonne Nouvelle)
Rapporteuse	Alice Turk (Professeure, University of Edinburgh)
Rapporteur	Yves Laprie (Directeur de recherche CNRS, Loria, Nancy)
Examinateur	Bryan Gick (Professeur, University of British Columbia)
Examinateur	John Kingston (Professeur, University of Massachusetts)
Examinatrice	Claire Pillot-Loiseau (Professeure des universités, Université Sorbonne Nouvelle)



**Titre :** Étude phonétique de la production du Human Beatbox : approche articulatoire, aérodynamique et acoustique

**Résumé :**

Avant d'être un objet d'étude scientifique, le Human Beatbox (HBB) est avant tout une pratique musicale. Les beatboxeurs et les beatboxeuses utilisent leur conduit vocal pour produire des sons musicaux : sons de batterie, instruments à vent ou à corde, musique électronique ... Comment est produit le Human Beatbox ? Quelles sont les capacités du conduit vocal humain ? Existe-t-il un lien entre la production du Beatbox et de la parole ? Cette thèse tente d'apporter des éléments de discussion sur ces questions. À partir de données articulatoires, aérodynamiques et acoustiques, nous proposons une analyse de la production du Human Beatbox à différentes vitesses (90, 120, 150 battements par minute). Nos résultats montrent que les sujets utilisent des mécanismes de production similaires aux locuteurs des langues. Toutefois, ils combinent un plus grand nombre de mécanismes différents. Le Human Beatbox possède les propriétés d'un système combinatoire discret (Proctor et al., 2013). Les résultats sur les effets du tempo suggèrent que les stratégies individuelles de réorganisation temporelle des gestes sont similaires aux stratégies proposées par Byrd et Tan (1996). En effet, les sujets manipulent la durée des gestes et/ou des intervalles séparant les gestes. On a observé chez 2 sujets, des contraintes de coordination entre les gestes d'initiation et d'articulation. Le Human Beatbox jette un regard différent sur la phonétique en permettant d'actualiser son cadre théorique et de passer d'une approche purement linguistique à une approche anthropophonique (Catford, 1977 ; Lindblom, 1990) des phénomènes phonétiques.

Mots-clés : Human Beatbox, Phonétique, Articulation, Aérodynamique, Acoustique, Tempo

**Title :** A phonetic study of Human Beatboxing production : an articulatory, aerodynamic and acoustic approach


**Abstract :**

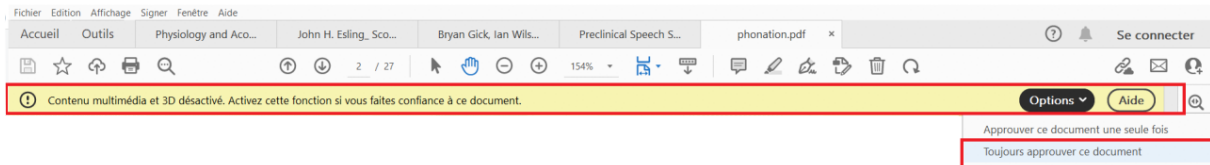
Before being a subject of scientific study, Human Beatboxing (HBB) is above all a musical practice. Beatboxers use their vocal tract to produce musical sounds : drum sounds, wind or string instruments, electronic music, and more. How is beatboxing produced? What are the capacities of the human vocal tract? Is there a connection between beatboxing and speech production? This thesis attempts to provide elements for discussion on these questions. Based on articulatory, aerodynamic, and acoustic data, we present an analysis of beatboxing production at different tempo (90, 120, 150 beats per minute). Our results show that subjects use similar production mechanisms to those found in linguistic systems. However, they combine a greater number of different mechanisms. Beatboxing possesses the properties of a discrete combinatorial system (Proctor et al., 2013). The results on tempo effects suggest that individual strategies for temporal reorganization of gestures are similar to the strategies proposed by Byrd et Tan (1996). Indeed, subjects modified the duration of gestures and the intervals between gestures. Coordination constraints between initiation and articulation gestures were observed for two subjects. Beatboxing offers a different perspective on phonetics, allowing for the update of its theoretical framework and a shift from a purely linguistic approach to an anthropophonic approach (Catford, 1977 ; Lindblom, 1990) of phonetic phenomena.

**Keywords :** Human Beatboxing, Phonetics, Articulation, Aerodynamics, Acoustics, Tempo

# Avertissement aux lecteurs et lectrices

## Contenu multimédia

Cette thèse intègre des fichiers audios afin de pouvoir entendre les sons de notre corpus. Les fichiers audio sont incrustés dans le PDF. Le symbole  indique qu'un fichier audio est intégré dans le PDF. Pour le lire, il suffit de cliquer sur jouer et d'approuver le document sur Adobe :



Autorisez le document à partir de l'exemple suivant :

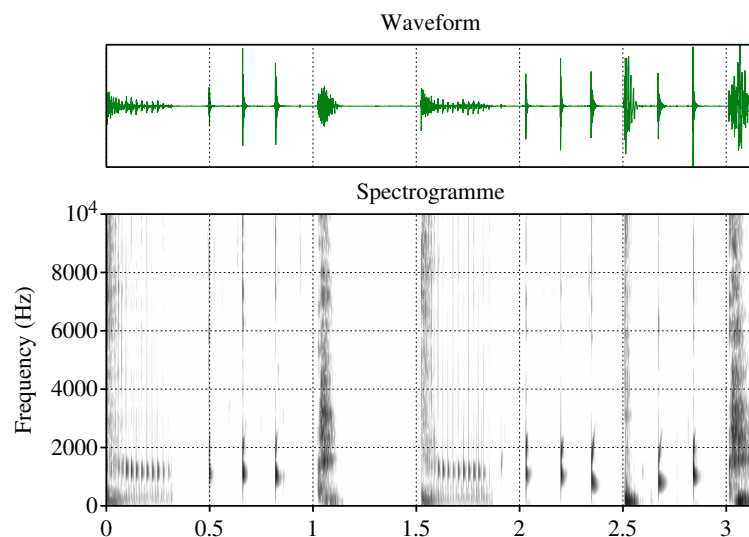
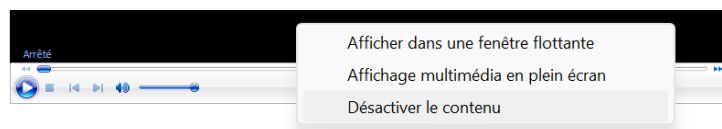


FIGURE : Description de la figure



Un lecteur va s'ouvrir. En changeant de page le lecteur disparaît. Pour le retirer manuellement : *clic droit > désactiver le contenu* :



## Langage inclusif

Cette thèse utilise le langage inclusif. Afin d'éviter l'utilisation excessive de points médians, j'ai fait le choix de rédupliquer les formes masculines et féminines et d'utiliser l'accord de proximité : *Les beatboxeurs et les beatboxeuses sont allées jusqu'au bout de l'expérience* ou

*Les beatboxeuses et les beatboxeurs sont allés jusqu'au bout de l'expérience.* De plus, nous ne ferons pas d'accord des participe passés avec le verbe avoir.  
Bonne lecture et bonne écoute !

# Remerciements

À mes grands mères Gisèle Dehais et  
Anne-Marie Underdown

---

Voilà 5 ans que mon aventure de jeune chercheur a débuté. J'ai finalement réussi à terminer ma thèse avant qu'elle n'en termine avec moi. J'ai à deux reprises failli arrêter ma thèse. Une première fois pendant le Covid après la perte de mes grand-mères et presque 2ans de retard dans l'acquisition des données. Une deuxième fois lorsque la précarité m'a frappé de plein fouet et que je me suis retrouvé sans revenus pendant 5 mois. Sans l'aide de mes co-directeurs, cette thèse n'aurait pas abouti.

Mes premiers remerciements vont donc à Lise Crevier-Buchman et Didier Demolin. Merci Lise pour les connaissances sur la physiologie et le fonctionnement du larynx qui est loin d'être un problème trivial. Merci à Didier pour les connaissances sur l'aérodynamique et les phénomènes phonétiques. Tu es une encyclopédie vivante. Maintenant profite de la retraite et termine ta pile de *unfished papers*. Merci à vous deux pour votre exigence et votre intégrité scientifique et merci à vous deux d'avoir soutenu mon projet de recherche dès 2017 avec une grande ouverture d'esprit. Ensuite, je tiens à remercier Paul Vignes, beatboxeur professionnel. Merci à toi Polo pour ton temps et tes connaissances. Merci à toi d'avoir participé à l'élaboration d'un protocole expérimental fonctionnel et adapté à l'étude phonétique du Human Beatbox.

Un énorme merci à ma famille qui m'a toujours soutenu moralement mais aussi financièrement lorsque ma situation est devenue très précaire. Une pensée à toi Gisèle, toi qui n'a pas voulu finir ta vie « au cul des vaches » et qui a quitté la Bretagne pour offrir un avenir meilleur à tes enfants et petits enfants. Une pensée à toi Anne-Marie qui a quitté le Berry pour offrir un avenir meilleur à tes filles et tes petits enfants. Reposez en paix. Un grand merci à Roland Trouville qui depuis le master m'aide sur les questions de programmation informatique. Un énorme merci à la *team coquille* : Arielle, Solène, Clémence, Amélie, Michaëla et Vincent. Merci à vous d'avoir relu attentivement mon manuscrit. Un grand merci aux doctorant-es du LPP pour les discussions scientifiques, pratiques, théoriques, méthodologiques mais aussi les discussions humaines sur les tracasseries des jeunes chercheurs et des jeunes chercheuses : syndrome de l'imposteur, précarité, perspectives d'avenir etc. Désolé pour les folies, les craquages et les fou-rires ... on décompresse comme on peut ! Un grand merci aux camarades de lutte, aux camarades de la CGT pour votre solidarité, votre humanisme et votre espoir pour une université et une recherche ouverte et émancipatrice plutôt qu'une université pratique le tri social et la casse du service public du supérieur et de la recherche.

Ce travail a reçu le soutien de plusieurs institutions françaises. Je tiens à remercier chaleureusement le Laboratoire de Phonétique et Phonologie, l'Université Sorbonne Nouvelle et le LABEX EFL (Empirical Foundation of Linguistics, ANR-10-LABX-0083) pour le soutien financier. Je remercie aussi l'Hôpital Foch pour le soutien administratif, humain et financier dans le dépôt d'une demande d'autorisation d'enregistrements physiologiques auprès d'un Comité de Protection des Personnes (CPP Île de France IV ID-RCB : 2020-A00246-33). Je remercie aussi le Centre Hospitalier de Recherche Universitaire de Nancy, le Centre d'Investigation Clinique – Innovation Technologique de Nancy (CIC-IT Nancy) et le centre d'Imagerie Adaptative Diagnostique et Interventionnelles (IADI). J'adresse mes remerciements à Pierre-André Vuissoz, Docteur et Ingénieur de Recherche (IADI), Karyna Isaieva post-doctorante (IADI), Marc Fauvel Ingénieur (CIC-IT Nancy), Jacques Felblinger, Professeur, Investigateur Principal et Directeur de l'IADI et du CIC-IT, ainsi qu'à Yves Laprie du Laboratoire Lorrain de Recherche

en Informatique et ses Applications (LORIA). La session d'acquisition IRM dépend du protocole "METHODO" (ClinicalTrials.gov Identifier : NCT02887053, RCB-ID n°CPP EST-III, 08.10.01) et a reçu des financements des fonds européens CPER "IT2MP", "LCHN" et FEDER.



# Table des matières

<b>Introduction</b>	<b>2</b>
<b>1 Cadre Théorique</b>	<b>5</b>
1.1 Anatomie du système de production vocale . . . . .	5
1.1.1 Système respiratoire . . . . .	6
1.1.2 Système laryngé . . . . .	8
1.1.3 Système oral . . . . .	12
1.2 Mécanismes de production des sons . . . . .	16
1.2.1 Principes Aérodynamiques . . . . .	16
1.2.2 Phase Musculaire . . . . .	18
1.2.3 Production du Human Beatbox . . . . .	22
1.3 Vitesse de production . . . . .	25
1.3.1 Effet du débit de parole sur la production . . . . .	25
1.3.2 Effet du tempo sur la production musicale . . . . .	28
1.4 Problématique, Objectifs et Hypothèses . . . . .	29
<b>2 Méthodologie</b>	<b>31</b>
2.1 Premiers Pas . . . . .	31
2.2 Principes de transcription . . . . .	32
2.3 Corpus . . . . .	35
2.4 Déroulement, sujets et protocole de l'étude . . . . .	35
2.4.1 Sujets . . . . .	37
2.4.2 Protocole expérimental . . . . .	38
2.4.3 Sujet Pilote . . . . .	38
2.5 Matériel Expérimental . . . . .	39
2.5.1 Fibroscopie laryngée . . . . .	39
2.5.2 Station d'enregistrement EVA2 . . . . .	39
2.5.3 Imagerie par Résonance Magnétique . . . . .	41
2.6 Analyse . . . . .	42
2.6.1 Analyse Qualitative . . . . .	42
2.6.2 Analyse Quantitative . . . . .	46
<b>3 Le Human Beatbox : un système combinatoire discret</b>	<b>50</b>
3.1 Initiation . . . . .	51
3.1.1 Mécanismes pulmoniques égressifs . . . . .	51
3.1.2 Mécanismes pulmoniques ingressifs . . . . .	54
3.1.3 Mécanismes laryngaux égressifs . . . . .	54
3.1.4 Mécanismes laryngaux ingressifs . . . . .	56
3.1.5 Mécanismes linguaux égressifs . . . . .	58

3.1.6	Mécanismes linguaux ingressifs . . . . .	60
3.1.7	Co-initiation . . . . .	62
3.2	Articulation . . . . .	65
3.2.1	Grosses caisses . . . . .	65
3.2.2	Charlestons . . . . .	67
3.2.3	Caisses claires . . . . .	67
3.2.4	Rolls . . . . .	72
3.3	Timbre . . . . .	73
3.3.1	Grosses Caisses . . . . .	73
3.3.2	Charlestons . . . . .	77
3.3.3	Caisses claires . . . . .	79
3.3.4	Rolls . . . . .	91
3.3.5	Des signatures acoustiques discrètes ? . . . . .	94
3.4	Le Human beatbox : un système combinatoire discret et complexe . . . . .	101
3.4.1	Voix beatboxée . . . . .	101
3.4.2	Mécanismes de modulation de la fréquence fondamentale . . . . .	110
3.4.3	Mécanismes sifflés . . . . .	114
<b>4</b>	<b>Des sons aux <i>Beat Patterns</i> : effet du tempo sur la production</b>	<b>118</b>
4.1	Analyse de la tâche de production . . . . .	119
4.1.1	Durée des patterns et vitesse de production . . . . .	119
4.1.2	Erreurs de production . . . . .	120
4.2	Analyse temporelle . . . . .	124
4.2.1	Effet du tempo sur la durée . . . . .	124
4.2.2	Effet du tempo selon les caractéristiques initiatrices des sons . . . . .	126
4.2.3	Effet du tempo selon les caractéristiques articulatoires des sons . . . . .	127
4.2.4	Effet du tempo selon la position des sons dans les patterns . . . . .	128
4.2.5	Effet du tempo selon le contexte phonétique adjacent . . . . .	130
4.2.6	Synthèse de l'analyse temporelle . . . . .	131
4.3	Analyse aérodynamique . . . . .	133
4.3.1	Effet du tempo sur la pression . . . . .	133
4.3.2	Effet du tempo sur l'aire de la constriction . . . . .	143
4.4	Analyse acoustique . . . . .	146
4.4.1	Effet du tempo sur les moments spectraux des sons . . . . .	146
4.4.2	Effet du tempo sur les moments spectraux en fonction du contexte adjacent	155
4.4.3	Synthèse de l'analyse acoustique . . . . .	157
<b>5</b>	<b>Discussion Générale</b>	<b>160</b>
5.1	Pratique artistique et objet d'étude scientifique . . . . .	160
5.2	Approche anthropophonique et primitives phonétiques . . . . .	161
5.3	Les sons du beatbox : des possibles sons de la parole ? . . . . .	164
5.4	Effort et contrainte de puissance . . . . .	166
5.5	La parole et le Human Beatbox . . . . .	168
	<b>Conclusion</b>	<b>170</b>
	<b>Annexes</b>	<b>172</b>
	<b>Bibliographie</b>	<b>230</b>



# Table des figures

1	Le Human Beatbox : une pratique vocale, musicale et sociale . . . . .	3
1.1	Vue antérieure (gauche) et postérieure (droite) de la structure osseuse de la cage thoracique (HIXON et al. (2020), Copyright 2005 by Thomas J. Hixon and Jeannette D. Hoit). . . . .	6
1.2	Vue antérieure du diaphragme (gauche), vue latérale des intercostaux externes (milieu) et vue postérieure des intercostaux internes (droite) (HIXON et al. (2020), Copyright 2005 by Thomas J. Hixon and Jeannette D. Hoit). . . . .	8
1.3	Vue latérale de la structure osseuse et cartilagineuse du larynx (HIXON et al. (2020), Copyright 2005 by Thomas J. Hixon and Jeannette D. Hoit). . . . .	9
1.4	Illustration des muscles laryngés intrinsèques (HIXON et al. (2020), Copyright 2005 by Thomas J. Hixon and Jeannette D. Hoit). . . . .	10
1.5	Muscles laryngés extrinsèques (HIXON et al. (2020), Copyright 2005 by Thomas J. Hixon and Jeannette D. Hoit). . . . .	12
1.6	Muscles Linguaux (HIXON et al. (2020), Copyright 2005 by Thomas J. Hixon and Jeannette D. Hoit). . . . .	13
1.7	Muscles faciaux : muscles de la mâchoire (images 1-3) et muscles labiaux (image 4) (HIXON et al. (2020), Copyright 2005 by Thomas J. Hixon and Jeannette D. Hoit). . . . .	15
1.8	Mécanismes d’initiation pulmonique (gauche), laryngal (droite-haut) et lingual (droite-bas). Les flèches indiquent une action de compression des organes initiant des flux d’air égressifs. . . . .	20
1.9	Hypothèses des stratégies de réduction temporelle des gestes de BYRD et TAN (1996) . . . . .	26
2.1	Illustration de la structure métrique d’un <i>Beat Pattern</i> où le tempo de référence est 90BPM. . . . .	35
2.2	Déroulement d’une session d’acquisition . . . . .	36
2.3	Nasofibroscope Souple . . . . .	39
2.4	Station d’enregistrement EVA2 . . . . .	40
2.5	Procédure de repositionnement du 0 des signaux aérodynamiques . . . . .	41
2.6	Acquisition en Imagerie par Résonance Magnétique . . . . .	42
2.7	Lieux d’articulation adapté de ESLING et al. (2019). . . . .	44
2.8	ESLING et al. (2019) : illustration des configurations laryngales pour plusieurs types de voix. À gauche la voix modale, en haut les voix produites sans constriction aryépiglottale, en bas les voix produites avec constriction aryépiglottale. . . . .	45
2.9	Caractérisation du timbre en musique (CASTELLENGO, 2015) . . . . .	45
2.10	Segmentation de 4 sons beatboxés du sujet CJ. Violet = occlusion, rose = burst, bleu = bruit d’explosion, vert = friction ou trille . . . . .	47

2.11	<i>Skewness et Kurtosis</i> . . . . .	48
3.1	<i>ESH Snare Outward</i> [ $\widehat{tʃ}$ ] - Sujet GA . . . . .	52
3.2	<i>K-Snare Inward</i> [ $\downarrow k_L$ ] - Sujet CJ . . . . .	53
3.3	<i>Classic kick</i> [p'] - Sujet VP . . . . .	55
3.4	Mode d'initiation par compression laryngale - [p'] . . . . .	56
3.5	<i>Throat kick</i> [b'] - Sujet GA . . . . .	57
3.6	Mode d'initiation par expansion laryngale - [b'] . . . . .	58
3.7	<i>PF snare humming</i> { $\widehat{pf}$ } - Sujet VP . . . . .	59
3.8	Mode d'initiation par compression - { $\widehat{pf}$ } . . . . .	60
3.9	<i>Lips Roll Humming</i> [ $\downarrow_B^1$ ] - Sujet VP . . . . .	61
3.10	Mode d'initiation par expansion - { $\downarrow_B^1$ } . . . . .	62
3.11	Deux exemples de co-initiation lors de la production de beatbox. Le premier exemple est tiré du corpus de GARRIGUES (2015) et analysé dans DEHAIS-UNDERDOWN et al. (2019). Le deuxième exemple provient de nos données. . . . .	63
3.12	Combinaison des mécanismes initiatoires et articulatoires des grosses caisses du sujet VP. La position au relâchement de [ $\downarrow p$ ] correspond au début de la trille du <i>Lips Roll</i> [ $\downarrow_B^1$ ]. . . . .	65
3.13	Production du <i>Classic Kick Inward</i> [ $\beta$ ] par expansion laryngale. Sujet AI : Fermeture ventriculaire puis recul du dos de la langue et abaissement du larynx. Sujet GA : Fermeture ventriculaire et abaissement du larynx. . . . .	66
3.14	Combinaison des mécanismes initiatoires et articulatoires des charlestons du sujet VP. . . . .	67
3.15	Combinaison des mécanismes initiatoires et articulatoires des caisses claires orales du sujet VP. . . . .	68
3.16	Mouvement latéral de la mâchoire pendant la production de [ $\downarrow k_L$ ] du sujet VP . . . . .	69
3.17	Production d'une <i>K-Snare Inward</i> [ $\downarrow k_L$ ] du sujet AI par expansion laryngale . . . . .	69
3.18	Combinaison des mécanismes initiatoires et articulatoires des caisses claires laryngales du sujet VP. . . . .	70
3.19	Configuration laryngale pendant l'occlusion et le relâchement de [ $\widehat{tʃ}$ ] . . . . .	71
3.20	Configuration laryngale pendant l'occlusion et le relâchement de [ $\downarrow \widehat{tʃ}$ ] . . . . .	71
3.21	Configuration laryngale pendant l'occlusion et le relâchement de [ $\widehat{tʃh}$ ] . . . . .	72
3.22	Combinaison des mécanismes initiatoires et articulatoires des <i>Rolls</i> . . . . .	72
3.23	Timbre du <i>Classic Kick</i> [p'ϕ] - Sujet GA . . . . .	73
3.24	Timbre du <i>Classic Kick humming</i> {p}- Sujet GA . . . . .	74
3.25	Timbre du <i>Classic Kick Inward</i> [ $\beta$ ] - Sujet GA . . . . .	75
3.26	Timbre du <i>Kick Lingual</i> [t'] - Sujet GA . . . . .	75
3.27	Comparaison du timbre des grosses caisses du sujet GA à partir de spectre LPC à 24 coefficients . . . . .	76
3.28	Timbre du <i>Hi-Hat fermé</i> - Sujet AI . . . . .	77
3.29	Timbre du <i>Hi-Hat Humming</i> { $\widehat{ts}$ } - Sujet AI . . . . .	78
3.30	Comparaison du timbre des charlestons du sujet AI à partir de spectre LPC à 24 coefficients . . . . .	79
3.31	Timbre de la <i>K-Snare inward</i> [ $\downarrow k_L$ ] - Sujet VP . . . . .	80
3.32	Timbre de la <i>K-Snare Humming</i> [ $\downarrow k_L$ ] - Sujet VP . . . . .	81
3.33	Timbre de la <i>PF Snare</i> [ $\widehat{pf}$ ] - Sujet VP . . . . .	82
3.34	Timbre de la <i>PF Snare humming</i> { $\widehat{pf}$ } - Sujet VP . . . . .	83
3.35	Timbre de la <i>Tish Snare</i> [tʃ'] - Sujet VP . . . . .	84

3.36	Comparaison des timbres de trois caisses claires orales au moment du relâchement	85
3.37	Comparaison des timbres au moment de la friction	86
3.38	Timbre de la <i>ESH Snare Outward</i> [ʔ̥] : - Sujet VP	87
3.39	Timbre de la <i>ESH Snare Inward</i> [↓ʔ̥] : - Sujet VP	88
3.40	Timbre de la <i>Cough Snare</i> [ʔ̥h] : - Sujet VP	89
3.41	Comparaison des timbres au moment du relâchement (gauche) et de la friction (droite) des trois caisses claires laryngées	90
3.42	Timbre du <i>Lips Roll</i> pulmonique [↓B <sup>1</sup> ] : - Sujet CJ	91
3.43	Comparaison d'un jeu de timbre entre deux <i>Lips Roll</i> du sujet VP	92
3.44	Timbre du <i>Lips Roll Humming</i> {↓B <sup>1</sup> } : - Sujet CJ	93
3.45	Timbre du <i>Clic Roll</i> [↓ɿ] : - Sujet CJ	93
3.46	Comparaison intersujet des signaux acoustiques et des spectrogrammes de la <i>PF Snare</i> [pf] :. <u>Attention, échelles de temps différentes.</u>	95
3.47	Comparaison intersujet des signaux acoustiques et des spectrogrammes de la <i>Tish Snare</i> [tʃ] :. <u>Attention, échelles de temps différentes.</u>	96
3.48	Comparaison intersujet des signaux acoustiques et des spectrogrammes de la <i>ESH Snare Inward</i> [↓ʔ̥] :. <u>Attention, échelles de temps différentes.</u>	97
3.49	Comparaison intersujet des signaux acoustiques et des spectrogrammes de la <i>Cough Snare</i> [ʔ̥h] :. <u>Attention, échelles de temps différentes.</u>	98
3.50	Comparaison intersujet des signaux acoustiques et des spectrogrammes du <i>Lips Roll</i> [↓B <sup>1</sup> ] :. <u>Attention, échelles de temps différentes.</u>	99
3.51	Comparaison intersujet des signaux acoustiques et des spectrogrammes du <i>Clic Roll</i> [↓ɿ] : (Pas de données pour le sujet VP). <u>Attention, échelles de temps différentes.</u>	100
3.52	Production de la Throat Bass par 4 sujets différents (Attention les échelles de temps sont différentes)	103
3.53	Spectre FFT de la structure harmonique des basses fréquences de la <i>Throat Bass</i> . (Le trait rouge indique H1)	104
3.54	<i>Inward Bass</i> - Sujet CJ	105
3.55	<i>Vibration Bass</i> - sujet GA. Sur les spectres les traits pleins correspondent à la partie orale du son ; les pointillés à la partie nasale.	107
3.56	<i>Inward Bass</i> - GA.	108
3.57	Production d'une « <i>double voix outward</i> » par le sujet GA : données aérodynamiques (gauche), spectre FFT (droite)	109
3.58	Trois images fibroscopiques extraites successivement lors de la production de la <i>double voix outward</i> . La ligne rouge symbolise la ligne médiane (milieu de la glotte). Un subtil décalage du mouvement de chaque pli vers la ligne médiane peut être observé.	110
3.59	Modulation de la F0 d'un scratch vocal du sujet VP	111
3.60	Modulation de la F0 de 494Hz à 45Hz par le sujet CJ	112
3.61	Vibration du pli aryépiglottique gauche lors de la production de la « <i>Chest Bass</i> » du sujet CJ	114
3.62	Mécanisme sifflé	115
3.63	Beat Pattern sifflé - sujet CJ.	116
4.1	Durée des BP en fonction du tempo de référence et du type de pattern (n= 483 BP).	119

4.2	Vitesse de production (BPM) en fonction du tempo de référence (BPM) pour chaque sujet. Les pointillés sur l'axe de la vitesse de production indiquent le tempo de référence. . . . .	120
4.3	Diagramme en camembert du taux d'erreurs de production des sujets. . . . .	121
4.4	Matrices d'erreurs de production des sujets. Les lignes = cibles attendues, les colonnes = sons tels que produits. Les matrices affichent uniquement les erreurs des beatboxeurs et ne sont pas directement comparables . . . . .	124
4.5	durée acoustique des sons (ms) en fonction du tempo de référence (BPM) pour tous les sons du corpus. Chaque point correspond à la moyenne de toutes les répétitions d'un son. Les barres d'erreur indiquent l'écart-type de la durée. . . . .	125
4.6	Durée des sons (ms) en fonction du tempo de référence (BPM). Les sons « _ » et « - » correspondent à l'intervalle entre les sons dont l'occlusion a été mesurée. Chaque point correspond à la moyenne de toutes les répétitions d'un son. Les barres d'erreur indiquent l'écart-type de la durée. . . . .	126
4.7	durée acoustique des sons (ms) en fonction du tempo de référence (BPM) selon les lieux et modes d'initiation. Seuls les sons affriqués sont illustrés ici. Chaque point correspond à la moyenne de toutes les répétitions d'un son. Les barres d'erreur indiquent l'écart-type de la durée. . . . .	127
4.8	durée acoustique des sons (ms) en fonction du tempo de référence (BPM) selon les lieux et modes d'initiation. Seuls les sons affriqués sont illustrés ici. Chaque point correspond à la moyenne de toutes les répétitions d'un son. Les barres d'erreur indiquent l'écart-type de la durée. . . . .	128
4.9	durée acoustique des sons (ms) en fonction du tempo de référence (BPM) selon la position métrique. Seuls 4 <i>beats patterns</i> sont illustrés ici. Chaque point correspond à la moyenne de toutes les répétitions d'un son. Les barres d'erreur indiquent l'écart-type de la durée. . . . .	129
4.10	Durée des sons (ms) en fonction du tempo de référence (BPM) selon la position métrique. Seul le son [ts'] est illustré ici. . . . .	131
4.11	Graphiques illustrant la différence entre valeurs positives et négatives (gauche) et les valeurs positives de pression où $P = \sqrt{(-P^2)}$ (droite) . . . . .	133
4.12	Pression intra-orale moyenne (hPa) en fonction du tempo de référence (BPM). . . . .	135
4.13	Pression intra-orale moyenne (hPa) en fonction du tempo de référence (BPM) selon les mécanismes d'initiation. . . . .	136
4.14	Pression intra-orale moyenne (hPa) en fonction du tempo de référence (BPM) selon les mécanismes d'articulation. . . . .	138
4.15	Pression intra-orale moyenne (hPa) en fonction du tempo de référence (BPM) selon les types de BP. . . . .	139
4.16	Illustration des signaux acoustique (vert), aérodynamiques (pression = cyan, débit d'air oral = bleu, débit d'air nasal = violet) et du spectrogramme de deux BP produits par le sujet VP. . . . .	139
4.17	Signal acoustique et images fibroscopiques correspondantes de la production par CJ de [ts'] en position 2 et de [ʧʃ:] en position 3 dans le BP [b̃ ts' ʧʃ : ts' b̃ b̃ ts' ʧʃ : ts']. Tempo de référence = 90BPM. . . . .	141
4.18	Pression intra-orale moyenne (hPa) en fonction du tempo de référence (BPM) du BP [p' ts' pf : ts' p' p' ts' pf : ts'] produit par GA. . . . .	142

4.19	Aire relative de constriction ( $cm^2$ ) en fonction du tempo de référence (BPM) selon les sons. Les pointillés à $0.03cm^2$ indiquent le seuil en-dessous duquel l'intensité de la turbulence est réduite. Les pointillés à $0.2cm^2$ indiquent le seuil au dessus duquel l'intensité de la turbulence est réduite. . . . .	144
4.20	Aire relative de constriction ( $cm^2$ ) en fonction du tempo de référence (BPM) pour le son $[\widehat{ts}']$ dans différents BP et différentes positions. Les pointillés à $0.03cm^2$ indiquent le seuil en-dessous duquel l'intensité de la turbulence est réduite. Les pointillés à $0.2cm^2$ indiquent le seuil au dessus duquel l'intensité de la turbulence est réduite. . . . .	145
4.21	Centre de Gravité Spectral (kHz) en fonction du tempo de référence (BPM) selon les sons, les événements acoustiques et les sujets. La durée des fenêtres d'analyse est indiquée par des formes de point différentes. Les lignes en pointillés délimitent les basses fréquences ( $F < 1.5kHz$ ), les moyennes fréquences ( $1.5 < F < 4kHz$ ) et les hautes fréquences ( $F > 4kHz$ ). . . . .	147
4.22	Écart-type du Centre de Gravité spectral (kHz) grosses caisses en fonction du tempo de référence. L'écart-type décrit l'étendue fréquentielle du Centre de Gravité spectral. . . . .	147
4.23	Centre de Gravité Spectral (kHz) des charlestons en fonction du tempo de référence (BPM) selon les sons, les événements acoustiques et les sujets. . . . .	148
4.24	Écart-type du Centre de Gravité spectral (kHz) charlestons en fonction du tempo de référence (BPM) selon les sons, les événements acoustiques et les sujets. . . . .	149
4.25	Centre de Gravité Spectral (kHz) des caisses claires orales en fonction du tempo de référence (BPM). . . . .	150
4.26	Écart-type du Centre de Gravité spectral (kHz) caisses claires orales en fonction du tempo de référence (BPM). . . . .	151
4.27	Centre de Gravité Spectral (kHz) en fonction du tempo de référence (BPM) selon les sons, les événements acoustiques et les sujets. . . . .	152
4.28	F0 moyenne (Hz) des caisses claires laryngales en fonction du tempo de référence (BPM). . . . .	153
4.29	Écart-type du Centre de Gravité spectral (kHz) des caisses claires laryngales en fonction du tempo de référence (BPM) selon les sons, les événements acoustiques et les sujets. . . . .	153
4.30	Centre de Gravité Spectral (kHz) des <i>Rolls</i> en fonction du tempo de référence (BPM). . . . .	154
4.31	Écart-type du Centre de Gravité spectral (kHz) <i>Rolls</i> en fonction du tempo de référence (BPM). . . . .	154
4.32	Centre de Gravité spectral (kHz) du bruit de friction de $[\widehat{ts}']$ en fonction du tempo. Le nombre d'échantillons correspond au nombre de fenêtres d'analyse de $0.05ms$ . . . . .	157
4.33	Écart-type du Centre de Gravité spectral (kHz) du bruit de friction de $[\widehat{ts}']$ en fonction du tempo. Le nombre d'échantillons correspond au nombre de fenêtres d'analyse de $0.05ms$ . . . . .	158
5.1	Approche anthropophonique de la production des sons (CATFORD, 1977; LINDBLOM, 1990) . . . . .	162
5.2	<i>Europa</i> , Santana . . . . .	164
5.3	Degrés de liberté de la pointe de la langue et du corps de la langue (Adapté de LINDBLOM, 1983) . . . . .	167

5.4	Approche anthropophonique du Human Beatbox . . . . .	168
-----	--	-----

# Liste des tableaux

1.1	Paramètres d'initiation selon CATFORD (1977) . . . . .	18
1.2	Gestes d'initiation . . . . .	21
2.1	Corpus des Beat Patterns (Audio intégré en cliquant sur le bouton lecture. Pour fermer un lecteur et jouer le son suivant : clic droit > désactiver le contenu). . .	33
2.2	Corpus des sons. Les sons dits pulmoniques <u>ou</u> laryngaux indiquent que nos données attestent les deux variantes d'initiation du son. . . . .	34
2.3	Profil des participant-es de l'étude . . . . .	37

# Introduction

Avant d'être un objet d'étude scientifique, le Human Beatbox (HBB) est avant tout une pratique musicale. Les musiciens utilisent leur conduit vocal pour produire des sons musicaux, c'est-à-dire, des sons dénués de sens. Ce n'est pas l'unique technique qui utilise le conduit vocal à des fins musicales. On pense à la chanteuse de *Jazz* Ella Fitzgerald, qualifiée sur France Musique d'« Acrobate du scat », c'est-à-dire l'art de chanter des onomatopées à la place des paroles (Betillouloux, 2020). On peut aussi mentionner la poésie sonore d'Henri Chopin qui produisait des sons dénués de signification. La production de combinaisons articulatoires dénuées de sens n'est donc pas le fait exclusif du HBB.

Le HBB est une discipline à part entière du Hip-Hop, qui a fait ses débuts dans les années 1980 à Brooklyn, New York (Martino, 2009 ; Park et al., 2005). À cette époque, on voit arriver l'utilisation de « drum machines » ou « beatbox » pour accompagner les MC (Maître de Cérémonie). Ces boîtes à rythme permettaient de jouer une base instrumentale de sons de batterie pour accompagner les rappeurs. L'achat des « beatbox » était coûteuse pour les artistes les plus modestes ; à l'inverse, le conduit vocal est gratuit et disponible à tout moment. Voici donc comment sont nés ces boîtes à rythme humaines faites de chair et d'os qui accompagnaient les rappeurs. Selon une hypothèse des pratiquant-es du Human Beatbox, la pratique se serait aussi développée en réponse à la confiscation des beatbox électroniques par la police.

Les premiers beatboxeurs sont ceux de la « *Old School* » (ancienne école) des années 80. On pense par exemple aux « *Fat Boys* » et leur tube *The human beatbox*<sup>1</sup>. C'est dans les années 90, et en particulier en 1999, que le HBB prend un virage important. C'est avec l'album « *Make the music 2000* », Razhel aka « *Godfather of Noyze* » (le parrain du bruit), que la discipline gagne en autonomie. En effet, un beatboxeur peut performer seul grâce à des *loopstation*<sup>2</sup>. Aujourd'hui, les beatboxeurs et les beatboxeuses de la « *New School* » (Nouvelle École) n'ont de cesse de perfectionner leur pratique musicale en explorant les capacités du conduit vocal. Internet a fortement participé à l'essor de la discipline et à la création d'une communauté internationale. Les tutoriels vidéos, échangés à travers le monde, permettent de s'initier, de progresser ou encore de perfectionner sa technique vocale. Le HBB n'est pas seulement un loisir, c'est une profession. Nombreuses sont les personnes qui exercent à mi-temps cette activité : ateliers d'apprentissage, partenariats associatifs, ateliers de rééducation orthophonique, projets artistiques ; en revanche les personnes exerçant à plein temps sont moins nombreuses.

L'essor de la pratique dans les années 2000 n'a pas manqué d'attirer l'attention des chercheurs et des chercheuses. Stowell et Plumbley (2008) donnent quelques caractéristiques vocales du Human Beatbox : patrons de production non-syllabiques, usage de sons pulmoniques ingressifs (*inhaled sounds*), qualités de voix variées ... Bien que leur étude se fonde sur une analyse « impressionniste » du Beatbox, les auteurs placent le beatbox parmi les phénomènes phonétiques. Depuis, deux approches principales ont été proposées pour étudier le beatbox :

1. <https://youtu.be/6c7b-Kfg0fU?si=5nPL0ERLBQUgq8v6>

2. Les *loopstation* sont des appareils électroniques permettant d'enregistrer et de jouer en boucle un enchaînement rythmique ou mélodique



l'approche paralinguistique (Proctor et al., 2013) et l'approche phonologique (Blaylock, 2022). Nous reviendrons sur ces approches dans le Chapitre 1. Entre musique et phonétique, le Human Beatbox est un objet d'étude scientifique qui n'a pas encore de cadre théorique clairement défini. Nous savons qu'il s'agit d'un système musical dont les notes sont produites par le conduit vocal. Il s'agit donc à la fois d'une pratique vocale et musicale. C'est le caractère atypique de cette technique qui a éveillé la curiosité des phonéticiens et phonéticiennes : comment est produit le Human Beatbox ? Quelles sont les capacités du conduit vocal humain ? Existe-t-il un lien entre la production du Beatbox et de la parole ? Cette thèse essaiera d'apporter des éléments de discussion sur ces questions.

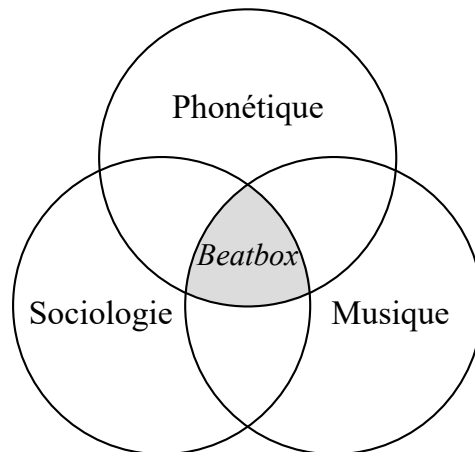


FIGURE 1 – Le Human Beatbox : une pratique vocale, musicale et sociale

Le Human Beatbox est aussi une pratique sociale (Martino, 2009). Le beatbox se trouve donc au carrefour de la phonétique, de la musique et de la sociologie. Les beatboxeurs et les beatboxeuses sont détentrices du savoir concernant les pratiques du Human Beatbox. C'est pour cela que cette thèse s'est faite en collaboration avec Paul Vignes (beatboxeur professionnel) afin de nourrir nos recherches de ses connaissances sur les multiples façons de pratiquer le HBB. Dans la perspective de comprendre à la fois le beatbox et les personnes qui pratiquent cette technique, nous avons adressé un questionnaire aux 5 sujets de notre étude. Nous reviendrons sur les questionnaires dans le Chapitre 5.

Cette thèse s'intéresse à la production du Human Beatbox à partir de données articulatoires, aérodynamiques et acoustiques. Afin de décrire les mécanismes de production de cette technique vocale, nous partirons des connaissances phonétiques sur les mécanismes de production de la parole. Nous voulons aussi savoir comment varient les sons beatboxés. Pour ce faire, nous analyserons des productions à l'isolé ainsi que des productions de sons dans des structures musicales que nous appelons *Beats Patterns*. Nous reviendrons sur la définition des *Beats Patterns* (BP) dans la méthodologie (Chapitre 2). Ces BP sont principalement composés d'une suite de sons occlusifs et affriqués. L'intervalle entre les sons se caractérise par un silence. Afin de s'assurer de bien induire de la variation, nous l'introduirons en utilisant un paradigme d'augmentation de la vitesse de production avec 3 tempos de référence : 90, 120 et 150 battements par minute (BPM). Cette thèse a deux objectifs :

1. En reprenant la proposition de Proctor et al. (2013) nous montrerons que le HBB est un système combinatoire discret. Nous tenterons de démontrer en quoi le HBB est un système de production (articulatoirement) plus complexe que la parole.
2. En reprenant le paradigme de variation de la vitesse de production, nous analyserons les effets du tempo de référence sur la production du HBB à partir de données temporelles,

aérodynamiques et acoustiques.

La présente thèse se compose de 5 chapitres. Le Chapitre 1 introduit le cadre théorique à partir duquel nous analyserons la production du Human Beatbox. Le Chapitre 2 développe la méthodologie adoptée. Le Chapitre 3 constitue une description de la production et de la combinaison de divers mécanismes de production. Le Chapitre 4 présente les résultats des effets du tempo sur la production du Human Beatbox. Enfin, le Chapitre 5 propose d'élargir le cadre théorique en phonétique.

# Chapitre 1

## Cadre Théorique

« in order to cope efficiently with the vocal sounds that constitute the sound-systems of particular languages, phonetics must proceed from the most general possible consideration of the human sound-producing potential »

---

*J. C. Catford*<sup>1</sup>

Ce Chapitre pose le cadre théorique dans lequel s'inscrit notre étude. Nous introduirons l'anatomie du système de production vocale. Nous mentionnerons ensuite les mécanismes de production de sons de la parole et du HBB. Enfin, nous terminerons par évoquer les effets de la vitesse de production sur la parole et la musique.

### 1.1 Anatomie du système de production vocale

L'activité langagière implique l'utilisation d'un nombre important de muscles afin de produire des sons. La fonction première des groupes de muscles est d'assurer les fonctions vitales du corps humain tel que la respiration ou la déglutition par exemple. La communication orale dans l'espèce humaine entraîne une réorganisation du fonctionnement physiologique de ces multiples systèmes dans le but d'assurer à la fois les fonctions vitales et la production de la parole. De même, lors de la production du HBB une autre organisation prend place pour assurer (1) les fonctions vitales et (2) une production vocale continue, c'est-à-dire ininterrompue<sup>2</sup>. Nous commencerons ce chapitre en passant en revue les différents sous-systèmes composant le système de production vocale. Nous commencerons par le système respiratoire, puis nous évoquerons le système laryngé. Enfin, nous terminerons en mentionnant le système oral dans lequel nous regroupons les sous-systèmes musculaires de la langue, des lèvres et de la mâchoire inférieure.

---

1. Catford, J. C. (1977). *Fundamental problems in phonetics*. Indiana University Press.

2. Rappelons que lors de battle de HBB ou bien de concerts, la personne qui beatboxe doit le faire pendant plusieurs minutes sans avoir à s'arrêter pour reprendre son souffle.

## 1.1.1 Système respiratoire

L'organe principal permettant la respiration sont les poumons. Les poumons assurent en premier lieu l'*hématose* c'est-à-dire « la transformation du sang veineux en sang artériel » pour permettre l'oxygénation du sang (Le Huche & Allali, 2010). Les propriétés élastiques des tissus des poumons agissent comme une force passive lors de l'expiration. Lors de la respiration calme, les phases inspiratoires et expiratoires sont plus ou moins équivalentes soit (Gick et al., 2013 ; Hixon et al., 2020). En revanche, lorsque l'on parle, on inspire un plus grand volume d'air sur une durée plus courte et on expire, pour parler, sur une durée plus longue (*ibid*). Concernant le HBB, l'activité respiratoire est réorganisée en intégrant des sons pulmoniques ingressifs (produits sur l'inspiration), en utilisant la respiration circulaire ou gérant son souffle par des phases d'inspiration ou d'expiration entre deux sons<sup>3</sup>.

### 1.1.1.1 Structures osseuses et cartilagineuses

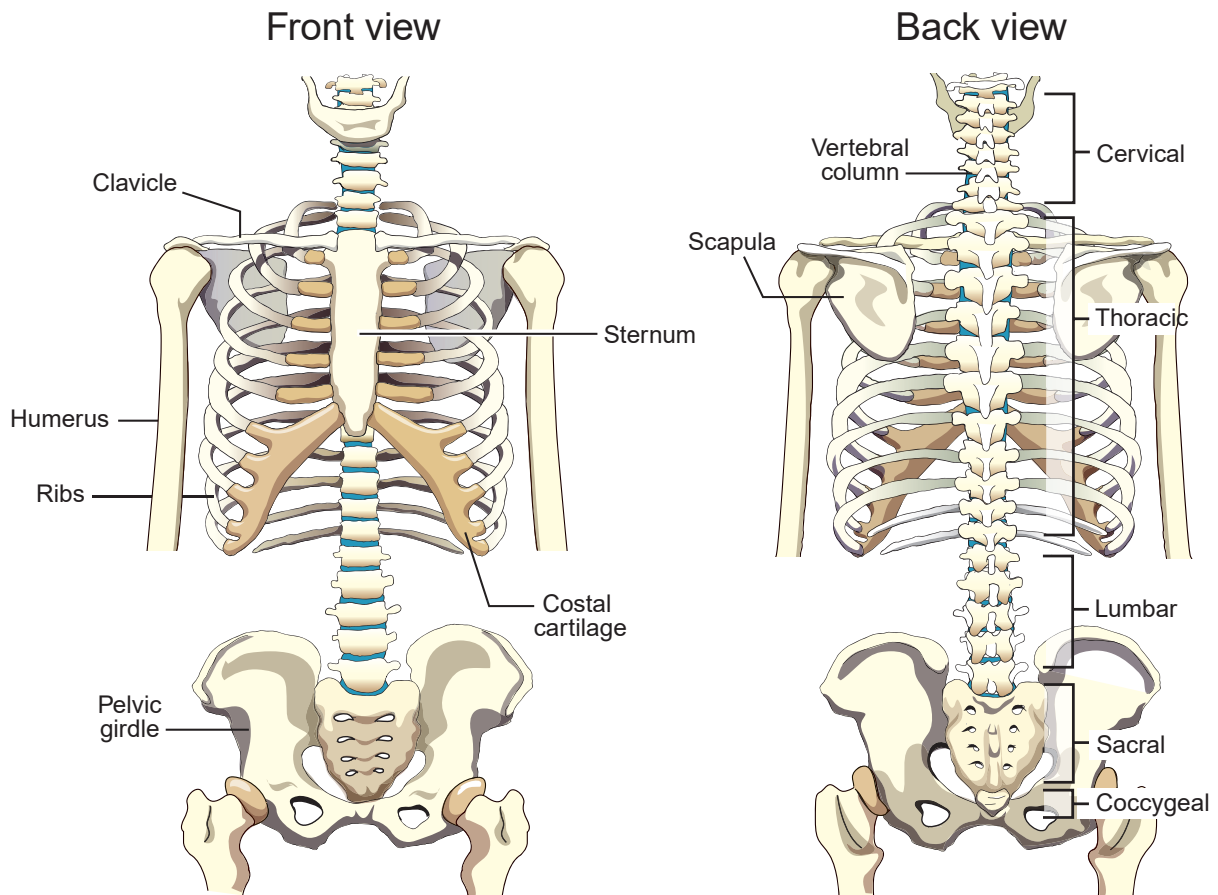


FIGURE 1.1 – Vue antérieure (gauche) et postérieure (droite) de la structure osseuse de la cage thoracique (Hixon et al. (2020), Copyright 2005 by Thomas J. Hixon and Jeanette D. Hoit).

Le système respiratoire se compose des poumons et de la paroi thoracique (cage thoracique, diaphragme, parois abdominales). Il se compose aussi de structures osseuses et cartilagineuses, de tendons et de muscles. La charpente osseuse est constituée de plusieurs vertèbres, qui forment

3. cf. Question 27 des questionnaires en Annexes.

la colonne vertébrale. Les côtes s'insèrent sur les vertèbres et forment la cage thoracique. La cage thoracique se compose de 12 côtes de chaque côté, qui s'insèrent dans le sternum via des cartilages costaux (Figure 1.1). Les 2 dernières côtes (« côtes flottantes ») ne s'insèrent pas sur le sternum.

Les côtes et la cage thoracique sont mises en mouvement par l'activité des muscles respiratoires. Les côtes peuvent se déplacer sur un axe horizontal, on parle d'un mouvement en « *poignée de pompe* » ou bien sur un axe oblique, on parle alors d'un mouvement en « *anse de seau* ». Ces deux types des mouvements permettent d'élever ou d'abaisser les côtes ou la cage thoracique dans son ensemble. L'élévation des côtes et de la cage thoracique augmentent le volume pulmonaire ce qui crée un flux inspiratoire ; l'abaissement provoque une diminution du volume pulmonaire et un flux expiratoire (Hixon et al., 2020 ; Le Huche & Allali, 2010).

Enfin, lorsque l'activité respiratoire est accrue, des mouvements vertébraux peuvent s'ajouter à ceux des côtes et de la cage thoracique. L'activation des muscles spinaux permet l'extension ou la flexion de la colonne vertébrale. L'extension de la colonne vertébrale éloigne les côtes en « éventail » et augmente le volume thoracique et permet une inspiration accrue. À l'inverse, le fléchissement réduit le volume thoracique et permet une expiration accrue (Le Huche & Allali, 2010).

### 1.1.1.2 Muscles respiratoires

Les muscles sont des fibres qui se contractent et se relâchent. D'un point de vue respiratoire, cela se traduit par une phase inspiratoire puis expiratoire. Il existe des muscles inspireurs et des muscles expirateurs. Ces muscles sont divisés en muscles dits « principaux », c'est-à-dire muscles impliqués dans la respiration calme et les muscles dits « accessoires » qui s'activent lorsque l'activité respiratoire est accrue.

#### 1.1.1.2.1 Muscles inspireurs

Le *diaphragme* (Figure 1.2) est un muscle inspireur principal. Il sépare le thorax de l'abdomen. L'activation de ce muscle complexe a plusieurs effets : il augmente le volume thoracique en s'abaissant, en déplaçant la paroi abdominale en avant et en élargissant les parois du thorax sur les côtés. L'augmentation du volume crée un appel d'air et donc un flux inspiratoire (Gick et al., 2013 ; Hixon et al., 2020 ; Le Huche & Allali, 2010 ; Marchal, 2011 ; Ratnovsky et al., 2008). Les *intercostaux externes* (Figure 1.2) sont aussi des muscles inspireurs principaux (*ibid.*). Activés collectivement, les onze intercostaux externes élèvent l'ensemble de la cage thoracique et l'élargissent sur les côtés. Cela a pour effet d'augmenter le volume thoracique et de créer un appel d'air.

Les muscles inspireurs accessoires sont les *scalènes*, le *sterno-cléido-mastoïdien* et le *grand et petit pectoral*. Leur contraction élève les premières côtes et le sternum ce qui a pour effet d'augmenter le volume thoracique en créant un appel d'air (*ibid.*). En revanche, Le Huche et Allali (2010) soulignent le manque de résultats expérimentaux quant au rôle des pectoraux dans l'activité respiratoire accrue.

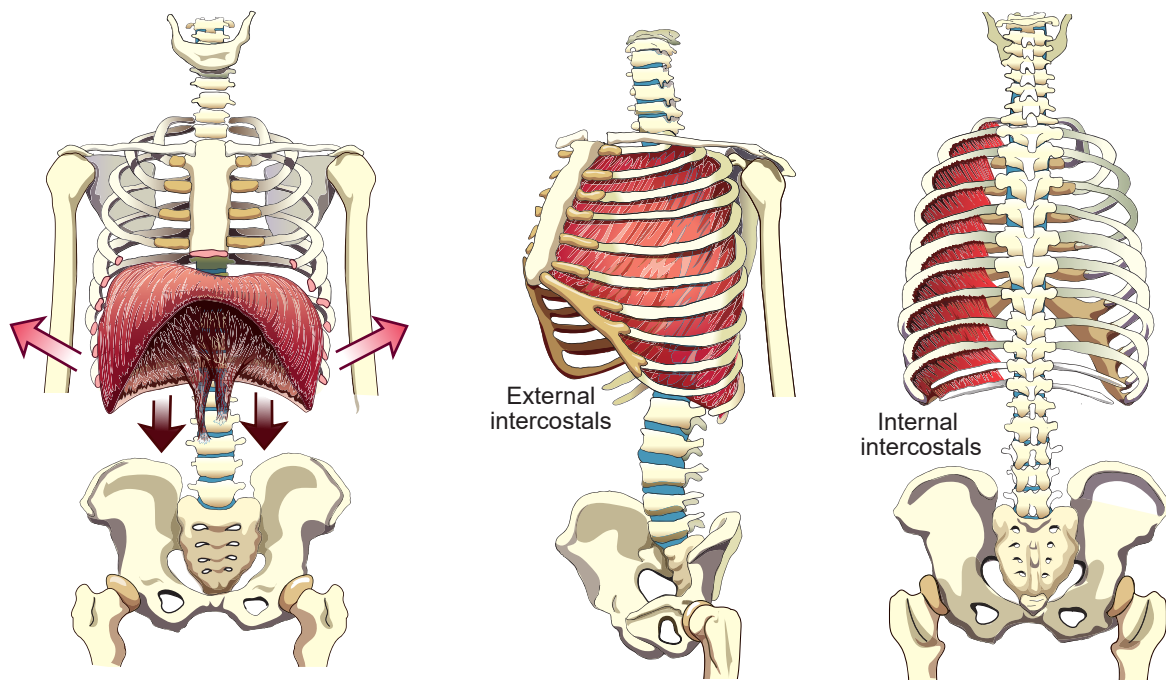


FIGURE 1.2 – Vue antérieure du diaphragme (gauche), vue latérale des intercostaux externes (milieu) et vue postérieure des intercostaux internes (droite) (Hixon et al. (2020), Copyright 2005 by Thomas J. Hixon and Jeannette D. Hoit).

#### 1.1.1.2.2 Muscles Expirateurs

*Les intercostaux internes* (Figure 1.2) sont les principaux muscles expirateurs (*ibid.*). Tout comme les intercostaux externes, ils sont au nombre de 11. Activés collectivement, ces muscles permettent d'abaisser la cage thoracique. Cela a pour effet de diminuer le volume de celle-ci.

*Les muscles abdominaux* font partie des muscles expirateurs accessoires. Le *grand droit* est un abaisseur du thorax, son activation diminue le volume thoracique. Les muscles *obliques internes et externes* ainsi que le *transverse abdominal* compriment l'abdomen et les viscères qui, en remontant le diaphragme vers le haut, diminuent le volume thoracique (*ibid.*)

#### 1.1.1.2.3 Muscles spinaux et dorsaux

Les muscles spinaux et dorsaux comme *les élévateurs costaux*, *l'iliocostal* (ou sacrolombaire), *le grand dorsal*, *le transversaire épineux* et *le muscle épineux* peuvent s'activer lors d'une activité respiratoire fortement accrue. *Les élévateurs costaux* et *l'iliocostal* peuvent participer à l'inspiration en élevant les côtes supérieures et en élargissant la partie inférieure du thorax ; ils peuvent aussi participer à l'expiration en abaissant les dernières côtes et en comprimant la partie inférieure du thorax. Les autres muscles sont des extenseurs de la colonne vertébrale qui augmentent le volume thoracique (Le Huche et Allali, 2010 ; Marchal, 2011 ; Gick et al., 2013 ; Hixon et al., 2020).

### 1.1.2 Système laryngé

Le système laryngé est un système complexe composé de muscles et de cartilages. Beaucoup d'attention a été portée au larynx comme source de voisement au niveau des plis vocaux (appelés couramment cordes vocales). Or, les travaux de John Esling puis Scott Moisik montrent qu'il est

réducteur de considérer le système laryngé comme simple source de voisement produit par les plis vocaux. Ils ont développé le « *Laryngeal Articulator Model (LAM)* » (Esling et al., 2019) afin de définir précisément le fonctionnement des structures laryngées pendant la production de la parole. De plus, tout comme le système respiratoire peut générer des flux d'air ingressifs et égressifs, le larynx peut lui aussi mettre l'air en mouvement. Nous reviendrons plus loin sur le larynx comme initiateur.

### 1.1.2.1 Structures osseuses et cartilagineuses

Le larynx se compose d'un seul os et de six cartilages (Figure 1.3) dans lesquels vont s'insérer les structures musculaires (Esling et al., 2019; Gick et al., 2013; Hixon et al., 2020; Le Huche & Allali, 2010; Marchal, 2011). En se contractant, les muscles laryngaux permettent de produire des gestes phonatoires, articulatoires ou encore initiateurs.

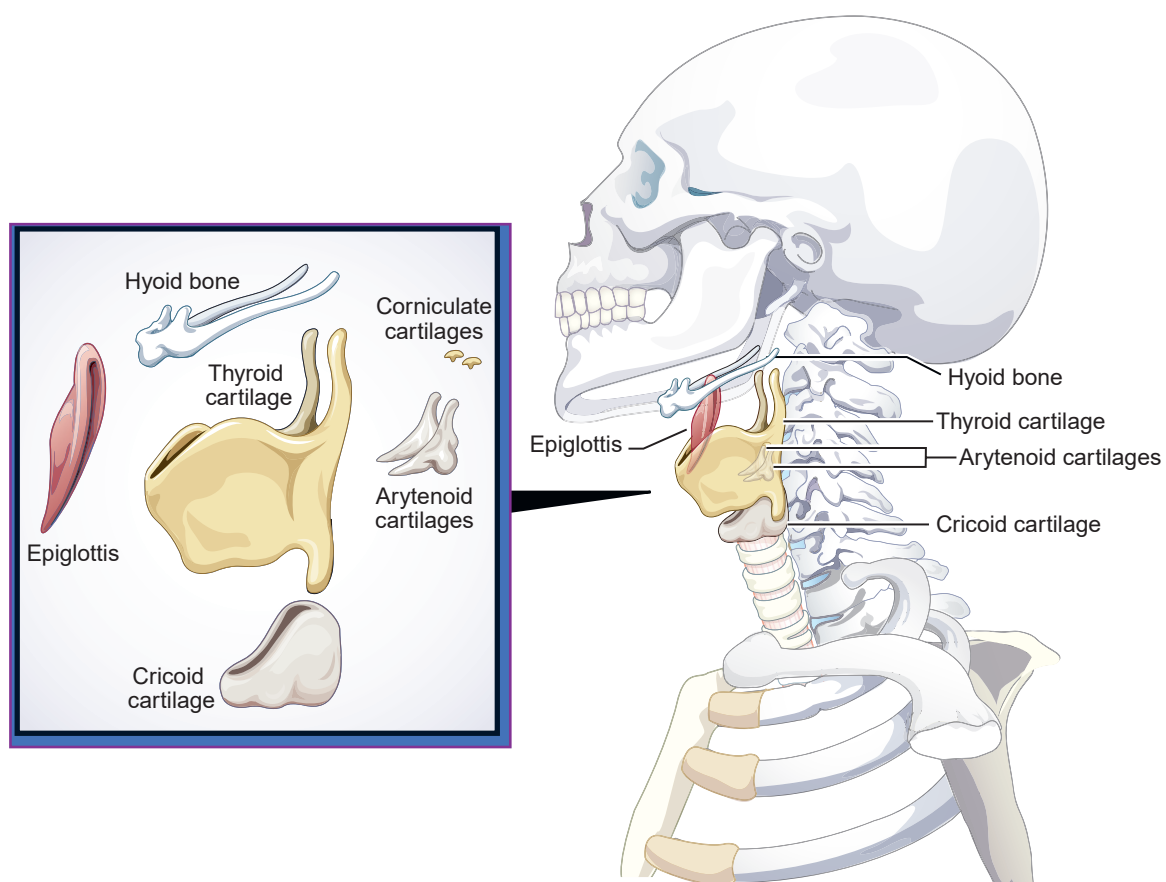


FIGURE 1.3 – Vue latérale de la structure osseuse et cartilagineuse du larynx (Hixon et al. (2020), Copyright 2005 by Thomas J. Hixon and Jeannette D. Hoit).

L'os hyoïde est un os en forme de U. C'est le seul os du corps humain qui ne s'insère dans aucun autre os. Cet os est aussi désigné sous le terme de « suspenseur du larynx » car les structures cartilagineuses qui forment le larynx sont suspendues à l'os hyoïde par la membrane thyro-hyoïdienne. Cet os a un rôle important dans la production de la parole car différents muscles linguaux et laryngaux s'y insèrent. L'activation des muscles trouvant une insertion sur l'hyoïde modifie la position de celui-ci et donc la position du larynx.

*Le cartilage thyroïde*, nous l'avons dit, est relié à l'hyoïde par une membrane. Deux faces composent ce cartilage et celles-ci se rejoignent en avant pour former la proéminence thyroïdienne (pomme d'adam / d'ève). La partie inférieure du thyroïde s'articule avec la face supérieure du cartilage cricoïde. *Le cartilage cricoïde* constitue la partie inférieure du larynx, il se trouve au-dessus du premier anneau trachéal. Sur sa partie postérieure s'insèrent les aryténoïdes.

*Les aryténoïdes* sont deux cartilages symétriques mobiles en forme de pyramide, ils s'insèrent sur la partie postérieure du *cricoïde*. La face antérieure des aryténoïdes forme le processus vocal sur lesquelles vont s'insérer les plis vocaux. La face postérieure forme le processus musculaires où s'insèrent des muscles laryngés intrinsèques qui permettent l'adduction (fermeture) ou l'abduction (ouverture) de la glotte. Ces cartilages sont de première importance dans les gestes d'ouverture et de fermeture de la glotte et de la mise en vibration des plis vocaux.

*L'épiglotte* est un cartilage de forme ovale. L'apex se situe au niveau de la racine de la langue et le cartilage s'attache vers le bas sur la face antérieure supérieure du cartilage thyroïde. Lors de la déglutition, l'épiglotte bascule en arrière pour protéger les voies respiratoires.

### 1.1.2.2 Muscles Laryngaux

Le larynx se compose de différents muscles que l'on distingue selon qu'ils sont intrinsèques, c'est-à-dire qu'ils s'insèrent sur les cartilages laryngés, ou extrinsèques, c'est-à-dire ceux qui s'insèrent sur le larynx et sur d'autres structures comme le sternum ou la mâchoire par exemple.

#### 1.1.2.2.1 Muscles Intrinsèques

Les muscles intrinsèques s'insèrent sur les cartilages laryngés. Ils contrôlent le voisement et la fréquence fondamentale. On distingue plusieurs types de muscles que l'on classe en fonction de leurs actions : les muscles tenseurs, les muscles adducteurs et les muscles abducteurs.

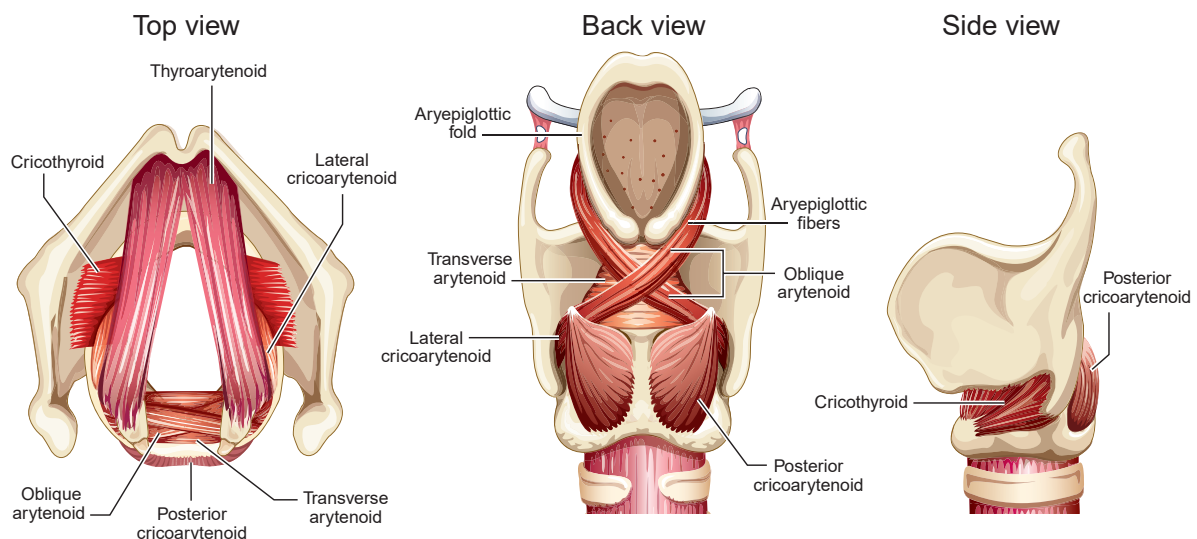


FIGURE 1.4 – Illustration des muscles laryngés intrinsèques (Hixon et al. (2020), Copyright 2005 by Thomas J. Hixon and Jeannette D. Hoit).

*Les muscles tenseurs* sont les muscles dont l'action augmente la tension musculaire des plis vocaux. *Le thyro-aryténoïdien* (Figure 1.4) est le muscle des plis vocaux. Il s'insère en avant sur la face interne du thyroïde et en arrière sur le processus vocal de chaque aryténoïde. *Le crico-thyroïdien* (Figure 1.4) est un muscle s'insérant sur le cricoïde et le thyroïde. Ce muscle



joue un rôle crucial dans le contrôle de la fréquence fondamentale. L'activation de ce muscle peut provoquer un abaissement du thyroïde sur le cricoïde et une bascule vers l'avant du cartilage thyroïde. Cela va augmenter la longueur et la tension des plis vocaux et la fréquence fondamentale (notée F0)

*Les muscles adducteurs* sont les muscles dont l'action ferme la glotte (l'espace entre les plis vocaux). *Le crico-aryténoïdien latéral* (Figure 1.4) est un muscle se trouvant sur la face postérieure du cartilage cricoïde et s'insérant sur les processus musculaires à l'arrière des aryténoïdes. La contraction de ce muscle fait pivoter les aryténoïdes vers l'intérieur et rapproche la partie membraneuse des plis vocaux (les deux tiers antérieurs de la glotte) de la ligne médiane. Son action rend possible le voisement à condition que les plis vocaux ne soient pas trop pressés l'un contre l'autre car cela empêcherait la vibration. L'interaryténoïdien (Figure 1.4) se compose de deux muscles : *l'inter-aryténoïdien transverse* et *l'inter-aryténoïdien oblique*. Le transverse s'insère d'un aryténoïde à l'autre sur un axe horizontal. Les inter-aryténoïdiens obliques s'insèrent en diagonale d'un aryténoïde à l'autre. Leur contraction permet de fermer la glotte cartilagineuse (le tiers postérieur de la glotte) et permet la vibration des plis vocaux.

*Les muscles abducteurs* sont les muscles qui ouvrent la glotte. Le muscle abducteur est *le crico-aryténoïdien postérieur* (Figure 1.4). C'est un muscle qui se trouve sur la face postérieure du cartilage cricoïde et qui s'insère sur les aryténoïdes. Sa contraction provoque un pivotement des aryténoïdes vers l'extérieur qui a pour effet d'éloigner les plis vocaux de la ligne médiane pour arrêter le voisement.

#### 1.1.2.2.2 Muscles Laryngés Extrinsèques

Les muscles extrinsèques du larynx (Figure 1.5) se divisent en trois groupes : les muscles infra-hyoïdiens (abaisseur du larynx), les muscles supra-hyoïdiens (élevateurs du larynx) et les muscles du pharynx (élevateurs et constricteurs)

Les muscles infra-hyoïdiens sont des muscles abaisseurs du larynx, on en compte trois plus le *thyro-hyoïdien* qui peut participer à l'élévation ou l'abaissement. *Le sterno-hyoïdien* et *le sterno-thyroïdien* sont deux muscles qui s'insèrent en bas sur le sternum et en haut respectivement sur l'hyoïde et sur le thyroïde. *L'omo-hyoïdien* s'insère sur l'omoplate et l'hyoïde. En se contractant ces trois muscles abaissent le larynx.

*Le thyro-hyoïdien* est un muscle s'insérant sur le cartilage thyroïde et l'hyoïde. Sa contraction réduit la distance entre le thyroïde et l'os hyoïde. Le thyro-hyoïdien a le potentiel d'assister à l'élévation du larynx lorsque l'hyoïde est fixe ; il peut aussi assister l'abaissement lorsque le cartilage thyroïde est fixe.

Les muscles supra-hyoïdiens (Figure 1.5) sont les muscles élevateurs du larynx. Ils s'attachent à l'hyoïde, à la base du crâne (*Ventre postérieur du digastrique, stylo-hyoïdien*) ou à la mâchoire inférieure (*ventre antérieur du digastrique, mylo-hyoïdien et génio-hyoïdien*). La contraction de ces muscles permet l'élévation du larynx.

Les muscles du pharynx peuvent comprimer et élever le laryngo-pharynx. *Le constricteur inférieur et moyen* permettent l'élévation du larynx en comprimant le laryngo-pharynx et l'oropharynx. *Le stylo-pharyngien* s'insère sur la base du crâne et sur le constricteur moyen et le thyroïde. En se contractant, il élève le pharynx et le larynx et comprime le pharynx. *Le palato-pharyngien* s'insère sur le palais mou et sur le thyroïde, il peut élever le larynx et le pharynx et abaisser le velum. Enfin, *le salpingo-pharyngien* s'insère sur la trompe d'eustache (conduit auditif) et sur le palato-pharyngien. En se contractant, il élève aussi le larynx et le pharynx.

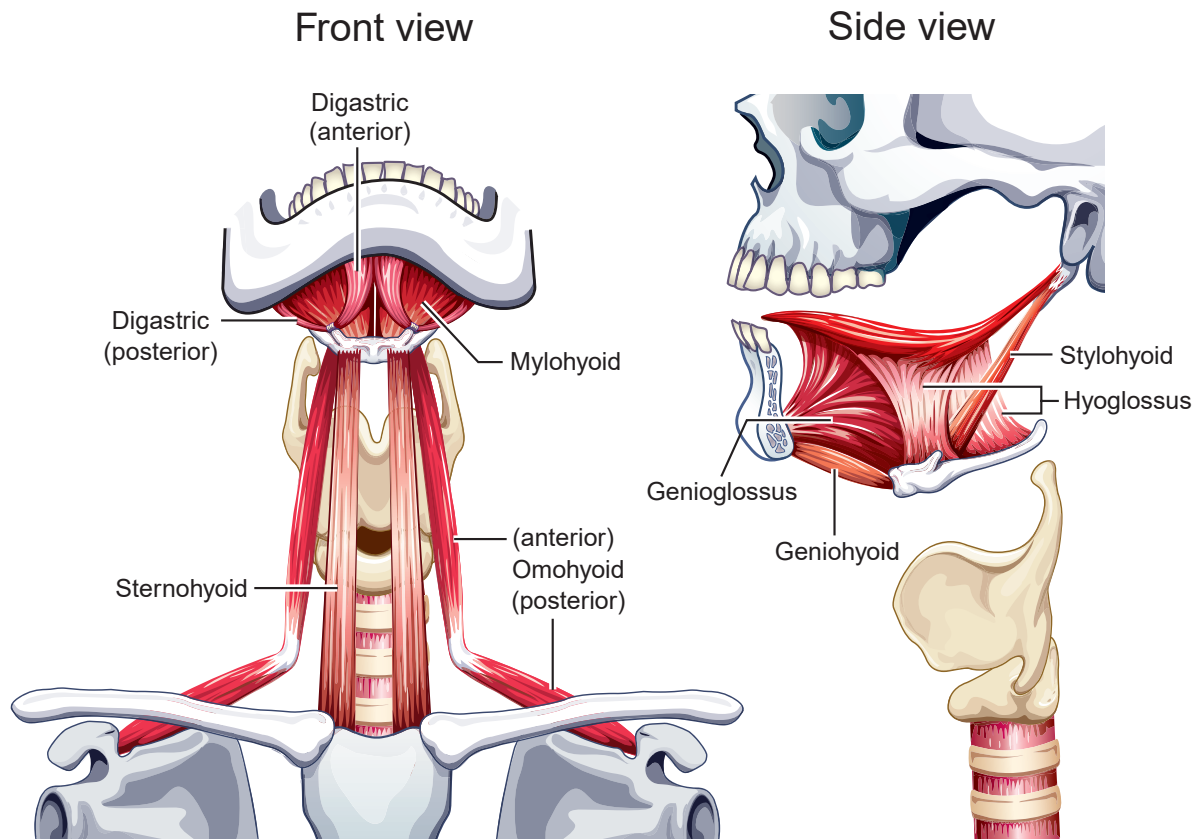


FIGURE 1.5 – Muscles laryngés extrinsèques (Hixon et al. (2020), Copyright 2005 by Thomas J. Hixon and Jeannette D. Hoit).

### 1.1.3 Système oral

Afin de simplifier la description du système oral dans la production des sons dans le conduit vocal, nous incluons ici le système lingual, mandibulaire et labial. Nous les traitons ensemble car ces différents sous-systèmes permettent la production des consonnes et des voyelles.

#### 1.1.3.1 Muscles linguaux

Les muscles du système lingual s'insèrent sur l'os de la mâchoire, sur l'*hyoïde* et sur l'os palatin. On peut diviser les muscles linguaux en deux catégories : les muscles intrinsèques et les muscles extrinsèques. Les muscles intrinsèques sont impliqués dans les changements de la forme de la langue. Les muscles extrinsèques sont impliqués dans le déplacement de la langue (Gick et al., 2013 ; Hixon et al., 2020 ; Marchal, 2011).

Les muscles intrinsèques sont au nombre de 4 : les deux longitudinaux supérieur et inférieur, le muscle transverse et le muscle vertical (Figure 1.6, bas). Ces muscles s'entrecroisent et forme une sorte de « tissage » musculaire lingual. La contraction du *muscle longitudinal* élève la pointe de la langue. Selon Stone et al. (2004), il est impliqué dans les manœuvres de rétroflexion. Le *muscle longitudinal* inférieur a une action antagonique en ce qu'il abaisse la pointe de la langue. En se contractant *le transverse* rapproche les bords latéraux de la langue vers l'intérieur ; *le vertical* a une action antagonique et permet d'éloigner les bords latéraux. Ils permettent de modifier l'épaisseur de la langue.

Les muscles extrinsèques sont le *généioglosse*, le *palatoglosse*, l'*hyoglosse* et le *styloglosse* (Figure 1.6, haut). Le *généioglosse* s'insère sur la mâchoire et dans la langue. C'est un muscle

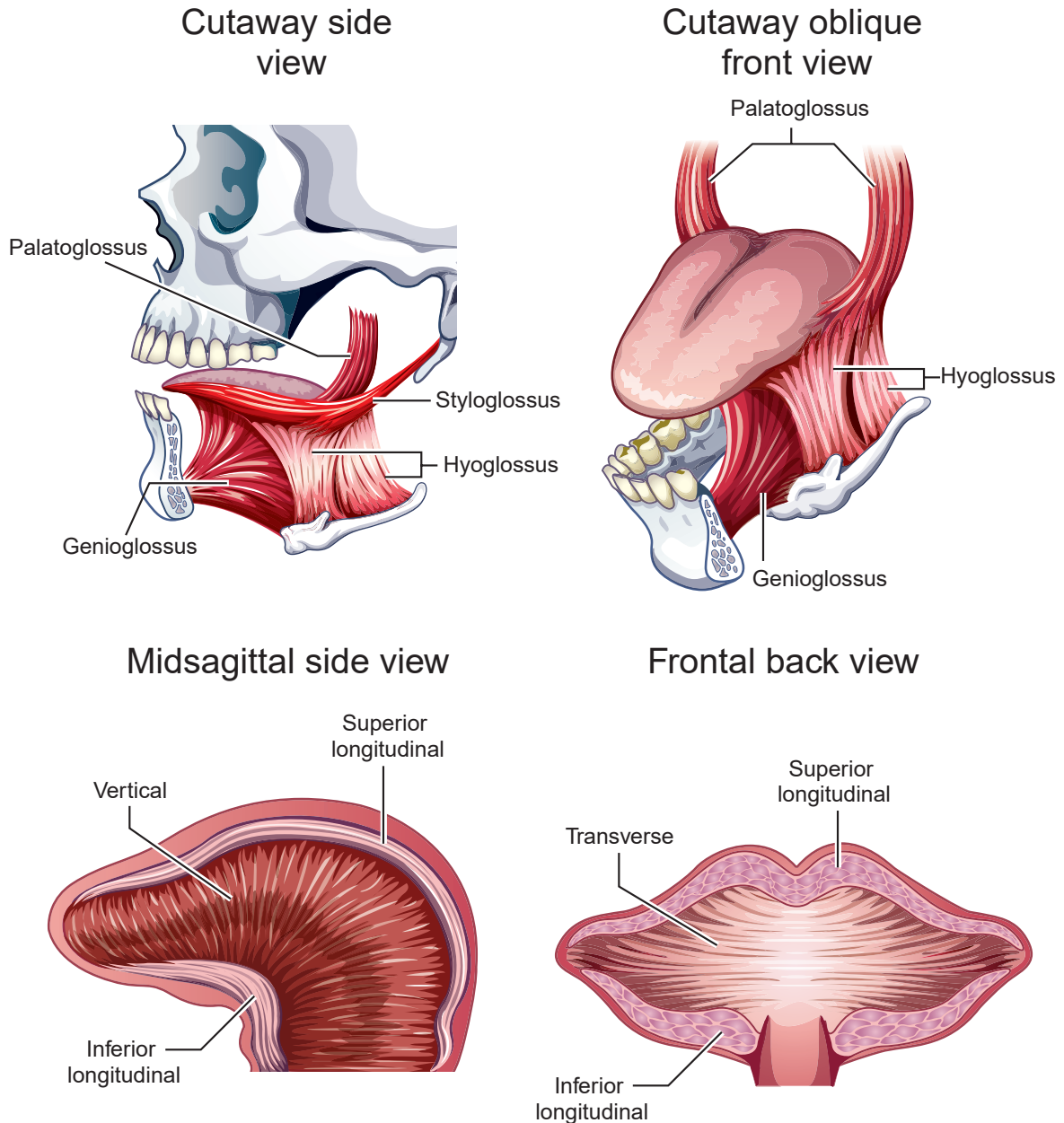


FIGURE 1.6 – Muscles Linguaux (Hixon et al. (2020), Copyright 2005 by Thomas J. Hixon and Jeannette D. Hoit).

complexe car il possède des fibres supérieures, des fibres moyennes et des fibres inférieures. La contraction de ces fibres provoque l'abaissement du corps de la langue et de l'avancement de la racine de la langue. Le *palatoglosse* s'insère dans l'os du palais jusque dans la langue. Lorsque le port vélopharyngé est fermé, il élève le dos de la langue. L'*hyoglosse* s'insère sur l'os hyoïde et dans la langue. Son action est d'abaisser le dos de la langue. Enfin, le *styloglosse* s'insère sur le processus styloïde jusque dans la langue. Il élève le dos de langue vers l'arrière.

### 1.1.3.2 Muscles Faciaux

Les muscles faciaux permettent de mettre en mouvement la mâchoire inférieure et les lèvres. Les muscles de la mâchoire (Figure 1.7) s'insèrent le crâne, la mâchoire inférieure et l'*hyoïde*. Le *masséter*, le *ptérygoïdien interne* et le *temporal* permettent l'élévation et les mouvements latéraux de la mâchoire inférieure. Bien que les mouvements latéraux de la mâchoire ne soient pas pertinents pour la production des sons de la parole, ils le sont pour la production du Human Beatbox comme pour la production de la caisse claire « *K-Snare Inward* » [↓k̂L̂]. Le *ptérygoïdien externe* abaisse la mâchoire. Enfin, quand l'os hyoïde est fixe, la contraction du *génio-hyoïdien*, du *mylohyoïdien* et du ventre antérieur du digastrique permettent l'abaissement de la mâchoire.

Le visage possède aussi de nombreux muscles qui agissent sur le mouvement des lèvres. Ils s'insèrent sur le crâne et la mâchoire (Figure 1.7). Nous allons mentionner les principaux muscles impliqués dans la production des gestes labiaux. L'*orbiculaire* est un muscle de forme ovale dont les fibres forment le pourtour des lèvres. Il s'agit d'un muscle complexe composé de fibres superficielles et profondes. Ce muscle est impliqué dans l'arrondissement et la fermeture des lèvres ainsi que dans la création des constrictions labiodentales. Les *grand et petit zygomatiques* (muscles du sourire sincère) et le *risorius* (muscles du faux sourire) sont impliqués dans l'étirement des lèvres. On peut aussi mentionner le *dépresseur de la lèvre inférieure* dont l'action abaisse la lèvre inférieure et participe donc aux gestes d'ouverture labiale.

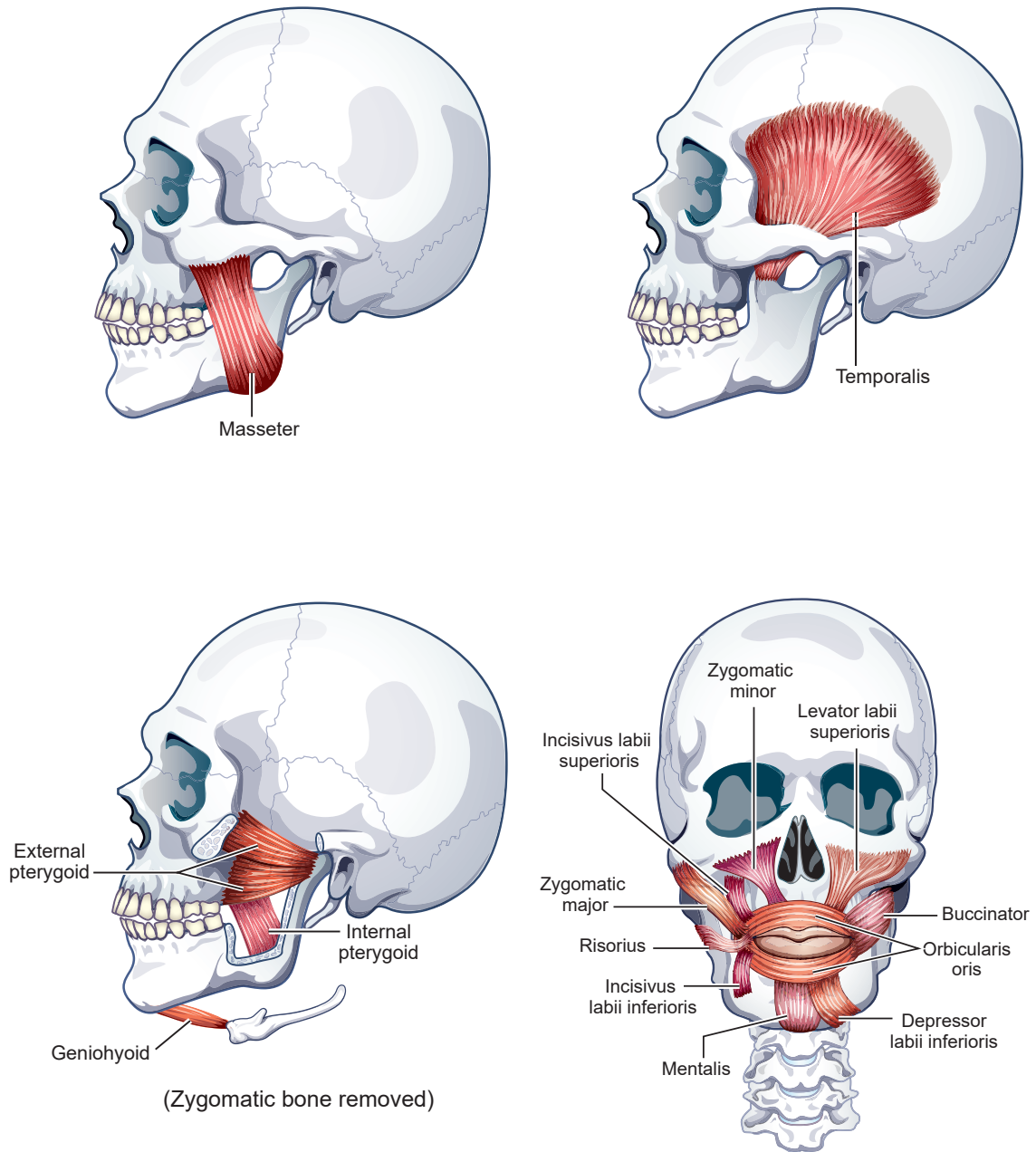


FIGURE 1.7 – Muscles faciaux : muscles de la mâchoire (images 1-3) et muscles labiaux (image 4) (Hixon et al. (2020), Copyright 2005 by Thomas J. Hixon and Jeannette D. Hoit).

## 1.2 Mécanismes de production des sons

Le système de production vocale décrit dans la partie précédente n'a pas pour fonction principale de produire des sons. Cependant, lors de la production vocale, qu'il s'agisse de la parole ou du HBB, l'activité de ces systèmes est réorganisée pour permettre la production des sons dans le conduit vocal. La production vocale se caractérise par une phase musculaire durant laquelle la contraction des muscles respiratoires, laryngaux et oraux vont mettre en mouvement les organes de production pour créer des variations de pression et générer un flux d'air sur lequel on va articuler pour produire des perturbations acoustiques, c'est-à-dire des sons. Nous allons introduire en premier lieu les principes aérodynamiques généraux. Nous présenterons ensuite l'état des connaissances sur les phases d'initiation et d'articulation dans la parole et le HBB. Enfin, nous évoquerons les études acoustiques sur le HBB.

### 1.2.1 Principes Aérodynamiques

L'aérodynamique est une branche de la mécanique des fluides qui étudie le mouvement de l'air. Bien que l'aérodynamique soit une technique de moins en moins utilisée en phonétique, elle n'en reste pas moins importante pour comprendre la production des sons dans le conduit vocal.

#### 1.2.1.1 Pression et volume

L'air est un gaz, il a une masse et il est compressible. Dans le conduit vocal, l'air occupe l'espace dans trois dimensions, on parle de volume d'air. Le volume varie continuellement pendant la phase musculaire. Les variations de volume entraînent des variations de pression de l'air dans le conduit vocal. Le volume et la pression d'un gaz dépendent de la masse  $n$  du gaz et de la température  $T$  :

$$PV = nRT \quad (1.1)$$

$P$  est la pression, c'est-à-dire la force appliquée par unité surface. Nous utiliserons dans ce travail l'unité internationale de pression : l'hectopascal ( $hPa$ )<sup>4</sup>. Dans l'équation 1.1,  $R$  est la constante universelle du gaz. Ce que dit l'équation, c'est qu'une augmentation de la température ou de la masse entrainera un changement de pression ou de volume. Les sons de la parole sont produits dans un environnement adiabatique, c'est-à-dire un environnement où la température reste constante. Selon la loi de Boyle-Mariotte, pression et volume sont en rapport inverse : « la pression d'une quantité donnée de gaz dont la température reste constante varie inversement à son volume » (Catford, 1977, p.29) :

$$P_1 \times V_1 = P_2 \times V_2 \quad (1.2)$$

$P_1$  est la pression atmosphérique de référence dans un conduit vocal ouvert, soit 1013 hPa.  $V_1$  est le volume du conduit vocal. Le volume des lèvres à la glotte est estimé par Catford à  $140cm^3$ . Si on inclut le volume des poumons, le volume est de  $3000cm^3$ .  $P_2$  correspond à la pression dans un conduit en mouvement, par exemple pendant la production d'un son occlusif.  $V_2$  correspond donc à la réduction du volume pendant la phase de production. L'équation 1.2 permet donc d'estimer les variations relatives de volume.

---

4. L'unité internationale de mesure de la pression est le Pascal (Pa) mais pour des raisons d'ordre de grandeur, il est plus pratique d'employer l'hectopascal en phonétique.

### 1.2.1.2 Vitesse du volume et vitesse des particules

Lorsque l'on comprime de l'air dans un conduit, on diminue le volume et on augmente la pression. L'air se déplace des régions de haute pression vers des régions où la pression est plus basse. C'est le principe d'égalisation de la pression. Le volume d'air se déplace sous forme de flux à une vitesse donnée et dans une direction donnée, c'est ce qu'on appelle la vitesse du volume ou débit d'air. La vitesse du volume est notée  $U$  ( $cm^3/s$ ). Il s'agit de la mesure du volume d'air déplacé par unité de temps. Dans le flux d'air, les particules se déplacent elles aussi à une vitesse et dans une direction donnée, c'est la vitesse des particules. On exprime la vitesse des particules  $v$  en centimètre par seconde ( $cm/s$ ). La vitesse des particules n'est pas la même partout dans le flux. Les particules s'écoulant sur les bords du flux se déplacent moins rapidement que les particules au centre du flux en raison du frottement avec les parois du tube dans lequel elles se déplacent. Il est très difficile de mesurer directement la vitesse des particules car cela nécessiterait des techniques de physique expérimentale avancées (mesure par laser). L'estimation mathématique de la vitesse des particules est difficile car elle dépend de nombreux facteurs comme la pression, la densité de l'air et l'accélération gravitationnelle (Shadle, 2010). Toutefois, Catford donne une équation simplifiée pour estimer la vitesse des particules :

$$v = 412\sqrt{P} \quad (1.3)$$

$v$  est la vitesse moyenne des particules car toutes les particules n'ont pas la même vitesse.  $P$  la pression mesurée dans le conduit. Plus la pression est élevée, plus la vitesse des particules est élevée. La vitesse maximale des particules à la sortie du conduit vocal est estimée par Catford à  $16840$   $cm/s$ . Rapportée en mètre par seconde, la vitesse maximale des particules est inférieure à celle du son  $168m/s < 352m/s$ . Les ondes se déplacent plus vite que la matière. La mesure de la vitesse est importante pour la suite car elle va nous permettre d'estimer l'aire des constriction dans le conduit vocal.

### 1.2.1.3 Aire de la constriction et turbulence aéroacoustique

L'air peut s'écouler dans un conduit de façon « laminaire » ou « turbulente ». Un flux est laminaire lorsque les trajectoires des particules ne se croisent pas et que leur vitesse reste constante. Les voyelles sont par exemple produites avec des flux laminaires. L'air peut aussi s'écouler de façon turbulente, par exemple lorsque celui-ci s'écoule dans une constriction. La turbulence est due au croisement des trajectoires des particules et à des variations soudaines de leur vitesse. Il est possible de savoir si le flux est laminaire ou turbulent en mesurant le *Nombre de Reynolds* ( $Re$ ). Pendant la production de son dans le conduit vocal, le caractère laminaire ou turbulent de l'écoulement de l'air dépend de la présence ou non d'une constriction en un point du conduit. En présence d'une constriction suffisamment étroite, le flux sera turbulent. C'est le cas des sons occlusifs et fricatifs par exemple. Il est possible de mesurer l'aire relative de la constriction  $A$  ( $cm^2/s$ ) à partir des mesures de débit d'air  $U$  ( $cm^3/s$ ) et de la vitesse des particules  $v$  ( $cm/s$ ) :

$$A = \frac{U}{v} \quad (1.4)$$

Lorsque l'air s'écoule dans une constriction, la vitesse des particules augmente et la pression dans la constriction diminue afin que le fluide conserve sa masse et son énergie. L'aire de la constriction influe sur le débit volumique et la vitesse des particules : plus la constriction est petite, plus le débit volumique devient faible. La relation entre pression, vitesse des particules et aire de la constriction est essentielle pour comprendre les caractéristiques acoustiques de la turbulence (Stevens, 1971). L'intensité du bruit de la turbulence dépend de l'interaction entre

*P*, *v* et *A*. Fant (1970) suggère que le niveau de pression sonore de la turbulence (*sound pressure level*) est proportionnel au carré de la vitesse des particules (qui elle même dépend de la pression et de l'aire de constriction). La répartition fréquentielle de l'énergie dépend du lieu de la constriction dans le conduit vocal.

## 1.2.2 Phase Musculaire

La phase musculaire est la phase pendant laquelle les muscles du conduit vocal se mettent en mouvement pour produire des sons. Il faut distinguer 2 types de gestes pendant cette phase. D'un côté, il y a les gestes initiatrices qui vont mettre l'air en mouvement. De l'autre côté, il y a les gestes articulatoires qui modifient la façon dont l'air s'écoule et la façon dont résonne le conduit vocal.

La phase d'initiation va mettre l'air en mouvement grâce à des changements de volume entre le lieu d'initiation et d'articulation. Les changements de volume induisent des changements de pression et peuvent générer des flux d'air dans des directions différentes. L'initiation peut se définir par deux paramètres principaux : la façon dont l'air est mis en mouvement et le lieu où l'air est mis en mouvement. Catford (1977) traduit ces deux paramètres en terme de direction du flux d'air égressif ou ingressif, et de lieu d'initiation, à savoir les poumons, le larynx ou la bouche (Tableau 1.1). Pour expliquer la direction du flux, il mentionne l'initiation par des gestes compressifs qui produisent des flux égressifs (flux s'écoulant vers l'extérieur du conduit vocal) et l'initiation par des gestes raréfactifs qui produisent des flux ingressifs (flux s'écoulant vers l'intérieur du conduit vocal).

TABLE 1.1 – Paramètres d'initiation selon Catford (1977)

Lieu	Direction	
	<i>compressif</i>	<i>raréfactif</i>
poumons	pression pulmonique	succion pulmonique
larynx	pression glottalique	succion glottalique
bouche	pression vélarique	succion vélarique

Concernant les mécanismes pulmoniques, leur fonctionnement est bien connu, notamment dans la production de flux égressif lors de la phase d'expiration. Nous avons vu dans la partie précédente l'anatomie du système respiratoire. L'initiation de flux d'air pulmonique est provoquée par l'action des muscles respiratoires faisant varier le volume des poumons. Les muscles inspirateurs permettent de générer des flux d'air ingressifs. Les muscles expirateurs génèrent des flux égressifs. Les mécanismes pulmoniques initiant des flux égressifs sont très communs dans les langues du monde. En revanche, les langues du monde n'intègrent pas de mécanismes pulmoniques ingressifs dans leurs inventaires phonologiques. Toutefois, Eklund (2008) souligne que ces mécanismes sont présents dans la communication humaine et animale. En ce qui concerne la communication de l'espèce humaine, il n'est pas anodin d'entendre une personne prononcer le mot « *oui* »<sup>5</sup> en inspirant.

Le larynx peut aussi faire office d'organe initiateur lorsque la glotte est fermée. Dans les langues du monde, lorsque qu'un son est produit par ce type de mécanisme, on parle de mécanismes éjectifs (flux égressif) et implatifs (flux ingressif). Lors de la production de ce type de mécanismes les variations de volume ont lieu entre le lieu d'articulation et la glotte fermée.

5. ou l'équivalent dans d'autres langues que le français : *yes, ya ...*



Pour générer un flux d'air laryngal égressif, la diminution du volume entre le lieu d'initiation et d'articulation est opérée par l'activation des muscles supra-hyoïdiens qui a pour effet d'élever le larynx et de diminuer le volume. À l'inverse, lorsque l'initiation est produite par expansion, l'augmentation du volume entre le lieu d'initiation et d'articulation est opérée par l'activation des muscles sous-hyoïdiens ce qui a pour effet de descendre le larynx et donc d'augmenter le volume. De plus, Kingston (1984) a proposé l'hypothèse selon laquelle un recul ou un avancement de la racine de la langue joue un rôle dans les variations de volume et de pression nécessaires à la production des mécanismes éjectifs et implosifs. Cela a été démontré dans plusieurs études IRM pour des implosives en Hendo (Demolin et al., 2002), pour des éjectives en Tigrinya (Hermes et al., 2016) et pour la production de sons beatboxés (Proctor et al., 2013).

Enfin, l'initiation peut avoir lieu dans la bouche. Pour initier un flux d'air dans la bouche, deux constriction sont créées dans le conduit vocal : une constriction antérieure (bilabiale, dentale, palatale) et une constriction postérieure (vélaire, uvulaire). Ce mécanisme apparaît sous le terme de *vélarique* dans la terminologie de Catford. Cependant, Miller et al. (2009) proposent le terme *lingual* car la constriction postérieure n'est pas forcément vélaire, celle-ci varie selon le lieu de la constriction antérieure. Pour générer un flux d'air avec un mécanisme lingual, il est nécessaire de varier le volume entre la constriction antérieure et la constriction postérieure. Dans les langues du monde, seul les mécanismes linguaux initiant des flux ingressifs sont attestés. On les appelle *clics*. De plus, pour produire les clics, on relâche en premier la constriction antérieure puis la constriction postérieure. Les mécanismes d'initiation linguale par compression buccale ne sont pas attestés dans les langues mais ils le sont dans la production du Human Beatbox (Blaylock et al., 2017). Les mécanismes pulmoniques, laryngaux et linguaux sont illustrés sur la Figure 1.8.

Bien que la terminologie de Catford soit motivée et reprise par les chercheurs de la discipline, elle n'est pas adaptée pour 3 raisons. En premier lieu, le terme « *glottalique* » pose problème. Catford (1977, p.247-8) explique dans les notes 1 et 2 les choix motivant l'utilisation du terme *glottalique* dans les années 1940. Il explique que le terme *glottalique* est préférable au terme *glottalisé* employé par Pike et plus généralement employé dans la terminologie américaine. Il indique que ces deux termes font référence à deux phénomènes différents : la phonation et l'initiation. En ce point, Catford a raison, cependant des études en IRM (Demolin et al., 2002 ; Hermes et al., 2016) montrent que l'initiation *glottalique* consiste en une combinaison de gestes laryngaux pour manipuler la pression : élévation/abaissement du larynx, rétraction/protraction de la racine de la langue et, dans certains cas, compression/expansion du pharynx. Nous proposons, et utiliserons par la suite, le terme d'« initiation laryngale » à la place de « glottalique ». Le terme *glottalique* porte à confusion. Il est commun d'entendre que les éjectives sont produites par des mouvements de la glotte. Or, la glotte ne génère pas de flux d'air étant donné qu'elle peut seulement être fermée par un rapprochement des plis vocaux vers la ligne médiane. Ensuite, le terme *pression*, en particulier *pression pulmonique*, n'est pas satisfaisant et porte aussi à confusion. En effet, il n'est pas clair si Catford fait référence à une pression aérodynamique ou à une pression de type « mécanique » dans quel cas il s'agirait d'une « compression ». Enfin, le modèle que propose Catford ne fait pas référence aux possibles contrastes acoustiques résultant de différents modes d'initiation.

Le Tableau 1.2 propose une reformulation du modèle de Catford en termes de gestes initiatrices. Par analogie à la phase articulatoire, nous définissons l'initiation par deux paramètres : le lieu et le mode d'initiation. Les lieux d'initiation restent les mêmes, nous avons remplacé le terme *glottalique* par *laryngal*. Nous proposons deux modes d'initiation : la compression et l'expansion. Le geste de compression provoque une diminution du volume et une augmentation de la pression. L'air va ensuite se déplacer de la zone où la pression est forte, c'est-à-dire

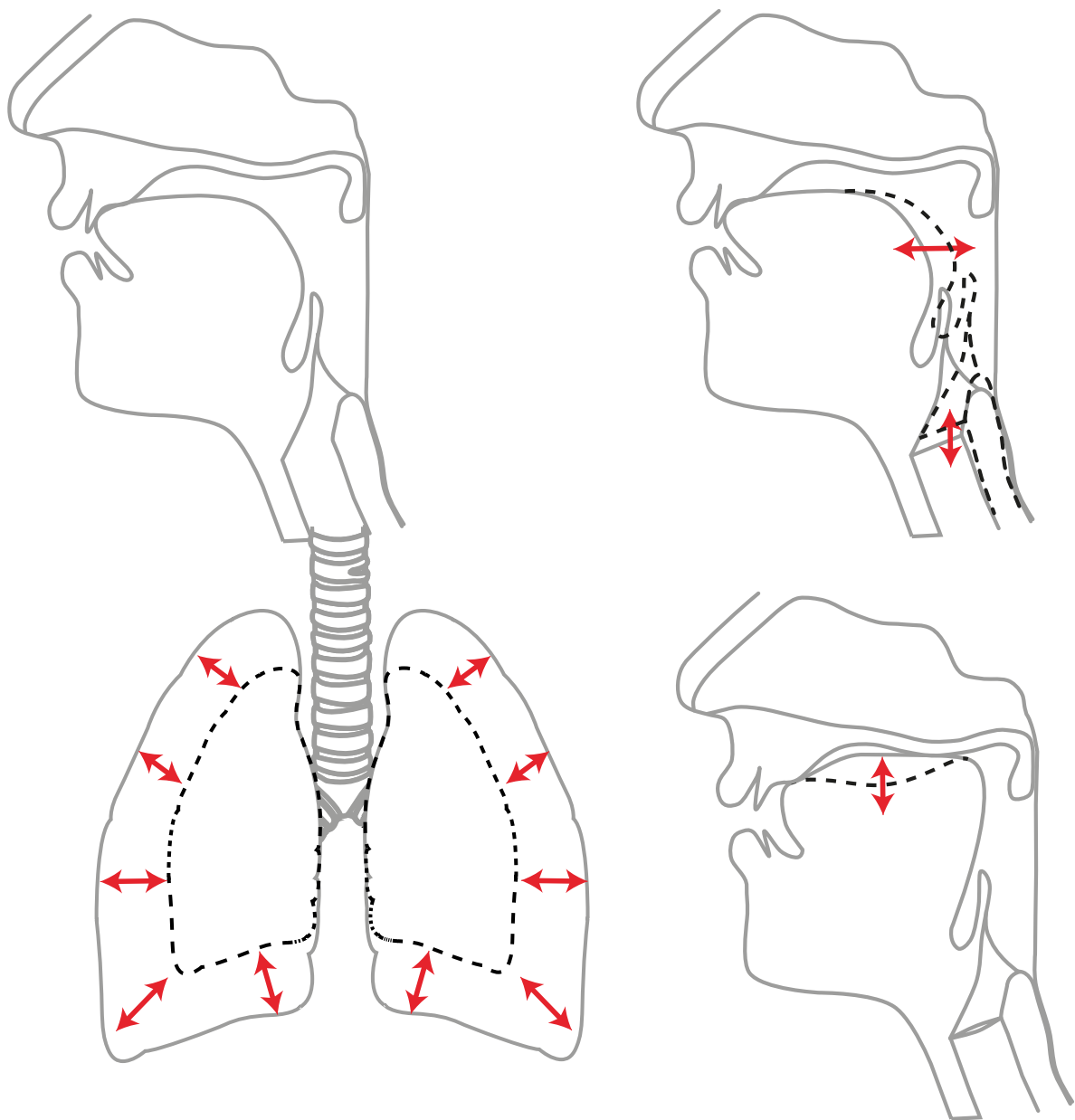


FIGURE 1.8 – Mécanismes d'initiation pulmonique (gauche), laryngal (droite-haut) et lingual (droite-bas). Les flèches indiquent une action de compression des organes initiant des flux d'air égressifs.

à l'intérieur du conduit, vers une zone où la pression est plus faible, c'est-à-dire à l'extérieur du conduit. Les mécanismes d'initiation par compression génèrent des flux égressifs. Le geste d'expansion à une action inverse. Celui-ci provoque une augmentation du volume entre le lieu d'initiation et d'articulation et une diminution de la pression. L'air va cette fois-ci se déplacer de l'extérieur du conduit vocal vers l'intérieur de celui-ci. Les mécanismes d'initiation par expansion génèrent des flux ingressifs. Pour les mécanismes d'initiation linguaux, le flux résultant est décrit comme étant buccal et non comme étant lingual.

TABLE 1.2 – Gestes d'initiation

<b>Lieu</b>	<b>Mode</b>	<b>Description du flux d'air résultant</b>
Poumons	Compression	Pulmonique Égressif
	Expansion	Pulmonique Ingressif
Larynx	Compression	Laryngal Égressif
	Expansion	Laryngal Ingressif
Bouche	Compression	Buccal Égressif
	Expansion	Buccal Ingressif

### 1.2.3 Production du Human Beatbox

Le Human Beatbox reste encore peu étudié et la littérature à ce sujet était assez restreinte jusqu'en 2022 où deux thèses de doctorat portant sur le Human Beatbox ont été soutenues (Blaylock, 2022; Paroni, 2022). Nous allons donc en premier lieu introduire les approches théoriques et proposer la notre. Nous passerons ensuite en revue la littérature sur la production du HBB.

#### Approches théoriques

Dans sa thèse, Paroni (2022) utilise une approche phonétique à partir de 5 sujets masculins. Elle a utilisé diverses méthodes expérimentales, principalement l'électro-magnéto-articulographie (EMA), la pléthysmographie par inductance (pour mesurer les variations d'expansion thoracique) et des enregistrements acoustiques. À partir des données du sujet pilote, elle donne une description de la dynamique articulatoire et ventilatoire de 12 sons beatboxés. Elle propose d'appeler les sons beatboxés des « *boxèmes* », par analogie aux phonèmes. Cependant, elle ne définit ni la nature ni la fonction de ces « *boxèmes* ». L'utilisation d'une machine EMA lui a permis de quantifier la vitesse maximale de déplacement des articulateurs. Ses données montrent que les lèvres se déplacent plus rapidement que la langue et que la mâchoire se déplace plus lentement que la langue. Ses données ne montrent pas de différences majeures entre les vitesses maximales de l'apex et du dos de la langue. Enfin, elle mentionne aussi la technique de *humming*. Le *humming* est une technique de beatbox consistant à produire des imitations de percussion avec des mécanismes d'initiation linguaux (mécanismes des clics) tout en produisant de la phonation. L'utilisation de ces mécanismes permet la dissociation du conduit vocal en deux parties et permet aux beatboxeurs et beatboxeuses d'utiliser leur voix pour produire une ligne mélodique ou rythmique supplémentaire. Ils peuvent aussi utiliser la technique de respiration circulaire<sup>6</sup> au lieu de leur voix. Les données de Paroni (2022) montrent clairement la dissociation du conduit vocal en deux parties, une orale et une pharyngale, et la production simultanée de phonation.

Dans sa thèse, Blaylock (2022) propose une approche cognitive et phonologique du Human Beatbox à partir des données IRM d'un sujet. Ce sujet a produit différents sons et patterns issus de son répertoire. Blaylock (2022) argue que le Human Beatbox possède une phonologie que l'on peut décrire en analysant les contrastes, les alternances et les processus phonologiques (harmonie). Il propose de définir les primitives du HBB en termes de gestes tel que le propose la théorie de la Phonologie Articulatoire (Browman & Goldstein, 1992). Il conclut son analyse des contrastes en indiquant que les sons beatboxés se composent d'un ensemble de constriction contrastives car les sons ont un sens « musical ». Il propose aussi une analyse de la technique de *humming*. Son analyse des alternances repose sur des sons produits avec *humming* et produits sans *humming*. Il interprète le *humming* comme un phénomène d'harmonie « dorsale », en référence au geste de constriction postérieur du mécanisme lingual. L'utilisation d'une approche cognitive s'insérant dans un cadre théorique phonologique est discutable. Nous reviendrons sur cette approche dans le Chapitre 5 lorsque l'on discutera de l'approche anthropophonique des phénomènes phonétiques.

Au long de cette thèse, nous utiliserons une approche « *paralinguistique* » adaptée de l'étude de Proctor et al. (2013). Dans leur étude des mécanismes de production du HBB, Proctor et al. (2013) s'appuient sur les données IRM d'un beatboxeur professionnel. L'objectif de l'étude

---

6. La respiration circulaire permet d'allonger la durée du souffle expiratoire chez les musiciens jouant des instruments à vent. Cette technique de respiration circulaire repose sur la dissociation entre cavité orale et cavité pharyngale et nasale.

était de (1) documenter et d'examiner les sons beatboxés, (2) comparer les caractéristiques des sons beatboxés avec les caractéristiques que l'on connaît des sons parlés et (3) proposer un système de notation du HBB. Leurs données suggèrent que « même si les buts [de production] [...] sont extralinguistiques, les locuteurs vont mobiliser les patrons de coordination articuloire des [systèmes phonologiques] des langues ». En effet, cette première étude en IRM a démontré l'utilisation d'un nombre important de mécanismes rares que l'on peut trouver dans les systèmes phonologiques (clics, éjectives). Ils notent que le Human Beatbox « présente beaucoup de propriétés d'un système combinatoire discret [...] que l'on peut à terme réduire à des structures coordinatives d'un ensemble réduit de primitives » (traduction personnelle). Que ces mécanismes soient attestés ou non dans les inventaires phonétiques des langues du monde, nous pensons en effet que le Human Beatbox présente les propriétés combinatoires d'un système discret. Nous ne parlerons pas de *boxèmes* ou de *gestes articulatoires* tel que proposé par la Phonologie articuloire; nous parlerons de sons beatboxés. Nous plaçons le Human Beatbox parmi les systèmes musicaux et nous considérons les sons beatboxés comme des notes de musique. Nous utiliserons un cadre hybride, c'est-à-dire que nous analyserons un système musical avec les méthodes de la phonétique expérimentale et la terminologie à la fois musicale et phonétique.

## État de l'art

Le Human Beatbox est un système sonore riche et complexe. Dans le HBB, on retrouve les sonorités de divers instruments de musique, de sons synthétiques et électroniques, des vocalisations d'animaux ou d'insectes ou encore des bruits divers (bruits de moteur, sirène de police). Pour acquérir ces sonorités, les beatboxeurs et les beatboxeuses explorent les capacités de leur conduit vocal pour se constituer un répertoire.

Les sons percussifs (imitations de grosses caisses, charlestons, caisses claires) se caractérisent par des stratégies articulatoires similaires à la production des consonnes. Cette catégorie de sons beatboxés se compose en grande majorité de mécanismes occlusifs, affriqués ou de trilles. Les imitations d'instruments à vent ou à corde (trompette, guitare) ainsi que les imitations de sons électroniques (e.g; basses, scratch) se caractérisent par des mécanismes phona-taires ou des trilles. On peut par exemple avoir la *tongue bass* produite comme [r] mais on peut aussi avoir des basses produites avec la voix comme la *Throat Bass*<sup>7</sup>.

Afin d'étudier les caractéristiques phonétiques du HBB, différentes techniques ont été utilisées. On retrouve en majorité des études s'intéressant aux caractéristiques articulatoires du HBB à partir d'imagerie par résonance magnétique et laryngoscopie. Il y a aussi quelques études s'intéressant aux caractéristiques aérodynamiques du beatbox. Concernant les caractéristiques acoustiques, il y a encore très peu d'études à ce sujet. Aucune étude en perception n'existe à notre connaissance.

Les méthodes d'investigation des mécanismes supralaryngaux sont l'Imagerie par Résonance Magnétique (IRM) (Blaylock, 2022; Blaylock et al., 2017; Patil et al., 2017; Proctor et al., 2013), l'Ultrason ou échographie linguale (Paroni, 2014) et l'électro-magnéto articulographie (EMA) (Paroni, 2022; Paroni et al., 2021). Une étude de l'activité électromyographique des lèvres a aussi été proposée par Paroni et al. (2022).

L'étude de Proctor et al. (2013) est pionnière dans la mesure où c'est la première étude décrivant minutieusement la production du HBB avec des données IRM permettant de visualiser le conduit vocal de la glotte jusqu'aux lèvres. Ils indiquent que le beatboxeur réutilise des mécanismes de production similaires aux mécanismes que l'on trouve dans les langues du monde

---

7. Pour la *Throat Bass*, se reporter au Chapitre 3

comme par exemple les éjectives [p'] ou les clics [[]]. Blaylock et al. (2017) soulignent que les mécanismes de production des artistes ne se bornent pas seulement à l'utilisation de mécanismes attestés dans les systèmes linguistiques. Il montre par exemple la combinaison d'une trille bilabiale avec un mécanisme d'expansion buccale (flux d'air buccal ingressif). Ce son est un *Lips Roll Humming*, nous l'avons par ailleurs inclus dans notre corpus.

Les études de Paroni (2022) et Paroni et al. (2022), Paroni et al. (2021) constituent un précieux apport à notre travail de recherche. En analysant la dynamique articulaire, ventilatoire et électromyographique les auteurs montrent une réorganisation des patrons de coordination articulaire en HBB par rapport aux routines articulaires utilisées en français. Par exemple, Lisa Paroni démontre la participation de la langue dans le mécanisme d'initiation de la grosse caisse bilabiale de base : [p']. Les données électromyographiques des lèvres de Paroni et al. (2022) montrent une activité accrue des lèvres pendant [p'] (beatboxé) et [p] (parlé). Toutefois, il est surprenant de voir que cette étude s'est intéressée à la production des sons [p t k] beatboxés et parlés alors que leurs données EMA ne montrent pas de participation des lèvres pour produire [t k] en parlant ou en beatboxant.

Plusieurs études se sont intéressées aux comportements laryngés pendant la production du Human Beatbox, principalement en utilisant la fibroscopie laryngée (De Torcy et al., 2014; Dehais-Underdown et al., 2019; Saphavee et al., 2014). Les beatboxeurs exploitent une part importante des capacités laryngées pour produire des sonorités musicales. De Torcy et al. (2014) montrent la production d'une variété importante de mécanismes laryngaux. Les artistes du HBB produisent des constrictions en différents points du larynx : constrictions glottales, ventriculaires ou encore aryépiglottiques. Les auteurs notent aussi des différences de configuration du tube épilaryngé et de la cavité pharyngale.

Saphavee et al. (2014) rapportent des comportements similaires chez des artistes de beatbox. Ils ont noté que les mécanismes de production ne reposent pas principalement sur des comportements phonatoires et sur la mise en vibration des plis vocaux. Ils soulignent l'agilité des stratégies de production des beatboxeurs qui utilisent de façon fine et indépendante différentes structures du système laryngal. Ces études soulignent l'absence de pathologie vocale ou de lésions des structures laryngales. Cela peut suggérer que, bien que les mécanismes de production soient extrêmes, la pratique du HBB n'est pas dangereuse pour la santé.

On trouve quelques études s'intéressant aux caractéristiques aérodynamiques de la production du HBB (Bourdin & Navion, 2013; Dehais-Underdown et al., 2019; Garrigues, 2015). Le mémoire d'orthophonie sur l'effort vocal de Bourdin et Navion (2013) proposait de comparer des productions beatboxées à des productions parlées, chantées et criées. Les autrices démontrent que le débit d'air oral lors de la production de beatbox est comparable aux valeurs de débit d'air de la voix criée. Ces résultats supposent que les beatboxeurs et les beatboxeuses utilisent des mécanismes plus puissants que dans la parole.

Le mémoire d'orthophonie de Garrigues (2015) est pionnier dans l'étude des caractéristiques aérodynamiques grâce à des mesures de pression intra-orale et de débit d'air oral et nasal pendant la production du HBB. Ce travail a montré l'utilisation de mécanismes articulaires et initiativaires variés dans la production de 7 sons : deux types de grosses caisses, trois types de caisses claires, un charleston et une guitare électrique. Ce mémoire rapporte l'utilisation de mécanismes d'initiation éjectifs, implusif et lingual (imitation de percussion de batterie) ainsi que la production de phonation pulmonique ingressive et nasale (imitation de guitare électrique). Cependant, cette étude s'appuie seulement sur des données de pression buccale qui permettent seulement d'analyser la pression intra-orale des sons produits par des constrictions labiales. De plus, l'étude porte seulement sur un seul sujet.

Concernant l'acoustique du beatbox, on trouve très peu d'études à ce sujet. Pillot-Loiseau

et al. (2020) ont comparé des sons de trompette avec des imitations de trompette beatboxée. En comparant la fréquence fondamentale et la répartition spectrale de l'énergie, l'auteur a montré de fortes similitudes acoustiques entre les sons de trompette et de beatbox. Une autre étude de Yeshoda et Raveendran (2021) s'est intéressée à la durée et l'intensité acoustique d'un petit échantillon de sons beatboxés comme le *Classic Kick drum* [p'], la *Throat Bass* ou une « oscillation uvulaire ». Les résultats montrent une intensité importante des sons (généralement > 80dB) et concordent avec les résultats de Bourdin et Navion (2013) sur la similitude des sons beatboxés à la voix criée. Ces résultats supposent une forte pression dans le conduit vocal pour générer des débits d'air et une intensité vocale aussi importante.

## 1.3 Vitesse de production

Nous venons d'introduire les mécanismes de production des sons dans le conduit vocal. Comment varie la production des sons selon la vitesse de production ? La vitesse de production est un terme regroupant ce qu'on appelle « débit de parole » et « tempo » en musique. Il faut entendre « vitesse de production » comme le nombre d'unités (syllabes, mots, notes de musique) produites par unité de temps. En musique le tempo peut s'exprimer en battement par minutes (BPM).

Nous allons en premier lieu parler de la variation de réalisation des gestes de production de la parole en fonction de la vitesse de production. Enfin, nous présenterons des études sur les effets du tempo sur la production musicale.

### 1.3.1 Effet du débit de parole sur la production

L'invention du *sonographe* par les laboratoires Bell dans les années 1940 et la *Théorie acoustique de la production de la parole* (Fant, 1970) ont permis un grand bond en avant de la recherche en phonétique expérimentale pendant la deuxième moitié du 20ème siècle. Pendant cette évolution, les phonéticiens se sont intéressés à la variation de la production de la parole en fonction de la vitesse de production. L'étude de Lindblom (1963) est une des premières études expérimentales, si ce n'est la première, à s'intéresser aux effets de la vitesse de production sur les sons de la parole. Il propose d'analyser les variations des 2 premiers formants (F1 et F2) des voyelles du suédois dans différents contextes accentuels, phonétiques et à différentes vitesses. Il montre que l'accélération de la production se traduit par une diminution de la durée des voyelles et un abaissement du F2. Il propose aussi une hypothèse de travail sur les variations de la vitesse de production et le lien avec les phénomènes de réduction : « En supposant que les événements nerveux correspondant aux phonèmes sont invariants, les organes de parole échouent, en raison des limites physiologiques, à atteindre les positions supposées ». Lindblom formule ici l'hypothèse que les phénomènes de réduction, d'assimilation ou d'élision, sont dus à des limites biomécaniques du système de production. Il explique que les cibles de production sont invariantes mais que le mouvement et le timing des articulateurs est variable. Il pose donc la question du compromis spatio-temporel de réalisation des gestes en fonction de la vitesse de production.

Des études postérieures se sont intéressées aux effets du débit de parole sur l'organisation temporelle des gestes de production mais aussi sur les mouvements articulatoires. Parler vite revient à produire plus de syllabes ou de mots sur un temps donné. À l'inverse, parler lentement revient à produire moins de syllabes ou de mots sur un même laps de temps. Lorsque la vitesse de production augmente la durée des sons tend à se réduire. Ceci aussi bien pour les consonnes que les voyelles et les diphtongues (Byrd & Tan, 1996 ; Gay, 1968, 1981 ; Gay et al., 1974 ;

Jannedy et al., 2010; Lindblom, 1963; Munhall & Löfqvist, 1992; Ostry & Munhall, 1985; Van Son & Pols, 1999). Les différentes études s'accordent à rejeter l'hypothèse d'une compression horizontale de la durée des gestes de production. Les effets du débit de parole sur la durée des gestes de production sont non-linéaires, c'est-à-dire que la réduction de la durée d'un son n'est pas proportionnelle à la vitesse de production (Gay, 1981; Gay et al., 1974; Ostry & Munhall, 1985). Il a été montré que la réduction de la durée des gestes de production varie selon les sujets, les sons et la position des sons dans les syllabes et les mots (Byrd & Tan, 1996; Gay, 1981).

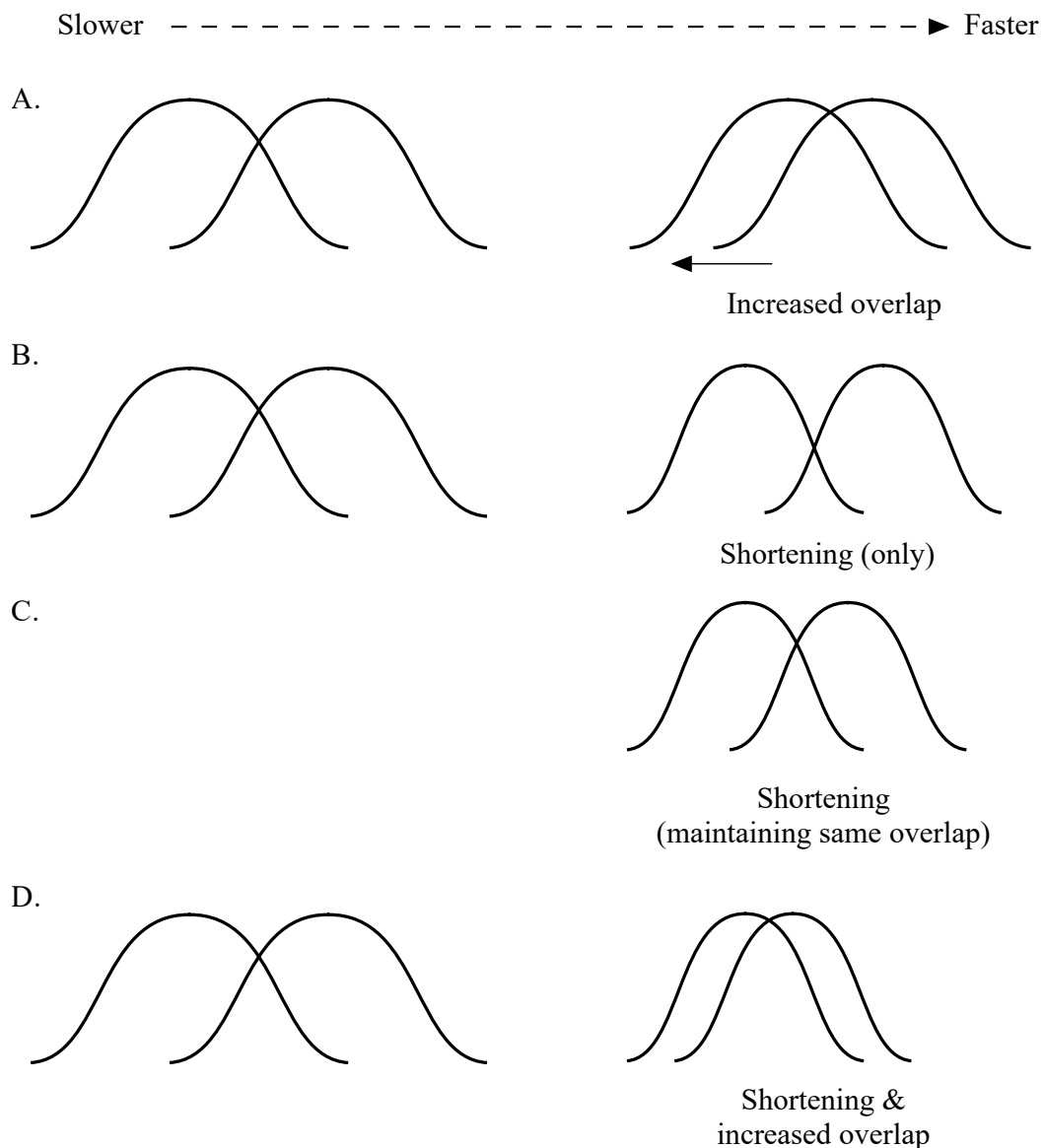


FIGURE 1.9 – Hypothèses des stratégies de réduction temporelle des gestes de Byrd et Tan (1996)

Dans l'étude de Byrd et Tan (1996), les auteurs soulignent que la réorganisation temporelle de la chaîne parlée ne concerne pas seulement la durée des phonèmes mais aussi la durée des transitions entre les gestes (intervalles intergestuels). L'intervalle intergestuel peut lui aussi



varier et se réduire. La réduction de la durée séparant les gestes se traduit par une augmentation du degré de chevauchement temporel des gestes de production. C'est d'ailleurs le degré de chevauchement entre les sons qui donne lieu aux phénomènes de « fusion<sup>8</sup> » ou d'élision des phonèmes. Le chevauchement temporel des gestes n'affecte pas seulement les articulations supralaryngales (voyelles et consonnes) mais aussi les gestes d'ouverture et de fermeture glottique (Munhall & Löfqvist, 1992). Enfin, Byrd et Tan (1996) proposent 4 hypothèses sur les possibles stratégies individuelles de réduction de la durée des sons lorsque le débit de parole augmente (Figure 1.9) :

- A. Les individus peuvent ne pas réduire la durée des sons et augmenter le chevauchement temporel des gestes.
- B. Les individus peuvent réduire la durée des sons et réduire le chevauchement temporel des gestes
- C. Les individus peuvent réduire la durée des sons en maintenant un chevauchement temporel constant
- D. Les individus peuvent réduire la durée des sons et augmenter le chevauchement

Les différentes études se sont aussi efforcées de comprendre et d'expliquer la variabilité des mouvements articulatoires en fonction du débit de parole. Nous venons de voir que parler plus vite implique de réduire la durée des sons ou des transitions entre les sons. Du point de vue musculaire, l'augmentation de la vitesse de production implique que les articulateurs se déplacent plus rapidement : ils accélèrent. Ostry et Munhall (1985) utilisent une approche biomécanique pour expliquer les différentes stratégies cinématiques de production à différentes vitesses. Ils quantifient les effets de la vitesse de production à partir du rapport de la vitesse maximale des articulateurs et de l'amplitude de leur mouvement (trajectoire plus ou moins longue). Ils montrent que lorsque la vitesse de production accélère, la durée des sons diminue. La réduction temporelle s'accompagne d'une augmentation du ratio vitesse/amplitude du geste.

Gay et al. (1974) ont mené une analyse en rayon X et en électromyographie des effets du débit de parole sur les sons labiaux [p w] et les voyelles [i a u]. Ils ont pu montrer différents patrons d'activation EMG et de variation du déplacement des articulateurs. L'utilisation des lèvres comme articulateurs actifs indique une activation musculaire intense de l'*orbiculaire* et une vitesse accrue du mouvement des lèvres à débit rapide. Concernant les mouvements de la langue pour la production des voyelles [i a u], l'étude montre une activation moins importante du *généioglosse*, en comparaison à l'*orbiculaire*, et une vitesse moins importante de la langue. La mâchoire se déplace plus lentement et le mouvement est très sensible au contexte phonétique.

Dans leur étude en électro-magnéto articulographie, Jannedy et al. (2010) ont analysé la production du locuteur qui a gagné le titre de « locuteur le plus rapide d'Allemagne ». Leur étude portait sur la production de virelangues à vitesse normale et rapide. Leur analyse de la vitesse des articulateurs a révélé des vitesses plus faibles pour les consonnes [s] et [k], [s] ayant la vitesse la plus faible, une vitesse intermédiaire pour les consonnes [t], [p], [tʃ] et une vitesse plus importante pour [l]. Les effets du débit de parole sur la trajectoire et l'amplitude des gestes de production dépendent des différents sous-systèmes du conduit vocal (système labial, système lingual).

La littérature des effets de la vitesse de production s'intéresse principalement aux effets sur les consonnes pulmoniques. Une étude de Miller et Elsner (2017) a montré que les clics du Mangetti Dune !Xung étaient aussi sujet à la réduction pendant la parole spontanée. C'est ce

---

8. La fusion est l'addition des caractéristiques de deux sons se chevauchant, elle diffère de l'assimilation phonologique en ce que la fusion des caractéristiques des deux sons n'est pas déclenchée par le contexte environnant.

qu'ont montré leurs analyses de durée et de répartition spectrale. En revanche, nous n'avons pas connaissance d'étude des effets du débit de parole sur les consonnes non-pulmoniques.

Enfin, il convient de mentionner les résultats expérimentaux de Hixon (1966) et Malécot (1969) des effets du débit de parole sur les paramètres aérodynamiques. Ces deux études se sont intéressées aux variations de la pression selon deux conditions : une condition variant le niveau d'intensité (*loudness*) et une condition variant le débit de parole. Hixon (1966) a analysé les valeurs de pression intra-orale, de débit d'air oral et de l'aire de la constriction de [s] chez 9 sujets selon les deux conditions. L'augmentation des niveaux d'intensité sonore se traduit par une augmentation de l'intensité de la turbulence acoustique, de la pression intra-orale et du débit d'air oral. L'aire de la constriction ne varie pas dans cette condition. Dans la deuxième condition, celle du débit de parole, aucun changement n'a été observé concernant la pression intra-orale, le débit d'air et l'aire de la constriction. Malécot (1969) a réalisé une expérimentation similaire en relevant seulement la durée et le pic de pression des consonnes [p b s z]. Il note aussi que l'augmentation de l'intensité vocale provoque une augmentation de la pression intra-orale. Concernant le débit de parole, il note un effet de la vitesse de production sur le pic de pression des fricatives mais pas sur celui le pic de pression des occlusives.

### 1.3.2 Effet du tempo sur la production musicale

Le tempo est un terme musical qui réfère à la vitesse de production d'un morceau de musique. Un morceau de musique est constitué de patrons rythmiques qui se répètent dans le temps. Toutefois, la répétition de ces patrons n'est pas nécessairement régulière (Patel, 2008). Les patrons rythmiques sont créés à partir d'alternances de notes de hauteur, d'intensité, de longueur et de timbre différents. Le tempo résulte de l'organisation temporelle des notes, c'est-à-dire du rapport entre la durée des notes et la distance temporelle qui les séparent. Des notes courtes et rapprochées donnent un tempo rapide par exemple. Des études se sont intéressées aux effets du tempo sur la production musicale de musiciens professionnels.

Des études se sont intéressées plus spécifiquement à des populations de pianistes (Dalla Bella & Palmer, 2011; Goebel & Palmer, 2008, 2013; van Vugt et al., 2014). L'analyse des effets du tempo sur la production de pianistes se fait grâce à des systèmes de capture de mouvements (*motion capture*). En positionnant des capteurs sur les phalanges, les doigts et les touches de piano, ces différentes études ont pu mesurer les variations de la dynamique temporelle et cinématique des doigts à différents tempos. Pour analyser la dynamique temporelle, les études utilisent des mesures de *timing*, c'est-à-dire de synchronisation de la production avec les pulsations du métronome. Pour analyser le *timing*, on mesure l'intervalle de temps entre l'attaque d'une note et l'attaque de la note suivante. En comparant l'intervalle de la production avec l'intervalle de référence, on peut quantifier la déviation temporelle entre l'intervalle du musicien et l'intervalle de référence. On peut ainsi savoir si un musicien joue plus rapidement ou plus lentement que le tempo de référence.

Dalla Bella et Palmer (2011) ont montré que le *timing* de production (par rapport au tempo de référence) devenait plus variable à mesure que le tempo accélérât. Concernant la dynamique cinématique, van Vugt et al. (2014) ont analysé les effets du tempo sur la production d'une population de pianistes à qui l'on a demandé de jouer des gammes descendantes et montantes avec la main droite et la main gauche. L'étude montre qu'à tempo rapide, des contraintes biomécaniques affectent les transitions entre les notes, c'est-à-dire les transitions d'une touche à une autre. En effet, les temps de transitions entre le pouce et l'index ne se réduisent pas en fonction du tempo. En revanche, les résultats ne montrent pas de différences majeures entre les deux mains et la direction de celles-ci pour les gammes descendantes et montantes. Goebel et

Palmer (2008) ont analysé le lien entre le timing de production et le feedback tactile des doigts au contact des touches. L'étude montre qu'à tempo rapide, le degré de contact entre la touche et les doigts augmente et montre une relation positive entre timing de production et le degré de feedback tactile : plus le feedback est important, moins le timing de la production varie.

D'autres études se sont intéressées à des populations de musiciens jouant des instruments à vent (Hofmann & Goebel, 2014 ; Palmer et al., 2009). Les instruments à vent demandent une coordination entre les mouvements des doigts, de la langue et du système respiratoire. Hofmann et Goebel (2014) se sont intéressés aux effets de la variation de tempo sur une population de saxophonistes. Les musiciens devaient produire des mélodies dans 3 conditions : mouvements des doigts, mouvements de la langue et mouvements coordonnés des doigts et de la langue. Les auteurs ont montré des différences de timing selon les trois conditions précédemment citées. À tempo lent, les résultats montrent que le timing est plus précis lorsque les musiciens utilisent les doigts en comparaison à la condition où les musiciens utilisent uniquement la langue. L'utilisation coordonnée de la langue et des doigts indiquent un timing plus précis que dans les deux autres conditions. À tempo rapide, les gestes linguaux ont un timing plus précis que celui des gestes digitaux. La condition de coordination langue-doigt indique un timing plus variable que l'utilisation de la langue. L'étude suggère un bénéfice de la coordination de différents systèmes musculaires dans le contrôle du timing des gestes de production. Enfin, l'étude de Palmer et al. (2009) sur une population de clarinettes confirme les résultats de Goebel et Palmer (2008) suggérant que les informations sensorielles améliorent le timing des gestes digitaux.

## 1.4 Problématique, Objectifs et Hypothèses

Les connaissances que nous avons sur les mécanismes de production des sons dans le conduit vocal viennent en grande majorité d'études linguistiques. La production des phonèmes est soumise à des contraintes phonologiques (inventaire phonémique fini, règles phonotactiques) et biomécaniques. Le Human Beatbox n'est pas un système linguistique mais musical. Par conséquent, les mécanismes de production ne sont pas contraints de la même façon. Étant donné que la production a lieu dans le conduit vocal, les contraintes biomécaniques sont les mêmes pour le HBB et la parole. Cependant, la structure et le fonctionnement du système dépendent de contraintes musicales. Ces contraintes dépendent des genres musicaux produits par les beatboxeurs et les beatboxeuses. Si l'on produit un morceau de *Jazz* ou de *drum n bass*, le choix des sons et leur combinaison dans une structure métrique dépendra des conventions et des règles d'usage du genre imité. Dès lors, les patrons de production, de combinaison et de coordination des unités (qu'il s'agisse de gestes ou de traits) du HBB seront différentes de ce que l'on observe dans les langues. Comment se réorganise la production lorsque l'on beatboxe ? Comment cette réorganisation varie selon le tempo de référence ?

Cette thèse a deux objectifs :

1. En reprenant la proposition de Proctor et al. (2013) nous montrerons que le HBB est un système combinatoire discret. Nous tenterons de démontrer en quoi le HBB est un système de production (articulatoirement) plus complexe que la parole.
2. En reprenant le paradigme de variation de la vitesse de production, nous analyserons les effets du tempo de référence sur la production du HBB à partir de données temporelles, aérodynamiques et acoustiques.

Nous faisons l'hypothèse que les primitives du HBB sont les mêmes que les systèmes phonologiques mais que les deux systèmes diffèrent dans les possibilités combinatoires des primitives qui structurent les unités. Nous faisons aussi l'hypothèse que l'augmentation de la vitesse de

production entrainera une diminution de la durée des gestes de production ainsi qu'une diminution de l'amplitude du geste d'initiation (undershoot). En l'absence de données articulatoires à différents tempo, nous ne faisons pas d'hypothèse sur l'amplitude des gestes articulatoires.

# Chapitre 2

## Méthodologie

« Si vous n’êtes pas capable — à longue échéance — d’expliquer à n’importe qui ce que vous avez fait, votre activité a été inutile. »

---

E. Schrödinger<sup>1</sup>

Avant d’introduire la méthodologie, je souhaiterais expliquer le chemin qui a permis l’élaboration du corpus et du protocole expérimental que je vais exposer dans cette partie ; le HBB constitue donc un nouveau champ d’expérimentation qu’il est intéressant d’explorer.

### 2.1 Premiers Pas

Les premiers mois de mon doctorat ont consisté à regarder des pages internet et des tutoriels vidéos à la recherche des sonorités produites par ces artistes afin d’en dresser un inventaire préliminaire. La tâche était trop ambitieuse, car il existe une quantité impressionnante de sonorités différentes : instruments de musique (percussion, instruments à cordes), sons synthétiques (musique électronique), vocalisations d’animaux ou d’insectes, ou bien des bruits divers (moteur, sirène de police). Rapidement, il a été décidé de créer un corpus des sons répandus dans le HBB que de nombreux artistes seraient en mesure de produire dans le cadre d’un protocole expérimental. Les sons choisis sont des imitations de batterie (grosse caisse, charleston, caisse claire) et des imitations de sons synthétiques (basses vocales, scratch, voix saturées, Lips Roll). La construction du corpus et du protocole s’est faite en collaboration avec Paul Vignes qui a mis à ma disposition, son temps, son réseau de connaissances, son expertise et son talent de beatboxeur professionnel. Il nous a fallu 1 an pour parvenir à un protocole expérimental fonctionnel en vue de déposer une demande d’autorisation d’enregistrements fibroscopiques et aérodynamiques à l’Hôpital Foch (Laryngoscopie, EVA2) auprès d’un Comité de Protection des Personnes. Le protocole initial et le nombre de sujets ont été modifiés et réduits à cause de l’épidémie de Covid-19.

---

1. Schrödinger, E. (1992). *Physique quantique et représentation du monde*. Éditions du Seuil.

## 2.2 Principes de transcription

Le HBB présente un défi certain quant à la transcription phonétique et à la notation musicale. À ce jour il n'existe aucun système consensuel de transcription permettant de rendre compte à la fois des aspects musicaux (rythme, mélodie, timbre) ainsi que des aspects articulatoires. Il convient tout de même de souligner qu'il existe des systèmes d'écriture. Des beatboxeurs ont mis au point la « Standard Beatbox Notation » (SBN) (TyTe & Splinter, 2014). Ce système utilise l'alphabet latin pour noter le HBB. Le « Beatbox Alphabet » de Stowell et Plumbley (2008) utilise des symboles de l'API ainsi que des symboles inventés pour les sons qui n'apparaissent pas dans l'API (Alphabet Phonétique International). Proctor et al. (2013) proposent un système hybride qui lie partition musicale et API. L'API est un frein majeur dans les systèmes d'écriture étant donné qu'il sert seulement aux descriptions linguistiques ; or le HBB n'est pas une langue et fait appel à des sonorités bien plus riches et variées (Blaylock et al., 2017 ; Proctor et al., 2013 ; Stowell & Plumbley, 2008). Enfin, il convient de mentionner le système d'écriture des français Antoine Pinchaud (beatboxeur professionnel) et Adrien Contesse (graphiste) « Vocal Grammatics »<sup>2</sup>. Leur système représente le lieu et le mode d'articulation avec des représentations graphiques similaires aux coupes médiosagittales utilisées en phonétique. Tout comme les notations phonétiques et phonologiques, ce système utilise un nombre de signes finis pouvant se combiner et écrire des séquences beatboxées. Il paraît adapté pour la transcription de sons supralaryngés, en revanche ni la symbolisation de sons laryngaux (basses vocales), ni celle des éléments de notation rythmiques et mélodiques ne semblent intégrés au système.



















Cette étude n'a pas pour objectif de créer un système de notation. Le système mis au point utilise l'API (Association Phonétique Internationale, 1999) ainsi que ses extensions pour la parole pathologique (Ball et al., 2018). Ce système n'a pas vocation à devenir un système de notation en ce qu'il a été mis au point pour les besoins de cette thèse et que celle-ci n'est pas représentative de la richesse et la diversité des sons beatboxés. Le système va nous permettre de représenter les paramètres phonétiques pertinents pour ce travail à savoir :

- Les lieux et modes : symboles des consonnes et des voyelles ;
- La latéralité : symbole diacritique [<sup>l</sup>] ([p<sup>l</sup>]) ;
- La direction du flux d'air : Les flux ingressifs sont notés d'une flèche « ↓ » précédant le son ([↓p<sup>l</sup>]), l'absence de flèches indique que le flux est égressif (ceci afin de ne pas surcharger les transcriptions) ;
- Le lieu d'initiation :
  - Lorsque le flux est initié par les poumons (initiation pulmonique) nous utiliserons les symboles des consonnes pulmoniques de l'API ;
  - Lorsque le flux d'air est initié par le larynx (initiation laryngale) on utilisera les symboles de l'API des consonnes dites « glottiques » ([p<sup>ʔ</sup>, b<sup>ʔ</sup>]) ;
  - Lorsque le flux d'air est initié par la langue (initiation linguale) nous utiliserons des accolades « {} » au début et à la fin des séquences utilisant ce mécanisme d'initiation ({p<sup>̂</sup> ts p<sup>̂</sup> ts}). Les symboles utilisés entre accolades seront les mêmes que les sons pulmoniques, ceci afin de ne pas avoir à inventer une multitude de nouveaux symboles (*Lips Roll Humming* {↓<sub>B</sub><sup>l</sup>})

TABLE 2.1 – Corpus des Beat Patterns (Audio intégré en cliquant sur le bouton lecture. Pour fermer un lecteur et jouer le son suivant : clic droit > désactiver le contenu).

Position	1	2	3	4	5	6	7	8	9	
Mélodie										
BP1	[p'	ts'	↓kL:	ts'	p'	p'	ts'	↓kL:	ts']	
BP2	{p	ts	↓kL:	ts	p	p	ts	↓kL:	ts}	
BP3	[t'	ts'	tʃ':	ts'	t'	t'	ts'	tʃ':	ts']	
BP4	[ḅ'	ts'	ʔf':	ts'	ḅ'	ḅ'	ts'	ʔf':	ts']	
BP5	[ḅ'	ts'	↓ʔf':	ts'	ḅ'	ḅ'	ts'	↓ʔf':	ts']	
BP6	[↓p	ts'	p̣f':	ts'	↓p	↓p	ts'	p̣f':	ts']	
BP7	[p'	ts'	p̣f':	ts'	p'	p'	ts'	p̣f':	ts']	
BP8	{p	ts	p̣f':	ts	p	p	ts	p̣f':	ts}	
BP9	[↓p↓Ḅ¹:	ts'	ʔh:	ts'	↓p	↓p↓Ḅ¹:	ts'	ʔh:	ts']	
BP10	{↓Ḅ¹:	ts	↓kL:	ts	↓Ḅ¹:	↓Ḅ¹:	ts	↓kL:	ts}	
BP11	[p'	↓ṭ:	p̣f':	↓ṭ:	p'	p'	↓ṭ:	p̣f':	↓ṭ:]	

TABLE 2.2 – Corpus des sons. Les sons dits pulmoniques ou laryngaux indiquent que nos données attestent les deux variantes d’initiation du son.

Son	IPA	Catégorie	Articulation		Initiation	
			Mode	Lieu	Mode	Lieu
<i>Classic Kick</i>	[p̰]		occlusif	bilabial	compression	laryngal
<i>Classic Kick Humming</i>	{p}		occlusif	bilabial	compression	lingual
<i>Classic Kick Inward</i>	[↓p]		occlusif	bilabial	expansion	pulmonique ou laryngal
<i>Kick Lingual</i>	[t̰]		occlusif	alvéolaire rétroflexe	compression	laryngal
<i>Throat Kick</i>	[b̰]		occlusif	bilabial	expansion	pulmonique
<i>Hi-Hat Fermé</i>	[ts̰]		affriqués	dental	compression	laryngal
<i>Hi-Hat Fermé Humming</i>	{ts̰}		affriqués	dental	compression	lingual
<i>K-Snare Inward</i>	[↓k̰l̰], [↓k̰l̰]		affriqués	vélaire latéral	expansion	pulmonique ou laryngal
<i>K-Snare Humming</i>	{↓k̰l̰}		affriqués	vélaire latéral	expansion	lingual
<i>PF Snare</i>	[p̰f̰], [p̰f̰]		affriqués	labiodental	compression	pulmonique ou laryngal
<i>PF Snare Humming</i>	{p̰f̰}		affriqués	labiodental	compression	lingual
<i>Tish Snare</i>	[t̰ʰ:]		affriqués	postalvéolaire	compression	laryngal
<i>ESH Snare Outward</i>	[ʔ̰:]		affriqués	laryngal et postalvéolaire	compression	pulmonique
<i>ESH Snare Inward</i>	[↓ʔ̰:]		affriqués	laryngal et postalvéolaire	compression	pulmonique
<i>Cough Snare</i>	[ʔ̰h̰:]		affriqués	laryngal	compression	pulmonique
<i>Lips Roll</i>	[↓b̰:]		trille	bilabial	expansion	pulmonique
<i>Lips Roll Humming</i>	{↓b̰:}		trille	bilabial	expansion	lingual
<i>Clic Roll</i>	[↓t̰:]		trille	rétroflexe	expansion	pulmonique



## 2.3 Corpus

Cette étude s'intéresse à la production de sons beatboxés dans différents contextes et à différents tempo : dans des structures beatboxées (Tableau 2.1), dans des productions à l'isolé (Tableau 2.2), et des productions libres (« *freestyle* »).

Les structures beatboxées sont désignées sous le terme de « Beat Patterns » (BP). Ces BP (Beat Patterns) sont créés à partir d'une même structure métrique, rythmique et mélodique dans le but de recueillir des données contrôlées et comparables entre les sujets. La métrique se compose de 4 pulsations et de 9 sons. Les paramètres rythmiques des Beat Patterns sont des alternances de hauteurs différentes (haut, bas), d'intensité différente (fort, faible) et de longueur différente (court, long). Enfin, les paramètres mélodiques réfèrent ici aux différents timbres propres à chaque instrument et à chaque catégorie de son électronique. La structure métrique, rythmique et mélodique est la même pour chaque pattern, seul le tempo change (Figure 2.1). La métrique se compose d'une suite de 9 sons répartis sur 4 temps de façon isochrone. Pour

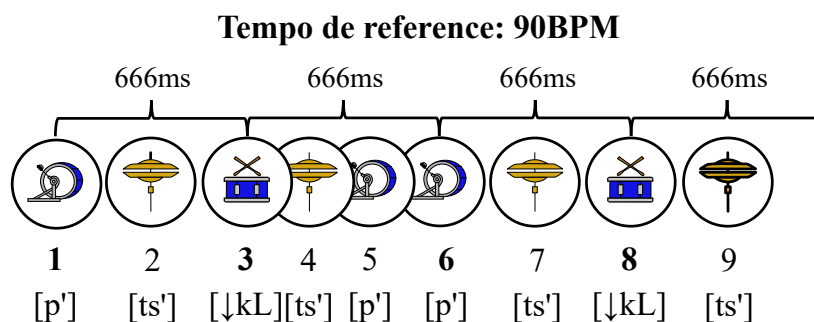


FIGURE 2.1 – Illustration de la structure métrique d'un *Beat Pattern* où le tempo de référence est 90BPM.

chaque instrument, il existe différentes variantes phonétiques. Par exemple, pour imiter une grosse caisse on peut produire [p'] ou [β]; pour imiter une caisse claire on peut produire [↓kL] ou [pf]. La catégorie ne change pas mais le contenu phonétique diffère. Ces variantes sont des variantes esthétiques ou stylistiques qui ne sont pas déterminées par le contexte phonétique adjacent. Nous avons créé 11 Beat Patterns différents en utilisant différentes variantes d'une même catégorie. Le Tableau 2.1 présente une synthèse des 11 BP créés pour les besoins de l'étude. La première ligne indique la position métrique de chaque son dans le pattern. La deuxième ligne indique les caractéristiques rythmiques associées à chaque son. Les paramètres notés en gras indiquent les positions où les 4 battements (ou pulsations) ont lieu, c'est-à-dire sur les grosses caisses et les caisses claires. La troisième ligne indique la structure mélodique des patterns. Les lignes suivantes sont des transcriptions phonétiques des Beat Patterns. Le tableau 2.2 donne ensuite la description phonétique des différentes variantes qui composent les 11 BP.

## 2.4 Déroulement, sujets et protocole de l'étude

La Figure 2.2 illustre le déroulement des séances d'enregistrement. La procédure est la même pour tout le monde et s'est déroulée dans le service d'oto-rhino-laryngologie de l'Hôpital Foch. Les examens ont été effectués par Lise Crevier Buchman (Médecin ORL) tels que

2. <https://www.vocalgrammatics.com/>

définis par le Protocole de recherche de catégorie 2 (ID-RCB n°2020-A00246-33) « Langage et Musique, la parole et le Human Beatbox : Enjeux théoriques pour la recherche en linguistique générale et appliquée ; étude exploratoire mono-centrique » ayant reçu l'avis favorable du Comité de Protection des Personnes – Ile de France IV en date du 15 avril 2020. Conformément à l'avis du comité, une indemnisation financière des artistes a été prévue à hauteur de 25€/h pour 4h d'expérimentation (pauses incluses) soit 100€ au total. En raison du Covid, les enregistrements ont été retardés et ont eu lieu en juillet 2021.

Chaque session d'enregistrement commençait par l'accueil des sujets à l'hôpital Foch par l'équipe de recherche. La mise en place du matériel expérimental se faisait en amont de l'accueil. Une fois l'accueil réalisé, nous avons remis la lettre d'information aux sujets (celle-ci leur avait déjà été envoyée par mail), expliqué la procédure et recueilli leur consentement écrit. Un questionnaire leur a été remis afin de recueillir des informations concernant leur profil socio-professionnel (âge, sexe, métier) mais aussi des informations quant à leur pratique du beatbox (niveau, nombre d'années de pratique).

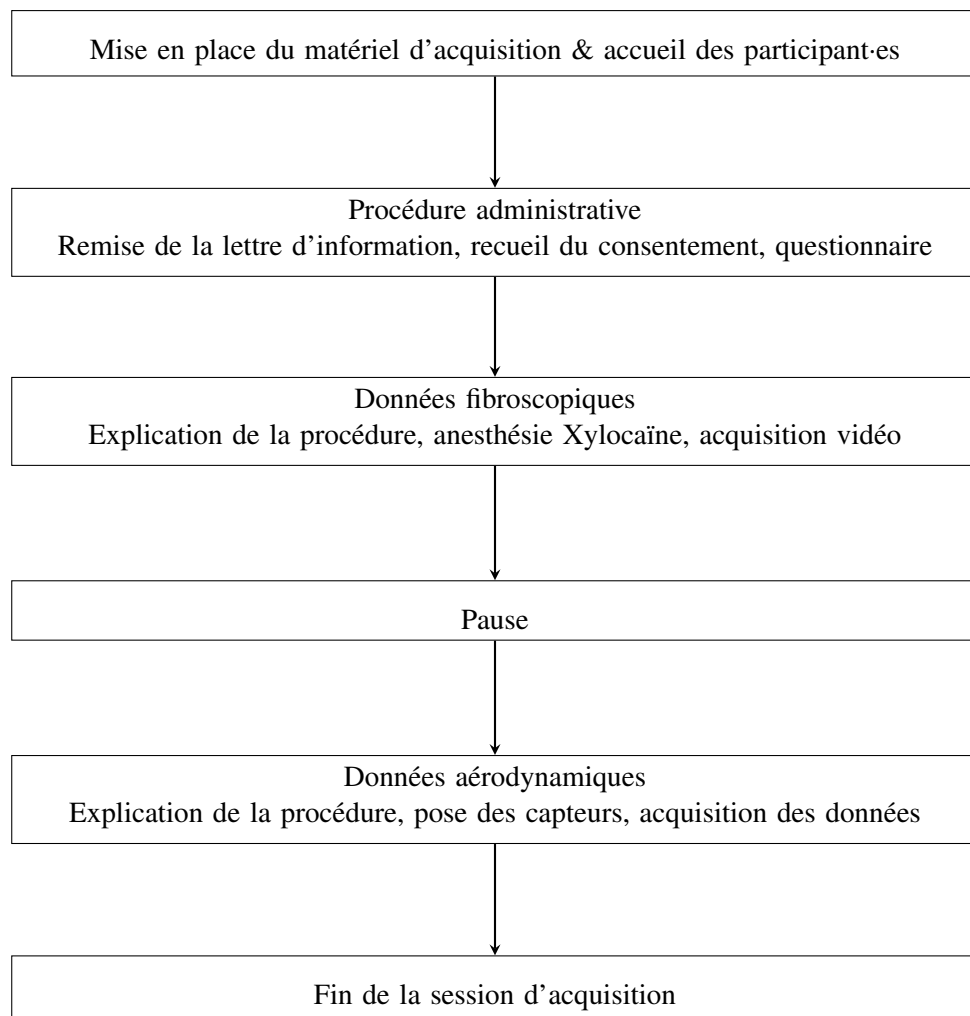


FIGURE 2.2 – Déroulement d'une session d'acquisition

TABLE 2.3 – Profil des participant-es de l'étude

	<b>Sujets</b>	<b>VP</b>	<b>AI</b>	<b>CJ</b>	<b>GA</b>
Sexe		M	F	M	M
Age		32	24	24	25
Niveau de Beatbox		Professionnel	Professionnel	Professionnel	Professionnel
Nombre d'années de pratique du Beatbox		15	11	12	13
Nombre d'années de pratique musicale		15	4	3	4
Entraînement		7j/semaine	3j/semaine	7j/semaine	3j/semaine

## 2.4.1 Sujets

### 2.4.1.1 Recrutement

Le sujet VP a été choisi comme artiste de référence car il avait déjà collaboré avec le laboratoire de phonétique et phonologie dans le cadre du projet européen i-Treasures (FP7-ICT-2011-9, #600676, Dimitropoulos et al., 2014). Les sujets de l'étude sont des beatboxeurs et des beatboxeuses professionnelles recrutées via le réseau du sujet VP faisant partie intégrante de l'équipe de recherche. Le nombre de sujets retenus est de 5 : un artiste de référence et 4 autres artistes (2 hommes, 2 femmes). Pour faire partie de l'étude il faut avoir 18 ans ou plus et ne présenter aucune pathologie vocale ou trouble du langage. Les Femmes enceintes, parturientes ou allaitantes ainsi que toutes personnes privées de liberté ne pouvaient pas participer. Les personnes privées de libertés sont des personnes mineures ou majeures faisant l'objet d'une mesure de protection légale, des personnes hors d'état d'exprimer leur consentement ou des personnes sous tutelle ou curatelle. Le critère d'exclusion était la découverte d'une pathologie vocale lors de l'observation des structures laryngées durant la nasofibroscopie. Les données d'une participante ont été exclues (cf. Section 2.5.2).

### 2.4.1.2 Profil des participant-es

Le Tableau 2.3 dresse le profil des 4 sujets de l'étude (3 hommes, 1 femme) à partir des questionnaires adressés aux sujets<sup>3</sup>. L'âge des sujets est de 32ans pour le sujet pilote VP, de 24 ans pour les sujets AI et CJ et 25 ans pour le sujet GA. Le sujet VP a 15 ans de pratique du Beatbox, les sujets AI, CJ et GA ont respectivement 11, 12 et 13 ans de pratique du HBB. Ils et elles sont toutes des professionnelles du Beatbox. Toutes et tous ont remporté au moins une fois le titre de champion de France de HBB et deux ont remporté un titre mondial. Toutes et tous rapportent avoir eu une pratique musicale autre que le HBB. Le sujet VP indique avoir commencé sa pratique musicale en même temps que le Beatbox à travers la pratique de la guitare, la batterie ou encore du chant. Le sujet AI indique avoir pratiqué le chant et la guitare pendant 4 ans. Le sujet CJ indique avoir pratiqué la batterie 3 ans et enfin le sujet GA indique avoir pratiqué le chant 4 ans. Les sujets AI et GA s'entraînent au Beatbox 3 jours par semaine et les sujets VP et CJ s'entraînent tous les jours.

3. Les questionnaires sont inclus dans les annexes 1 à 5

## 2.4.2 Protocole expérimental

La première partie de l'expérience consistait à répéter les exemples audios des *Beat Patterns* (BP) du corpus. Pour chaque BP, chaque son le composant devait être répété 8 fois à l'isolé; ensuite les sujets devaient répéter 4 fois le BP à 90, 120, 150 battements par minute (BPM). Afin de donner le même tempo de référence à tout le monde, chaque sujet était équipé d'un métronome à vibration placé sur le poignet à la façon d'une montre. Chaque augmentation du tempo de référence était marquée par une courte pause afin de donner la nouvelle référence. Cette 1ère partie du protocole constitue une base commune à toutes les personnes participant à l'étude.

La 2ème partie du protocole s'intéresse à l'exploration du répertoire des artistes concernant des sonorités utilisant la voix. Pour cela, nous avons demandé aux sujets de produire des sons issus de la musique électronique comme les « Scratches » ou les basses. Les « scratches », par analogie au Scratching, sont des sons imitant les DJ (Disc-Jockey) faisant tourner les vinyles sur des platines pour altérer la sonorité des sons et les rendre plus graves ou aigus. Les « basses vocales » sont des imitations de sons synthétiques caractérisés par des basses fréquences; leur production repose sur des stratégies phonatoires complexes. Tous les sujets ont des compétences différentes en terme de style de Beatbox, par conséquent le corpus exploratoire n'est pas le même chez chaque sujet.

Enfin, pour conclure la séance d'enregistrement, il a été demandé aux personnes participant à l'étude de jouer un morceau d'une minute de leur composition personnelle.

Le même protocole a été utilisé pendant l'acquisition des données aérodynamiques, fibroscopiques et d'Imagerie par Résonance Magnétique. Cependant, la résolution temporelle des enregistrements fibroscopiques et IRM étant contrainte à respectivement 25 images/s et 50 images/s, nous avons fait le choix de ne pas demander aux sujets de produire les patterns à trois tempo différents mais seulement à 90 BPM; ceci afin d'alléger le protocole invasif en fibroscopie qui peut présenter de la gêne pour les sujet et afin de raccourcir la durée du protocole pour les enregistrements IRM.

## 2.4.3 Sujet Pilote

Le premier sujet est l'artiste de référence de cette étude, il est aussi l'expert en HBB de l'équipe de recherche. Avant le début de l'étude, nous nous sommes assurés de faire un enregistrement sonore du protocole en chambre sourde afin de vérifier la faisabilité de celui-ci. Cette étude pré-pilote nous a permis de préparer l'étude pilote en condition expérimentale. À l'issue de cette séance, nous avons vérifié la faisabilité, la fiabilité et la reproductibilité du protocole. La faisabilité a été évaluée pour chaque item du protocole de façon binaire : le chanteur arrive à produire l'item ou le chanteur n'arrive pas à produire l'item. La reproductibilité a été évaluée en lien avec la faisabilité mais cette fois sur une échelle à 4 niveaux : 1) aucune difficulté à produire l'item cible (sons, BP), 2) difficultés mineures à produire l'item cible, 3) difficultés majeures à produire l'item cible et 4) impossibilité de produire l'item cible. Tous les items impossibles ont été retirés. Aucun item ne présentait de difficultés majeures. La fiabilité du protocole, a été confirmée par l'analyse des données brutes. Les sujets AI, CJ et GA ont reproduit à l'identique le protocole expérimental tel qu'il a été modifié à la suite de l'étude pilote.

## 2.5 Matériel Expérimental

### 2.5.1 Fibroscopie laryngée

Le Docteur Lise Crevier Buchman a réalisé les examens fibroscopiques avec le système d'enregistrement vidéo stroboscopique numérique Xion (Figure 2.3) qui comprend une source de lumière pour l'endoscope souple et un enregistreur acoustique avec microphone sur la poignée du fibroscope. La distance entre le micro de la caméra et les lèvres du sujet était de 10 cm. Le nasofibroscope souple était muni d'une gaine de protection à usage unique. Il a été introduit par le nez jusqu'à la cavité pharyngale. Une anesthésie locale a été pratiquée par administration de Xylocaïne à 2%. Dès que le nasofibroscope a été positionné, le sujet a réalisé quelques sons pour vérifier la bonne visibilité du larynx puis le protocole a été reproduit.

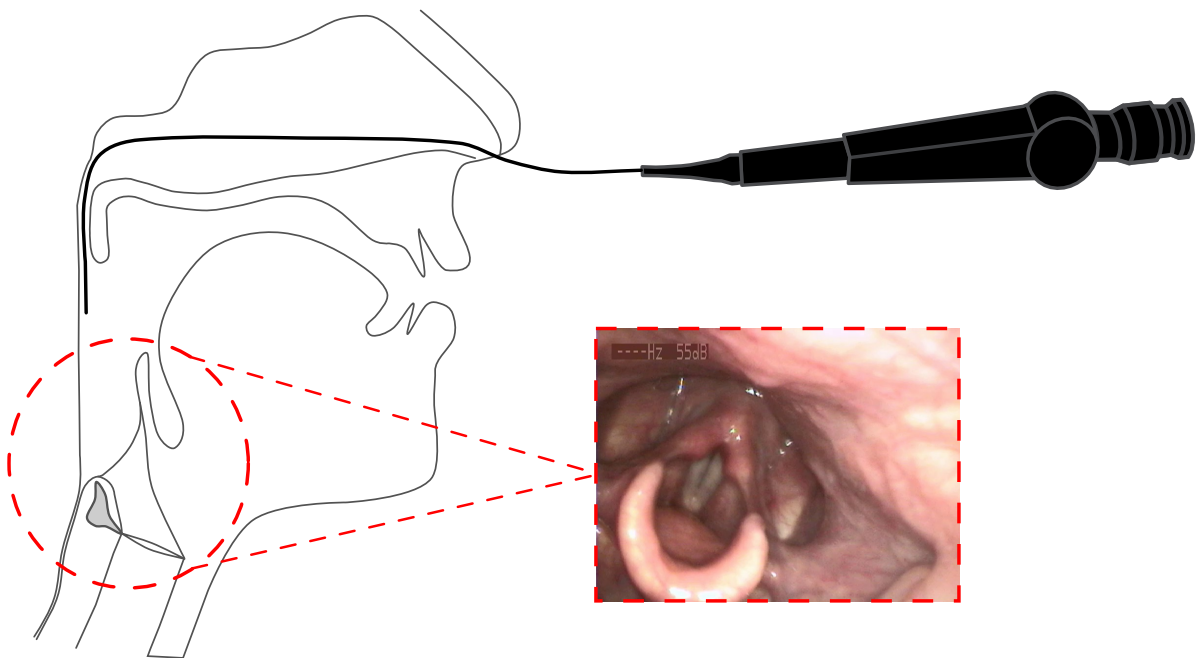


FIGURE 2.3 – Nasofibroscope Souple

### 2.5.2 Station d'enregistrement EVA2

Les données aérodynamiques ont été recueillies avec la station d'enregistrement EVA 2 (Figure 2.4) qui permet l'obtention de données aérodynamiques et acoustiques automatiquement synchronisées et calibrées. Elle est équipée de deux canaux pour mesurer le débit d'air (oral via un masque en silicone et nasal via des embouts placés à l'entrée des narines) ainsi qu'une sonde pour la pression intra-orale, introduite dans le nez jusqu'à l'oropharynx. Le débit d'air oral a été mesuré par un pneumotachographe (Ghio & Teston, 2004). La station EVA2 est également munie d'un microphone intégré se trouvant derrière le pneumotachographe. Nous avons enregistré des données synchronisées de pression intra-orale, de débit d'air buccal, de débit d'air nasal ainsi que le signal acoustique. Le masque en silicone a été pressé contre les joues du sujet pour éviter les fuites d'air et les perturbations sur le signal de débit d'air oral. Une sonde de pression a été introduite par une narine jusqu'à l'oropharynx pour mesurer la pression intra-orale ; dans l'autre narine, un petit embout a été placé à l'entrée de la narine et qui permet de mesurer le débit d'air nasal.

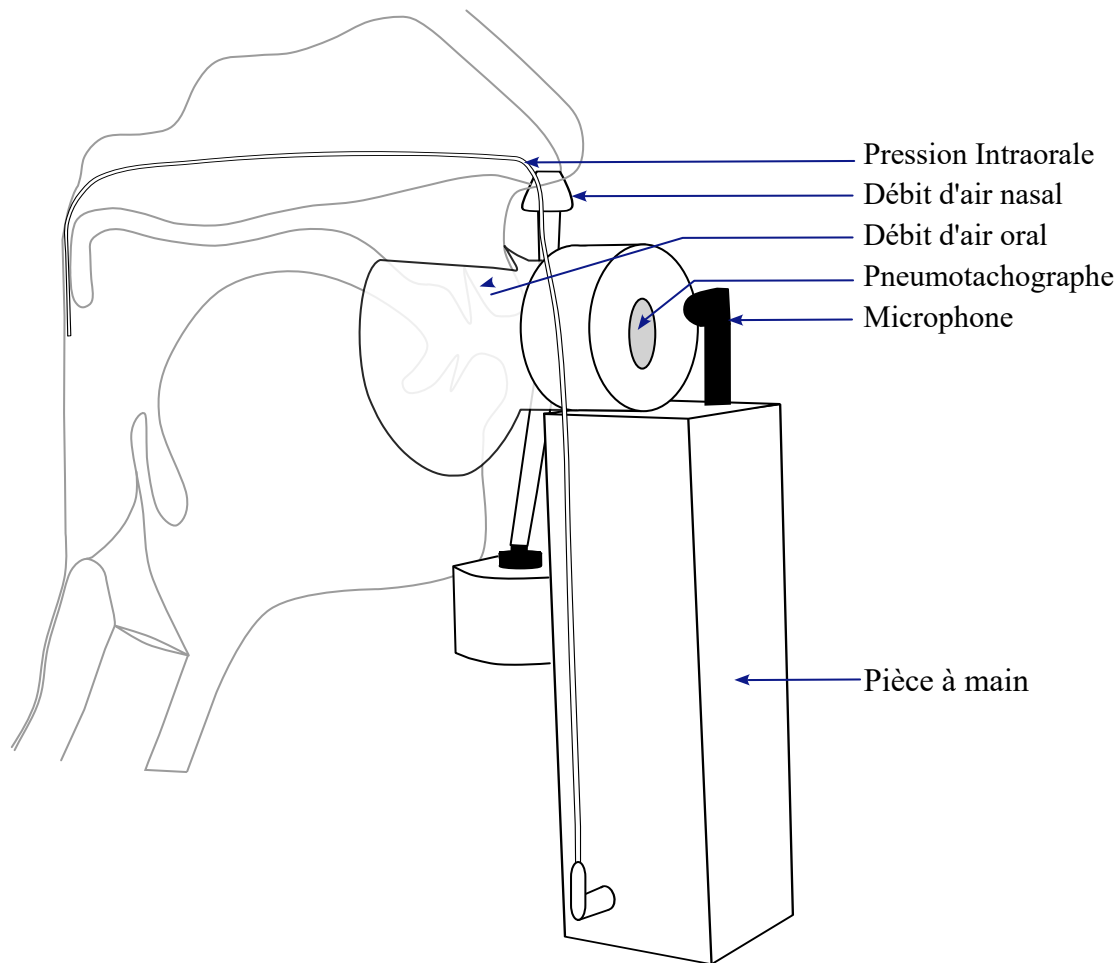


FIGURE 2.4 – Station d’enregistrement EVA2

Nos données présentent certaines limites. En premier lieu, nous avons décidé de ne pas inclure le sujet LJ pour qui la pression a été mesurée dans la bouche à cause de la gêne occasionnée par la sonde dans l’oropharynx. Le tube souple n’était pas fixé et sortait par moment de la cavité buccale. De plus, nous avons noté des valeurs de pression intraorale aberrantes allant jusqu’à 200hPa. Nous avons aussi observé des fluctuations aléatoires des signaux de débit d’air nasal et buccal ; nous en avons conclu qu’il y avait donc du bruit et nous avons préféré exclure les données de ce sujet.

Ensuite, nous avons rencontré des problèmes de saturation des signaux de pression concernant certains sons produits. Ces occurrences ne seront pas incluses dans les analyses où la pression est nécessaire. Nous avons rencontré des problèmes de saturation du débit d’air buccal au-delà de  $2dm^3/s$  (!). Enfin, nous avons rencontré des problèmes de saturation du signal acoustique. Certains sons sont plus intenses que d’autres et lors des enregistrements nous avons dû parfois réduire le niveau d’amplification du signal pour éviter la saturation (et ce même pendant la production). Cela rend donc impossible toute mesure de l’intensité du signal acoustique. Les futurs protocoles s’intéressant aux caractéristiques aérodynamiques du Human Beatbox devront anticiper ces problèmes techniques lors de la paramétrisation des seuils de saturation. Il est recommandé de choisir des seuils de pression d’au moins  $-100hPa$  &  $+100hPa$  et des seuils de débit d’air buccal d’au moins  $-2dm^3/s$  &  $+2dm^3/s$ .

La pièce à main de la station d’enregistrement EVA 2 est sensible aux variations de tempéra-

ture qui a pour effet de faire varier la position du 0 sur les signaux. Afin de presser le masque sur la bouche des sujets, ces derniers avaient la pièce à main dans leur main, ce qui a occasionné des problèmes de positionnement du 0. La Figure 2.5 montre un signal avant et après la procédure de correction du 0. La correction des signaux a été réalisée sur Praat de façon semi-automatique à l'aide d'un script que nous avons créé. Le niveau du 0 a été identifié visuellement pour chaque fichier ( $-26.9hPa$  dans l'exemple de la figure). La différence entre le 0 de référence ( $0hPa$ ) et le 0 mal positionné ( $-26.9hPa$ ) a été ajouté à chaque échantillon du signal correspondant.

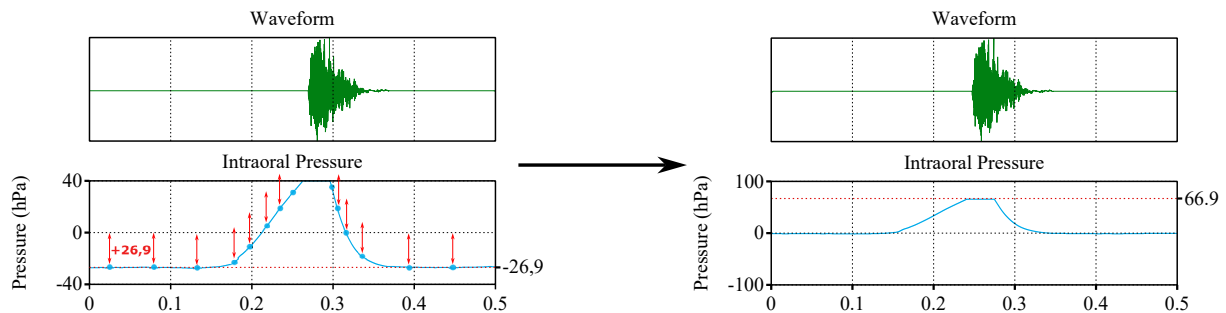


FIGURE 2.5 – Procédure de repositionnement du 0 des signaux aérodynamiques

### 2.5.3 Imagerie par Résonance Magnétique

Les données aérodynamiques et fibroscopiques permettent d'étudier les gestes laryngaux et supralaryngaux pendant la production du HBB. Cependant, la nature exacte des gestes supralaryngaux, c'est-à-dire les points de contact et la configuration du conduit vocal, ne peut être confirmée sans des données articulatoires. Pour confirmer les hypothèses émises à l'issue de l'analyse de la production du sujet VP à partir des données aérodynamiques et fibroscopiques, nous avons décidé d'enregistrer des données IRM. Un protocole simplifié a été mis en place en vue de l'acquisition de données IRM. Le corpus se composait de 9 des BP du corpus (les BP 8 et 11 ont été retirés) que le sujet VP devait répéter à 90 BPM seulement.

Les enregistrements IRM ont eu lieu au Centre Hospitalier Universitaire de Nancy sous la supervision d'André Vuissoz, Ingénieur de recherche de l'Université de Lorraine<sup>4</sup>. Le sujet VP a été allongé dans un scanner Siemens Prisma 3T, Erlangen, Germany (Figure 2.6). Il a été équipé d'une antenne tête/cou Siemens à 64 canaux. Le protocole consistait en l'acquisition d'une séquence modifiée de type radial RF-spoiled FLASH où  $TR = 2.22$  ms,  $TE = 1.47$  ms, FoV 192x192mm, angle de rotation =  $5^\circ$ . La coupe médiosagittale a une épaisseur de 8mm, la largeur de bande des pixels est de 1670 Hz/pixel, une matrice de 136x136 et la résolution temporelle est de 50 images par seconde (Uecker et al., 2010) et reconstruites par l'équipe du CHRU en suivant la technique d'inversion non-linéaire présentée dans Uecker et al. (2008). Le signal audio a été enregistré via un microphone à fibre optique (FOMRI III, Optoacoustics Ltd., Mazor, Israel). La fréquence d'échantillonnage de 16kHz. L'équipe du CHRU a pris soin de débruiter le signal acoustique en utilisant un algorithme proposé par Ozerov et al. (2012).

4. La session d'acquisition IRM dépend du protocole "METHODO" (ClinicalTrials.gov Identifier : NCT02887053, RCB-ID n°CPP EST-III, 08.10.01) et a reçu des financements des fonds européens CPER "IT2MP", "LCHN" et FEDER.

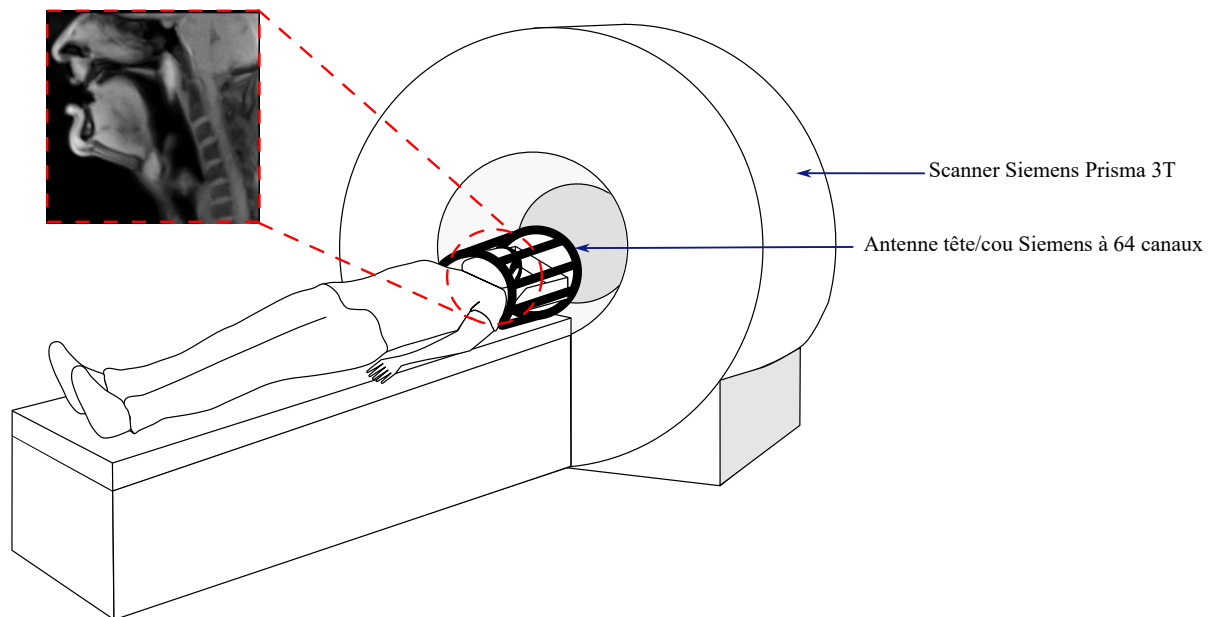


FIGURE 2.6 – Acquisition en Imagerie par Résonance Magnétique

## 2.6 Analyse

L'analyse des données occupera les deux prochains chapitres. Le Chapitre 3 a pour objectif de décrire les possibilités combinatoires du HBB. Cette description n'a pas vocation à être exhaustive. Nous évoquerons la variation intra et interindividuelle dans le Chapitre 4.

Dans le Chapitre 4, nous procéderons à la quantification des données aérodynamiques et acoustiques des sons produits dans les BP à 90, 120 et 150 BPM. Les données fibroscopiques et IRM (VP seulement) des patterns produits à 90 BPM nous serviront à interpréter les résultats aérodynamiques et acoustiques.

### 2.6.1 Analyse Qualitative

L'analyse qualitative reprend les principes de la description phonétique et physiologique. Nous faisons le choix ici de décrire les données à l'isolé afin que le contexte phonétique, c'est-à-dire les sons adjacents, ne puisse pas influencer sur la réalisation des sons analysés. Dans cette description, nous aborderons les productions à l'isolé sous quatre aspects : l'initiation, l'articulation, la phonation (à partir du corpus exploratoire de voix beatboxée) et l'acoustique (le timbre).

#### 2.6.1.1 Initiation

La description du lieu et du mode d'initiation se fera à partir des données aérodynamiques, fibroscopiques et IRM. Les données de pression intra-orale prises dans la cavité pharyngale sont précieuses dans la mesure où elles nous informent sur les diminutions et augmentations de volume du conduit vocal. La sonde de pression se trouvant dans la cavité pharyngale, nous ne disposons pas de données de pression intra-orale pour les mécanismes linguaux. En effet, ces mécanismes sont produits par compression/expansion du volume compris entre la constriction vélaire/uvulaire et la constriction antérieure. De ce fait, la pression mesurée est la pression dans le pharynx. En l'absence de données articulatoires au début du recueil des données, il a été



demandé aux sujets de produire du « *humming* » (base mélodique fredonnée où l'air s'écoule par les cavités nasales) à fréquence fondamentale constante. Ainsi, nous pouvons inférer la présence des mécanismes linguaux à partir d'indices aérodynamiques systématiques : une pression intra-orale nulle (ou très faible dans certains cas), un débit d'air nasal positif (ou négatif si les sujets reprennent leur souffle) et un débit d'air oral positif pour les sons égressifs ( $\{p\}$ ) ou négatif pour les sons ingressifs ( $\{\downarrow B^1\}$ ).

Les données aérodynamiques permettent aussi de définir ce que nous appelons « mode et lieu d'initiation ». Le mode d'initiation est le paramètre permettant de caractériser la façon dont le flux d'air est initié : par un geste de compression ou par un geste d'expansion. Le mode d'initiation par compression du volume entre le lieu d'initiation et d'articulation provoque une augmentation de pression et crée un flux d'air égressif. Inversement, le mode d'initiation par expansion du volume entre le lieu d'initiation et d'articulation provoque une diminution de pression et crée un flux d'air ingressif.

Pour ce qui est du lieu d'initiation, nous avons déjà mentionné comment nous inférons les mécanismes linguaux. Pour les mécanismes pulmoniques et laryngaux, nous utiliserons les données aérodynamiques et fibroscopiques. Les mécanismes laryngaux sont facilement identifiables sur le signal de pression et sur les images fibroscopiques. Lorsque la pression est élevée, que la glotte est fermée et qu'il y a des indices de mouvements verticaux du larynx et de compression ou d'expansion pharyngale, alors nous pouvons affirmer que le mécanisme d'initiation est laryngal. Dans le cas où la glotte resterait ouverte et que la pression serait plus faible, alors nous serons en mesure d'affirmer que l'initiation est pulmonique. Les différences de pression s'expliquent par les différences de volume. Si la température est constante, dans un grand volume (des poumons aux lèvres) la pression sera plus faible que dans un plus petit volume (de la glotte aux lèvres). Pour confirmer nos interprétations, nous utiliserons les clichés IRM en particulier pour les mécanismes laryngaux et linguaux.

### 2.6.1.2 Articulation

Concernant l'articulation, nous décrivons les lieux ainsi que les modes d'articulation. Pour définir le lieu d'articulation des sons beatboxés, nous nous appuyerons essentiellement sur les données IRM du sujet VP pour ce qui est des gestes supra-laryngaux. En revanche pour les gestes laryngaux, nous nous appuyerons sur les images laryngoscopiques de tous les sujets. Les lieux d'articulation sont illustrés sur la Figure 2.7. Nous y avons ajouté les lieux d'articulation pharyngaux et laryngaux en suivant le modèle de l'articulateur laryngé (Esling et al., 2019). Pour chacun des sons du corpus du sujet VP, nous identifierons de façon précise les articulateurs jouant un rôle dans l'articulation des sons beatboxés. En analysant les points de contact entre les articulateurs, nous définiront les lieux d'articulation. Ensuite, nous regarderons la signature acoustique et aérodynamique des occurrences produites par VP en les comparant aux signatures des sons correspondants produits par les autres sujets. Cela permettra de généraliser l'interprétation des lieux d'articulation à l'ensemble des sujets. Le mode d'articulation traduit la manière dont les articulateurs vont façonner le flux d'air en un point du conduit vocal. Le mode sera identifié à partir des données articulatoires, aérodynamiques et acoustiques. Notre corpus se compose de 3 modes d'articulation : occlusif, trille et affriqué.

Le mode occlusif (sourd) correspond à une fermeture totale en un point du conduit vocal suivi d'un relâchement. Cela se traduit par une pression positive ou négative (seulement lorsque l'occlusion se trouve en avant de la sonde de pression, c'est-à-dire pour les gestes supralaryngaux), un débit oral et nasal nul et un silence sur le signal acoustique.

Le mode trille correspond à un cycle d'ouverture et fermeture en un point du conduit vocal

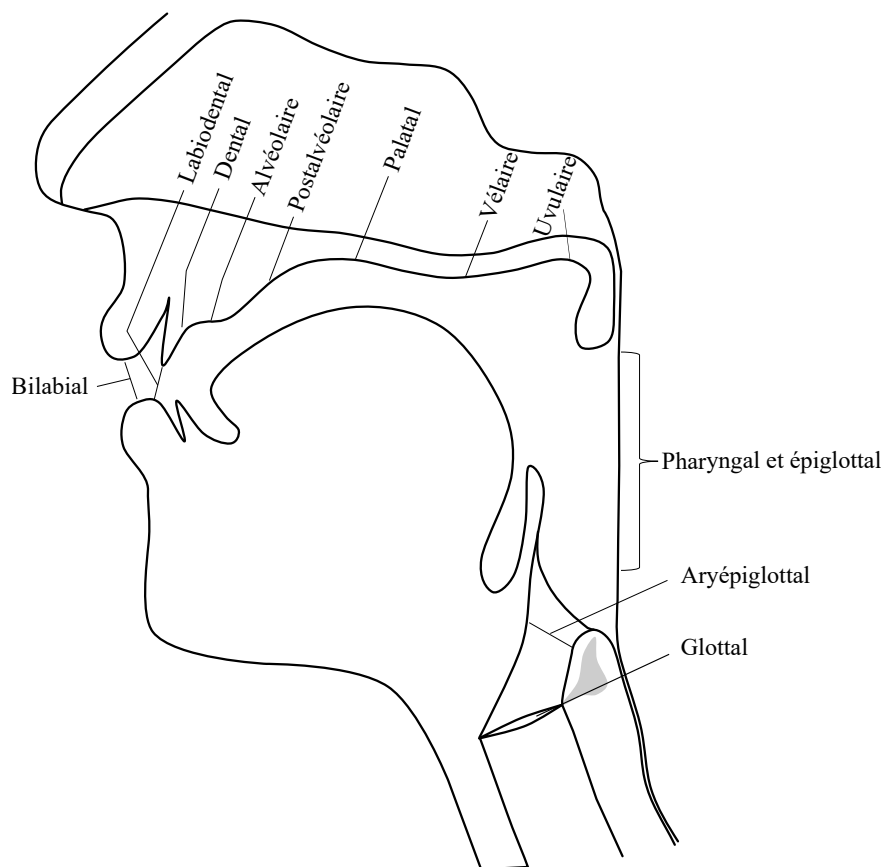


FIGURE 2.7 – Lieux d'articulation adapté de Esling et al. (2019).

à une fréquence donnée et qui est reflété par une oscillation périodique sur le signal acoustique et sur les signaux de pression et de débit d'air oral.

Notre corpus ne contient pas assez d'occurrences de sons latéraux et nasaux pour les analyser de la même façon que les modes d'articulations occlusifs, trille et affriqués.

### 2.6.1.3 Phonation

L'analyse de la voix beatboxée se fera à partir des données aérodynamiques, fibroscopiques et acoustiques. Rappelons ici que nous utiliserons le corpus exploratoire des répertoires des artistes à qui l'on a demandé de produire des « *scratches* », des « *basses vocales*<sup>5</sup> » ou divers effets de voix (voix saturée). Nous décrirons dans un premier temps les configurations laryngales observées sur les images fibroscopiques. Nous nous appuierons sur les catégories articutoires du *Laryngeal Articulator Model* (Esling et al. (2019), Figure 2.8). Nous mettrons en lien les différentes configurations avec les évènements aérodynamiques présent sur les signaux de pression et de débit d'air oral et nasal. Ensuite, nous décrirons les caractéristiques acoustiques à partir des spectres FFT et LPC. Sauf indication contraire, les spectres FFT sont extraits sur des fenêtres d'analyse de 25ms au milieu du son et les spectres LPC sont obtenus par un lissage

5. Les basses vocales réfèrent à l'imitation des basses électroniques produites avec du voisement. Il existe aussi des basses qui ne sont pas vocales et qui sont dites « *basses labiales ou linguales* ».

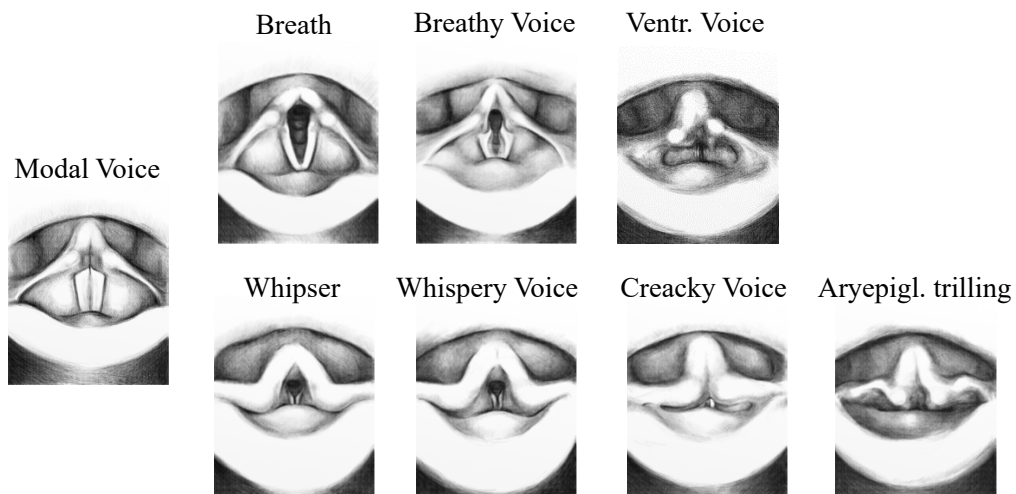


FIGURE 2.8 – Esling et al. (2019) : illustration des configurations laryngales pour plusieurs types de voix. À gauche la voix modale, en haut les voix produites sans constriction aryépiglottale, en bas les voix produites avec constriction aryépiglottale.

à 24 coefficients du spectre FFT. Nous décrivons trois paramètres : la structure harmonique, la différence d'amplitude entre les harmoniques (notamment H1 et H2) et la pente spectrale.

#### 2.6.1.4 Timbre

Le timbre permet de différencier deux sons de même intensité et hauteur. Avec la dynamique temporelle, le timbre permet de catégoriser les sons musicaux.

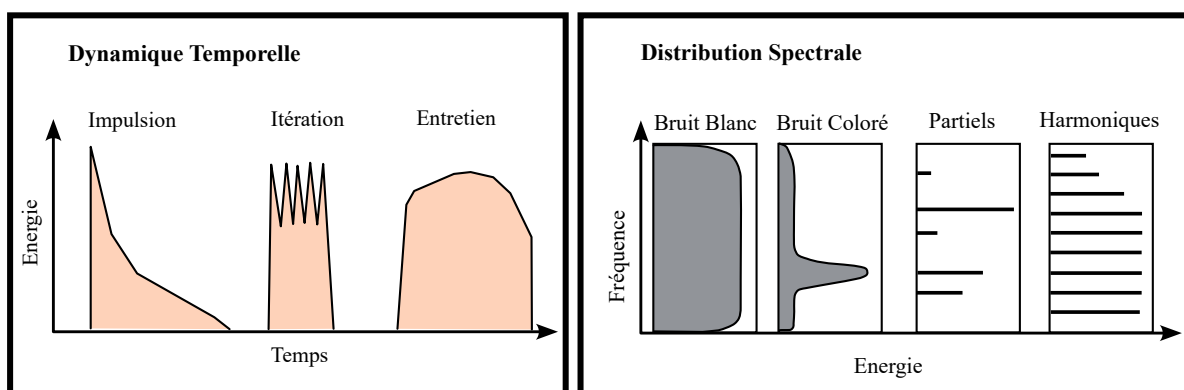


FIGURE 2.9 – Caractérisation du timbre en musique (Castellengo, 2015)

La dynamique temporelle (Castellengo, 2015) ou les transitoires dans la terminologie de Leipp (1971) se rapporte à la façon dont un son musical commence et se termine. Ainsi, nous distinguerons les sons « impulsifs » (percussions) des sons « continus » (basses vocales Beatboxées). Les sons impulsifs se caractérisent par une diminution plus ou moins rapide de l'énergie après l'attaque (moment de l'impact). Ils ont un caractère plus instantané. On notera que les sons de type « itératif » (Figure 2.9) sont compris dans la classe des impulsions étant donné qu'ils constituent une suite d'impulsions. En opposition, nous avons les sons « continus » (ou entretenus dans la terminologie de Castellengo, 2015), qui ne présentent pas de diminution abrupte et graduelle de l'énergie après l'attaque.

Le timbre se caractérise aussi par la distribution spectrale des fréquences. On distinguera les spectres à harmoniques ou partiels des spectres de bruit. Les sons caractérisés par la présence d'harmoniques réguliers ou partiels dépendent du fondamental, c'est-à-dire du premier harmonique. La hauteur et l'intensité du fondamental et de ses harmoniques jouent un rôle dans la sensation sonore d'un son. Les spectres de bruits se caractérisent par une perturbation acoustique aperiodique et par une prééminence spectrale dont la fréquence dépend de la longueur de la cavité antérieure, et donc du lieu d'articulation et dont l'amplitude dépend de la pression, de la vitesse des particules et de la taille de la constriction. L'analyse portera uniquement sur les imitations de percussion du corpus des *Beat Patterns*. La qualification du timbre des sons se fera à partir de la visualisation du signal acoustique, du spectrogramme et des spectres qui ont été paramétrés de la même façon que celle exposée dans méthodologie d'analyse de la phonation.

## 2.6.2 Analyse Quantitative

L'analyse quantitative nous permettra d'étudier les effets du tempo sur la production des *Beat Patterns* à 90, 120 et 150 BPM.

### 2.6.2.1 Segmentation

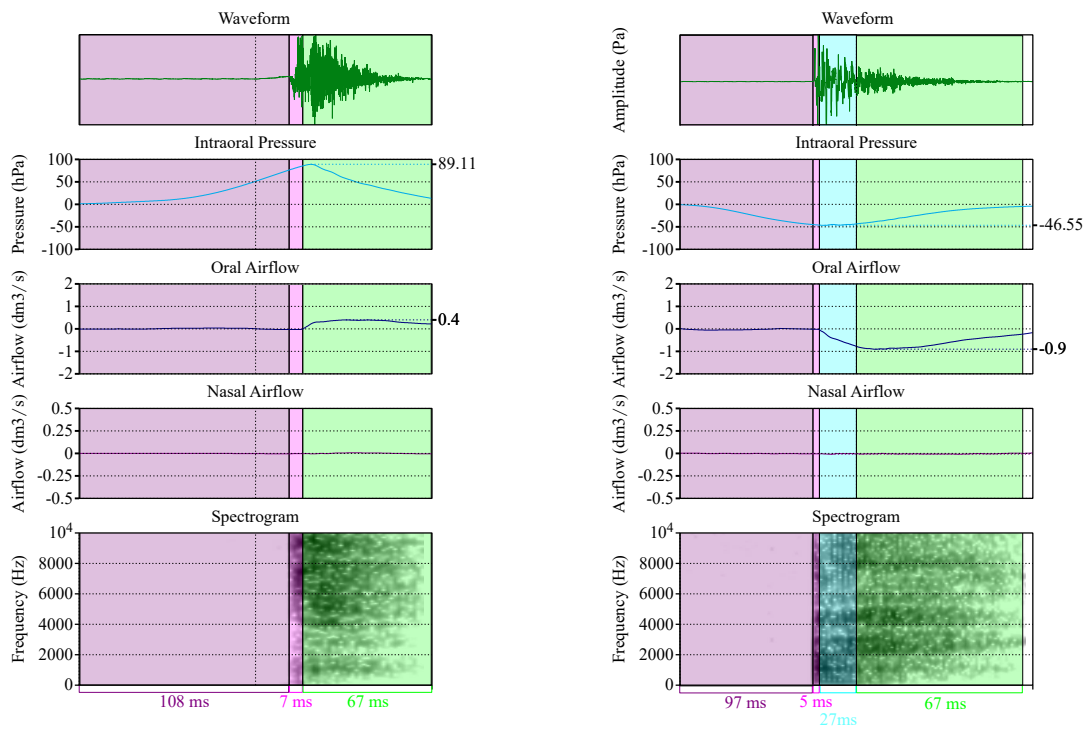
Les données acoustiques et aérodynamiques ont fait l'objet d'une segmentation manuelle. Celle-ci s'est faite sur le logiciel Praat (Boersma & Weenink, 2001). La Figure 2.10 illustre la segmentation de 4 sons beatboxés par le sujet CJ.

Les fichiers *textgrid* des enregistrements comprenaient plusieurs *tiers* d'annotation, la première *tier* phonétique correspond à la segmentation des événements aérodynamiques et acoustiques à partir d'indices phonétiques identifiés sur les signaux et le spectrogramme. Dans notre corpus, on trouve des occlusions, des bursts suivis d'un bruit d'explosion, des bruits de friction, des trilles et du voisement. Nous avons décidé de distinguer le burst du bruit d'explosion car le burst correspond à la collision des particules de part et d'autre de la constriction au moment de l'ouverture du conduit, alors que le bruit d'explosion correspond à une courte période de friction due à l'écoulement de l'air entre les articulateurs s'éloignant progressivement du lieu d'occlusion. Si l'on regarde les Figures 2.10a, 2.10b et 2.10c, le burst est indiqué en rose. On note du bruit sur le signal acoustique et le spectrogramme alors que le débit d'air oral reste nul. Ce dernier augmente quelques millisecondes après. Le burst se caractérise donc par la présence de bruit acoustique et l'absence de débit d'air oral alors que le bruit d'explosion se caractérise par la présence de bruit acoustique et la présence de débit d'air oral.

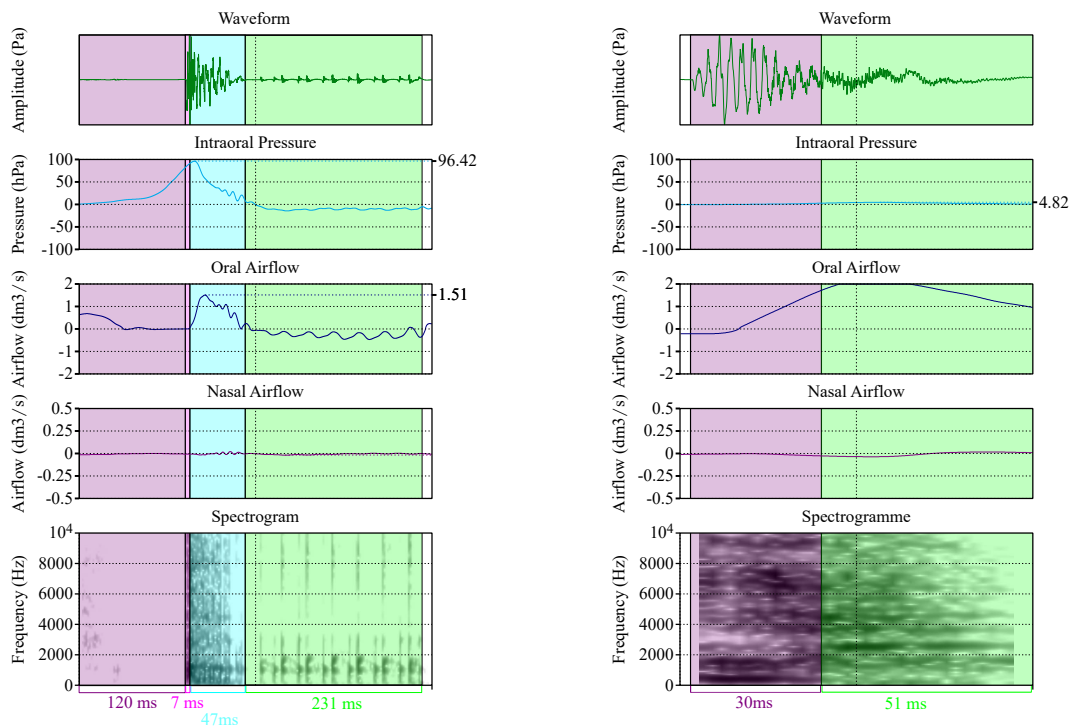
La deuxième ligne sert à annoter les sons tels qu'ils ont été produits et la troisième ligne correspond aux sons cibles c'est-à-dire tels qu'ils devraient être produits. Le choix d'une ligne des sons produits et des sons cibles nous permettra de relever toutes les erreurs de production comme les élisions (le son n'est pas produit) ou encore les transformations diverses (affriquée > occlusive).

### 2.6.2.2 Analyse Temporelle

Dans cette analyse, nous avons mesuré la durée des événements aérodynamiques et acoustiques : la durée des BP, des sons et des événements acoustiques. La durée des événements aérodynamiques et acoustiques a été mesurée comme la différence entre la fin et le début de l'événement (occlusion, friction, trille). La durée des sons a été mesurée comme la différence entre la fin et le début du son. Enfin, la durée des *Beat Patterns* est calculée comme la différence entre la fin du dernier son et le début du premier son.



(a) Segmentation de l'occlusion, du burst et de la friction de  $[ts']$  (b) Segmentation de l'occlusion, du burst, du bruit d'explosion et de la friction de  $[\downarrow kl:]$



(c) Segmentation de l'occlusion, du burst, du bruit d'explosion et de la trille de  $[p' \downarrow B^1:]$  (d) Segmentation du voisement et de la friction de  $[\updownarrow h:]$

FIGURE 2.10 – Segmentation de 4 sons beatboxés du sujet CJ. Violet = occlusion, rose = burst, bleu = bruit d'explosion, vert = friction ou trille

### 2.6.2.3 Analyse Aérodynamique

L'analyse aérodynamique a consisté à extraire automatiquement les pics de pression, de débit d'air oral et nasal pour chaque son lorsque ceux-ci ne présentent aucun signe de saturation. En l'absence de sons nasals dans notre corpus, nous n'analyserons pas le débit d'air nasal dans le Chapitre 4. À partir de ces paramètres, il est possible d'estimer la vitesse des particules et de calculer l'aire relative de la constriction.

Afin d'estimer l'aire relative de la constriction, nous devons calculer la vitesse des particules ( $cm/s$ ) à partir de l'équation 1.3 :  $v = 412\sqrt{P}$  où  $P$  est la pression intraorale mesurée en hectopascal et ramenée en décapascal au lieu des millimètres d'eau suggérés par Catford. Concernant l'aire de la constriction ( $cm^2$ ), reprenons la formule 1.4. Celle-ci nous précise que l'aire de constriction est le rapport du débit d'air oral sur la vitesse des particules :  $A = U/v$  où  $U$  est le débit d'air oral mesuré en  $dm^3/s$  puis converti en  $cm^3/s$  et  $v$  la vitesse des particules estimée à partir de la pression intraorale.

### 2.6.2.4 Analyse Acoustique

L'analyse acoustique portera sur les percussions, les *Lips Rolls* et les *clics Rolls*. Nous utiliserons les quatre moments spectraux que sont le Centre de Gravité Spectral, sa variance, le *Skewness* et le *Kurtosis*. Les signaux acoustiques ont été rééchantillonnés à 20kHz afin d'observer la répartition de l'énergie entre 0 et 10kHz.

Le Centre de Gravité Spectral est la moyenne calculée à partir d'une fonction intégrale d'un spectre. Le Centre de Gravité spectral et sa variance renseignent sur la plage de fréquences où l'énergie est concentrée. On exprime ces deux moments spectraux en Hertz.

Le *Skewness* est le coefficient d'asymétrie (Figure 2.11), contrairement au deux premiers moments, il n'a pas d'unité standardisée. Ce coefficient indique si la répartition fréquentielle est symétrique (*Skewness* = 0) ou si l'énergie est concentrée dans les basses fréquences (*Skewness* positif) ou dans les hautes fréquences (*Skewness* négatif). Le *Kurtosis* (sans unité standardisée) est le coefficient d'aplatissement (Figure 2.11). Il indique si la structure du spectre est plate ou pointue. Les moments spectraux sont des indices redondants qui concernent l'enveloppe spectrale. Le Centre de Gravité spectral et le *Skewness* (ou coefficient d'asymétrie) renseignent sur la répartition de l'énergie entre 0 et 10kHz. L'écart-type du Centre de Gravité spectral et le *Kurtosis* (ou coefficient d'aplatissement) informent sur la forme du spectre. Par exemple, un spectre dont l'enveloppe est plate aura un *Kurtosis* faible et un écart-type élevé.

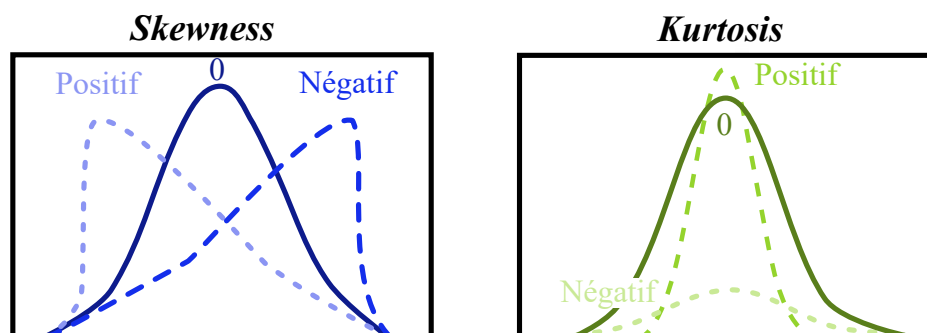


FIGURE 2.11 – *Skewness* et *Kurtosis*

Ces quatre moments spectraux ont été mesurés pour chaque événement acoustique (excepté les occlusions) à partir de spectres FFT dont la fenêtre temporelle est soit de 25 millisecondes

si la durée de l'événement est égale ou supérieur à la durée de la fenêtre ou sur une fenêtre de 5 millisecondes lorsque la durée de l'évènement est inférieure à 25ms. Ce choix est motivé par le fait qu'une fenêtre de 25ms lorsque l'évènement dure 5ms (comme c'est le cas pour les bursts) sera positionnée sur 5ms de bruit et 20ms de silence. Nous avons fait coulisser la fenêtre sur toute la durée du son (excepté sur l'occlusion) sans que les fenêtres ne se chevauchent dans le temps. Si la fin de la fenêtre s'étend au delà du son, aucune mesure n'a été faite afin de ne pas analyser le silence. L'analyse par fenêtre coulissante est motivée par la volonté de capturer à la fois la répartition spectrale mais aussi la dynamique temporelle.

Pour ce qui est du voisement, des mesures de fréquence fondamentale (F0) ont été réalisées par l'algorithme de cross-corrélation de Praat. Les paramètres par défaut de détection ont été modifiés de façon à ce que l'algorithme soit en mesure de détecter la fréquence fondamentale sur une plage de 25Hz à 1000Hz ainsi que la fenêtre temporelle de détection que nous avons fixé à 30ms dans les cas où la F0 serait très basse.

### **2.6.2.5 Statistiques descriptives**

Afin de faciliter la présentation de la quantification des résultats, nous utiliserons des statistiques descriptives. Nous nous intéressons aux effets du tempo sur la production du Beatbox. Le tempo est une variable indépendante avec 3 valeurs : 90BPM, 120BPM et 150BPM. Les différents sons de notre corpus constituent eux aussi des variables catégorielles. Les variables dépendantes sont les mesures temporelles, aérodynamiques et acoustiques présentées précédemment, à savoir : la durée des BP et des sons ; les pics de pression, de débit d'air oral et l'aire relative de la constriction ; et enfin, les moments spectraux et la F0. Nous analyserons la moyenne et l'écart-type des variables dépendantes en fonction de la variable indépendante, c'est-à-dire le tempo de référence.

Afin de vérifier la présence ou non de données aberrantes, nous procéderons à une analyse de la distribution de celles-ci grâce à des boîtes à moustaches générées sur le logiciel R. Les boîtes à moustaches nous permettront de connaître la moyenne, la médiane et les quartiles de nos variables numériques, par exemple, la durée des sons.

# Chapitre 3

## Le Human Beatbox : un système combinatoire discret

« The devil [is] in the details. »

---

John J. Ohala<sup>1</sup>

### Introduction

Le Chapitre 3 a pour objectif de montrer que le Human Beatbox est un système combinatoire discret (Proctor et al., 2013), c'est-à-dire un système sonore composé que l'on peut identifier et délimiter. Nous tenterons en plus de démontrer en quoi la production vocale du beatbox est plus complexe que la production de la parole. L'hypothèse que nous défendons est que les primitives du Human Beatbox sont les mêmes qu'en parole car ces deux systèmes manipulent les mêmes organes de production (système respiratoire, laryngal, oral). Ce qui diffère entre les deux systèmes sont les possibilités combinatoires des primitives. Pour confirmer ou infirmer cette hypothèse, nous proposons une description phonétique des mécanismes de production des sons de notre corpus à partir de données articulatoires, aérodynamiques et acoustiques.

Le Chapitre 3 se compose de 4 parties. Les trois premières traiteront des sons issus du corpus des *Beats Patterns* (Tableau 2.1 et 2.2). Nous montrerons dans un premier temps quels sont les mécanismes d'initiation utilisés par les sujets. Ensuite, nous regarderons comment les beatboxeurs et les beatboxeuses combinent ces mécanismes avec des mécanismes articulatoires. Enfin, nous montrerons comment ces combinaisons donnent lieu à des signatures acoustiques discrètes. Ces trois parties sont essentielles pour comprendre les effets du tempo sur les sons beatboxés (Chapitre 4). La dernière partie s'intéressera aux répertoires individuels des participant-es. Nous présenterons les caractéristiques physiologiques et acoustiques de la voix beatboxée. Nous tenterons de voir comment les beatboxeurs et les beatboxeuses combinent différents mécanismes phonatoires (phonation égressive, phonation ingressive). Enfin, nous mentionnerons les mécanismes sifflés et nous montrerons comment ils peuvent se combiner à d'autres mécanismes.

---

1. Ohala, J. J. (2009). Languages' sound inventories : the devil in the details. *Approaches to phonological complexity*, 16, 47.



## 3.1 Initiation

Nous commençons ce chapitre par l'analyse des mécanismes d'initiation dans nos données. Les sujets ont produit les 6 mécanismes d'initiation possible (cf. Tableau 1.2).

### 3.1.1 Mécanismes pulmoniques égressifs

La Figure 3.1 illustre la production d'une *ESH Snare Outward*<sup>2</sup> produite par le sujet GA. La figure se compose du signal acoustique (vert), de pression intra-orale (cyan), de débit d'air oral (bleu), de débit d'air nasal (violet) ainsi que le spectrogramme correspondant. Nous garderons les mêmes échelles de pression ( $-100, +100hPa$ ), de débit oral ( $-2dm^3/s, +2dm^3/s$ ), de débit nasal ( $-0.5dm^3/s, +0.5dm^3/s$ ), de fréquences ( $0kHz, 10kHz$ ) et de temps ( $0, 500ms$ ) tout au long du manuscrit. Il conviendra dans certains cas de changer l'ordre de grandeur des échelles, dès que ce sera le cas nous l'indiquerons dans la légende de la figure. Lorsque nous disposons d'images fibroscopiques et/ou IRM, nous les insérerons en dessous des signaux acoustiques et/ou aérodynamiques.

La *ESH Snare Outward* fait partie des caisses claires « oro-laryngales » de notre corpus. Nous définissons les caisses claires oro-laryngales comme des affriquées dont l'occlusion a lieu dans le larynx et dont la partie fricative peut-être glottale ou orale. Notre corpus en comporte trois : la *ESH Snare Outward* [ $\hat{ʃ}f$ :], la *ESH Snare Inward* [ $\downarrow\hat{ʃ}f$ :], et la *Cough Snare* [ $\hat{ʃ}h$ :]. L'utilisation du diacritique [ $\hat{ʃ}$ ] de voisement est motivée par la présence de voisement au moment du relâchement. Cette caisse claire est initiée par un geste de compression pulmonique. Le signal de pression indique qu'il y a une constriction dans le conduit oral. La pression augmente au moment où la friction apparaît sur le signal acoustique et non pendant l'occlusion. L'occlusion est illustrée par l'image fibroscopique de gauche où l'on voit une fermeture des bandes ventriculaires. Au moment du relâchement, il y a des indices de voisement. Nous interprétons donc cette occlusive comme une occlusive aryépiglottale voisée. Nous sommes ici en train de décrire le lieu d'articulation et non le lieu d'initiation. L'initiation a lieu au niveau des poumons. Dès que l'occlusion est relâchée, le débit d'air oral augmente. L'absence de débit d'air nasal indique que le port vélo-pharyngé est fermé. Notons que la pression ( $+40hPa$ ) et le débit ( $+2dm^3/s$ ) sont supérieurs aux valeurs courantes de 3 à  $40hPa$  et de  $+0.2$  à  $1dm^3/s$  des mécanismes d'initiation pulmonique de la parole.

Une pression intra-orale et un débit d'air oral positifs indiquent un flux d'air égressif. On sait par la loi de Boyle-Mariotte (cf. équation 1.2) que pression et volume sont en rapport inverse. On peut donc estimer la réduction du volume pulmonique  $V_2$  à partir du rapport entre le pression initiale dans un conduit ouvert ( $P_1$ ), volume pulmonique initial ( $V_1$ ) et la pression intraorale mesurée après compression ( $P_2$ ). Si l'on considère que le volume des poumons est de  $3000cm^3$  alors on peut écrire :

$$V_2 = \frac{P_1 \times V_1}{P_2} = \frac{1013 \times 3000}{1053} = 2886 (cm^3)$$

Le volume pulmonique a donc diminué de  $114cm^3$

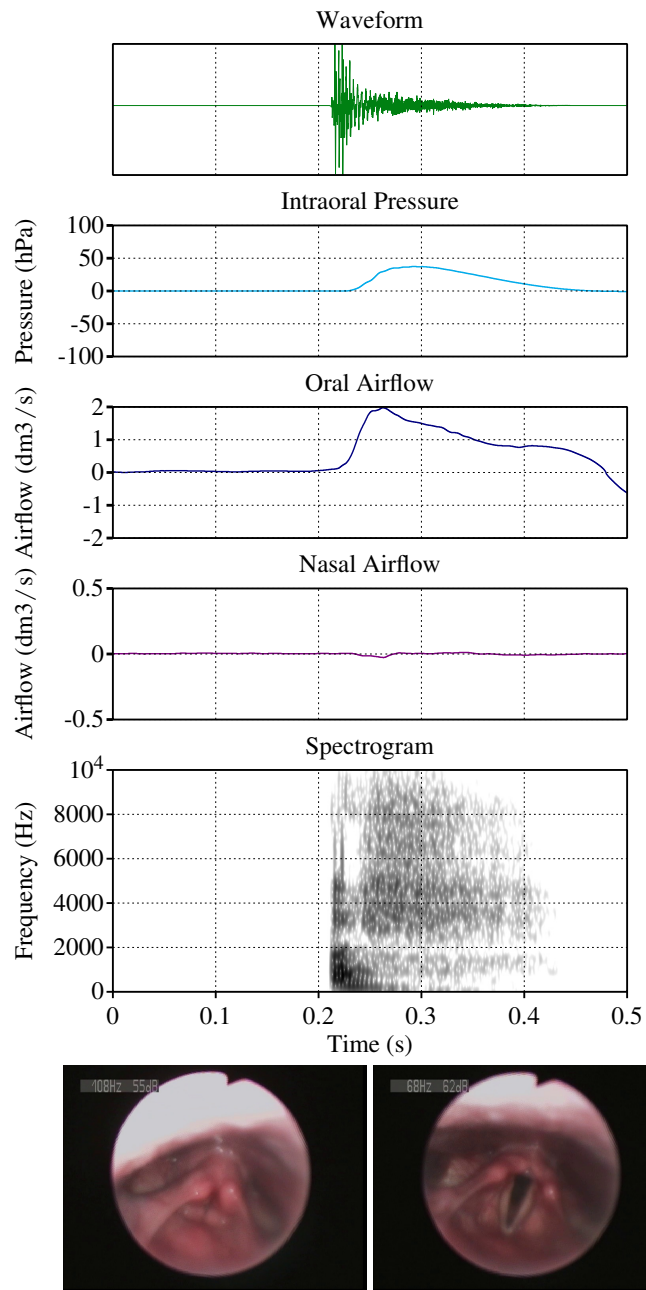


FIGURE 3.1 – *ESH Snare Outward* [ʔ:] - Sujet GA

⊙

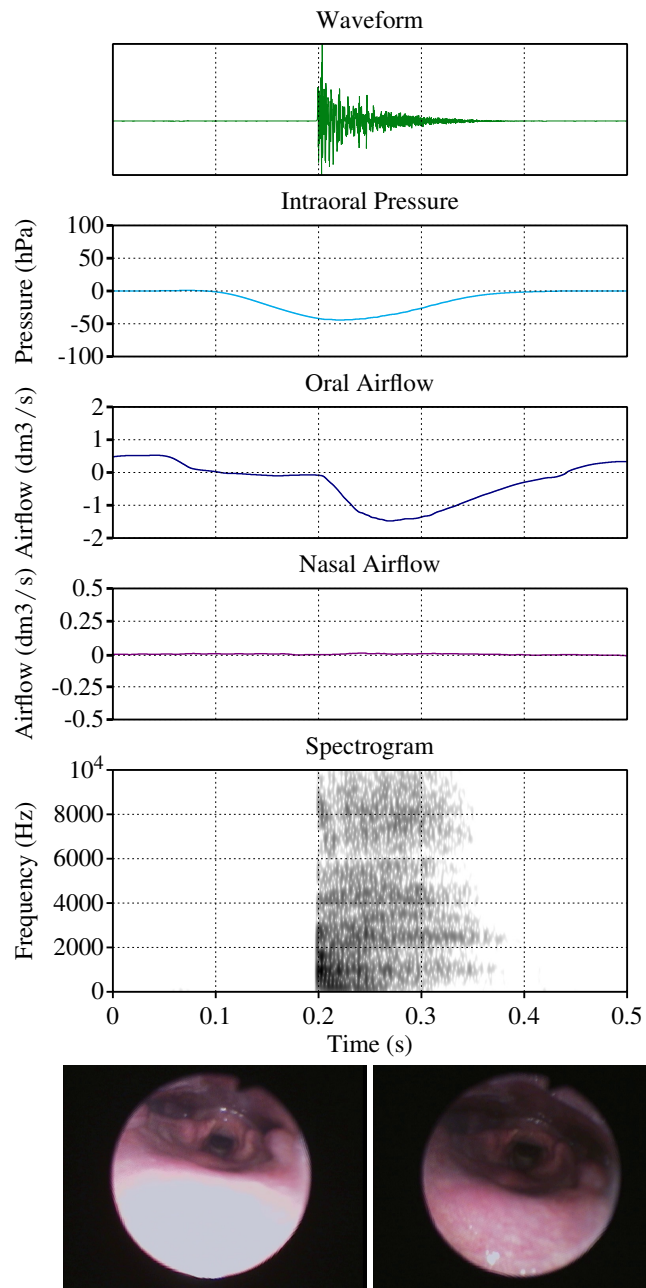


FIGURE 3.2 – *K-Snare Inward* [ $\downarrow\widehat{kL}$ ] - Sujet CJ

⊙

### 3.1.2 Mécanismes pulmoniques ingressifs

La Figure 3.2 illustre la production de la *K-Snare Inward* (caisse claire) par le sujet CJ. Cette caisse claire fait partie des 3 sons de base du beatbox avec le *Classic Kick* [p'] (grosse caisse) et le *Hi-Hat fermé* [ts'] (charleston fermé). La caisse claire [↓k̂L:] est une affriquée vélaire latérale initiée par expansion pulmonique (pulmonique ingressif). La pression durant l'occlusion vélaire est négative et atteint  $-40hPa$  au relâchement. Le port vélo-pharyngé est fermé (pas de débit nasal). Le tube de pression se trouvant dans le pharynx, si l'initiation avait été linguale, nous n'aurions pas mesuré de pression. Le débit d'air oral descend à  $-1dm^3/s$ . Ceci suggère encore une fois que l'initiation n'est pas linguale car le débit d'air aurait été plus faible. Les images fibroscopiques ont été extraites pendant l'occlusion et pendant la friction. La glotte reste ouverte tout du long, ce qui exclut une initiation laryngale. La descente du larynx s'accompagne d'une expansion pharyngale. Il s'agit donc ici d'une initiation pulmonique ingressive. Une diminution de la pression implique une augmentation du volume (expansion pulmonique). En reprenant la loi de Boyle-Mariotte on peut mesurer l'augmentation volumique, soit une augmentation de  $116 cm^3$ . Notons que les signatures acoustiques de [t̂j:] (sujet GA) et [↓k̂L:] (sujet CJ) sont remarquablement similaires bien que différentes dans leurs structures acoustiques.

### 3.1.3 Mécanismes laryngaux égressifs

Le larynx peut aussi initier des flux d'air. Dans le cas du *Classic Kick* (Figure 3.3), l'air est mis en mouvement par un mécanisme de compression laryngale, c'est-à-dire par un ensemble de manœuvres du système laryngal qui ont pour but d'augmenter la pression dans le conduit vocal.

La pression atteint presque  $+50hPa$  et le débit oral  $+0.8dm^3/s$ . Le flux d'air est égressif. Les images fibroscopiques ont été extraites au début et à la fin du geste d'initiation, c'est-à-dire à partir du moment où la glotte se ferme jusqu'à l'instant où le larynx a fini son ascension. La glotte n'est pas totalement fermée, on remarque une très légère ouverture. Le geste de fermeture glottique s'accompagne d'une constriction aryépiglottique. Le recul de la racine de la langue permet de compresser l'air présent dans le pharynx et l'épilarynx. La pression de [p'] est plus de 2x supérieure à ce que l'on mesure dans les langues pour une éjective bilabiale.

Dans sa thèse, Kingston (1984) a fait l'hypothèse que l'augmentation de la pression pendant des mécanismes éjectifs était le résultat d'une combinaison de manœuvres pour manipuler la pression. Selon lui, l'élévation du larynx ne peut expliquer à elle seule la forte pression observée dans le conduit vocal. Il évoque le possible rôle de la racine de la langue ainsi que les ajustements de la raideur des parois du conduit vocal. Dans le cas du *Classic Kick*, une augmentation de la rigidité des parois du conduit vocal permettrait d'augmenter la pression pendant la compression.

Les oscillations sur le signal acoustique et sur le débit d'air oral ne sont pas dues à la production de voisement mais plutôt à une oscillation labiale lors d'un relâchement. Il est compliqué de savoir exactement ce qui oscille. Il pourrait s'agir de la lèvre ou bien de la muqueuse labiale.

---

2. Le saviez-vous? Les beatboxeurs et les beatboxeuses se saluent en se disant [t̂j:]. Cette utilisation paralinguistique de la *ESH snare* est dérivée du français « WESH ma gueule ! ». L'expression a été introduite par Faya Braz et Swish Beatbox : <https://www.humanbeatbox.com/glossary/esh/>.

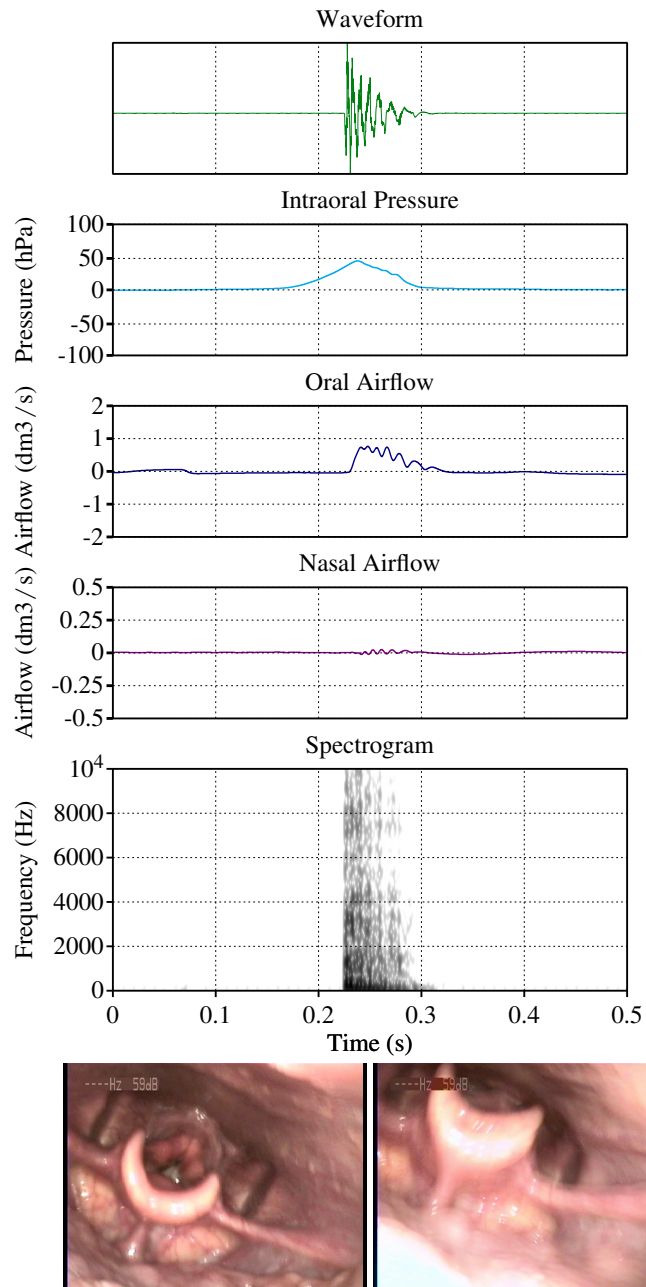


FIGURE 3.3 – *Classic kick [p']* - Sujet VP

⊙

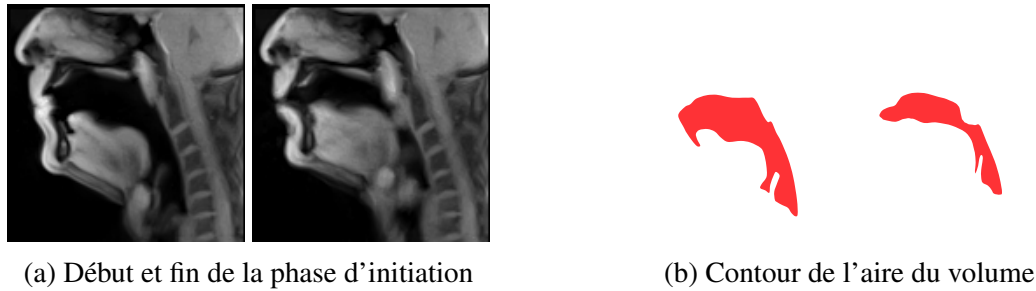


FIGURE 3.4 – Mode d'initiation par compression laryngale - [p']  
sujet VP

La figure 3.4 illustre la réduction du volume entre le lieu d'initiation et d'articulation lors de la production du *Classic Kick* [p']. Sur l'image IRM de droite, la petite boule grise en dessous de la luette entre la langue et la paroi postérieure du pharynx est l'amygdale du sujet VP. Ceci a été confirmé lors de l'examen fibroscopique. Pour chaque image IRM, nous avons dessiné manuellement le contour de l'aire au début et à la fin de la phase d'initiation. Si l'on regarde le contour initial et le contour final, on note une réduction de l'aire. Le calcul de la diminution du volume supralaryngal, entre les lèvres et la glotte, peut se faire à partir des données de pression intra-orale de la figure 3.3, où le pic de pression se trouve à 50 hPa environ. Pour calculer le volume  $V_2$  après compression on prend cette fois le rapport entre la pression initiale  $P_1$  (pression atmosphérique), le volume supralaryngal initial  $V_1$  ( $140\text{cm}^3$ ) et la pression intraorale mesurée  $P_2$  :

$$V_2 = \frac{P_1 \times V_1}{P_2} = \frac{1013 \times 140}{1063} = 133 (\text{cm}^3)$$

Nous savons à présent qu'entre le début et la fin de la phase d'initiation une réduction du volume de  $7\text{cm}^3$  s'est produite.

Encore une fois, la réduction du volume ne se fait pas uniquement par la montée du larynx. Le geste d'initiation combine un geste d'élévation laryngale suivi d'une rétraction de la racine de la langue.

### 3.1.4 Mécanismes laryngaux ingressifs

La Figure 3.5 illustre la production d'un *throat kick* [6'] produit par le sujet GA. Le *throat kick* est produit par un mécanisme d'expansion laryngale. En effet, la pression intra-orale approche les  $-50\text{hPa}$ . Il y a donc une augmentation du volume. La pression négative s'explique par une descente du larynx pendant la fermeture des lèvres. Le début du geste laryngé rappelle la configuration de la *ESH Snare Outward* : la glotte et les bandes ventriculaires sont en adduction. À la fin du geste d'initiation la glotte est légèrement ouverte et le larynx est en position basse. Les oscillations sur le signal acoustique correspondent à une phase de voisement qui coïncide avec le pic de pression intra-orale. Nous pensons que le relâchement a lieu au niveau des plis vocaux alors que les lèvres restent fermées. En effet, le signal de débit d'air oral est faible ( $\pm 0.1\text{dm}^3/\text{s}$ ). Or, avec une pression de  $-50\text{hPa}$ , on s'attendrait à voir un débit d'air oral plus important. Les images IRM (Figure 3.6a) confirment que les lèvres restent fermées.

On remarque une augmentation de l'aire du conduit vocal sur les images IRM (Figure 3.6a) et sur les contours dessinés (Figure 3.6b). Sur les données aérodynamiques, cette l'expansion se traduit par une pression et un débit d'air oral négatif. Reprenons notre équation avec une valeur de  $-50\text{hPa}$  où  $V_2$  est le volume après expansion,  $P_1$  la pression dans un conduit vocal ouvert et  $V_1$  le volume initial entre la glotte et les lèvres :

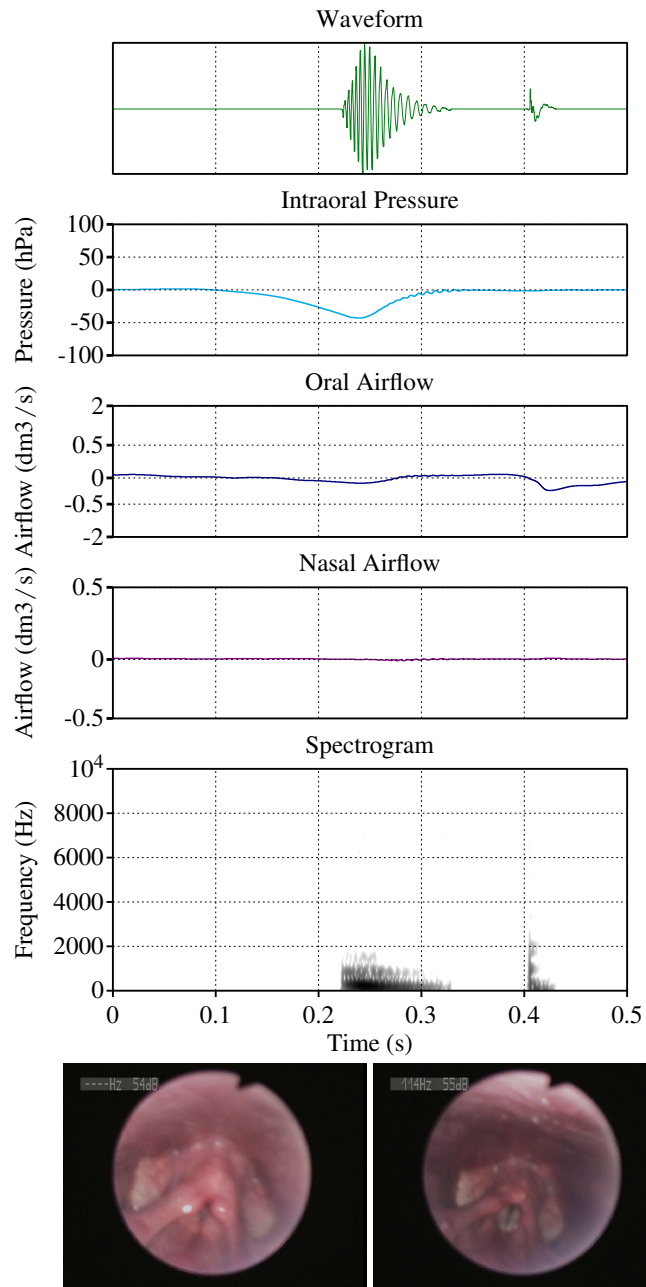


FIGURE 3.5 – *Throat kick* [ɬ] - Sujet GA

⊙

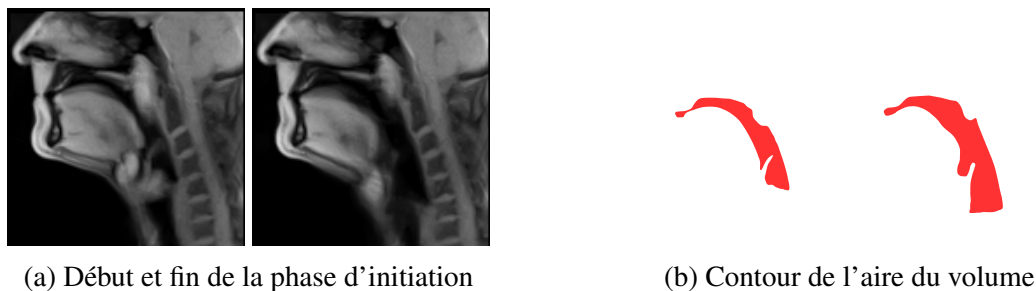


FIGURE 3.6 – Mode d'initiation par expansion laryngale - [b̃]  
sujet VP

$$V_2 = \frac{P_1 \times V_1}{P_2} = \frac{1013 \times 140}{963} = 147(\text{cm}^3)$$

Le volume a augmenté de  $7\text{cm}^3$ . Les données IRM confirment que la stratégie de réduction de la pression consiste à abaisser le larynx et à avancer la racine de la langue pour augmenter le volume.

### 3.1.5 Mécanismes linguaux égressifs

Enfin, le flux d'air peut être initié dans la cavité buccale. La Figure 3.7 illustre un mécanisme d'initiation linguale lors de la production de la *PF Snare humming* {p̃f:}. Le mécanisme d'initiation n'est pas directement mesurable à partir de nos données aérodynamiques car la sonde de pression se trouve derrière la constriction postérieure, dans cavité pharyngale. Cependant, nous sommes en mesure d'inférer l'utilisation de ces mécanismes. En effet, la pression est nulle lors de la production, ceci s'explique par la consigne donnée aux sujets. Nous leur avons demandé de produire du *humming* c'est-à-dire de la voix nasale. Le mécanisme lingual permet de séparer le conduit vocal en 2 parties : la cavité buccale d'une part et la cavité pharyngo-laryngale d'autre part. Ceci implique la possibilité de produire de multiples sources acoustiques. Ici le débit d'air nasal est positif. Le port vélopharyngé est donc ouvert.

Le débit d'air oral est positif. La présence de débit d'air oral est due à une activité articulaire dans la cavité buccale. Tout cela est confirmé par les images IRM de la Figure 3.8a. Sur les clichés IRM, on peut observer la formation de deux strictions dans le conduit vocal : une constriction labiale et une constriction uvulaire. On note un geste de protraction de la langue qui vient réduire le volume entre les deux strictions (compression buccale).



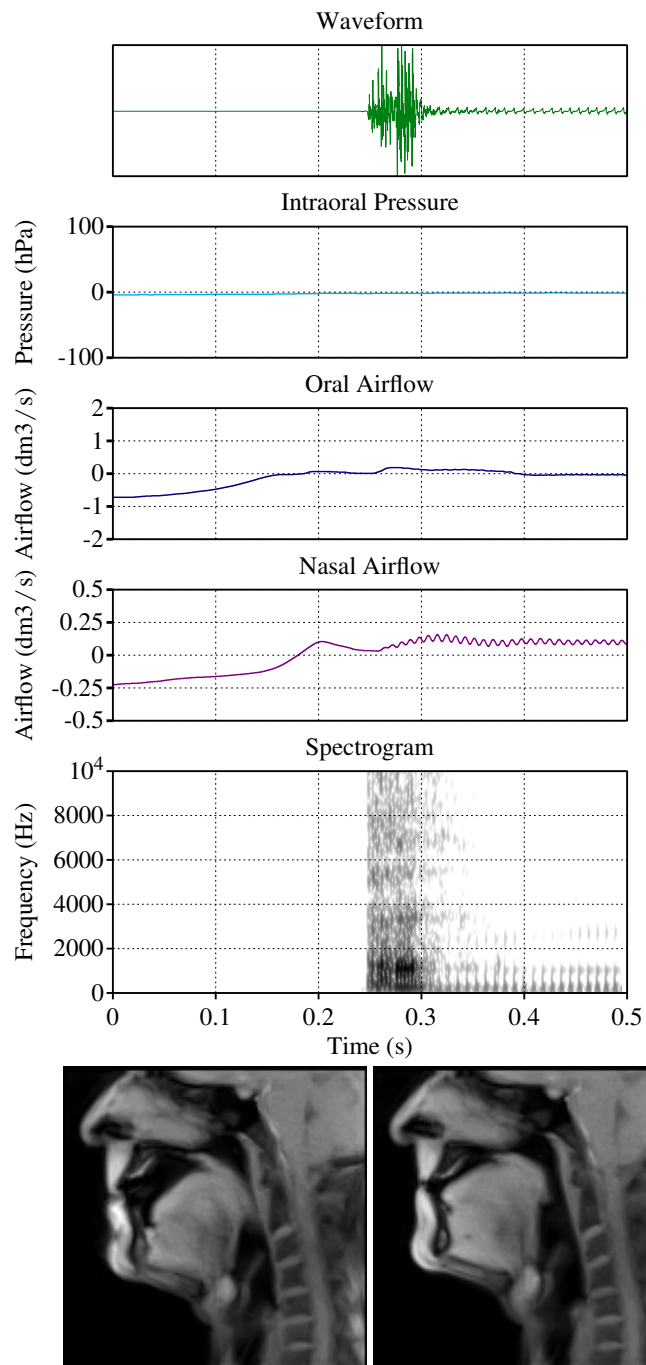
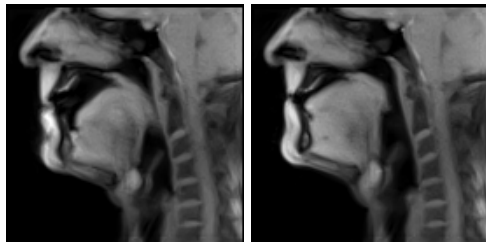


FIGURE 3.7 – *PF snare humming*  $\{\widehat{pf}\}$  - Sujet VP

⊙



(a) Début et fin de la phase d'initiation



(b) Contour de l'aire du volume

FIGURE 3.8 – Mode d'initiation par compression -  $\{\widehat{pf}:\}$   
sujet VP

### 3.1.6 Mécanismes linguaux ingressifs

Nous allons terminer l'analyse de l'initiation en mentionnant le geste d'expansion buccale. La Figure 3.9 illustre ce mécanisme lors de la production du *Lips Roll Humming*  $\{\downarrow_B^1:\}$ . On remarque une constriction labiale et une constriction « palatale » qui s'étend du palais dur dans la région palatale, jusqu'au palais mou dans la région vélaire. Cette fois, on observe un geste de rétraction linguale qui augmente le volume de la cavité entre les deux constriction. L'augmentation du volume se voit assez nettement sur la figure 3.10b. On notera que la différence de l'aire finale est 3 fois supérieure à l'initiation linguale par compression. Cette différence peut s'expliquer par le fait que le geste d'initiation de la trille linguale  $\{\downarrow_B^1:\}$  est plus long que pour la *pf-snare humming*  $\{\widehat{pf}:\}$ . Cela ne serait guère surprenant étant donné que pour initier la trille, des conditions aérodynamiques spécifiques doivent être respectées comme par exemple le différentiel de pression de part et d'autre de la constriction.

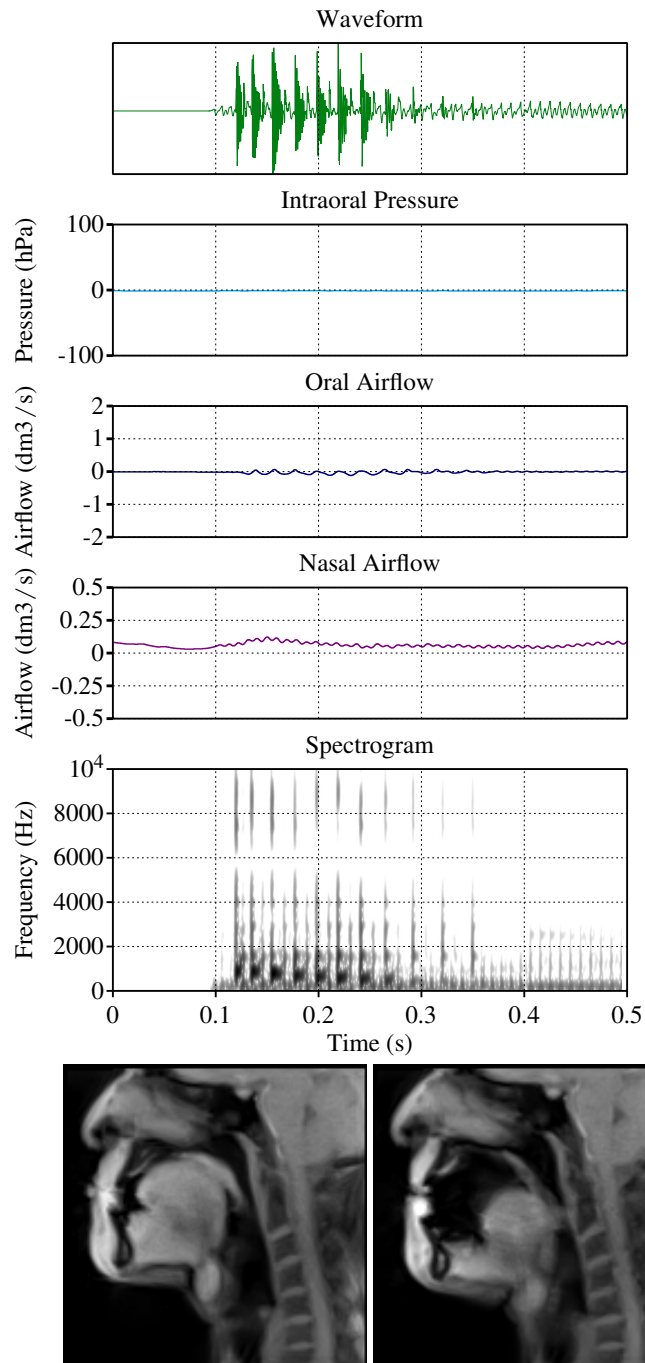


FIGURE 3.9 – *Lips Roll Humming* {↓B<sup>1</sup>:} - Sujet VP

⊙

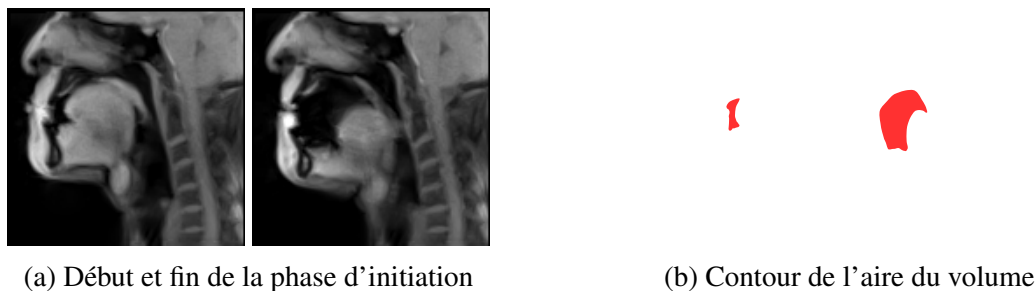


FIGURE 3.10 – Mode d'initiation par expansion -  $\{\downarrow_B^1\}$   
sujet VP

### 3.1.7 Co-initiation

Tout comme les sons peuvent-être co-articulés lorsque l'on parle, on peut co-initier un son, c'est-à-dire de produire deux mécanismes différents se chevauchant dans le temps. Cela peut paraître contre-intuitif mais la coordination de deux mécanismes d'initiation existe aussi dans la parole. Par exemple dans des langues à clics comme le Xóõ, on peut produire un mécanisme lingual ( $[Ø]$ ) avec un mécanisme laryngal ( $[Øq']$ ) ou avec un mécanisme pulmonique ( $[Øq^h]$ ). Nous avons dans nos données deux exemples de co-initiation. Le premier vient du corpus de Garrigues (2015) et est analysé dans Dehais-Underdown et al. (2019), le second a été enregistré dans le corpus de cette thèse.

La Figure 3.11a illustre la production d'un *Classic Kick* affriquée  $[p^h\phi]$ . Sur la figure, il correspond au moment où la pression est à  $+20hPa$ . Il faut rappeler ici que la pression dans l'étude de Garrigues (2015) a été mesurée dans la cavité buccale. La signature sur le signal acoustique et sur le spectrogramme est similaire à l'exemple du « *Classic Kick* » de la Figure 3.3. La première image fibroscopique correspond au début du geste d'initiation laryngale. Le débit d'air oral et le signal acoustique indiquent la présence d'une composante de friction visible sur le spectrogramme à un moment où la glotte est ouverte. Notons l'ouverture du port vélo-pharyngé à la fin du geste qui a pour effet de supprimer les fréquences supérieures à 2kHz. La fricative étant produite immédiatement après le relâchement le mécanisme de compression pulmonique doit déjà avoir activement débuté sans quoi le débit oral n'atteindrait pas  $+1dm^3/s$ .

La Figure 3.11b illustre la production par le sujet CJ d'un *Lips Roll* amorcé par un *Classic Kick*<sup>3</sup>. Le *Lips Roll* est une trille bilabiale ingressive produite à la commissure des lèvres. Contrairement à la variante humming  $\{\downarrow_B^1\}$ , le *Lips Roll* illustré ici est pulmonique. Les deux sons sont clairement distinguables sur le signal acoustique et le spectrogramme. Selon le sujet pilote VP, ces deux sons peuvent fonctionner séparément ou bien comme une seule unité. Le *Classic Kick* est facilement identifiable par les caractéristiques que nous avons mentionné plus haut. Le sujet CJ utilise un mécanisme de fermeture aryépiglottique. La deuxième image fibroscopique a été extraite juste après le *Classic Kick* à 0.2s. L'ouverture de la glotte a des effets immédiats comme le montrent les signaux de pression et de débit oral. Les signaux passent en dessous de 0 et une série d'oscillations est observée sur le signal acoustique, de pression intra-orale, de débit d'air oral et sur le spectrogramme. On peut même observer de légères oscillations sur le débit d'air nasal. Ce que nous observons ici c'est le passage d'un mécanisme d'initiation par compression laryngale à un mécanisme d'initiation par expansion pulmonique. Le passage d'un mécanisme à l'autre se traduit par un silence très court entre le *kick* et le *Lips Roll*. Le changement rapide de mécanisme retarde la mise en vibration des

3. Dans le protocole, le *Lips Roll* devait être amorcé par un *Classic Kick Inward*  $[\downarrow p]$ , le sujet CJ a substitué  $[\downarrow p]$  par  $[p']$

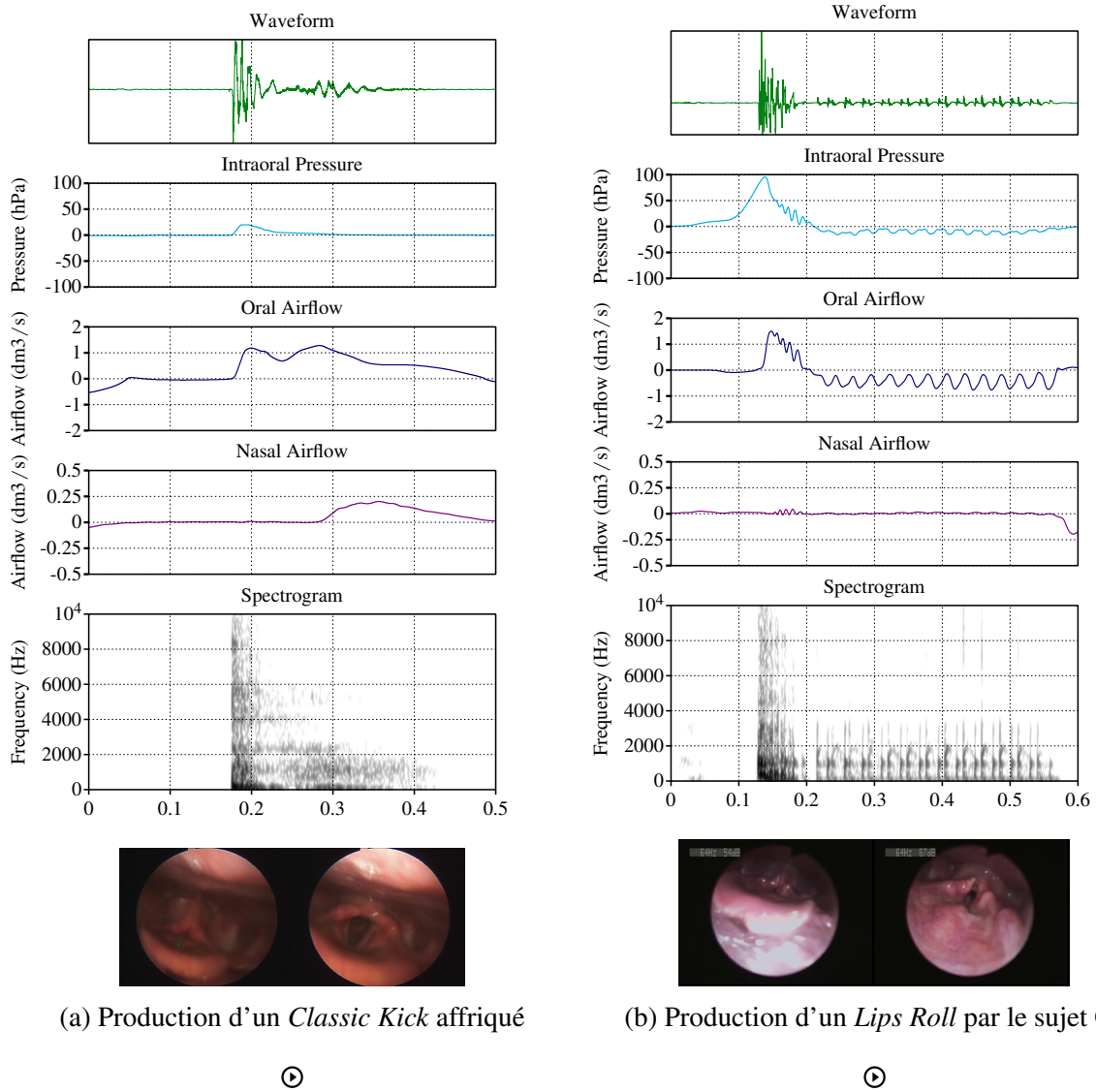


FIGURE 3.11 – Deux exemples de co-initiation lors de la production de beatbox. Le premier exemple est tiré du corpus de Garrigues (2015) et analysé dans Dehais-Underdown et al. (2019). Le deuxième exemple provient de nos données.

lèvres. D'un point de vue aérodynamique, cette latence s'explique par le même principe que la « contrainte aérodynamique sur le voisement » de Ohala (1997). Pour que l'effet Bernoulli ait lieu, il faut que la différence entre la pression sous-glottique et intraorale soit maximale  $\Delta P_{glottique} = P_{sous-glottique} - P_{intraorale} = max^4$ . Dans notre cas, le différentiel est quelque peu différent car, la glotte étant ouverte, la pression sous-glottique est la même que la pression intra-orale, nous la désignerons donc comme pression intra-orale. La deuxième pression du différentiel est la pression en dehors du conduit vocal, c'est-à-dire la pression atmosphérique à  $1030hPa$ . La pression intra-orale lors de la trille est en moyenne de  $-8.5hPa$  soit  $1021.5hPa$ . Rappelons-nous que le 0 sur nos signaux correspond à la pression atmosphérique donc  $\Delta P_{labiale} = P_{atmosphérique} - P_{intraorale} = -8.5hPa$ . Ce différentiel est maximal mais il faut 20ms pour mettre en vibration les lèvres.

Ces deux exemples illustrent l'interaction entre les mécanismes laryngaux et pulmoniques. Le dernier illustre de façon spectaculaire un mécanisme jusqu'ici non étudié qui motive et justifie la co-initiation.

---

4. Dans le cas des plis vocaux, le  $\Delta P$  est de 2hPa (Ohala, 1997).

## 3.2 Articulation

Nous venons de voir comment étaient produits les mécanismes d'initiation. Nous allons à présent les données IRM du sujet VP pour décrire comment il combine les 6 mécanismes d'initiation avec différentes stratégies articulatoires. Des données laryngoscopiques des autres sujets seront aussi présentées. Cette section se compose de 4 parties, une par catégorie d'instruments : les grosses caisses, les charlestons, les caisses claires et les *Rolls*.

### 3.2.1 Grosses caisses

La Figure 3.12 illustre les images IRM des 5 grosses caisses du corpus. Les images ont été extraites au début et à la fin des sons. Les 4 premières grosses caisses sont des occlusives bilabiales, la dernière est une occlusive alvéolaire.

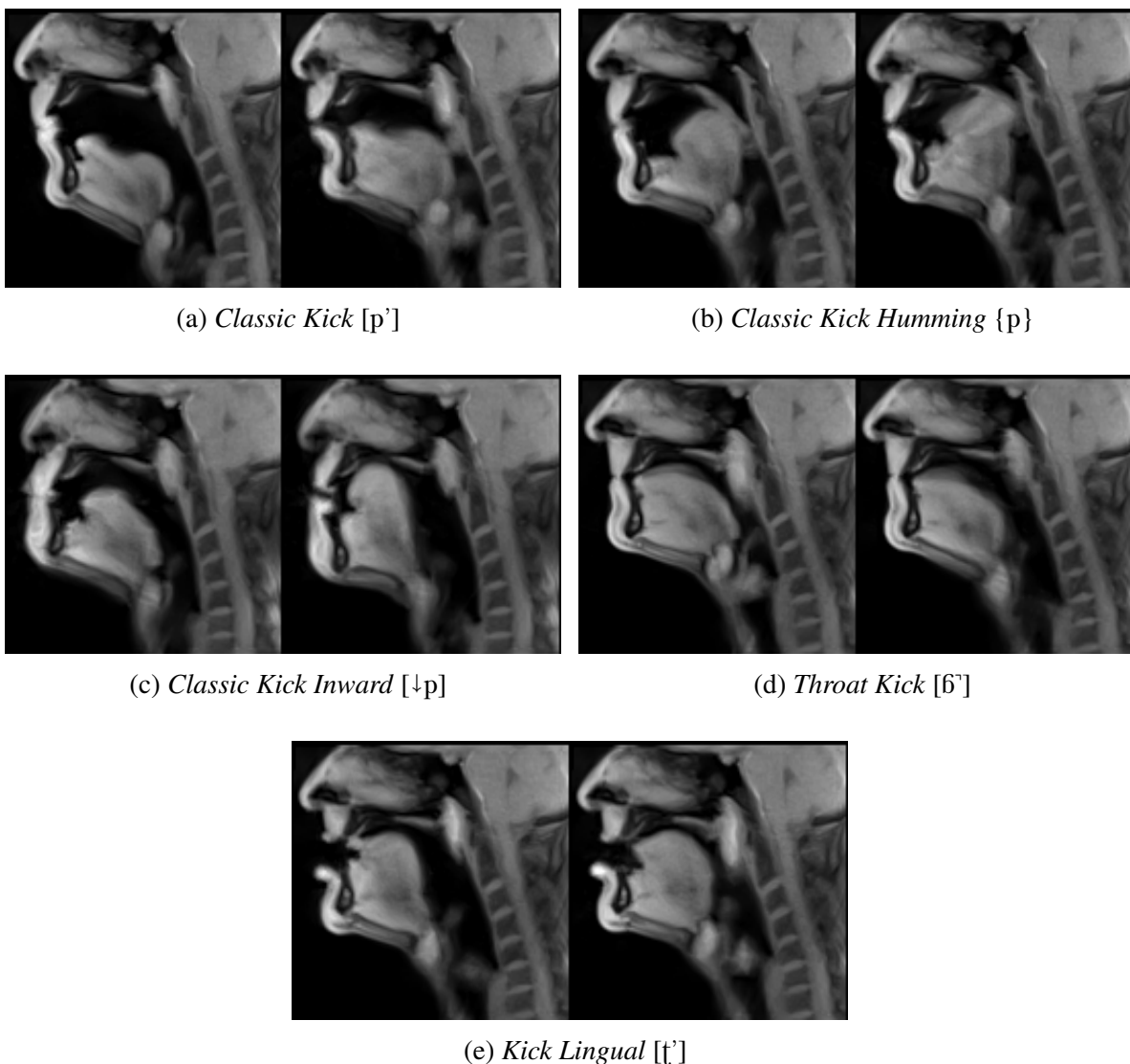
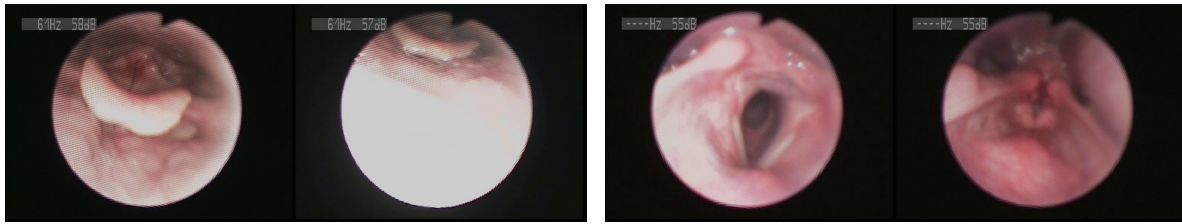


FIGURE 3.12 – Combinaison des mécanismes initiatifs et articulatoires des grosses caisses du sujet VP. La position au relâchement de [↓p] correspond au début de la trille du *Lips Roll* [↓B̥<sup>1</sup>].

Concernant le *Classic Kick* [p'], on remarque une fermeture totale des lèvres; on observe que les lèvres sont pincées en « dedans ». En effet, on remarque que la constriction se trouve



(a) Sujet AI

(b) Sujet GA

FIGURE 3.13 – Production du *Classic Kick Inward* [β] par expansion laryngale. Sujet AI : Fermeture ventriculaire puis recul du dos de la langue et abaissement du larynx. Sujet GA : Fermeture ventriculaire et abaissement du larynx.

entre les incisives. Derrière la constriction se trouve une seule cavité dont le volume est assez important. Le geste d'occlusion bilabial s'accompagne d'un geste de compression laryngal. La variante humming du *Classic Kick* {p}, montre de fortes similitudes dans le geste de fermeture des lèvres. On observe qu'elles sont pincées en dedans. On note que le geste articulatoire se combine avec un geste de compression buccale. Le *Throat Kick* [β̥] combine un geste d'occlusion bilabiale (non relâchée) et un geste d'expansion laryngale. Les données fibroscopiques et aérodynamiques des sujets AI, CJ et GA suggèrent qu'ils utilisent les mêmes combinaisons articulatoires et initiatoires pour produire les grosses caisses [p̥] {p} [β̥].

En revanche, concernant *Classic Kick Inward* [↓p], nous avons noté des différences entre les sujets. Les données IRM du sujet VP, illustre un *Classic Kick Inward*, la deuxième image correspond à la transition de [↓p] à [↓β̥<sup>1</sup>]. On note pour ce son que VP combine un geste d'occlusion labial avec un geste d'expansion pulmonique. L'expansion pulmonique n'est pas visible à l'IRM ; cependant on note l'absence de constriction glottale et laryngale sur les données IRM et fibroscopiques. La pression est en moyenne de 10hPa chez VP. Les autres sujets ont produit un *Classic Kick Inward* implosif [β] (Figure 3.13).

Sur la Figure 3.13, nous pouvons constater que la glotte est fermée et que le larynx s'abaisse. Chez le sujet AI on note un recul du dos de la langue. Nous n'avons pas de données fibroscopiques pour le sujet CJ. Cependant, la pression est en moyenne de 50hPa ce qui suggère l'utilisation d'un mécanisme d'expansion laryngale. La pression a saturé pour le sujet AI. Chez GA la pression moyenne est de 50-60hPa, soit au même niveau que CJ.

Enfin, il y a le *Kick Lingual*. Cette grosse caisse est initiée par une compression laryngale comme le montre la différence de hauteur du larynx entre le début et la fin de la production et la pression élevée ( $\pm 30$ hPa). On note un contact de la langue avec le palais dans la région alvéolaire. Notons aussi que c'est l'apex de la langue qui est en contact. La langue présente des signes de rétroflexion, de protrusion labiale et de recul du dos de la langue : la rétroflexion s'observe sur la partie antérieure de la langue dont la forme crée une cavité sublinguale. La labio-vélarisation est visible au niveau des lèvres qui sont protrusées et à l'arrière du conduit où le dos de la langue présente des similitudes avec une configuration de type [u]. On note l'apparition d'une cavité antérieure due à la rétroflexion et à la protrusion des lèvres. L'allongement de la cavité antérieure a pour effet de baisser la fréquence de résonance du bruit. Une transcription fine de ce son serait en fait le symbole rétroflexe et le diacritique de protrusion labio-vélarisation : [t<sup>w'</sup>].



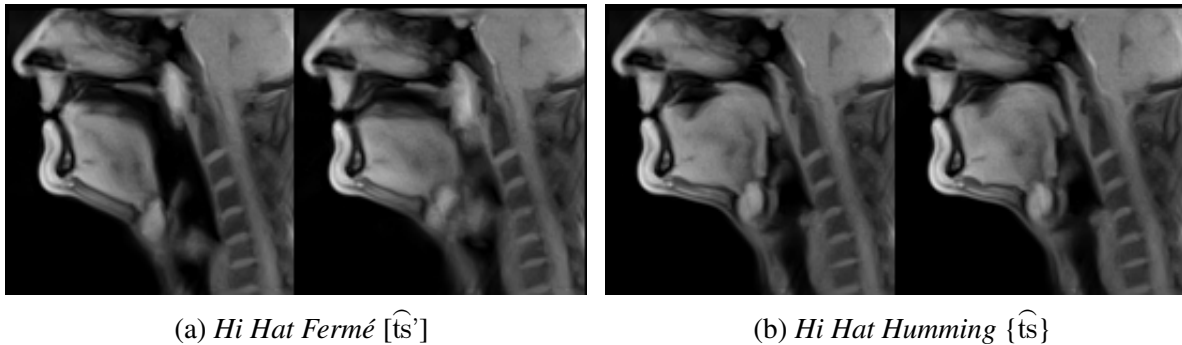


FIGURE 3.14 – Combinaison des mécanismes initiateurs et articulatoires des charlestons du sujet VP.

### 3.2.2 Charlestons

La Figure 3.14 illustre la production d'un charleston fermé [ts̥] et de sa variante humming {ts̥}. Concernant le « *Hi-Hat fermé* » le lieu d'articulation est dental. L'image à gauche montre le moment de constriction maximale, c'est-à-dire l'occlusion. On note que la langue est en contact avec les dents. Derrière la constriction dentale, nous n'observons aucun signe de constriction. Au début du geste d'initiation, le larynx se trouve au niveau de la 3ème vertèbre cervicale alors qu'à la fin du geste, nous le retrouvons en position plus haute au niveau de la 2ème vertèbre cervicale. Si l'on compare la position initiale et la position finale des articulateurs, hormis le relâchement de la langue et l'élévation du larynx, nous n'observons aucun mouvement.

Les données du *Hi-Hat Humming* montrent que l'occlusion est dentale. La configuration dans la cavité orale confirme bien l'utilisation d'un mécanisme lingual. On remarque en effet la création d'une constriction postérieure, dans la zone vélaire et uvulaire. Le corps de langue s'élève et comprime le volume se trouvant entre les deux constriction.

### 3.2.3 Caisses claires

Nous allons présenter les caisses claires dites « orales » et les caisses claires dites « laryngales ». Pour rappel, les caisses claires laryngales sont des caisses claires combinant une occlusion laryngale et une phase de friction glottale ou orale. Les caisses claires orales combinent un geste occlusif et fricatif sur un même lieu d'articulation oral.

#### 3.2.3.1 Caisses claires orales

La Figure 3.15 illustre les images IRM de 5 caisses claires orales de notre corpus. On trouve les deux caisses claires vélares [↓k̠L:] et {↓k̠L:}; les deux caisses claires labiodentales [p̠f:] et {p̠f:} et une caisse claire post-alvéolaire [t̠ʃ:]. Pour chaque son, la première image illustre l'occlusion. La deuxième image illustre la phase de friction.

Il est difficile d'analyser les *K-Snare Inward* [↓k̠L:] et {↓k̠L:} à partir des données IRM étant donné que le point d'articulation est latéral. Nos données sont prises dans un plan médio-sagittal ce qui permet seulement d'observer les articulations centrales. Le caractère latéral de ces sons est expliqué par le sujet pilote VP et confirmé par les autres sujets de l'étude. Nous l'avons vu précédemment, la *K-Snare Inward* [↓k̠L:] est initiée par expansion pulmonique. On observe une constriction palatale qui s'étend des dents jusqu'au palais dur. Selon le sujet VP, il s'agit

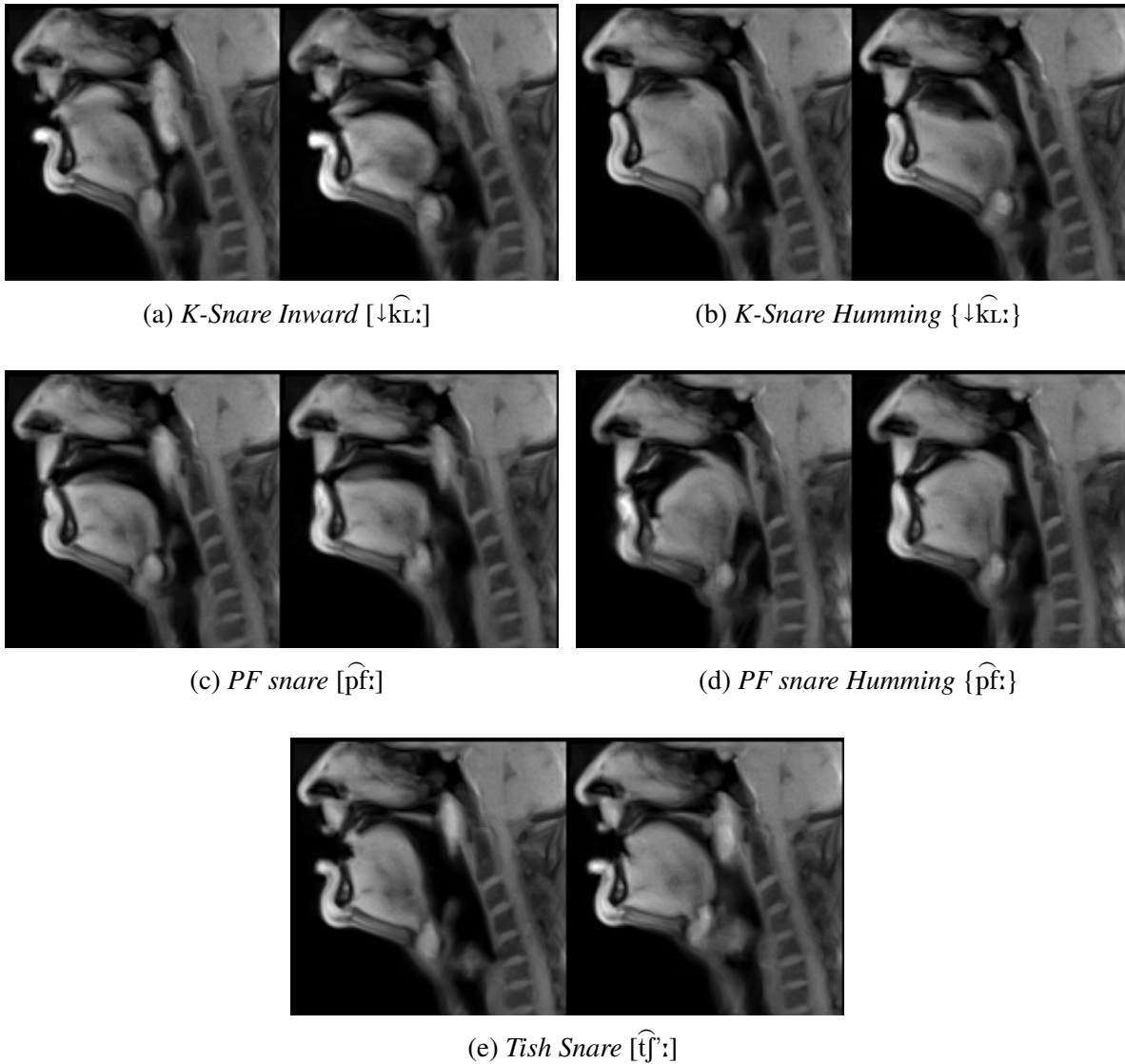


FIGURE 3.15 – Combinaison des mécanismes initiativés et articulatoires des caisses claires orales du sujet VP.

d'un « point d'appui » pour produire le son. Le lieu d'articulation (c'est-à-dire le lieu où le flux d'air est modifié) se trouve dans la zone vélaire au niveau des molaires. Selon VP, le dos de la langue crée une occlusion au niveau des molaires droites. Au relâchement, il maintient la constriction en ce lieu pour générer un bruit de friction. Tout comme les incisives peuvent former un obstacle, les molaires peuvent aussi faire office d'obstacle permettant de rendre un flux turbulent. On retrouve le caractère latéral non seulement au niveau de l'occlusion vélaire mais aussi au niveau de la mandibule. Nos données montrent un changement de position de la mâchoire entre le début et la fin du son (Figure 3.16). En effet, on remarque que la mâchoire se déplace vers le bas et la droite. Le déplacement de la mâchoire doit avoir un effet sur la position de la langue dont les effets acoustiques doivent être expliqués. L'instrumentation utilisée ne permet pas d'analyser la dynamique articulatoire des sons latéraux mais nous tenterons tout de même d'en analyser les effets acoustiques.

Le sujet AI n'a pas produit la *K-Snare Inward* par expansion pulmonique mais par expansion laryngale (Figure 3.17). Les trois images ont été extraites avant l'occlusion, au moment du



FIGURE 3.16 – Mouvement latéral de la mâchoire pendant la production de  $[\downarrow\widehat{k}L:]$  du sujet VP

relâchement de l'occlusion vélaire et pendant la phase de friction. On remarque que les configurations sur la première et la dernière image sont similaires, mais sur la dernière image, le larynx s'est abaissé. L'image du milieu montre clairement l'utilisation d'un mécanisme implusif : insertion ventriculaire et descente du larynx.

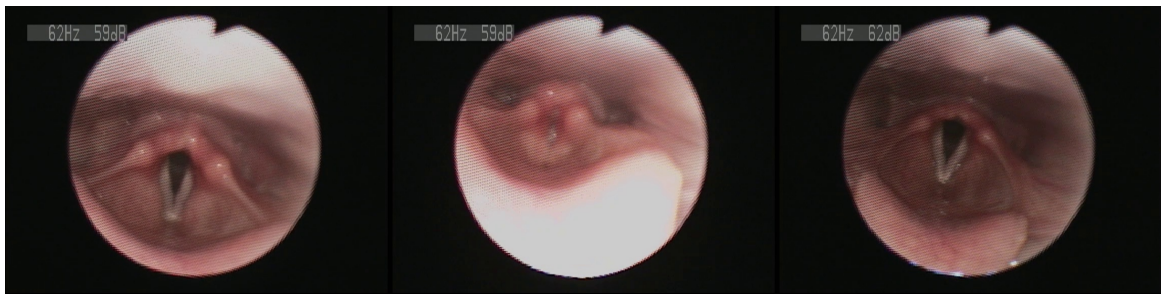


FIGURE 3.17 – Production d'une *K-Snare Inward*  $[\downarrow\widehat{k}L]$  du sujet AI par expansion laryngale

La *K-Snare Humming*  $\{\downarrow\widehat{k}L:\}$  est initiée par expansion buccale, comme l'indiquent les images IRM. On remarque deux points d'occlusion dans le conduit vocal : une constriction dentale et une constriction vélaire/uvulaire. Notons que, contrairement aux clics, la constriction antérieure est maintenue tout du long et que la constriction postérieure est relâchée en premier<sup>5</sup>. L'occlusion vélaire que nous observons sur l'image de gauche correspond au geste d'initiation. Au moment du relâchement, la langue se trouve dans la région post-vélaire voire uvulaire.

Chez le sujet VP, la *PF Snare* est initiée par une compression pulmonique. La lèvre inférieure est pressée sur les incisives supérieures (celles-ci n'apparaissent pas sur les images IRM). Derrière la constriction, nous n'observons pas la même configuration que pour  $[p']$ . Étant donné que l'initiation est pulmonique dans le cas de  $[\widehat{p}f:]$ , le larynx ne descend pas (2<sup>e</sup> vertèbre cervicale). Si l'on compare le début et la fin du geste articulatoire, on note un déplacement de la mâchoire vers le haut et l'avant. La production de la caisse claire  $[\widehat{p}f:]$  implique donc un déplacement du point d'articulation de la zone labiodentale à la zone bilabiale pendant la production. Chez les sujets CJ et GA, la pression intra-orale est supérieure à 50hPa et indique l'utilisation d'un mécanisme éjectif  $[\widehat{p}f:]$ . La variante humming  $\{\widehat{p}f:\}$  présente les mêmes caractéristiques

5. Tuhuse et Traill (1999) ont rapporté l'utilisation d'un clic similaire où la constriction postérieure est relâchée avant la constriction palatale. Les auteurs transcrivent le clic ainsi :  $[+\ast]$ .

articulatoires que sa correspondante pulmonique [p̄f:] mais est initié par compression buccale.

La « *Tish Snare* » [tʃ̄:] est initiée par compression laryngale comme le montrent les différences de position du larynx entre la phase occlusive (image de gauche) et la phase fricative (image de droite). La phase d'occlusion est similaire à celle que l'on vient d'observer pour le *Kick Lingual* [t̄]. On remarque que la pointe de la langue crée une constriction dans la zone alvéolaire. La forme de la langue présente également des signes de rétroflexion et les lèvres sont protrusées. La cavité antérieure se retrouve donc allongée et provoque un abaissement de la fréquence du bruit de relâchement de l'occlusion. Cette fois-ci, aucun signe de vélarisation n'est observé sur nos données. En revanche, on note un signe de palatalisation présent à la fois sur la partie occlusive et sur la partie fricative. La partie fricative se caractérise par une constriction post-alvéolaire une fois que la pointe de la langue s'abaisse. La constriction est réalisée avec la lame de la langue tandis que le corps de la langue reste massé dans la zone palatale. La palatalisation des consonnes a pour conséquence d'augmenter la hauteur du pic spectral.

### 3.2.3.2 Caisses claires laryngales

La Figure 3.18 illustre les images IRM des caisses claires laryngales de notre corpus : la *ESH Snare Outward* [ʃ̄:], la *ESH Snare Inward* [↓ʃ̄:] et la *Cough Snare* [ʃ̄h:].

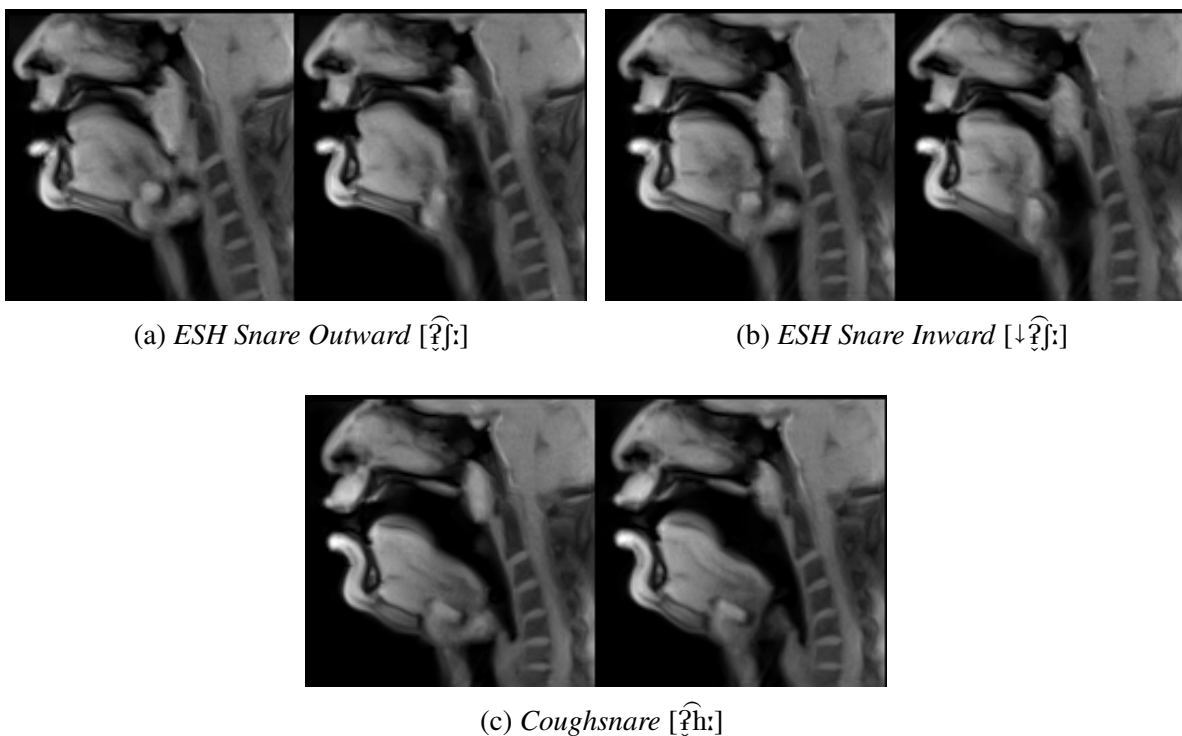


FIGURE 3.18 – Combinaison des mécanismes initiateurs et articulatoires des caisses claires laryngales du sujet VP.

La *ESH Snare Outward* du sujet VP combine trois mécanismes : une occlusion épilaryngale, une fricative post-alvéolaire et un geste de compression pulmonique. On peut aussi noter un geste de protrusion des lèvres. La production de la *ESH Snare Outward* débute avec une occlusion du tube épilaryngal combinant une fermeture aryépiglottique, un recul de la racine de la langue, une compression des parois latérales du pharynx et une tubularisation de l'épiglotte (Figure 3.19 - Sujet VP). Le relâchement des caisses claires laryngales est visé. La phase fricative se caractérise par un tube épilaryngal ouvert et une constriction post-alvéolaire.

La Figure 3.19 montre des différences individuelles systématiques de la configuration laryngale pendant l'occlusion. Chez le sujet AI, les plis vocaux et les bandes ventriculaires sont en adduction, le tube épilaryngal est ouvert. On note aucun signe de constriction aryépiglottique. Chez CJ, on note une constriction aryépiglottique comme le montre bascule antérieure des aryénoïdes sur la base de l'épiglotte. On ne note pas de signe de compression du tube épilaryngal par les parois latérales du pharynx. En revanche, pour GA on note une participation des parois latérales du pharynx en plus de la fermeture aryépiglottale.

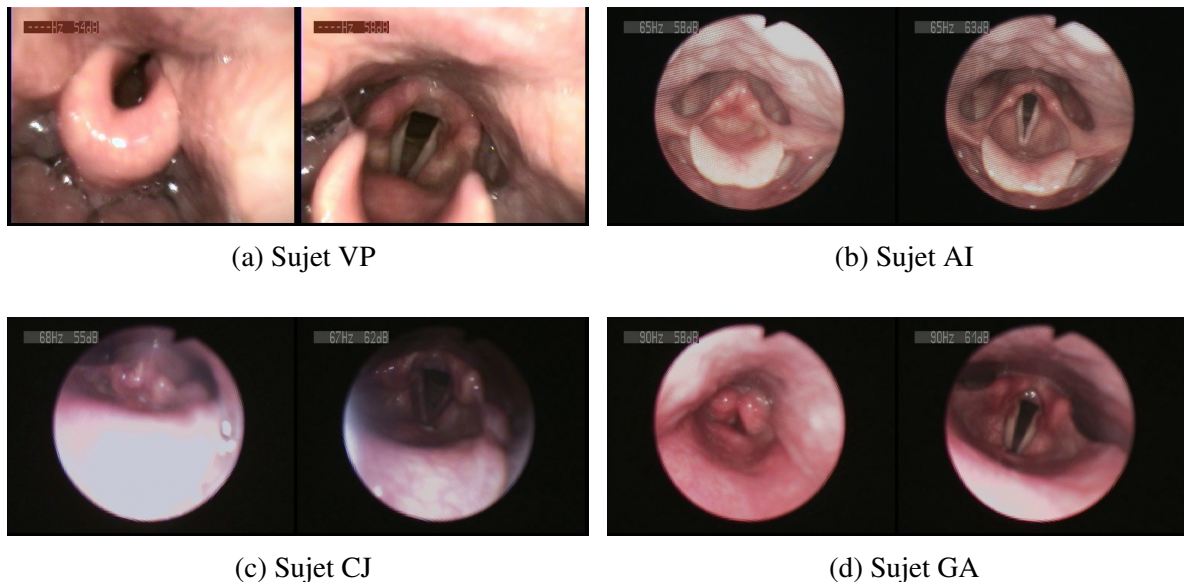


FIGURE 3.19 – Configuration laryngale pendant l'occlusion et le relâchement de  $[\downarrow\hat{t}]:$

La *ESH Snare Inward*  $[\downarrow\hat{t}]:$  présente une configuration similaire à sa correspondante égressive sur les images IRM. Nous ne disposons pas de données fibroscopiques pour les sujets VP et AI. Les données fibroscopiques des sujets CJ et GA (Figure 3.20) montrent que les configurations sont presque identiques à leur correspondante égressive  $[\hat{t}]:$ . Le larynx est plus bas dans le cas de la variante ingressive que dans le cas de la variante égressive.

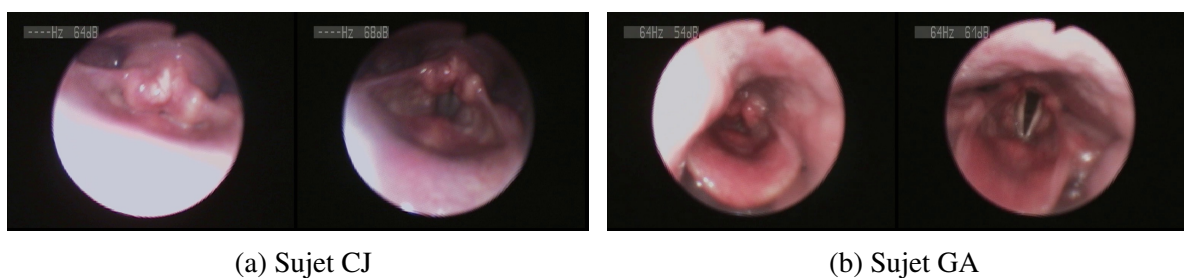


FIGURE 3.20 – Configuration laryngale pendant l'occlusion et le relâchement de  $[\downarrow\hat{t}]:$

Enfin, la *Cough Snare* se caractérise par la combinaison d'un geste d'occlusion et de friction épilaryngal et d'un geste de compression pulmonique. De même que pour les caisses claires laryngales précédentes, il y a des indices de protrusion labiale. On n'observe pas de constriction supralaryngale.

La Figure 3.21 indique encore une fois des différences systématiques de la configuration laryngale entre les sujets. Les configurations sont similaires à celles décrites pour la *ESH Snare Outward* et *inward*.



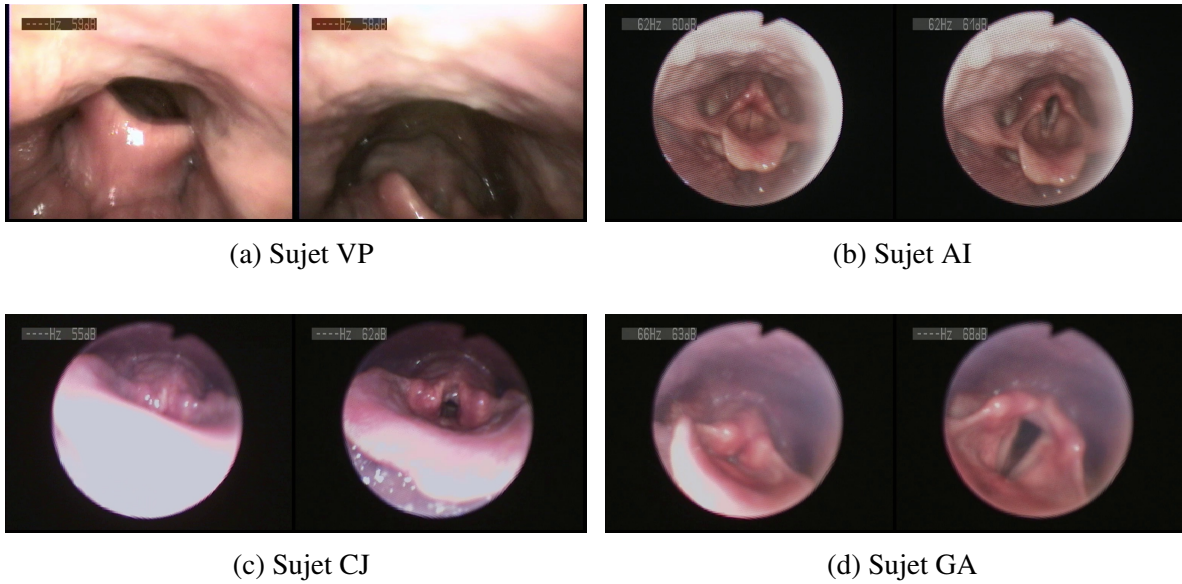


FIGURE 3.21 – Configuration laryngale pendant l’occlusion et le relâchement de [ʔh:].

### 3.2.4 Rolls

La Figure 3.22 illustre 2 des 3 *Rolls* de notre corpus. Le *Clic Roll* n’a pas été produit à l’IRM par VP. Le *Lips Roll* [ $\downarrow_B^1$ ] est produit sur un flux d’air pulmonique ingressif. On observe que les lèvres et la langue sont impliquées dans la production du son chez ce sujet. Les lèvres sont accolées afin de permettre leur oscillation sous l’effet de Bernoulli. La position de la langue joue en fait un rôle dans la manipulation du timbre résultant. On peut masser la langue à l’avant du conduit pour obtenir un timbre clair ou on peut la masser à l’arrière du conduit pour obtenir un timbre sombre. Le *Lips Roll Humming* combine une trille bilabiale et un mécanisme d’expansion buccale que nous avons déjà analysé précédemment.

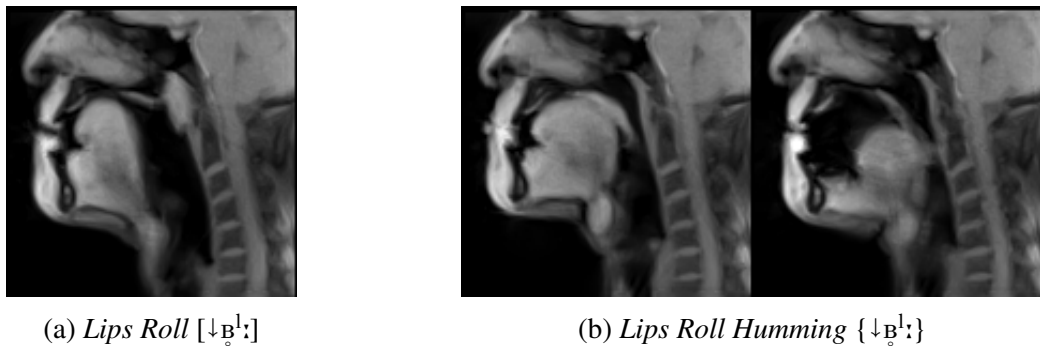


FIGURE 3.22 – Combinaison des mécanismes initiateurs et articulatoires des *Rolls*.

### 3.3 Timbre

Nous venons de voir que les sons de notre corpus combinaient différents gestes de production : un ou plusieurs gestes articulatoires et un ou plusieurs gestes initiatifs ([↓K̂L]). Nous allons maintenant analyser le timbre des sons de notre corpus pour montrer que les combinaisons précédemment décrites donnent lieu à des signatures acoustiques discrètes. Nous reprendrons l'analyse par catégorie de sons. Pour chaque catégorie nous présenterons les productions d'un seul sujet.

#### 3.3.1 Grosses Caisses

Pour analyser les grosses caisses, nous analyserons les productions du sujet CJ. Nous n'analyserons pas le *throat kick*. Ce dernier se caractérise par un pic unique (le fondamental) vers 200-250Hz. Les grosses caisses produites par les boîtes à rythme (les « *beatbox* ») et les *beatboxeurs* sont des imitations des grosses caisses de la batterie. Les caractéristiques acoustiques de cet instrument percussif sont bien connues. Il s'agit d'un instrument percussif à deux peaux que l'on frappe. Le timbre résultant est grave et bruyant (Leipp, 1971).

#### Le Classic Kick

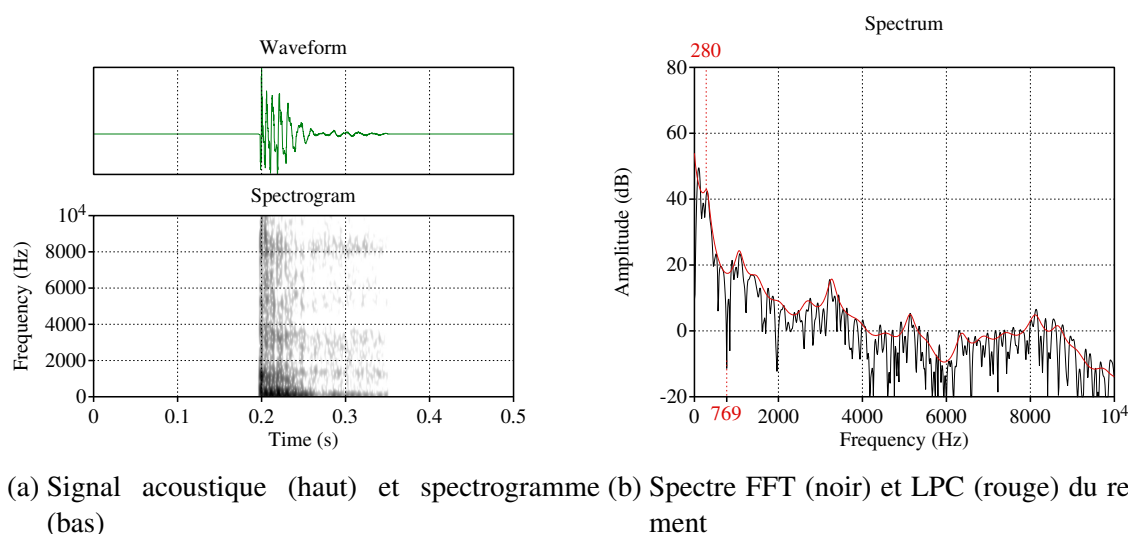
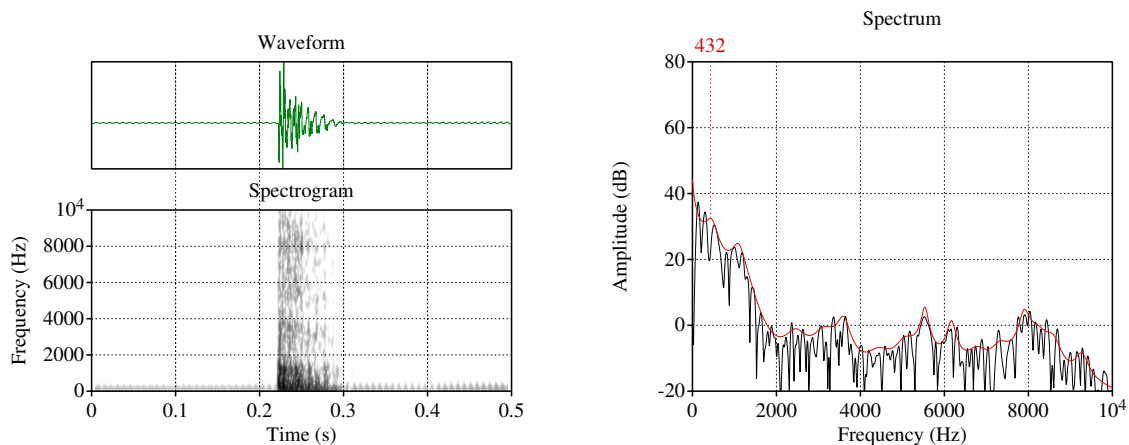


FIGURE 3.23 – Timbre du *Classic Kick* [p'ϕ] - Sujet GA



La Figure 3.23 illustre les caractéristiques acoustiques de la variante affriquée [p'ϕ] du *Classic Kick*. À gauche se trouve le signal acoustique et le spectrogramme correspondant ; à droite le spectre FFT (noir) et un lissage LPC à 24 coefficients (rouge) des 50 premières millisecondes du relâchement. Le spectre est diffus et descendant avec de l'énergie de 0 à 10kHz. Il y a une résonance intense à 280Hz. Les résonances suivantes sont de moins en moins intenses ; elles se trouvent à 1kHz, 3kHz et 5kHz. On note aussi une première anti-résonance à 769Hz et à 6kHz. Les caractéristiques spectrales concordent avec celles des occlusives labiales, c'est-à-dire de l'énergie en-dessous de 1.5kHz. Le timbre résultant est similaire à celui d'une grosse caisse.

## Le Classic Kick humming



(a) Signal acoustique (haut) et spectrogramme (bas)  
(b) Spectre FFT (noir) et LPC (rouge) du relâchement

FIGURE 3.24 – Timbre du *Classic Kick humming* {p}- Sujet GA

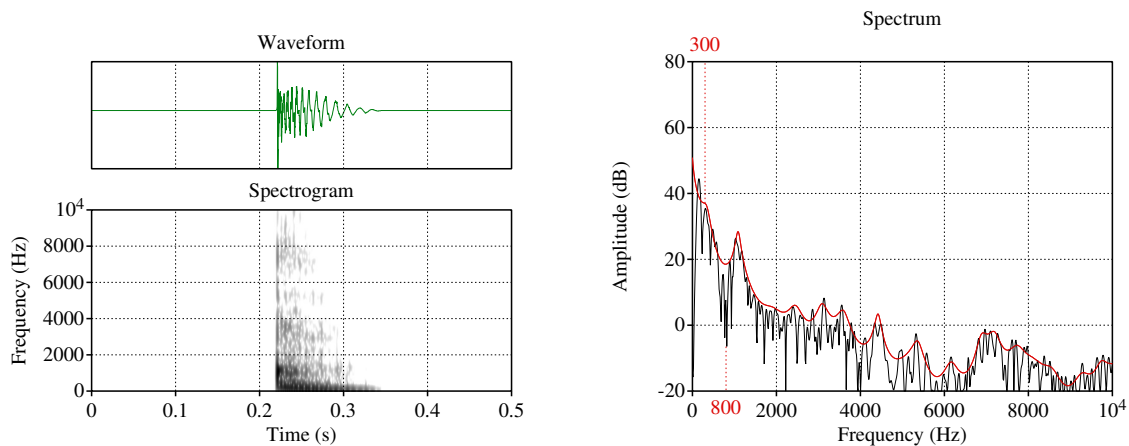


La Figure 3.24 illustre les caractéristiques spectrales du *Classic Kick humming* produit par le sujet GA. La barre de voisement sur le signal et le spectrogramme est due au *humming*. On remarque une première résonance à 432Hz. Les autres résonances se trouvent à 1kHz, 3.6kHz et 5.5kHz. Le spectre FFT indique une répartition diffuse et descendante avec de l'énergie de 0 à 10kHz. Comme nous l'avons souligné dans le chapitre méthodologique, nous avons rencontré des problèmes de paramétrage des seuils d'amplitude de nos signaux et les différences d'amplitude sont peut-être dues à différents seuils d'amplification du signal. Cependant, le niveau d'amplitude globale du spectre du *Classic Kick humming* semble plus bas que sa correspondante laryngale. Nous pensons qu'il s'agit d'une différence entre les deux mécanismes d'initiation. Paroni et al. (2021) ont aussi trouvé des différences d'intensité entre ces deux variantes.

## Le Classic Kick Inward

La Figure 3.25 illustre les caractéristiques du *Classic Kick Inward* [β]. Tout comme pour [p'] et {p}, le spectre est diffus et descendant avec de l'énergie entre 0 et 10kHz. Cette variante se caractérise ici par une atténuation plus importante de l'énergie après 4kHz. Il y a une première résonance à 300Hz, une résonance à 1kHz et des résonances affaiblies entre 2kHz et 4kHz. La première résonance est plus plate que les autres variantes et la résonance à 1kHz est plus pointue en comparaison avec les variantes [p'] et {p}. Il y a une anti-résonance très marquée à 800Hz. Cette anti-résonance est similaire à celle mentionnée pour la variante laryngale égressive du *Classic Kick*. Cette anti-résonance n'est pas introduite par la cavité postérieure étant donné que la direction du flux est inversée et qu'il n'y a pas de cavité derrière les lèvres. Elle est introduite par la cavité antérieure, c'est-à-dire la cavité entre les lèvres et la glotte.



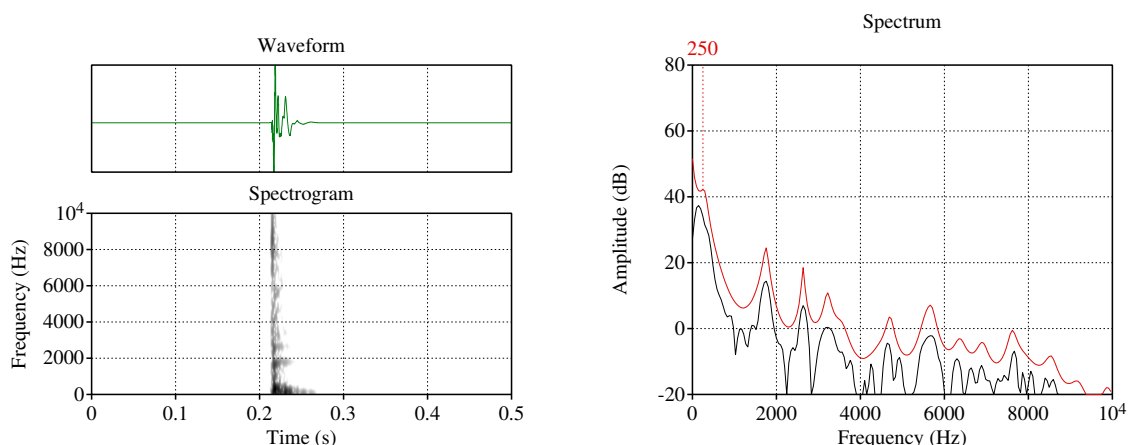


(a) Signal acoustique (haut) et spectrogramme (bas) (b) Spectre FFT (noir) et LPC (rouge) du relâchement

FIGURE 3.25 – Timbre du *Classic Kick Inward* [β] - Sujet GA



### Le Kick Lingual



(a) Signal acoustique (haut) et spectrogramme (bas) (b) Spectre FFT (noir) et LPC (rouge) du relâchement

FIGURE 3.26 – Timbre du *Kick Lingual* [t'] - Sujet GA



La Figure 3.26 illustre les caractéristiques du *Kick Lingual* produit par le sujet GA. Nous avons vu que le sujet VP utilisait une stratégie de labio-vélarisation et de rétroflexion. Cette stratégie permet d'abaisser les fréquences de résonance dans le conduit vocal. Bien que nous n'ayons pas d'images IRM pour le sujet GA, nous pensons que les stratégies de labio-vélarisation et de rétroflexion s'appliquent ici. Le spectre FFT est encore une fois diffus et descendant. Le spectre se caractérise par une faible présence de composantes fréquentielles assez espacées. Le spectre a été mesuré sur une fenêtre de 25ms, soit deux fois moins que la taille des fenêtres précédentes. Les résonances se trouvent à 250Hz, 1.7kHz et 2.6kHz et leur amplitude diminue

en fonction de la fréquence. La signature acoustique de cette occlusive coronale ne correspond donc pas au spectre diffus ascendant des consonnes coronales.

### Timbre des grosses caisses

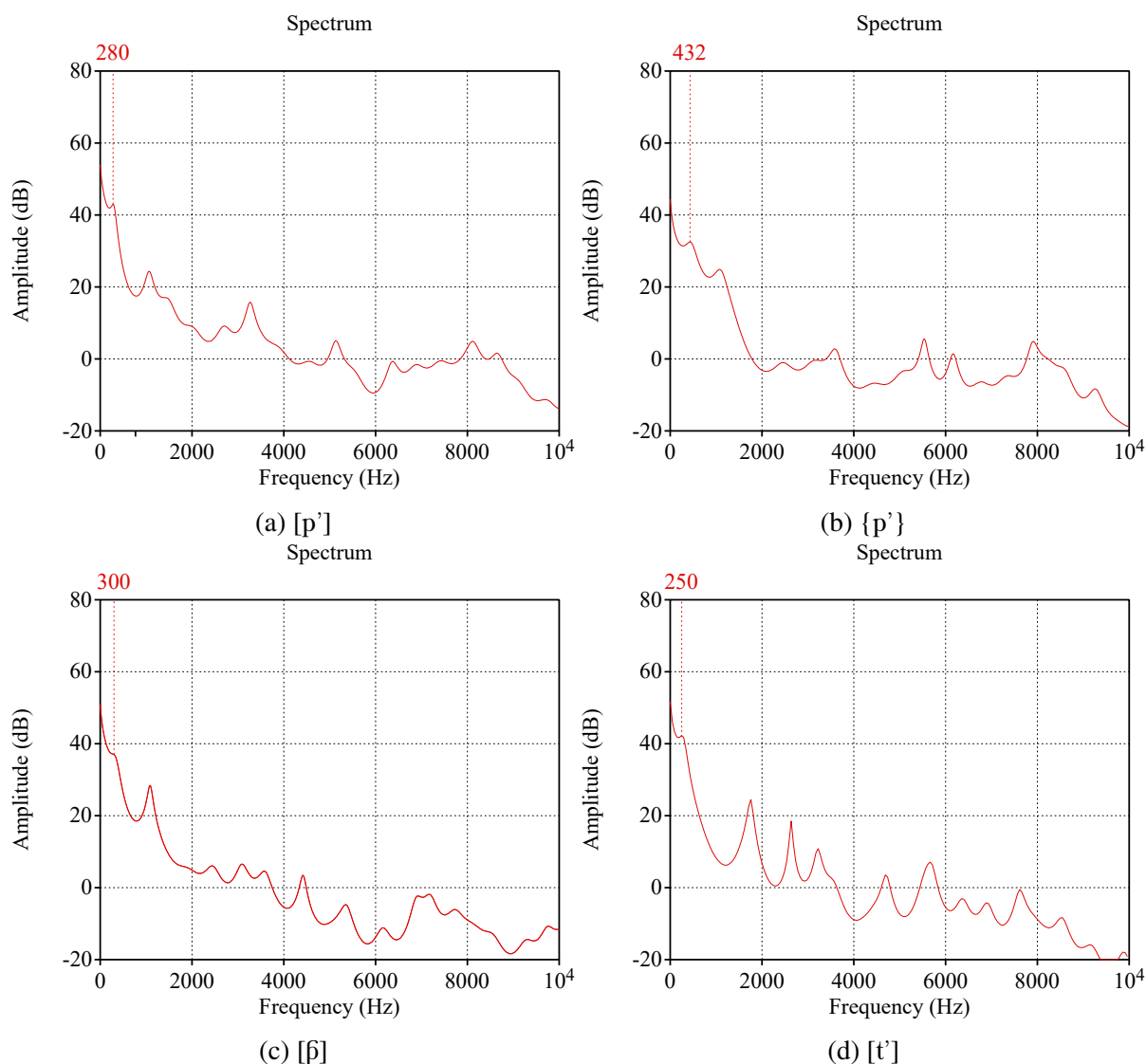


FIGURE 3.27 – Comparaison du timbre des grosses caisses du sujet GA à partir de spectre LPC à 24 coefficients

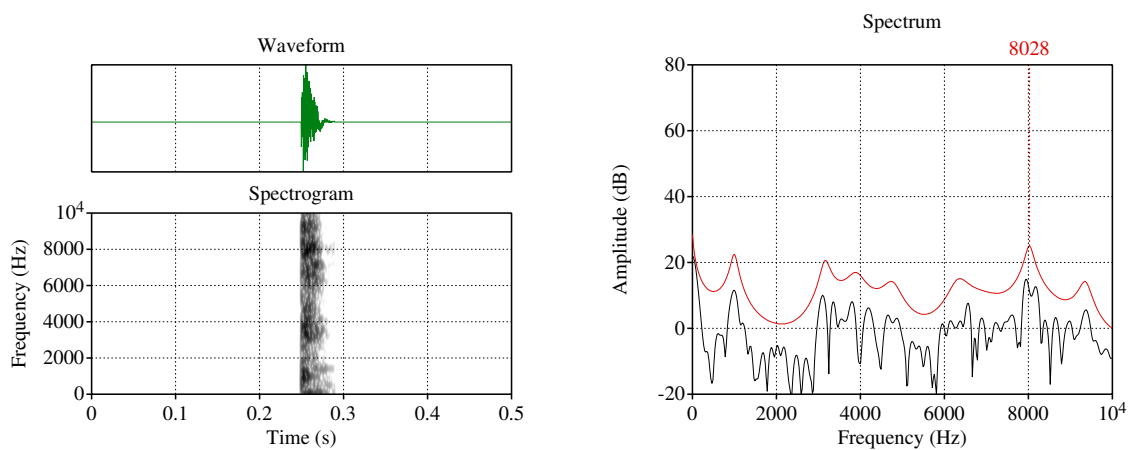
Pour caractériser le timbre des grosses caisses, comparons les spectres LPC à 24 coefficients (Figure 3.27). Bien que les stratégies d’initiation et d’articulation soient diverses, toutes les grosses caisses sont caractérisées par un spectre diffus et descendant, c’est-à-dire que plus la fréquence augmente, plus l’amplitude diminue. En d’autres termes, l’énergie des grosses caisses se trouve dans les basses fréquences afin de se rapprocher (perceptuellement) au mieux du timbre d’une grosse caisse. Perceptuellement, il peut être difficile pour un non-beatboxeur de différencier à l’oreille ces différentes grosses caisses, à l’exception peut-être de la variante linguale. Toutefois, les beatboxeurs et beatboxeuses arrivent à les différencier. Il doit donc y avoir des indices acoustiques permettant de catégoriser ces différentes grosses caisses. Le *Kick Lingual* [t'] se distingue assez nettement des grosses caisses labiales, elle possède une résonance à 250Hz suivie d’une antirésonance assez large et un pic à 1.7kHz.

En comparant les spectres a) et b) on remarque une pente spectrale abrupte à 1kHz pour {p} et un spectre plat jusqu'à 10kHz. Ceci donne un aspect plus compact des fréquences en dessous de 2kHz. Enfin, la différence entre [p'] et [β] est plus ténue. Les deux spectres sont très similaires. On note que pour [p'] il y a 3 résonances assez nettes entre 1kHz et 5kHz alors que pour [β] on remarque une succession de plusieurs pics assez faibles. Une étude perceptive à ce sujet apporterait sans aucun doute des éléments de réflexion sur la catégorisation des différents types de grosse caisse et sur les différents mécanismes d'initiation et d'articulation.

### 3.3.2 Charlestons

Contrairement aux grosses caisses, le charleston n'est pas un instrument à peau mais deux cymbales métalliques fixées sur un système de pied à pédale. En appuyant sur la pédale les deux cymbales viennent se percuter. Sur un spectrogramme, la signature acoustique des cymbales comporte un choc suivi de raies. Contrairement à la cymbale, l'amortissement de l'onde du charleston fermé est plus rapide.

#### Le Hi-Hat fermé



(a) Signal acoustique (haut) et spectrogramme (bas) (b) Spectre FFT (noir) et LPC (rouge) de la friction

FIGURE 3.28 – Timbre du *Hi-Hat fermé* - Sujet AI



La Figure 3.28 illustre le timbre du *Hi-Hat fermé* [ts'] du sujet AI. Les consonnes dentales et alvéolaires se caractérisent par un spectre diffus-montant et un pic spectral à environ 4-5kHz. On ne remarque pas ici un spectre diffus-montant. En effet, le niveau d'amplitude des pics spectraux est assez similaire sur l'ensemble du spectre avec une amplitude plus importante du pic à 8kHz ( $\pm 4$ dB). Le caractère plat du spectre peut s'expliquer par la résistance à l'écoulement de l'air. En effet, intensité acoustique et débit d'air sont corrélés (Hixon, 1966; Isshiki, 1965; Stevens, 1971, 1989). Fant (1970) suggère que l'intensité acoustique des flux turbulents est proportionnelle au carré de la vitesse des particules et dépend aussi de la taille de la constriction.

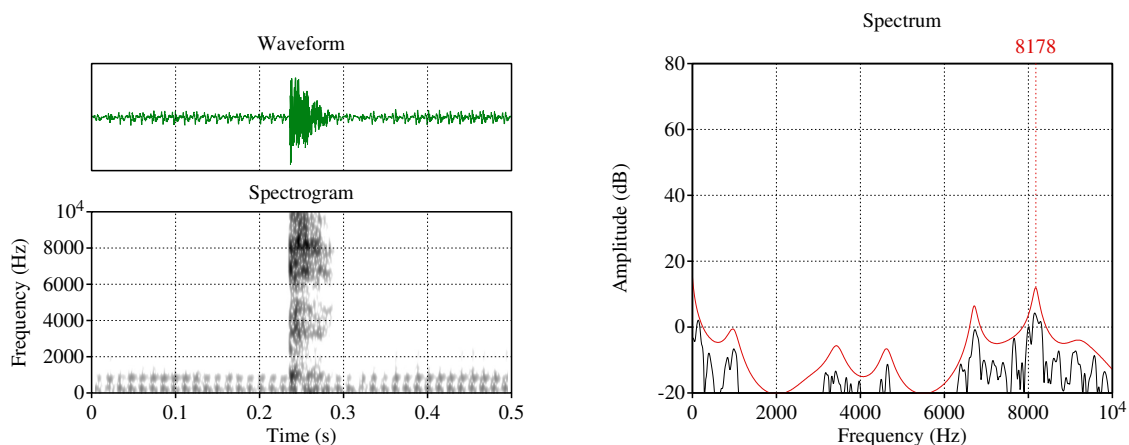
La résistance orale est le ratio de la pression (hPa) sur le débit ( $cm^3/s$ )  $P/U$ . Pour [ts'] la pression intraorale est de 60hPa et le débit d'air oral de 200  $cm^3/s$ , ici la résistance à l'écoulement d'air est forte  $60/200 = 0.3$ . Ce qui pourrait expliquer la résistance à l'écoulement de

l'air peut, selon nous, être la taille de la constriction. L'aire relative de la constriction ( $cm^2$ ) est estimée en faisant le rapport du débit sur la vitesse des particules  $A = U/v$  où  $U$  est le débit d'air oral mesuré ( $cm^3/s$ ) et  $v$  la vitesse des particules ( $cm/s$ ) estimée à partir de la pression intraorale mesurée  $v = 412\sqrt{P_{intraorale}}$ . La constriction linguale est donc de petite taille :  $0.02cm^2$ . À propos des constriction inférieures à  $0.03cm^2$ , Stevens (1989, p. 21-22) suggère que « l'amplitude de la source se réduit à cause de la réduction du débit d'air ». Il paraît donc possible que ce soit la taille de la constriction qui induise une résistance à l'écoulement de l'air et une réduction du débit d'air et de l'intensité acoustique.

Le *Hi-Hat fermé* se caractérise par de l'énergie allant de 0kHz à 10kHz dont l'amplitude est légèrement plus importante dans les hautes fréquences. L'énergie dans les hautes fréquences n'est pas surprenante pour un son dental ou alvéolaire car la hauteur du Centre de Gravité spectral dépend de la longueur de la cavité antérieure. La cavité antérieure de  $[\widehat{ts}]$  est petite ce qui explique que l'énergie prédomine dans les hautes fréquences.

### Le Hi-Hat Humming

La Figure 3.29 illustre les caractéristiques du charleston lingual  $\{\widehat{ts}\}$ . Les oscillations présentes sur le spectrogramme sont dues au *humming*. Le spectre a été extrait sur la totalité du son soit une fenêtre de 30ms. Nous faisons trois observations. La première concerne le niveau global d'amplitude du spectre plus faible que la variante laryngale. Il s'agit de la même différence entre  $[\widehat{ts}]$  et  $\{\widehat{ts}\}$  qu'entre  $[p]$  et  $\{p\}$ . Le pic à 8kHz suggère une cavité antérieure de petite taille.



(a) Signal acoustique (haut) et spectrogramme (bas) (b) Spectre FFT (noir) et LPC (rouge) de la friction

FIGURE 3.29 – Timbre du *Hi-Hat Humming*  $\{\widehat{ts}\}$  - Sujet AI



### Timbre des charleston

Regardons à présent la comparaison des spectres LPC de la Figure 3.30. On note que la répartition de l'énergie est la même entre les deux variantes et que seul le niveau d'amplitude globale du spectre diffère. En effet, dans le cas de la variante linguale, le niveau d'amplitude du spectre est bien plus faible. Les charlestons beatboxés se caractérisent donc par de l'énergie dans les hautes fréquences.

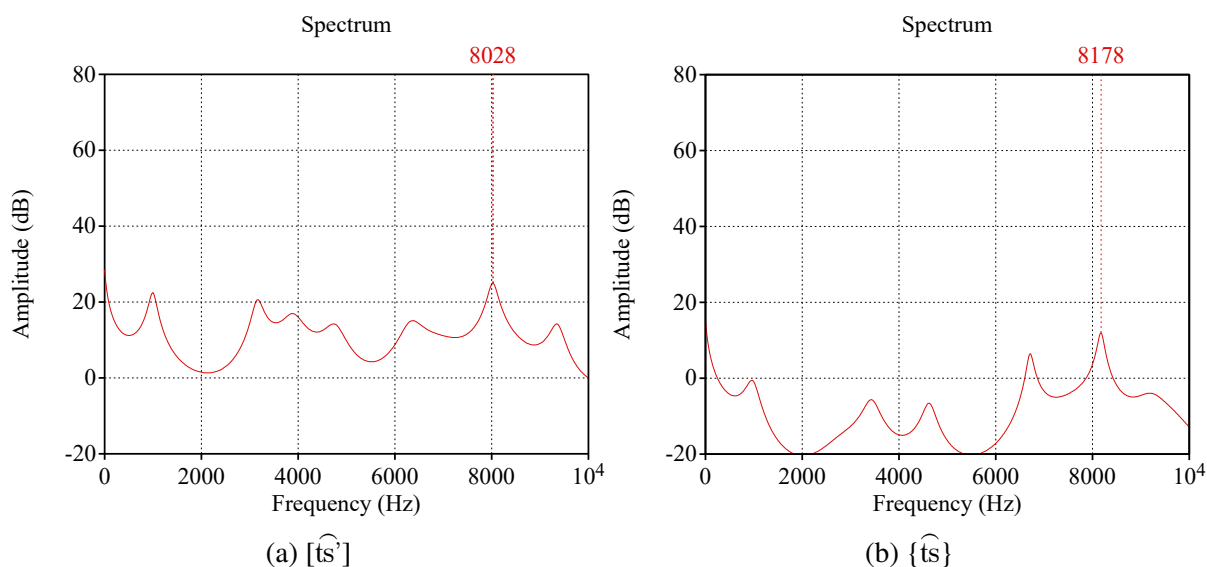


FIGURE 3.30 – Comparaison du timbre des charlestons du sujet AI à partir de spectre LPC à 24 coefficients

### 3.3.3 Caisses claires

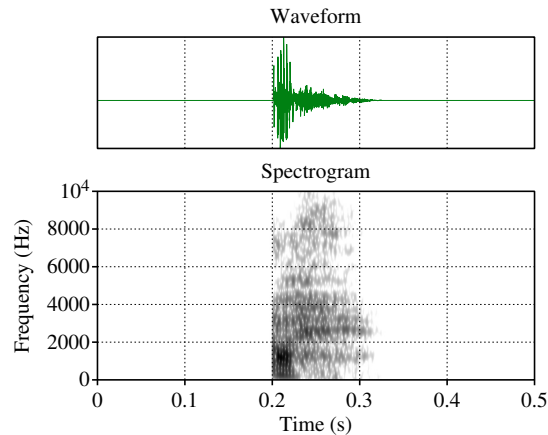
Nous allons présenter les caisses claires du sujet VP. Concernant les caisses claires, les dimensions de la membrane de la caisse claire sont plus petites que celles de la grosse caisse. De plus, en dessous de la peau, à l'intérieur de la caisse, on trouve des bandes de fils métalliques qui, si utilisées, peuvent modifier le timbre. Enfin, il existe aussi différentes façons de frapper la peau ou le bord de la caisse. On peut aussi utiliser différents types de baguettes pouvant aussi modifier le timbre résultant. Ces différents aspects expliquent la diversité des timbres de caisses claires beatboxées. Notons que les caisses claires représentent à elles seules 44% des occurrences dans le corpus des *beat patterns*.

#### 3.3.3.1 Caisses claires orales

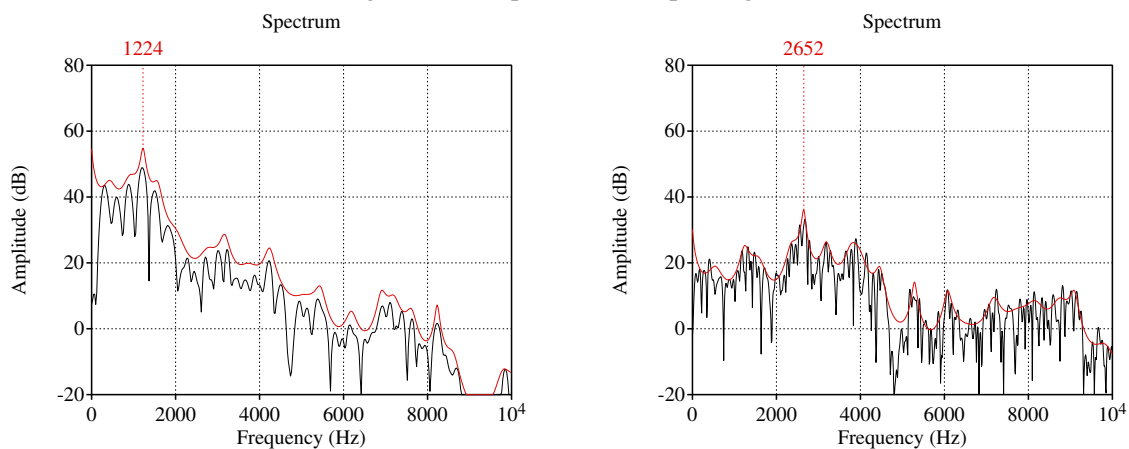
##### La K-Snare Inward

La Figure 3.31 illustre les caractéristiques spectrales de la *K-Snare Inward* [ $\downarrow\text{KL}$ ] du sujet VP. La fricative est une prolongation du relâchement vélaire. Nous l'avons vu plus haut, le relâchement est latéralisé et les molaires fonctionnent comme un obstacle. La partie occlusive est similaire à ce qu'on s'attend à trouver pour une occlusive vélaire, c'est-à-dire un pic proéminent dans les fréquences moyennes entre 1kHz et 4kHz. Le spectre de la *K-Snare Inward* possède une prééminence plutôt compacte aux environs d'1.2 kHz. On note qu'au-dessus de 2kHz l'énergie est diffuse et l'amplitude diminue progressivement. Notons la présence d'une anti-résonance marquée entre 4 et 6kHz. Cette anti-résonance est sans doute introduite par la création d'une cavité latérale.

Le bruit de friction possède un pic moins ample à 2.6kHz. Notons que le pic d'1.2kHz est toujours présent bien que moins ample. Le mouvement du pic de résonance peut s'expliquer par le mouvement de la mâchoire (cf. Figure 3.16). Un déplacement latéral de la mâchoire inférieure implique *de facto* un déplacement de la langue. Les fricatives vélares égressives des langues du monde se caractérisent généralement par de l'énergie dans les basses et moyennes fréquences  $\pm 2\text{kHz}$  (Gordon et al., 2002). Ici, le pic spectral est dans les moyennes fréquences car le flux est ingressif. Supposons que la constriction se trouve à 6cm de la glotte, nous avons



(a) Signal acoustique (haut) et spectrogramme (bas)



(b) Spectre FFT (noir) et LPC (rouge) du relâchement  
(c) Spectre FFT (noir) et LPC (rouge) du bruit de friction

FIGURE 3.31 – Timbre de la *K-Snare* inward [ $\downarrow_{\text{KL}}$ ] - Sujet VP

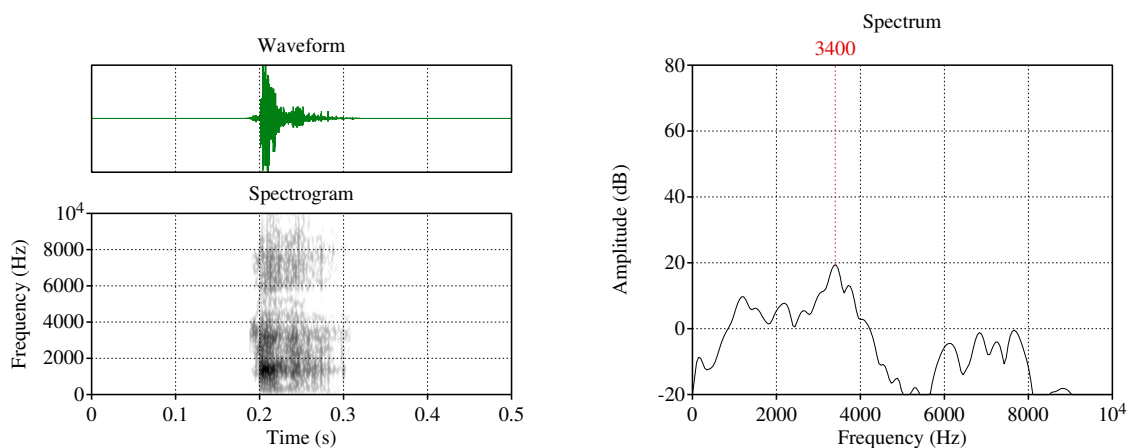


une grande cavité de 10cm entre les lèvres et la constriction et une plus petite cavité de 6cm à l'arrière du conduit entre la glotte et la constriction. La cavité en avant de la constriction est la plus petite ce qui explique pourquoi l'énergie se concentre dans les fréquences moyennes et non dans les fréquences moyennes-basses. On note toujours l'antirésonance majeure à 4.8kHz que l'on avait déjà sur le relâchement. La mâchoire et la langue étant en mouvement, la taille de la cavité latérale en est affectée.

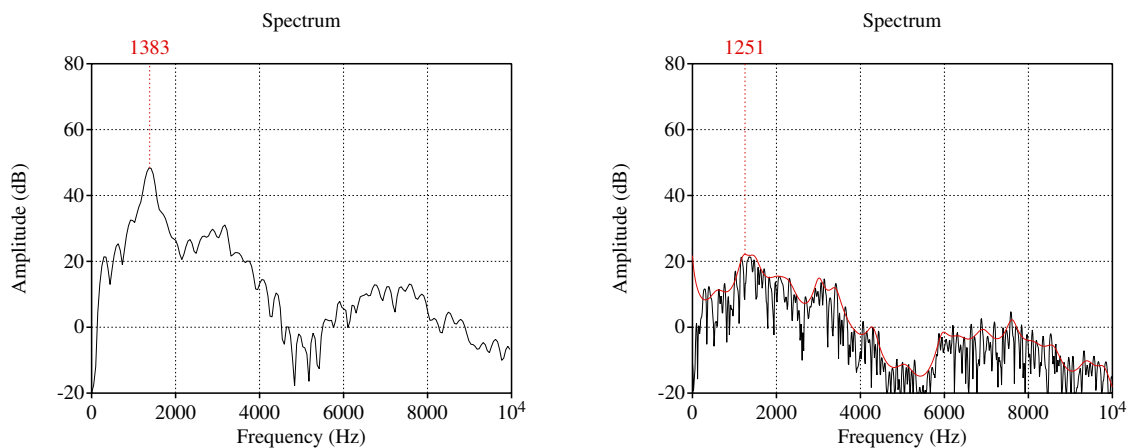
### La *K-Snare* Inward humming

La Figure 3.32 illustre les caractéristiques spectrales de la variante humming  $\{\downarrow_{\text{KL}}\}$  de la *K-Snare* ingressive. Sur le signal acoustique et le spectrogramme on remarque un burst plutôt intense et un bruit court avant le burst et un bruit long après le burst. Il y a aussi un phénomène de latéralité dans le relâchement. En effet, selon les discussions avec le sujet VP, il semblerait que le relâchement ne se fasse pas de façon uniforme sur toute la longueur du dos de la langue. Le relâchement est suivi d'un bruit de friction plus long que le relâchement. Nous avons extrait les spectres FFT du bruit de début (fenêtre de 10ms) et du burst (fenêtre de 18ms) ainsi que le

spectre FFT et LPC à 24 coefficients du bruit de friction (fenêtre de 65ms). Étant donné la taille réduite des fenêtres d'analyse des deux parties du relâchement, il n'est pas nécessaire de générer de lissage. Les 3 structures sont en fait les mêmes du point de vue de la répartition fréquentielle. Tous les spectres présentent un pic à 1.3kHz, 3.4kHz, de l'énergie entre 6kHz et 8kHz et une anti-résonance importante à 5kHz environ. Ce qui semble varier, c'est l'amplitude des deux premières résonances. Dans le bruit du début, les basses fréquences sont fortement atténuées et la deuxième résonance est du même niveau d'amplitude que le bruit de friction mais moins ample d'environ 15dB par rapport au burst. La différence d'amplitude est possiblement due aux molaires agissant comme un obstacle lors de la seconde phase du relâchement. L'enveloppe globale de ces trois spectres est similaire aux caractéristiques des occlusives vélares. Le double relâchement présenté ici n'est pas présent de façon systématique chez le sujet VP. On l'observe très rarement chez les autres sujets dont la signature acoustique burst + friction suggère un relâchement uniforme du dos de la langue.



(a) Signal acoustique (haut) et spectrogramme (bas) (b) Spectre FFT (noir) et LPC (rouge) du bruit de début



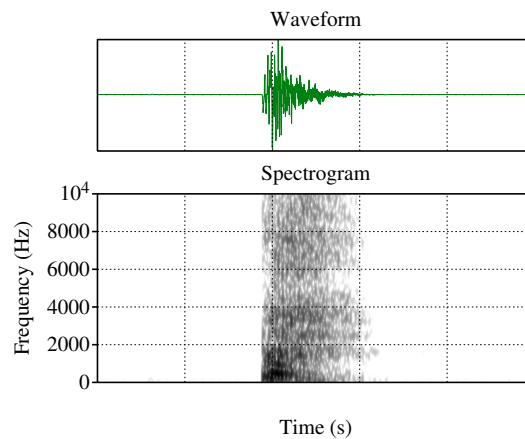
(c) Spectre FFT (noir) et LPC (rouge) du bruit de début (d) Spectre FFT (noir) et LPC 24 coefficients (rouge) du bruit de friction

FIGURE 3.32 – Timbre de la *K-Snare Humming*  $\{\downarrow_{kL}\}$  - Sujet VP

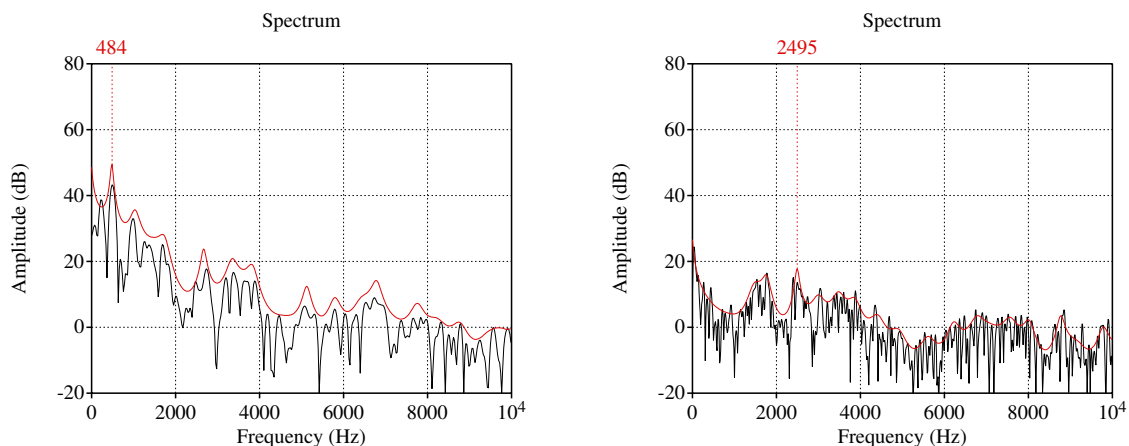


## La PF Snare classique

La Figure 3.33 illustre les caractéristiques de la *PF snare* [pf:]. Il s'agit d'une affriquée. On peut distinguer deux phases sur le signal et le spectrogramme : un relâchement suivi de friction. Le relâchement se traduit parfois par la présence d'une trille plutôt que par de la friction comme c'est le cas dans les occlusives des langues du monde. Ce n'est pas surprenant étant donné que la trille est un mode d'articulation intermédiaire entre les occlusives et les fricatives. C'est-à-dire une constriction un peu plus fermée qu'une fricative mais plus ouverte que l'occlusive afin de générer une oscillation. Le spectre du relâchement labial (b) présente une enveloppe spectrale similaire aux grosses caisses décrites plus haut (cf. Figure 3.27) à la différence qu'ici la pente spectrale est moins raide. Contrairement à la grosse caisse [p'], on observe une pente moins abrupte entre 0 et 4kHz. Concernant la fricative, les caractéristiques spectrales du bruit sont représentées en (c). L'enveloppe spectrale est très similaire à la consonne fricative [f]. En effet, le bruit se caractérise par un spectre plat (ni ascendant, ni descendant) et diffus. L'énergie est cependant plus ample entre 1.6kHz et 4kHz.



(a) Signal acoustique (haut) et spectrogramme (bas)



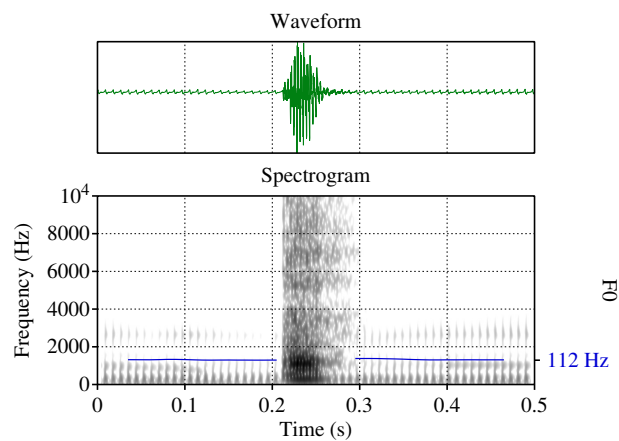
(b) Spectre FFT (noir) et LPC 12 coefficients (rouge) du bruit de relâchement (rouge) du bruit de friction

FIGURE 3.33 – Timbre de la *PF Snare* [pf:] - Sujet VP

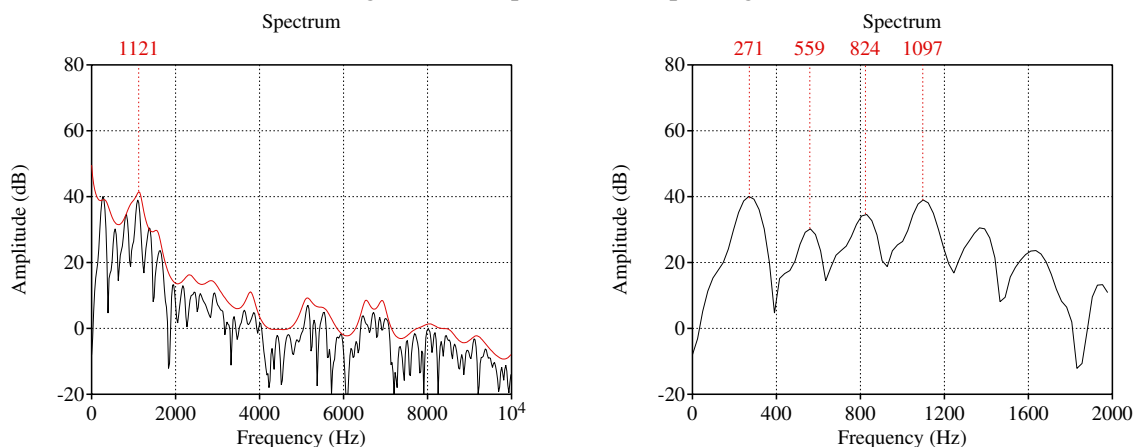


## La PF Snare humming

La Figure 3.34 illustre les caractéristiques de la *PF Snare humming*  $\{\widehat{pf}\}$ . Nous la traitons comme une affriquée dont le relâchement est produit comme une trille suivie d'une fricative. On peut aussi observer une phase d'oscillation et de friction. Toutefois, chez les autres sujets on peut observer du bruit à la place de l'oscillation. Les spectres présentés en (b) et (c) sont les mêmes; le spectre (c) illustre seulement la répartition de l'énergie en-dessous de 2kHz. Ils ont été générés à partir de la phase d'oscillation sur le signal. La friction suivant l'oscillation n'est pas représentée ici. Notons que sa signature sur le spectrogramme est similaire à la phase d'oscillation. L'oscillation labiale se caractérise par une énergie plus importante dans les basses fréquences avec un pic à 1kHz et par une enveloppe descendante et diffuse. Entre 0 et 2kHz on remarque une structure harmonique irrégulière avec des pics qui sont presque des multiples entiers de 271Hz. Cette structure harmonique n'est pas liée au voisement observé sur le spectrogramme. En effet, ni le pic à 270Hz ni les autres pics ne sont des multiples entiers de 112Hz.



(a) Signal acoustique (haut) et spectrogramme (bas)



(b) Spectre FFT (noir) et LPC 12 coefficients (rouge) du bruit de friction

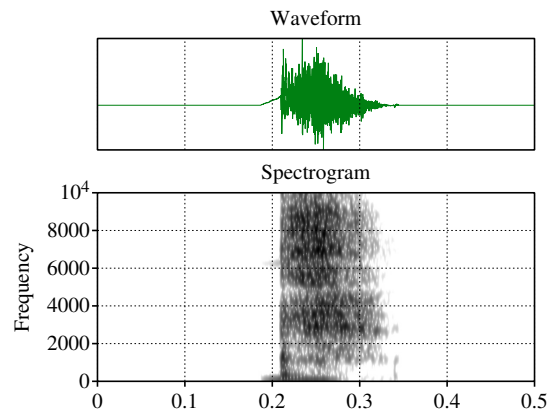
(c) Spectre FFT à 2kHz

FIGURE 3.34 – Timbre de la *PF Snare humming*  $\{\widehat{pf}\}$  - Sujet VP

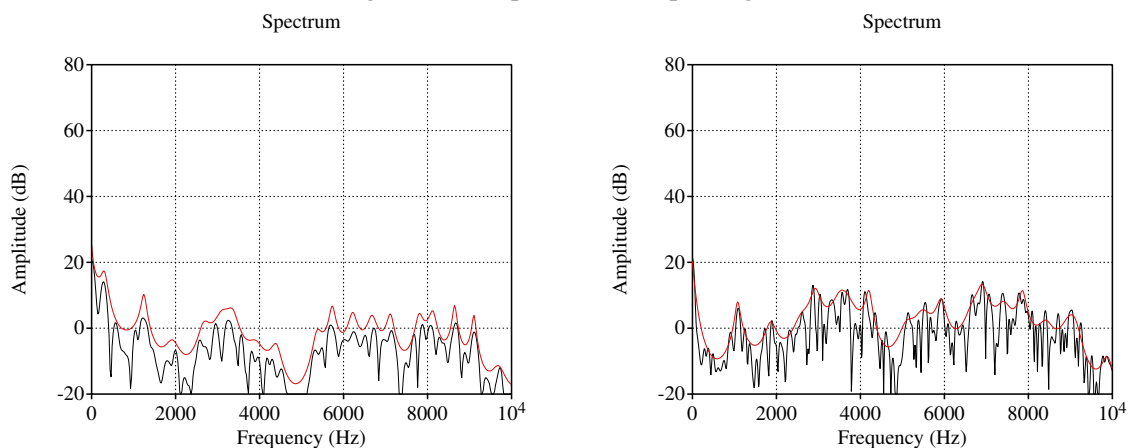


## La Tish Snare

Enfin, la *Tish Snare* [tʃ̥:] est illustrée sur la Figure 3.35. On note une partie occlusive avec un relâchement et une partie fricative. La répartition de l'énergie est similaire entre les deux phases. On note de l'énergie à la fois dans les moyennes fréquences entre 2-4kHz et dans les hautes fréquences au-dessus de 5-6kHz. L'enveloppe spectrale est plate et la différence d'amplitude entre l'énergie dans les moyennes et hautes fréquences est de niveau comparable.



(a) Signal acoustique (haut) et spectrogramme (bas)



(b) Spectre FFT (noir) et LPC 24 coefficients (rouge) du bruit du relâchement (c) Spectre FFT (noir) et LPC 24 coefficients (rouge) du bruit de friction

FIGURE 3.35 – Timbre de la *Tish Snare* [tʃ̥:] - Sujet VP



## Timbre des caisses claires orales

Concernant le relâchement, la Figure 3.36 illustre les résonances de 4 des 5 caisses claires précédemment analysées. Nous avons volontairement laissé de côté la *K-Snare Humming* étant donné que nous n'avons pas effectué de lissage sur le spectre FFT. À l'exception de la *Tish Snare*, l'aspect global des spectres est similaire. Il s'agit de spectres descendants avec un pic en dessous de 2kHz. On observe un pic à environ 1kHz au relâchement de la vélaire [ɫ̥L:] et de la labiale [p̥f:]. On observe un pic bien plus bas, à 480Hz, concernant [p̥f:]. Nous avons

souligné que la *K-Snare Humming* possédait aussi un pic à 1kHz. Contrairement à l'attaque des grosses caisses, la diminution de l'amplitude est graduelle et moins abrupte et de ce fait les moyennes et hautes fréquences sont moins atténuées. La pente spectrale pourrait donc avoir un rôle dans l'identification des caisses claires orales. Il n'est pas clair ici si un contraste acoustique s'établit sur l'attaque entre les différentes caisses claires pour les différencier entre elles. Des tests de perception où l'on manipule les transitoires d'attaque et d'extinction permettraient de formuler des hypothèses plus approfondies quant au poids de ce type d'indices acoustiques sur la perception et la catégorisation acoustique des différents types de caisse claire.

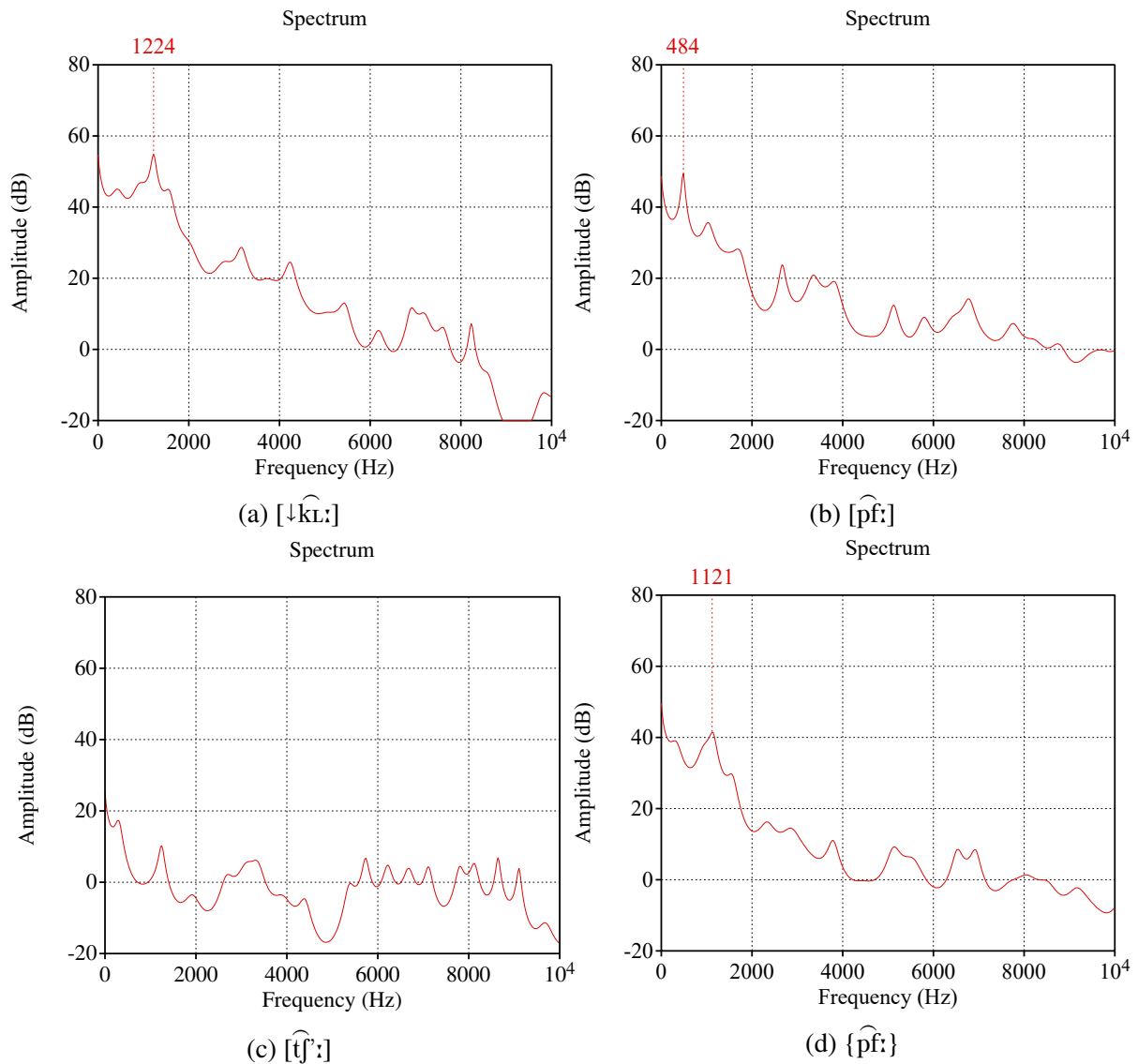


FIGURE 3.36 – Comparaison des timbres de trois caisses claires orales au moment du relâchement

Concernant la phase de friction, 4 des 5 caisses claires sont illustrées dans la Figure 3.37). La *K-Snare Inward* possède le niveau d'amplitude global le plus élevé avec des pics très marqués. Le spectre de la *PF Snare* est quant à lui plus plat et possède seulement deux pics aux alentours de 2kHz. Les autres pics sont très affaiblis ce qui met en relief les deux pics à 2kHz. Le spectre de la *K-Snare Humming* est plus affaibli que sa correspondante pulmonique mais possède tout de même une structure assez proche de cette dernière. On note sur les deux spectres ont plusieurs résonances de même fréquence mais dont l'amplitude est réduite dans le cas de la variante de humming. Enfin, la *Tish Snare* se distingue des autres caisses claires par la présence d'énergie de même amplitude dans les moyennes et hautes fréquences. Pour toutes ces caisses claires nous observons que l'énergie se trouve dans les moyennes fréquences concernant le transitoire d'extinction.

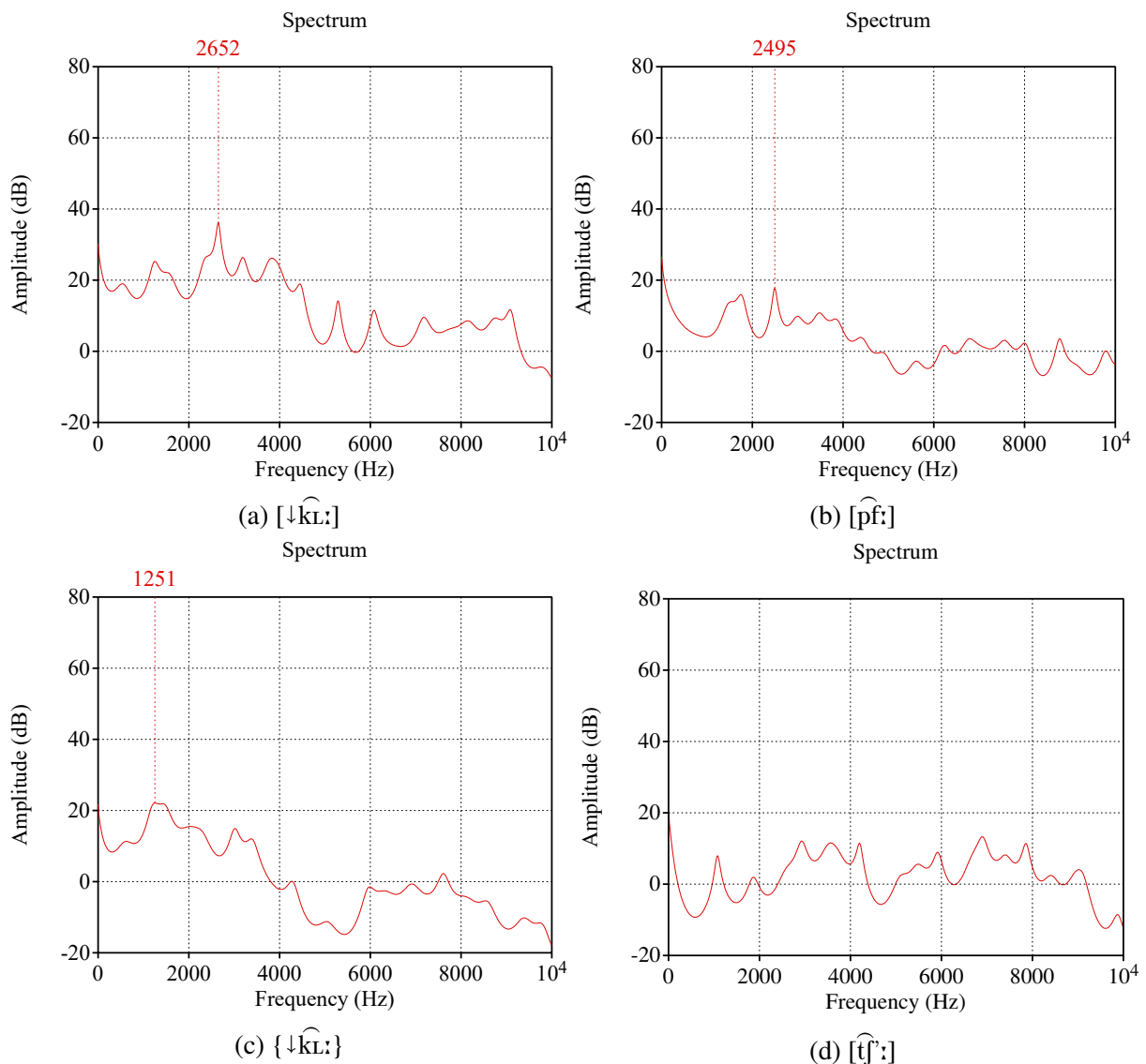
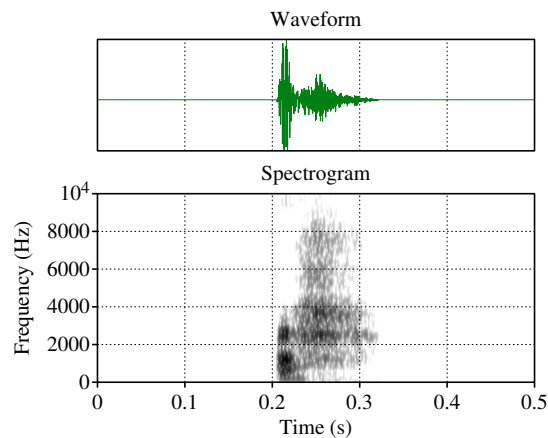


FIGURE 3.37 – Comparaison des timbres au moment de la friction

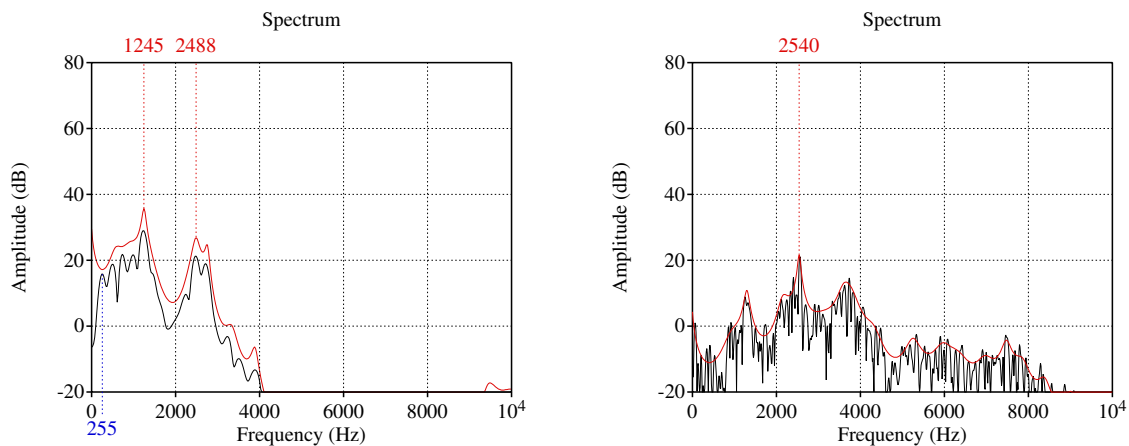
### 3.3.3.2 Caisses claires oro-laryngales

#### La ESH snare

La Figure 3.38 illustre les caractéristiques de la *ESH snare* égressive [ʔ̥]. Concernant le relâchement, on distingue des résonances à 1.2kHz et 2.5kHz sur le spectrogramme. Si l'on regarde le spectre attentivement, on retrouve ces résonances et on note au passage un fondamental à 255Hz et des multiples entiers de celui-ci, ce qui confirme bel et bien que le relâchement est voisé. L'enveloppe spectrale indique un spectre compact dont toute l'énergie se trouve entre 0 et 4kHz. La friction se caractérise par un spectre plat et diffus entre 1kHz et 8kHz. Trois résonances à 1.2kHz, 2.5kHz (pic spectral) et 3.8kHz peuvent être observées sur les spectres.



(a) Signal acoustique (haut) et spectrogramme (bas)



(b) Spectre FFT (noir) et LPC (rouge) du bruit de relâchement (c) Spectre FFT (noir) et LPC (rouge) du bruit de friction

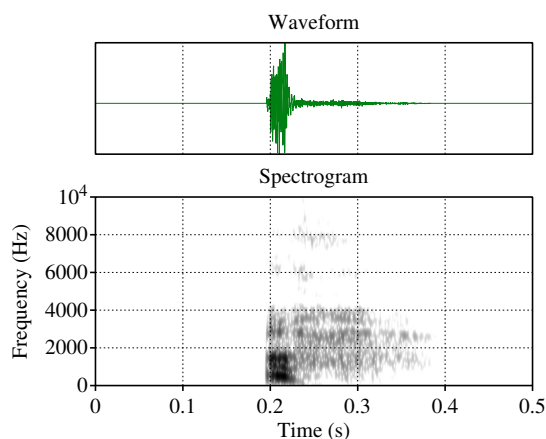
FIGURE 3.38 – Timbre de la *ESH Snare Outward* [ʔ̥] - Sujet VP



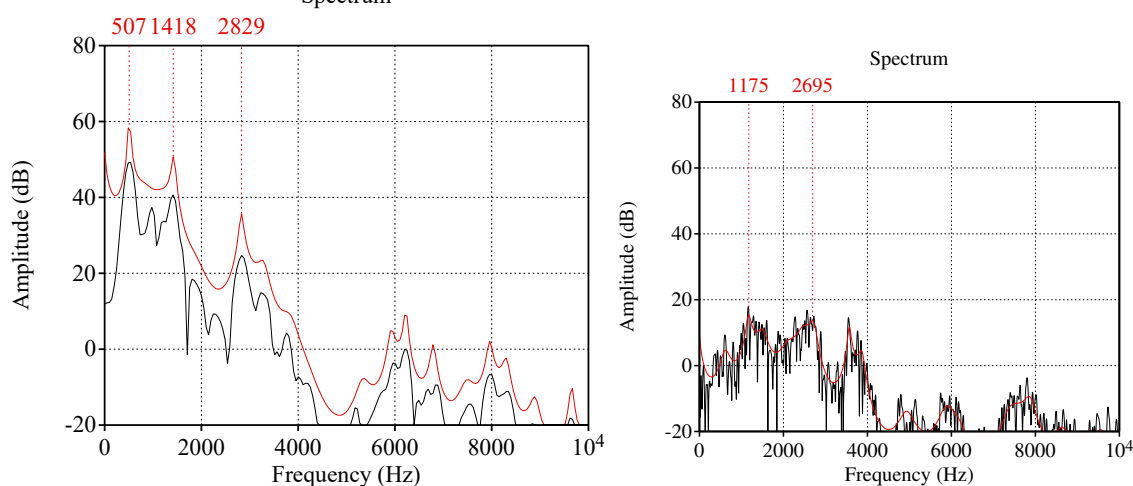
#### La ESH Snare Inward

La Figure 3.39 illustre la correspondante ingressive [ʔ̥̚] de la *ESH snare*. Une fois encore on observe des résonances au moment du relâchement. Les deux premières sont plus intenses

que la troisième. Sur le spectre (b), ces résonances se trouvent à 500Hz, 1.4kHz et 2.8kHz. Notons que la première résonance correspond à la F0. Le spectre du relâchement est compact et descendant de 0kHz à 4kHz. Concernant la friction, le spectre est aussi compact entre 0kHz et 4kHz; cependant il est plutôt plat. Contrairement à sa correspondante égressive, la variante ingressive possède (à quelques Hertz près) des résonances moins intenses que la variante égressive avec un pic légèrement plus élevé pour cette dernière (2.5kHz *versus* 2.7kHz).



(a) Signal acoustique (haut) et spectrogramme (bas)



(b) Spectre FFT (noir) et LPC (rouge) du bruit de relâchement (c) Spectre FFT (noir) et LPC (rouge) du bruit de friction

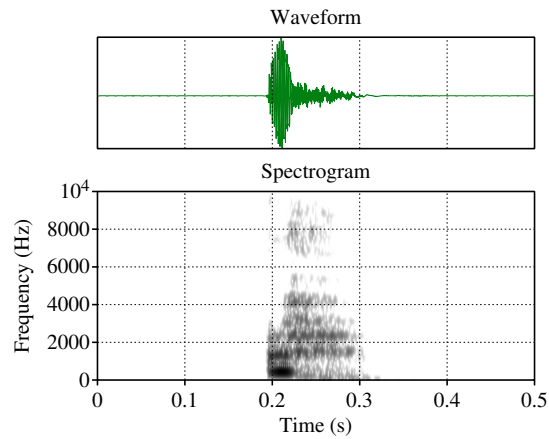
FIGURE 3.39 – Timbre de la *ESH Snare Inward* [ $\downarrow \hat{z}$ ]: - Sujet VP



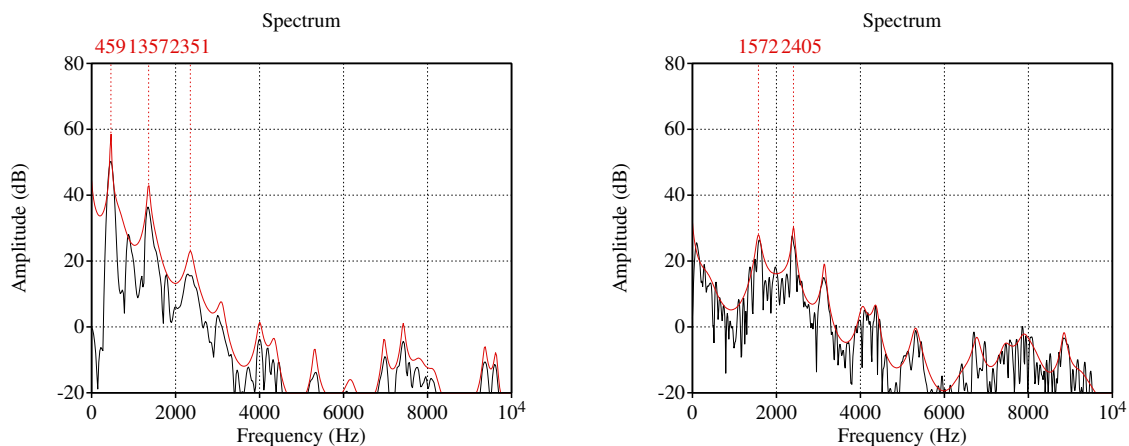
### La Cough Snare

La Figure 3.40 illustre les caractéristiques de la *Cough Snare* [ $\hat{z}h$ ]. Le relâchement comporte deux résonances intenses à 459Hz et 1.3kHz et une troisième résonance moins intense à 2.3 kHz. Notons une fois encore que la première résonance correspond à la F0. Le spectre est compact et descendant et ne comprend pas d'énergie au-dessus de 3.5kHz. La fricative glottale se caractérise par un spectre compact et plat entre 1.5kHz et 4kHz. Hormis la faible énergie aux

alentours de 8kHz, nous ne notons pas d'énergie au-dessus de 4kHz. Les résonances du relâchement et de la friction sont similaires, ce qui n'est pas surprenant étant donné que la source est générée au niveau glottique et que la configuration du conduit vocal sert de filtre. On peut en déduire que la configuration reste la même pour l'occlusive et la fricative. C'est d'ailleurs ce que montraient les clichés IRM de la Section 3.2 précédente (cf. Figure 3.18).



(a) Signal acoustique (haut) et spectrogramme (bas)



(b) Spectre FFT (noir) et LPC (rouge) du bruit de relâchement (c) Spectre FFT (noir) et LPC (rouge) du bruit de friction

FIGURE 3.40 – Timbre de la *Cough Snare* [ʔh:] - Sujet VP



### 3.3.3.3 Timbre des caisses claires oro-laryngales

Sur la Figure 3.41, chaque ligne représente une des trois caisses claires laryngales ; la colonne de gauche illustre le relâchement voisé et la colonne de droite la friction. Les résonances observées au relâchement se trouvent toutes en dessous de 4kHz. Elles semblent correspondre à la plage de 1kHz à 4Hz des caisses claires orales. L'énergie est fortement diminuée voire supprimée au-dessus de 4kHz. La phase de friction est présentée sur les figures de droite. Les spectres des *ESH snare* diffèrent du spectre de la *Cough Snare* étant donné que les lieux d'articulation sont différents. Notons encore une fois la répartition de l'énergie entre 1kHz et 4kHz. De plus,

il semblerait que la direction du flux de la *ESH Snare Inward* ait des effets assez nets du point de vue acoustique. Premièrement le pic de la variante ingressive est plus basse en comparaison avec sa correspondante égressive. Il est possible que cela soit dû à la présence des dents à la sortie de la constriction étant donné que les différences de direction du flux. Deuxièmement, les fréquences supérieures à 4kHz sont fortement atténuées dans le cas de la variante ingressive. Il est possible qu'un couplage avec la trachée ait lieu ce qui modifierait les caractéristiques du spectre de sortie. Ces deux effets rendent  $[\downarrow\hat{t}]:$  plus compacte et plus plat entre 1kHz et 4kHz.

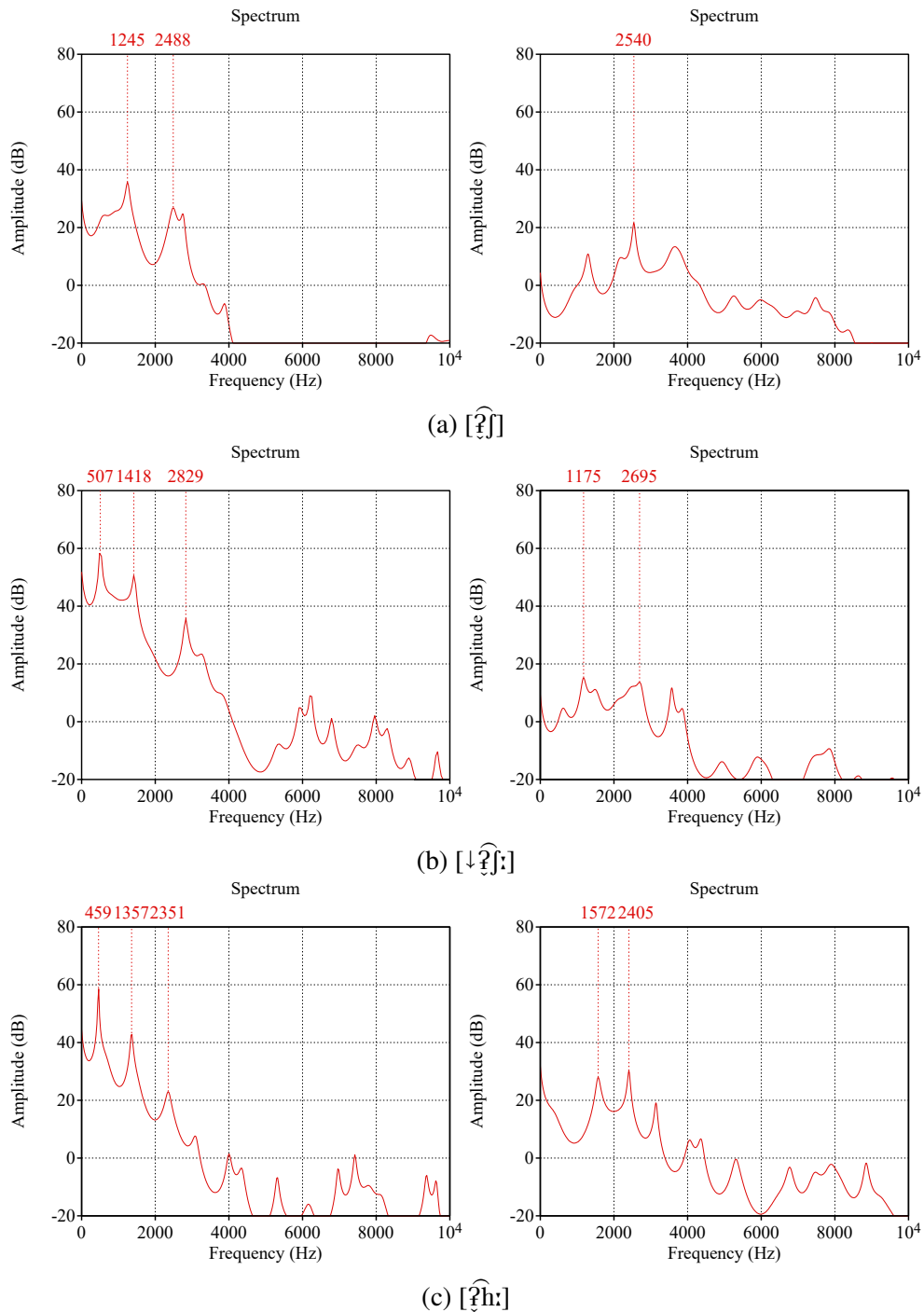


FIGURE 3.41 – Comparaison des timbres au moment du relâchement (gauche) et de la friction (droite) des trois caisses claires laryngées



### 3.3.4 Rolls

Les Rolls ne sont pas des sons percussifs, ils sont plutôt classés par les beatboxeurs parmi les sons électroniques. Ce sont des trilles. Le *Lips Roll* [ $\downarrow_B^1$ ], { $\downarrow_B^1$ } peut occuper la même position que les grosses caisses dans les *beat patterns* et le *Clic Roll* [ $\downarrow_C$ ] peut occuper la même position que les charlestons. Pour illustrer ces deux sons nous avons décidé de tirer des exemples du sujet CJ.

#### Lips Roll pulmonique

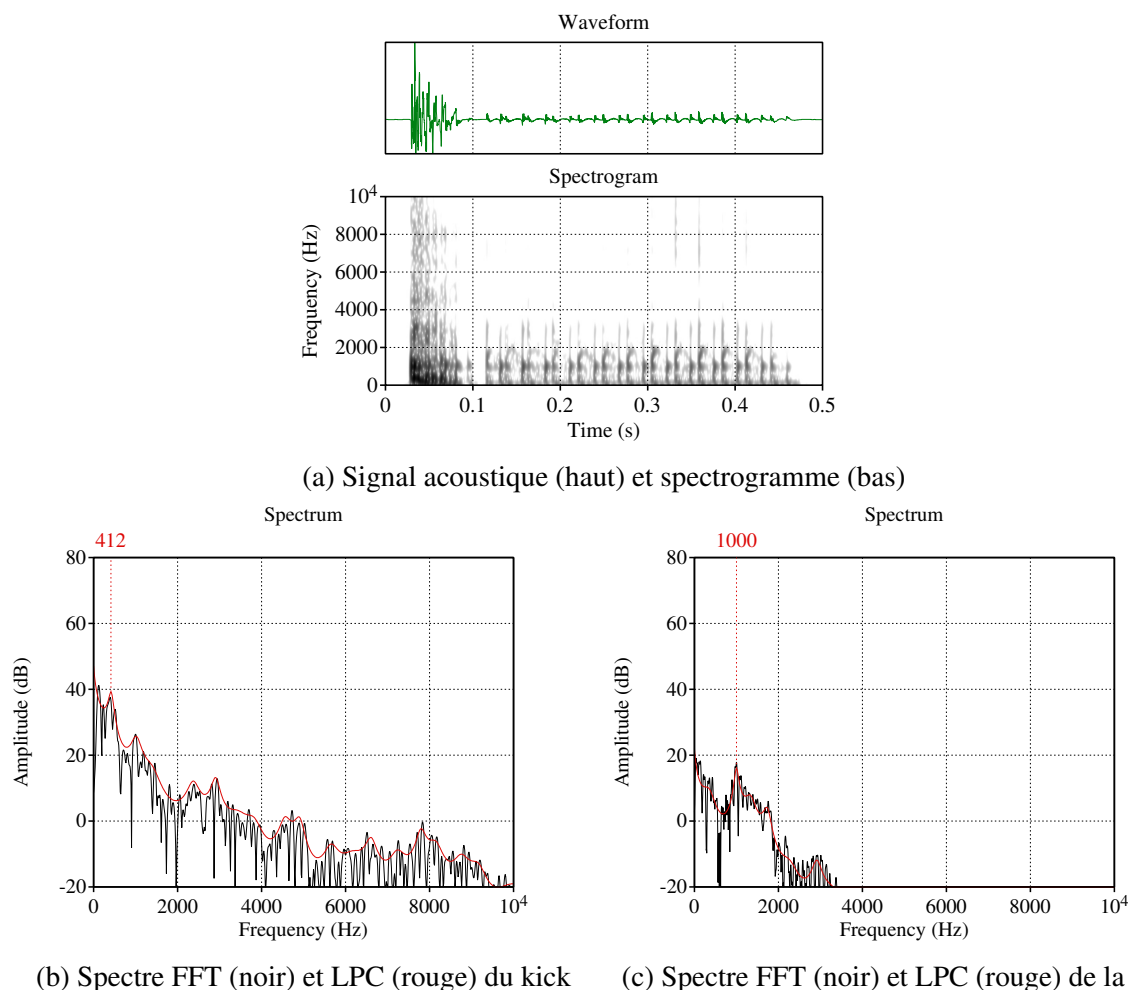


FIGURE 3.42 – Timbre du *Lips Roll* pulmonique [ $\downarrow_B^1$ ] - Sujet CJ

⊙

La Figure 3.42 illustre les caractéristiques du *Lips Roll* [ $\downarrow_B^1$ ] produit par le sujet CJ. Notons qu’il s’agit du même exemple donné plus haut lorsque nous proposons la notion de co-initiation. La trille est donc précédée d’un *kick* éjectif. L’occlusive possède les mêmes caractéristiques que les *kicks* analysés au début de cette sous-partie. Les caractéristiques de la trille sont intéressantes. Notons que la trille dure 300ms environ et que nous avons extrait le spectre sur une fenêtre de 100ms au milieu de la trille labiale. Les lèvres oscillent à 35Hz et on note une résonance à 1kHz. Dans des communications passées (par ex. Dehais-Underdown et al., 2021),

nous avons émis l'hypothèse d'une oscillation de la muqueuse labiale à 1kHz. Nous revenons sur cette hypothèse et la contredisons avec les arguments suivants.

Tous les beatboxeurs ne possèdent pas cette résonance mais certains arrivent à la produire et même à la modifier en changeant la configuration linguale. C'est le cas du sujet VP dans la Figure 3.43. Les résonances observées à 1.5 et 4kHz ne varient pas beaucoup. En revanche on remarque l'abaissement du pic de 1kHz à 500Hz concernant la variante en [o] ce qui a pour effet de donner un timbre plus « grave ». S'il s'agissait d'une composante de type périodique (oscillation de la muqueuse labiale), nous n'aurions pas observé de changements liés à la configuration linguale.

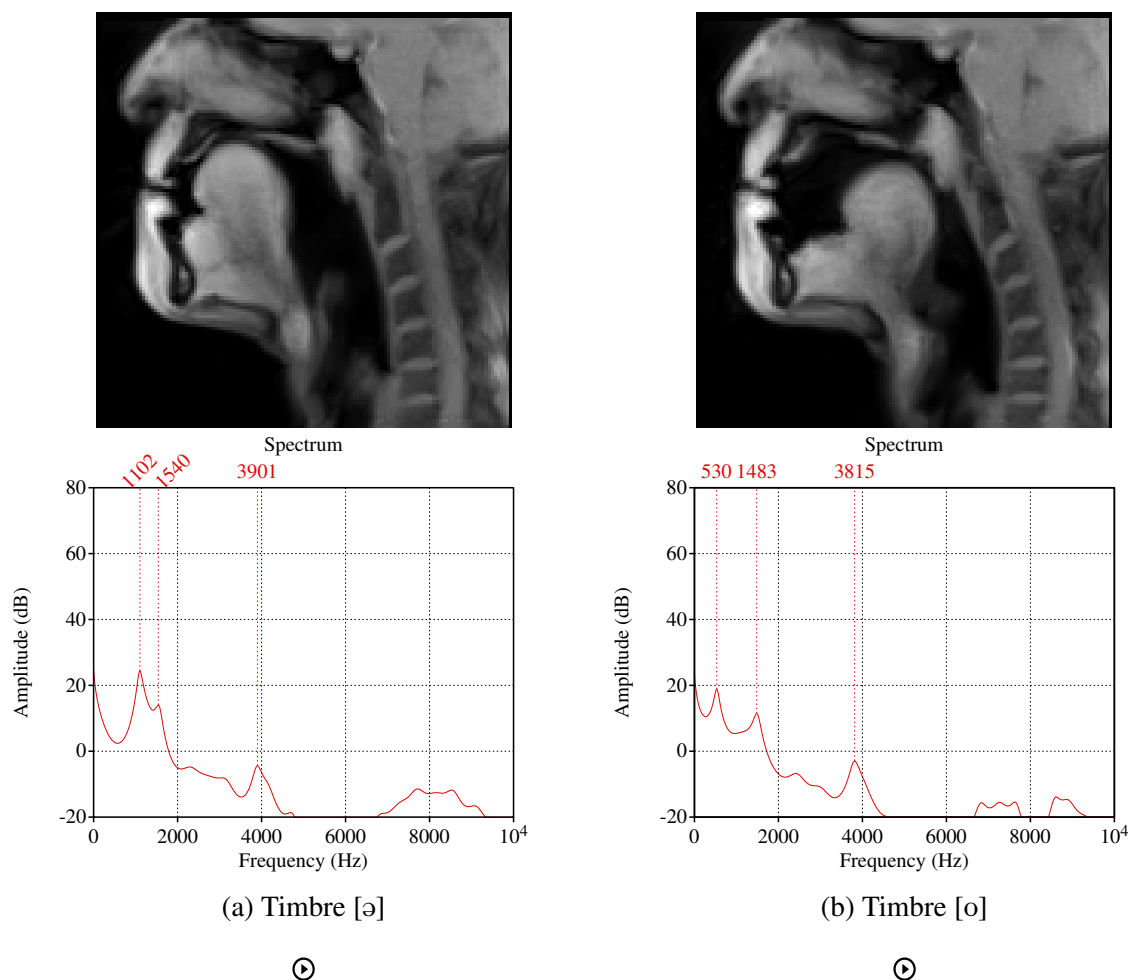


FIGURE 3.43 – Comparaison d'un jeu de timbre entre deux *Lips Roll* du sujet VP

### Lips Roll Humming

Le *Lips Roll* possède une variante humming  $\{\downarrow_{\text{B}}^{\text{!}}\}$ . La Figure 3.44 illustre les caractéristiques de la variante linguale. Les lèvres oscillent à 29.5Hz. Nous sommes face à un spectre très compact avec un pic à 690 Hz. Celui-ci est plus bas et plus ample que le pic de la variante pulmonique. On note sur le spectrogramme un mouvement de la résonance. La résonance s'abaisse car le dos de la langue recule pour permettre l'expansion de la cavité buccale et générer un flux ingressif. Le recul résulte en un agrandissement de la cavité buccale et implique donc une diminution de la résonance.

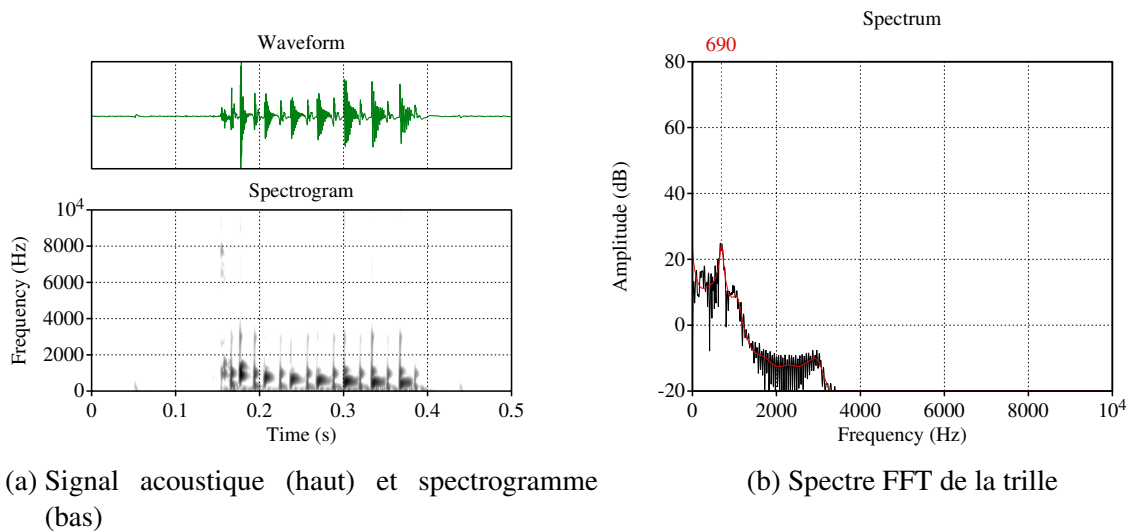


FIGURE 3.44 – Timbre du *Lips Roll Humming*  $\{\downarrow_{\text{B}}^1\}$  - Sujet CJ

⊙

### Clic roll

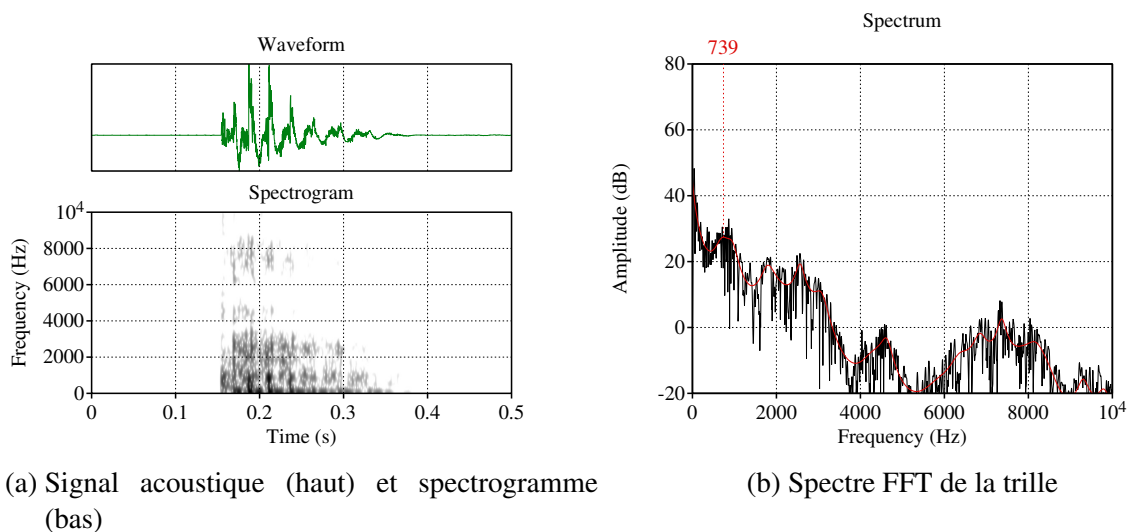


FIGURE 3.45 – Timbre du *Clic Roll*  $[\downarrow_{\text{r}}]$  - Sujet CJ

⊙

Enfin, la Figure 3.45 illustre les caractéristiques du *Clic Roll*  $[\downarrow_{\text{r}}]$ . Selon plusieurs sujets, le *Clic Roll* est une trille où pour produire l'oscillation linguale, les sujets replient la langue sur le palais. Leur description semble indiquer une articulation rétroflexe. Blaylock et al. (2017) décrivent aussi une articulation rétroflexe de ce même son à partir de données IRM. Nous n'avons malheureusement pas de données du sujet VP sur ce son. Le *Clic Roll* produit par le sujet CJ comporte une composante de bruit assez importante. En effet, le spectre indique la présence d'énergie diffuse de 0 à 10kHz. Cependant, il apparaît assez clairement que l'énergie au-dessus de 4kHz est fortement atténuée. Nous avons ici un spectre diffus et descendant entre 0 et 4kHz.

Lors de la segmentation des données aérodynamiques et acoustiques, nous avons noté des différences en termes de composante de bruit et de régularité de la composante périodique. Nous reviendrons sur ces différences dans le prochain chapitre.

### 3.3.5 Des signatures acoustiques discrètes ?

Au début du chapitre, nous nous étions fixés comme objectif de savoir si les sons beatboxés possédaient des signatures acoustiques discrètes, c'est-à-dire des signatures identifiables et délimitables. L'analyse du timbre s'est faite à partir des données conjointes des 4 sujets retenus pour cette étude. Bien qu'un seul sujet ait été analysé pour chacune des catégories de sons pour simplifier la présentation des résultats, les signatures acoustiques des sons du corpus semblent être discrètes. En effet, pour chaque sujet, nous avons trouvé des similitudes concernant le timbre des sons produits. Les signaux, spectrogrammes et sons pour chaque sujet peuvent être consultés en ANNEXE 6. On peut toutefois noter quelques différences individuelles, en particulier en ce qui concerne les caisses claires laryngales.

La *PF Snare* [p̄f:] du sujet GA se distingue auditivement de celle des autres sujets (Figure 3.46). La durée de cette caisse claire est plus longue que chez les autres sujets. De plus, il y a plus d'énergie dans les hautes fréquences. De plus, le débit d'air oral expiré semble bien plus important que chez les autres sujets. On note chez CJ de l'énergie dans les hautes fréquences.

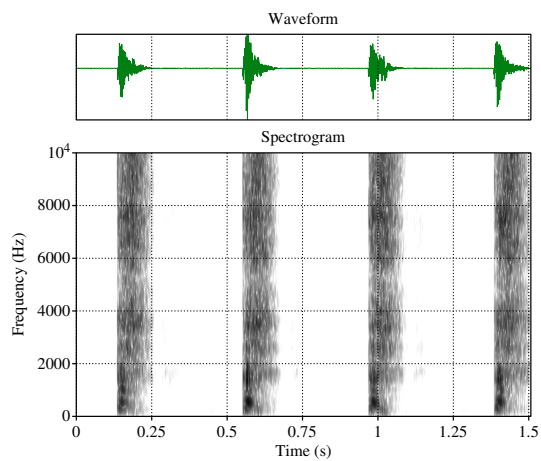
La *Tish Snare* [tʃ̄:] du sujet AI possède moins d'énergie sur la partie fricative, le spectre de la friction est plus aplati et celle-ci est moins audible que le relâchement (Figure 3.47).

La *ESH Snare Inward* [↓ʃ̄:] des sujets AI et CJ est différente de celle des sujets VP et GA (Figure 3.48). Chez le sujet AI, la fricative [↓ʃ] semble auditivement plus postérieure, en l'occurrence elle ressemble à la fricative palatale [ç]. Concernant le sujet CJ, le voisement est peu perceptible par rapport à la fricative [↓ʃ] et le son présente une composante de bruit plus importante que pour les autres sujets.

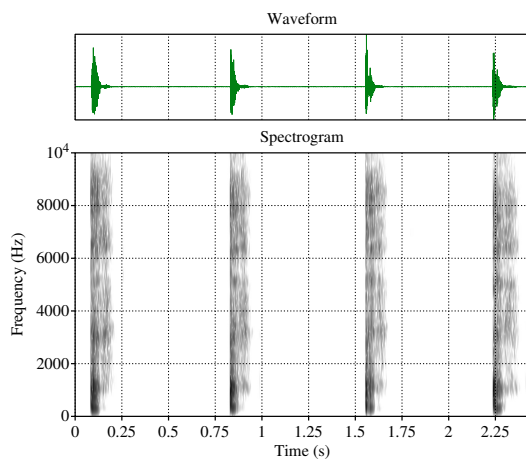
La *Cough Snare* [ʔ̄h:] des sujets CJ et GA semble auditivement plus soufflée que les sujets VP et AI (Figure 3.49). Chez CJ, le relâchement voisé [ʔ̄] se compose de 16ms de voix de type modal et de 23ms de voix de type soufflée. Chez GA, la totalité du relâchement voisé est soufflé.

Enfin, on note des différences concernant les rolls. Les *Lips Rolls* [↓p̄<sup>1</sup>:] des sujets AI et GA ne possèdent pas de résonance à 1kHz ce qui donne un timbre plus grave au son (Figure 3.50). De plus, les trilles sont plus longues que pour les sujets VP et CJ. Les *clics rolls* [↓ɾ:] des sujets CJ et GA ont une composante de bruit plus importante en comparaison avec le sujet AI (Figure 3.51).

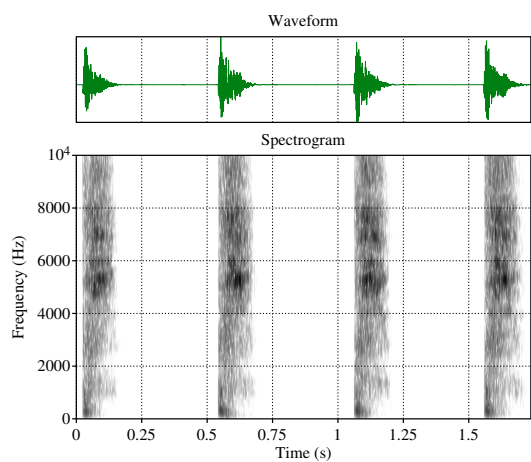
Nous ne savons pas si ces différences constituent des différences du niveau de maîtrise de la pratique du HBB ou des différences de style entre les sujets. Aussi, nous ne savons si ces différences sont suffisamment importantes pour changer le « *percept* » c'est-à-dire la catégorisation faite par les beatboxeurs et les beatboxeuses pour identifier les différentes grosses caisses, caisses claires etc. Ces différences semblent être marginales et comme nous n'avons pas observé de différences majeures pour les autres sons, nous pensons en effet que les sons beatboxés sont des unités discrètes. Des études perceptives permettraient de confirmer le caractère discret des sons beatboxés.



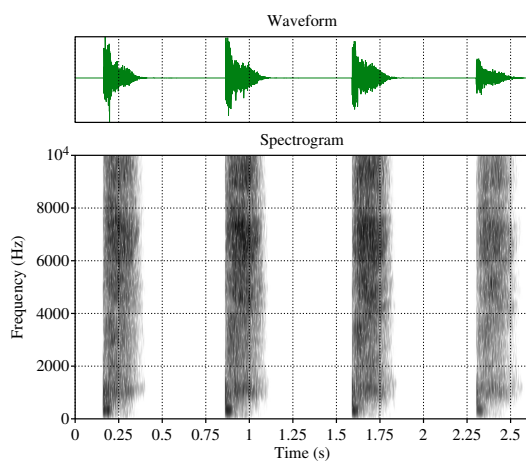
(a) Sujet VP



(b) Sujet AI



(c) Sujet CJ



(d) Sujet GA



FIGURE 3.46 – Comparaison intersujet des signaux acoustiques et des spectrogrammes de la *PF Snare* [pf:]. Attention, échelles de temps différentes.

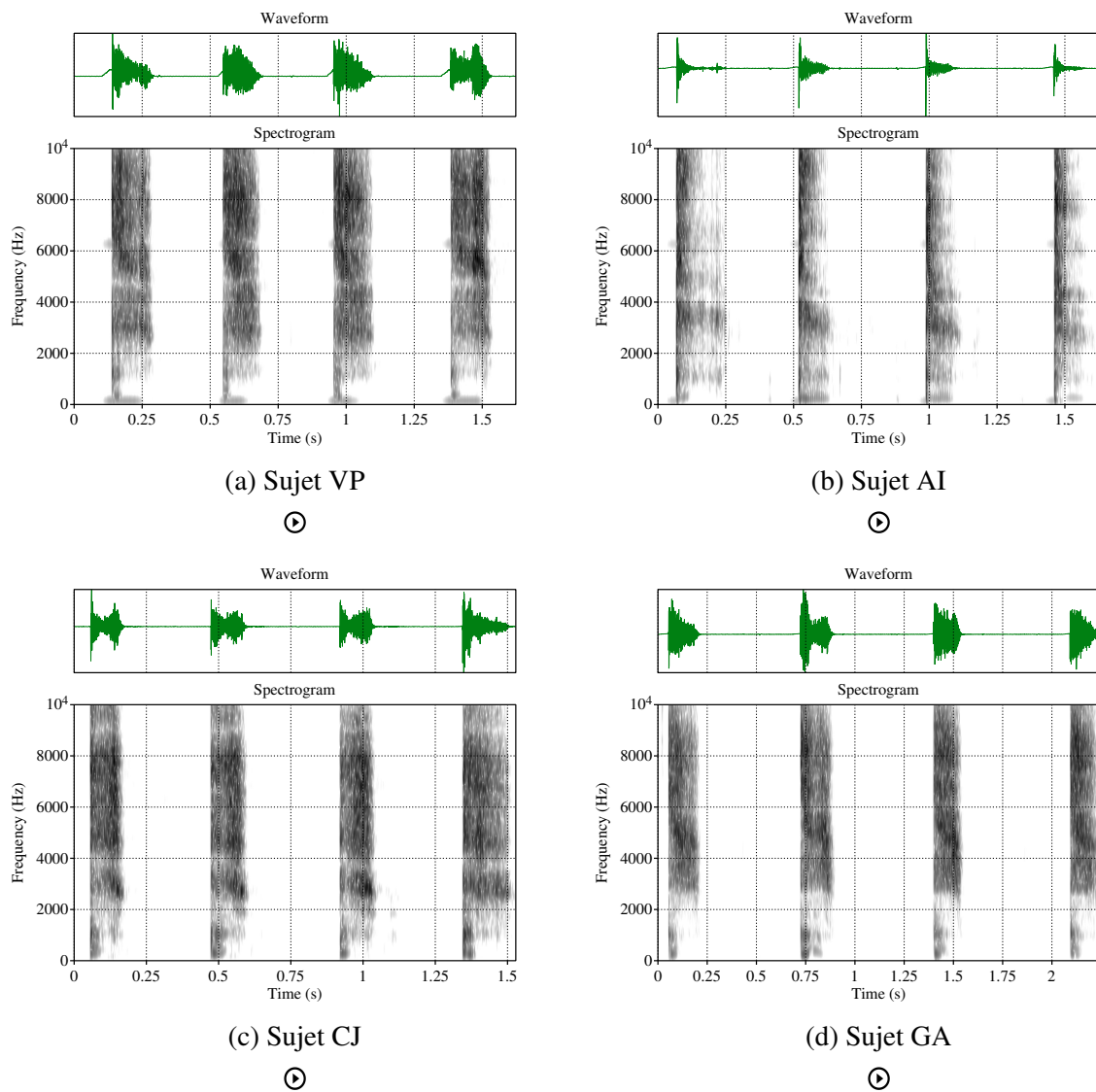


FIGURE 3.47 – Comparaison intersujet des signaux acoustiques et des spectrogrammes de la *Tish Snare* [tʃː]. Attention, échelles de temps différentes.

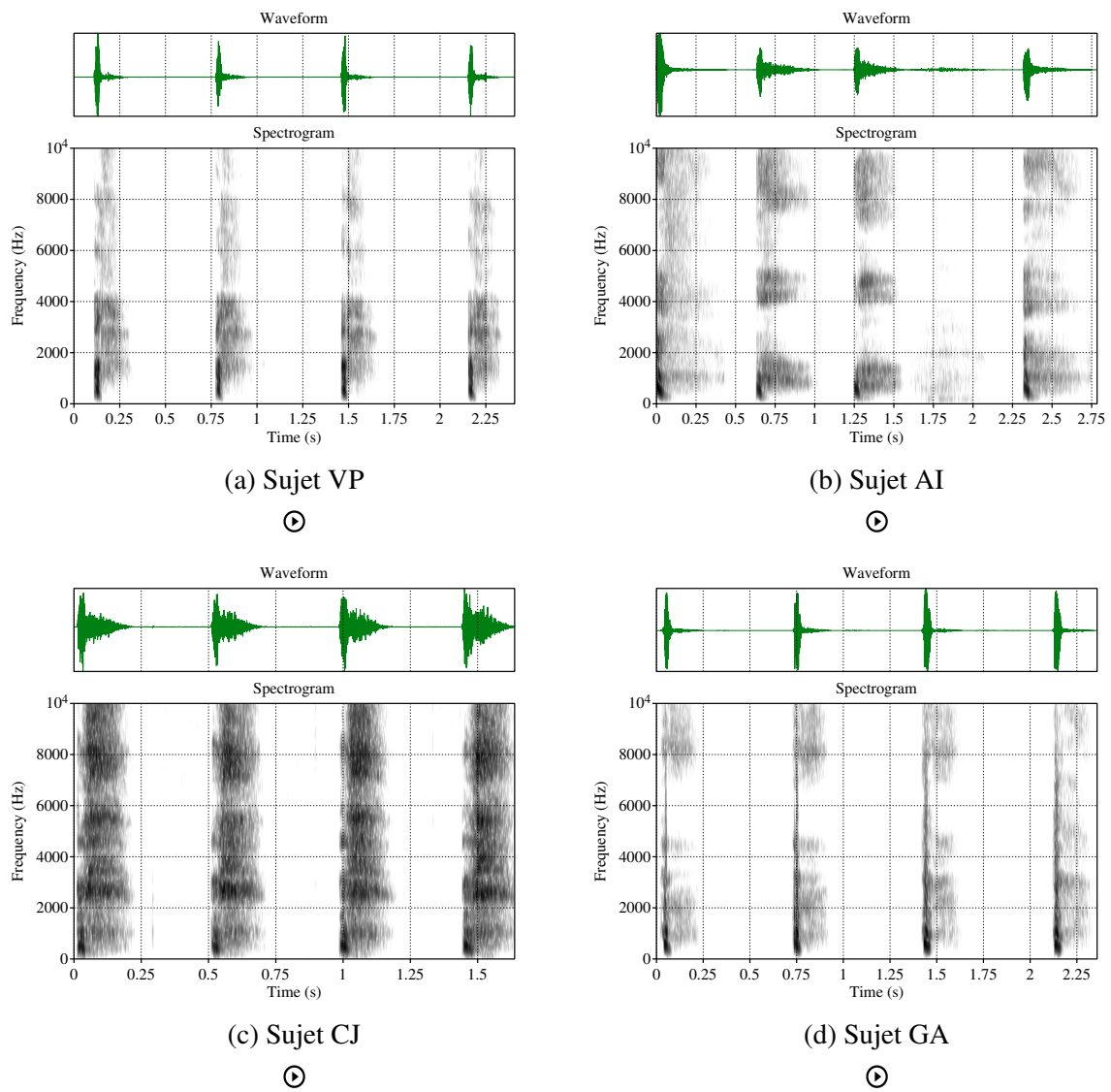


FIGURE 3.48 – Comparaison intersujet des signaux acoustiques et des spectrogrammes de la *ESH Snare Inward* [↓ $\hat{?}$ ]. Attention, échelles de temps différentes.

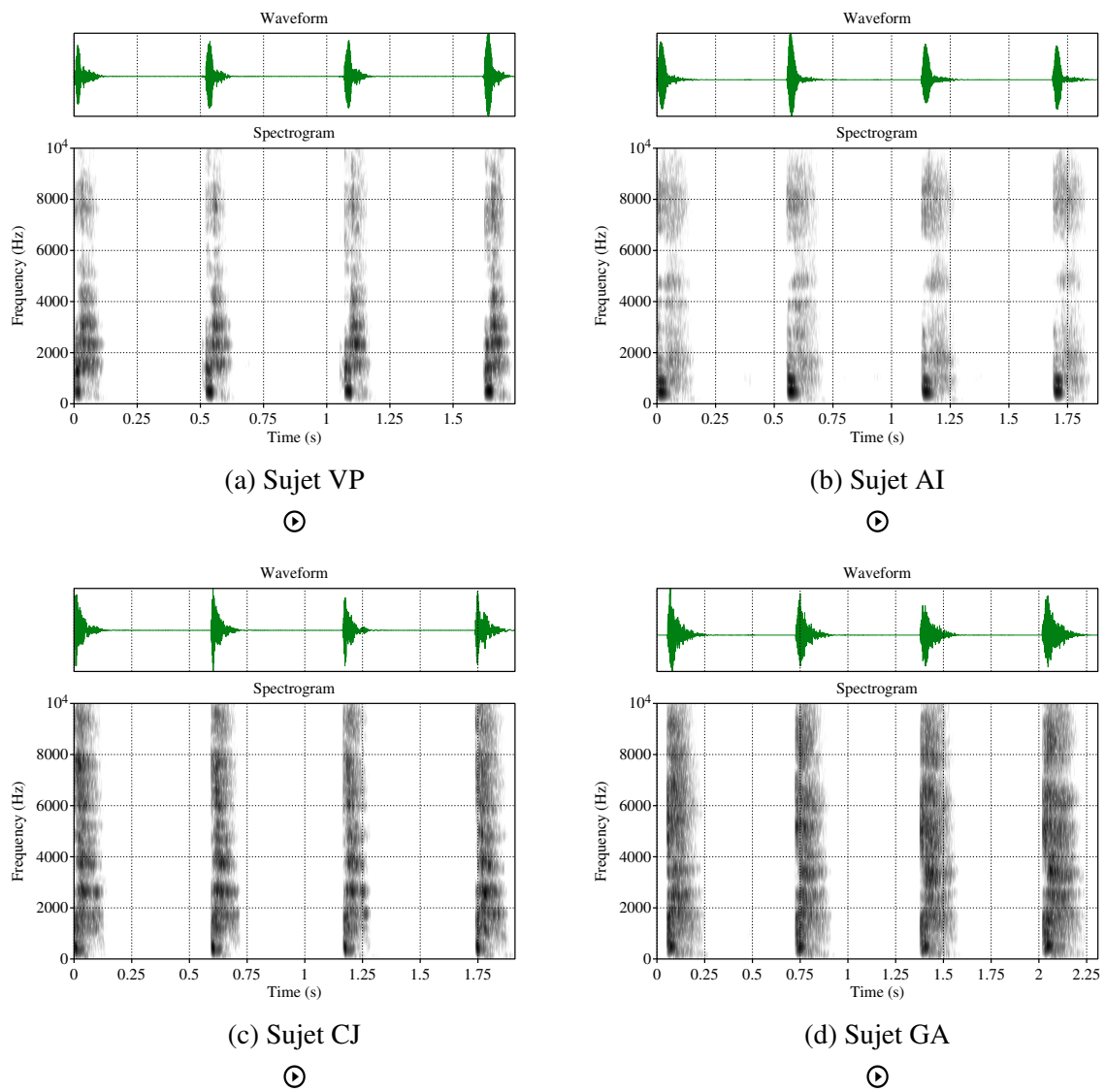
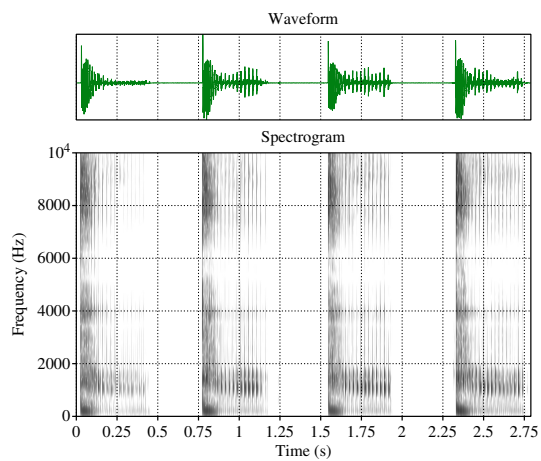
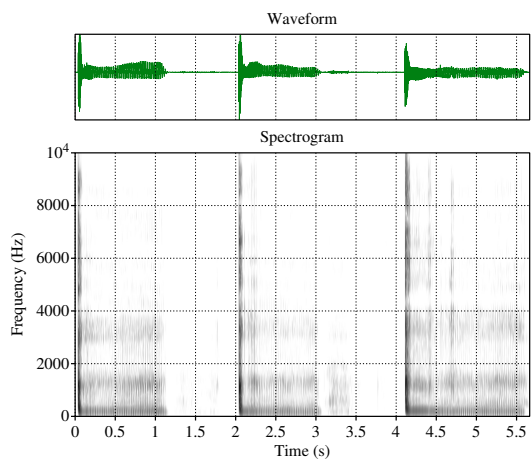


FIGURE 3.49 – Comparaison intersujet des signaux acoustiques et des spectrogrammes de la *Cough Snare* [ʔh]. Attention, échelles de temps différentes.

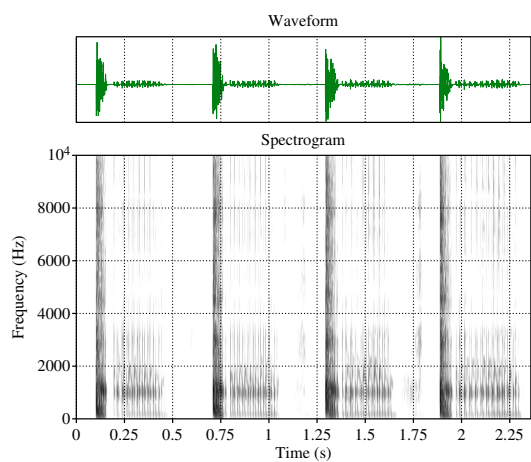




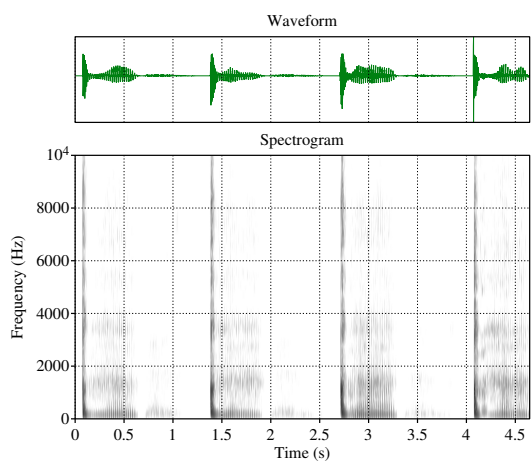
(a) Sujet VP



(b) Sujet AI



(c) Sujet CJ



(d) Sujet GA



FIGURE 3.50 – Comparaison intersujet des signaux acoustiques et des spectrogrammes du *Lips Roll* [ $\downarrow_B^1$ ]. Attention, échelles de temps différentes.

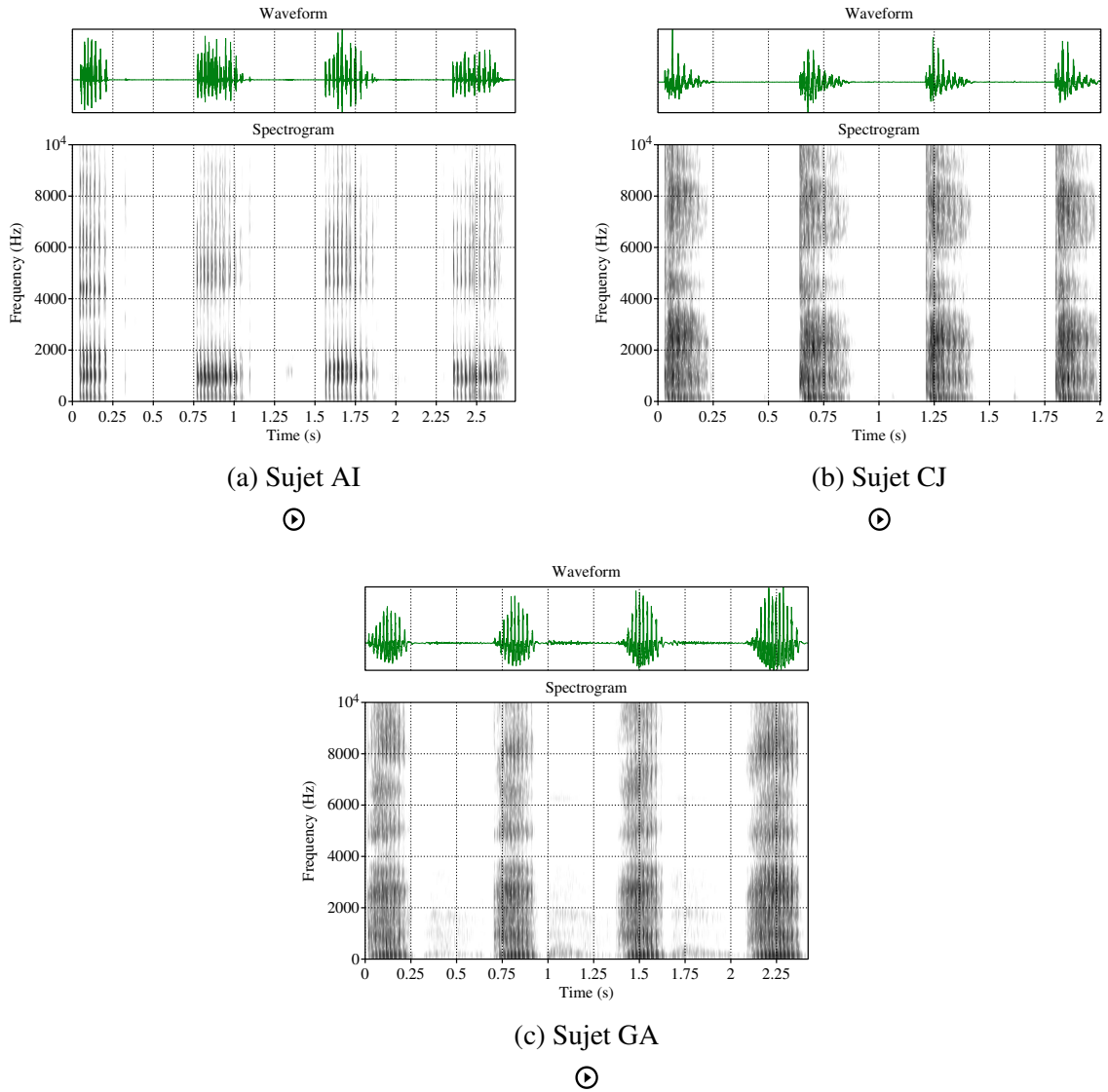


FIGURE 3.51 – Comparaison intersujet des signaux acoustiques et des spectrogrammes du *Clic Roll* [ $\downarrow\uparrow$ ] (Pas de données pour le sujet VP). Attention, échelles de temps différentes.

## 3.4 Le Human beatbox : un système combinatoire discret et complexe

Les sections précédentes présentaient les caractéristiques combinatoires des sons. Nous avons vu que la combinaison de mécanismes initiativés et articulatoires donnaient lieu à des signatures acoustiques qui semblent discrètes, du moins en ce qui concerne les imitations de percussions. Nous allons maintenant évoquer la deuxième partie du protocole où nous sommes intéressés à la voix beatboxée et aux répertoires individuels des sujets. Nous commencerons par la voix beatboxée et les mécanismes de modulation de la fréquence fondamentale puis nous conclurons ce chapitre en évoquant les mécanismes sifflés.

### 3.4.1 Voix beatboxée

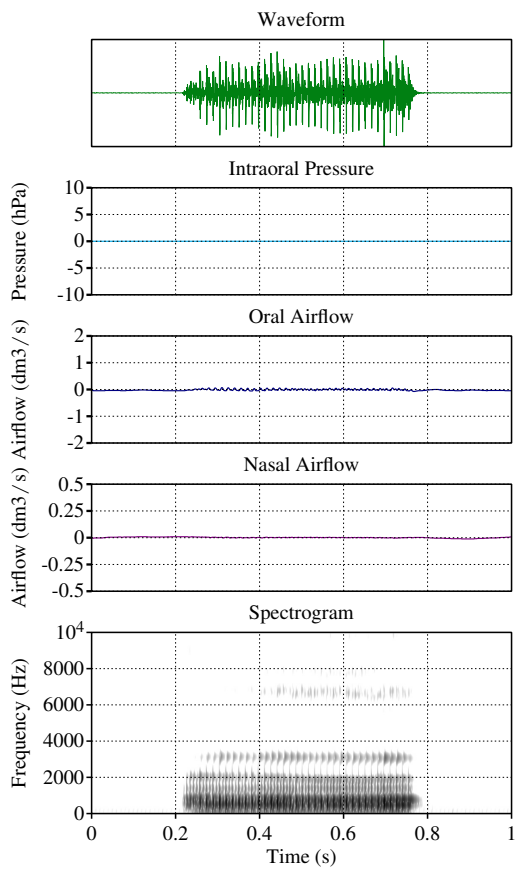
Nous allons maintenant présenter quelques aspects intéressants de la phonation dans le HBB qui n'ont pas pour but d'être exhaustifs. Pour cela, nous illustrerons nos propos par des productions beatboxées de type électronique (scratches vocaux, basses, effets ou « FX »). Nous analyserons à la fois la combinaison des gestes de production et la signature acoustique.

#### 3.4.1.1 Throat Bass

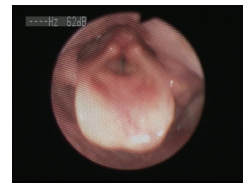
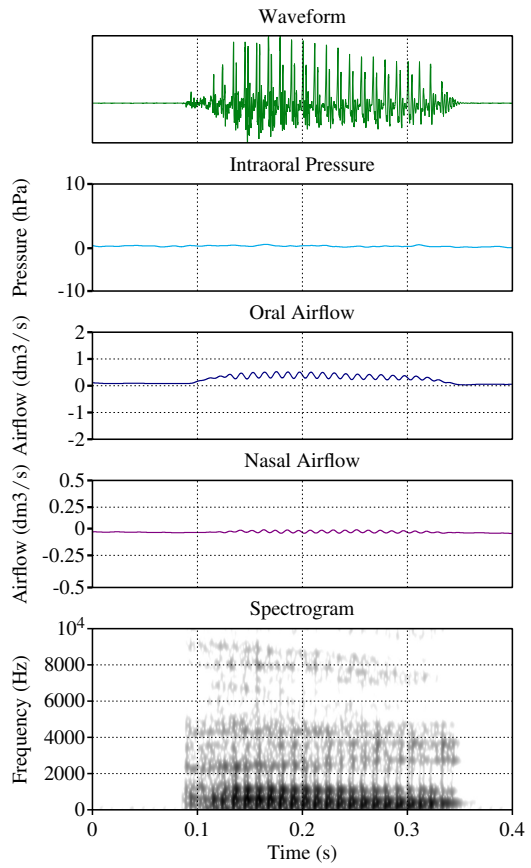
Tous les sujets ont pu produire une *Throat Bass* (Figure 3.52). Les quatre sujets ont produit de la phonation sur un flux d'air pulmonique égressif. Les données fibroscopiques des sujets VP et CJ montrent une fermeture aryépiglottique presque totale. Les sujets AI et GA utilisent de la phonation ventriculaire. Ces configurations sont similaires à celles de la voix craquée. Tous les spectrogrammes affichent une structure formantique et de l'énergie dans les fréquences basses. On note une composante de bruit supplémentaire en ce qui concerne les productions des sujets AI et CJ.

Le sujet VP est le sujet ayant produit la  $F_0$  la plus basse (67Hz). Cette  $F_0$  basse s'accompagne d'un débit d'air faible ( $0.1dm^3/s$ ) indiquant que la phase de fermeture du cycle de vibratoire est plus longue que la phase d'ouverture. Le sujet AI a produit une *Throat Bass* à 92Hz. Le débit d'air oral est à peu près de  $0.5dm^3/s$ , ce qui traduit une durée de fermeture moins longue que VP. Le sujet CJ a produit une *Throat Bass* à 75Hz avec un débit d'air oral de  $0.9dm^3/s$ , ce qui est plus élevé que les sujets VP et AI. On note que la pression intra-orale est positive, il doit y avoir une constriction ou un resserrement dans le conduit vocal. Le sujet GA a produit une *Throat Bass* à 88Hz avec un débit d'air oral diminuant de  $0.8dm^3/s$  à  $0.2dm^3/s$ . La diminution du débit ne s'accompagne pas d'une diminution de la  $F_0$  et n'est donc pas lié à un changement du mode de vibration des bandes ventriculaires. La diminution du débit d'air oral est inversement proportionnelle à l'augmentation de pression intra-orale. Cela indique un resserrement dans le conduit vocal.

La Figure 3.53 montre que la *Throat Bass* possède une structure harmonique. La fréquence fondamentale est notée par un trait rouge. On note qu'elle est plus basse chez VP et CJ pour qui nous avons remarqué une constriction aryépiglottale. De plus, les spectres diffèrent dans leur amplitude et dans leur pente spectrale. Chez CJ et GA, l'amplitude de l'enveloppe spectrale est plus intense en comparaison à VP et AI. On remarque que chez VP et CJ, l'énergie diminue progressivement à partir de 1kHz alors que chez AI et GA, l'énergie diminue à partir de 1.2kHz. Il est possible que les différentes pentes spectrales soient dues à l'engagement du complexe aryépiglottique par les sujets VP et CJ.



(a) Sujet VP



(b) Sujet AI



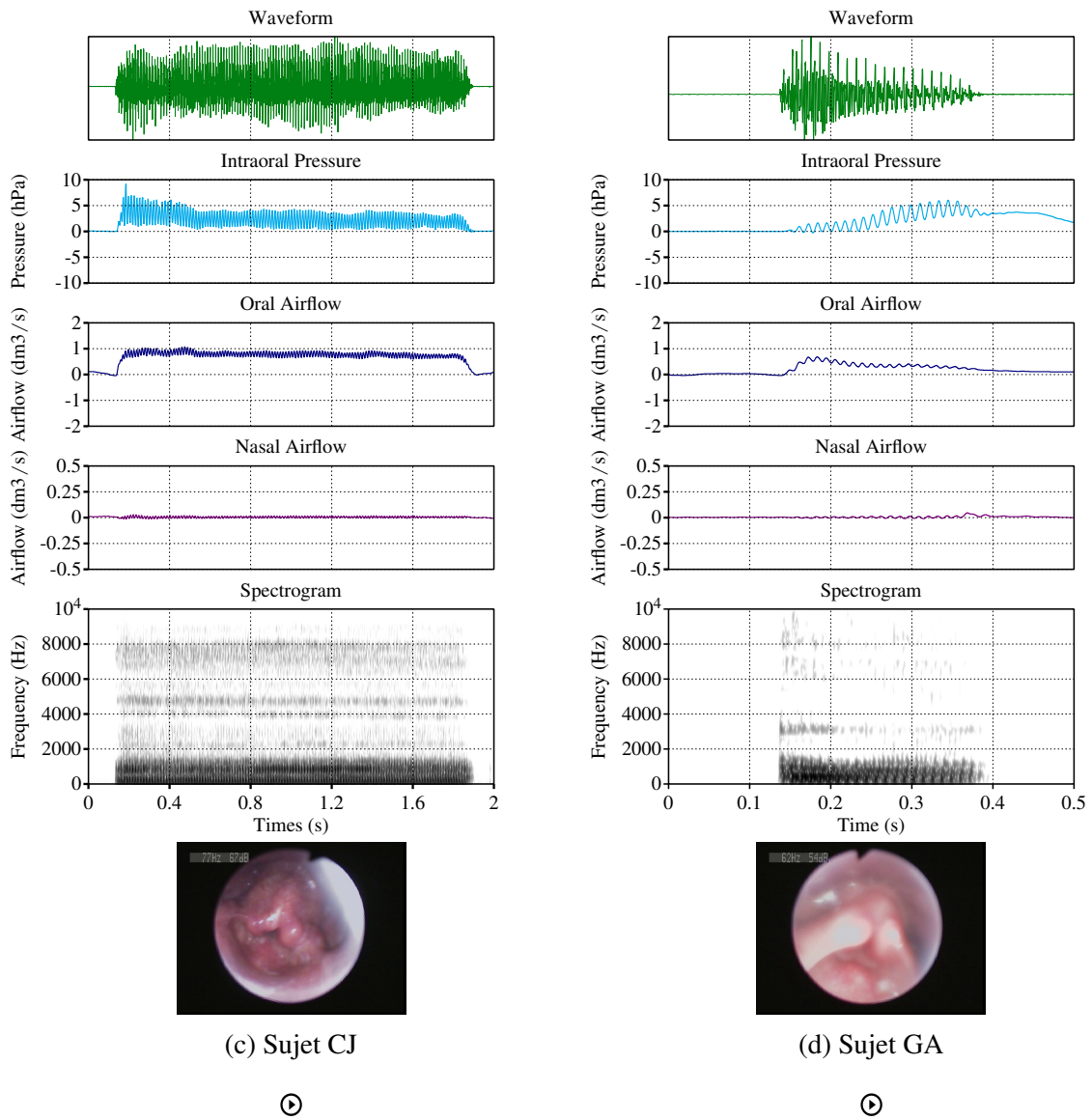
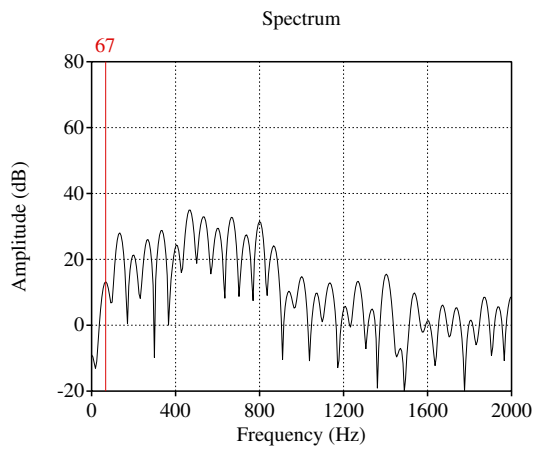
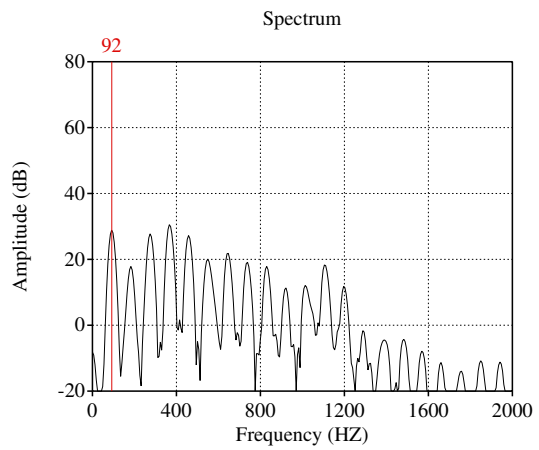


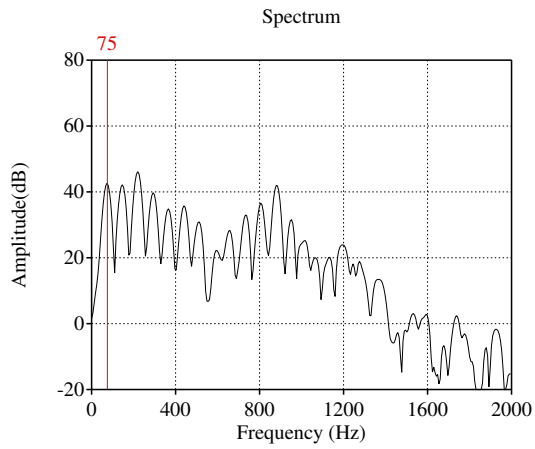
FIGURE 3.52 – Production de la Throat Bass par 4 sujets différents (Attention les échelles de temps sont différentes)



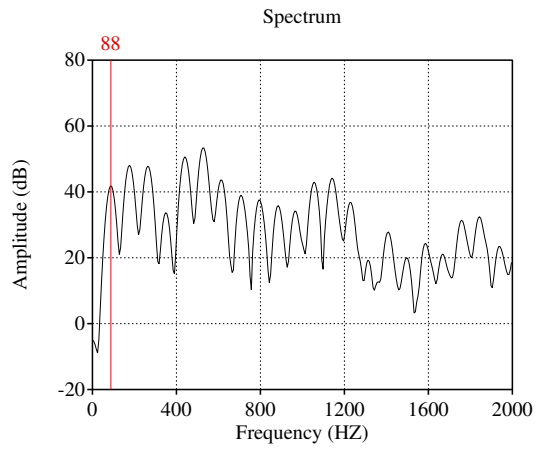
(a) Sujet VP



(b) Sujet AI



(c) Sujet CJ



(d) Sujet GA

FIGURE 3.53 – Spectre FFT de la structure harmonique des basses fréquences de la *Throat Bass*. (Le trait rouge indique H1)

Les différences acoustiques individuelles posent la question du caractère discret de la *Throat Bass*. Un test perceptif de dénomination, où l'on demanderait à des beatboxeurs et des beatboxeuses de nommer le son perçu, permettrait de trancher la question du caractère discret ou de ce son. Si les sujets s'accordaient à dire qu'il s'agit d'une *Throat Bass* alors la signature acoustique pourrait être qualifiée de discrète.

### 3.4.1.2 Inward Bass - Sujet CJ

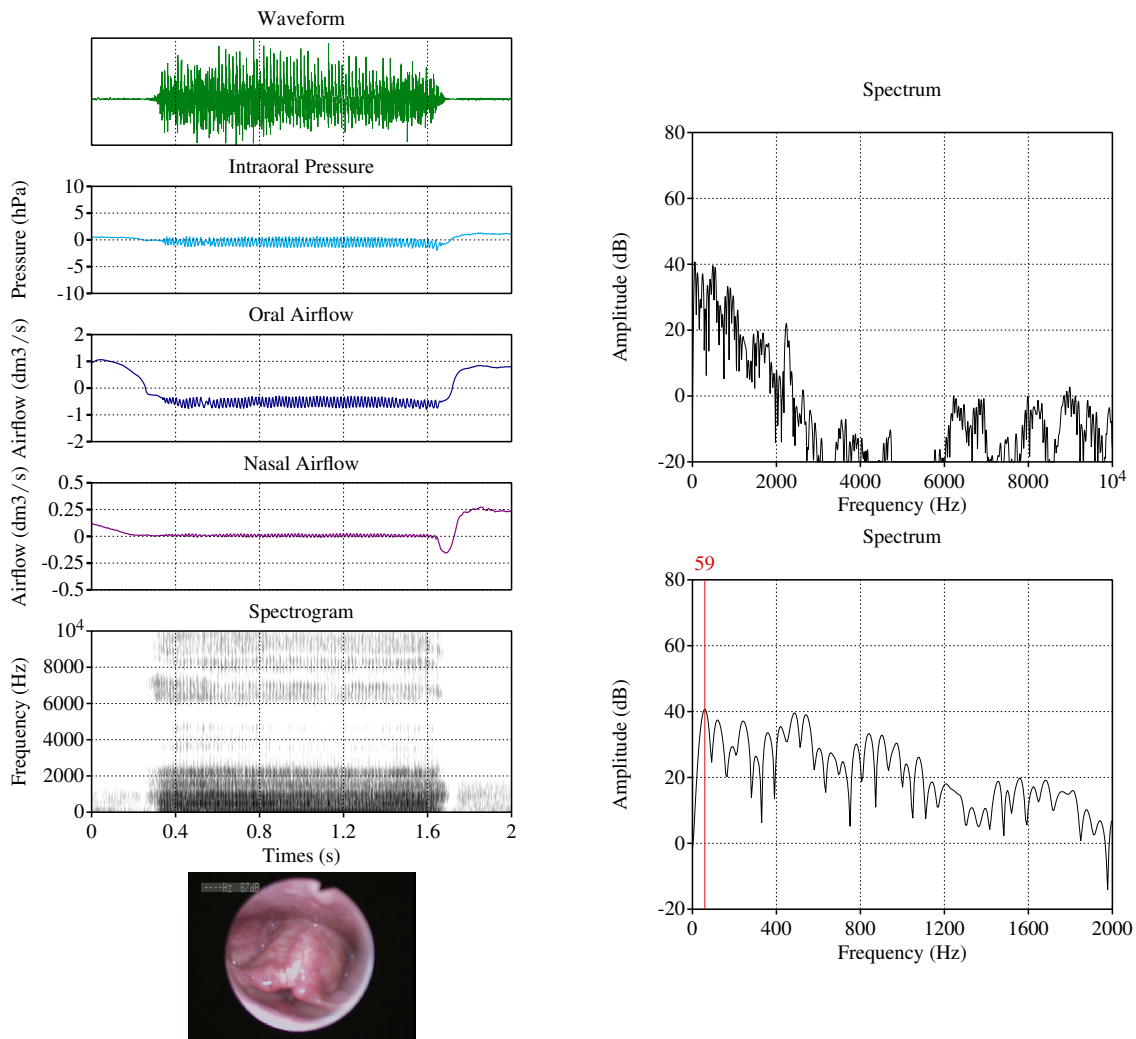


FIGURE 3.54 – *Inward Bass* - Sujet CJ

Ⓢ

L'*Inward Bass* du sujet CJ est similaire à sa *Throat Bass* (Figure 3.52c). On remarque en premier lieu une configuration laryngée similaire à la différence qu'ici, la constriction aryépiglottique est plus importante, possiblement à cause d'un recul de la racine de la langue. La pression intra-orale oscille autour de 0. Le débit d'air oral est négatif (flux pulmonique ingressif) concernant l'*Inward Bass* et celui-ci est légèrement moins important comparé à la *throat Bass*. Ceci est dû au fait que la constriction aryépiglottique est plus importante pour l'*Inward Bass*. Aucune différence n'est observée concernant le débit d'air nasal.

En revanche, des différences acoustiques peuvent être observées sur les spectrogrammes. Pour visualiser ces différences, nous allons présenter les spectres de ces sons (Figure 3.54). Les spectres de ces deux sons présentent des différences quant au fondamental et à la répartition fréquentielle. La fréquence fondamentale est de 59Hz. Cette différence s'explique par le degré plus important de constriction aryépiglottique. Concernant la répartition fréquentielle, on observe très peu d'énergie entre 3 et 6 kHz mais il y a de l'énergie au-dessus de 6kHz. Sur le spectre FFT de 2kHz, la pente spectrale est moins raide par rapport à la *Throat Bass*. Une étude acoustique plus approfondie permettrait de savoir si les différences de pente spectrale et de répartition de l'énergie entre la *Throat Bass* et l'*Inward Bass* sont des indices acoustiques de la direction du flux d'air.

### 3.4.1.3 Vibration Bass

La Figure 3.55 illustre la production d'une basse vocale avec une importante composante de bruit. Le sujet GA a produit une basse vocale qu'il n'a pas nommé, cette basse est perceptuellement similaire à une « *Vibration Bass* ». La *Vibration Bass* du sujet GA combine des composantes périodiques et apériodiques. Ce son est produit sur un flux pulmonique égressif à la fois oral et nasal. On note une partie à prédominance orale au début et nasale à la fin. De plus, au fur et à mesure que le débit nasal augmente, le débit oral diminue.

Les changements de débit oral et nasal affectent les caractéristiques acoustiques. La *Vibration Bass* commence par une occlusion aryépiglottique et la phonation débute au moment du relâchement. La composante périodique est en fait irrégulière. Elle est due à la vibration des tissus du complexe aryténoïdien qui se propage aux plis aryépiglottiques. Les composantes apériodiques sont dues à la glotte ouverte. C'est la combinaison de ces deux composantes qui donne lieu à un signal acoustique irrégulier et d'aspect apériodique. Contrairement aux trilles aryépiglottiques « classique » où un seul pli (parfois les deux) oscille sur la base de l'épiglotte, ici les plis sont tendus. Il s'agit d'une trille aryépiglottique sans fermeture aryépiglottique. La fréquence de vibration passe de 64Hz (partie orale) à 45Hz (partie nasale).

La Figure 3.55 montre aussi deux spectres de la partie orale (trait plein) et nasale (pointillés). Sur le spectre du haut (0-2kHz), la structure n'est pas régulière mais présente des composantes « pseudo harmoniques ». Les pics se succèdent à intervalle irréguliers avec quelques hertz de différence. Le spectre du bas illustre les effets du couplage entre cavité orale et nasale par un lissage LPC à 12 coefficients du FFT de 10kHz. Sur la partie orale, on note des résonances à 542Hz, 1.8kHz, 3kHz, 5.5kHz, 7kHz et 9.7kHz. La nasalisation a des effets très nets sur les résonances. On voit l'introduction d'anti-résonances et des nouvelles résonances. Les anti-résonances sont nettement visibles à 2.4kHz et 5kHz et annulent les résonances de la partie orale se trouvant à 1.8kHz et 5.5kHz. On note l'apparition de résonances nasales à 3.9kHz et 6.8kHz. La première résonance de la partie orale est à 542Hz alors que pour la partie nasale elle est à 800Hz. Nous pensons que cette augmentation reflète un mouvement lingual.



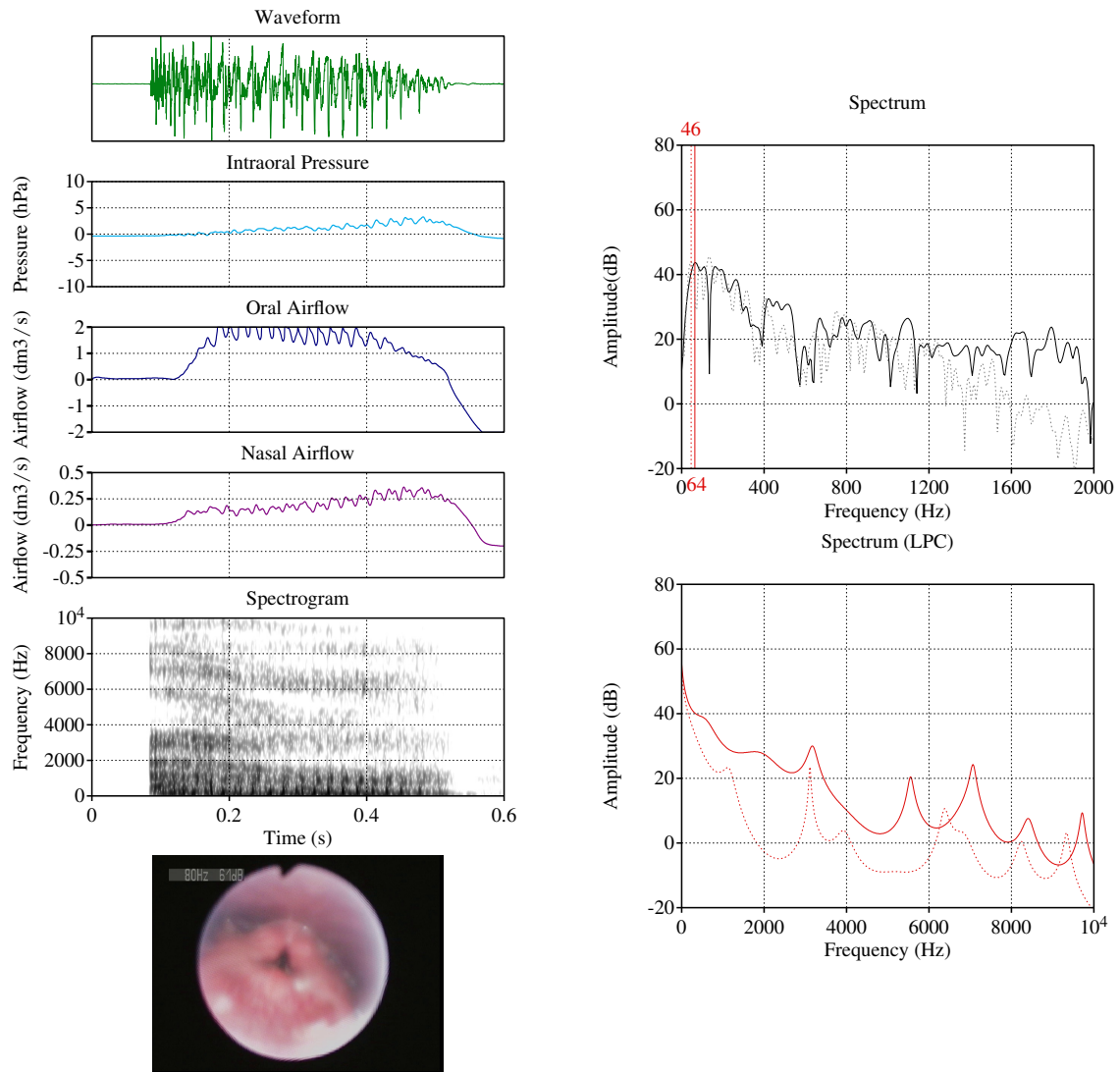


FIGURE 3.55 – *Vibration Bass* - sujet GA. Sur les spectres les traits pleins correspondent à la partie orale du son ; les pointillés à la partie nasale.

⊙

### 3.4.1.4 Inward Bass - Sujet GA

L'*Inward Bass* du sujet GA est bien ingressive (Figure 3.56). Le débit d'air nasal indique que le port vélopharyngé est fermé. On observe des oscillations en dessous de 0 sur le signal de pression intra-orale. Aucun signe de constriction aryépiglottique n'est visible sur l'image fibroscopique. Les plis vocaux restent ouverts tout du long. En regardant attentivement l'image fibroscopique au niveau de l'espace inter-aryténoïdien (juste au-dessus de la glotte) on remarque que les tissus sont « aspirés » vers l'intérieur de la glotte. Sur les vidéos, on remarque plusieurs cycles d'« aspiration ».

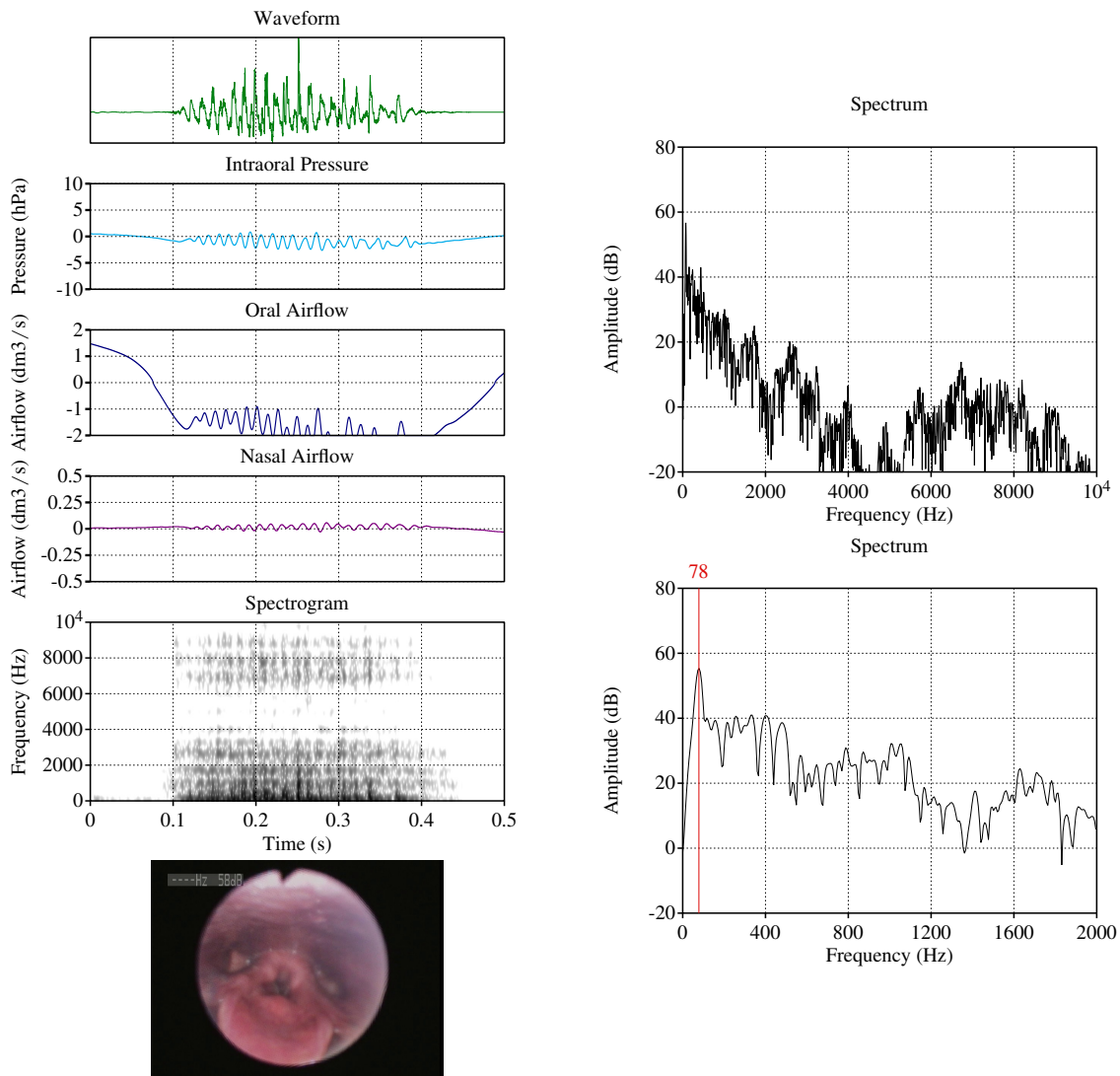


FIGURE 3.56 – *Inward Bass* - GA.

Ⓢ

L'enveloppe spectrale correspond à un type phonatoire de voix soufflée comme le montre la répartition diffuse de l'énergie de 0 à 10kHz. Il y a une anti-résonance à 4kHz. On note qu'il y a de l'énergie entre 6 et 10kHz. Nous avons observé de l'énergie entre 6 et 10kHz pour l'*Inward Bass* du sujet CJ. Il est possible que la présence d'énergie dans les hautes fréquences caractérise la phonation ingressive. Il y a des résonances à 400Hz, 975Hz et 1.6kHz. L'irrégularité du cycle

vibratoire se traduit par des harmoniques irréguliers. Cette *Inward Bass* utilise un mécanisme phonatoire proche de la voix soufflée.

### 3.4.1.5 Double Voix

Dans nos données, nous avons trouvé des exemples de phonation où les composantes périodiques ne sont pas en phase. Le sujet GA appelle cette voix la « *double voix outward* ». Il ne s'agit pas à proprement parler d'un cas *biphonation* comme le nom pourrait le suggérer. En effet, dans le cas de la *double voix outward*, il n'y a pas plusieurs sources ou de sous-harmoniques. Il s'agit plutôt d'un phénomène similaire au battement. Castellengo (2015, p. 123) explique que « les battements se produisent lorsqu'on émet simultanément deux sons de fréquences voisines, par exemple  $f_1$  et  $f_2$ . En se combinant, les deux ondes primaires produisent une onde résultante dont l'amplitude globale a pour fréquence  $f_1 - f_2$  ». Le battement peut aussi se manifester par un décalage de la phase entre des sinusoïdes ce qui se traduit par des pics en phase et des pics en anti-phase (Hixon et al., 2020). Plus le battement est élevé, plus on aura l'impression d'entendre deux tonalités. La double voix du sujet GA est initiée par un flux pulmonique égressif à la fois oral et nasal.

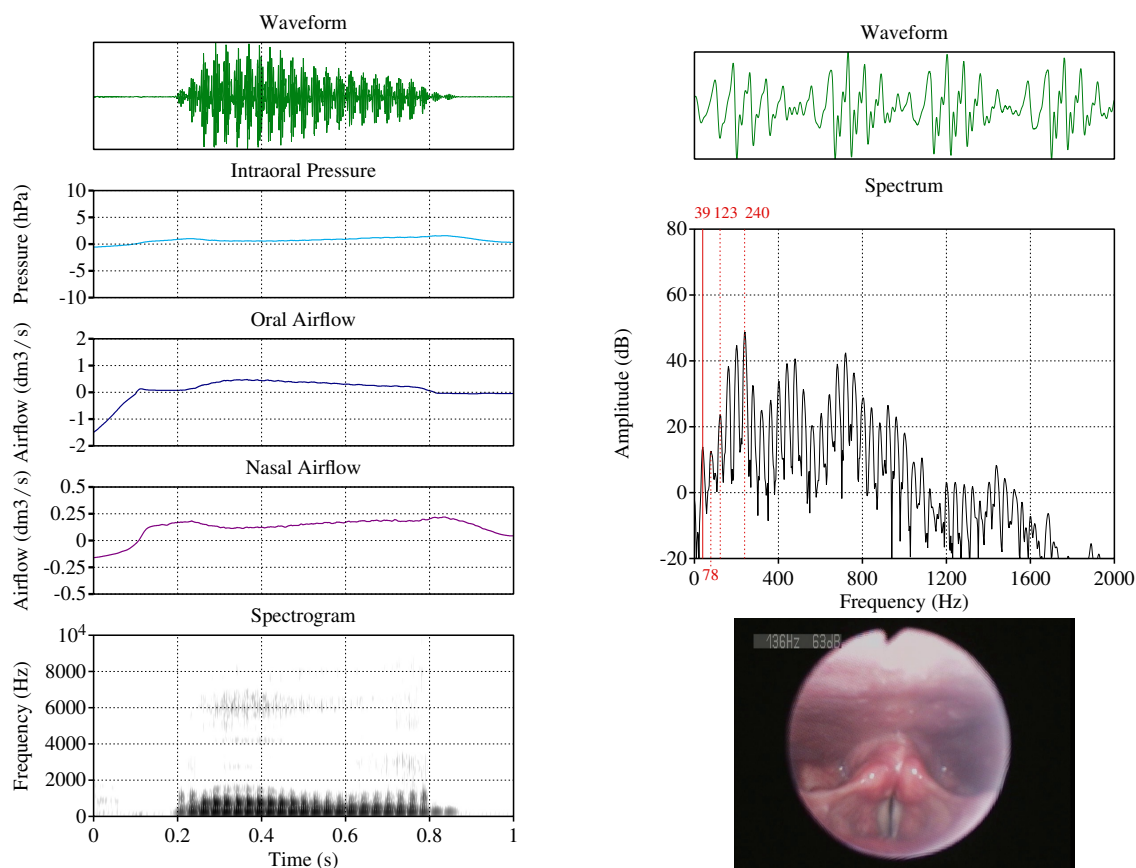


FIGURE 3.57 – Production d'une « *double voix outward* » par le sujet GA : données aérodynamiques (gauche), spectre FFT (droite)

⊙

Concernant la structure du spectre, il présente une structure harmonique régulière. Sur le spectre, on note que  $H_1 = 39\text{Hz}$  et  $H_2 = 78\text{ Hz}$ .  $H_1$  et  $H_2$  sont ici des harmoniques au sens littéral

c'est-à-dire des multiples entiers. Toutefois,  $H3 = 123\text{Hz}$ ,  $H4 = 162\text{Hz}$ ,  $H5 = 201\text{Hz}$  soit 6Hz de plus que ce qu'on s'attendrait à observer. En divisant ces harmoniques par le fondamental on remarque qu'ils sont des multiples non-entiers : 3.15, 4.15, 5.15 etc. Nous avons affaire ici à ce qui s'apparente à un « *battement* ». Est-ce que le décalage observé est un décalage fréquentiel ou temporel ? Sur les images fibroscopiques dont nous disposons (Figure 3.58), nous notons un décalage temporel entre le mouvement du pli gauche et du pli droit vers la ligne médiane (milieu de la glotte). En effet, sur trois images successives espacées de 25ms, nous observons que le pli gauche se déplace vers la ligne médiane et 25ms plus tard, le pli droit se rapproche à son tour de la ligne médiane. Il y a un décalage de la mise en vibration des plis vocaux.

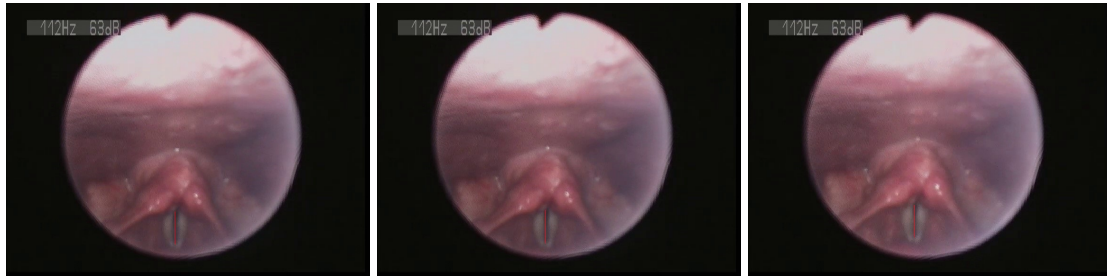


FIGURE 3.58 – Trois images fibroscopiques extraites successivement lors de la production de la *double voix outward*. La ligne rouge symbolise la ligne médiane (milieu de la glotte). Un subtil décalage du mouvement de chaque pli vers la ligne médiane peut être observé.

### 3.4.2 Mécanismes de modulation de la fréquence fondamentale

Nos données montrent que les beatboxeurs et les beatboxeuses peuvent produire des  $F0$  extrêmement basses (25Hz) et extrêmement hautes (733Hz).

Commençons par la production d'un scratch vocal par le sujet VP (Figure 3.59). Les productions de scratches vocaux imitent diverses techniques de « *scratching* ». Dans la terminologie des DJ, le *scratching* consiste à faire varier manuellement la vitesse de lecture d'un vinyle sur une platine. Lorsque la vitesse est rapide, la hauteur du son est plus élevée ; lorsque que la vitesse est lente la hauteur est basse. Pour ce son, nous disposons de données aérodynamiques, acoustiques et articulatoires (IRM et fibroscopie). Le scratch illustré ici est produit sur un flux pulmonique égressif oral comme l'indique la faible pression ( $P_{intraorale} < 5hPa$ ). Le premier trait rouge sur la figure indique où la première image IRM et fibroscopique a été extraite, c'est-à-dire sur la partie fricative. À cet instant, la  $F0$  est à 160Hz. L'impression auditive aigüe vient de la résonance intense vers 7kHz et non de la  $F0$ . L'image IRM montre la configuration du conduit à cet instant : les lèvres sont quelque peu arrondies et protruses (en comparaison au frame 3), la langue forme une constriction dans la zone alvéolaire et la cavité pharyngale ne montre aucun signe de constriction. L'image fibroscopique correspondante montre des signes d'élongation des plis vocaux. Cette élongation est due à la contraction du muscle thyro-cricoïdien (cf. Figure 1.4) dont l'action provoque une bascule vers l'avant du cartilage thyroïde et allonge les plis vocaux.

Le deuxième trait rouge indique l'extraction des images IRM et fibroscopie au moment où la fréquence fondamentale est au plus bas (25Hz). Si l'on regarde le signal acoustique, on peut distinguer des périodes de 35ms environ. À cet instant précis, le conduit vocal a une configuration très différente. On note de la labio-vélarisation accompagnée de rétroflexion rappelant la configuration du *Kick Lingual* (Figure 3.12). Cette stratégie est utilisée ici pour donner un timbre grave à la production. Au niveau pharyngo-laryngé on observe un recul de la racine de

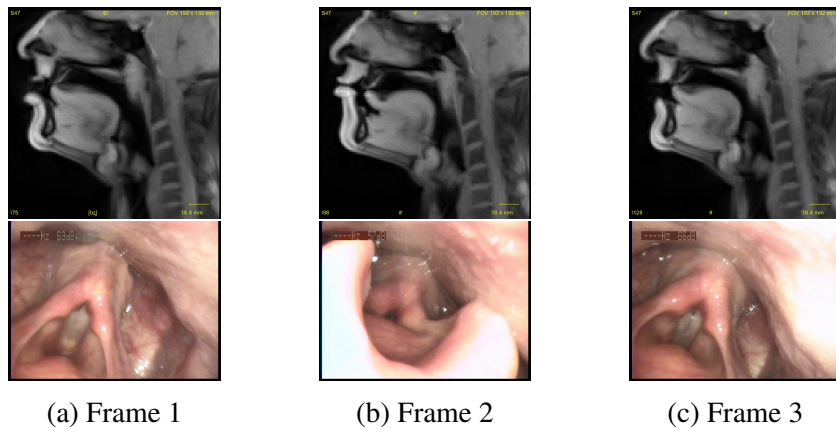
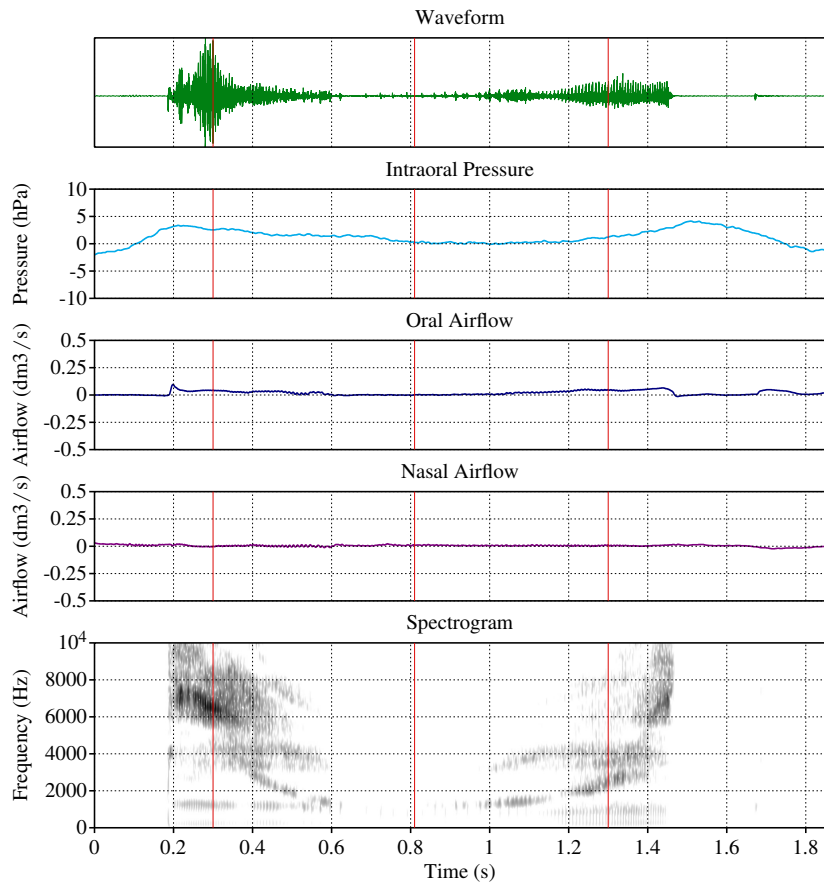


FIGURE 3.59 – Modulation de la F0 d'un scratch vocal du sujet VP



la langue ainsi qu'une bascule de l'épiglotte. L'image fibroscopique correspondante confirme la bascule postérieure de l'épiglotte venant réduire la profondeur du tube épilaryngé. La longueur des plis vocaux est fortement réduite au point de ne presque plus être visibles. Cette configuration est similaire à la voix craquée avec constriction aryépiglottique et au « Fry ». Cette action de pliage laryngé permet de diminuer grandement la fréquence de vibration des plis vocaux. La constriction aryépiglottique implique une action conjointe des muscles laryngés intrinsèques (thyro-aryténoïdien, crico-aryténoïdien latéral, l'inter-aryténoïdien) et de muscles extrinsèques (hyo-glosse, supra-hyoïdien) (Esling et al., 2019).

Enfin, les données à l'instant 3 sont similaires à l'instant 1 à la différence qu'on n'observe plus de labialisation mais on observe de la vélarisation. La Figure 3.60 illustre une autre straté-

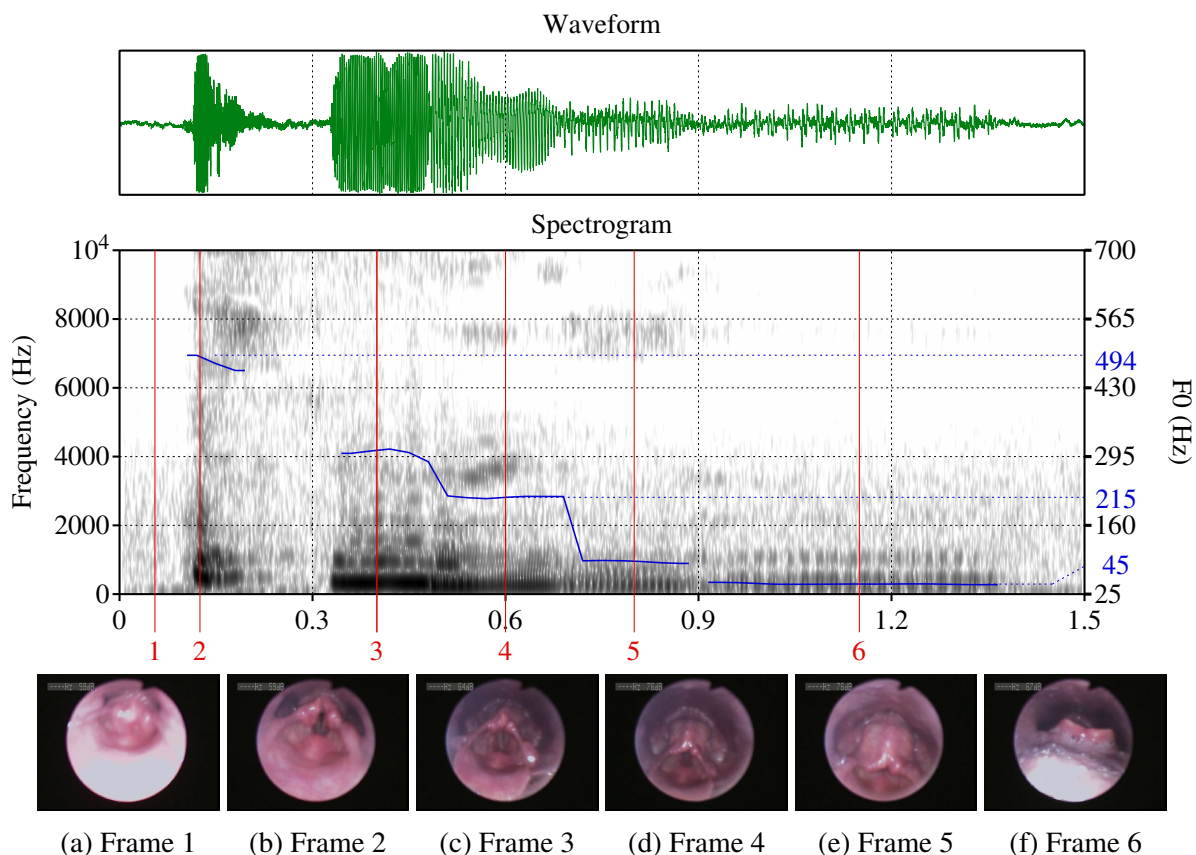


FIGURE 3.60 – Modulation de la F0 de 494Hz à 45Hz par le sujet CJ

Ⓢ

gie de modulation de la fréquence fondamentale. Pour cet exemple, nous disposons seulement de données fibroscopiques et acoustiques. Le bruit présent sur le spectrogramme est généré par le système d'acquisition du fibroscope. Les images fibroscopiques ont été extraites à 6 instants différents représentés par des tracés rouges.

Le premier instant ne correspond pas à de la phonation mais à une occlusion aryépiglottique. Cette occlusion doit permettre d'augmenter la pression sous-glottique dans le but de produire une F0 élevée (500Hz). Cette pression élevée doit bien entendu s'accompagner des ajustements laryngés nécessaires. La première observation qui peut être faite concerne la position du larynx. Sur l'image 1 le larynx est en position haute et descend au fur et à mesure. La configuration est la même que pour celle de la *ESH Snare Inward* décrite dans la Figure 3.20. La position basse accroît la tension des tissus laryngés sur l'axe vertical et augmente la fréquence de vibration des plis vocaux. Cependant, cette observation ne peut pas expliquer à elle seule pourquoi ce sujet masculin produit une F0 de presque 500Hz. Nous avons vu qu'il était possible d'augmenter la F0 en allongeant les plis vocaux. Ici les plis vocaux ne semblent pas allongés. Il est possible de produire une F0 élevée par une tension isométrique. La tension isométrique d'un muscle réfère à la contraction de ce dernier sans mobiliser les structures auxquelles il se rattache (tendons, cartilages, os). Dans notre cas, il est possible qu'une tension isométrique s'applique aux plis vocaux : leur tension augmente sans que leur longueur ne change. Ils vibrent à haute fréquence



sans activer le crico-thyroïdien pour allonger les plis vocaux. Ce mécanisme a été décrit pour la production de la *Cough Snare* chez deux beatboxeurs (CJ et GA) dans Dehais-Underdown et al. (2021). Enfin, nous pensons qu'il s'agit d'une production ingressive. La configuration du larynx rappelle celle de la *ESH Snare Inward*. De plus, l'impression auditive résultante de la production paraît semblable à un réflexe de type hoquet<sup>6</sup>. Cependant en l'absence de données aérodynamiques cette hypothèse est purement spéculative.

Au cours des instants 3 et 4, la F0 diminue à 300 et 215 Hz. Les images fibroscopiques correspondantes indiquent une diminution de la longueur des plis vocaux ainsi qu'un contact sur toute leur longueur. Les effets de cette action conjointe s'observent sur le signal acoustique : on note des périodes plus espacées et une diminution de l'amplitude du signal. Les instants 5 et 6 correspondent respectivement à la *Throat Bass* (cf. Figure 3.52c) et à la *Chest Bass* (Figure 3.61 - voir après). La constriction aryépiglottique a des effets très nets sur la F0 et l'amplitude du signal acoustique. En effet, la F0 tombe à 90Hz, les périodes sur le signal sont plus longues et l'amplitude diminue de façon importante. L'instant 6 illustre une stratégie de diminution de la F0 (45Hz) en mobilisant la langue. Sur l'image 6, nous distinguons un net recul à la fois de la racine de la langue et du dos de la langue, ce dernier entrant dans le champs de vision du fibroscope. Cette rétraction est l'illustration de ce que Edmondson et Esling (2006) décrivaient comme l'action de la valve 4 (« *epiglottopharyngeal constriction* »). Le recul de la langue et la constriction épiglotto-pharyngale fonctionne comme un niveau supplémentaire de serrage laryngé (en plus de la fermeture glottique, ventriculaire et aryépiglottique). L'air expulsé des poumons passant dans le tube laryngé doit maintenant avoir une pression suffisamment élevée afin de produire de la phonation. Dans le cas de la *Chest Bass*, il est assez difficile de savoir si en plus d'une oscillation du pli aryépiglottique il y a aussi une oscillation des plis vocaux et des bandes ventriculaires. Étant donné le niveau important de fermeture à différents niveaux, une oscillation supplémentaire pourrait être le résultat du couplage des plis vocaux et des bandes ventriculaires vibrant comme une seule structure. Cependant, rien dans les données ne laisse penser qu'il y ait de la biphonation.

Regardons en détail les données concernant la *Chest Bass*. La Figure 3.61 illustre la production de la *Chest Bass* par des signaux acoustique, de pression, de débit d'air oral et nasal, des spectres ainsi que par des images fibroscopiques de la vibration du pli aryépiglottique gauche. Sur le signal acoustique on remarque à l'oeil nu les périodes de la *Chest Bass* sur 400ms. Les périodes durent 23ms et la F0 est donc de 42Hz environ. Cette basse est produite sur un flux pulmonique égressif. Le port vélo-pharyngé est fermé.

La Figure 3.61 illustre aussi les caractéristiques spectrales de la *Chest Bass*. Celle-ci possède seulement de l'énergie en dessous de 2kHz mais possède tout de même des résonances à 400 et 700Hz. Contrairement aux phonations de type craquée analysées plus haut, l'énergie est compacte de 0 à 800Hz puis diminue soudainement. Comme nous l'avons mentionné H1 = 42Hz. H2 se trouve à 84Hz.

---

6. Le hoquet est un réflexe physiologique dû à une contraction involontaire des muscles inspiratoires résultant en un « hic » aigu

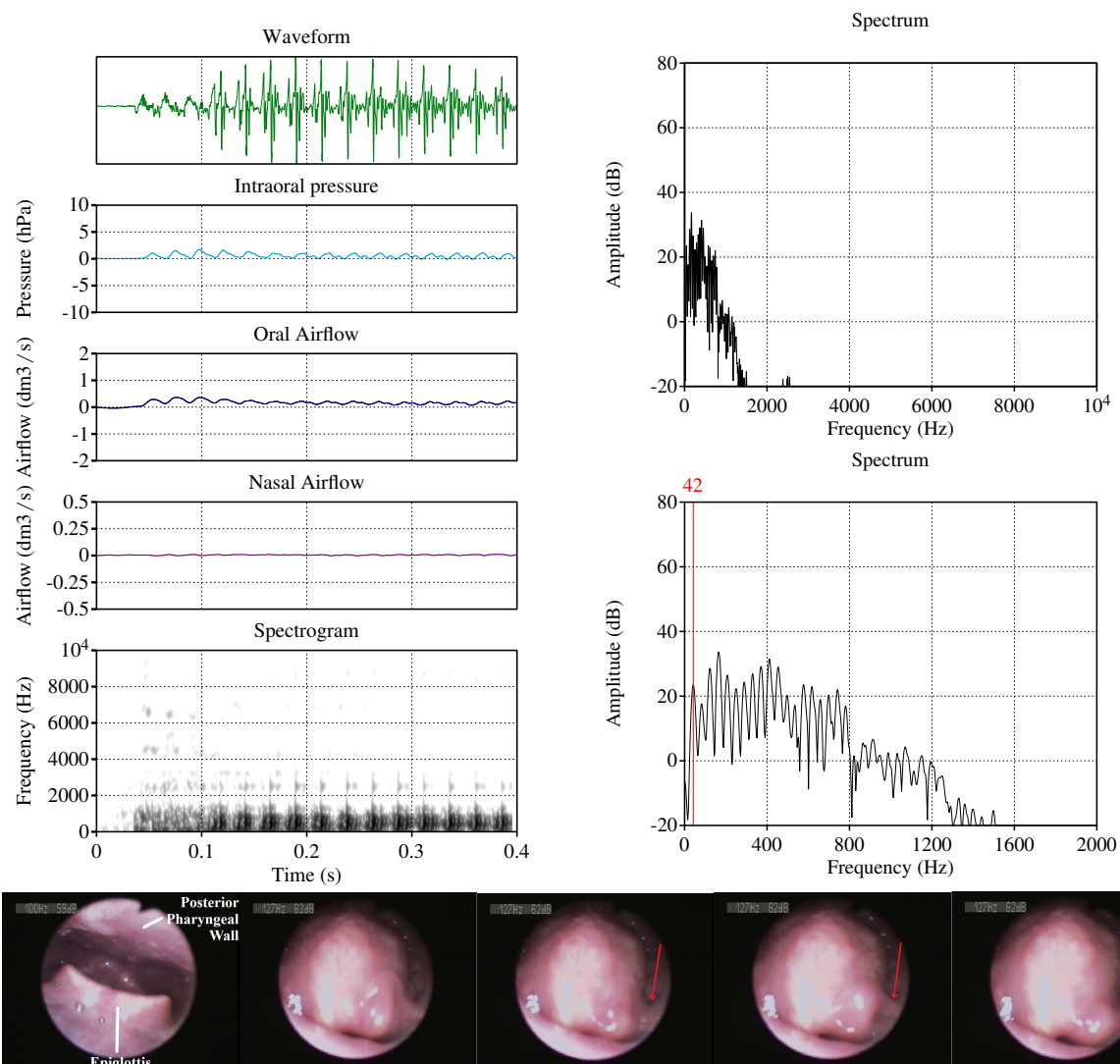


FIGURE 3.61 – Vibration du pli aryépiglottique gauche lors de la production de la « *Chest Bass* » du sujet CJ



### 3.4.3 Mécanismes sifflés

Les mécanismes sifflés ne sont pas rares dans la communication humaine. Ils peuvent établir des contrastes phonologiques comme dans certaines langues Bantoues (Shosted, 2006). Le mécanisme de sifflement peut aussi se substituer aux mécanismes d'articulation et de phonation pour produire de la parole. On parle alors dans ce cas précis de langues sifflées (Meyer, 2015). Notre corpus contient des mécanismes sifflés.

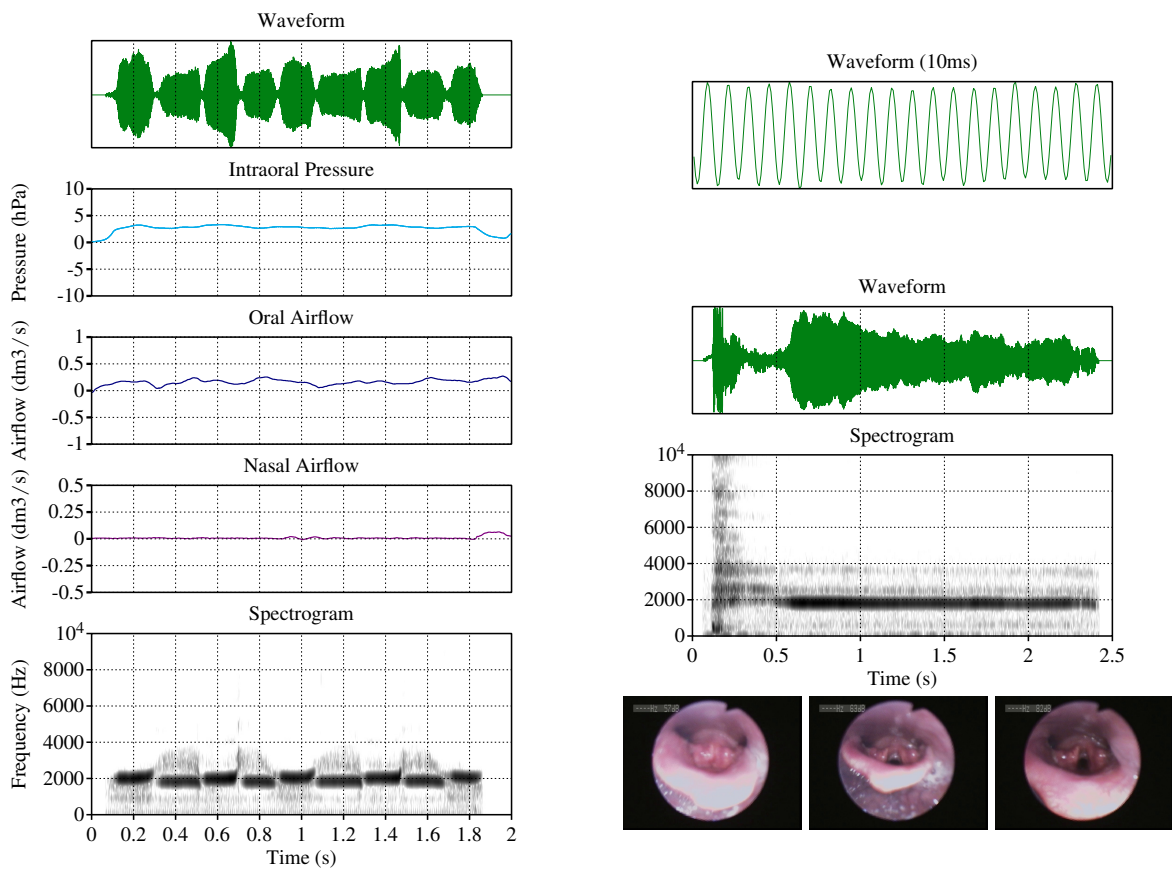
La Figure 3.62a illustre la production du « *Tongue Whistle* » et la Figure 3.62b illustre un sifflement glottal avec tout en haut un zoom sur le signal sur une fenêtre de 10ms. Nous ne disposons pas de données articulatoires concernant le sifflement lingual, cependant il existe un tutoriel sur l'apprentissage du *tongue whistle* par le beatboxeur Colaps : <https://youtu.be/3orPSrOMe7o>. Colaps indique que les paramètres articulatoires pour produire le *Tongue Whistle* sont ceux des fricatives [s ʃ]. Le sifflement dépend de la géométrie de la constriction et de la vélocité du flux (le débit d'air) afin de satisfaire les conditions aérodynamiques nécessaires à générer des vortex périodiques et des résonances dans la cavité antérieure. L'aire de la constriction



est calculée par le rapport du débit  $U$  sur la vitesse des particules  $v$ . L'aire de la constriction est donc égale à  $165/713 = 0.2\text{cm}^2$ . Stevens (1989) indique qu'une aire de cette dimension résulte en un spectre d'amplitude maximale.

En regardant la configuration laryngale du sifflement glottal, on s'aperçoit que la configuration est assez différente des autres constriction glottales de type [h]. Alors que pour la *Cough Snare* la glotte du sujet CJ présentait une forme de « V », ici la constriction forme plutôt un « O ». Nous avons noté aussi une constriction de ce type pour le sifflement lingual du sujet CJ lorsque nous lui avons demandé comment il produisait le *tongue whistle*.

Les mécanismes sifflés peuvent être combinés avec d'autres mécanismes Figure 3.63. Il y a des indices de sifflement assez nets aux environs de 0.5s, 0.75s et 1.75s. Ceci démontre que le mécanisme de sifflement lingual peut se superposer avec des trilles labiales [β]. Il ne s'agit pas de sifflement labial mais de la superposition d'un geste lingual et labial. La combinaison d'une trille et d'un sifflement partageant le même lieu d'articulation paraît être une hypothèse peu plausible étant donné les contraintes aérodynamiques des mécanismes d'oscillation et de sifflement.



(a) Production d'un « *Tongue whistle* » par le sujet CJ

(b) Sifflement Glottal du sujet CJ

FIGURE 3.62 – Mécanisme sifflé

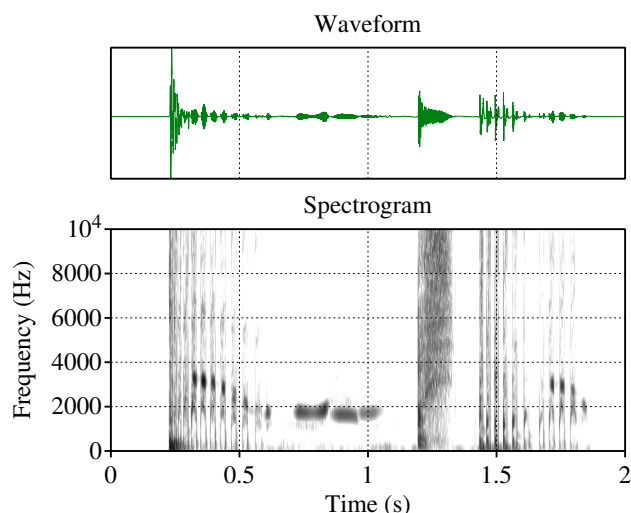


FIGURE 3.63 – Beat Pattern sifflé - sujet CJ.

⊙

## Synthèse

Tout au long de ce chapitre, nous avons tenté de démontrer que le Human Beatbox est un système combinatoire discret et complexe. Nous avons vu que les beatboxeurs et les beatboxeuses utilisaient les mêmes mécanismes de production qu'en parole. Toutefois, la combinaison des mécanismes dans le HBB diffèrent des combinaisons observées dans les systèmes sonores des langues du monde. Par exemple, [b̥] est produit de façon similaire aux implosives des langues du monde, à la différence que dans le cas du *Throat Kick*, le relâchement est glottal et les lèvres sont maintenues fermées.

Nos données ont montré l'utilisation des 6 mécanismes d'initiation possibles. Ces mécanismes ne sont pas spécifiques au HBB car on les trouve aussi dans la parole. L'initiation se combine ensuite avec différents mécanismes articulatoires. Ces derniers ne sont pas non plus spécifiques au HBB. Nos données montrent l'utilisation de 5 modes d'articulation : occlusif, fricatif, trille, nasal et latéral. Nous avons observé des constriction produites dans 11 lieux différents du conduit vocal : bilabial, labiodental, dental, alvéolaire, post-alvéolaire, vélaire, pharyngal, épiglottal, aryépiglottal, ventriculaire et glottal. Toutes ces combinaisons donnent lieu à des signatures acoustiques discrètes qui semblent permettre de distinguer les catégories entre elles (grosses caisses, caisses claires) et les sons d'une même catégorie (*K-Snare Inward*, *pf snare*).

Nous avons aussi montré que la production de la voix beatboxée reposait sur des stratégies phonatoires complexes combinant des éléments périodiques réguliers et irréguliers, des éléments aperiodiques (friction) et des éléments périodiques déphasés (double voix). Les stratégies phonatoires sont complexes en ce qu'elles sont produites par l'action coordonnées de l'ensemble des structures de l'articulateur laryngé. Les stratégies phonatoires s'accompagnent de mouvements linguaux et de l'ouverture du port vélo-pharyngé qui ont un effet sur la pression intra-orale, le débit d'air oral et nasal ainsi que sur la signature acoustique. Enfin, nous avons montré l'utilisation de mécanismes sifflés et que ces derniers pouvaient se superposer à d'autres mécanismes (trilles par ex.).

Tous ces éléments confirment que le Human Beatbox présente les caractéristiques d'un sys-

tème combinatoire discret et complexe. Notre hypothèse initiale est validée en ce qui concerne le caractère discret des imitations de percussions. Toutefois, les différences individuelles observées lors de la production de la *Throat Bass* pose la question du caractère discret des imitations de basses électroniques. Une étude, à la fois en production et en perception, de cette catégorie de sons permettrait de répondre à la question du caractère discret ou idiosyncratique des basses beatboxées.

Les primitives ne semblent pas différer selon qu'on parle ou qu'on beatboxe. Ce qui diffère, ce sont les possibilités de combinaison plus nombreuses dans le Beatbox que dans la parole. L'activité de production vocale est donc organisée différemment. Nous allons à présent voir comment varie la production en fonction du tempo de référence.

## Chapitre 4

# Des sons aux *Beat Patterns* : effet du tempo sur la production

« [La] répétition a pour effet de détruire la réflexion et de livrer l'organisme à ses tendances naturelles. »

---

L'Abbé Rousselot <sup>1</sup>

Le chapitre 4 a pour objectif de décrire, d'analyser et d'expliquer les effets du tempo de référence donné par le métronome (90, 120, 150 BPM) sur la production du Human Beatbox. Le paradigme d'accélération de la vitesse de production a été choisi pour induire de la variation des sons dans des séquences beatboxées. Ainsi, on pourra analyser la variabilité des sons beatboxés et plus généralement la variabilité de la production chez une population de sujets experts et entraînés. Nous avons fait l'hypothèse que l'augmentation de la vitesse de production entraînerait une diminution de la durée des gestes de production ainsi qu'une diminution de l'amplitude du geste d'initiation. En l'absence de données articulatoires à différents tempo, nous n'avons pas fait d'hypothèse sur l'amplitude des gestes articulatoires <sup>2</sup>. La production du HBB est analysée à travers plusieurs types de mesures phonétiques. Les mesures temporelles visent à analyser la durée des structures beatboxées, les *Beats Patterns*, ainsi que la durée des sons (durée des gestes de production). Les mesures aérodynamiques permettront d'étudier les effets du tempo sur la pression et sur l'aire relative de la constriction. Enfin, nous utiliserons les mesures des moments spectraux pour étudier leur variation en fonction du tempo. Les sons de ce chapitre ont déjà fait l'objet d'une analyse dans le Chapitre 3.

Avec les analyses temporelles, aérodynamiques et acoustiques, nous tenterons d'identifier comment des facteurs comme les caractéristiques intrinsèques des sons et des facteurs externes (position métrique, type de BP) influent sur les effets du tempo. Notre corpus présente une importante diversité de sons dont les caractéristiques phonétiques sont systématiquement différentes. La difficulté majeure de ce chapitre est l'analyse des effets du tempo sur les sons d'un système encore très peu étudié et compris. Le défi de ce travail est de démêler les facteurs qui influencent la variabilité de production des sons. Afin d'identifier ces facteurs, nous essaierons, dans la mesure du possible, de mener nos analyses sur des sous-ensembles contrôlés afin de rendre les analyses comparables.

---

1. Rousselot, P.-J. (1922). *La phonétique expérimentale : Leçon d'ouverture du cours professé au Collège de France*. Paris : Boivin.

2. Pour rappel l'amplitude d'un geste se rapporte à sa trajectoire plus ou moins longue (Ostry & Munhall, 1985)

Ce chapitre se compose de 4 parties. La première partie abordera l'analyse de la tâche de production en présentant des résultats généraux concernant les variations de vitesse de production des sujets ainsi que le taux d'erreurs de production. La deuxième partie présentera les résultats des effets du tempo sur la durée des gestes de production. La troisième partie analysera la variation des paramètres aérodynamiques en fonction du tempo de référence. La quatrième portera sur l'analyse acoustique des effets du tempo sur la production du HBB.

## 4.1 Analyse de la tâche de production

Pour introduire ce chapitre, nous souhaitons analyser rapidement la tâche de production. Nous commencerons par vérifier que la durée des *Beat Patterns* se réduise bien en fonction du tempo de référence. Nous évaluerons les différences selon les sujets et les patterns. Enfin, nous analyserons les erreurs de production.

### 4.1.1 Durée des patterns et vitesse de production

La Figure 4.1 illustre la durée des *Beats Patterns* en fonction du tempo de référence pour les sujets VP, AI, CJ et GA. À l'exception du sujet AI, les BP se réduisent de façon similaire. Chez le sujet AI, on observe des différences et une variabilité plus importante à mesure que le tempo accélère. On note que les BP les moins réduits par le sujet AI sont les BP {p ts pf: ts p p ts pf: ts} et [t' ts' t̃': ts' t' t' ts' t̃': ts']. De plus, le BP [↓p↓B<sup>1</sup>: ts' ʔh: ts' ↓p ↓p↓B<sup>1</sup>: ts' ʔh: ts'] suit la même tendance à 120BPM que les patterns précédents mais se réduit fortement à 150BPM. Une durée plus longue de ces BP implique que la vitesse de production est plus lente que le tempo de référence. En convertissant la durée des patterns en battements par minute, on peut estimer la vitesse de production des sujets. Les sujets beatboxent-ils aussi rapidement que le tempo de référence donné par le métronome ?

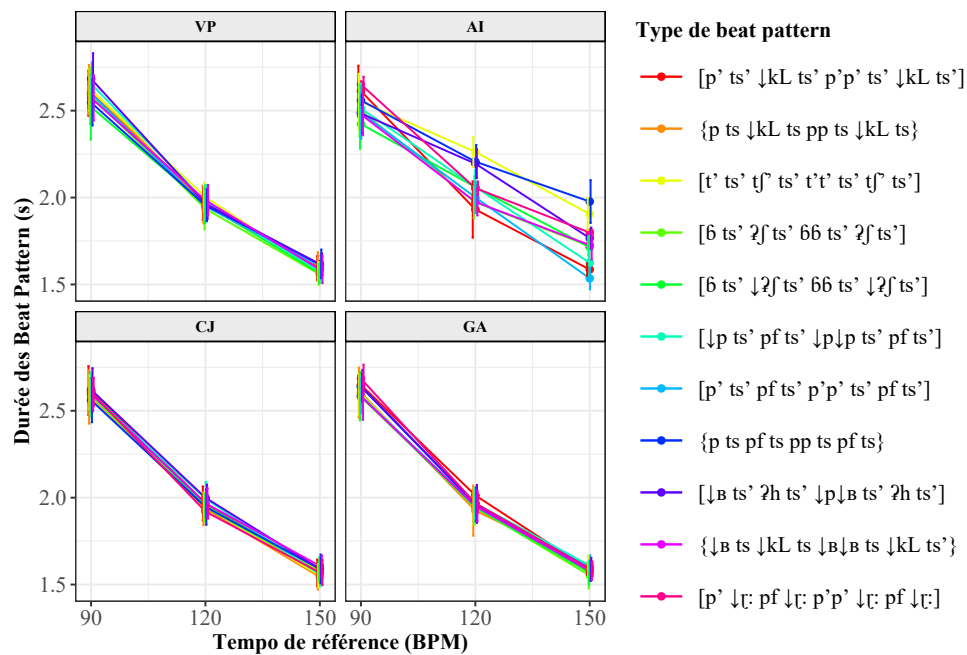


FIGURE 4.1 – Durée des BP en fonction du tempo de référence et du type de pattern (n= 483 BP).

Sur la Figure 4.2b, on a représenté la vitesse de production en fonction du tempo de référence. Chaque couleur correspond à un sujet. Les pointillés noirs indiquent le tempo de référence. On note que la vitesse de production est variable selon les sujets. Les sujets GA et VP ont une vitesse de production proche du tempo de référence. À l'inverse, le sujet AI a une vitesse de production plus éloignée du tempo de référence. On note que les vitesses de production sont parfois plus rapides, parfois plus lentes que le tempo de référence. Enfin, on note la présence de « données aberrantes » (points noirs) systématiquement plus rapides que le tempo de référence.

Si l'on regarde la Figure 4.2a, ces données n'ont rien d'aberrant, il s'agit en fait d'une accélération de la vitesse de production lors de la quatrième répétition des patterns<sup>3</sup> à l'exception du sujet AI à 150BPM qui semble ralentir.

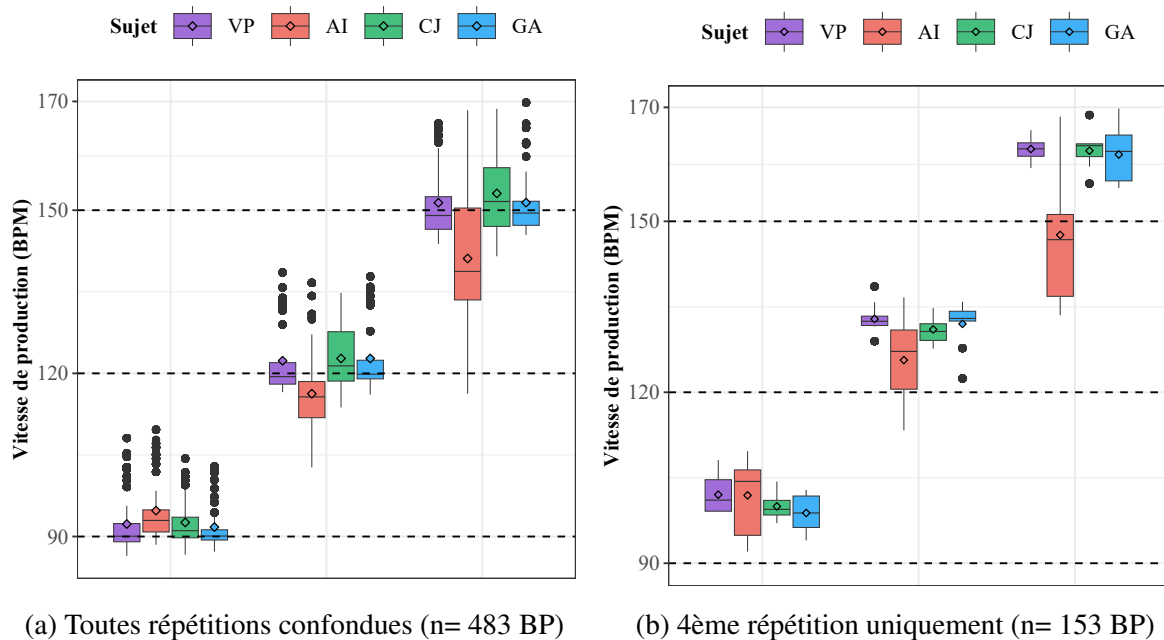


FIGURE 4.2 – Vitesse de production (BPM) en fonction du tempo de référence (BPM) pour chaque sujet. Les pointillés sur l'axe de la vitesse de production indiquent le tempo de référence.

#### 4.1.2 Erreurs de production

La tâche de production était une tâche de répétition. En effet, les sujets devaient répéter un exemple audio à différent tempo. Lors de la segmentation, nous avons utilisé une *tier* d'annotation des cibles de chaque pattern qu'il fallait produire. Dans une autre *tier*, nous avons annoté les sons tels qu'ils avaient été produits. On annotait quand un sujet produisait un autre son que la cible ( $[p'] \rightarrow \{p\}$ ). Ce type d'erreur de production regroupe les diverses transformations phonétiques illustrées plus bas. Deux autres types d'erreurs ont été annotées : les sons produits avec distorsion (notés « err. »), rendant l'identification, des sons impossibles et les élisions (sons non produits, notés  $\emptyset$ ). Ces trois types d'erreurs (transformation, distorsion et élision) sont définis comme des erreurs de production et ont été exclues de nos analyses<sup>4</sup>. Les sons fortement

3. Pour rappel, chaque pattern a été répété 4 fois à chaque tempo de référence.

4. Le terme « erreur » n'est en aucune façon un jugement concernant la technique, le style ou les stratégies de production des sujets.

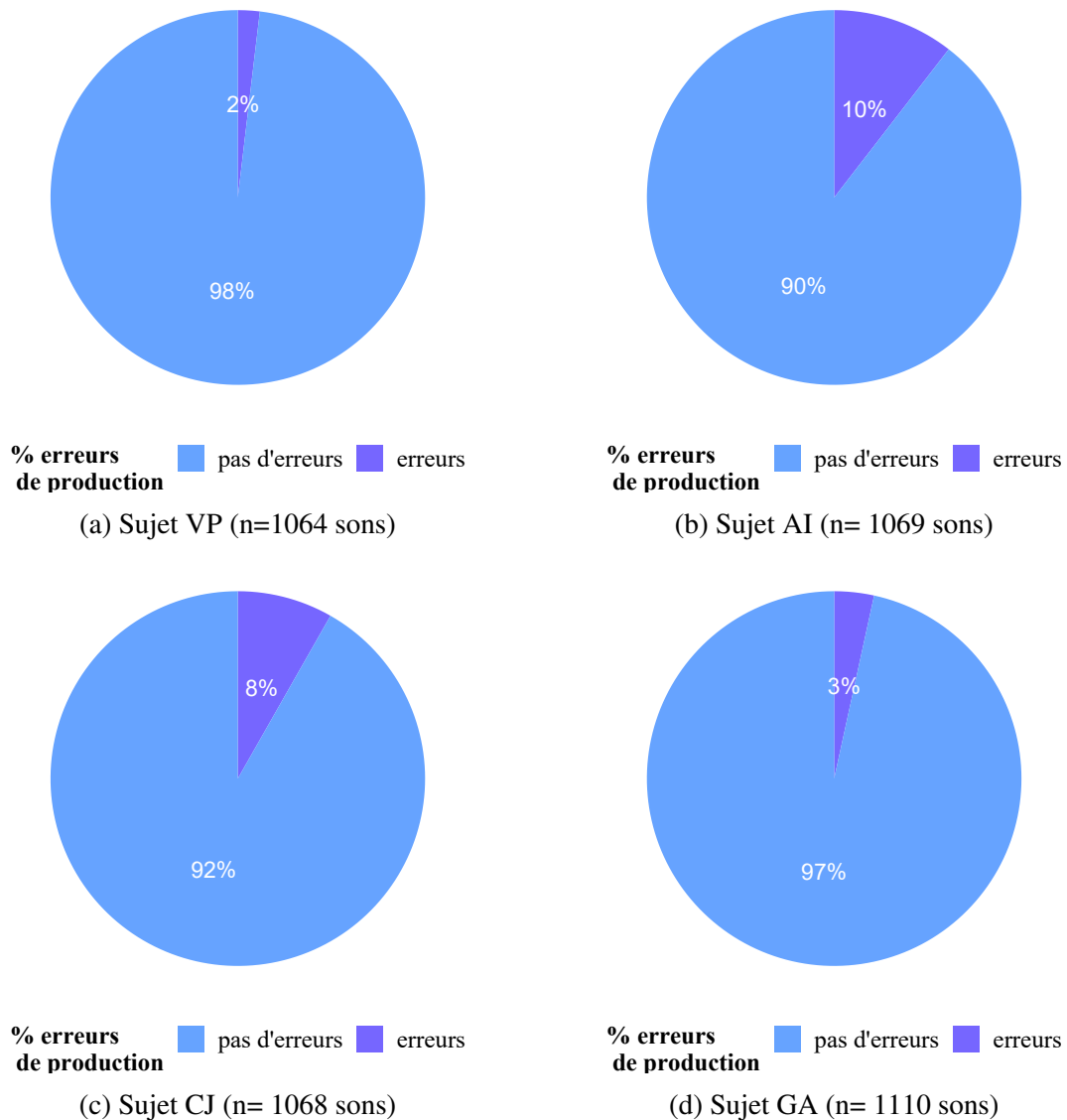


FIGURE 4.3 – Diagramme en camembert du taux d’erreurs de production des sujets.

réduits ( $[\widehat{ts}'] \rightarrow [\widehat{t}^s'] \rightarrow [t']$ ) n’ont pas été classés parmi les erreurs de production et ont été inclus dans les analyses.

La Figure 4.3 présente des diagrammes en « camembert » du taux d’erreurs de production pour chaque sujet. Le taux d’erreur est relativement faible, en effet, le corpus a été produit correctement à au moins 90%. On note que les sujets VP et GA ont le taux d’erreurs le plus faible. À l’inverse, les sujets AI et CJ ont un taux d’erreurs supérieur. Quels sont les types d’erreurs de production ?

La Figure 4.4 présente les matrices d’erreurs de production. Ces matrices s’inspirent directement des matrices de confusion utilisées dans les études de perception. Nous les reprenons pour les appliquer aux erreurs de production. Chaque ligne correspond à une cible et chaque colonne correspond aux sons produits. Les cases chiffrées indiquent le nombre d’occurrences pour lesquelles une cible a été produite différemment. Les matrices présentées plus bas ne sont pas directement comparables car elles représentent uniquement les 2 à 10% d’erreurs de production qui ne sont pas les mêmes pour chaque sujet. La description des sons du corpus est disponible dans le Tableau 2.2 du Chapitre 2.

Le sujet VP a un taux d’erreurs de production faible (2%). On note seulement 9 cas d’éli-sions de *Lips Roll* [ $\downarrow_{\text{B}}^{\text{p}}$ ] et un cas d’élision d’une *Tish Snare* [ $\text{t}^{\text{h}}$ ]. Les erreurs les plus communes sont des transformations notamment de charlestons humming égressifs en charlestons ingres-sifs  $\{\widehat{\text{ts}}\} \rightarrow \{\downarrow\widehat{\text{ts}}\}$ , des *Classic Kick Inward* [ $\downarrow\text{p}$ ]  $\rightarrow$  [ $\downarrow_{\text{B}}^{\text{p}}$ ] et de façon plus surprenante, la substitution du charleston  $\{\widehat{\text{ts}}\}$  par une grosse caisse  $\{\text{p}\}$ . La transformation de [ $\downarrow\text{p}$ ]  $\rightarrow$  [ $\downarrow_{\text{B}}^{\text{p}}$ ] suggère une réduction de la tension des lèvres pendant l’occlusion. En effet, une réduction de la tension des lèvres augmenterait les chances de générer une oscillation sous l’effet de Bernoulli.

Le sujet AI a seulement fait 2 omissions. L’erreur la plus fréquente est la transformation de mécanismes d’initiation linguaux par des mécanismes laryngaux  $\{\text{p}\}$ ,  $\{\widehat{\text{ts}}\}$ ,  $\{\downarrow\widehat{\text{KL}}\} \rightarrow [\text{p}^{\text{h}}]$ ,  $[\text{ts}^{\text{h}}]$ , [ $\downarrow\widehat{\text{KL}}$ ]. Similairement, on observe qu’elle transforme [ $\downarrow\text{p}$ ]  $\rightarrow$  [ $\text{p}^{\text{h}}$ ].

Le sujet CJ a omis 5 *clic Rolls* [ $\downarrow\text{r}$ ]. De même que pour le sujet AI, l’erreur la plus commune est la transformation de  $\{\widehat{\text{ts}}\} \rightarrow [\text{ts}^{\text{h}}]$  et de [ $\downarrow\text{p}$ ]  $\rightarrow$  [ $\text{p}^{\text{h}}$ ].

Le sujet GA a omis 1 charleston [ $\widehat{\text{ts}}^{\text{h}}$ ]. De même que le sujet VP, l’erreur la plus commune est la transformation de  $\{\widehat{\text{ts}}\} \rightarrow \{\downarrow\widehat{\text{ts}}\}$ . On retrouve la transformation de [ $\downarrow\text{p}$ ]  $\rightarrow$  [ $\text{p}^{\text{h}}$ ], générale-ment dans les mêmes contextes que CJ, à savoir avant le *Lips Roll* en position 5. Enfin, on note que  $[\text{ts}^{\text{h}}] \rightarrow [\text{ts}^{\text{h}}:]$  en position 9 dans  $[\text{p}^{\text{h}} \widehat{\text{ts}}^{\text{h}} \downarrow\widehat{\text{KL}}: \widehat{\text{ts}}^{\text{h}} \text{p}^{\text{h}} \text{p}^{\text{h}} \widehat{\text{ts}}^{\text{h}} \downarrow\widehat{\text{KL}}: \widehat{\text{ts}}^{\text{h}}]$ .

Il y a des similitudes entre les erreurs de production chez les différents sujets. Ces similitudes concernent principalement des substitutions de mécanismes d’initiation linguaux par des méca-nismes laryngaux ( $\{\text{p}\} \rightarrow [\text{p}^{\text{h}}]$ ). Ces choix, aussi involontaires qu’ils puissent être, dénotent une préférence pour les mécanismes laryngaux. En discutant avec le sujet VP, il est ressorti que pour beatboxer rapidement, il était préférable d’utiliser les « variantes non-humming ». Cela peut suggérer l’existence de contraintes de coordination selon les mécanismes d’initiation.

Pour terminer, nous mentionnerons la transformation de  $[\text{p}^{\text{h}}] \rightarrow [\text{p}^{\text{h}}\phi]$  des sujets CJ et GA. C’est un son que nous avons déjà rencontré dans une autre étude (Dehais-Underdown et al., 2019). Dans cette autre étude, nous avons montré que la partie fricative était pulmonique égres-sive. Ici, il est difficile de se prononcer sur le caractère pulmonique ou non-pulmonique de la fricative bilabiale en l’absence d’occurrence de cette variante dans nos données fibroscopiques. Dans l’hypothèse qu’elle soit pulmonique comme dans Dehais-Underdown et al. (2019), est-ce que la friction fait partie du son, dans quel cas il s’agirait d’une fonction stylistique, ou est-ce que la phase de friction possède une fonction respiratoire ? Ces deux hypothèses ne sont pas contradictoires. Les beatboxeurs peuvent produire cette fricative dans un but à la fois stylistique et respiratoire pour permettre l’illusion d’une production ininterrompue.







On note aussi des différences individuelles pour un même son. Prenons le cas de la *Cough Snare* [ʔh:]. Les productions de AI et CJ sont plus longues et se réduisent plus fortement ( $\pm 50$ ms) que celles des sujets VP et GA qui sont plus courtes et semblent moins affectées par les variations de tempo. Notons aussi les différences individuelles concernant la réduction temporelle de la *Cough Snare*. Chez les sujets AI et CJ on observe deux trajectoires temporelles différentes : AI maintient une durée stable de 90 à 120 BPM puis réduit fortement la durée du son à 150BPM; alors que CJ réduit la durée de 90 à 120BPM puis maintient une durée stable entre 120 et 150BPM. VP et GA montrent la même trajectoire entre 90 et 120 BPM mais à 150BPM on note une réduction de la durée chez GA et une augmentation de la durée chez VP. Les stratégies de réduction temporelles sont donc individuelles. Qu'en est-il de la durée des sons, c'est-à-dire lorsque l'occlusion a été mesurée ?

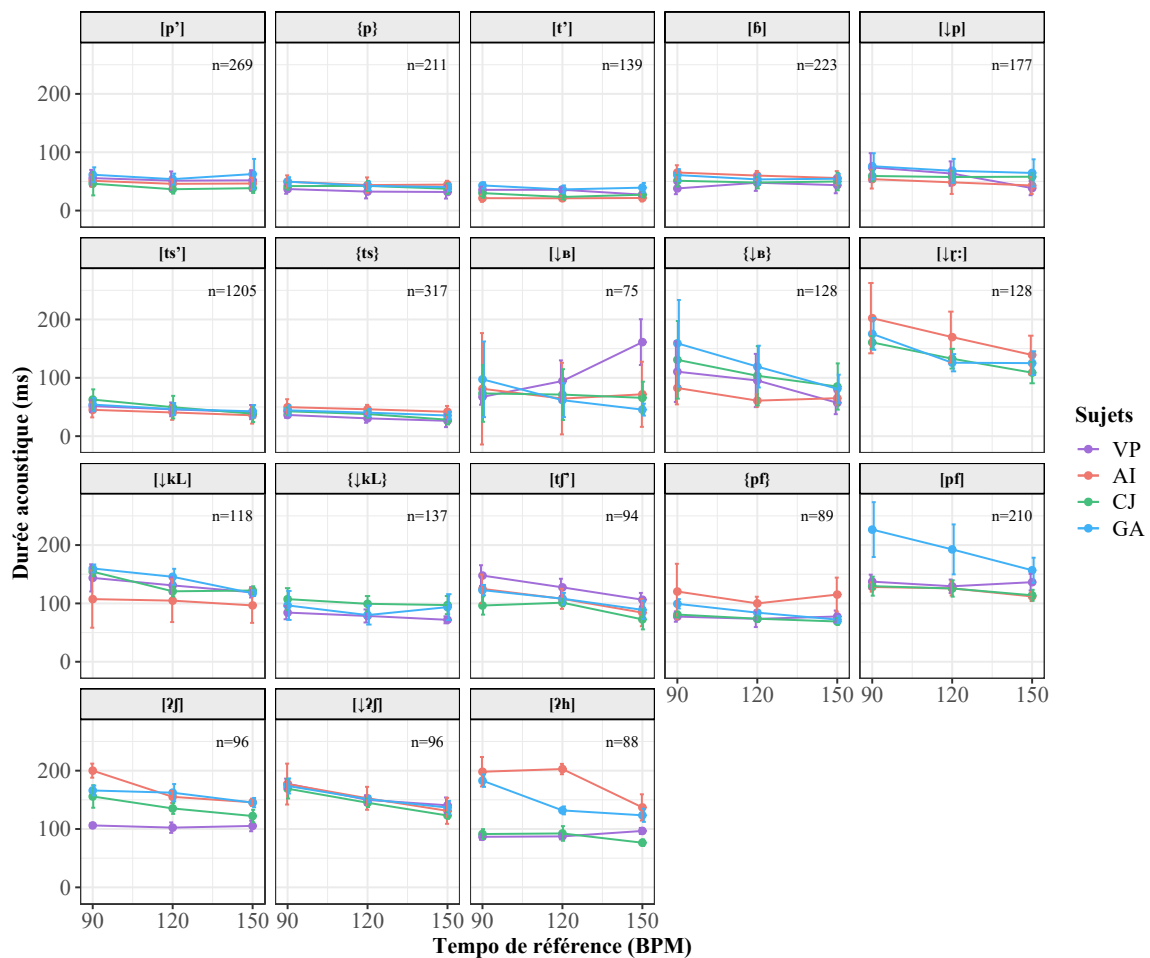


FIGURE 4.5 – durée acoustique des sons (ms) en fonction du tempo de référence (BPM) pour tous les sons du corpus. Chaque point correspond à la moyenne de toutes les répétitions d'un son. Les barres d'erreur indiquent l'écart-type de la durée.

La Figure 4.6 illustre la durée des sons dont l'occlusion a été mesurée en fonction du tempo de référence. Les sons « \_ » et « - » correspondent à l'intervalle temporel entre les sons dont l'occlusion a été mesurée. On observe encore une fois des différences selon les sons et les sujets. On observe que les caisses claires et le charleston [tsʰ] se réduisent davantage en comparaison aux grosses caisses. On note que la durée des sons tend à se réduire. Cela suggère donc que les gestes d'occlusion réduisent davantage leur durée que les gestes de relâchement (reflétée par la durée acoustique). Concernant la durée des silences entre les sons, « - » réfère au silence mesuré

entre les deux grosses caisses en positions 5 et 6. Il est plus court que dans le reste des positions et tend à moins réduire sa durée, à l'exception du sujet VP. La réduction temporelle des *Beat Patterns* ne s'explique pas uniquement par la réduction temporelle des gestes de production mais aussi par la réduction de l'intervalle de temps qui sépare les gestes de production.

Nous venons de voir que les effets du tempo diffèrent selon les sons. Nous savons que lors de la production de la parole, les caractéristiques articulatoires influent sur le degré de réduction temporelle. Qu'en est-il pour le HBB ? Est-ce que les caractéristiques initiatrices et articulatoires des sons influent sur la réduction temporelle des sons ?

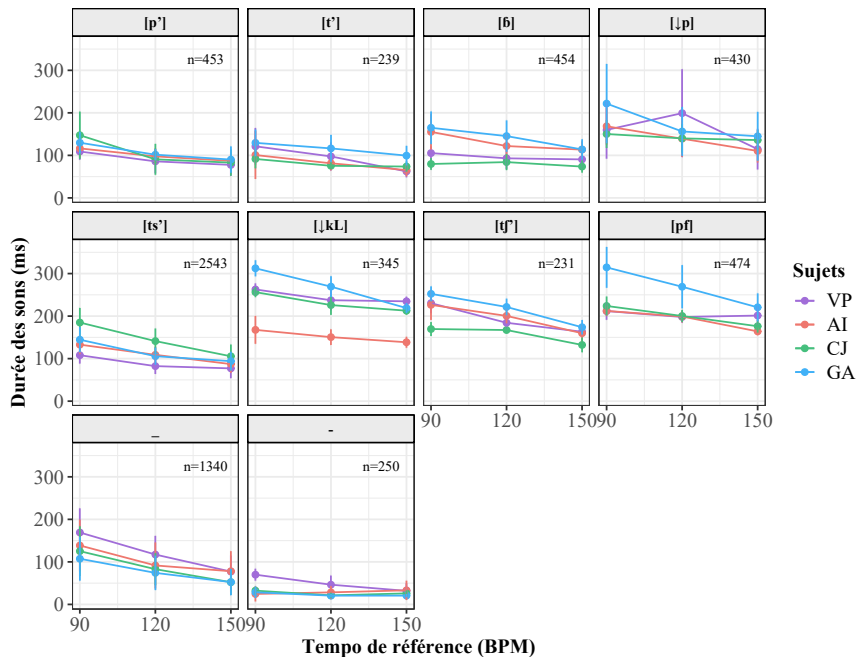


FIGURE 4.6 – Durée des sons (ms) en fonction du tempo de référence (BPM). Les sons « \_ » et « - » correspondent à l'intervalle entre les sons dont l'occlusion a été mesurée. Chaque point correspond à la moyenne de toutes les répétitions d'un son. Les barres d'erreur indiquent l'écart-type de la durée.

#### 4.2.2 Effet du tempo selon les caractéristiques initiatrices des sons

La Figure 4.7 illustre la durée acoustique des sons en fonction du tempo de référence. Afin de rendre les analyses comparables, nous avons fait le choix d'analyser uniquement les sons affriqués. À notre connaissance, aucune étude ne s'est intéressée aux effets de la vitesse de production sur les paramètres d'initiation. Nous nous demandons donc si les mécanismes d'initiation influent sur la réduction des sons beatboxés.

La durée acoustique des sons produits avec des mécanismes pulmoniques est plus longue que la durée des sons produits avec des mécanismes laryngaux et linguaux. Nous n'observons pas de différences majeures de durée entre mécanismes laryngaux et linguaux. Parmi les affriquées produites avec un geste pulmonique, nous observons des différences selon le mode d'initiation. En effet, les sons produits par compression pulmonique ([tʃ]) suggèrent un comportement plus variable concernant leur réduction temporelle. Le sujet VP augmente la durée acoustique moyenne alors que les autres sujets la réduisent. À l'inverse, les sons produits par expansion pulmonique ([ɹkL]) indiquent une trajectoire de réduction plutôt linéaire et similaire entre les différents sujets.

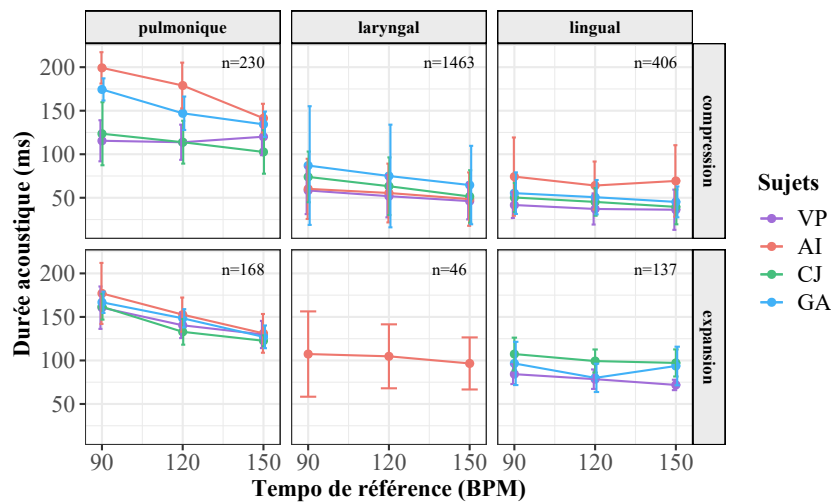


FIGURE 4.7 – durée acoustique des sons (ms) en fonction du tempo de référence (BPM) selon les lieux et modes d’initiation. Seuls les sons affriqués sont illustrés ici. Chaque point correspond à la moyenne de toutes les répétitions d’un son. Les barres d’erreur indiquent l’écart-type de la durée.

Nous observons une variation importante de la durée acoustique des sons produits par compression laryngale ( $[\widehat{tʃ}ʔ]$ ), en particulier chez le sujet GA pour qui l’écart-type, indiqué par les barres d’erreur, est à peu près deux fois supérieur aux autres sujets. Le degré de réduction est plus faible chez les sujets VP et AI en comparaison aux sujets CJ et GA. Nous n’avons qu’une seule affriquée produite par expansion laryngale : la *K-Snare Inward* [ $\downarrow k_L$ ] du sujet AI. Pour ce son, elle réduit modérément la durée moyenne ainsi que l’écart-type. Toutefois, la durée acoustique du son est très variable.

Enfin, des différences peuvent être observées entre les variantes linguales par compression et expansion. En effet, les sons produits par compression linguale ( $\{\widehat{ts}\}$ ) ont un relâchement qui dure en moyenne 50 à 75ms. Les sons produits par expansion buccale ( $\{\downarrow k_L\}$ ) ont une durée moyenne de relâchement qui varie de 75ms à 100ms selon les sujets. Les trajectoires de réduction temporelle n’indiquent pas de tendance commune à tous les sujets.

### 4.2.3 Effet du tempo selon les caractéristiques articulatoires des sons

Contrairement aux effets du tempo sur l’initiation, nous savons que les mécanismes articulatoires influencent la réduction des sons parlés. En effet sur la Figure 4.8, nous pouvons voir, sans surprise, que les sons affriqués sont plus longs que les sons occlusifs. Cependant, nous n’observons pas d’effet du mode d’articulation sur la réduction. En effet, les comportements individuels en ce qui concerne les affriquées ne laissent pas entrevoir de tendance commune à tous les sujets.

Il y a une différence de durée acoustique entre les occlusives bilabiales et coronales, ces dernières étant plus courtes. Cependant, il ne semble pas y avoir d’effet du lieu d’articulation sur la durée acoustique des occlusives. Concernant les affriquées laryngales, elles sont toutes pulmoniques et les différences observées sont dues aux mécanismes d’initiation plutôt qu’aux mécanismes d’articulation. Nous regroupons les sons dentaux ( $[\widehat{ts}ʔ]$ ), alvéolaires ( $[\widehat{t}ʔ]$ ) et post-alvéolaires ( $[\widehat{tʃ}ʔ]$ ) sous le terme « coronal » pour indiquer les sons produits par contact de la pointe ou la lame de la langue dans les trois zones précédemment citées. Les sons affriqués « coronaux » ont une durée acoustique plus courte en comparaison avec les autres lieux d’ar-

ticulation. De plus, ces sons tendent à se réduire modérément de façon linéaire pour tous les sujets. Les sons labiodentaux et vélaire indiquent une durée stable voir légèrement réduite, à l'exception du sujet GA. Les données de ce sujet indiquent qu'il réduit fortement la durée des labiodentales en fonction du tempo de référence. Pour les vélaire, on note que la durée moyenne est la même à 90 et 150 BPM mais elle est réduite à 120BPM. Les effets du tempo sur la durée acoustique dépendent en partie des caractéristiques phonétiques des sons. Est-ce que les effets du tempo sur les durées dépendent en partie de facteurs externes aux sons comme par exemple la position dans la métrique ou le contexte phonétique adjacent ?

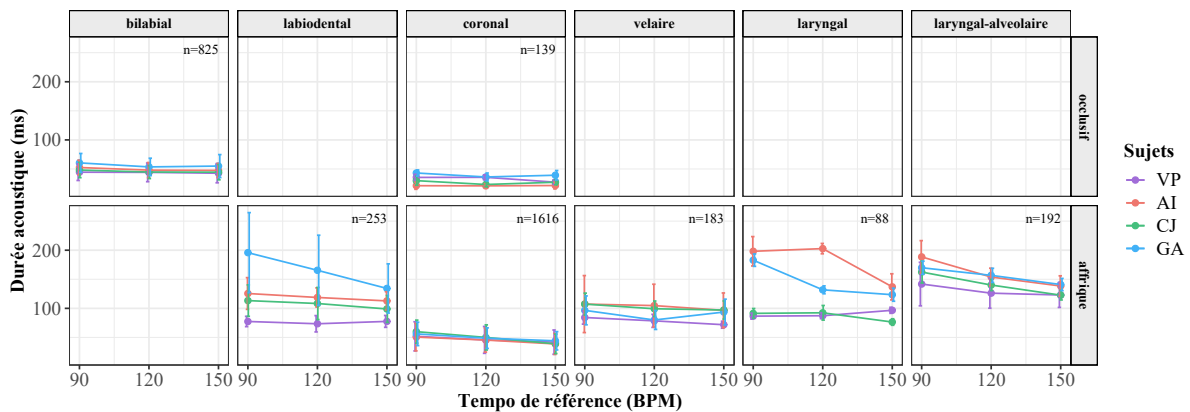


FIGURE 4.8 – durée acoustique des sons (ms) en fonction du tempo de référence (BPM) selon les lieux et modes d'initiation. Seuls les sons affriqués sont illustrés ici. Chaque point correspond à la moyenne de toutes les répétitions d'un son. Les barres d'erreur indiquent l'écart-type de la durée.

#### 4.2.4 Effet du tempo selon la position des sons dans les patterns

Nous avons vu que l'initiation et l'articulation pouvaient influencer la réduction temporelle des sons beatboxés. La position des sons dans la métrique est susceptible d'influer sur le degré de réduction.

Chaque catégorie de sons apparaît dans des positions précises dans la structure métrique des *Beat Patterns*. Les grosses caisses et le *Lips Roll Humming* ( $\{\downarrow\mathbb{B}^1\}$ ) apparaissent dans les positions 1, 5 et 6. Le *Lips Roll* [ $\downarrow\mathbb{B}^1$ ] apparaît seulement en position 1 et 6. Dans ce pattern la position 5 est occupée par le *Classic Kick Inward* [ $\downarrow p$ ]. Les charlestons apparaissent dans les positions 2, 4, 7 et 9 et les caisses claires dans les positions 3 et 8. Pour rendre nos analyses comparables, nous avons choisi d'analyser 4 patterns dont les caractéristiques articulatoires sont comparables : les BP [ $p' \widehat{ts}' \downarrow \widehat{kl} : \widehat{ts}' p' p' \widehat{ts}' \downarrow \widehat{kl} : \widehat{ts}'$ ] et  $\{p \widehat{ts} \downarrow \widehat{kl} : p p \widehat{ts} \downarrow \widehat{kl} : \widehat{ts}\}$  et les BP [ $p' \widehat{ts}' \widehat{pf} : \widehat{ts}' p' p' \widehat{ts}' \widehat{pf} : \widehat{ts}'$ ] et  $\{p \widehat{ts} \widehat{pf} : \widehat{ts} p p \widehat{ts} \widehat{pf} : \widehat{ts}\}$ . Chaque paire de patterns se compose des mêmes sons mais se différencie par le lieu d'initiation (lingual *versus* non-lingual).

La Figure 4.9 illustre les durées acoustiques des sons selon la position métrique : à gauche les patterns produits sans humming, à droite les patterns produits avec humming. Nous observons de nouveau de légères différences en fonction du lieu d'initiation. Les durées acoustiques en positions 3 et 8 sont plus longues que les autres. Ce que montrent les données, c'est que même si ces sons sont longs, cela n'implique pas forcément qu'ils se réduisent davantage. Si nous regardons la Figure 4.9a, nous remarquons qu'en position 3 les sujets VP et AI ne réduisent pas aussi fortement la durée acoustique que les sujets CJ et GA. En position 8, les tendances sont les mêmes à l'exception de VP qui semble réduire légèrement la durée.

À l'inverse, on sait que les grosses caisses et les charlestons sont plus courts que les caisses claires. On observe que dans les autres positions les durées moyennes sont plus faibles. Dans la métrique des patterns, les positions 3, 4 et 5 appartiennent au même intervalle. Si l'on regarde, à gauche, la durée moyenne à 90BPM, on observe que les durées sont légèrement plus faibles ( $t < 50ms$ ) dans les positions 4 ([ts']) et 5 ([p']) en comparaison avec les mêmes sons en position 2 pour [ts'] et 1 ou 6 pour [p']. En position 4, on observe une différence entre les deux graphiques, en particulier pour le sujet CJ. La durée acoustique de [ts'] chez le sujet CJ est en moyenne plus longue que pour les autres sujets en position 4 et 9 et se réduit fortement à 150BPM. En comparaison, les autres sujets réduisent presque pas (GA) ou peu (VP et AI) la durée acoustique. Pour les patterns humming (droite), on note moins de variation individuelle dans cette position et on note aussi une légère tendance à la réduction. Afin de confirmer ou non une possible influence de la position sur la réduction temporelle, il faudrait inclure les durées d'occlusion de tous les sons et pas seulement les durées acoustiques.

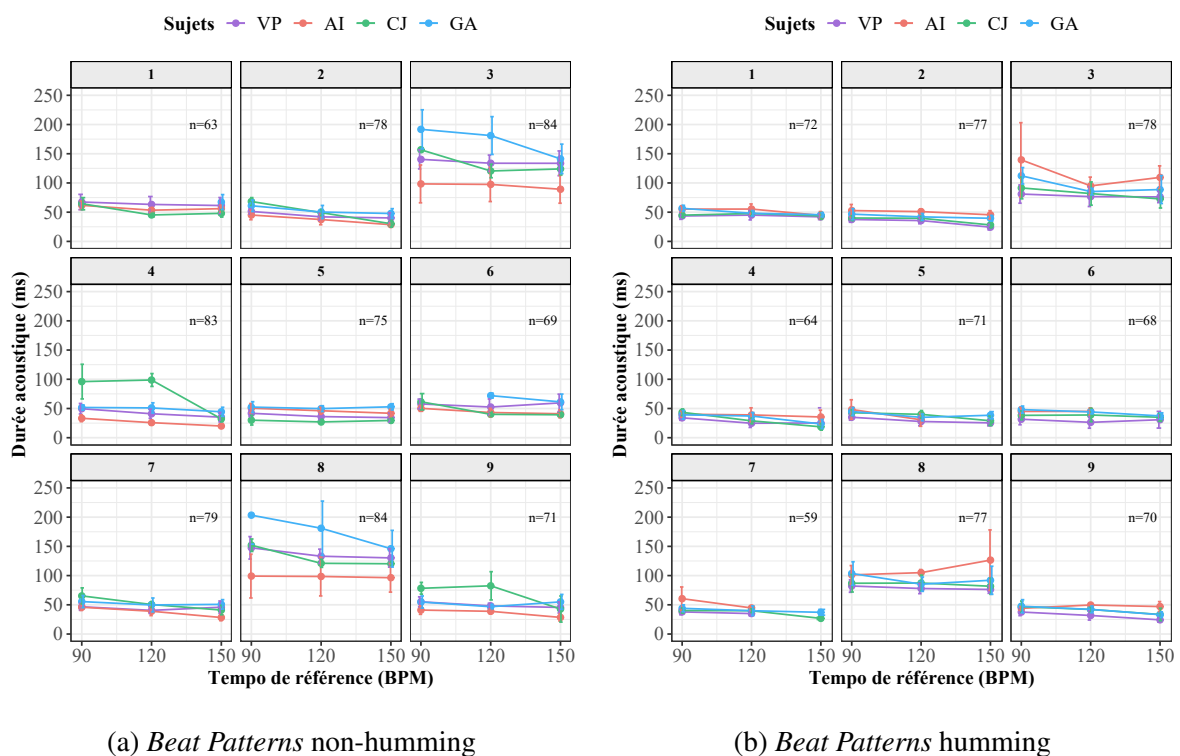


FIGURE 4.9 – durée acoustique des sons (ms) en fonction du tempo de référence (BPM) selon la position métrique. Seuls 4 *beats patterns* sont illustrés ici. Chaque point correspond à la moyenne de toutes les répétitions d'un son. Les barres d'erreur indiquent l'écart-type de la durée.

## 4.2.5 Effet du tempo selon le contexte phonétique adjacent

Pour conclure l'analyse temporelle, nous allons analyser la possible influence du contexte phonétique. Est-ce que la durée des sons et leur réduction est conditionnée par les sons qui précèdent et qui suivent? Pour répondre à cette question, nous avons décidé de porter notre analyse sur un sous-ensemble composé exclusivement de charlestons fermés éjectifs. Le son [ts'] est un son qui ne change pas entre les patterns (sauf pour les patterns humming et le pattern avec le *Clic Roll* [↓ɾ:]). Ainsi, [ts'] est en contact avec la quasi-totalité des grosses caisses, des caisses claires et du *Lips Roll* pulmonique. De plus, la durée de l'occlusion de [ts'] a été mesurée, nous analyserons donc la durée du son plutôt que la durée acoustique.

La Figure 4.10 illustre la durée de [ts'] en fonction du tempo de référence. Chaque colonne représente une position métrique, chaque ligne un sujet et chaque couleur un *Beat Pattern* différent. La première observation que nous faisons se rapporte à la variation individuelle des sujets. Y-a-t-il tout de même des tendances partagées par les sujets?

Nous pouvons observer en premier lieu que [ts'] tend à être plus court lorsqu'il est produit dans le BP [↓p↓B': ts' ʔh: ts' ↓p ↓p↓B': ts' ʔh: ts'], plus particulièrement dans les positions 2 et 7 lorsqu'il suit le *Lips Roll*. Il est possible que la production du charleston soit retardée et, qu'en plus, les sujets réduisent la durée du son afin de maintenir une durée de l'intervalle régulière. De plus, on note que le charleston tend à se réduire systématiquement dans le BP [p' ts' ↓KL: ts' p' p' ts' ↓KL: ts']. De même, dans les patterns [b' ts' ʔj: ts' b' b' ts' ʔj: ts'] et [b' ts' ↓ʔj: ts' b' b' ts' ↓ʔj: ts'], on observe une tendance similaire concernant la réduction de la durée du son.

Le « zoom » sur les données de [ts'] montre que la réduction peut avoir lieu dans toutes les positions. Cependant, certaines positions semblent influencer différemment sur la réduction temporelle. En regardant attentivement la Figure 4.10 on peut par exemple noter des comportements similaires dans les positions 2 et 7 et des comportements similaires dans les positions 4 et 9. C'est par exemple le cas chez les sujets VP et CJ pour qui on observe une tendance à la réduction temporelle de [ts'] dans les positions 4 et 9 quel que soit le pattern. À l'inverse dans les positions 2 et 7 la réduction dépend du pattern dans lequel [ts'] est produit. Par exemple, chez le sujet GA la durée de [ts'] augmente en position 2 dans le BP [↓p ts' pf: ts' ↓p ↓p ts' pf: ts'] mais se réduit par exemple dans le BP [b' ts' ʔj: ts' b' b' ts' ʔj: ts']. Notons qu'en position 4, les durées tendent à se réduire quasi-systématiquement. Dans les positions 2 et 7 le son précédent est toujours une grosse caisse et le son suivant une caisse claire alors qu'en position 4 et 9 c'est l'inverse.



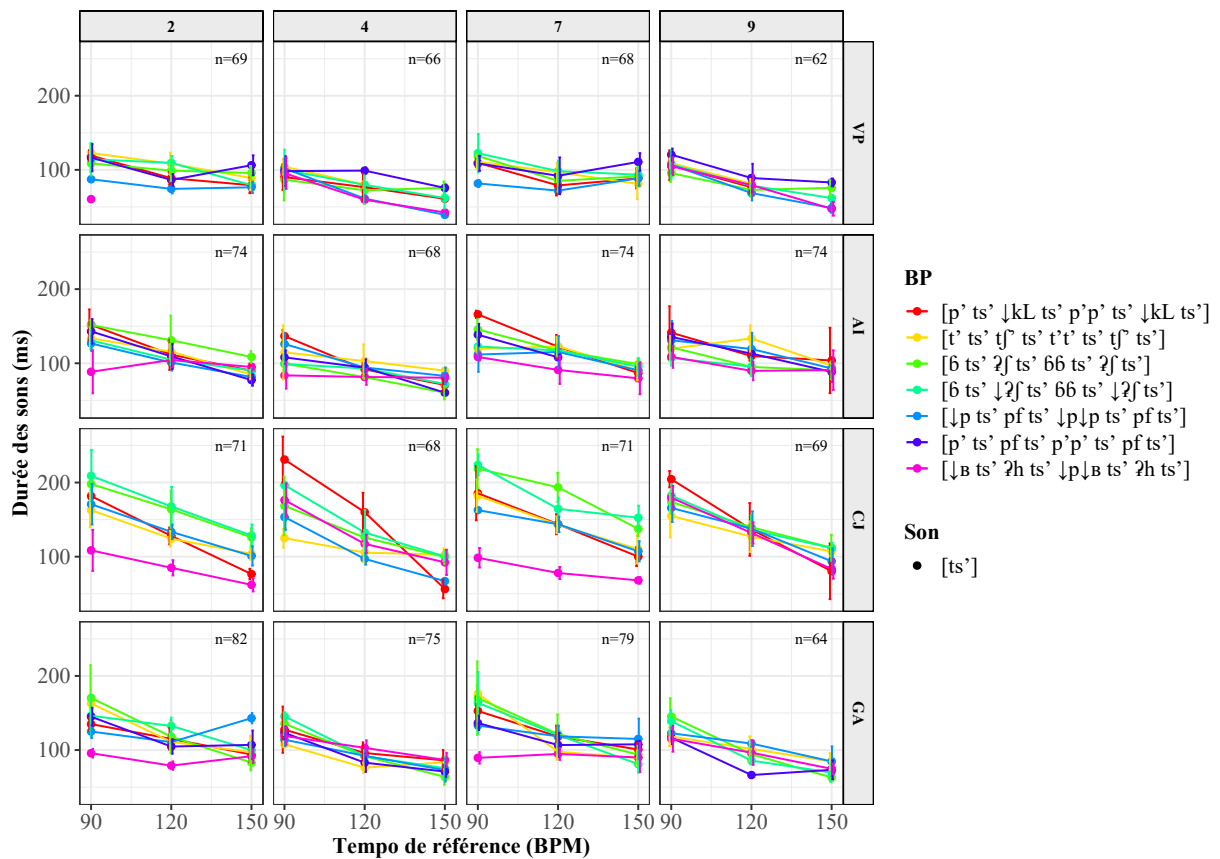


FIGURE 4.10 – Durée des sons (ms) en fonction du tempo de référence (BPM) selon la position métrique. Seul le son [ts'] est illustré ici.

#### 4.2.6 Synthèse de l'analyse temporelle

Pour résumer, nous avons montré que les effets du tempo différaient selon les sujets, les sons, la position métrique et le contexte adjacent. Nous avons montré qu'il y avait une tendance des sons et des intervalles intergestuels à réduire leur durée. Nous validons une partie de notre hypothèse initiale qui stipulait que les sons réduisaient leur durée en fonction du tempo de référence. La durée des sons ainsi que leur réduction temporelle dépend du type de mécanisme d'initiation utilisé. En effet, le lieu d'initiation (pulmonique, laryngal et lingual) influence la réduction des sons, en particulier lorsque le mécanisme d'initiation a lieu dans les poumons. De même, le mode d'initiation influence le degré de réduction comme nous l'avons montré pour les mécanismes pulmoniques et linguaux. Il y a un effet du mode d'articulation sur la durée et la réduction temporelle des sons. Sans surprise les affriquées (et les trilles illustrées sur la Figure 4.5) sont plus longues et plus propices à se réduire temporellement en comparaison aux occlusives. Concernant les différents lieux d'articulation, on a observé que les sons coronaux sont plus courts et se réduisent de façon similaires et linéaires chez les différents sujets. L'analyse du rôle de la position métrique dans la réduction de la durée acoustique des sons n'a pas révélé d'influence majeure d'une ou plusieurs positions sur la réduction temporelle. Cependant, notre analyse porte sur seulement un sous-ensemble de sons du corpus (6/20 sons, 4/11 BP) où seules les durées acoustiques ont été analysées. Outre les importantes différences individuelles, nous avons trouvé que d'une part, certains BP ( $[\downarrow p \downarrow B^1: \widehat{ts'} \text{ ʒh: } \widehat{ts'} \downarrow p \downarrow p \downarrow B^1: \widehat{ts'} \text{ ʒh: } \widehat{ts'}]$ ) influençaient la durée de [ts'] à 90BPM et que d'autre part, certains BP ( $[\widehat{p'} \widehat{ts'} \downarrow kL: \widehat{ts'} p' p' \widehat{ts'} \downarrow kL: \widehat{ts'}]$ ) ont une influence sur le degré de réduction. Nous avons aussi montré des similitudes entre

les comportements observés dans les positions 2 et 7 d'une part et les positions 4 et 9 d'autres part. Ces différences dépendent du contexte phonétique adjacent car en position 2 et 7 le son précédent est toujours une grosse caisse et le son suivant une caisse claire alors qu'en position 4 et 9 c'est l'inverse. Cette fois, l'analyse portait sur la durée des sons (occlusion comprise) et suggère qu'il y a un effet de la position sur la réduction temporelle.

### 4.3 Analyse aérodynamique

Nous cherchons à présent à connaître les effets du tempo sur la pression et l'aire de la constriction. Nous traiterons dans un premier temps les effets sur la pression puis ensuite sur l'aire de la constriction. Concernant la pression, notre corpus se compose de sons égressifs et de sons ingressifs. Il y a donc des valeurs de pression positives et négatives comprises entre +100hPa et -100hPa, soit une plage de 200hPa. Afin de réduire la plage de pression et faciliter la comparaison des sons, nous avons transformé toutes les valeurs de pression et en nombres positifs ( $P = \sqrt{-P^2}$ ) où P est la pression intraorale mesurée. Le signe négatif renseigne uniquement sur la direction du geste d'initiation, c'est-à-dire un mouvement de compression ou d'expansion. Qu'un mécanisme ait une pression de +50 hPa ou -50 hPa tous deux sont égaux à une déviation de 50hPa par rapport à la pression atmosphérique de référence. Dans un cas, la déviation est provoquée par une compression et dans l'autre par une expansion. La Figure 4.11 illustre la différence entre les valeurs positives et négatives et les valeurs rendues positives. Par exemple, pour le son [↓ʔ]: de GA, on voit sur la figure de gauche que la pression se rapproche de zéro, on dirait qu'elle augmente alors qu'un rapprochement de zéro signifie que l'expansion du volume est plus faible à 150BPM. Sur la figure de droite, le même son du même sujet suggère exactement le même comportement : une expansion pulmonique plus faible à 150BPM. Dans cette partie, par augmentation ou diminution de la pression intraorale, nous entendons une augmentation de la déviation de la pression intraorale mesurée par rapport à la pression de référence que le son soit ingressif ou égressif.

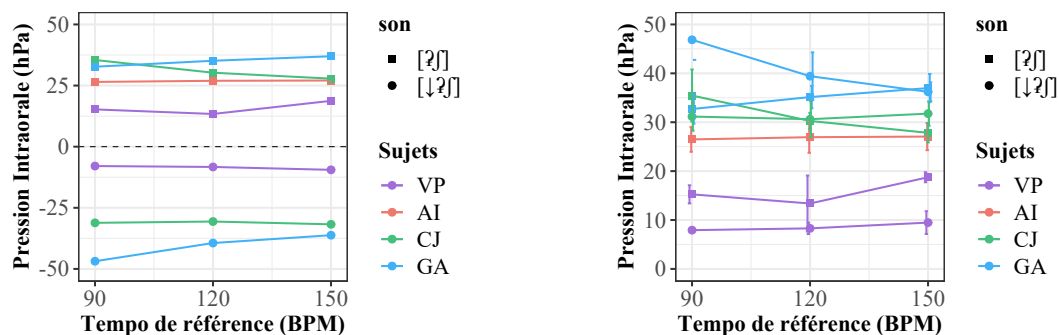


FIGURE 4.11 – Graphiques illustrant la différence entre valeurs positives et négatives (gauche) et les valeurs positives de pression où  $P = \sqrt{-P^2}$  (droite)

#### 4.3.1 Effet du tempo sur la pression

Quel est l'effet du tempo sur la pression intra-orale ? Pour répondre à cette question, il va nous falloir une fois de plus démêler les multiples facteurs susceptibles d'influencer les variations de pression. Nous évoquerons l'influence des caractéristiques initiatrices et articulatoires des sons ainsi que les effets des différents contextes sur la pression de [ts'] dans différentes positions. Les variantes humming ne seront pas analysées car la pression n'a pas été mesurée pour ces sons.

#### 4.3.1.1 Effet du tempo sur la pression selon les sons

La Figure 4.12 illustre la pression intra-orale en fonction du tempo de référence. En premier lieu, nous pouvons remarquer que les sujets CJ et GA génèrent des niveaux de pression plus importants que les sujets VP et AI.

Contrairement à l'hypothèse que nous postulions, la pression ne tend pas à se réduire. En effet, il arrive que des sujets diminuent la pression. C'est par exemple le cas pour le *Classic Kick* [p'] du sujet AI, le *Kick Lingual* [t'] des sujets VP et GA. Le sujet CJ ne tend pas à réduire la pression. Nous possédons peu de données pour le sujet AI en raison de la saturation.

Toutefois, il arrive que les sujets augmentent la pression à mesure que le tempo accélère. C'est par exemple le cas du sujet CJ pour le son [β], du sujet GA pour le *Lips Roll* [↓<sub>B</sub><sup>1</sup>], du sujet AI pour le son [t'] ou encore du sujet VP pour le son [↓<sub>2</sub><sup>1</sup>].

Dans de nombreux cas, la pression reste plus ou moins stable. Parfois, on peut noter des augmentations ou diminutions à 120BPM. Par exemple, si l'on regarde les valeurs de pression des sujets CJ et GA pour produire [↓<sub>1</sub>], on note que la pression à 90BPM est comparable à la pression à 150 BPM. Cependant à 120BPM, le sujet CJ augmente légèrement la pression alors que le sujet GA la réduit. Toutefois, les variations de pression ont lieu sur une plage réduite.

La réduction de la pression intra-orale, c'est-à-dire le fait que la pression soit de plus en plus proche de 0, implique que la variation du volume est de plus faible amplitude à 150 BPM qu'à 90 BPM. En d'autres termes, les sujets ne vont pas « jusqu'au bout » du geste d'initiation par rapport à 90 BPM. Inversement, l'augmentation de la pression intra-orale suggère que les sujets vont « au-delà » de la cible du geste d'initiation. Quelles peuvent-être les raisons de ces augmentations de pression ?

Du point de vue physiologique, cette augmentation de la pression peut refléter une augmentation de la raideur des articulateurs et/ou une augmentation de la vélocité du geste d'initiation. Par exemple, GA tend à augmenter la pression de la *Tish Snare* [t<sup>h</sup>']. On sait aussi qu'il réduit la durée du son. On peut donc penser qu'une augmentation de la raideur des articulateurs couplée à une augmentation de la vélocité du geste permettent une compression plus forte, c'est-à-dire une importante réduction du volume et donc une augmentation de la pression.

Du point de vue musical et acoustique, ces augmentations de la pression peuvent refléter des stratégies de régulation de l'intensité acoustique. En effet, la pression acoustique dépend des variations de pression dans le conduit vocal. En augmentant la pression, les sujets sont aussi susceptibles d'augmenter le niveau d'intensité. Bien entendu, cette hypothèse serait à vérifier avec des mesures d'intensité que nous ne pouvons mener en raison de problèmes de saturation acoustique.

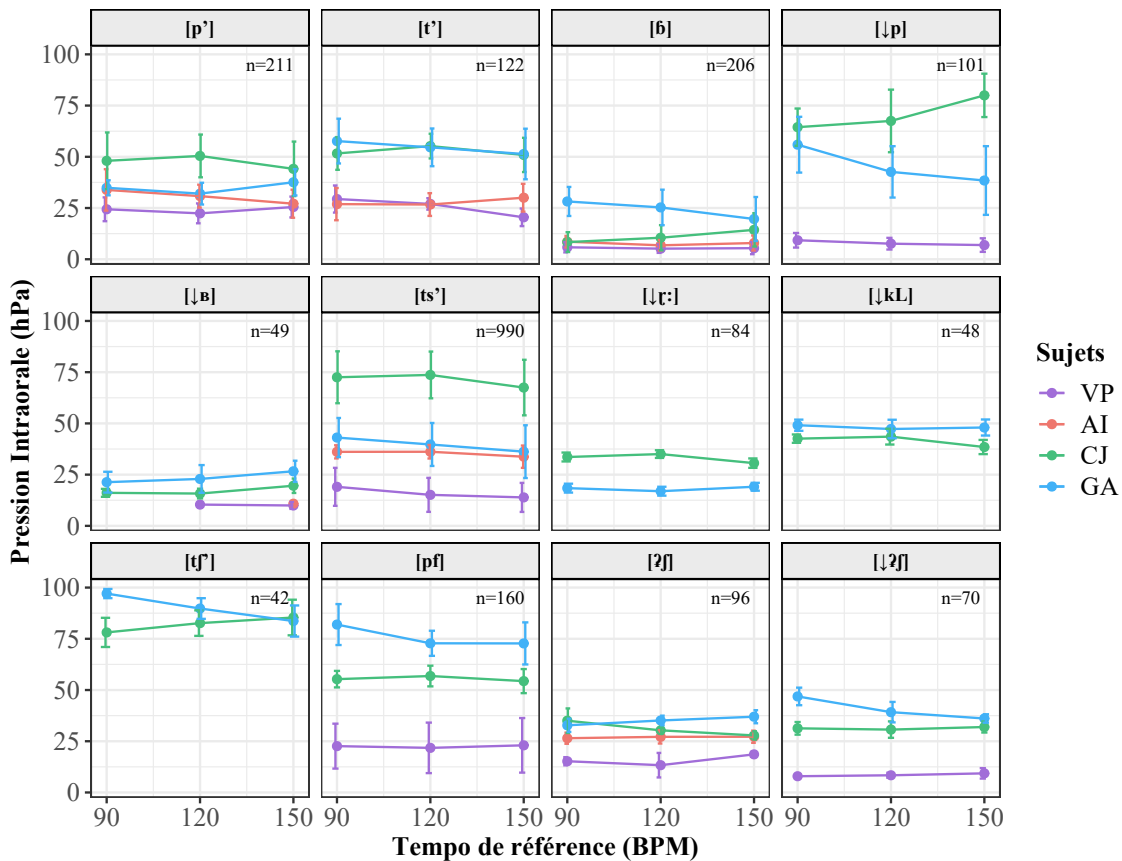


FIGURE 4.12 – Pression intra-orale moyenne (hPa) en fonction du tempo de référence (BPM).

#### 4.3.1.2 Effet du tempo sur la pression selon les caractéristiques initiatrices

Nous venons de voir que le tempo ne provoquait pas de réduction de la pression. Nous voulons à présent savoir si les caractéristiques initiatrices ont une influence sur les diminutions et augmentations de la pression intra-orale. On sait que la pression dans le conduit vocal dépend du lieu d'initiation. Un plus grand volume aura une pression plus faible qu'un petit volume. Une consonne pulmonique aura une pression plus faible qu'une consonne éjective. En revanche, nous ne savons pas si la pression de différents mécanismes d'initiation varie en fonction de la vitesse de production.

La Figure 4.13 illustre la pression intra-orale en fonction du tempo de référence pour 4 des 6 mécanismes d'initiation. Les mécanismes d'initiation pulmoniques montrent des différences entre les sujets et les modes d'initiation par compression et expansion. Concernant les mécanismes pulmoniques produits par compression, les valeurs de pression intra-orale varient entre 10 et 40hPa. La pression générée par les sujets CJ (40-25hPa) et GA (30-40hPa) est en moyenne supérieure à celle générée par les sujets VP (10-35hPa) et AI (25hPa). À l'exception du sujet CJ qui réduit la pression, on observe une très légère tendance à l'augmentation de la pression.

Lorsque le mécanisme pulmonique est produit par expansion, on observe des différences du niveau de la pression chez VP et une variabilité plus importante de la pression chez les sujets CJ et GA (cf. barres d'erreur). Chez VP, la pression est d'environ 10hPa et concerne seulement le *Classic Kick Inward* [ɹp] et le *Lips Roll* [ɹɸʰ:]. Chez le sujet CJ, la pression varie entre 20 et 40hPa, sur une plage de pression similaire au mécanisme par compression. Chez le sujet GA, la pression varie entre 15 et 45hPa, sur une plage de pression plus importante que pour le mécanisme de compression pulmonique. On n'observe pas d'effet du tempo sur la pression

pour des mécanismes d'expansion pulmonique.

Nous avons beaucoup plus de données sur les compressions laryngales que sur les expansion laryngales. Concernant les mécanismes de compression, on remarque bien les distinctions de pression moyenne des sujets. Le sujet VP utilise des mécanismes de compression laryngale dont la pression varie entre 10 et 35hPa. De façon surprenante, la plage de variation de pression est presque identique à celle des mécanismes pulmoniques. Le sujet AI varie la pression entre 25 et 40hPa, soit légèrement plus haut que l'initiation par compression pulmonique. Le sujet CJ varie la pression entre 50 et 80hPa, soit légèrement plus haut que l'initiation par compression pulmonique. Le sujet GA varie la pression entre 50 et 80 hPa et le sujet GA entre 30 et 65hPa. Pour ces sujets, la pression est plus forte pour les mécanismes de compression laryngale que les mécanismes de compression pulmonique. Le sujet GA réduit légèrement la pression. Nous n'observons pas d'effet majeur du tempo sur la pression des mécanismes de compression laryngale. Enfin, les différences observées entre les sujets concernant les mécanismes d'expansion laryngale sont dues au fait que les sujets CJ et GA ont produit le *Classic Kick Inward* comme [β]. Donc les valeurs de pression de ces sujets regroupent celles des sons [β̞] et [↓p]. Pour les sujets AI et VP seul le *Throat Kick* [β̞] est représenté sur la figure. Chez VP et AI, la pression est similaire, plus ou moins 5hPa. Les valeurs de pression chez CJ varient entre 10hPa pour [β̞] et 65hPa pour [β]; chez GA, on est plutôt aux alentours de 20hPa et 40hPa. On note une fois la tendance de GA à diminuer la pression contrairement aux autres sujets.

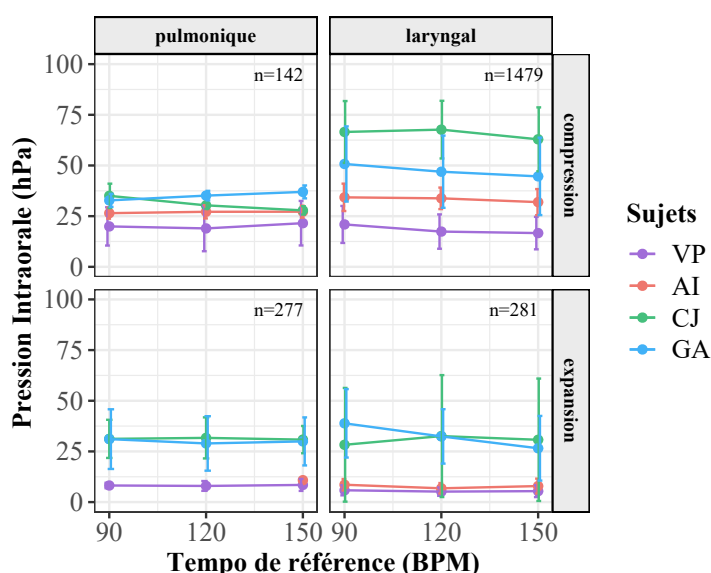


FIGURE 4.13 – Pression intra-orale moyenne (hPa) en fonction du tempo de référence (BPM) selon les mécanismes d'initiation.

Pour résumer, l'analyse des données de pression intra-orale selon le type de mécanisme initiateur utilisé ne montre pas d'effets du tempo de référence sur les variations de pression. On note des différences entre les mécanismes de compression pulmonique et laryngale. La pression tend à être plus faible pour les mécanismes de compression pulmonique. De plus, il est intéressant de souligner que les valeurs moyennes de pression peuvent être similaires d'un lieu d'initiation à l'autre. Ceci nuance la conception communément admise, et quelque peu simpliste, que la pression dépend du lieu d'initiation. En réalité, elle dépend à la fois du lieu et de l'amplitude des gestes de compression ou d'expansion.

#### 4.3.1.3 Effet du tempo sur la pression selon les caractéristiques articulatoires

Nous venons de voir que l'analyse des variations de pression en fonction du tempo et des mécanismes d'initiation n'était pas concluante. Qu'en est-il des lieux et modes d'articulation ? Notre corpus est composé de différents modes d'articulation (occlusif, affriqués, trille) et de différents lieux d'articulation (labial, coronal, dorsal, laryngal). Afin de rendre comparables nos données, nous avons fait le choix d'analyser seulement un sous-ensemble composé de sons produits avec des mécanismes d'initiation laryngaux car ils sont majoritaires dans notre corpus. Malheureusement, les problèmes de saturation empêchent une comparaison systématique entre tous les sujets, lieux et modes d'articulation.

La Figure 4.14 illustre la pression intra-orale moyenne en fonction du tempo et des lieux et modes d'articulation. Nous comparerons seulement les lieux labiaux et coronaux et seulement les affriqués et les occlusives (les trilles du corpus sont seulement produites par expansion). On observe deux tendances, celle des sujets VP et GA et celle des sujets AI et CJ. Chez les sujets VP et GA, on remarque que la pression des sons occlusifs est supérieure à la pression des sons affriqués. Pour ces deux sujets, on remarque une tendance à la diminution de la pression de 5 à 15hPa. Chez les sujets AI et CJ, on observe l'inverse, à savoir une pression plus élevée des sons affriqués par rapport aux sons occlusifs. En revanche, nous n'observons pas de tendance similaire de variation de pression en fonction du tempo entre ces deux sujets.

Concernant les lieux d'articulation, on sait qu'un lieu d'articulation antérieur aura un plus grand volume derrière la constriction qu'un lieu postérieur et que la pression sera donc plus faible pour un lieu antérieur que pour un lieu postérieur. Concernant les sons occlusifs, on observe que pour tous les sujets, la pression est plus faible pour les labiales que pour les coronaux. Pour les affriqués, nous observons d'un côté que pour les sujets VP, AI et CJ, la pression des sons coronaux est supérieure à la pression des sons labiaux. Pour le sujet GA, on observe l'inverse. Les trilles ne sont pas représentées ici, cependant nous avons des données des sujets CJ et GA pour les trilles pulmoniques [ $\downarrow$ B<sup>1</sup>:] (labial) et [ $\downarrow$ r:] (coronal). Pour ces trilles, sur la Figure 4.12, on note les mêmes tendances que pour les sons affriqués, à savoir une pression plus élevée pour la labiale que pour la coronale chez CJ et on note l'inverse chez le sujet GA. Quelque soit le lieu ou le mode, nous n'observons pas de tendance commune à tous les sujets concernant les variations de pression en fonction du tempo et des caractéristiques articulatoires.

En résumé, l'analyse des effets du tempo selon les caractéristiques articulatoires des sons n'est pas concluante et n'explique pas les variations de pression selon le tempo. On note tout de même plus de similarités entre la pression des occlusives coronaux. Quant aux effets du tempo, nous n'observons pas de tendance commune à tous les sujets en fonction des caractéristiques articulatoires.

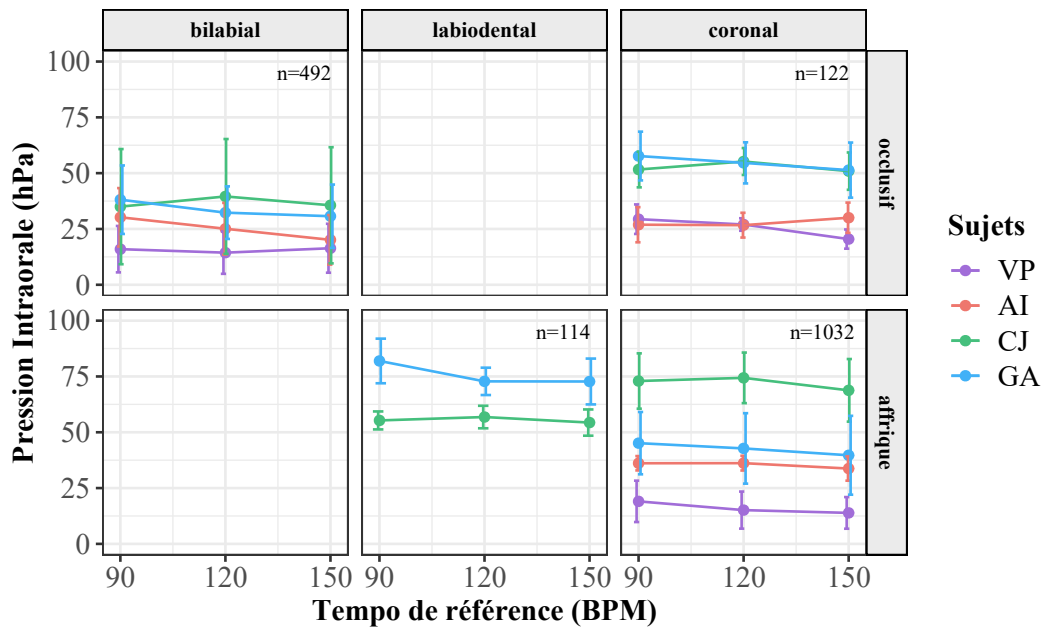


FIGURE 4.14 – Pression intra-orale moyenne (hPa) en fonction du tempo de référence (BPM) selon les mécanismes d’articulation.

#### 4.3.1.4 Effet du tempo sur la pression selon la position et le contexte phonétique adjacent

Nous allons à présent conclure l’analyse des variations de pression intra-orale en discutant de l’influence de la position et du contexte phonétique adjacent. Nous reprendrons le même sous-ensemble de charleston fermé  $[\widehat{ts}']$ . Y-a-t-il des différences de pression de  $[\widehat{ts}']$  selon la position dans la structure métrique et selon le contexte adjacent ?

La Figure 4.15 illustre les variations de pression de  $[\widehat{ts}']$  en fonction du tempo selon les différents patterns du corpus. Il y a un taux de saturation élevé de la pression pour le sujet AI (cf. le nombre d’échantillons  $n$ ). Nous reviendrons sur les données du sujet AI en fin d’analyse. Pour le moment, nous focaliserons notre analyse sur les données des autres sujets. Nos données montrent une fois de plus des comportements similaires entre les positions 4 et 9. Chez VP, on note une réduction plus importante de la pression dans les positions 4 et 9 que dans les positions 2 et 7. En effet, on remarque qu’à 150BPM la pression est plus faible qu’à 90 BPM sauf pour ce qui est du BP  $[\widehat{b} \widehat{ts}' \widehat{ʔ}]; \widehat{ts}' \widehat{b} \widehat{b} \widehat{ts}' \widehat{ʔ}]; \widehat{ts}'$  en position 4. Chez le sujet CJ, on observe un comportement similaire à VP dans ces positions. Toutefois, pour certains patterns, par exemple celui du *Kick Lingual*  $[\widehat{t}' \widehat{ts}' \widehat{tʃ}'; \widehat{ts}' \widehat{t}' \widehat{t}' \widehat{ts}' \widehat{tʃ}'; \widehat{ts}']$ , on observe une augmentation de la pression. Il y a donc un effet de la position et cet effet dépend du pattern dans lequel  $[\widehat{ts}']$  est produit. Enfin, on note que chez le sujet GA les réductions de pression sont plus importantes en position 4 seulement.

Il y a un effet du type de pattern sur la pression moyenne des charlestons fermés. L’effet peut notamment s’observer chez le sujet VP. On peut voir dans les données que la pression de  $[\widehat{ts}']$  est systématiquement plus faible lorsqu’il est produit dans les BP  $[\widehat{b} \widehat{ts}' \downarrow \widehat{ʔ}]; \widehat{ts}' \widehat{b} \widehat{b} \widehat{ts}' \downarrow \widehat{ʔ}]; \widehat{ts}']$ ,  $[\downarrow p \widehat{ts}' \widehat{pf}: \widehat{ts}' \downarrow p \downarrow p \widehat{ts}' \widehat{pf}: \widehat{ts}']$  et  $[\downarrow p \downarrow \widehat{B}^1: \widehat{ts}' \widehat{ʔ}h: \widehat{ts}' \downarrow p \downarrow \widehat{B}^1: \widehat{ts}' \widehat{ʔ}h: \widehat{ts}']$  quelle que soit la position. Concernant ce dernier pattern, il manque les données de pression dans les positions 2 et 7 (position où  $[\widehat{ts}']$  est produit après  $[\downarrow \widehat{B}^1:]$ ). L’analyse des signaux de pression a révélé une pression extrêmement faible ( $P_o < 1\text{hPa}$ ) mais un débit d’air oral positif et une phase de friction sur le spectrogramme. Le son est donc produit mais la pression est proche de zéro.

La Figure 4.16 montre les signaux acoustiques et aérodynamiques entre le BP  $[\widehat{b} \widehat{ts}' \widehat{ʔ}];$



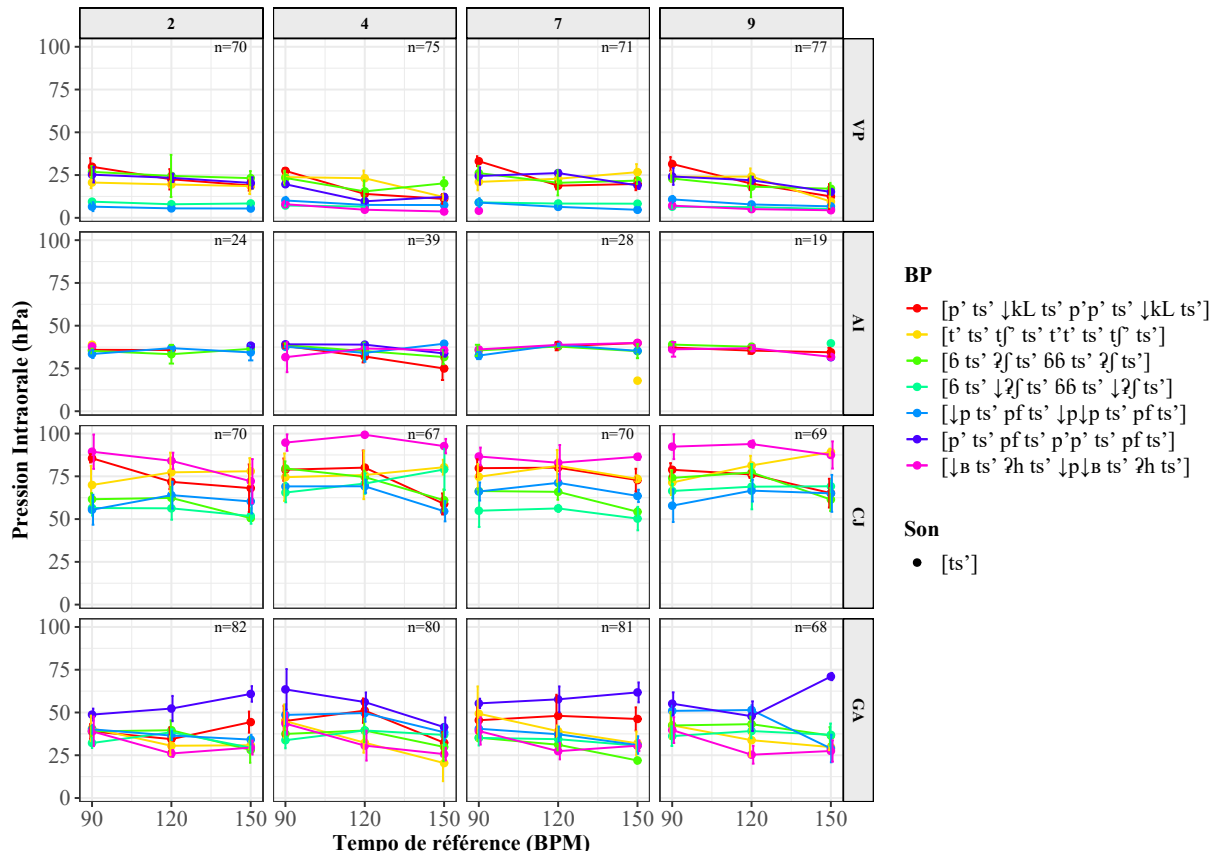
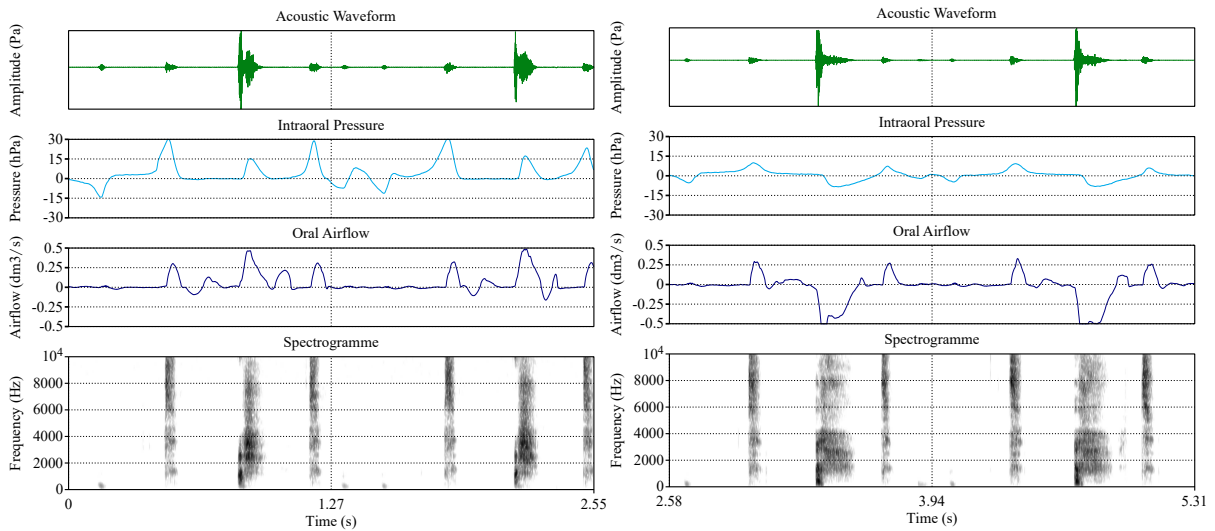


FIGURE 4.15 – Pression intra-orale moyenne (hPa) en fonction du tempo de référence (BPM) selon les types de BP.



(a) [ʃ ts' ʒʃ ts' ʃʃ ts' ʒʃ ts']

(b) [ʃ ts' ↓ʒʃ ts' ʃʃ ts' ↓ʒʃ ts']

FIGURE 4.16 – Illustration des signaux acoustique (vert), aérodynamiques (pression = cyan, débit d'air oral = bleu, débit d'air nasal = violet) et du spectrogramme de deux BP produits par le sujet VP.

[ts̄ ʃ̄ ts̄ ʃ̄] (Figure 4.16a) et le BP [ʃ̄ ts̄ ↓ʃ̄: ts̄ ʃ̄ ts̄ ↓ʃ̄: ts̄] (Figure 4.16b). On note une différence frappante concernant les niveaux de pression. En effet, sur la figure de droite on observe une pression systématiquement plus faible des charlestons et des *Throat Kick* [ʃ̄]. En revanche, les pics de débit d'air oral ne s'aplatissent pas.

Qu'est-ce qui peut expliquer les différences de pression entre les deux patterns ? On sait que les mécanismes pulmoniques de VP ont une faible pression, on pourrait penser que la glotte reste ouverte tout du long. Or, les données fibroscopiques et IRM montrent qu'à 90BPM la glotte est fermée pendant la production de [ts̄]. Il doit donc s'agir d'un phénomène de contrainte de coordination gestuelle. Il est très probable que, tout comme les gestes articulatoires, les gestes initiatoires s'influencent mutuellement. De plus, il est aussi possible que les gestes d'initiation interagissent aussi avec les gestes articulatoires (compression laryngale [ts̄] suivi d'une affriquée épilaryngale [ʃ̄h:]). De telles interactions, de type articulation\*initiation, dépendraient du contexte. Or, chez VP ce n'est pas ce que l'on observe. La réduction de la pression moyenne ( $\pm 15$ hPa) est probablement due à une interaction de type initiation\*initiation car cette réduction n'est pas localisée dans un contexte précis mais s'applique à l'intégralité du pattern. Nous pensons que les gestes initiatoires de compression/expansion réduisent leur amplitude afin de faciliter la coordination intergestuelle et limiter le chevauchement lors du passage d'un mécanisme à l'autre. Ce qu'on note sur la figure, c'est que les durées des charlestons sont similaires entre ces deux BP (cf. Figure 4.10) mais que la forme des pics de pression diffère. À gauche, les pics ressemblent à des courbes gaussiennes alors qu'à droite elles sont aplaties (aspect platykurtique). Ceci montre bien que dans le pattern de droite l'amplitude des gestes est réduite.

Chez le sujet CJ, l'effet est moins important et localisé dans les positions 2 et 7. Les patterns où la pression de [ts̄] est la plus faible sont les BP [ʃ̄ ts̄ ↓ʃ̄: ts̄ ʃ̄ ts̄ ↓ʃ̄: ts̄], [ʃ̄ ts̄ ʃ̄: ts̄ ʃ̄ ts̄ ʃ̄: ts̄], [↓p ts̄ p̄f: ts̄ ↓p ↓p ts̄ p̄f: ts̄] lorsque celui-ci est produit dans les positions 2 et 7. Contrairement à VP, dans le BP [↓p↓B̄: ts̄ ʃ̄h: ts̄ ↓p ↓p↓B̄: ts̄ ʃ̄h: ts̄] la pression de [ts̄] est plus forte que dans les autres BP. Il est possible que le sujet CJ augmente significativement la pression pour éviter une pression trop faible et ne pas risquer d'élider le son<sup>5</sup>. La différence de pression moyenne dans les positions 2 et 7 suggère qu'il peut s'agir d'une contrainte de coordination gestuelle de type initiation\*initiation ou bien initiation\*articulation. La contrainte de type initiation\*articulation peut expliquer pourquoi la différence de pression moyenne est localisée dans un contexte où [ts̄] précède [ʃ̄:] ou [↓ʃ̄:]. Ce que montrent les données fibroscopiques de ce sujet (Figure 4.17), c'est qu'à la fin du geste de compression laryngale du charleston, le sujet CJ prépare déjà le geste d'occlusion épilaryngale en maintenant la constriction épilaryngale, en descendant le larynx et en reculant la racine de la langue et l'épiglotte. Le larynx continue de descendre jusqu'au relâchement de [ʃ̄:]. Cependant, cette explication ne saurait expliquer pourquoi dans le pattern de *Classic Kick Inward* (composé de [ʃ̄], [ts̄] et [p̄f:]), nous observons le même effet de réduction de pression moyenne localisée dans les position 2 et 7. Il est donc possible que le geste d'expansion laryngale précédant [ts̄] influence le niveau de pression intra-orale du son suivant. Cette explication vaudrait aussi pour les BP [ʃ̄ ts̄ ʃ̄: ts̄ ʃ̄ ts̄ ʃ̄: ts̄] et [ʃ̄ ts̄ ↓ʃ̄: ts̄ ʃ̄ ts̄ ↓ʃ̄: ts̄] où [ts̄] est précédé de [ʃ̄].

Les seules données que nous illustrons pour le sujet AI sont des données non-saturées, c'est-à-dire des données dont la pression est inférieure à 40hPa. On remarque que les patterns dont la pression des charlestons est inférieure à 40hPa sont des BP composés de sons pulmoniques à savoir les BP [ʃ̄ ts̄ ʃ̄: ts̄ ʃ̄ ts̄ ʃ̄: ts̄], [↓p ts̄ p̄f: ts̄ ↓p ↓p ts̄ p̄f: ts̄], [↓p↓B̄: ts̄ ʃ̄h: ts̄ ↓p ↓p↓B̄: ts̄ ʃ̄h: ts̄] et [p̄ ts̄ ↓kL: ts̄ p̄ p̄ ts̄ ↓kL: ts̄]. On note aussi que dans la position 4 l'échantillon est plus important donc qu'il y a plus de charlestons produits avec une pression

5. Par élision, nous entendons la non-production d'un son.

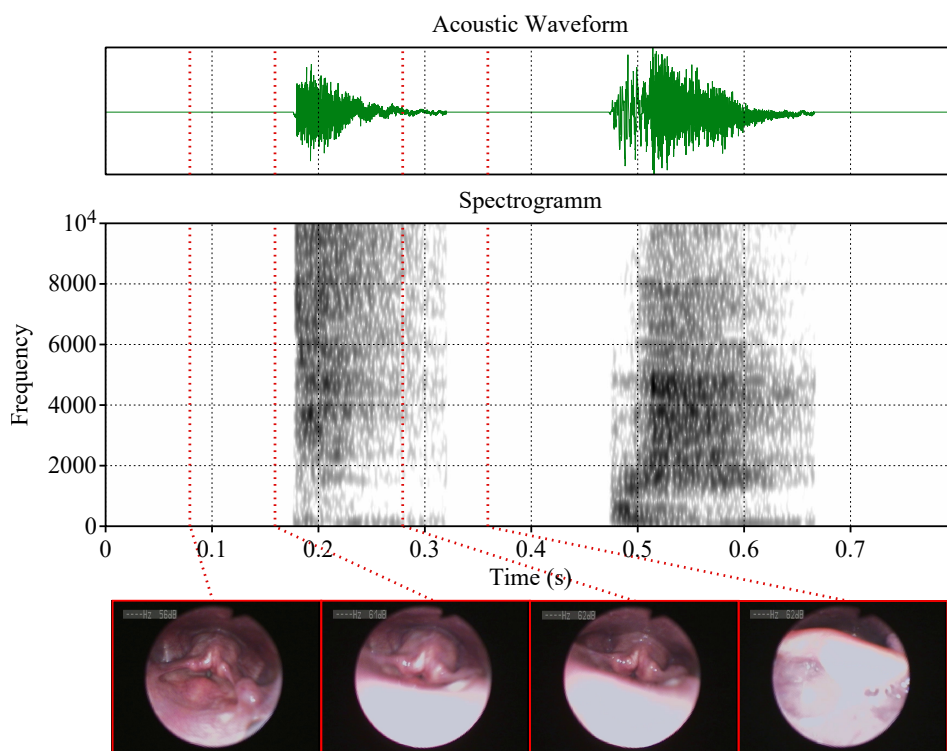


FIGURE 4.17 – Signal acoustique et images fibroscopiques correspondantes de la production par CJ de  $[\widehat{ts}]$  en position 2 et de  $[\widehat{tʃ}]$  en position 3 dans le BP  $[\widehat{b} \widehat{ts} \widehat{tʃ} : \widehat{ts} \widehat{b} \widehat{b} \widehat{ts} \widehat{tʃ} : \widehat{ts}]$ . Tempo de référence = 90BPM.

inférieure à 40hPa dans cette position. Il est possible que la pression moyenne des charlestons se réduise au contact de sons pulmoniques.

Qu'en est-il du sujet GA ? Les données du sujet GA n'indiquent aucun effet des sons pulmoniques sur la pression moyenne des charlestons. En revanche, on note une différence concernant le BP  $[\widehat{p} \widehat{ts} \widehat{pf} : \widehat{ts} \widehat{p} \widehat{p} \widehat{ts} \widehat{pf} : \widehat{ts}]$  comme illustré sur la Figure 4.18. Dans ce BP la pression de  $[\widehat{ts}]$  augmente systématiquement dans les positions 2, 7 et 9 à mesure que le tempo accélère. Nous avons noté que chez GA, la pression de la *PF snare* se réduisait ce qui suggère que l'effet du contextuel ne touche pas que le charleston. En regardant les données de pression (incomplètes) du *Classic Kick*  $[\widehat{p}]$  de ce pattern, on note que la pression augmente légèrement. Il y a donc un effet du contexte sur la grosse caisse. La pression moyenne du *Classick Kick* et du charleston fermé sont plus faibles que la pression générée pour produire la *PF snare*. De plus, la pression moyenne de la grosse caisse et du charleston tend à augmenter alors que la pression moyenne de la *PF snare* tend à se réduire. On peut donc en déduire que les pressions moyennes tendent à varier sur une plage de pression plus réduite lorsque le tempo accélère. À 90BPM la pression varie entre 30 et 95hPa, soit une plage de 65hPa alors qu'à 150BPM la plage de variation est réduite entre 40 et 85hPa, soit une plage de 45hPa. Toutefois, le BP  $[\widehat{p} \widehat{ts} \widehat{pf} : \widehat{ts} \widehat{p} \widehat{p} \widehat{ts} \widehat{pf} : \widehat{ts}]$  n'est pas l'unique pattern composé uniquement de mécanismes éjectifs. Il y a aussi le BP  $[\widehat{t} \widehat{ts} \widehat{tʃ} : \widehat{ts} \widehat{t} \widehat{t} \widehat{ts} \widehat{tʃ} : \widehat{ts}]$  pour lequel nous n'observons pas de tendance similaire, cependant ce BP est particulier en ce qu'il est exclusivement composé de sons coronaux éjectifs ce qui contraint fortement les patterns de coordination gestuelle.

Pour résumer, nous avons montré deux phénomènes de variation de la pression intra-orale. En premier lieu, il y a un effet de la position 4 sur la pression des charlestons où celle-ci tend

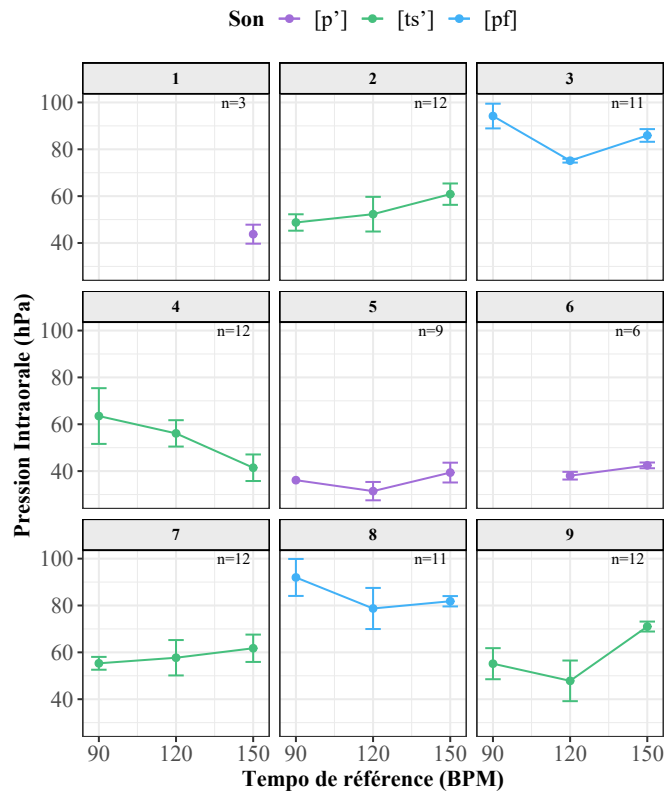


FIGURE 4.18 – Pression intra-orale moyenne (hPa) en fonction du tempo de référence (BPM) du BP [p' ts' pf: ts' p' p' ts' pf: ts'] produit par GA.

à se réduire. Enfin, il y a aussi des phénomènes de variation contextuelle qui peuvent dépendre à la fois du lieu et du mode d'initiation ainsi que du lieu d'articulation. Cependant, le BP [p' ts' ↓KL: ts' p' p' ts' ↓KL: ts'] n'est pas affecté par ces contraintes de coordination gestuelle alors qu'il est composé de la caisse claire pulmonique [↓KL:]. Ceci peut s'expliquer par le fait que les sons de ce pattern constituent le répertoire d'entrée dans la pratique du beatbox. On s'entraîne à produire ces sons de manière isolée puis on s'entraîne à les coordonner. Ce n'est que par la suite que l'on acquiert d'autres sonorités et d'autres structures plus complexes de coordination entre les sons du répertoire musical d'un artiste. Il est fort possible que ce pattern soit mieux maîtrisé que les autres BP, qui parfois ce sont révélés être des nouveautés pour le sujet sans pour autant qu'ils soient difficiles à produire.

#### 4.3.1.5 Synthèse de l'analyse de la pression intra-orale

L'analyse de la pression intra-orale en fonction du tempo a montré des stratégies individuelles d'adaptation qui ne suivaient pas la même tendance que la réduction temporelle des gestes. Par conséquent, la durée et la pression des sons est régulée différemment en fonction du tempo. Notre hypothèse selon laquelle la pression se réduirait à cause d'une réduction de l'amplitude du geste d'initiation est donc réfutée par nos données. En effet, parfois la pression se réduit, parfois elle augmente sans que ces différences ne s'expliquent par les variations du tempo de référence. En revanche, nous avons noté des phénomènes de variation contextuelle. En position 4, la pression se réduit systématiquement. De plus, il semblerait y avoir des contraintes de coordination intergestuelle entre initiation pulmonique et laryngale d'une part, et entre initiation et articulation d'autre part. Toutefois, la contrainte a été observée principalement chez VP,

dans une moindre mesure chez CJ. Elle n'a pas été observée chez le sujet GA. Cela suggère que ces contraintes biomécaniques et aéromécaniques peuvent être dépassées. Cela pose la question du rôle de l'apprentissage et de l'entraînement des beatboxeurs.

Si la durée des gestes de production et la pression intra-orale, et par extension l'amplitude des gestes initiatrices, ne sont pas régulées de la même façon, de quel(s) facteur(s) dépendent donc les variations de pression ? Nous pensons que les variations de pression vont principalement dépendre de l'intensité intrinsèque des sons. Dans une structure musicale, tous les éléments qui la composent se différencient par des variations de hauteur, de longueur et d'intensité ce qui va créer le rythme. L'intensité d'un son est corrélée à la vitesse des particules (Fant, 1970). La vitesse des particules dépend elle-même de la pression et de l'aire de la constriction. On peut donc penser qu'à taille de constriction constante, plus la pression est forte, plus l'intensité sera forte. Les variations d'intensité sonore des sons dépendraient des variations de pression. Cela supposerait que le contrôle de la pression se fasse par une boucle de feedback auditif et possiblement proprioceptif ou aéro-tactile.

### 4.3.2 Effet du tempo sur l'aire de la constriction

Nous savons que les constriction jouent un rôle primordial pendant l'articulation car elles permettent de transformer des flux laminaires en flux turbulents. Il est important de discuter des variations de l'aire de constriction avant d'aborder les effets du tempo sur l'acoustique des sons du corpus. En effet, l'aire de la constriction peut, en plus des variations de pression, affecter l'intensité de la turbulence. Nous allons travailler sur un échantillon très réduit de sons : les affriquées. Les affriquées se caractérisent par le maintien de la constriction après le relâchement de l'occlusion sur un temps donné. À l'inverse, les occlusives se caractérisent par un éloignement rapide des articulateurs après le relâchement pour éviter une phase de friction tenue. L'aire de la constriction des affriquées est donc plus pertinente à analyser que celle des occlusives. Nous n'incluons pas les trilles que l'on peut considérer comme une suite d'occlusives.

#### 4.3.2.1 Effet du tempo sur l'aire de la constriction selon les sons

La Figure 4.19 illustre l'aire relative de la constriction en fonction du tempo de référence. L'échantillon est d'environ 1200 sons dont la très grande majorité sont des charlestons fermés. La faiblesse des échantillons est due à la saturation de la pression ou du débit d'air oral. Les pointillés indiquent la zone où les variations d'aire de la constriction induisent des variations d'intensité de la turbulence de  $\pm 3\text{dB}$ . En-dessous et au-delà de ces seuils, l'intensité se réduit de façon plus importante. On note que la majeure partie de nos données sont comprises dans l'intervalle estimé par Stevens (1989). On peut en conclure que les variations de l'aire de la constriction n'ont qu'un effet mineur sur l'intensité de la turbulence.

Aucune donnée, à l'exception peut-être de quelques charlestons fermés, ne se situent en dessous de  $0.03\text{cm}^2$ . En revanche, il y a des constriction dont les aires dépassent le seuil de  $0.2\text{cm}^2$ . C'est le cas de VP et CJ pour  $[\text{pf}]$ , de GA pour  $[\text{tʃ}]$  et de CJ pour  $[\text{dʒ}]$ . Notons la différence entre l'aire de la constriction laminaire de la *Tish Snare*  $[\text{tʃ}]$  et des *ESH Snare*  $[\text{tʃ}]$  et  $[\text{dʒ}]$ . La constriction de la *Tish Snare* est comparable à celle des charlestons fermés  $[\text{ts}]$ . Cela concorde avec la description que donne VP du son  $[\text{tʃ}]$ , un son produit en « écrasant la langue sur les alvéoles ».

En ce qui concerne les effets du tempo, nous observons que la constriction des charlestons, de la *K-Snare Inward* de CJ et de la *Tish Snare* des sujets CJ et GA tend à s'agrandir. À l'inverse, les sujets GA et CJ tendent à réduire l'aire de la constriction de  $[\text{dʒ}]$ . Pour les autres sons, on

note des tendances individuelles seulement. Les aires (relatives) de constriction sont comprises dans l'intervalle de  $\pm 3\text{dB}$  et ne sont donc pas susceptibles de provoquer des variations importantes d'intensité de la turbulence. Est-ce que ces variations sont causées par des contraintes de coordination gestuelle ?

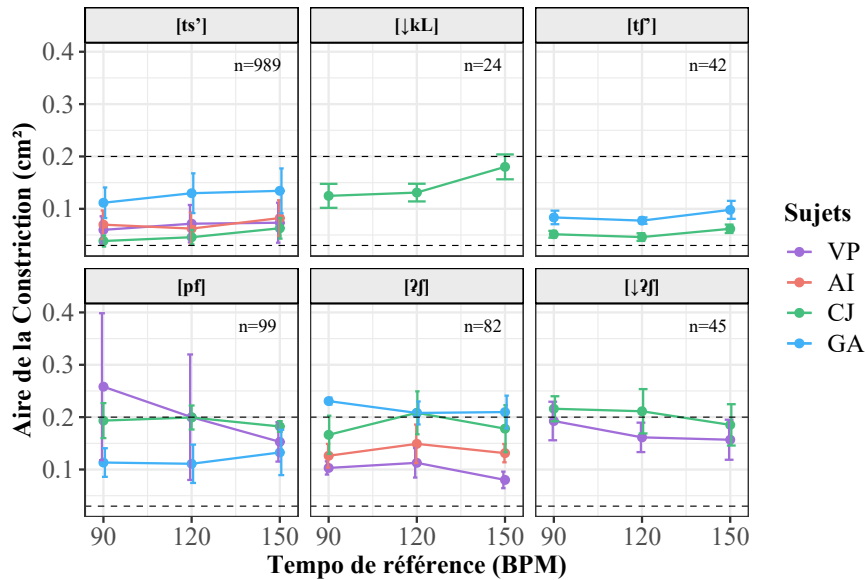


FIGURE 4.19 – Aire relative de constriction ( $\text{cm}^2$ ) en fonction du tempo de référence (BPM) selon les sons. Les pointillés à  $0.03\text{cm}^2$  indiquent le seuil en-dessous duquel l'intensité de la turbulence est réduite. Les pointillés à  $0.2\text{cm}^2$  indiquent le seuil au-dessus duquel l'intensité de la turbulence est réduite.

#### 4.3.2.2 Effet du tempo sur l'aire de la constriction selon le contexte phonétique adjacent

Nous allons à présent analyser les variations de l'aire de la constriction de  $[\widehat{\text{ts}}']$  en fonction du tempo dans différents patterns et différentes positions. La Figure 4.20 représente l'aire relative de la constriction en fonction du tempo de référence. On note que l'aire de la constriction tend à augmenter en fonction du tempo. Dans quelques cas, on remarque que la constriction reste stable (sujet AI - BP  $[\widehat{\text{b}}' \widehat{\text{ts}}' \widehat{\text{ʔ}}] : \widehat{\text{ts}}' \widehat{\text{b}}' \widehat{\text{b}}' \widehat{\text{ts}}' \widehat{\text{ʔ}}] : \widehat{\text{ts}}' \widehat{\text{ts}}' \widehat{\text{ʔ}}]$ ). Dans très peu de cas, l'aire se réduit en fonction du tempo. Encore une fois, on note des comportements similaires dans les positions 2 et 7 et dans les positions 4 et 9. Les trajectoires des moyennes de l'aire de constriction montrent des stratégies individuelles différentes. Par exemple, en position 7, CJ garde une aire de constriction stable entre 90 et 120BPM puis augmente fortement l'aire de la constriction. On observe une stratégie inverse chez GA.

On note des différences de l'aire de la constriction selon le BP, ce qui suggère des effets du contexte adjacent sur le geste articulatoire. Chez le sujet VP, on note que la constriction de  $[\widehat{\text{ts}}']$  est plus large dans les BP  $[\downarrow \text{p} \widehat{\text{ts}}' \text{pf} : \widehat{\text{ts}}' \downarrow \text{p} \downarrow \text{p} \widehat{\text{ts}}' \text{pf} : \widehat{\text{ts}}']$ ,  $[\widehat{\text{b}}' \widehat{\text{ts}}' \widehat{\text{ʔ}}] : \widehat{\text{ts}}' \widehat{\text{b}}' \widehat{\text{b}}' \widehat{\text{ts}}' \widehat{\text{ʔ}}] : \widehat{\text{ts}}' \widehat{\text{ts}}' \widehat{\text{ʔ}}]$  et  $[\downarrow \text{p} \downarrow \text{b}^1 : \widehat{\text{ts}}' \widehat{\text{ʔ}}] : \widehat{\text{ts}}' \downarrow \text{p} \downarrow \text{p} \downarrow \text{b}^1 : \widehat{\text{ts}}' \widehat{\text{ʔ}}] : \widehat{\text{ts}}' \widehat{\text{ʔ}}]$ . Chez GA, on observe qu'en position 4 la constriction de  $[\widehat{\text{ts}}']$  est significativement plus large que dans les autres patterns. On observe aussi, chez GA, qu'en position 7, particulièrement à 120 et 150BPM, la constriction de  $[\widehat{\text{ts}}']$  est plus grande dans les BP  $[\widehat{\text{b}}' \widehat{\text{ts}}' \widehat{\text{ʔ}}] : \widehat{\text{ts}}' \widehat{\text{b}}' \widehat{\text{b}}' \widehat{\text{ts}}' \widehat{\text{ʔ}}] : \widehat{\text{ts}}' \widehat{\text{ts}}' \widehat{\text{ʔ}}]$ ,  $[\widehat{\text{b}}' \widehat{\text{ts}}' \downarrow \widehat{\text{ʔ}}] : \widehat{\text{ts}}' \widehat{\text{b}}' \widehat{\text{b}}' \widehat{\text{ts}}' \downarrow \widehat{\text{ʔ}}] : \widehat{\text{ts}}' \widehat{\text{ts}}' \widehat{\text{ʔ}}]$ ,  $[\widehat{\text{b}}' \widehat{\text{ts}}' \text{pf} : \widehat{\text{ts}}' \widehat{\text{b}}' \widehat{\text{b}}' \widehat{\text{ts}}' \text{pf} : \widehat{\text{ts}}' \widehat{\text{ts}}' \widehat{\text{ʔ}}]$  et  $[\widehat{\text{b}}' \widehat{\text{ts}}' \widehat{\text{ʔ}}] : \widehat{\text{ts}}' \widehat{\text{b}}' \widehat{\text{b}}' \widehat{\text{ts}}' \widehat{\text{ʔ}}] : \widehat{\text{ts}}' \widehat{\text{ts}}' \widehat{\text{ʔ}}]$ .

En résumé, nous avons montré qu'il y a un effet du tempo sur l'aire de la constriction pour le son  $[\widehat{\text{ts}}']$ . En effet, nous avons montré que la constriction devenait plus grande à mesure que

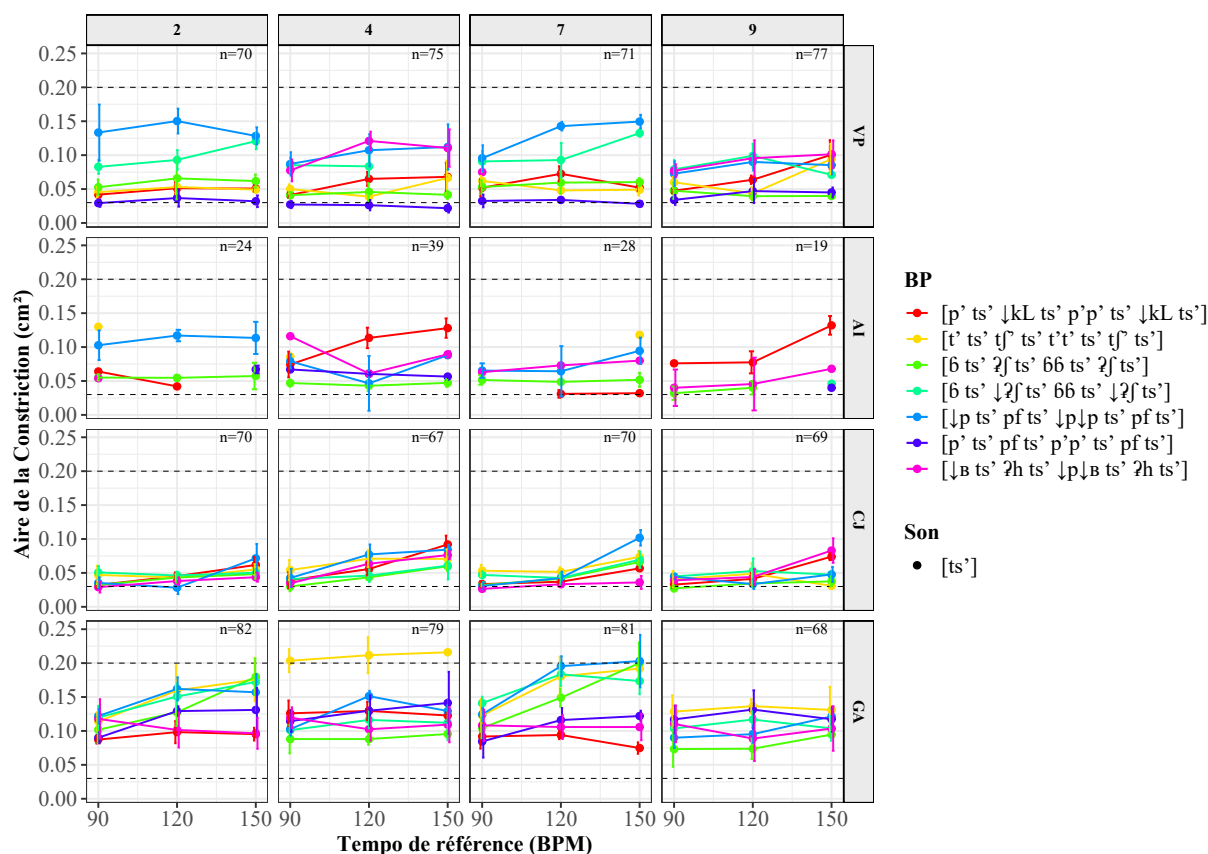


FIGURE 4.20 – Aire relative de constriction ( $cm^2$ ) en fonction du temps de référence (BPM) pour le son  $[ts']$  dans différents BP et différentes positions. Les pointillés à  $0.03cm^2$  indiquent le seuil en-dessous duquel l'intensité de la turbulence est réduite. Les pointillés à  $0.2cm^2$  indiquent le seuil au dessus duquel l'intensité de la turbulence est réduite.

le tempo accélère. Nous avons aussi montré que la durée acoustique de  $[ts']$  se réduisait en fonction du tempo, cela signifie que la durée de maintien de la constriction se réduit. À 90BPM la constriction est maintenue et à 150BPM la constriction est très rapidement relâchée. Il arrive aussi que la phase de friction disparaisse et laisse place à un bruit d'explosion, ce qui suggère que plus on réduit les charlestons plus ils risquent d'être produits comme des occlusives. Nous pensons donc que les variations de l'aire de la constriction reflètent la dynamique temporelle de coordination articulaire et qu'à 150BPM nous serions plus susceptibles de mesurer différents degrés d'éloignement de l'apex de la langue plutôt qu'une constriction maintenue.

## 4.4 Analyse acoustique

Nous allons à présent conclure le chapitre 4 avec l'analyse des moments spectraux. Les moments spectraux permettent de décrire la répartition (Centre de Gravité Spectral et *Skewness*) et la forme spectrale (écart-type du Centre de Gravité et *Kurtosis*) des sons. Nous illustrerons seulement les données concernant le Centre de Gravité Spectral et son écart-type ; nous mentionnerons le *Skewness* et le *Kurtosis* dans le texte. L'analyse de la variation des moments spectraux nous permettra par la suite d'inférer des changements de configuration articuloire en fonction du tempo. Nous commencerons par analyser les variations de la répartition puis de forme spectrale des différents sons du corpus. Enfin, nous tenterons de voir s'il y a des effets du contexte adjacent sur les moments spectraux de [tʰ].

### 4.4.1 Effet du tempo sur les moments spectraux des sons

Les Figures de cette section représentent le Centre de Gravité Spectral (CdGS) en fonction du tempo de référence. Avec le *Skewness*, cette mesure permet de décrire la répartition dans les basses, moyennes et hautes fréquences. Pour chaque son, nous avons représenté chacun des évènements acoustiques composant les sons. De plus, les durées des fenêtres d'analyses sont différentes selon l'évènement acoustique (durée du burst *versus* durée du voisement) ; nous avons donc indiqué la durée des fenêtres d'analyse avec différentes formes de point. Nous analyserons les variations de répartition spectrale des sons par catégories.

#### 4.4.1.1 Grosses Caisses

##### 4.4.1.1.1 Répartition Spectrale

Nous l'avons mentionné dans le Chapitre 3, les grosses caisses se caractérisent par de l'énergie dans les basses fréquences. Concernant les variations du Centre de Gravité en fonction du tempo, on remarque qu'il est stable (Figure 4.21). Celle-ci ne varie pas en fonction du tempo ce qui suggère que la configuration du conduit vocal reste la même quel que soit le tempo. On peut tout de même noter que chez le sujet VP, le Centre de Gravité spectral de [tʰ] est plus élevé que chez les autres sujets. On note que le coefficient d'asymétrie est positif, l'énergie prédomine donc dans les basses fréquences. On remarque dans certains cas une tendance à l'augmentation du *Skewness* pour les bruits d'explosion ([tʰ]). Cela suggère que le spectre présente moins d'énergie dans les moyennes et hautes fréquences à mesure que le tempo accélère.

##### 4.4.1.1.2 Forme Spectrale

La Figure 4.22 illustre l'écart-type du Centre de Gravité Spectral en fonction du tempo. Il indique l'étendue du Centre de Gravité. L'écart-type est faible, il se situe aux alentours de 0.5kHz. De plus, les données de *Kurtosis* indiquent un *coefficient d'aplatissement* élevé en moyenne de 160(±500). Ces deux mesures confirment les observations faites dans le Chapitre 3 à propos du timbre des grosses caisses : les grosses caisses présentent un pic spectral étroit dans les basses fréquences et une pente spectrale abrupte.

Les données du sujet VP indiquent parfois un écart-type plus important, parfois d'1kHz ([tʰ]) et même 2kHz si l'on regarde les données du burst de [↓p]. Pour ce son, nous avons montré la variabilité de production qui parfois tendait à être produit comme un *Lips Roll* (cf. Matrice des erreurs de VP - Figure 4.4a). Des valeurs élevées de l'écart-type et des valeurs faibles de



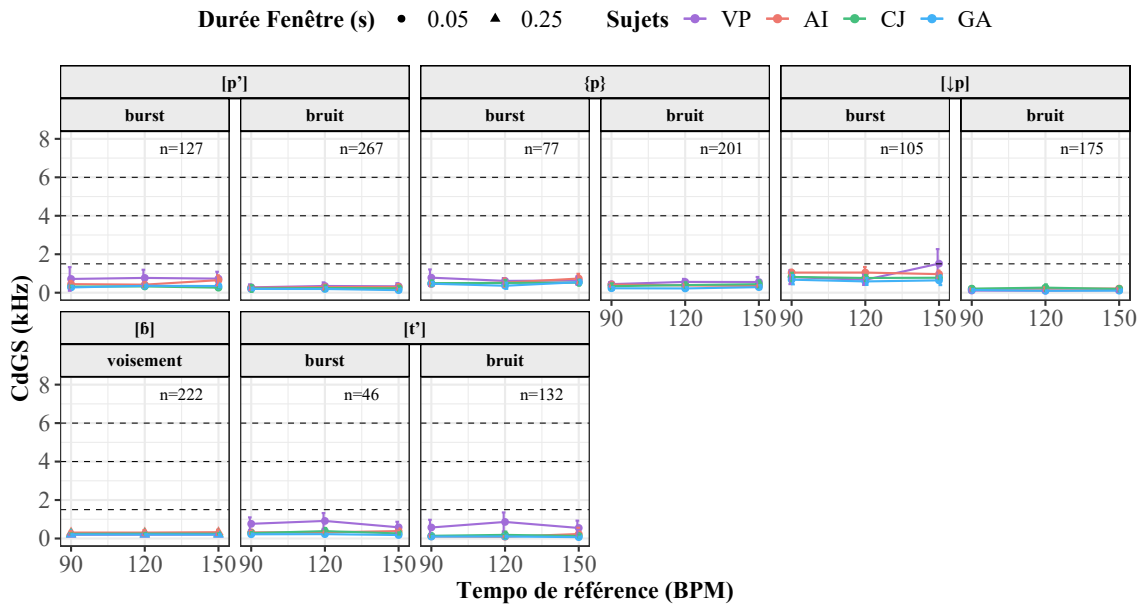


FIGURE 4.21 – Centre de Gravité Spectral (kHz) en fonction du tempo de référence (BPM) selon les sons, les événements acoustiques et les sujets. La durée des fenêtres d’analyse est indiquée par des formes de point différentes. Les lignes en pointillés délimitent les basses fréquences ( $F < 1.5\text{kHz}$ ), les moyennes fréquences ( $1.5 < F < 4\text{kHz}$ ) et les hautes fréquences ( $F > 4\text{kHz}$ ).

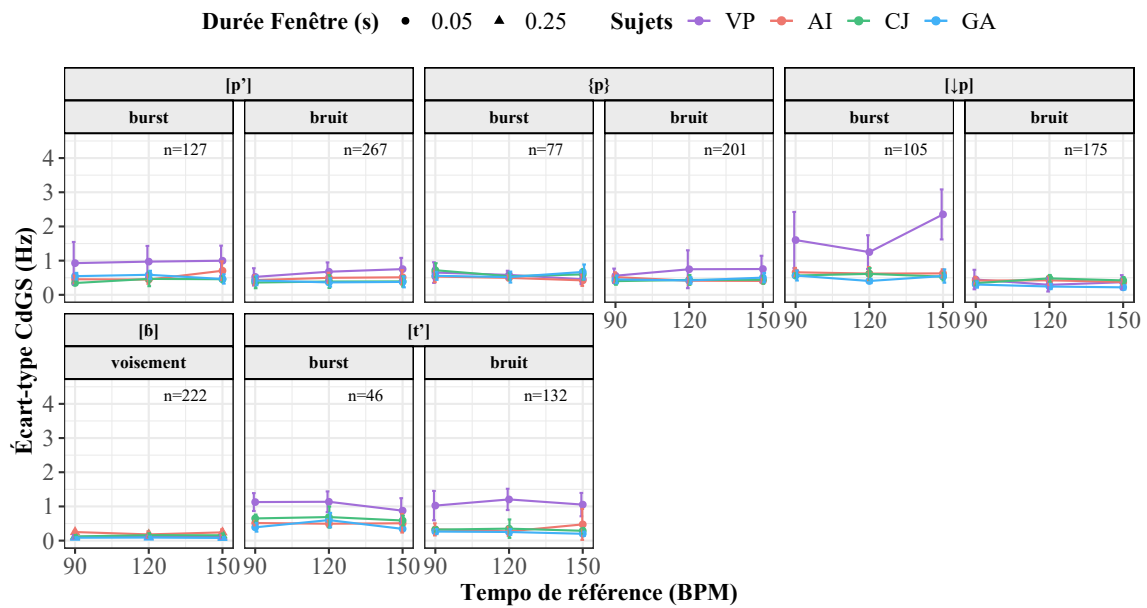


FIGURE 4.22 – Écart-type du Centre de Gravité spectral (kHz) grosses caisses en fonction du tempo de référence. L’écart-type décrit l’étendue fréquentielle du Centre de Gravité spectral.

*Kurtosis* ( $\pm 5\text{-}15$  unités du coefficient) indiquent une pente spectrale moins abrupte et un spectre plus plat. Nous n’observons pas d’effet majeur du tempo sur la pente et la forme spectrale.

## 4.4.1.2 Charlestons

### 4.4.1.2.1 Répartition Spectrale

Comme mentionné dans le Chapitre 3, les charlestons possèdent de l'énergie dans les hautes fréquences. Les charlestons ont un Centre de Gravité qui varie en fonction du temps. De plus, les barres d'erreur indiquent une variabilité importante de la répartition spectrale. En effet, lors de la phase de friction, ce dernier diminue à mesure que le tempo accélère, à l'exception du sujet GA pour  $\{\widehat{ts}\}$  où le Centre de Gravité spectral reste stable. Notons que l'énergie de ce son se trouve dans les moyennes fréquences chez ce sujet. Chez les sujets VP et AI, le Centre de Gravité spectral moyen de  $\{\widehat{ts}'\}$  se déplace dans les moyennes fréquences à 150BPM. L'abaissement fréquentiel du Centre de Gravité suggère un allongement de la cavité antérieure. L'allongement peut-être dû à un recul du lieu de constriction ou bien à une protrusion des lèvres.

Le coefficient d'asymétrie tend à augmenter et le coefficient moyen du bruit de friction devient positif à mesure que le tempo accélère. Cette tendance s'observe pour tous les sujets. Cette augmentation reflète l'abaissement de la répartition fréquentielle vers les moyennes fréquences. En revanche, pour la variante *humming*  $\{\widehat{ts}\}$ , les variations du coefficient, en particulier pour la phase de friction, se concentrent autour de 0 (distribution symétrique) et ne montrent pas de différences majeures entre les valeurs à 90BPM et 150BPM.

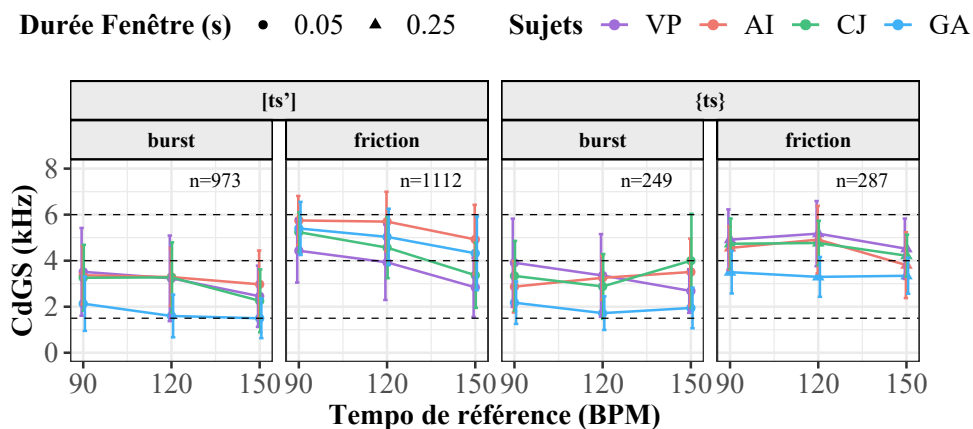


FIGURE 4.23 – Centre de Gravité Spectral (kHz) des charlestons en fonction du tempo de référence (BPM) selon les sons, les événements acoustiques et les sujets.

### 4.4.1.2.2 Forme Spectrale

La Figure 4.24 illustre les variations de l'écart-type du Centre de Gravité spectral en fonction du tempo. On note que pour les charlestons, l'écart-type est élevé. Le Centre de Gravité spectral s'étend sur une plage de fréquence qui varie entre 1.5kHz et 4kHz. Les données de *Kurtosis* indiquent un coefficient qui varie entre  $-1$  et  $10$ . Un écart-type élevé et un faible *Kurtosis* indiquent une pente spectrale, et par extension une forme spectrale, aplatie. Toutefois, l'écart-type et le *Kurtosis* sont variables comme le montrent les barres d'erreur. Les mesures varient selon les sujets, le tempo et la variante du charleston. Nous reviendrons plus bas sur les possibles causes de ces variations de forme spectrale.

Durée Fenêtre (s) • 0.05 ▲ 0.25 Sujets — VP — AI — CJ — GA

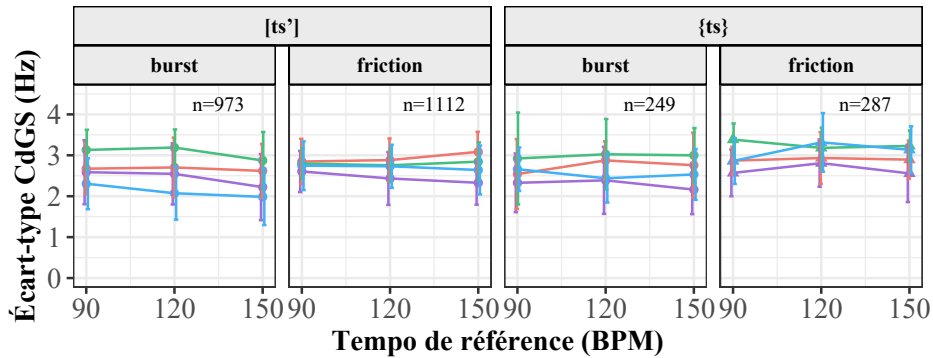


FIGURE 4.24 – Écart-type du Centre de Gravité spectral (kHz) charlestons en fonction du tempo de référence (BPM) selon les sons, les événements acoustiques et les sujets.

### 4.4.1.3 Caisses Claires Orales

#### 4.4.1.3.1 Répartition Spectrale

La Figure 4.25 représente le Centre de Gravité Spectral (kHz) des caisses claires orales en fonction du tempo de référence. On note que l'énergie du burst et du bruit d'explosion prédomine dans les fréquences moyennes-basses alors que le bruit de friction concentre l'énergie dans les fréquences moyennes, voire hautes dans le cas de  $[\widehat{tj}']$  chez les sujets CJ et GA. Les données de *Skewness* indiquent aussi une concentration de l'énergie dans les moyennes et basses fréquences. En effet, le coefficient moyen varie entre 0 et 5 ce qui suggère une asymétrie vers les moyennes et basses fréquences du spectre. Un coefficient de 0 indiquerait une énergie répartie symétriquement autour de 5kHz. En revanche, pour le relâchement labial de  $[\widehat{pf}]$  et  $\{\widehat{pf}\}$ , le coefficient est supérieur à 5 ce qui suggère que l'énergie se concentre dans les basses fréquences. Sur les deux figures, nous avons d'une part des tendances similaires entre les sujets pour les caisses claires vélares et d'autre part, des tendances différentes pour les autres caisses claires orales.

Concernant  $[\downarrow\widehat{kl}]$ ,  $[\downarrow\widehat{kl}]$  et  $\{\downarrow\widehat{kl}\}$ , on note que la répartition de l'énergie du burst et du bruit d'explosion de la variante *humming* prédomine dans les moyennes fréquences en comparaison aux variantes classiques où l'énergie se concentre dans les basses fréquences. Quand au bruit de friction, dans les deux cas, l'énergie se concentre dans les moyennes fréquences. Notons que chez AI, qui a produit  $[\downarrow\widehat{kl}]$ , l'énergie se concentre dans les fréquences moyennes et basses. Le tempo ne semble pas affecter la répartition spectrale. On remarque quand même que chez GA, il y a une augmentation du Centre de Gravité Spectral pour la variante *humming*. Cependant, la variation du coefficient d'asymétrie ne suggère pas de changement majeur de la répartition spectrale en fonction du tempo.

La *Tish Snare*  $[\widehat{tj}']$  présente des différences de répartition spectrale selon les sujets. Chez VP et AI, l'énergie se trouve dans les moyennes fréquences alors que chez CJ et GA, l'énergie se trouve dans les hautes fréquences. Cela suggère que la cavité antérieure chez CJ et GA est plus courte que chez VP et AI. On note des effets du tempo sur le Centre de Gravité chez les sujets VP et AI. L'abaissement du Centre de Gravité des fréquences moyennes-hautes vers les fréquences moyennes-basses chez VP indique que la cavité antérieure s'allonge. Il est peu plausible que l'allongement de la cavité antérieure soit dû à un recul du lieu d'articulation étant

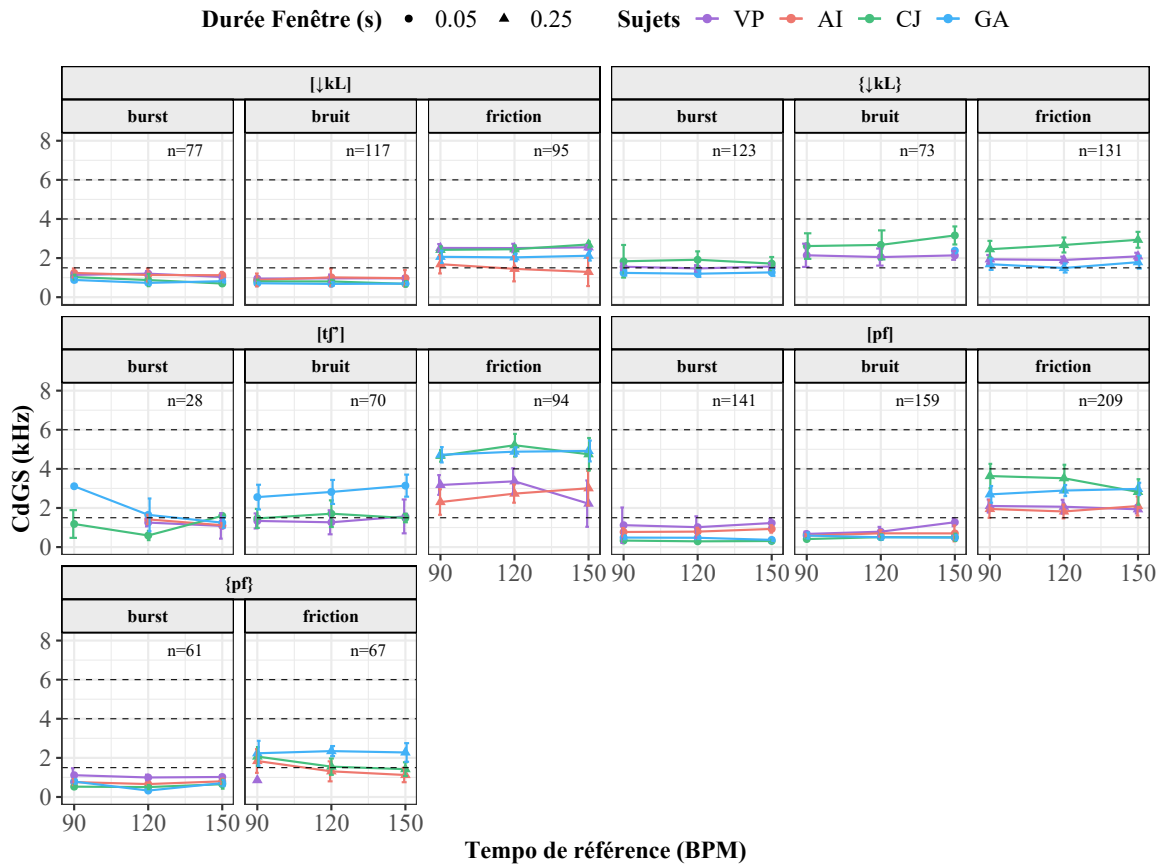


FIGURE 4.25 – Centre de Gravité Spectral (kHz) des caisses claires orales en fonction du tempo de référence (BPM).

donné que ce son a été produit dans le pattern  $[t' \widehat{ts'} \widehat{tj'} : \widehat{ts'} t' t' \widehat{ts'} \widehat{tj'} : \widehat{ts'}]$  et que le contexte adjacent est composé de sons dentaux et alvéolaires. Il peut s'agir d'un indice de chevauchement de la *Tish Snare* avec le *Kick Lingual* suivant. Le *Kick Lingual* a la particularité d'être produit avec rétroflexion et protrusion des lèvres qui ont pour effet d'allonger la cavité antérieure et d'abaisser l'énergie des hautes fréquences vers les basses fréquences. Chez le sujet AI, on note la tendance inverse, à savoir une augmentation du Centre de Gravité Spectral des fréquences moyennes-basses vers les fréquences moyennes-hautes. Cela indique une diminution de la longueur de la cavité antérieure. Il est possible que chez ce sujet, la *Tish Snare* soit influencée par le fait que  $[\widehat{ts'}]$  et  $[t']$  ont un lieu d'articulation plus antérieur. En d'autres termes, il est possible que la constriction de  $[\widehat{tj'}]$  soit produite plus en avant dans le conduit, dans la région dentale ou alvéolaire. Un étirement des lèvres pourrait réduire, en plus de l'avancement du lieu d'articulation, la longueur de la cavité antérieure.

Enfin, pour la *PF snare* et sa variante humming, on note aussi des comportements individuels différents. Pour  $[\widehat{pf}]$  on note que l'énergie tend à se concentrer autour de 2-3kHz à 150 BPM. Le tempo semble affecter la production des sujets CJ et GA. En effet, on note une diminution de 0.5kHz chez CJ et une augmentation de 0.5kHz chez GA. Ces tendances suggèrent un changement du geste labiodental en fonction du tempo. Pour la variante *humming*, on remarque chez AI et CJ un abaissement dans les basses fréquences du Centre de Gravité en fonction du tempo. Des données vidéos des lèvres avec une caméra haute vitesse permettraient de savoir si cette variabilité est liée au geste labiodental.

#### 4.4.1.3.2 Forme spectrale

La Figure 4.26 représente les variations de l'écart-type du Centre de Gravité Spectral des caisses claires orales en fonction du tempo. L'étendue du Centre de Gravité spectral est plus petite pour les bursts et bruits d'explosion en comparaison avec les bruits de frictions. En effet, pour les bursts et bruits d'explosion, l'écart-type varie entre 0.5-1kHz. Le coefficient d'aplatissement varie entre 5 et 160 et s'élève en moyenne à 40. Ceci suggère une pente spectrale abrupte. Le Centre de Gravité des bruits de friction s'étend sur une plage de fréquences entre 1-3kHz. Le *Kurtosis* est plus faible, il s'élève en moyenne à 9. Le spectre des phases de friction est donc plat.

On note toutefois des exceptions et des sons dont la pente spectrale est variable. On note par exemple une variabilité de l'écart-type plus important pour les sons  $[tʃ:]$ ,  $[pʃ:]$  et  $\{pʃ:\}$ . On note que l'écart type du Centre de Gravité de la *Tish Snare* est élevé, indiquant une enveloppe spectrale plate. Enfin, on remarque que pour la  $\{pʃ:\}$ , l'écart type se réduit en fonction du tempo chez les sujets CJ et AI indiquant une enveloppe spectrale plus abrupte.

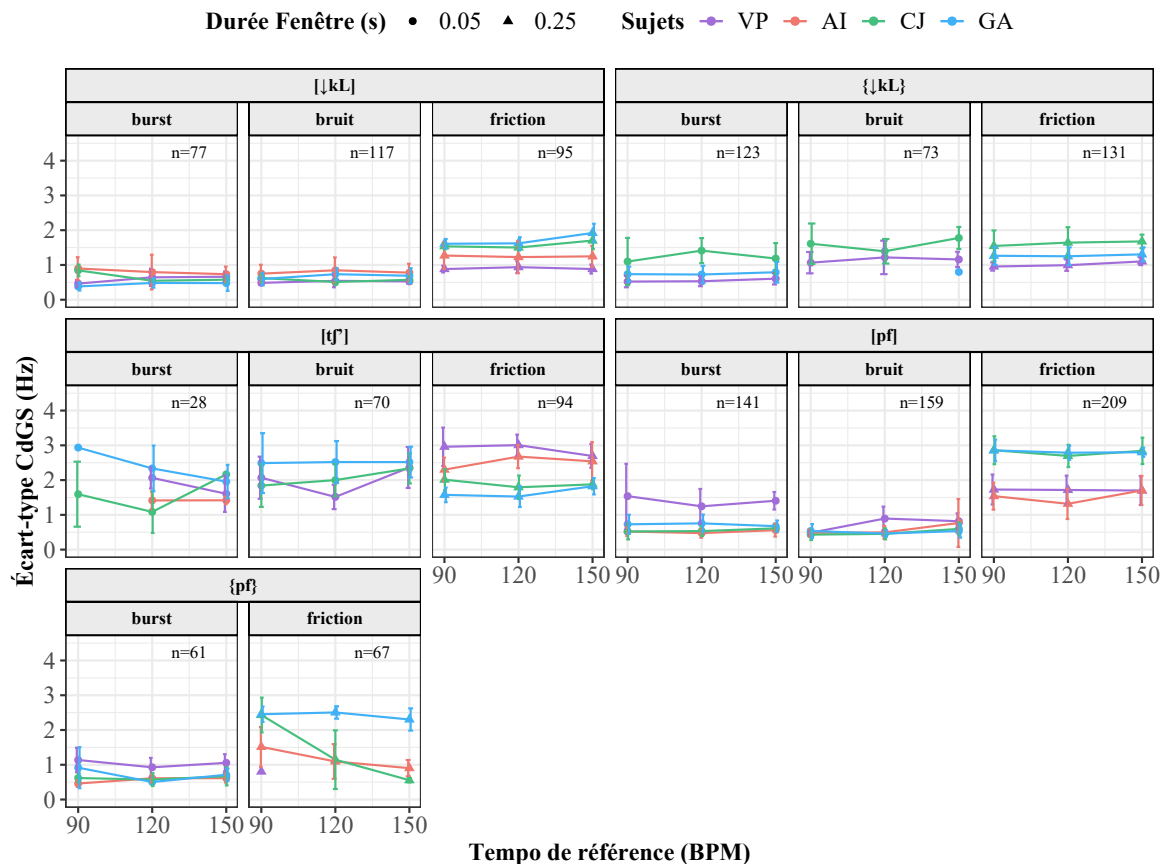


FIGURE 4.26 – Écart-type du Centre de Gravité spectral (kHz) caisses claires orales en fonction du tempo de référence (BPM).

#### 4.4.1.4 Caisses Claires Laryngales

La Figure 4.27 représente le Centre de Gravité Spectral en fonction du tempo. Les phases de voisement et de friction montrent une répartition différente. Le voisement prédomine dans les basses fréquences et la friction dans les moyennes et basses fréquences. Nous commencerons par comparer les phases de friction et nous terminerons par l'interprétation du Centre de Gravité du voisement et des données de F0 mesurée.

La répartition de l'énergie des bruits de friction diffère selon les sons. Pour  $[\hat{ʔ}]:$  l'énergie se trouve dans les fréquences moyennes, pour  $[\downarrow\hat{ʔ}]:$  l'énergie se trouve dans les fréquences moyennes-basses et pour  $[\hat{ʔ}h]:$  l'énergie se trouve dans les fréquences basses. La différence de répartition des variantes égressives et ingressives de la *ESH Snare* peut s'expliquer par la différence de directionnalité du flux d'air (cf. Chapitre 3). La cavité antérieure est plus longue dans le cas de la variante ingressive et le Centre de Gravité est donc plus bas comparé à la variante égressive. Les données de *Skewness* reflètent aussi cette différence. Quant à la *Cough Snare*, l'énergie se concentre dans les basses fréquences car, en l'absence de constriction supralaryngée, la répartition spectrale dépend de la configuration du conduit vocal au relâchement. Les valeurs de *Skewness* de la *Cough Snare* de AI indiquent une répartition plus asymétrique que les autres sujets. Ceci suggère que les spectres possèdent moins de composantes de hautes fréquences que les autres sujets. La même observation peut-être faite concernant  $[\downarrow\hat{ʔ}]:$  pour ce même sujet.

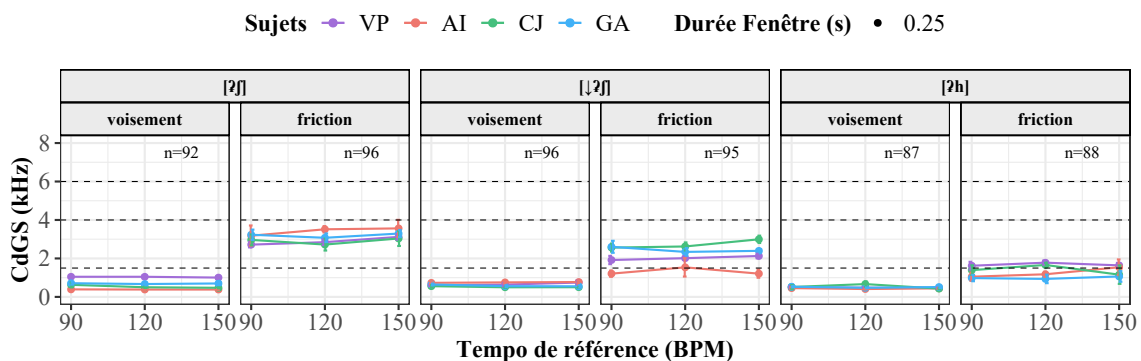


FIGURE 4.27 – Centre de Gravité Spectral (kHz) en fonction du tempo de référence (BPM) selon les sons, les événements acoustiques et les sujets.

Regardons à présent les données du Centre de Gravité (Figure 4.27) concernant le voisement ainsi que les données de F0 (Figure 4.28). Les valeurs de Centre de Gravité suggèrent une variabilité interindividuelle plus importante de la distribution spectrale au relâchement de  $[\hat{ʔ}]:$ . En effet, les valeurs varient entre 0.2 et 1kHz selon les sujets. Les données de F0 indiquent une fréquence fondamentale aux alentours de 250Hz pour les sujets VP, CJ et GA et aux alentours de 350Hz pour AI. Cependant, il y a un chevauchement des valeurs des sujets VP et AI. À 150BPM, la fréquence fondamentale moyenne des deux sujets est de  $\pm 350$ Hz. Le relâchement de  $[\downarrow\hat{ʔ}]:$  montre un Centre de Gravité similaire entre les sujets à  $\pm 0.5$ kHz. En revanche, les valeurs de F0 diffèrent selon les sujets, la F0 est de 500Hz pour les sujets VP, CJ et GA et de 750Hz pour AI. Enfin, le Centre de Gravité au relâchement de  $[\hat{ʔ}h]:$  indique une concentration de l'énergie dans les basses fréquences. Les valeurs moyennes de F0 varient entre 400-500Hz et les valeurs les plus élevées sont celles du sujet CJ. On note que les valeurs des sujets masculins et féminins sont proches voire égales à l'exception des valeurs de F0 du sujet AI pour

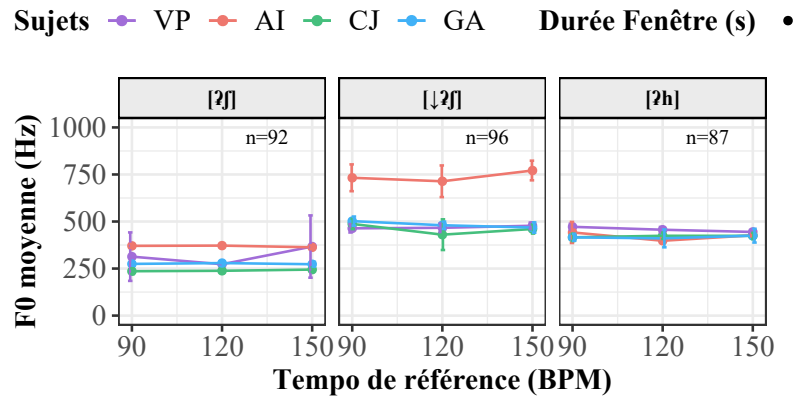


FIGURE 4.28 – F0 moyenne (Hz) des caisses claires laryngales en fonction du tempo de référence (BPM).

[↓ʔ]:]. Nous pensons cependant que ces différences sont liées principalement à des différences de configuration laryngale au relâchement. Alors que les sujets VP, CJ et GA produisent une occlusion épilaryngale, le sujet AI produit une occlusion ventriculaire sans constriction du tube épilaryngé. La longueur des bandes ventriculaires et des plis vocaux au relâchement est plus longue que chez les autres sujets.

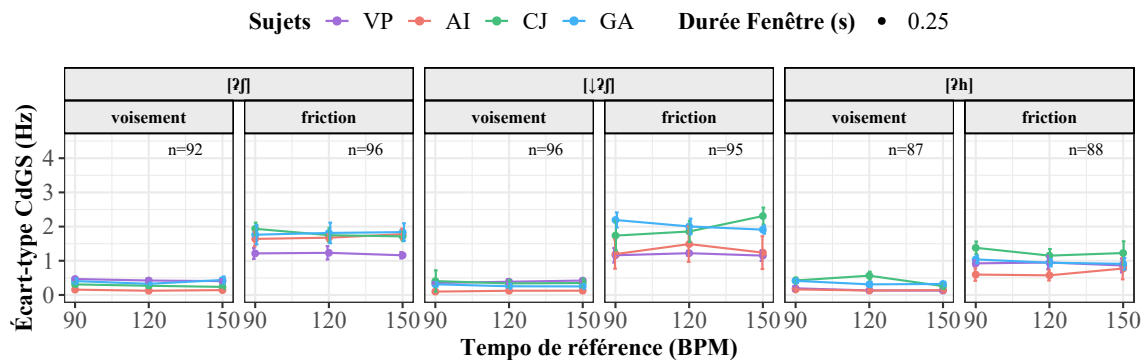


FIGURE 4.29 – Écart-type du Centre de Gravité spectral (kHz) des caisses claires laryngales en fonction du tempo de référence (BPM) selon les sons, les événements acoustiques et les sujets.

La Figure 4.29 illustre les variations de l'écart-type des caisses claires laryngales en fonction du tempo. On note en premier lieu que l'écart-type est plus faible pour le voisement en comparaison au bruit de friction. Le coefficient d'aplatissement est de 400. La pente spectrale de la phase de voisement est donc abrupte. La variation de l'écart-type du bruit de friction diffère selon les sons et les sujets. En effet, on note des différences entre [ʔ] et [↓ʔ]. Dans le cas de la variante ingressive, l'écart-type est plus variable entre les sujets, dans les deux cas il indique une répartition de l'énergie entre 1 et 2.5kHz. Pour la *Cough Snare*, l'écart-type est plus faible. Le coefficient d'aplatissement moyen est de 10. Pris ensemble, ces résultats montrent que le spectre a une enveloppe plate. On note un léger effet du tempo sur la variation de l'écart-type de la *ESH Snare Inward* des sujets CJ et GA. Cependant, cette variation est faible ( $\pm 0.5$ kHz).

#### 4.4.1.5 Rolls

Pour conclure l'analyse de la répartition spectrale, regardons les données des *Rolls* sur les Figures 4.30 et 4.31. L'énergie des *Rolls* se concentre en dessous d'1kHz et ne semble pas varier sous l'effet du tempo. Le coefficient d'asymétrie suggère des différences de répartition de l'énergie selon les sons et les sujets. Chez les sujets CJ et GA, [↓B<sup>1</sup>] possède plus d'énergie dans les basses fréquences en comparaison au *Lips Roll Humming* et au *Clic Roll*. La répartition est plus variable que chez les autres sujets.

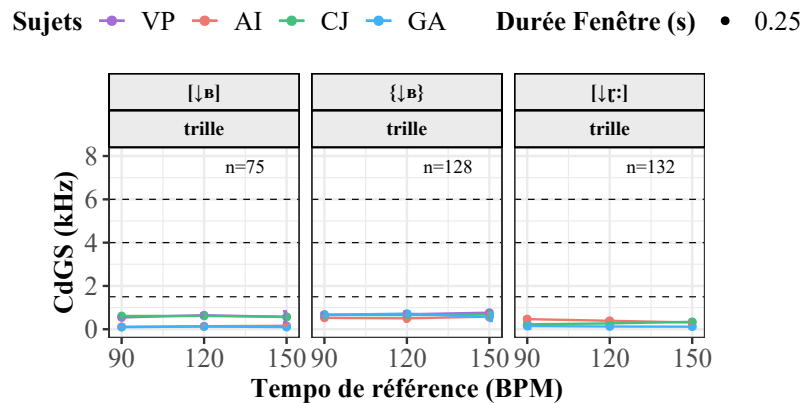


FIGURE 4.30 – Centre de Gravité Spectral (kHz) des *Rolls* en fonction du tempo de référence (BPM).

La Figure 4.31 montre les variations de l'écart-type du Centre de Gravité spectral en fonction du tempo. L'écart-type est faible, suggérant que l'énergie se concentre sur une plage de  $\pm 0.5$ kHz autour du Centre de Gravité.

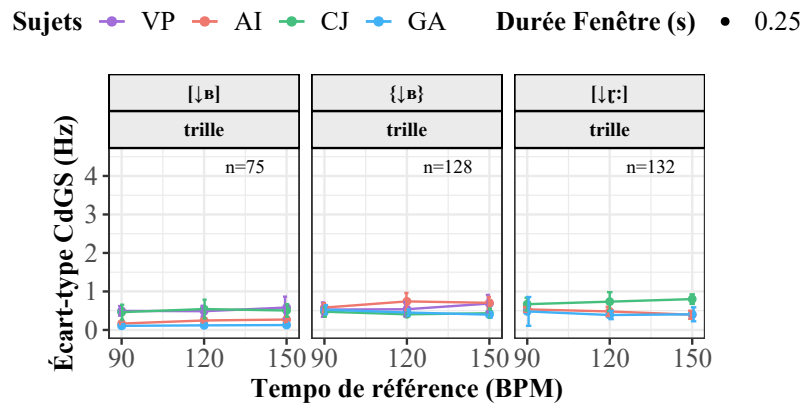


FIGURE 4.31 – Écart-type du Centre de Gravité spectral (kHz) *Rolls* en fonction du tempo de référence (BPM).



#### 4.4.2 Effet du tempo sur les moments spectraux en fonction du contexte adjacent

Nous allons conclure l'analyse des effets du tempo sur la production en regardant comment varie, en fonction du tempo, les moments spectraux de  $[\widehat{ts}']$  dans différentes positions et différents patterns. Pour faciliter la comparaison entre les sujets, les positions et les patterns, nous focaliserons l'analyse sur le bruit de friction de  $[\widehat{ts}']$ . La Figure 4.32 représente les variations du Centre de Gravité spectral du charleston  $[\widehat{ts}']$  en fonction du tempo. La Figure 4.33 représente les variations de l'écart-type du Centre de Gravité en fonction du tempo.

Chez le sujet VP, on note déjà à 90BPM des différences systématiques du Centre de Gravité spectral de  $[\widehat{ts}']$  dans différents patterns. En effet, si l'on regarde seulement les données à 90BPM, on peut remarquer que dans les BP  $[\widehat{b} \widehat{ts}' \widehat{t} \widehat{f} : \widehat{ts}' \widehat{b} \widehat{b} \widehat{ts}' \widehat{t} \widehat{f} : \widehat{ts}']$ ,  $[\widehat{p}' \widehat{ts}' \downarrow \widehat{kl} : \widehat{ts}' \widehat{p}' \widehat{p}' \widehat{ts}' \downarrow \widehat{kl} : \widehat{ts}']$ ,  $[\downarrow \widehat{p} \widehat{ts}' \widehat{pf} : \widehat{ts}' \downarrow \widehat{p} \downarrow \widehat{p} \widehat{ts}' \widehat{pf} : \widehat{ts}']$  et  $[\downarrow \widehat{p} \downarrow \widehat{b}^1 : \widehat{ts}' \widehat{t} \widehat{h} : \widehat{ts}' \downarrow \widehat{p} \downarrow \widehat{b}^1 : \widehat{ts}' \widehat{t} \widehat{h} : \widehat{ts}']$ , le Centre de Gravité spectral se trouve dans les moyennes fréquences quelle que soit la position. À l'inverse, on note que dans le  $[\widehat{t}' \widehat{ts}' \widehat{t} \widehat{f} : \widehat{ts}' \widehat{t}' \widehat{t}' \widehat{ts}' \widehat{t} \widehat{f} : \widehat{ts}']$ , le Centre de Gravité est toujours plus haut que dans les autres patterns.

Les données de VP suggèrent que la taille de la cavité antérieure au relâchement de  $[\widehat{ts}']$  varie selon contexte adjacent. On peut en conclure que l'organisation spatio-temporelle des gestes de production, chez VP, dépend de la structure phonétique des patterns produits. Les différences d'écart-type (Figure 4.33) sont moins marquées entre les patterns. On note tout de même chez VP que dans le BP  $[\widehat{t}' \widehat{ts}' \widehat{t} \widehat{f} : \widehat{ts}' \widehat{t}' \widehat{t}' \widehat{ts}' \widehat{t} \widehat{f} : \widehat{ts}']$ , l'écart-type tend à être plus faible suggérant une forme spectrale abrupte et montante. On observe un effet du tempo sur les variations de Centre de Gravité spectral. En effet, à quelques exceptions, le Centre de Gravité spectral s'abaisse dans les moyennes fréquences voir, comme par exemple en position 4, dans les basses fréquences. En position 4, il est clair que l'abaissement du Centre de Gravité est plus important que dans les autres positions. Les données de VP montrent l'influence du contexte adjacent sur la réalisation acoustique du son  $[\widehat{ts}']$  et des effets du tempo sur les réalisations acoustiques.

Chez le sujet AI, on ne note pas de différences de la répartition spectrale à 90BPM aussi systématiques que chez VP. En effet, on note qu'à 90BPM, le Centre de Gravité moyen de  $[\widehat{ts}']$  est similaire entre les BP à l'exception des charlestons en position 4. Dans cette position, on note que l'énergie prédomine dans les très hautes fréquences ( $F > 6\text{kHz}$ ) pour le BP  $[\widehat{b} \widehat{ts}' \widehat{t} \widehat{f} : \widehat{ts}' \widehat{b} \widehat{b} \widehat{ts}' \widehat{t} \widehat{f} : \widehat{ts}']$ , dans les hautes fréquences pour les BP  $[\widehat{b} \widehat{ts}' \downarrow \widehat{t} \widehat{f} : \widehat{ts}' \widehat{b} \widehat{b} \widehat{ts}' \downarrow \widehat{t} \widehat{f} : \widehat{ts}']$ ,  $[\widehat{p}' \widehat{ts}' \widehat{pf} : \widehat{ts}' \widehat{p}' \widehat{p}' \widehat{ts}' \widehat{pf} : \widehat{ts}']$  et  $[\widehat{p}' \widehat{ts}' \downarrow \widehat{kl} : \widehat{ts}' \widehat{p}' \widehat{p}' \widehat{ts}' \downarrow \widehat{kl} : \widehat{ts}']$  et dans les moyennes fréquences pour les BP  $[\downarrow \widehat{p} \widehat{ts}' \widehat{pf} : \widehat{ts}' \downarrow \widehat{p} \downarrow \widehat{p} \widehat{ts}' \widehat{pf} : \widehat{ts}']$  et  $[\downarrow \widehat{p} \downarrow \widehat{b}^1 : \widehat{ts}' \widehat{t} \widehat{h} : \widehat{ts}' \downarrow \widehat{p} \downarrow \widehat{b}^1 : \widehat{ts}' \widehat{t} \widehat{h} : \widehat{ts}']$ . On note cependant un effet du tempo sur les variations de Centre de Gravité qui diminue vers les moyennes fréquences mais qui dans certains cas augmente vers les très hautes fréquences (cf. position 4 - BP  $[\downarrow \widehat{p} \widehat{ts}' \widehat{pf} : \widehat{ts}' \downarrow \widehat{p} \downarrow \widehat{p} \widehat{ts}' \widehat{pf} : \widehat{ts}']$ ). Un déplacement vers les très hautes fréquences suggère une diminution de la longueur de la cavité antérieure. On note des comportements particuliers en position 9 où à 120BPM, les valeurs de Centre de Gravité se concentrent vers 6kHz puis les valeurs diminuent différemment à 150BPM. Les patterns  $[\widehat{t}' \widehat{ts}' \widehat{t} \widehat{f} : \widehat{ts}' \widehat{t}' \widehat{t}' \widehat{ts}' \widehat{t} \widehat{f} : \widehat{ts}']$  et  $[\downarrow \widehat{p} \downarrow \widehat{b}^1 : \widehat{ts}' \widehat{t} \widehat{h} : \widehat{ts}' \downarrow \widehat{p} \downarrow \widehat{b}^1 : \widehat{ts}' \widehat{t} \widehat{h} : \widehat{ts}']$  ne suivent pas cette tendance. Notons la forme en « éventail » des trajectoires du Centre de Gravité à 150BPM. Les données d'écart-type du Centre de Gravité spectral tendent à augmenter et indiquent un aplatissement de l'enveloppe spectrale à mesure que le tempo accélère.

Chez le sujet CJ, on ne note pas de différence du Centre de Gravité en fonction du BP à 90BPM. Toutes les valeurs se situent dans les hautes fréquences. Le Centre de Gravité s'abaisse à mesure que le tempo accélère, cette tendance diffère selon les patterns et les positions. Par exemple en position 4, l'abaissement du Centre de Gravité spectral est systématique. Pour les

BP [ṭ' tṣ' tʃ' : tṣ' ṭ' ṭ' tṣ' tʃ' : tṣ'] et [↓p tṣ' pf̣ : tṣ' ↓p ↓p tṣ' pf̣ : tṣ'], le Centre de Gravité s'abaisse dans les basses fréquences ce qui suggère une augmentation plus importante de la longueur de la cavité antérieure par rapport aux autres patterns. L'écart-type du Centre de Gravité spectral tend à augmenter en fonction du tempo. L'augmentation est minimale ( $\pm 0.5\text{kHz}$ ) et suggère une forme spectrale légèrement plus aplatie à 150BPM. Notons que l'écart-type se réduit fortement en position 4 pour les patterns [ṭ' tṣ' tʃ' : tṣ' ṭ' ṭ' tṣ' tʃ' : tṣ'] et [↓p tṣ' pf̣ : tṣ' ↓p ↓p tṣ' pf̣ : tṣ'], patterns pour lesquels nous avons montré un abaissement du Centre de Gravité des hautes fréquences vers les basses fréquences. Ces changements majeurs de la répartition fréquentielle et de la forme du spectre suggèrent un chevauchement plus important avec les sons adjacents à mesure que le tempo accélère.

Pour terminer chez GA, hormis le charleston en position 4 dans le pattern [ṭ' tṣ' tʃ' : tṣ' ṭ' ṭ' tṣ' tʃ' : tṣ'], à 90BPM l'énergie se concentre dans les hautes fréquences. Encore une fois, on remarque un abaissement important du Centre de Gravité dans les moyennes fréquences. Les BP [ḅ' tṣ' ↓ʃ' : tṣ' ḅ' ḅ' tṣ' ↓ʃ' : tṣ'] et [↓p ↓ḅ' : tṣ' ʃḥ : tṣ' ↓p ↓p ↓ḅ' : tṣ' ʃḥ : tṣ'] ne suivent pas cette tendance. Cela suggère qu'à mesure que le tempo accélère, GA varie peu dans la réalisation du geste articulatoire et de la coordination avec les gestes adjacents. Contrairement aux autres sujets, GA maintient l'énergie dans les hautes fréquences du spectre à mesure que le tempo accélère. En revanche, on note plus de variabilité de l'étendue du Centre de Gravité spectral. Toutefois, les variations d'écart-type se situent au-dessus de 2kHz. Cela suggère des degrés d'aplatissement différents selon la position et le pattern.

Nous ne pouvons pas entrer dans les détails des variations de réalisation acoustique car il s'agirait d'un travail de recherche sur les phénomènes de coordination articulatoire à part entière. De plus, nous ne disposons seulement de données IRM à 90BPM pour VP. Toutefois, les variations de longueur de la cavité antérieure peuvent être causées par un nombre limité d'ajustements articulatoires. Il peut s'agir d'une influence des gestes labiaux adjacents étant donné que les grosses caisses sont des occlusives bilabiales sauf le *Kick Lingual* qui est produit avec une protrusion labiale. Il est possible que l'allongement de la cavité antérieure soit dû à un recul du lieu d'articulation. Par exemple, lorsque [tṣ'] est suivi de [↓kḷ:], particulièrement à 150BPM, on peut penser que pour réduire la distance spatiale entre les cibles dentales et vélaires, les sujets reculent le lieu de la constriction dentale. Dans les cas où la longueur de la cavité antérieure se réduit et où le Centre de Gravité augmente, on peut penser à des ajustements inverses, c'est-à-dire un avancement du point d'articulation et un étirement des lèvres.

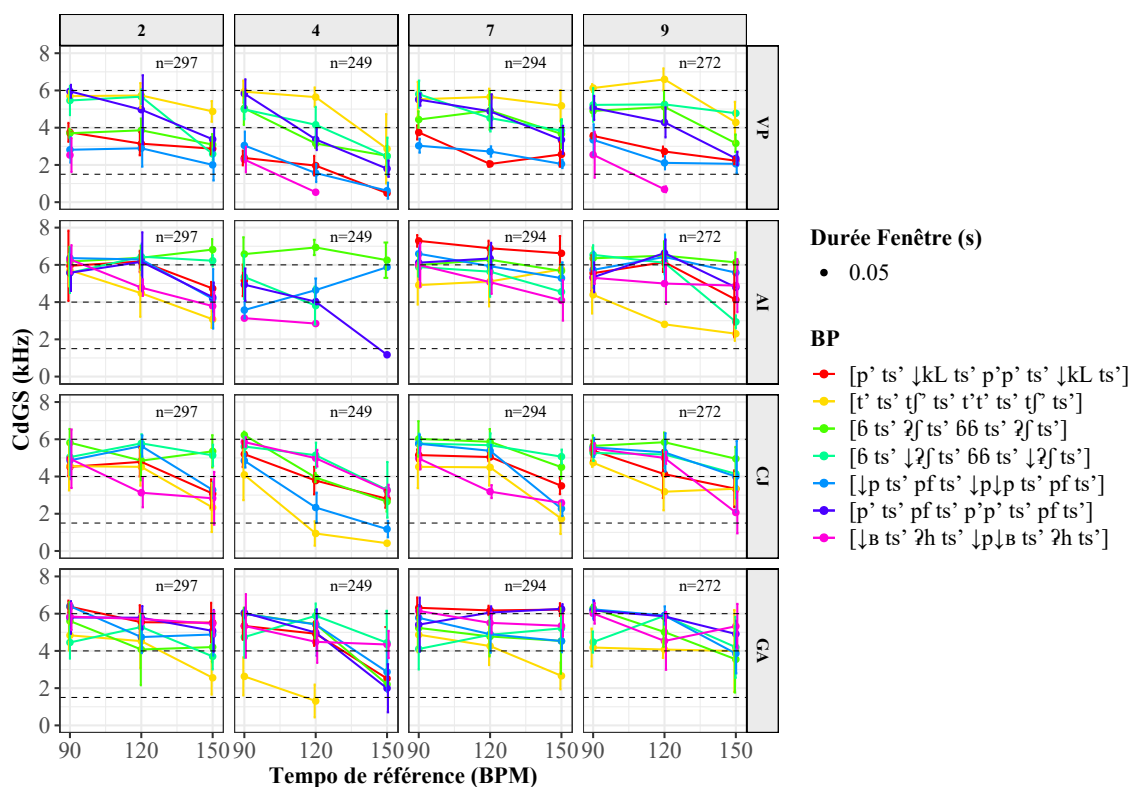


FIGURE 4.32 – Centre de Gravité spectral (kHz) du bruit de friction de  $[\widehat{ts}]$  en fonction du tempo. Le nombre d'échantillons correspond au nombre de fenêtres d'analyse de 0.05ms.

#### 4.4.3 Synthèse de l'analyse acoustique

L'analyse acoustique s'est intéressée aux variations des moments spectraux en fonction du tempo de référence. Nous avons quantifié la répartition spectrale des sons grâce au Centre de Gravité Spectral et au *Skewness* (ou coefficient d'asymétrie). Nous avons ensuite quantifié la forme du spectre grâce à l'écart-type du Centre de Gravité Spectral et du *Kurtosis* (ou coefficient d'aplatissement). Enfin, nous avons analysé les effets du contexte adjacent en analysant les moments spectraux des charlestons fermés.

La quantification de la répartition spectrale a confirmé la description acoustique des catégories de sons du Chapitre 3. Les grosses caisses et les *Rolls* possèdent de l'énergie dans les basses fréquences, les caisses claires dans les moyennes fréquences et les charlestons dans les hautes fréquences. De plus, nous avons vu que pour les grosses caisses et les *Rolls*, le Centre de Gravité était stable à mesure que le tempo accélérât. Ces catégories sont principalement constituées de sons produits avec les lèvres, à l'exception de  $[t']$  et  $[\downarrow t:]$ . L'analyse des charlestons a révélé un abaissement du Centre de Gravité vers les moyennes fréquences. Nous avons aussi noté des variations de répartition des caisses claires. Les caisses claires vélares montraient moins de variation (en fonction du tempo) de la répartition spectrale que les caisses claires labiodentales  $[\widehat{pf}]$  et  $\{\widehat{pf}\}$ . Nous avons trouvé des indices acoustiques de chevauchement de  $[\widehat{t'}:]$  avec  $[t']$  et  $[\widehat{ts}]$  chez les sujets VP et AI. Enfin, nous avons aussi montré des différences de répartition acoustique et de F0 entre les caisses claires laryngales  $[\widehat{ʔ}:]$  et  $[\downarrow \widehat{ʔ}:]$ .

L'analyse de l'écart-type de Centre de Gravité spectral et le *Kurtosis* nous ont permis d'inférer des différences de pente spectrale et de quantifier la forme du spectre. Les grosses caisses et les *Rolls* ont une forme spectrale stable qui ne varie pas en fonction du tempo. Les charles-

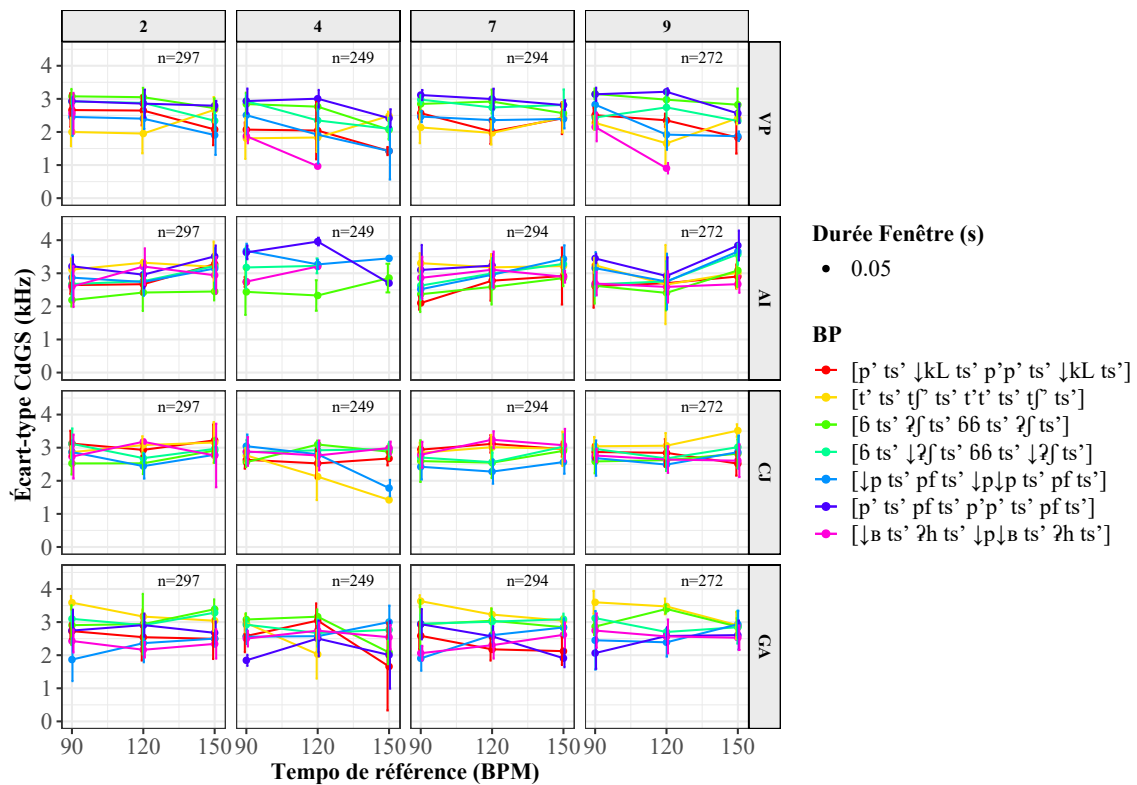


FIGURE 4.33 – Écart-type du Centre de Gravité spectral (kHz) du bruit de friction de  $[\widehat{ts}']$  en fonction du tempo. Le nombre d'échantillons correspond au nombre de fenêtres d'analyse de 0.05ms.

tons et les caisses claires orales, à l'exception de  $[\downarrow kL:]$ , ont montré plus de variations de forme spectrale. Enfin, concernant les caisses claires laryngales, nous avons trouvé que le voisement avait une enveloppe et une répartition spectrale moins variable que le bruit de friction. Ces résultats suggèrent une variation plus importante du geste supralaryngal, en comparaison au geste laryngal, pour ces sons.

## Synthèse

Le Chapitre 4 avait pour objectif de décrire, quantifier et expliquer la variation de la production en fonction du tempo. Nous avons trouvé peu d'erreur de production et nous avons vu que la production était plutôt stable. De prochaines études sur ce paradigme devront complexifier la structure des *Beat Patterns*, par exemple en ajoutant des basses vocales entre les percussions.

Notre hypothèse initiale est partiellement confirmée. Nous avons montré une tendance des sons et des intervalles intergestuels à se réduire. Dans l'analyse acoustique du charleston fermé  $[\widehat{ts}']$ , nous avons montré des phénomènes de variation contextuels possiblement déclenchés par un chevauchement temporel des gestes. Le tempo a donc un effet sur l'organisation temporelle des gestes dans le pattern. Cependant, les variations d'amplitude des gestes d'initiation ne présentaient pas un pattern de réduction similaire à celui de la durée. Concernant la pression, nous avons vu qu'elle ne se réduisait pas forcément à mesure que le tempo accélérât. Nous avons suggéré que la régulation de la pression dépendait de stratégies individuelles de régulation de l'intensité acoustique.

Nos données ont aussi montré l'influence d'autres facteurs que le tempo sur la variation de la

réalisation de sons. En premier lieu, nous avons démontré des différences de durée des sons et de coordination gestuelle entre mécanismes d'initiation. L'analyse du sous-ensemble des charlestons [tʰs'] a révélé, chez au moins 2 sujets, des interactions entre les gestes d'initiation d'une part, et entre le geste d'initiation et d'articulation d'autre part. La pression tend à s'affaiblir quand ces mécanismes interagissent. Il est toutefois possible de surpasser cette contrainte comme l'ont montré les données du sujet GA. Nous avons aussi observé un effet systématique de la position 4 sur le degré de réduction de la durée et de la pression des charlestons.

Toutes nos analyses se sont efforcées de faire ressortir les patterns de variation individuels. La stratégie d'accélération de la vitesse de production se traduit par la réduction temporelle des sons et des intervalles intergestuels à divers degrés. Cependant, dans l'analyse acoustique on a noté que chez VP, principalement, et chez CJ, dans une moindre mesure, la réduction temporelle des sons n'était pas suffisante pour limiter le chevauchement temporel entre les gestes. Chez le sujet AI, on a observé moins d'indices acoustiques de chevauchement temporel. On a aussi noté que la réduction de la durée des sons et des intervalles intergestuels était plus variable. Cela signifie que le degré de chevauchement des gestes est plus variable. Cela pourrait expliquer pourquoi le sujet AI est plus variable dans sa vitesse de production. GA semble avoir beaucoup d'agilité dans sa stratégie de production. C'est d'ailleurs le sujet qui a la vitesse de production la plus proche du tempo de référence. Il réduit la durée des sons et des intervalles intergestuels mais nous n'avons pas pour autant observé d'indices de chevauchement temporel chez ce sujet. Il arrive donc à limiter le chevauchement temporel des gestes tout en accélérant.

# Chapitre 5

## Discussion Générale

« The antropophonic perspective invites us to look for the constraints that operate in the evolutionary selection of speech sounds. »

---

Björn Lindblom<sup>1</sup>

Nous allons à présent conclure cette thèse en discutant des implications théoriques de nos résultats. Les 2 précédents chapitres ont montré comment le système de production réorganisait son activité à partir de nouveaux patrons combinatoires et de nouveaux patrons de coordination articulatoire (Chapitre 3). Nous avons aussi montré comment variaient les *Beat Patterns* en fonction du tempo de référence (Chapitre 4). La description du HBB dans le cadre actuel n'est pas sans poser problèmes. Par exemple, nous avons proposé dès l'introduction un modèle gestuel de l'initiation plus à même de décrire et expliquer les mécanismes qui initient les flux d'air. Les limites du cadre actuel sont dues au fait que les connaissances accumulées en phonétique viennent de la recherche en linguistique. Or, si l'on veut que la phonétique puisse expliquer l'ensemble des phénomènes de production vocale de l'espèce humaine, il faut élargir le cadre aux productions non-linguistiques. Nous proposons de placer la phonétique dans un cadre « *anthropophonique* » comme l'ont proposé Catford (1977) puis Lindblom (1990). De même, nous plaçons le Human Beatbox dans un cadre anthropophonique.

### 5.1 Pratique artistique et objet d'étude scientifique

Bien avant d'être un objet d'étude scientifique, le Human Beatbox est une pratique artistique. Dans cette thèse, nous nous sommes principalement intéressés à l'objet scientifique. Toutefois, nous avons rapporté les propos des sujets à plusieurs reprises pour éclairer nos résultats de leurs connaissances. Ces connaissances nous viennent aussi des questionnaires qui leur ont été adressés. Les questionnaires des 5 sujets initiaux sont disponibles dans les ANNEXES 1 à 5. Avec ces questionnaires, on cherchait à collecter des informations sur le profil des participant-es, leur pratique pratique du HBB, leurs connaissances méta-phonétiques et les risques potentiels pour la santé.

Un point qui nous intéressait particulièrement était la conscience phonétique de leur pratique. Il en ressort que les sujets sont conscients de mobiliser l'ensemble de leurs articulateurs : larynx, langue, lèvres ... Selon des échanges oraux, il semblerait qu'ils aient une connaissance

---

1. Lindblom, B. (1990). On the notion of "possible speech sound". *Journal of phonetics*, 18(2), 135-152.

moins fine de l'activité articulatoire du larynx en comparaison avec l'activité articulatoire linguale ou labiale par exemple. Chaque artiste travaille différemment. Certain-es travaillent principalement sur les sensations et retours auditifs (*feedback auditif*) et d'autres travaillent davantage les sensations et retours kinesthésiques (*feedback proprioceptif*).

Beatboxer ce n'est pas seulement bouger ses articulateurs. C'est aussi se mouvoir soi-même. On peut lire dans les questionnaires qu'aucun des sujets ne reste statique en performant. Ils accompagnent le rythme et la mélodie avec leurs mains et peuvent dans certaines occasion imiter des musiciens jouant d'un instrument. La posture et la position de la tête sont tout aussi importantes. Par exemple, on peut projeter sa tête vers l'avant ou rentrer la tête dans ses épaules. À ce propos, les sujets VP et CJ nous ont indiqué qu'il était plus facile de produire des basses vocales en rentrant la tête dans ses épaules. La posture peut aider pour la phonation et la respiration selon eux. Le Human Beatbox est une pratique incarnée (*embodied beatboxing*). C'est une pratique scénique à la fois vocale et corporelle.

Une question brulait nos lèvres : pouvez-vous produire tous les sons beatboxés ? Aucun des 5 sujets n'a répondu oui. Certain-es ont répondu qu'il était impossible de produire tous les sons du Human Beatbox. En discutant avec les sujets, il est ressorti que la pratique est en permanente évolution depuis la *old school* à la *new school*. Les artistes ne cessent d'explorer les capacités du conduit vocal humain en composant, décomposant, recomposant des sonorités riches et variées. C'est toute une fanfare que les beatboxeurs ont dans la gorge.

Est-ce que la pratique intensive et professionnelle du Human Beatbox présente des risques pour la santé ? C'est une question susceptible d'intéresser médecins, orthophonistes, chercheurs et artistes. C'est une question légitime à la vue des données que nous présentons dans cette thèse. En effet, les beatboxeurs et les beatboxeuses déploient des efforts physiques intenses pour produire des sons. Nous avons posé la question à nos sujets. 3 sujets sur 5 rapportent avoir consulté un médecin en lien avec leur pratique vocale. Il ressort que la production des basses vocales, en particulier les variantes *inward*, peut provoquer des douleurs et des irritations « vocales ». Il est aussi fait état de douleurs « musculaires » et « articulatoires », notamment en ce qui concerne les mouvements latéraux de la mâchoire ([↓k̂L:] [↓k̂L] [↓B̂<sup>l</sup>:]). Ces différents éléments suggèrent que la pratique du Human Beatbox n'est pas totalement sans danger pour la santé vocale. Comme le suggère Paroni (2022) (orthophoniste de formation), il est primordial de connaître les risques potentiels de cette pratique pour la santé avant d'utiliser le beatbox comme un outil de rééducation orthophonique.

Une approche purement scientifique du Human Beatbox ne nous aurait pas permis de comprendre cette pratique vocale et sociale. C'est d'ailleurs grâce à la coopération avec le sujet VP que nous avons pu mettre sur pied un protocole expérimental fonctionnel et adapté aux contraintes musicales et biomécaniques du Human Beatbox. Il a été en quelque sorte notre interprète en langage beatboxé. L'étude scientifique du Human Beatbox ne saurait se faire sans ses beatboxeurs et ses beatboxeuses. Mais revenons-en à nos moutons scientifiques car le Human Beatbox n'est pas dénué d'implications théoriques pour la phonétique.

## 5.2 Approche anthropophonique et primitives phonétiques

Dans l'introduction de *Fundamental problems in phonetics*, Catford écrit que le titre de son ouvrage aurait dû être *Anthropophonics : the phonetic categorization of human sounds*. Il ajoute avoir été « dissuadé » de choisir un titre « excentrique » qui pourrait décourager certains lecteurs. Catford emprunte le terme d'*anthropophonie* au linguiste Jan Ignacy Niecislaw Bau-

douin de Courtenay (1845-1929)<sup>2</sup>. Pour Baudouin de Courtenay la phonétique ou phonologie (il emploie les 2 termes comme synonymes) se compose de 2 parties : l'une anthropophonique (physiologique), l'autre psycho-phonétique (cf l'anthologie de Stankiewicz, 1972<sup>3</sup>). Pour lui, contrairement à la psychophonétique, l'anthropophonie (ou anthropophonétique) n'appartient pas exclusivement à la linguistique. Il propose même d'élargir l'anthropophonétique à la communication animale (zoophonétique) ce qui aujourd'hui correspondrait à la bioacoustique.

Le concept d'anthropophonie peut se définir comme l'étude du potentiel humain de production vocale. Cette approche ne s'intéresse donc pas seulement aux sons et phonèmes d'une langue mais à tous les sons qu'il est possible de produire avec le conduit vocal humain. Lindblom (1990) revient sur cette approche dans son article *On the notion of possible speech sounds*. La Figure 5.1 illustre la conception de Lindblom sur les sons possibles de la parole.

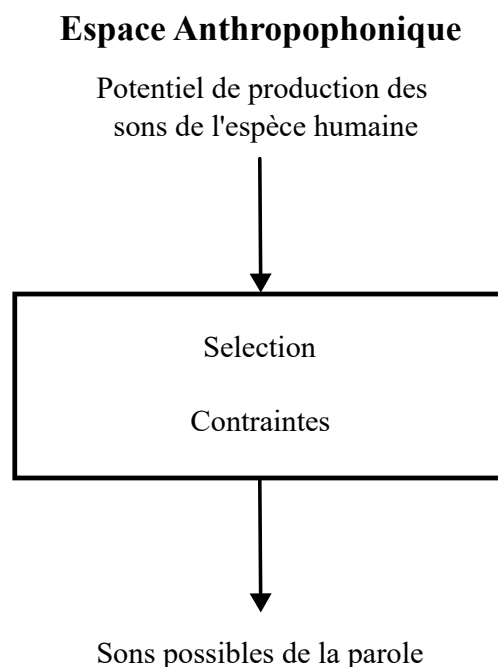


FIGURE 5.1 – Approche anthropophonique de la production des sons (Catford, 1977; Lindblom, 1990)

Dans l'approche anthropophonique les considérations phonologiques et linguistiques ne sont pas centrales lorsqu'on étudie des sons (qu'ils soient ou non linguistiques). Ce qui est central c'est la théorie phonétique fondée sur des lois de la physique. Cette approche invite à repenser les liens entre la phonétique et la phonologie ainsi qu'à repenser les primitives des systèmes sonores linguistiques ou non-linguistiques comme le Human Beatbox.

« Dans la langue il n'y a que des différences » écrivait Ferdinand de Saussure dans son cours de linguistique générale (Saussure, 1916, p.166). Ce sont ces différences qui permettent aux locuteurs d'une langue de communiquer à l'aide d'un système de différences. Phonéticiens et phonologues ont cherché à décrire la nature et le fonctionnement des différents contrastes dans les langues du monde. Les données linguistiques ont ensuite permis de s'interroger sur la nature

2. Pour connaître les positions et les contributions de Baudouin de Courtenay aux sciences du langage voir Adamska-Sałaciak (1998) et Alpatov (2021)

3. Stankiewicz, E. (Éd.). (1972). *A Baudouin de Courtenay Anthology : The Beginnings of Structural Linguistics*. Indiana University Press. <https://muse.jhu.edu/book/84822>.



des primitives phonologiques. Aujourd'hui, deux approches s'opposent : les traits distinctifs et les gestes articulatoires.

La notion de traits distinctifs a été proposée par Jakobson, Fant et Halle (Jakobson et al., 1951). Dans leur approche, les traits sont des indices acoustiques qui permettent d'établir des contrastes lexicaux. Depuis les années 1950, les traits distinctifs ont connus des évolutions. D'un côté, Chomsky et Halle (1968) proposent une conception linguistique et universelle des traits distinctifs. De l'autre côté, Stevens (1989) en proposant la théorie quantale replace les traits dans leur contexte acoustique. Dans la théorie quantale, les traits ne sont seulement pas définis par des observations empiriques mais par la théorie phonétique de production et de perception de la parole.

La notion de geste articulatoire (comme primitive phonologique et non dans le sens d'un mouvement articulatoire) a été introduite par Catherine Browman et Louis Goldstein dans leur théorie de la *phonologie articulatoire* (Browman & Goldstein, 1989, 1992). Cette théorie s'inscrit dans la lignée de la *théorie motrice de la perception de la parole* de Liberman et Mattingly (1985). La théorie motrice fait l'hypothèse d'un module (perceptif) spécialisé dans la détection des gestes articulatoires. Dans la *phonologie articulatoire*, les primitives sont définies comme des gestes articulatoires. Ces unités sont abstraites et représentent des constriction statiques dans le conduit vocal. La dynamique de la production est modélisée mathématiquement avec le modèle *Task Dynamic* (Saltzman & Kelso, 1987).

Et pour le Human Beatbox, traits ou gestes ? Le Chapitre 3 a démontré que les beatboxeurs et les beatboxeuses manipulent des mécanismes de production similaires à ceux observés dans les langues (mode d'articulation, lieu d'articulation). Étant donné que les deux systèmes utilisent le conduit vocal, il n'est pas surprenant de retrouver les mêmes mécanismes de production. Cela implique, selon nous, que les primitives ne sont pas spécifiques aux systèmes linguistiques : elles sont phonétiques et universelles.

Il est difficilement concevable que le HBB et la parole utilisent des primitives différentes. Prenons l'exemple des différences entre la planète Terre et la planète Saturne. La Terre est une planète rocheuse dont l'atmosphère est principalement composée d'azote et d'oxygène. Saturne est une planète gazeuse dont l'atmosphère est principalement composée de dihydrogène et d'hélium. Devons-nous en conclure que les composantes chimiques des deux planètes proviennent de deux ensembles différents ? La réponse est évidemment non. Les éléments chimiques des deux planètes font partie du même ensemble mais la combinaison des éléments chimiques est différente pour diverses raisons. Il revient aux chercheurs d'en expliquer les raisons.

Il en est de même pour le HBB et la parole. On ne saurait expliquer les deux systèmes indépendamment. Le HBB est un système musical où des personnes imitent des *beatbox* (boîtes à rythme) qui, elles-mêmes, imitent des sonorités d'instruments. Elles peuvent aussi créer de nouvelles sonorités comme le montrent les morceaux de musique électronique. Pillot-Loiseau et al. (2020) a montré la proximité acoustique de la trompette avec des imitations de trompettes beatboxées. Dans l'émission de vulgarisation scientifique française *E=M6*, nous avons mis au défi 2 beatboxeurs. Le défi consistait à imiter l'introduction de deux morceaux connus : *Walk this way* d'Aerosmith (batterie) et *Europa* de Santana (guitare électrique). Les signatures acoustiques sont très similaires entre la guitare du morceau original et l'imitation beatboxée (Figure 5.2). Évidemment l'acoustique de la guitare électrique n'est pas comparable avec la voix humaine, d'où les différences de la structure acoustique. Ces deux exemples rappellent la nature acoustique des cibles de production dans le Human Beatbox.

Dans nos données, les gestes de production manipulent l'air dans le conduit vocal et changent la configuration du conduit vocal ce qui modifie la signature acoustique des sons. Les gestes de production ne sont donc pas une finalité mais un moyen d'atteindre les cibles acoustiques. De

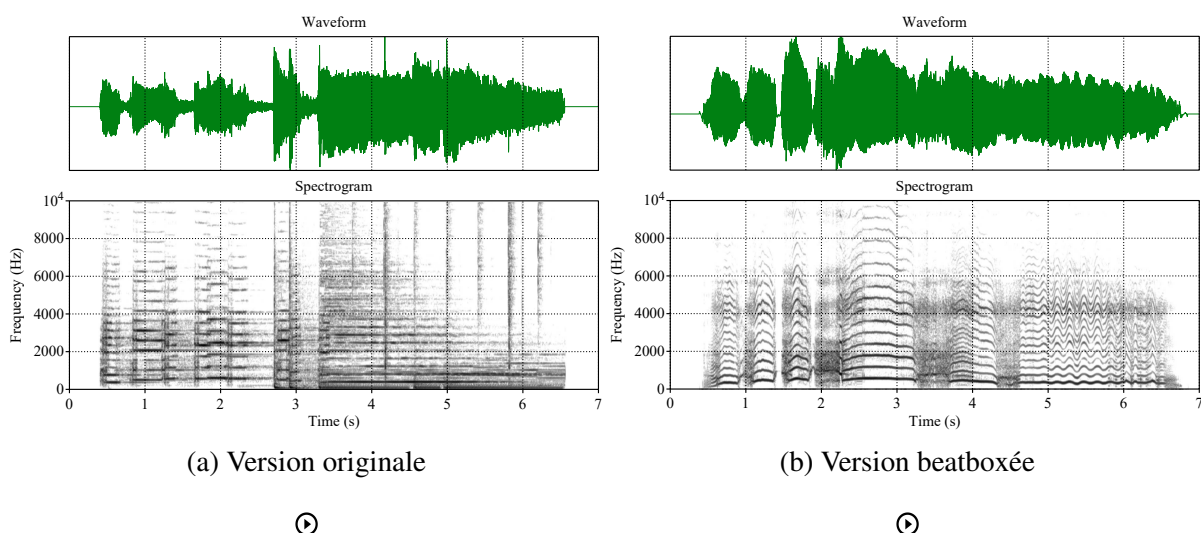


FIGURE 5.2 – *Europa*, Santana

plus, nous avons vu que différentes stratégies articulatoires individuelles étaient possibles pour produire un même son avec une signature acoustique similaire entre individus (*K-Snare Inward*, *ESH Snare*, *Cough Snare*, cf. chapitre 3). L’articulation et l’initiation peuvent donc varier sans que la signature ne change. Cela rejoint la théorie quantale de Stevens (1989) et suggère qu’une approche par trait (issus de la théorie phonétique) est plus pertinente qu’une approche par gestes.

Toutefois, la théorie (phonologique) des traits (en l’état) n’est pas satisfaisante non plus car elle est spécifique aux systèmes linguistiques et ne permet pas de représenter les unités beatboxées (notes de musique). C’est pourquoi il faudrait faire un pas en arrière et revenir à une théorie acoustique des traits et de leur corrélats phonétiques dont la pertinence proviendrait d’indices articulatoires, aérodynamiques et acoustiques comme le suggéraient Jakobson, Fant et Halle puis Stevens. Cette approche acoustique des traits doit se faire en premier temps dans une perspective anthropophonique. Ensuite seulement, les traits doivent être envisagés dans une perspective linguistique pour comprendre quels sont les traits exploités ou non par les langues et comment ces derniers se combinent. Il faut donc envisager les primitives phonétiques à un niveau macroscopique dans une perspective biologique et sociale.

Notre corpus de *Beat Patterns* comportait 3 caisses claires laryngales : la *ESH Snare Outward*, la *ESH Snare Inward* et la *Cough Snare*. Ces trois caisses claires présentent des similitudes avec des mécanismes biologiques comme la toux ou le hoquet. La *Cough Snare* (caisse claire « toussée ») fait une allusion directe au mécanisme de la toux. Y-a-t-il un lien entre ces mécanismes biologiques et les mécanismes de production utilisés pour beatboxer ? Les beatboxeuses et les beatboxeurs utilisent-ils leurs connaissances phonétiques pour créer de nouvelles sonorités ?

### 5.3 Les sons du beatbox : des possibles sons de la parole ?

Nous venons de proposer une approche *anthropophonique* des phénomènes phonétiques des langues. Si on manipule les mêmes primitives lorsque l’on parle et que l’on beatboxe, qu’est-ce qui explique que ça ne « sonne » pas pareil ? Pour répondre à cette interrogation, on va s’intéresser à la notion de « *possible speech sound* » proposée par Lindblom (1990). En reprenant l’approche de Ian Catford, il questionne le caractère fini des inventaires phonétiques. La fini-

tude des inventaires phonétiques est en fait un artéfact de l'approche purement linguistique de la phonétique. C'est d'ailleurs pour cela qu'il est difficile d'utiliser l'alphabet phonétique international pour transcrire le Human Beatbox. Les inventaires phonétiques ne sont pas finis : ils sous-exploitent les capacités du conduit vocal.

Pour expliquer cette sous-exploitation, Björn Lindblom discute de certains principes et de certaines contraintes qui opèrent lors de la sélection des sons pertinents pour les systèmes linguistiques. Il évoque les phénomènes de co-occurrence des phonèmes d'une langue. Il remarque que les inventaires de la base de données UPSID (UCLA Phonological Segment Inventory Database) sont plutôt uniformes. La structuration des inventaires phonétiques des langues suit des tendances similaires. Ces tendances ont été explicitées et discutées par Ian Maddieson (Maddieson, 1984). Or, dans notre corpus de *beat patterns* nous avons des phénomènes de co-occurrence de sons. La contrainte de co-occurrence n'est donc pas pertinente pour définir les sons de la parole.

Il mentionne aussi la simplicité/complexité articulatoire comme une contrainte lors de la sélection des segments linguistiquement pertinents. La question de la complexité en phonétique et phonologie a été abordée par différentes approches (Chitoran & Cohn, 2009) : la marque (Troubetzkoy, 1976), le nombre de traits par segment (Lehmann, 1974), la phonologie naturelle (Donegan & Stampe, 1979), la complexité articulatoire (Lindblom & Maddieson, 1988) ou encore une approche de la complexité par la théorie de l'information (Hume, 2006). Avant de comprendre la complexité phonologique, il faut définir la complexité phonétique.

Lindblom et Maddieson (1988) suggèrent que les systèmes linguistiques se structurent selon différents degrés de complexité articulatoire. Les petits inventaires tendent à sélectionner des articulations simples alors que les grands inventaires tendent à sélectionner des articulations simples et complexes. Les articulations simples seraient sélectionnées en premier puis en fonction des besoins du système en termes de contrastes. Les systèmes sélectionneraient ensuite des articulations élaborées et complexes. Les articulations simples sont les séries occlusives [p t k b d g], la série de fricatives sourdes [f s h] et l'affriquée [tʃ] ainsi que les sonnantes [m n ŋ l r w j]. Les articulations élaborées sont produites sur les autres lieux d'articulation, avec d'autres mécanismes de production (types phonatoires, mécanismes d'initiation) ou avec des articulations secondaires par exemple. Les articulations complexes sont des combinaisons de plusieurs dimensions des articulations élaborées.

Leur proposition est intéressante mais ne permet pas de définir la complexité phonétique dans une perspective anthropophonique. Les sons de notre corpus tombent dans les catégories élaborées ([p']) ou dans aucune des catégories concernant les sons pulmoniques ingressifs ([↓<sub>B</sub>↑]). Il est cependant clair que le *Lips Roll* est une articulation complexe. C'est d'ailleurs un son très complexe à acquérir pour un beatboxeur car l'apprentissage du son peut prendre plusieurs mois voire plusieurs années<sup>4</sup>. À l'inverse, l'acquisition du kit basique de la batterie [p' ts' k'] se fait dans les tous premiers mois de pratique. Il y a donc des articulations simples et complexes dans le Human Beatbox.

La durée d'acquisition d'un segment est aussi une contrainte des systèmes linguistiques. Des articulations simples seront plus facilement acquises chez des enfants que des articulations aussi complexes que ceux du Human Beatbox. Il paraît assez clair que les sons beatboxés ne constituent pas des candidats idéaux pour les systèmes linguistiques en raison de leur complexité et du temps d'apprentissage qu'ils nécessitent. Cela montre que le beatboxeur travaille le conduit vocal comme un musicien travaille son instrument.

Comment définir la complexité phonétique dans une perspective anthropophonique ? Au-

---

4. Un beatboxeur de la communauté française nous a confié avoir travaillé pendant 6 ans pour acquérir le *Lips Roll*!

tremement dit, comment définir une échelle de complexité macroscopique qui puisse s'appliquer aussi bien au Human Beatbox qu'à la parole ? Il faudrait définir la complexité de façon graduelle et non de façon ternaire comme le proposent Lindblom & Maddieson. L'idée que ces derniers proposent concernant la distance des lieux d'articulation par rapport à la position de repos des organes est très intéressante. Une des sources de complexité émergerait donc des distances à parcourir par les organes par rapport à une position de repos. En d'autres termes, on pourrait voir les articulations simples comme des « paramètres par défaut » (*default settings*) et la complexité articulatoire émanerait du passage des paramètres par défaut à des paramètres de plus en plus complexes. Ces paramètres ne sont pas seulement articulatoires mais aussi aérodynamiques. On peut émettre l'hypothèse que les occlusives sont plus simples par rapport aux trilles, aux fricatives et aux affriquées dont la production est plus contrainte par des facteurs aérodynamiques. Par exemple pour les trilles (ou pour la phonation), on sait qu'elles sont contraintes par les conditions aérodynamiques nécessaires à générer une suite d'oscillation sur un temps donné. On peut aussi évoquer les contraintes aérodynamiques sur le voisement (Ohala, 1997).

La caractère pulmonique ingressif est complexe, plus particulièrement dans les cas de phonation ingressive. En effet, les sujets de notre étude mentionnent des douleurs lors de l'apprentissage des *bass inward*. Deux contraintes sont évoquées pour expliquer l'absence de ces mécanismes dans les langues du monde<sup>5</sup> : le caractère conflictuel des flux pulmoniques ingressifs et égressifs ainsi que la forme des plis vocaux (Hewlett & Beck, 2006 ; Ladefoged, 1971). La contrainte respiratoire est pertinente pour la parole mais pas pour le Human Beatbox. Les beatboxeurs et les beatboxeuses réorganisent l'activité respiratoire par différents moyens : intégration de sons pulmoniques ingressifs, micro-cycles respiratoires et respiration circulaire (cf. questionnaires en annexe - Question 27).

La question de la complexité des mécanismes d'initiation doit être posée. Est-ce que les mécanismes linguaux et laryngaux sont aussi complexe l'un que l'autre ? Nous avons vu que les mécanismes laryngaux interagissaient avec les mécanismes pulmoniques et avec les gestes articulatoires laryngaux ; de même, on sait que les mécanismes d'initiation linguaux contraignent les articulations linguales (*back vowel constraint*, Traill, 1985). Nos données montrent qu'on peut aussi co-initier des sons (cf. Chapitre 3). Pour expliquer la complexité phonétique, il nous semble important d'intégrer à la fois les contraintes biomécaniques, physiologiques et les possibles conflits de coordination entre les gestes de production. C'est ce que propose aussi Lindblom (1983) en postulant une contrainte d'effort, c'est-à-dire le coût en énergie de la production par unité de temps. Bien que polémique, la notion d'effort physique nous paraît intéressante pour définir les sons possibles de la parole.

## 5.4 Effort et contrainte de puissance

La notion d'« effort », en particulier celle du moindre effort fait polémique en linguistique. Proposée par Zipf (1949) (et les néo-grammairiens avant lui), le principe du « moindre effort » se voulait une explication (parmi d'autres) du changement linguistique. La langue changerait parce que nous sommes fainéants. Cette approche est trop simpliste pour être généralisée à l'ensemble des pratiques langagières hétérogènes. Martinet (1960) propose le principe d'« économie des changements linguistiques » plutôt qu'un principe (quasi-absolu) du moindre effort. Il écrit que l'économie des changements linguistiques est la « recherche permanente de l'équilibre entre des besoins contradictoires qu'il faut satisfaire : besoins communicatifs d'une part, inertie mémorielle et inertie articulatoire d'autre part » (p.181). L'intuition de Martinet est correcte :

---

5. pour une revue exhaustive de la phonation ingressive voir Eklund (2008)

les systèmes linguistiques présentent des conflits entre le besoin social de communication<sup>6</sup> et coût cognitif et physiologique de la parole. En proposant le principe d'économie des changements linguistiques, Martinet place la linguistique et la phonétique dans une perspective sociale et biologique.

Lindblom (1983) pose le problème d'économie des gestes de production. En supposant que les sons de la parole résultent des adaptations aux contraintes biologiques, il tente d'explicitier les contraintes qui opèrent sur les systèmes de production et de perception. Il propose deux principes d'« optimisation motrice » : la contrainte de synergie (*synergy constraint*) et la contrainte de vitesse (*rate constraint*). La première opère dans une dimension statique et spatiale des gestes de production ; la deuxième opère sur la dimension temporelle et dynamique des gestes de production. Il souligne que ces contraintes ne sont pas spécifiques à la parole mais aux mouvements de façon générale (écriture, langue des signes).

La contrainte de *synergie* opère à partir des degrés de liberté articulaire des organes de production. Il propose une modélisation des (nombreux) degrés de liberté des articulateurs et s'appuie en particulier sur les restrictions de co-articulation linguale des consonnes et des voyelles (Figure 5.3). La coarticulation vient de l'interaction des différents degrés de liberté de la pointe de la langue et du corps de la langue. Une étude de la totalité des degrés de liberté par des modélisations biomécaniques permettrait d'approfondir l'hypothèse de la contrainte de synergie et de son rôle dans le principe d'économie des gestes de production.



FIGURE 5.3 – Degrés de liberté de la pointe de la langue et du corps de la langue (Adapté de Lindblom, 1983)

Il développe la contrainte de vitesse (*rate constraint*) en s'appuyant sur les phénomènes de réduction (hyper/hypo-articulation). Cette contrainte est une proposition pour résoudre le problème de l'invariance des phonèmes et la variabilité de leur réalisation phonétique. L'hyper-articulation et l'hypoarticulation se rapportent à la dynamique de la production articulaire. Si l'on applique une force pour mettre en mouvement un articulateur et que le temps d'application est fortement réduit alors la réponse du système peut se révéler insuffisante pour que l'articulateur atteigne la cible avant que la commande suivante soit exécutée. Il y a donc un risque

6. Besoin parfois vital comme le suggère les situations de privation de communication humaine (cf. Empereur Frédéric II de Prusse (1194-1250), Victor l'« enfant sauvage »)

d'hypo-articulation (*undershoot*). Toutefois, l'*undershoot* n'est pas inévitable. Pour s'assurer d'atteindre la cible au bon moment, on peut accélérer, c'est-à-dire modifier la vitesse et la trajectoire de l'articulateur. Il existe donc un équilibre subtil entre différentes forces musculaires qui permet le contrôle du timing (atteindre la cible dans les temps) et de la vitesse de production.

Il émet l'hypothèse que ces deux contraintes découlent d'une contrainte plus globale de puissance de production, c'est-à-dire du coût énergétique de la production par unité de temps. Les systèmes linguistiques sélectionneraient des sons au coût énergétique faible. La production du Human Beatbox nécessite un effort intense. Les sujets de notre étude nous ont rapporté qu'il leur arrivait de ressentir des douleurs musculaires. De plus, les beatboxeurs et les beatboxeuses ne pourraient pas beatboxer plusieurs heures comme on pourrait le faire pour parler. Nous pensons que l'hypothèse de Lindblom est correcte, les sons de la parole (contrairement aux sons beatboxés) sont restreints par une contrainte de puissance. Toutefois, le principe d'économie ne saurait être réduit à sa dimension purement physiologique. La dimension sociale est tout aussi importante car les énoncés linguistiques varient selon les personnes et les situations de communication (qui est-ce qui vient ? *versus* qui vient ?).

## 5.5 La parole et le Human Beatbox

Qu'il s'agisse de possibles sons beatboxés ou de possibles sons parlés, la sélection des unités dans l'espace anthropophonique (*alphabet anthropophonique*) est soumise à des contraintes systémiques définies par les besoins des systèmes sonores pour établir des contrastes lexicaux (dans le cas de la parole) ou phonétiques (dans le cas du HBB) (Figure 5.4).

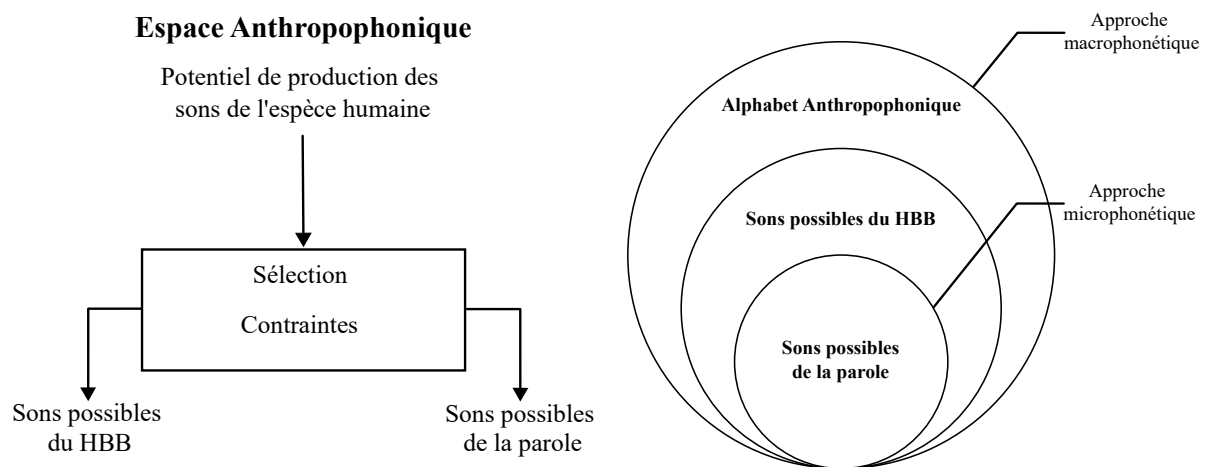


FIGURE 5.4 – Approche anthropophonique du Human Beatbox

La *contrainte d'effort* consiste à éviter les mouvements articulaires extrêmes (*synergy constraint*) et les vitesses extrêmes (*rate constraint*). L'effort contraint la complexité et les patrons de coordination articulaires pendant la parole. Il est possible de dépasser ces contraintes lorsque l'on beatboxe. Tout comme un musicien travaillerait pour améliorer ses performances musicales, les beatboxeurs et les beatboxeuses doivent s'entraîner pour exploiter pleinement les capacités du conduit vocal humain. Dépasser la contrainte d'effort demande du temps et de l'entraînement (temps dont notre espèce ne dispose pas pour apprendre à parler). C'est pour cette raison que, selon nous, les sons beatboxés font de mauvais candidats pour les systèmes linguistiques. Les systèmes linguistiques choisissent des articulations simples en nombre limité qui

contrastent suffisamment entre elles (*sufficient perceptual contrast* plutôt que *maximal perceptual contrast*, à ce propos voir Lindblom (1990)). Pour le HBB, les beatboxeurs sélectionnent les sons en nombre illimité en fonction de besoins musicaux et esthétiques. La sélection des sons se fait en fonction des choix artistiques des beatboxeurs et des beatboxeuses : pour produire des sons de batterie ou de percussion, on va sélectionner des sons occlusifs, fricatifs et affriqués ; pour produire des sons de trompette ou de saxophone, on va sélectionner des sons avec une structure harmonique ; et ainsi de suite. Cela suggère donc que la sélection des unités pertinentes de la parole ou du Beatbox est motivée par les buts de production.

Si la sélection des unités pertinentes du HBB dépend des buts de production alors l'approche théorique du Human Beatbox doit se faire sans considérations linguistiques ou phonologiques. Les unités beatboxées, et plus largement la production du Beatbox, ne peuvent être analysées au prisme de la théorie phonologique. On ne saurait analyser les sons beatboxés comme des phonèmes ou la technique de « *humming* » comme un processus phonologique d'harmonie car cela impliquerait qu'un système musical, ici le HBB, posséderait les caractéristiques d'un système phonologique. Il est clair que linguistique et musique partagent des points communs car ce sont tous les deux des systèmes de communication acoustique et que tous deux dépendent de contraintes de perception auditive. Toutefois, les systèmes linguistiques et musicaux diffèrent dans leurs structures et dans leurs fonctions (Jackendoff, 2009 ; Patel, 2008). Il convient donc, jusqu'à preuve du contraire, de distinguer le HBB et la parole comme deux objets théoriques distincts. Cela n'empêche pas d'analyser les points communs et les différences entre ces deux objets. La distinction en 2 objets théoriques différents n'entre pas en contradiction avec une approche anthropophonétique du moment qu'on se borne à l'analyse de la production et de la perception des sons sans chercher à établir le fonctionnement de ces derniers dans une perspective phonologique ou cognitive.

# Conclusion

Il est à présent temps de conclure ce travail. Nous nous étions donnés 2 objectifs et nous avons formulé deux hypothèses :

— Objectifs :

1. En reprenant la proposition de Proctor et al. (2013) nous montrerons que le HBB est un système combinatoire discret. Nous tenterons de démontrer en quoi le HBB est un système de production (articulatoirement) plus complexe que la parole.
2. En reprenant le paradigme de variation de la vitesse de production, nous analyserons les effets du tempo de référence sur la production du HBB à partir de données temporelles, aérodynamiques et acoustiques.

— Hypothèses :

1. Les primitives du HBB sont les mêmes que les systèmes linguistiques mais que les deux systèmes diffèrent dans les possibilités combinatoires des primitives pour former des unités.
2. L'augmentation de la vitesse de production entraînera une diminution de la durée des gestes de production ainsi qu'une diminution de l'amplitude du geste d'initiation. En l'absence de données articulatoires à différents tempo, nous ne faisons pas d'hypothèse sur l'amplitude des gestes articulatoires.

Le Chapitre 3 cherchait à montrer que le Human Beatbox est un système combinatoire discret comme le proposent Proctor et al. (2013). L'hypothèse défendue était que les primitives du Human Beatbox sont les mêmes qu'en parole car ces deux systèmes manipulent les mêmes organes de production (système respiratoire, laryngal, oral). Ce qui diffère entre les deux systèmes sont les possibilités combinatoires de ces primitives. Pour confirmer ou infirmer l'hypothèse, nous avons proposé une description phonétique des mécanismes de production des sons de notre corpus à partir de données articulatoires, aérodynamiques et acoustiques.

Le Chapitre 3 a montré que les beatboxeurs et les beatboxeuses utilisaient les mêmes mécanismes de production qu'en parole et que la combinaison des mécanismes dans le HBB diffèrent des combinaisons observées dans les systèmes sonores des langues du monde. Nous avons observé l'utilisation des 6 mécanismes d'initiation possibles. Les mécanismes d'initiation se combinent ensuite avec différents mécanismes articulatoires et donnent lieu à des signatures acoustiques discrètes qui distinguent les catégories entre elles (grosses caisses, caisses claires) mais aussi les sons d'une même catégorie (*K-Snare Inward*, *pf snare*). Nous avons aussi montré que la production de la voix beatboxée reposait sur des stratégies phonatoires complexes combinant des éléments périodiques réguliers et irréguliers, des éléments apériodiques (friction) et des éléments périodiques déphasés (double voix). Enfin, nous avons montré que l'utilisation de mécanismes sifflés pouvait se superposer à d'autres mécanismes. Tous ces éléments suggèrent que le Human Beatbox possède en effet les propriétés d'un système combinatoire discret et complexe. Notre hypothèse initiale est validée. Les primitives du beatbox sont les mêmes que



les primitives des systèmes sonores mais dans le HBB, les possibilités de combinaison et les patrons de coordination sont différents de la parole.

Le chapitre 4 avait pour objectif de décrire, d'analyser et d'expliquer les effets du tempo de référence (90, 120, 150 BPM) sur la production du Human Beatbox à partir d'une analyse temporelle, aérodynamique et acoustique des sons et des *Beat Patterns*. Nous avons fait l'hypothèse que l'augmentation de la vitesse de production entraînerait une diminution de la durée des gestes de production ainsi qu'une diminution de l'amplitude du geste d'initiation. En l'absence de données articulatoires à différents tempo, nous n'avons pas fait d'hypothèse sur l'amplitude des gestes articulatoires. L'hypothèse initiale est partiellement confirmée. Nous avons montré une tendance des sons et des intervalles intergestuels à se réduire. L'analyse acoustique a révélé des phénomènes de chevauchement temporel des gestes articulatoires. La pression ne se réduit pas forcément à mesure que le tempo accélère. Nous avons suggéré que la régulation de la pression dépendait de stratégies individuelles de contrôle de l'intensité acoustique. L'analyse du sous-ensemble des charlestons [ts'] a révélé, chez au moins 2 sujets, des interactions entre les gestes d'initiation d'une part, et entre le geste d'initiation et d'articulation d'autre part. La pression tend à s'affaiblir quand ces mécanismes interagissent. Il est toutefois possible de surpasser cette contrainte comme l'ont montré les données du sujet GA.

Pris dans leur ensemble, les résultats montrent que les sujets utilisent des mécanismes de production similaires aux locuteurs des langues mais que les beatboxeurs utilisent un nombre plus important de combinaison de mécanismes différents ([↓kL: ʔj: ↓B¹:]). Les résultats sur les effets du tempo suggèrent que les stratégies individuelles de réorganisation temporelles des gestes sont similaires aux stratégies proposées par Byrd et Tan (1996). En effet, les sujets ont manipulé la durée des gestes et des intervalles séparant les gestes. De futures études sont nécessaires pour connaître les stratégies de régulation de la pression et l'intensité acoustique à mesure que le tempo accélère. Cette thèse ouvre la voie à de futures études des effets du tempo sur la production du Human Beatbox. Il serait intéressant par exemple de reprendre les données IRM pour analyser les patrons de coordination gestuelle.

Qu'est-ce que le Human Beatbox peut apporter à la phonétique ? Il est assez clair que les beatboxeurs et les beatboxeuses exploitent une plus grande partie des capacités du conduit vocal humain. Le Human Beatbox pose la question du potentiel vocal de notre espèce et interroge sur la sous-exploitation du conduit vocal par les systèmes phonétiques des langues. Pourquoi les langues se structurent comme elles le font et pas autrement ? Si l'on veut savoir pourquoi les langues sous-exploitent les capacités vocales de notre espèce, il faudrait en premier lieu tenter de définir l'étendue des capacités du conduit vocal. Définir l'étendue des capacités nécessite de passer d'une approche *microscopique* (phonétique linguistique) des phénomènes phonétiques, à une approche *macroscopique* (anthropophonie). L'approche macroscopique ne peut se faire sans que le cadre phonétique soit élargi. Le Human Beatbox invite à regarder différemment la phonétique. Le HBB invite aussi à actualiser le cadre théorique de la phonétique afin de passer d'une approche purement linguistique à une approche anthropophonique (Catford, 1977 ; Lindblom, 1990) des phénomènes phonétiques. Une telle approche ne se révélerait pas seulement bénéfique pour les études sur le Human Beatbox mais aussi pour les études sur l'acquisition et les pathologies du langage.

# ANNEXE 1 : Questionnaire VP

## Questionnaire HBB

A remplir par l'investigateur avant de donner le questionnaire

N° du sujet : |V|P|-|0|1|

Consignes : merci de répondre aux questions suivantes. Les ronds (o) indiquent que plusieurs réponses sont possibles, les carrés (☐) indiquent qu'une seule réponse est attendue. Si vous ne comprenez pas une question, vous pouvez demander une explication au médecin investigateur ou à un membre de son équipe.

### 1. Profession :

- Libéral
- Artiste
- Artisan
- Enseignant
- Informatique
- 
- Autres

### 2. Langue(s) parlée(s) couramment :

- o Langue 1 : FRANCAIS
- o Langue 2 : ANGLAIS
- o Langue 3 : \_\_\_\_\_
- o Langue 4 : \_\_\_\_\_
- o Langue 5 : \_\_\_\_\_

### 3. Niveau d'éducation :

- Collège
- Lycée
- Licence
- Master
- Doctorat
- Autre

### 4. Nombre d'année(s) de pratique du beatbox :

- 1-2 ans
- 2-3 ans
- 3-4 ans
- 5-6 ans
- Plus de 6 ans

# ANNEXE 1 : Questionnaire VP

## 5. Expérience musicale :

- oui
- non

### 5.1. Si oui : liste d'instrument, style ou technique musicale

- Guitare
- Basse
- Batterie
- Violon
- Contrebasse
- Piano
- Trompette
- Saxophone
- Chant
- Synthétiseur
- Autre

### 5.2. Nombre d'année(s) de pratique:

## 6. Utilisez-vous le beatbox à des fins

- Pédagogiques
- Thérapeutiques
- Aucun des deux

*Par pédagogique, nous entendons si vous enseignez le beatbox ; par thérapeutique, nous voulons savoir si vous participez à des programmes de rééducation orthophonique, si vous intervenez dans des associations animant des ateliers pour un public en situation de handicap (autisme, surdit , dysphasie...).*

## 7. A quel  ge avez-vous d but  le Beatbox ? |1|7|

## 8. Vous  tes

- D butant
- Interm diaire
- Avanc 
- Professionnel / expert

## 9. Vous pratiquez

- Seul
- En duo
- En groupe
- De plusieurs mani res

## 10. Vous vous entrainez

# ANNEXE 1 : Questionnaire VP

- Une fois par jour
- 3 fois par semaine
- Moins de 3 fois par semaine

## 10.1. Si vous vous entraînez de façon quotidienne, indiquez la durée approximative par jour :

- Moins de 2h
- Entre 2h et 4h
- Entre 4h et 6h
- Plus de 6 heures

## 10.2. Si vous vous entraînez moins de 3 fois par semaine, indiquez la fréquence approximative sur le mois

- 1 jour par mois
- Entre 1 et 3 jours par mois
- Entre 3 et 5 jours par mois
- Entre 5 et 10 jours par mois
- Plus de 10 jours par mois

## 11. Indiquez la durée de vos entraînements (en minutes ou en heures) : |3|0| minutes à |1 H 30| heures

## 12. De quel(s) style(s) musical(aux) vous inspirez-vous pour Beatboxer ?

<input checked="" type="checkbox"/> Disco	<input checked="" type="checkbox"/> Country	<input checked="" type="checkbox"/> R'n'B
<input checked="" type="checkbox"/> Blues	<input checked="" type="checkbox"/> Reggae	<input checked="" type="checkbox"/> Electro
<input checked="" type="checkbox"/> Funk	<input checked="" type="checkbox"/> Afro	<input type="checkbox"/> K-pop
<input checked="" type="checkbox"/> Jazz	<input type="checkbox"/> Raï	<input type="checkbox"/> Zouk
<input type="checkbox"/> Metal	<input checked="" type="checkbox"/> Ska	<input type="checkbox"/> makossa
<input checked="" type="checkbox"/> Pop	<input type="checkbox"/> Gospel	<input checked="" type="checkbox"/> Musique Indi
<input checked="" type="checkbox"/> Punk	<input checked="" type="checkbox"/> Soul	<input checked="" type="checkbox"/> Musique Latino
<input checked="" type="checkbox"/> Rap	<input type="checkbox"/> Kompa	
<input checked="" type="checkbox"/> Rock 'n' roll	<input checked="" type="checkbox"/> Classique	

## 13. Vous maîtrisez les groupes de sons suivants :

(une réponse par ligne)

Maitrise :	Pas du tout	Moyenne	Bonne	Excellente
Grosse caisse	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Charleston	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Caisse claire	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Instrument à vent (trompettes, saxo)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Scratches	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

## ANNEXE 1 : Questionnaire VP

Lips roll	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Clic roll	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Bass	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Par groupe de sons, nous voulons avoir, par exemple, votre maîtrise des différentes déclinaisons de caisse claires (PF, Inward, PCH etc). Si vous avez une bonne maîtrise de l'ensemble des déclinaisons, mettez « excellent ». A l'inverse, si vous ne maîtrisez pas, mettez « pas du tout »

### 14. Vous utilisez et travaillez beaucoup les ...

- Grosse caisse
- Charleston
- Caisse claire
- Instrument à vent (trompettes, saxo...)
- Scratches
- Lips roll
- Clic roll
- Bass

### 15. Savez-vous produire tous les sons existants à ce jour ?

- Oui
- Non
- Ne sait pas
- C'est impossible

#### 15.1. Si c'est impossible, pourquoi ?

### 16. Votre maîtrise des aspects suivants :

(une réponse par ligne)

Maitrise :	Pas du tout	Moyenne	Bonne	Excellente
Précision articulatoire	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Voix	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Le ressenti tactile	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Le ressenti auditif	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Rythme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Puissance	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Endurance	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Respiration	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Fluidité (enchaînement des sons)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Vitesse	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### 17. Vous travaillez principalement

- Précision articulatoire
- Voix
- Le ressenti tactile
- Le ressenti auditif

# ANNEXE 1 : Questionnaire VP

- Rythme
- Puissance
- Endurance
- Respiration
- Fluidité (enchaînement des sons)
- Vitesse

## 18. Utilisez-vous un métronome pour rester en rythme ?

- Oui
- Non

## 19. Pour beatboxer il faut avoir une prédisposition articulatoire et/ou vocale\*

- D'accord
- Plus ou moins d'accord
- Pas d'accord

*\*prédisposition articulatoire = facilité à bouger les articulateurs comme la langue ou la mâchoire pour coordonner une suite de son de Beatbox*

## 20. Quand je beatboxe je m'appuie surtout sur des voyelles

- D'accord
- Plus ou moins d'accord
- Pas d'accord

## 21. Quand je beatboxe je m'appuie surtout sur des consonnes

- D'accord
- Plus ou moins d'accord
- Pas d'accord

## 22. Quand je beatboxe j'articule plus que quand je parle

- D'accord
- Plus ou moins d'accord
- Pas d'accord

## 23. Quand je beatboxe :

*(une réponse par ligne)*

	Jamais	Parfois	Souvent	Toujours
Je suis assis	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je suis debout	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je suis statique (aucun mouvement)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je balance mon buste	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je suis en mouvement (danse)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'ai les pieds bien ancrés au sol	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

## ANNEXE 1 : Questionnaire VP

Je projette ma tête en avant	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je garde les épaules à la même hauteur que quand je parle	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je garde mes épaules plus hautes que quand je parle	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je garde mes épaules plus basses que quand je parle	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### 24. J'utilise mes mains/bras pour

(une réponse par ligne)

	Jamais	Parfois	Souvent	Toujours
Accompagner le rythme	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Donner le rythme	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Accompagner la mélodie	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Imiter un instrument	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Imiter un DJ	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### 25. Visage. Quand je beatboxe

(une réponse par ligne)

	Oui	Non
J'ai l'impression de faire des grimaces	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je fronce les sourcils	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge le nez	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge les oreilles	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Je bouge ma mâchoire de haut en bas	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge ma mâchoire d'avant en arrière	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge ma mâchoire sur les côtés	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je gonfle les joues	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### 26. Appareil phonatoire ; Quand je beatboxe

(une réponse par ligne)

	Oui	Non
J'utilise ma mâchoire pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mes lèvres pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mes joues pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise ma langue pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise ma gorge (arrière de la langue) pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mon nez pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mon larynx (au niveau des cordes vocales) pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du visage	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau des lèvres	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau de la langue	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du palais	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau de la gorge (arrière de la gorge)	<input checked="" type="checkbox"/>	<input type="checkbox"/>

# ANNEXE 1 : Questionnaire VP

Je ressens des sensations au niveau du nez	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du larynx (au niveau des cordes vocales) (pendant la production de bass par exemple)	<input checked="" type="checkbox"/>	<input type="checkbox"/>

## 27. Respiration

(une réponse par ligne)

	Jamais	Parfois	Souvent	Toujours
Je suis essoufflé(e) quand je beatboxe	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pour reprendre mon souffle, j'arrête de beatboxer	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pour reprendre mon souffle, j'utilise des sons inspirés (ex : inward bass, inward k-snare)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Pour reprendre mon souffle, j'utilise mon nez mais je continue à beatboxer avec ma bouche	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

## 28. Parmi ces sons, quels sont ceux qui sont difficiles à produire, douloureux et quel(s) type(s) de douleurs provoquent-ils ?

Sons :	Difficile à produire ?	Douloureux ?	Type de douleur :
Grosse caisse	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Charleston	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Caisse claire	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	Douloureux uniquement pour la spit snare (douleurs aux muscles du masque) par manque de pratique
Instrument à vent (trompettes, saxo..)	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Scratches	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Lips roll	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Clic roll	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Bass	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	Douloureux en début d'apprentissage. Peut parfois être capricieuse à froid et générer des gênes à cause des muqueuses.

## 29. Une fois que vous maîtrisez ces sons, les douleurs persistent-elles ?

- Oui  
 Non

## 30. Avez-vous déjà consulté un médecin pour un problème lié à votre pratique ?

- Oui  
 Non



# ANNEXE 2 : Questionnaire LJ

## Questionnaire HBB

A remplir par l'investigateur avant de donner le questionnaire

N° du sujet : |L|J|-|0|2|

Consignes : merci de répondre aux questions suivantes. Les ronds (o) indiquent que plusieurs réponses sont possibles, les carrés (☐) indiquent qu'une seule réponse est attendue. Si vous ne comprenez pas une question, vous pouvez demander une explication au médecin investigateur ou à un membre de son équipe.

### 1. Profession :

- Libéral
- Artiste
- Artisan
- Enseignant
- Informatique
- 
- Autres

### 2. Langue(s) parlée(s) couramment :

- o Langue 1 : FRANCAIS
- o Langue 2 : ANGLAIS
- o Langue 3 : \_\_\_\_\_
- o Langue 4 : \_\_\_\_\_
- o Langue 5 : \_\_\_\_\_

### 3. Niveau d'éducation :

- Collège
- Lycée
- Licence
- Master
- Doctorat
- Autre

### 4. Nombre d'année(s) de pratique du beatbox :

- 1-2 ans
- 2-3 ans
- 3-4 ans
- 5-6 ans
- Plus de 6 ans

## ANNEXE 2 : Questionnaire LJ

### 5. Expérience musicale :

- oui
- non

#### 5.1. Si oui : liste d'instrument, style ou technique musicale

- Guitare
- Basse
- Batterie
- Violon
- Contrebasse
- Piano
- Trompette
- Saxophone
- Chant
- Synthétiseur
- Autre

#### 5.2. Nombre d'année(s) de pratique: 3

### 6. Utilisez-vous le beatbox à des fins

- Pédagogiques
- Thérapeutiques
- Aucun des deux

*Par pédagogique, nous entendons si vous enseignez le beatbox ; par thérapeutique, nous voulons savoir si vous participez à des programmes de rééducation orthophonique, si vous intervenez dans des associations animant des ateliers pour un public en situation de handicap (autisme, surdit , dysphasie...).*

### 7. A quel  ge avez-vous d but  le Beatbox ? |1|2|

### 8. Vous  tes

- D butant
- Interm diaire
- Avanc 
- Professionnel / expert

### 9. Vous pratiquez

- Seul
- En duo
- En groupe
- De plusieurs mani res

### 10. Vous vous entrainez

## ANNEXE 2 : Questionnaire LJ

- Une fois par jour
- 3 fois par semaine
- Moins de 3 fois par semaine

### 10.1. Si vous vous entraînez de façon quotidienne, indiquez la durée approximative par jour :

- Moins de 2h
- Entre 2h et 4h
- Entre 4h et 6h
- Plus de 6 heures

### 10.2. Si vous vous entraînez moins de 3 fois par semaine, indiquez la fréquence approximative sur le mois

- 1 jour par mois
- Entre 1 et 3 jours par mois
- Entre 3 et 5 jours par mois
- Entre 5 et 10 jours par mois
- Plus de 10 jours par mois

11. Indiquez la durée de vos entraînements (en minutes ou en heures) : |\_\_|\_\_| minutes ou |\_\_|\_\_| heures

### 12. De quel(s) style(s) musical(aux) vous inspirez-vous pour Beatboxer ?

<input checked="" type="checkbox"/> Disco	<input type="checkbox"/> Country	<input checked="" type="checkbox"/> R'n'B
<input type="checkbox"/> Blues	<input type="checkbox"/> Reggae	<input checked="" type="checkbox"/> Electro
<input checked="" type="checkbox"/> Funk	<input checked="" type="checkbox"/> Afro	<input type="checkbox"/> K-pop
<input type="checkbox"/> Jazz	<input type="checkbox"/> Raï	<input checked="" type="checkbox"/> Zouk
<input type="checkbox"/> Metal	<input type="checkbox"/> Ska	<input type="checkbox"/> makossa
<input checked="" type="checkbox"/> Pop	<input type="checkbox"/> Gospel	<input checked="" type="checkbox"/> Musique Indi
<input type="checkbox"/> Punk	<input checked="" type="checkbox"/> Soul	<input checked="" type="checkbox"/> Musique Latino
<input checked="" type="checkbox"/> Rap	<input type="checkbox"/> Kompa	
<input type="checkbox"/> Rock 'n' roll	<input checked="" type="checkbox"/> Classique	

### 13. Vous maîtrisez les groupes de sons suivants :

(une réponse par ligne)

Maitrise :	Pas du tout	Moyenne	Bonne	Excellente
Grosse caisse	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Charleston	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Caisse claire	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Instrument à vent (trompettes, saxo)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Scratches	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

## ANNEXE 2 : Questionnaire LJ

Lips roll	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Clic roll	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Bass	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Par groupe de sons, nous voulons avoir, par exemple, votre maîtrise des différentes déclinaisons de caisse claires (PF, Inward, PCH etc). Si vous avez une bonne maîtrise de l'ensemble des déclinaisons, mettez « excellent ». A l'inverse, si vous ne maîtrisez pas, mettez « pas du tout »

### 14. Vous utilisez et travaillez beaucoup les ...

- Grosse caisse
- Charleston
- Caisse claire
- Instrument à vent (trompettes, saxo...)
- Scratches
- Lips roll
- Clic roll
- Bass

### 15. Savez-vous produire tous les sons existants à ce jour ?

- Oui
- Non
- Ne sait pas
- C'est impossible

#### 15.1. Si c'est impossible, pourquoi ?

### 16. Votre maîtrise des aspects suivants :

(une réponse par ligne)

Maitrise :	Pas du tout	Moyenne	Bonne	Excellente
Précision articulatoire	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Voix	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le ressenti tactile	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Le ressenti auditif	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Rythme	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Puissance	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Endurance	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Respiration	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Fluidité (enchaînement des sons)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Vitesse	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### 17. Vous travaillez principalement

- Précision articulatoire
- Voix
- Le ressenti tactile
- Le ressenti auditif

## ANNEXE 2 : Questionnaire LJ

- Rythme
- Puissance
- Endurance
- Respiration
- Fluidité (enchaînement des sons)
- Vitesse

### 18. Utilisez-vous un métronome pour rester en rythme ?

- Oui
- Non

### 19. Pour beatboxer il faut avoir une prédisposition articulatoire et/ou vocale\*

- D'accord
- Plus ou moins d'accord
- Pas d'accord

*\*prédisposition articulatoire = facilité à bouger les articulateurs comme la langue ou la mâchoire pour coordonner une suite de son de Beatbox*

### 20. Quand je beatboxe je m'appuie surtout sur des voyelles

- D'accord
- Plus ou moins d'accord
- Pas d'accord

### 21. Quand je beatboxe je m'appuie surtout sur des consonnes

- D'accord
- Plus ou moins d'accord
- Pas d'accord

### 22. Quand je beatboxe j'articule plus que quand je parle

- D'accord
- Plus ou moins d'accord
- Pas d'accord

### 23. Quand je beatboxe :

*(une réponse par ligne)*

	Jamais	Parfois	Souvent	Toujours
Je suis assis	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je suis debout	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je suis statique (aucun mouvement)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je balance mon buste	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je suis en mouvement (danse)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
J'ai les pieds bien ancrés au sol	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## ANNEXE 2 : Questionnaire LJ

Je projette ma tête en avant	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Je garde les épaules à la même hauteur que quand je parle	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je garde mes épaules plus hautes que quand je parle	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Je garde mes épaules plus basses que quand je parle	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 24. J'utilise mes mains/bras pour

(une réponse par ligne)

	Jamais	Parfois	Souvent	Toujours
Accompagner le rythme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Donner le rythme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Accompagner la mélodie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Imiter un instrument	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Imiter un DJ	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 25. Visage. Quand je beatboxe

(une réponse par ligne)

	Oui	Non
J'ai l'impression de faire des grimaces	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je fronce les sourcils	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge le nez	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge les oreilles	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Je bouge ma mâchoire de haut en bas	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge ma mâchoire d'avant en arrière	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge ma mâchoire sur les côtés	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je gonfle les joues	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### 26. Appareil phonatoire ; Quand je beatboxe

(une réponse par ligne)

	Oui	Non
J'utilise ma mâchoire pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mes lèvres pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mes joues pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise ma langue pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise ma gorge (arrière de la langue) pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mon nez pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mon larynx (au niveau des cordes vocales) pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du visage	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau des lèvres	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau de la langue	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du palais	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau de la gorge (arrière de la gorge)	<input checked="" type="checkbox"/>	<input type="checkbox"/>

## ANNEXE 2 : Questionnaire LJ

Je ressens des sensations au niveau du nez	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du larynx (au niveau des cordes vocales) (pendant la production de bass par exemple)	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### 27. Respiration

(une réponse par ligne)

	Jamais	Parfois	Souvent	Toujours
Je suis essoufflé(e) quand je beatboxe	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pour reprendre mon souffle, j'arrête de beatboxer	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pour reprendre mon souffle, j'utilise des sons inspirés (ex : inward bass, inward k-snare)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Pour reprendre mon souffle, j'utilise mon nez mais je continue à beatboxer avec ma bouche	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

### 28. Parmi ces sons, quels sont ceux qui sont difficiles à produire, douloureux et quel(s) type(s) de douleurs provoquent-ils ?

Sons :	Difficile à produire ?	Douloureux ?	Type de douleur :
Grosse caisse	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Charleston	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Caisse claire	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Instrument à vent (trompettes, saxo..)	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Scratches	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Lips roll	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Clic roll	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Bass	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	

### 29. Une fois que vous maîtrisez ces sons, les douleurs persistent-elles ?

- Oui  
 Non

### 30. Avez-vous déjà consulté un médecin pour un problème lié à votre pratique ?

- Oui  
 Non

# ANNEXE 3 : Questionnaire AI

## Questionnaire HBB

A remplir par l'investigateur avant de donner le questionnaire

N° du sujet : |A||-|0|3|

Consignes : merci de répondre aux questions suivantes. Les ronds (o) indiquent que plusieurs réponses sont possibles, les carrés (☐) indiquent qu'une seule réponse est attendue. Si vous ne comprenez pas une question, vous pouvez demander une explication au médecin investigateur ou à un membre de son équipe.

### 1. Profession :

- Libéral
- Artiste
- Artisan
- Enseignant
- Informatique
- 
- Autres

### 2. Langue(s) parlée(s) couramment :

- Langue 1 : FRANCAIS
- Langue 2 : PORTUGAIS
- Langue 3 : ANGLAIS
- Langue 4 : ESPAGNOL
- Langue 5 : \_\_\_\_\_

### 3. Niveau d'éducation :

- Collège
- Lycée
- Licence
- Master
- Doctorat
- Autre

### 4. Nombre d'année(s) de pratique du beatbox :

- 1-2 ans
- 2-3 ans
- 3-4 ans
- 5-6 ans
- Plus de 6 ans



## ANNEXE 3 : Questionnaire AI

### 5. Expérience musicale :

- oui
- non

#### 5.1. Si oui : liste d'instrument, style ou technique musicale

- Guitare
- Basse
- Batterie
- Violon
- Contrebasse
- Piano
- Trompette
- Saxophone
- Chant
- Synthétiseur
- Autre

#### 5.2. Nombre d'année(s) de pratique: 4

### 6. Utilisez-vous le beatbox à des fins

- Pédagogiques
- Thérapeutiques
- Aucun des deux

*Par pédagogique, nous entendons si vous enseignez le beatbox ; par thérapeutique, nous voulons savoir si vous participez à des programmes de rééducation orthophonique, si vous intervenez dans des associations animant des ateliers pour un public en situation de handicap (autisme, surdité, dysphasie...).*

### 7. A quel âge avez-vous débuté le Beatbox ? |1|3|

### 8. Vous êtes

- Débutant
- Intermédiaire
- Avancé
- Professionnel / expert

### 9. Vous pratiquez

- Seul
- En duo
- En groupe
- De plusieurs manières

### 10. Vous vous entraînez

## ANNEXE 3 : Questionnaire AI

- Une fois par jour
- 3 fois par semaine
- Moins de 3 fois par semaine

### 10.1. Si vous vous entraînez de façon quotidienne, indiquez la durée approximative par jour :

- Moins de 2h
- Entre 2h et 4h
- Entre 4h et 6h
- Plus de 6 heures

### 10.2. Si vous vous entraînez moins de 3 fois par semaine, indiquez la fréquence approximative sur le mois

- 1 jour par mois
- Entre 1 et 3 jours par mois
- Entre 3 et 5 jours par mois
- Entre 5 et 10 jours par mois
- Plus de 10 jours par mois

### 11. Indiquez la durée de vos entraînements (en minutes ou en heures) : |\_\_|\_\_| minutes ou |\_\_|\_\_| heures

### 12. De quel(s) style(s) musical(aux) vous inspirez-vous pour Beatboxer ?

<input checked="" type="checkbox"/> Disco	<input type="checkbox"/> Country	<input type="checkbox"/> R'n'B
<input type="checkbox"/> Blues	<input type="checkbox"/> Reggae	<input checked="" type="checkbox"/> Electro
<input checked="" type="checkbox"/> Funk	<input type="checkbox"/> Afro	<input type="checkbox"/> K-pop
<input checked="" type="checkbox"/> Jazz	<input type="checkbox"/> Raï	<input checked="" type="checkbox"/> Zouk
<input type="checkbox"/> Metal	<input type="checkbox"/> Ska	<input type="checkbox"/> makossa
<input type="checkbox"/> Pop	<input type="checkbox"/> Gospel	<input type="checkbox"/> Musique Indi
<input type="checkbox"/> Punk	<input type="checkbox"/> Soul	<input type="checkbox"/> Musique Latino
<input checked="" type="checkbox"/> Rap	<input type="checkbox"/> Kompa	
<input type="checkbox"/> Rock 'n' roll	<input type="checkbox"/> Classique	

### 13. Vous maîtrisez les groupes de sons suivants :

(une réponse par ligne)

Maitrise :	Pas du tout	Moyenne	Bonne	Excellente
Grosse caisse	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Charleston	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Caisse claire	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Intrument à vent (trompettes, saxo)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Scratchs	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## ANNEXE 3 : Questionnaire AI

Lips roll	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Clic roll	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Bass	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Par groupe de sons, nous voulons avoir, par exemple, votre maîtrise des différentes déclinaisons de caisse claire (PF, Inward, PCH etc). Si vous avez une bonne maîtrise de l'ensemble des déclinaisons, mettez « excellent ». A l'inverse, si vous ne maîtrisez pas, mettez « pas du tout »

### 14. Vous utilisez et travaillez beaucoup les ...

- Grosse caisse
- Charleston
- Caisse claire
- Instrument à vent (trompettes, saxo...)
- Scratches
- Lips roll
- Clic roll
- Bass

### 15. Savez-vous produire tous les sons existants à ce jour ?

- Oui
- Non
- Ne sait pas
- C'est impossible

**15.1.** Si c'est impossible, pourquoi ? Le beatbox ne cesse d'évoluer, on découvre de plus en plus de sons.

### 16. Votre maîtrise des aspects suivants :

(une réponse par ligne)

Maîtrise :	Pas du tout	Moyenne	Bonne	Excellente
Précision articulatoire	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Voix	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Le ressenti tactile	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Le ressenti auditif	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Rythme	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Puissance	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Endurance	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Respiration	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fluidité (enchaînement des sons)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vitesse	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 17. Vous travaillez principalement

- Précision articulatoire
- Voix
- Le ressenti tactile

## ANNEXE 3 : Questionnaire AI

- Le ressenti auditif
- Rythme
- Puissance
- Endurance
- Respiration
- Fluidité (enchaînement des sons)
- Vitesse

### 18. Utilisez-vous un métronome pour rester en rythme ?

- Oui
- Non

### 19. Pour beatboxer il faut avoir une prédisposition articulatoire et/ou vocale\*

- D'accord
- Plus ou moins d'accord
- Pas d'accord

*\*prédisposition articulatoire = facilité à bouger les articulateurs comme la langue ou la mâchoire pour coordonner une suite de son de Beatbox*

### 20. Quand je beatboxe je m'appuie surtout sur des voyelles

- D'accord
- Plus ou moins d'accord
- Pas d'accord

### 21. Quand je beatboxe je m'appuie surtout sur des consonnes

- D'accord
- Plus ou moins d'accord
- Pas d'accord

### 22. Quand je beatboxe j'articule plus que quand je parle

- D'accord
- Plus ou moins d'accord
- Pas d'accord

### 23. Quand je beatboxe :

*(une réponse par ligne)*

	Jamais	Parfois	Souvent	Toujours
Je suis assis	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je suis debout	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je suis statique (aucun mouvement)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je balance mon buste	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Je suis en mouvement (danse)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

## ANNEXE 3 : Questionnaire AI

J'ai les pieds bien ancrés au sol	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Je projette ma tête en avant	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je garde les épaules à la même hauteur que quand je parle	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je garde mes épaules plus hautes que quand je parle	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je garde mes épaules plus basses que quand je parle	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 24. J'utilise mes mains/bras pour

(une réponse par ligne)

	Jamais	Parfois	Souvent	Toujours
Accompagner le rythme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Donner le rythme	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Accompagner la mélodie	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Imiter un instrument	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Imiter un DJ	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 25. Visage. Quand je beatboxe

(une réponse par ligne)

	Oui	Non
J'ai l'impression de faire des grimaces	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je fronce les sourcils	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge le nez	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge les oreilles	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Je bouge ma mâchoire de haut en bas	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge ma mâchoire d'avant en arrière	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge ma mâchoire sur les côtés	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je gonfle les joues	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### 26. Appareil phonatoire ; Quand je beatboxe

(une réponse par ligne)

	Oui	Non
J'utilise ma mâchoire pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mes lèvres pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mes joues pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise ma langue pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise ma gorge (arrière de la langue) pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mon nez pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mon larynx (au niveau des cordes vocales) pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du visage	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau des lèvres	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau de la langue	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du palais	<input checked="" type="checkbox"/>	<input type="checkbox"/>

## ANNEXE 3 : Questionnaire AI

Je ressens des sensations au niveau de la gorge (arrière de la gorge)	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du nez	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du larynx (au niveau des cordes vocales) (pendant la production de bass par exemple)	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### 27. Respiration

(une réponse par ligne)

	Jamais	Parfois	Souvent	Toujours
Je suis essoufflé(e) quand je beatboxe	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pour reprendre mon souffle, j'arrête de beatboxer	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pour reprendre mon souffle, j'utilise des sons inspirés (ex : inward bass, inward k-snare)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Pour reprendre mon souffle, j'utilise mon nez mais je continue à beatboxer avec ma bouche	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 28. Parmi ces sons, quels sont ceux qui sont difficiles à produire, douloureux et quel(s) type(s) de douleurs provoquent-ils ?

Sons :	Difficile à produire ?	Douloureux ?	Type de douleur :
Grosse caisse	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Charleston	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Caisse claire	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	Douleurs musculaires ou articulaires
Instrument à vent (trompettes, saxo..)	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	Douleurs vocales
Scratches	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Lips roll	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	mâchoire car décalage de celle-ci; articulaire; musculaire
Clic roll	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	
Bass	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	Douleurs différentes en fonction de la basse

### 29. Une fois que vous maîtrisez ces sons, les douleurs persistent-elles ?

- Oui  
 Non

### 30. Avez-vous déjà consulté un médecin pour un problème lié à votre pratique ?

- Oui  
 Non

# ANNEXE 4 : Questionnaire CJ

## Questionnaire HBB

A remplir par l'investigateur avant de donner le questionnaire

N° du sujet : |C|J|-|0|4|

Consignes : merci de répondre aux questions suivantes. Les ronds (o) indiquent que plusieurs réponses sont possibles, les carrés (☐) indiquent qu'une seule réponse est attendue. Si vous ne comprenez pas une question, vous pouvez demander une explication au médecin investigateur ou à un membre de son équipe.

### 1. Profession :

- Libéral
- Artiste
- Artisan
- Enseignant
- Informatique
- 
- Autres

### 2. Langue(s) parlée(s) couramment :

- Langue 1 : FRANCAIS
- Langue 2 : ANGLAIS
- Langue 3 : ESPAGNOL
- Langue 4 : \_\_\_\_\_
- Langue 5 : \_\_\_\_\_

### 3. Niveau d'éducation :

- Collège
- Lycée
- Licence
- Master
- Doctorat
- Autre

### 4. Nombre d'année(s) de pratique du beatbox :

- 1-2 ans
- 2-3 ans
- 3-4 ans
- 5-6 ans
- Plus de 6 ans

## ANNEXE 4 : Questionnaire CJ

### 5. Expérience musicale :

- oui
- non

#### 5.1. Si oui : liste d'instrument, style ou technique musicale

- Guitare
- Basse
- Batterie
- Violon
- Contrebasse
- Piano
- Trompette
- Saxophone
- Chant
- Synthétiseur
- Autre

#### 5.2. Nombre d'année(s) de pratique: 3

### 6. Utilisez-vous le beatbox à des fins

- Pédagogiques
- Thérapeutiques
- Aucun des deux

*Par pédagogique, nous entendons si vous enseignez le beatbox ; par thérapeutique, nous voulons savoir si vous participez à des programmes de rééducation orthophonique, si vous intervenez dans des associations animant des ateliers pour un public en situation de handicap (autisme, surdit , dysphasie...).*

### 7. A quel  ge avez-vous d but  le Beatbox ? |1|2|

### 8. Vous  tes

- D butant
- Interm diaire
- Avanc 
- Professionnel / expert

### 9. Vous pratiquez

- Seul
- En duo
- En groupe
- De plusieurs mani res

### 10. Vous vous entrainez



## ANNEXE 4 : Questionnaire CJ

- Une fois par jour
- 3 fois par semaine
- Moins de 3 fois par semaine

### 10.1. Si vous vous entraînez de façon quotidienne, indiquez la durée approximative par jour :

- Moins de 2h
- Entre 2h et 4h
- Entre 4h et 6h
- Plus de 6 heures

### 10.2. Si vous vous entraînez moins de 3 fois par semaine, indiquez la fréquence approximative sur le mois

- 1 jour par mois
- Entre 1 et 3 jours par mois
- Entre 3 et 5 jours par mois
- Entre 5 et 10 jours par mois
- Plus de 10 jours par mois

### 11. Indiquez la durée de vos entraînements (en minutes ou en heures) : |\_\_|\_\_| minutes ou |\_\_|\_\_| heures

### 12. De quel(s) style(s) musical(aux) vous inspirez-vous pour Beatboxer ?

<input checked="" type="checkbox"/> Disco	<input type="checkbox"/> Country	<input checked="" type="checkbox"/> R'n'B
<input checked="" type="checkbox"/> Blues	<input checked="" type="checkbox"/> Reggae	<input checked="" type="checkbox"/> Electro
<input checked="" type="checkbox"/> Funk	<input checked="" type="checkbox"/> Afro	<input checked="" type="checkbox"/> K-pop
<input checked="" type="checkbox"/> Jazz	<input type="checkbox"/> Raï	<input checked="" type="checkbox"/> Zouk
<input type="checkbox"/> Metal	<input type="checkbox"/> Ska	<input type="checkbox"/> makossa
<input checked="" type="checkbox"/> Pop	<input type="checkbox"/> Gospel	<input type="checkbox"/> Musique Indi
<input type="checkbox"/> Punk	<input checked="" type="checkbox"/> Soul	<input checked="" type="checkbox"/> Musique Latino
<input checked="" type="checkbox"/> Rap	<input type="checkbox"/> Kompa	
<input checked="" type="checkbox"/> Rock 'n' roll	<input checked="" type="checkbox"/> Classique	

### 13. Vous maîtrisez les groupes de sons suivants :

(une réponse par ligne)

Maitrise :	Pas du tout	Moyenne	Bonne	Excellente
Grosse caisse	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Charleston	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Caisse claire	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Intrument à vent (trompettes, saxo)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Scratches	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## ANNEXE 4 : Questionnaire CJ

Lips roll	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Clic roll	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Bass	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Par groupe de sons, nous voulons avoir, par exemple, votre maîtrise des différentes déclinaisons de caisse claires (PF, Inward, PCH etc). Si vous avez une bonne maîtrise de l'ensemble des déclinaisons, mettez « excellent ». A l'inverse, si vous ne maîtrisez pas, mettez « pas du tout »

### 14. Vous utilisez et travaillez beaucoup les ...

- Grosse caisse
- Charleston
- Caisse claire
- Instrument à vent (trompettes, saxo...)
- Scratches
- Lips roll
- Clic roll
- Bass

### 15. Savez-vous produire tous les sons existants à ce jour ?

- Oui
- Non
- Ne sait pas
- C'est impossible

#### 15.1. Si c'est impossible, pourquoi ?

### 16. Votre maîtrise des aspects suivants :

(une réponse par ligne)

Maitrise :	Pas du tout	Moyenne	Bonne	Excellente
Précision articulatoire	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Voix	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le ressenti tactile	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le ressenti auditif	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Rythme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Puissance	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Endurance	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Respiration	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Fluidité (enchaînement des sons)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Vitesse	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

### 17. Vous travaillez principalement

- Précision articulatoire
- Voix
- Le ressenti tactile
- Le ressenti auditif

## ANNEXE 4 : Questionnaire CJ

- Rythme
- Puissance
- Endurance
- Respiration
- Fluidité (enchaînement des sons)
- Vitesse

### 18. Utilisez-vous un métronome pour rester en rythme ?

- Oui
- Non

### 19. Pour beatboxer il faut avoir une prédisposition articulatoire et/ou vocale\*

- D'accord
- Plus ou moins d'accord
- Pas d'accord

*\*prédisposition articulatoire = facilité à bouger les articulateurs comme la langue ou la mâchoire pour coordonner une suite de son de Beatbox*

### 20. Quand je beatboxe je m'appuie surtout sur des voyelles

- D'accord
- Plus ou moins d'accord
- Pas d'accord

### 21. Quand je beatboxe je m'appuie surtout sur des consonnes

- D'accord
- Plus ou moins d'accord
- Pas d'accord

### 22. Quand je beatboxe j'articule plus que quand je parle

- D'accord
- Plus ou moins d'accord
- Pas d'accord

### 23. Quand je beatboxe :

*(une réponse par ligne)*

	Jamais	Parfois	Souvent	Toujours
Je suis assis	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je suis debout	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je suis statique (aucun mouvement)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je balance mon buste	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Je suis en mouvement (danse)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'ai les pieds bien ancrés au sol	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

## ANNEXE 4 : Questionnaire CJ

Je projette ma tête en avant	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je garde les épaules à la même hauteur que quand je parle	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je garde mes épaules plus hautes que quand je parle	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je garde mes épaules plus basses que quand je parle	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 24. J'utilise mes mains/bras pour

(une réponse par ligne)

	Jamais	Parfois	Souvent	Toujours
Accompagner le rythme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Donner le rythme	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Accompagner la mélodie	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Imiter un instrument	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Imiter un DJ	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 25. Visage. Quand je beatboxe

(une réponse par ligne)

	Oui	Non
J'ai l'impression de faire des grimaces	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je fronce les sourcils	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge le nez	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge les oreilles	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge ma mâchoire de haut en bas	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge ma mâchoire d'avant en arrière	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge ma mâchoire sur les côtés	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je gonfle les joues	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### 26. Appareil phonatoire ; Quand je beatboxe

(une réponse par ligne)

	Oui	Non
J'utilise ma mâchoire pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mes lèvres pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mes joues pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise ma langue pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise ma gorge (arrière de la langue) pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mon nez pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mon larynx (au niveau des cordes vocales) pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du visage	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau des lèvres	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau de la langue	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du palais	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau de la gorge (arrière de la gorge)	<input checked="" type="checkbox"/>	<input type="checkbox"/>

## ANNEXE 4 : Questionnaire CJ

Je ressens des sensations au niveau du nez	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du larynx (au niveau des cordes vocales) (pendant la production de bass par exemple)	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### 27. Respiration

(une réponse par ligne)

	Jamais	Parfois	Souvent	Toujours
Je suis essoufflé(e) quand je beatboxe	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pour reprendre mon souffle, j'arrête de beatboxer	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pour reprendre mon souffle, j'utilise des sons inspirés (ex : inward bass, inward k-snare)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Pour reprendre mon souffle, j'utilise mon nez mais je continue à beatboxer avec ma bouche	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 28. Parmi ces sons, quels sont ceux qui sont difficiles à produire, douloureux et quel(s) type(s) de douleurs provoquent-ils ?

Sons :	Difficile à produire ?	Douloureux ?	Type de douleur :
Grosse caisse	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Charleston	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Caisse claire	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Instrument à vent (trompettes, saxo..)	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Scratches	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Lips roll	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Clic roll	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Bass	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	Irritations

### 29. Une fois que vous maîtrisez ces sons, les douleurs persistent-elles ?

- Oui  
 Non

### 30. Avez-vous déjà consulté un médecin pour un problème lié à votre pratique ?

- Oui  
 Non

# ANNEXE 5 : Questionnaire GA

## Questionnaire HBB

A remplir par l'investigateur avant de donner le questionnaire

N° du sujet : |G|A|-|0|5|

Consignes : merci de répondre aux questions suivantes. Les ronds (o) indiquent que plusieurs réponses sont possibles, les carrés (☐) indiquent qu'une seule réponse est attendue. Si vous ne comprenez pas une question, vous pouvez demander une explication au médecin investigateur ou à un membre de son équipe.

### 1. Profession :

- Libéral
- Artiste
- Artisan
- Enseignant
- Informatique
- 
- Autres

### 2. Langue(s) parlée(s) couramment :

- o Langue 1 : FRANCAIS
- o Langue 2 : ANGLAIS
- o Langue 3 : \_\_\_\_\_
- o Langue 4 : \_\_\_\_\_
- o Langue 5 : \_\_\_\_\_

### 3. Niveau d'éducation :

- Collège
- Lycée
- Licence
- Master
- Doctorat
- Autre

### 4. Nombre d'année(s) de pratique du beatbox :

- 1-2 ans
- 2-3 ans
- 3-4 ans
- 5-6 ans
- Plus de 6 ans

## ANNEXE 5 : Questionnaire GA

### 5. Expérience musicale :

- oui
- non

#### 5.1. Si oui : liste d'instrument, style ou technique musicale

- Guitare
- Basse
- Batterie
- Violon
- Contrebasse
- Piano
- Trompette
- Saxophone
- Chant
- Synthétiseur
- Autre

#### 5.2. Nombre d'année(s) de pratique: 3

### 6. Utilisez-vous le beatbox à des fins

- Pédagogiques
- Thérapeutiques
- Aucun des deux

*Par pédagogique, nous entendons si vous enseignez le beatbox ; par thérapeutique, nous voulons savoir si vous participez à des programmes de rééducation orthophonique, si vous intervenez dans des associations animant des ateliers pour un public en situation de handicap (autisme, surdité, dysphasie...).*

### 7. A quel âge avez-vous débuté le Beatbox ? |1|2|

### 8. Vous êtes

- Débutant
- Intermédiaire
- Avancé
- Professionnel / expert

### 9. Vous pratiquez

- Seul
- En duo
- En groupe
- De plusieurs manières

### 10. Vous vous entraînez

## ANNEXE 5 : Questionnaire GA

- Une fois par jour
- 3 fois par semaine
- Moins de 3 fois par semaine

### 10.1. Si vous vous entraînez de façon quotidienne, indiquez la durée approximative par jour :

- Moins de 2h
- Entre 2h et 4h
- Entre 4h et 6h
- Plus de 6 heures

### 10.2. Si vous vous entraînez moins de 3 fois par semaine, indiquez la fréquence approximative sur le mois

- 1 jour par mois
- Entre 1 et 3 jours par mois
- Entre 3 et 5 jours par mois
- Entre 5 et 10 jours par mois
- Plus de 10 jours par mois

11. Indiquez la durée de vos entraînements (en minutes ou en heures) : |\_\_|\_\_| minutes ou  
|\_\_|\_\_| heures

### 12. De quel(s) style(s) musical(aux) vous inspirez-vous pour Beatboxer ?

<input type="checkbox"/> Disco	<input type="checkbox"/> Country	<input checked="" type="checkbox"/> R'n'B
<input checked="" type="checkbox"/> Blues	<input checked="" type="checkbox"/> Reggae	<input checked="" type="checkbox"/> Electro
<input checked="" type="checkbox"/> Funk	<input checked="" type="checkbox"/> Afro	<input type="checkbox"/> K-pop
<input checked="" type="checkbox"/> Jazz	<input type="checkbox"/> Raï	<input type="checkbox"/> Zouk
<input checked="" type="checkbox"/> Metal	<input checked="" type="checkbox"/> Ska	<input type="checkbox"/> makossa
<input checked="" type="checkbox"/> Pop	<input type="checkbox"/> Gospel	<input checked="" type="checkbox"/> Musique Indi
<input type="checkbox"/> Punk	<input checked="" type="checkbox"/> Soul	<input type="checkbox"/> Musique Latino
<input checked="" type="checkbox"/> Rap	<input type="checkbox"/> Kompa	
<input type="checkbox"/> Rock 'n' roll	<input checked="" type="checkbox"/> Classique	

### 13. Vous maîtrisez les groupes de sons suivants :

(une réponse par ligne)

Maitrise :	Pas du tout	Moyenne	Bonne	Excellente
Grosse caisse	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Charleston	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Caisse claire	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Intrument à vent (trompettes, saxo)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Scratchs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>



## ANNEXE 5 : Questionnaire GA

Lips roll	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Clic roll	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Bass	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Par groupe de sons, nous voulons avoir, par exemple, votre maîtrise des différentes déclinaisons de caisse claires (PF, Inward, PCH etc). Si vous avez une bonne maîtrise de l'ensemble des déclinaisons, mettez « excellent ». A l'inverse, si vous ne maîtrisez pas, mettez « pas du tout »

### 14. Vous utilisez et travaillez beaucoup les ...

- Grosse caisse
- Charleston
- Caisse claire
- Instrument à vent (trompettes, saxo...)
- Scratches
- Lips roll
- Clic roll
- Bass

### 15. Savez-vous produire tous les sons existants à ce jour ?

- Oui
- Non
- Ne sait pas
- C'est impossible

#### 15.1. Si c'est impossible, pourquoi ?

### 16. Votre maîtrise des aspects suivants :

(une réponse par ligne)

Maitrise :	Pas du tout	Moyenne	Bonne	Excellente
Précision articulatoire	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Voix	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Le ressenti tactile	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Le ressenti auditif	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Rythme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Puissance	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Endurance	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Respiration	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Fluidité (enchaînement des sons)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Vitesse	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

### 17. Vous travaillez principalement

- Précision articulatoire
- Voix
- Le ressenti tactile
- Le ressenti auditif

## ANNEXE 5 : Questionnaire GA

- Rythme
- Puissance
- Endurance
- Respiration
- Fluidité (enchaînement des sons)
- Vitesse

### 18. Utilisez-vous un métronome pour rester en rythme ?

- Oui
- Non

### 19. Pour beatboxer il faut avoir une prédisposition articulatoire et/ou vocale\*

- D'accord
- Plus ou moins d'accord
- Pas d'accord

*\*prédisposition articulatoire = facilité à bouger les articulateurs comme la langue ou la mâchoire pour coordonner une suite de son de Beatbox*

### 20. Quand je beatboxe je m'appuie surtout sur des voyelles

- D'accord
- Plus ou moins d'accord
- Pas d'accord

### 21. Quand je beatboxe je m'appuie surtout sur des consonnes

- D'accord
- Plus ou moins d'accord
- Pas d'accord

### 22. Quand je beatboxe j'articule plus que quand je parle

- D'accord
- Plus ou moins d'accord
- Pas d'accord

### 23. Quand je beatboxe :

*(une réponse par ligne)*

	Jamais	Parfois	Souvent	Toujours
Je suis assis	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je suis debout	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je suis statique (aucun mouvement)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je balance mon buste	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je suis en mouvement (danse)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
J'ai les pieds bien ancrés au sol	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## ANNEXE 5 : Questionnaire GA

Je projette ma tête en avant	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Je garde les épaules à la même hauteur que quand je parle	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je garde mes épaules plus hautes que quand je parle	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je garde mes épaules plus basses que quand je parle	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 24. J'utilise mes mains/bras pour

(une réponse par ligne)

	Jamais	Parfois	Souvent	Toujours
Accompagner le rythme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Donner le rythme	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Accompagner la mélodie	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Imiter un instrument	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Imiter un DJ	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### 25. Visage. Quand je beatboxe

(une réponse par ligne)

	Oui	Non
J'ai l'impression de faire des grimaces	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je fronce les sourcils	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge le nez	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge les oreilles	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Je bouge ma mâchoire de haut en bas	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge ma mâchoire d'avant en arrière	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je bouge ma mâchoire sur les côtés	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je gonfle les joues	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### 26. Appareil phonatoire ; Quand je beatboxe

(une réponse par ligne)

	Oui	Non
J'utilise ma mâchoire pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mes lèvres pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mes joues pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise ma langue pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise ma gorge (arrière de la langue) pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mon nez pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
J'utilise mon larynx (au niveau des cordes vocales) pour produire certains sons	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du visage	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau des lèvres	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau de la langue	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du palais	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau de la gorge (arrière de la gorge)	<input checked="" type="checkbox"/>	<input type="checkbox"/>

## ANNEXE 5 : Questionnaire GA

Je ressens des sensations au niveau du nez	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Je ressens des sensations au niveau du larynx (au niveau des cordes vocales) (pendant la production de bass par exemple)	<input checked="" type="checkbox"/>	<input type="checkbox"/>

### 27. Respiration

(une réponse par ligne)

	Jamais	Parfois	Souvent	Toujours
Je suis essoufflé(e) quand je beatboxe	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pour reprendre mon souffle, j'arrête de beatboxer	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pour reprendre mon souffle, j'utilise des sons inspirés (ex : inward bass, inward k-snare)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Pour reprendre mon souffle, j'utilise mon nez mais je continue à beatboxer avec ma bouche	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### 28. Parmi ces sons, quels sont ceux qui sont difficiles à produire, douloureux et quel(s) type(s) de douleurs provoquent-ils ?

Sons :	Difficile à produire ?	Douloureux ?	Type de douleur :
Grosse caisse	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Charleston	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Caisse claire	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Instrument à vent (trompettes, saxo..)	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Scratches	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Lips roll	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Clic roll	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	
Bass	<input type="checkbox"/> oui <input checked="" type="checkbox"/> non	<input checked="" type="checkbox"/> oui <input type="checkbox"/> non	Irritations variantes inward

### 29. Une fois que vous maîtrisez ces sons, les douleurs persistent-elles ?

- Oui  
 Non

### 30. Avez-vous déjà consulté un médecin pour un problème lié à votre pratique ?

- Oui  
 Non

## ANNEXE 6 : Signaux et spectrogrammes des sons du corpus

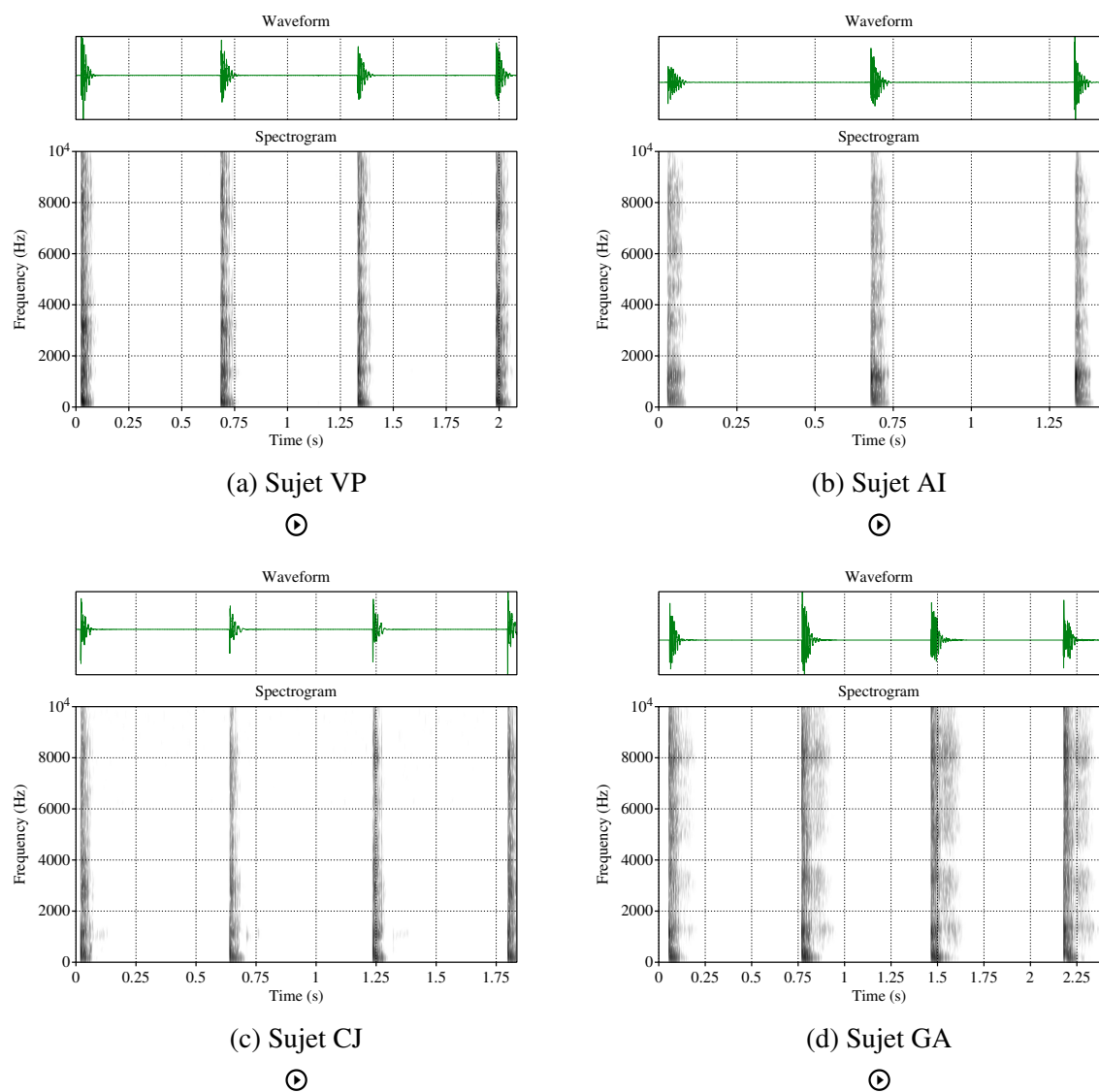
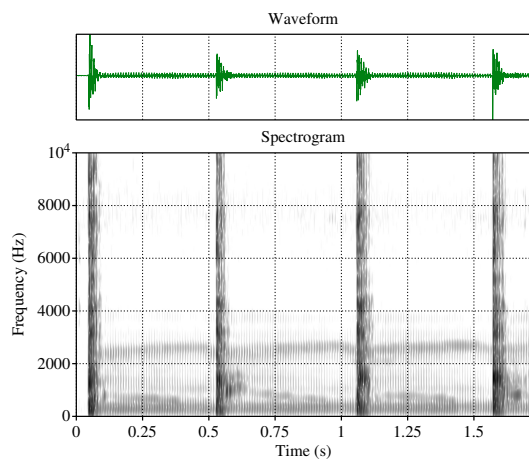
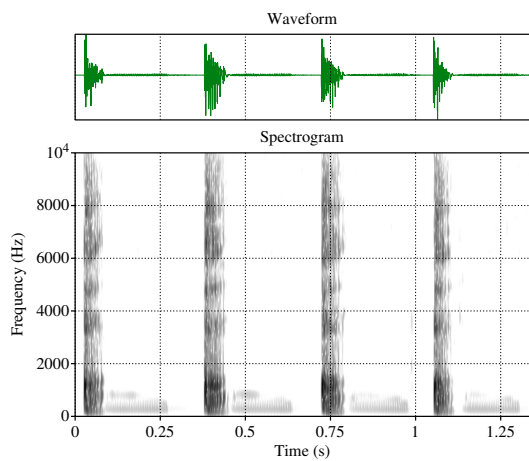


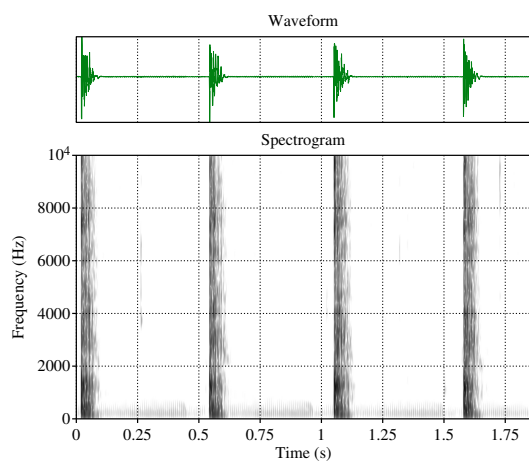
FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes du *Classic Kick [p']*. Attention, échelles de temps différentes.



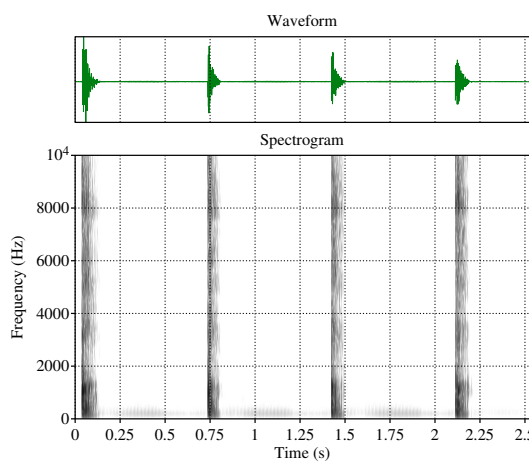
(a) Sujet VP



(b) Sujet AI



(c) Sujet CJ



(d) Sujet GA



FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes du *Classic Kick Humming* {p}. Attention, échelles de temps différentes.

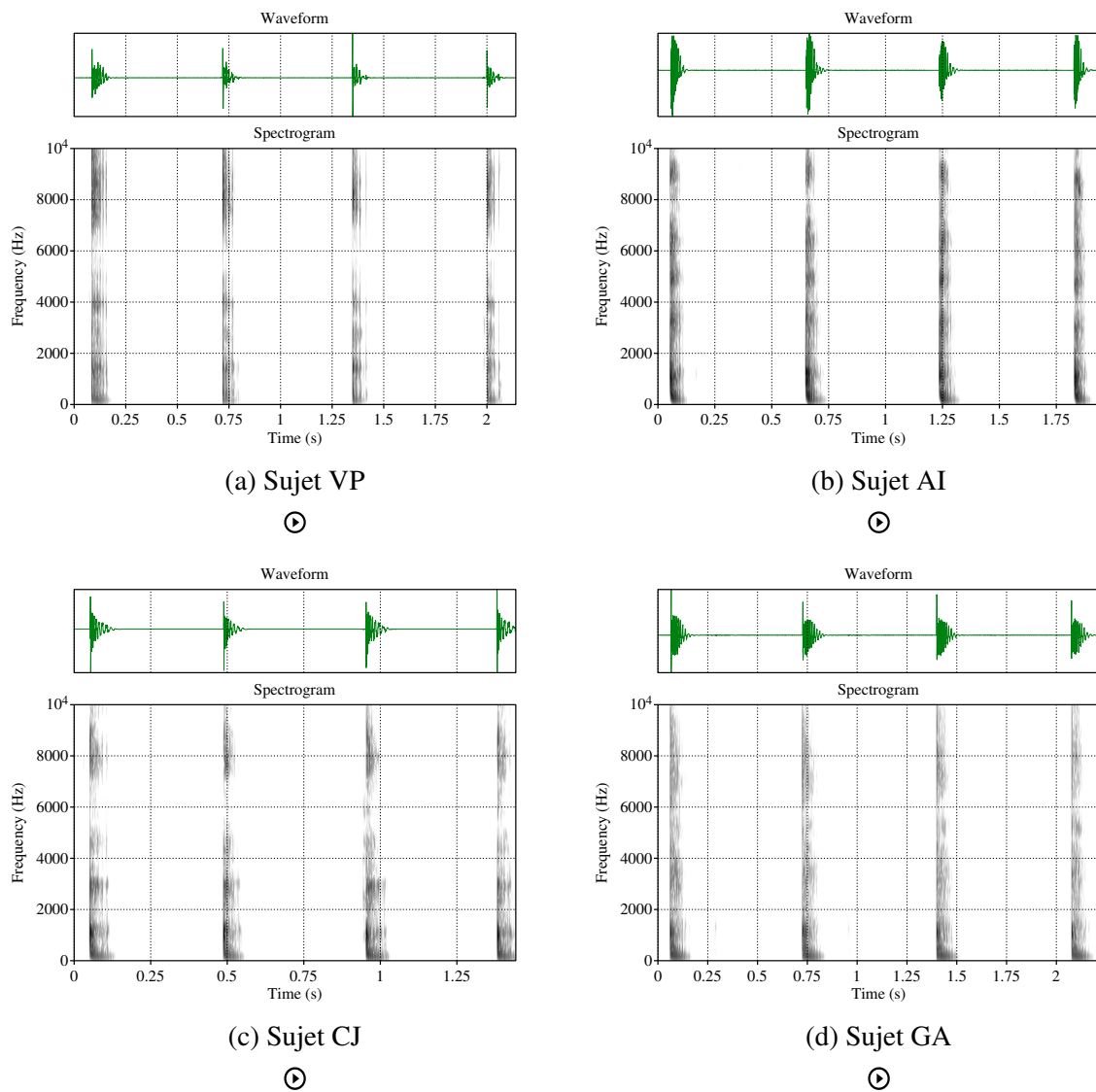
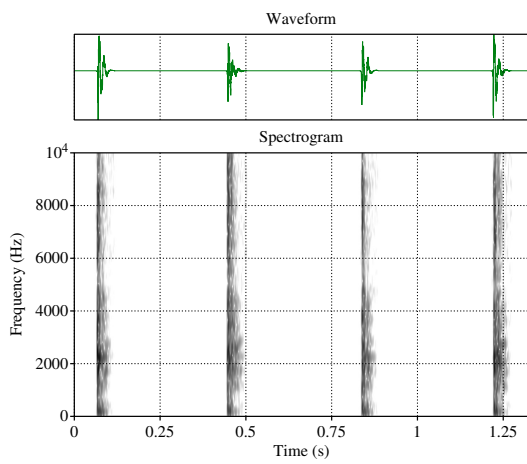
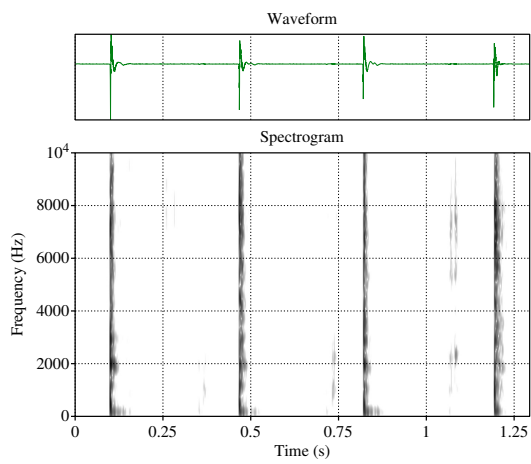


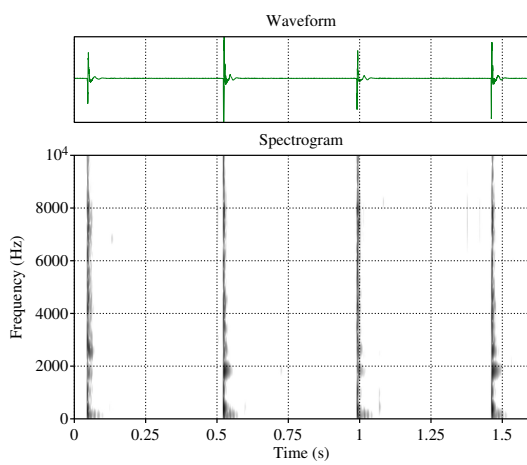
FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes du *Classic Kick Inward* [ $\downarrow p$ ]. Attention, échelles de temps différentes.



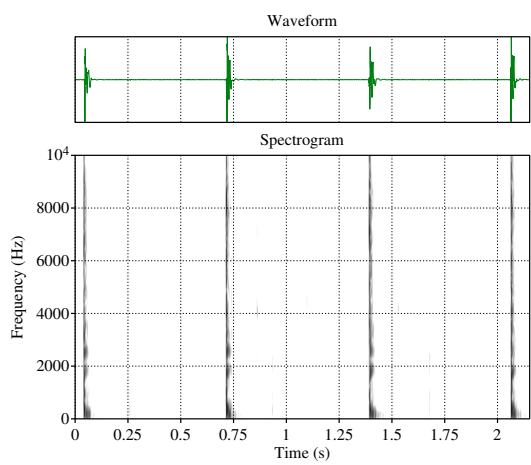
(a) Sujet VP



(b) Sujet AI



(c) Sujet CJ

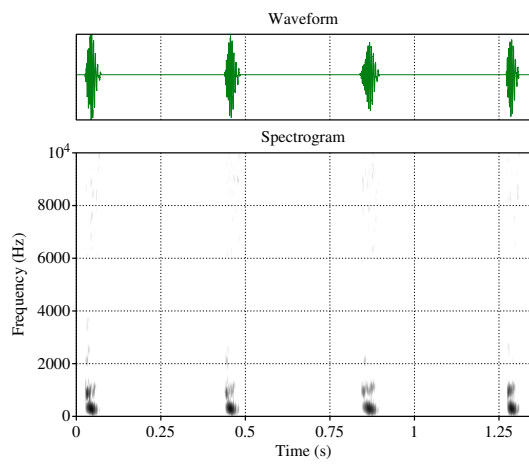


(d) Sujet GA

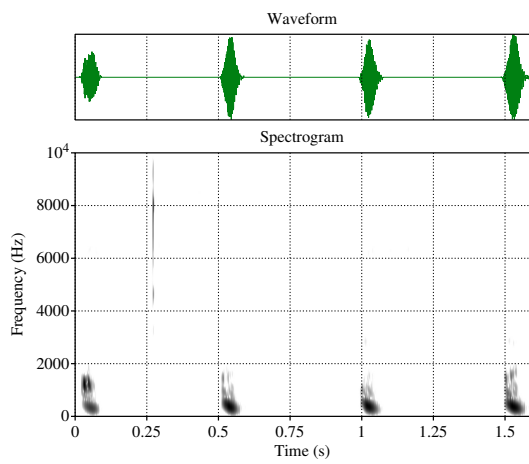


FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes du *Kick Lingual* [t']. Attention, échelles de temps différentes.

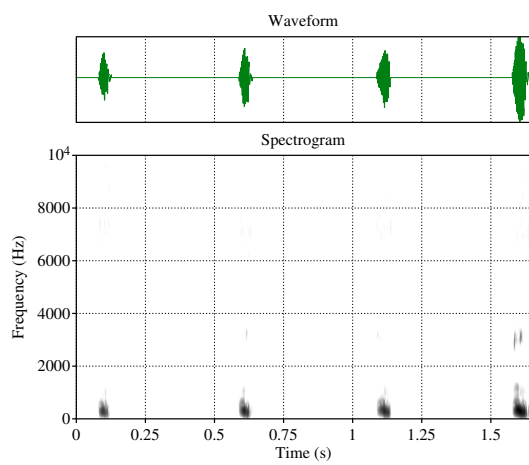




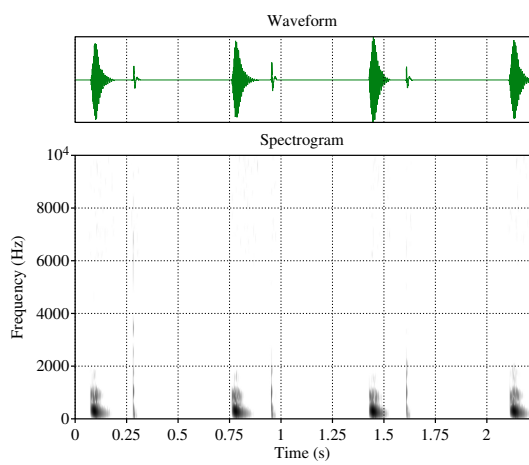
(a) Sujet VP



(b) Sujet AI



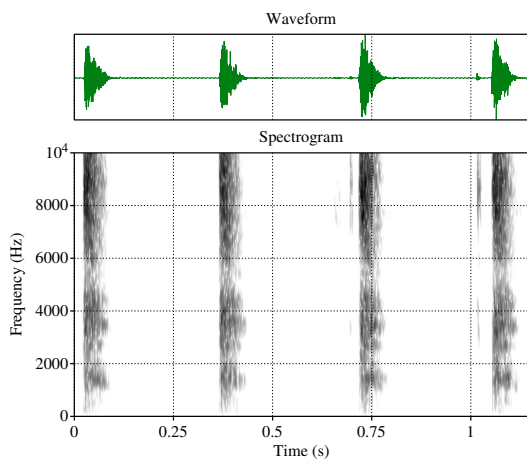
(c) Sujet CJ



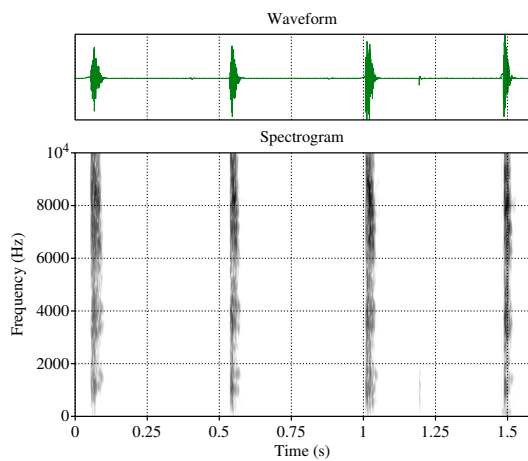
(d) Sujet GA



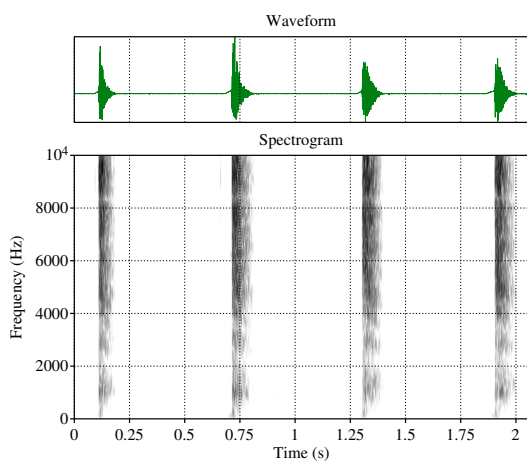
FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes du *Throat Kick* [67]. Attention, échelles de temps différentes.



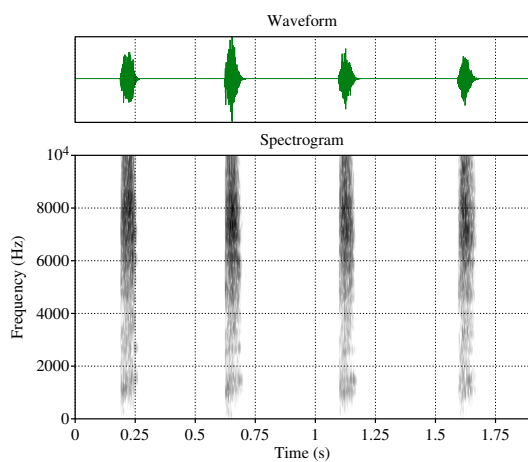
(a) Sujet VP



(b) Sujet AI



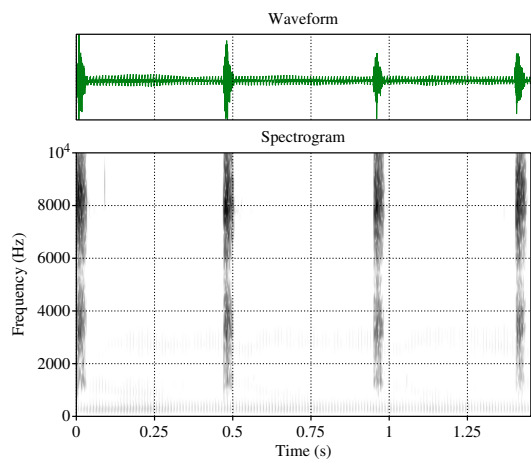
(c) Sujet CJ



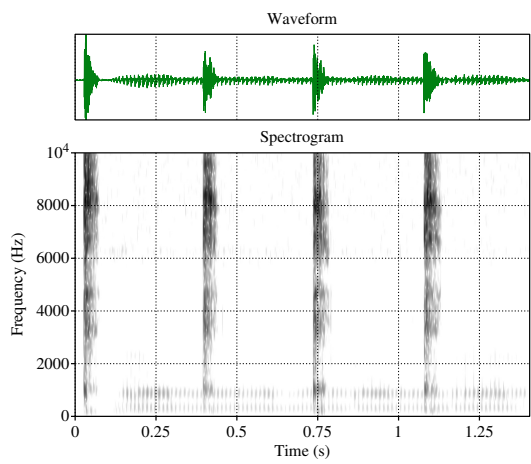
(d) Sujet GA



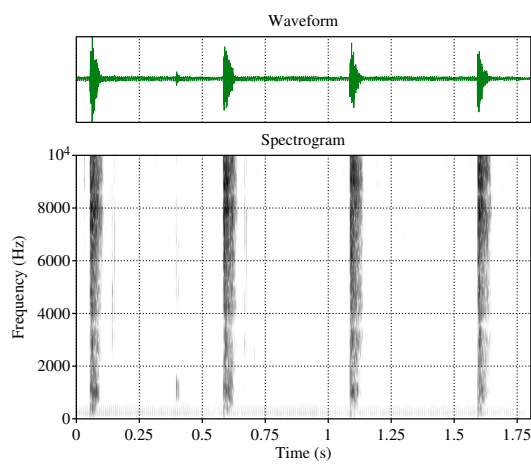
FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes du *Hi Hat Fermé* [ts']. Attention, échelles de temps différentes.



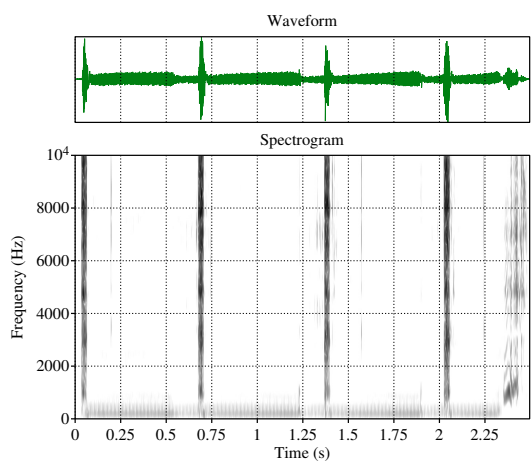
(a) Sujet VP



(b) Sujet AI



(c) Sujet CJ



(d) Sujet GA



FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes du *Hi Hat Fermé Humming*  $\{\hat{ts}\}$ . Attention, échelles de temps différentes.

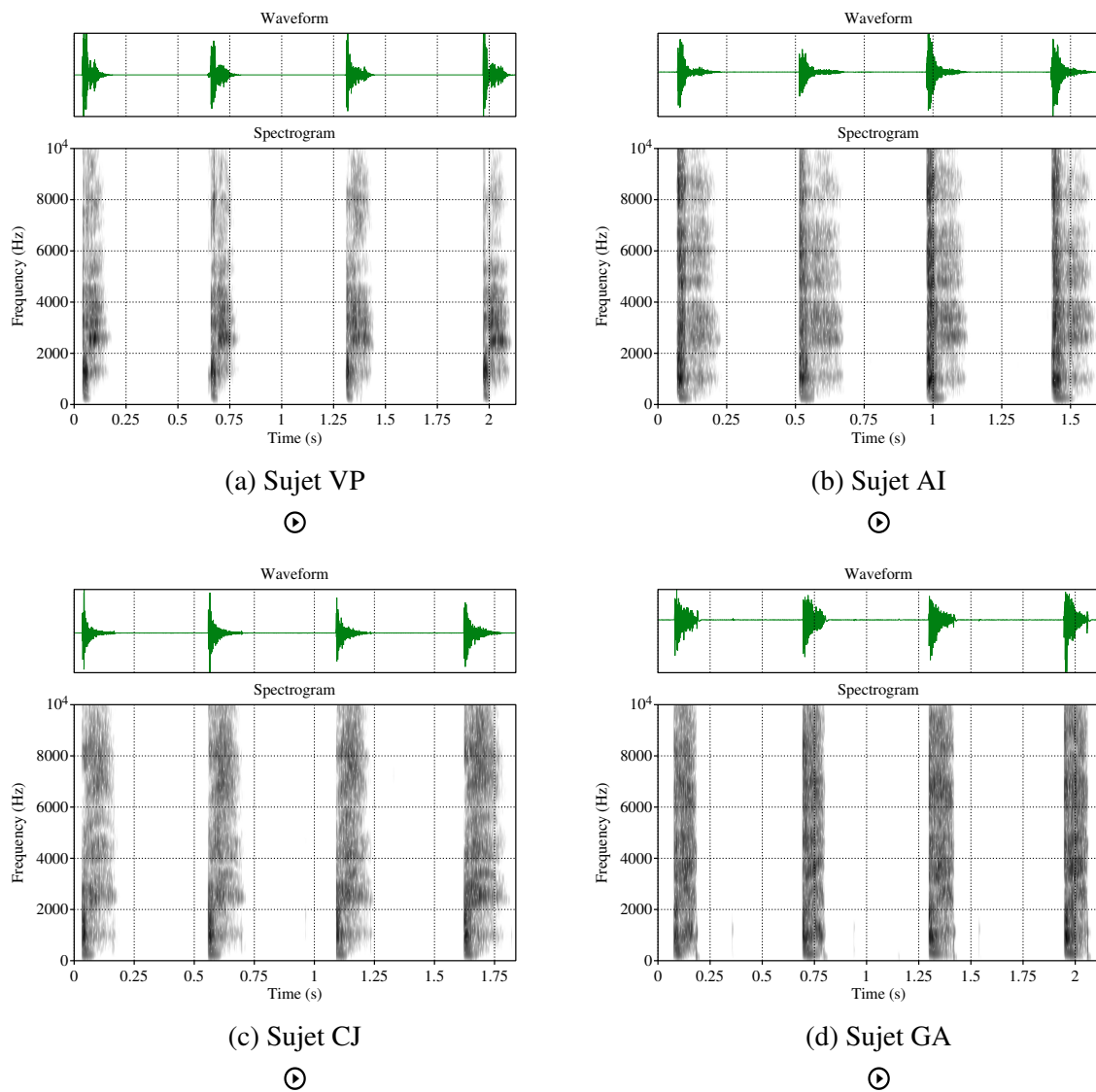
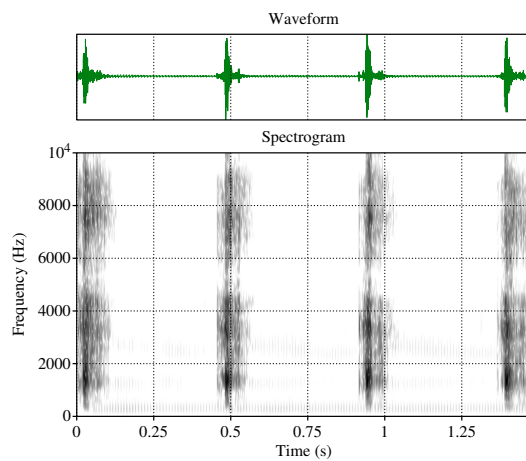
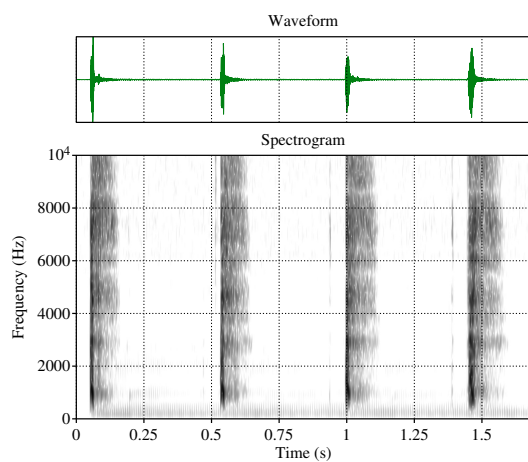


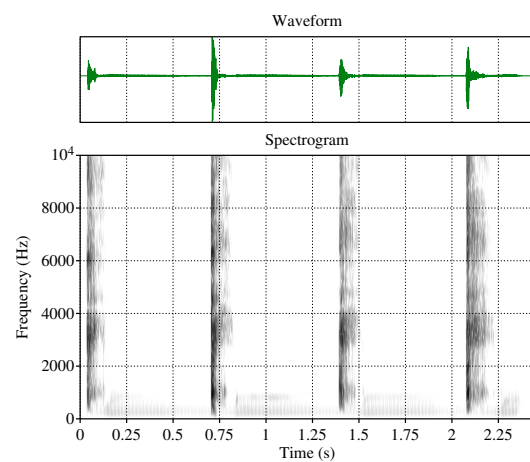
FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes de la *K-Snare Inward* [ $\downarrow$ KL]. Attention, échelles de temps différentes.



(a) Sujet VP



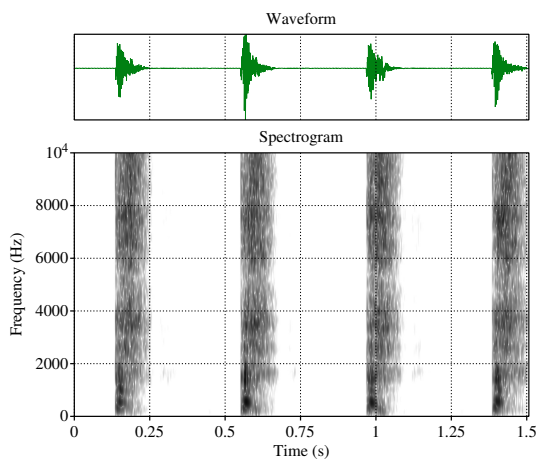
(b) Sujet CJ



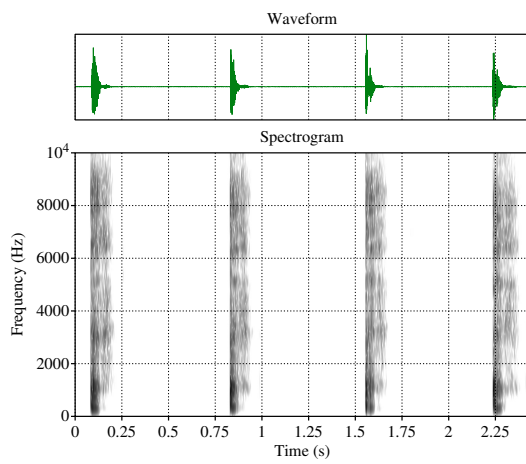
(c) Sujet GA



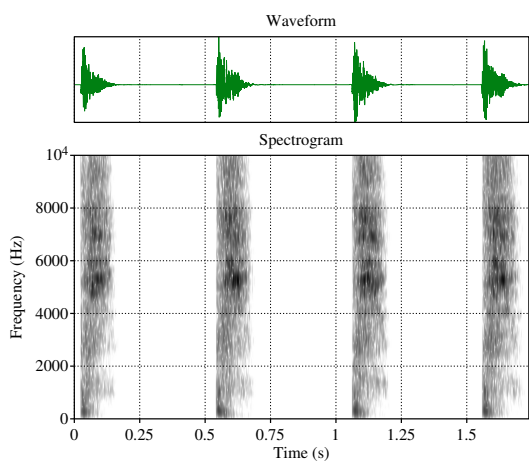
FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes de la *K-Share Humming*  $\{\downarrow_{\text{KL}}\}$  (Pas de données pour le sujet AI). Attention, échelles de temps différentes.



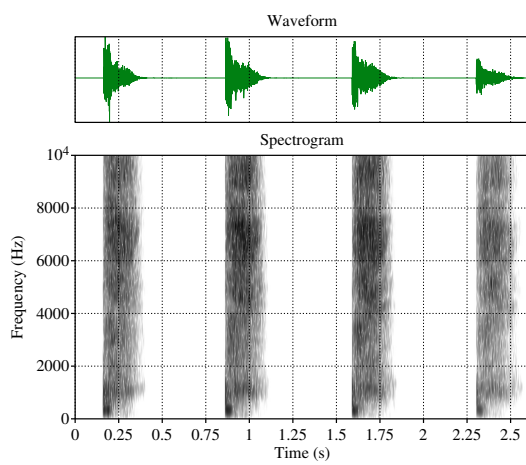
(a) Sujet VP



(b) Sujet AI



(c) Sujet CJ



(d) Sujet GA



FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes de la *PF Snare* [pf:]. Attention, échelles de temps différentes.

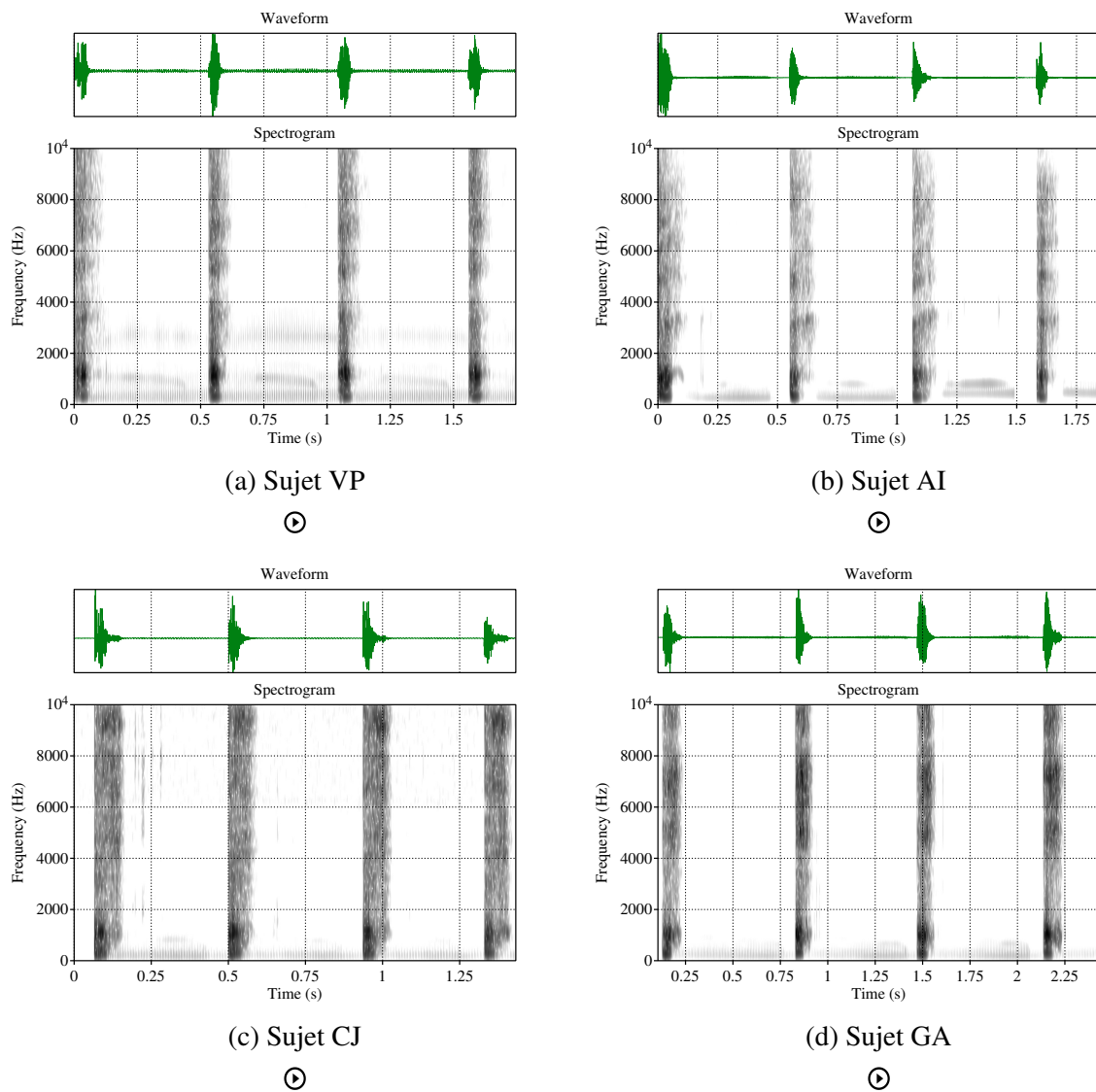
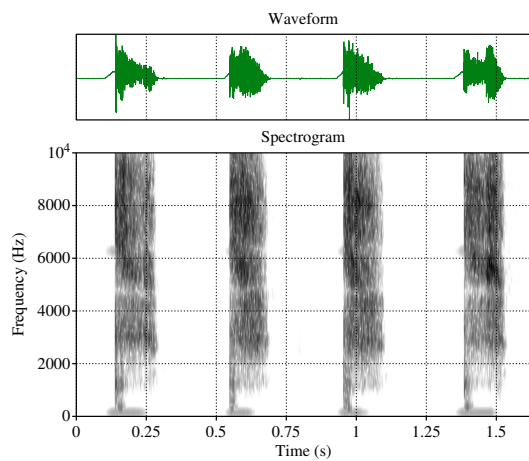
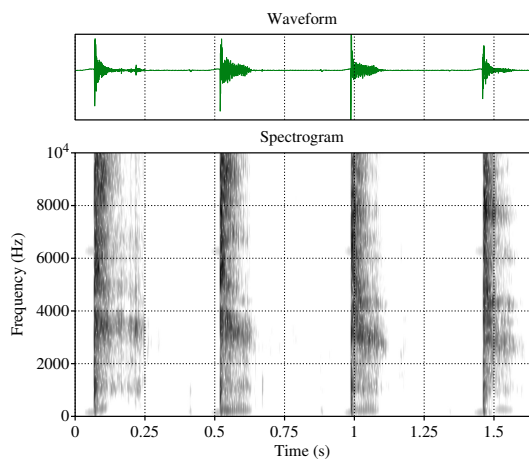


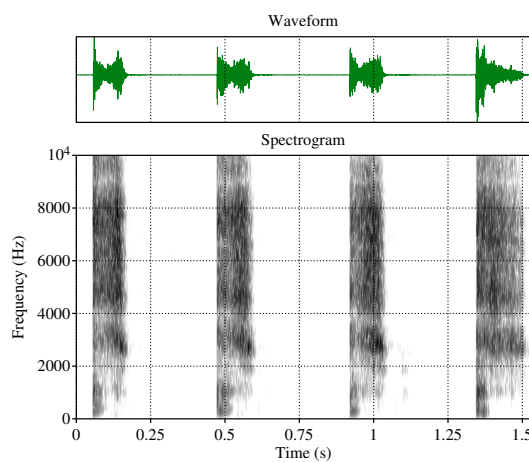
FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes de la *PF Snare Humming* { $\widehat{pf}$ :}. Attention, échelles de temps différentes.



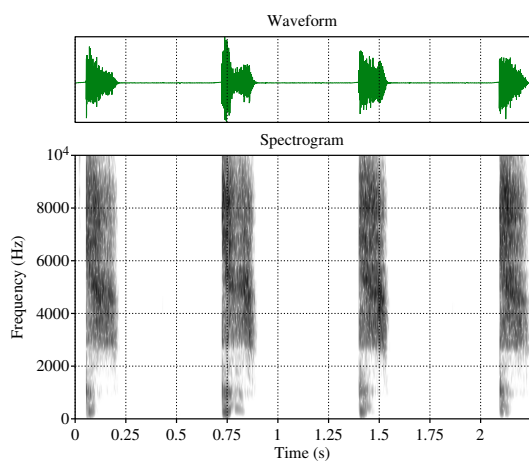
(a) Sujet VP



(b) Sujet AI



(c) Sujet CJ



(d) Sujet GA



FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes de la *Tish Snare* [tʃː]. Attention, échelles de temps différentes.



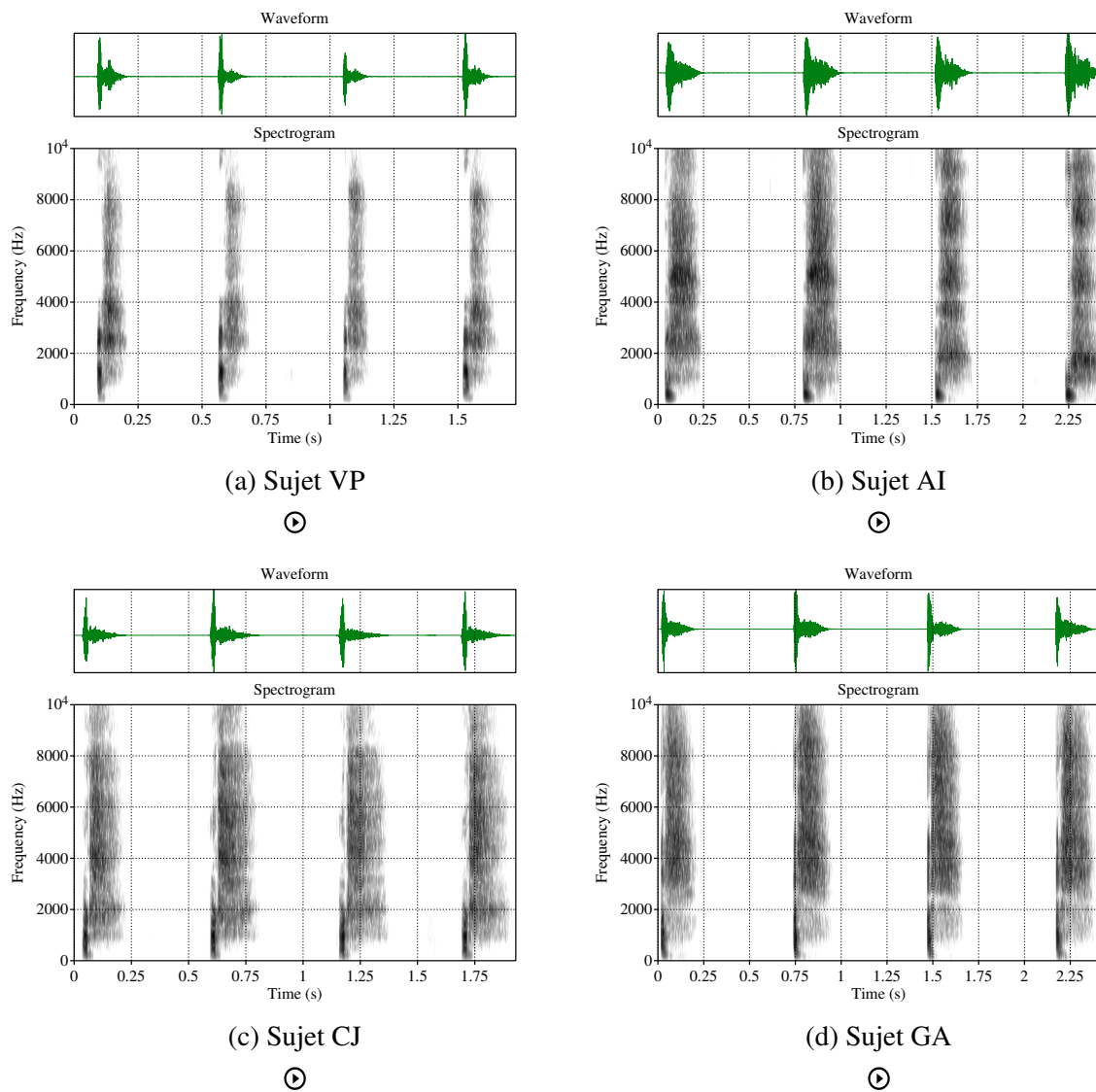


FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes de la *ESH Snare Outward* [ʔf:]. Attention, échelles de temps différentes.

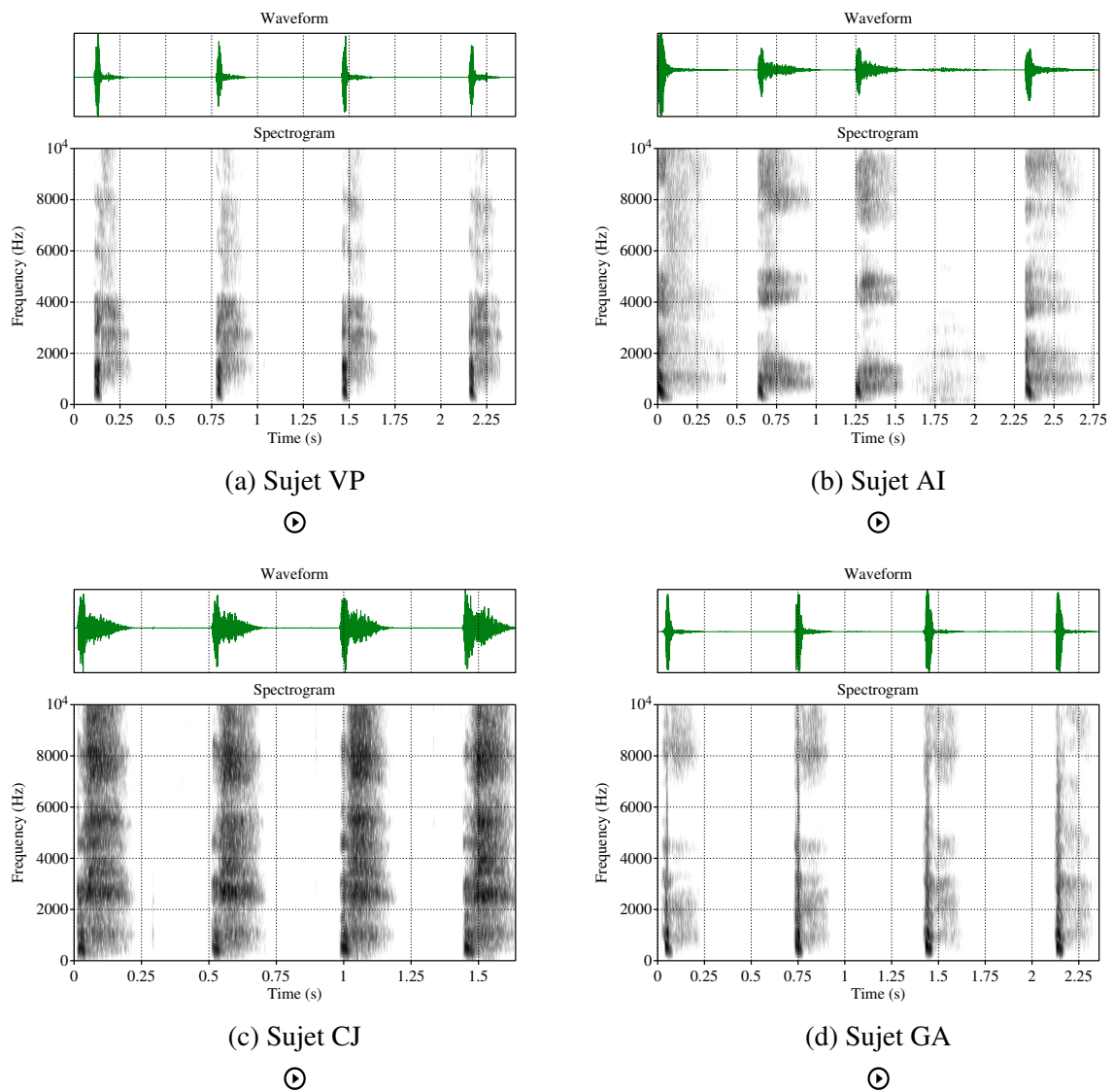


FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes de la *ESH Snare Inward* [↓↑]:. Attention, échelles de temps différentes.

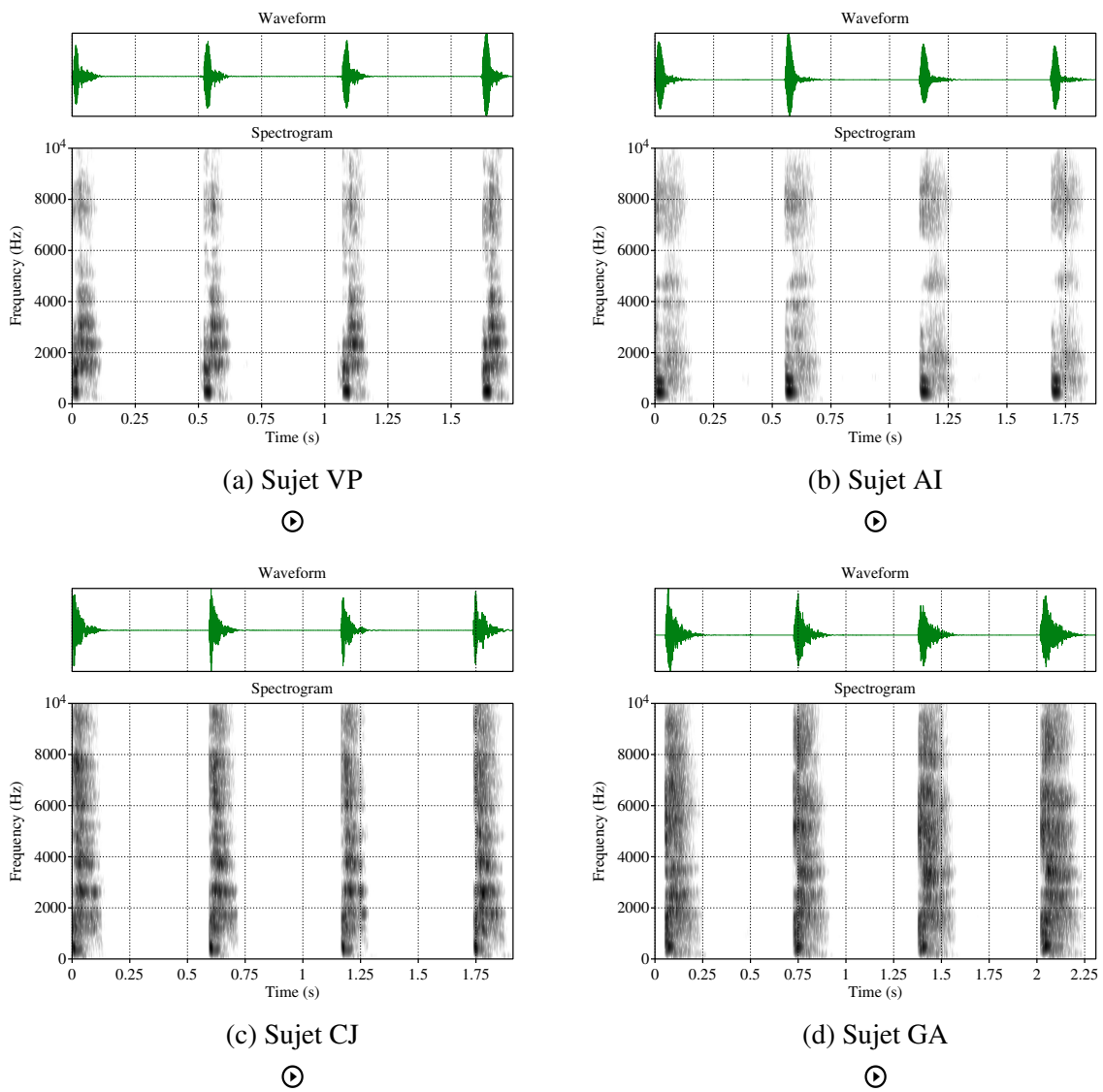
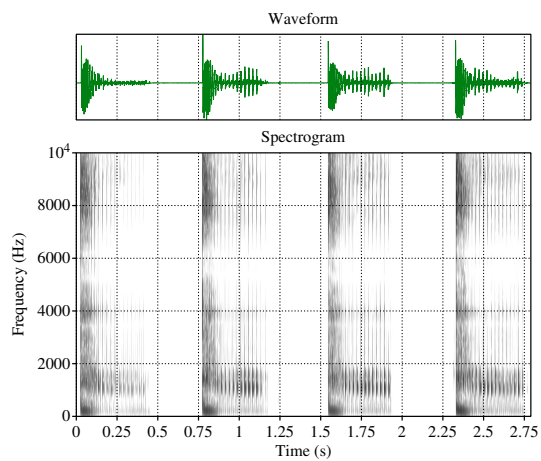
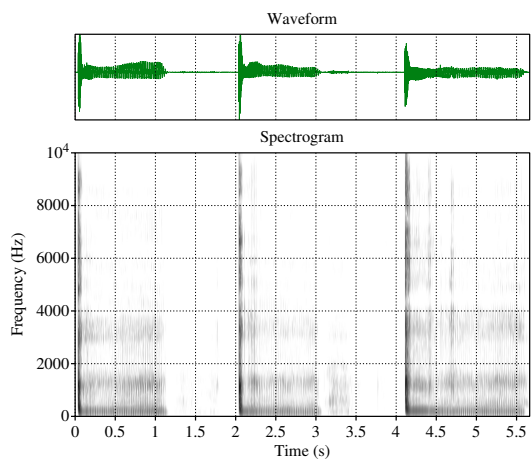


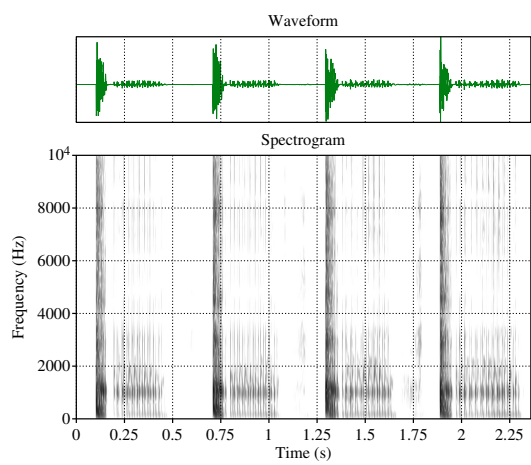
FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes de la *Cough Snare* [ʔh]. Attention, échelles de temps différentes.



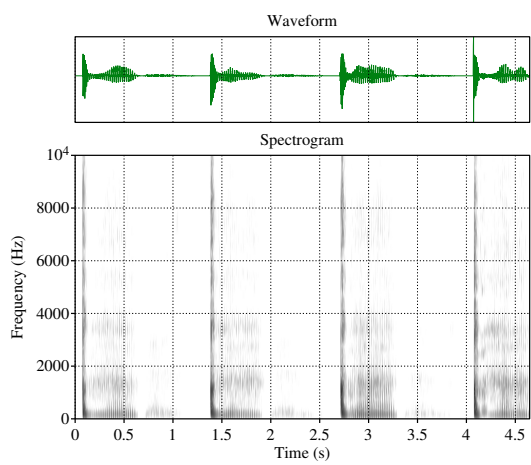
(a) Sujet VP



(b) Sujet AI



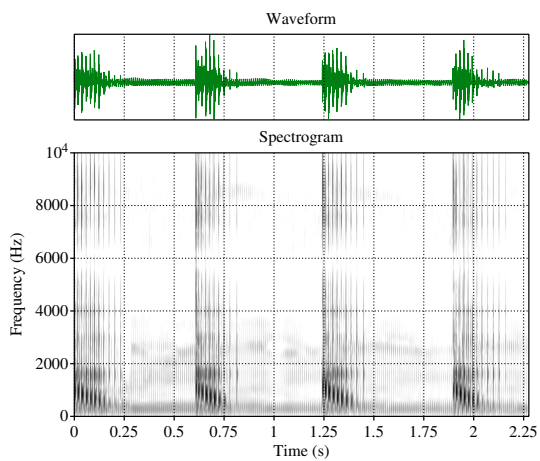
(c) Sujet CJ



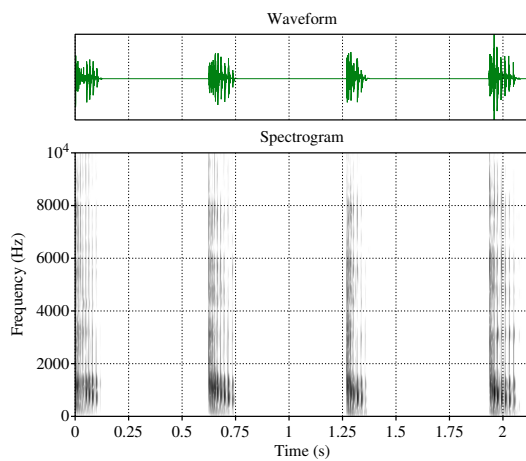
(d) Sujet GA



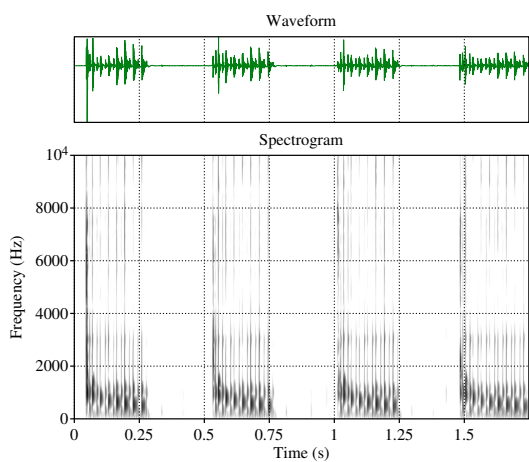
FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes du *Lips Roll* [ $\downarrow_B^1$ ]. Attention, échelles de temps différentes.



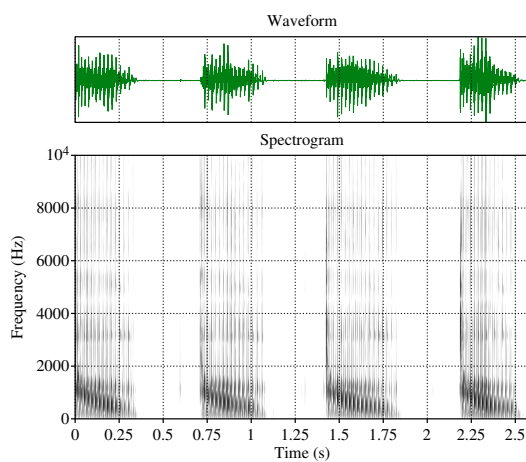
(a) Sujet VP



(b) Sujet AI



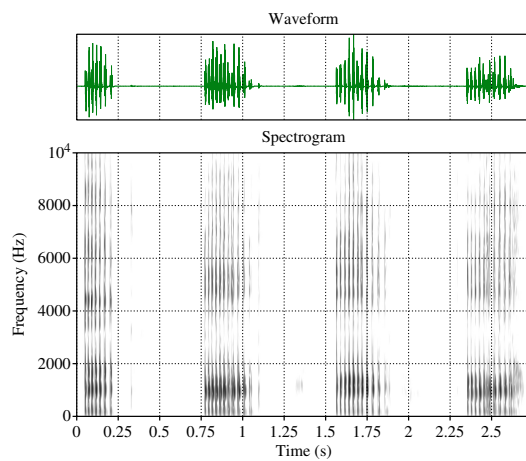
(c) Sujet CJ



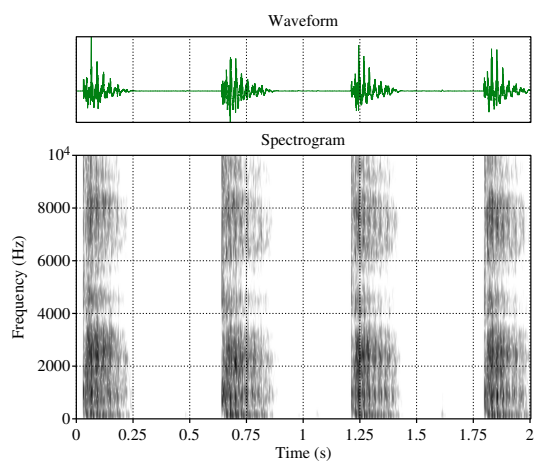
(d) Sujet GA



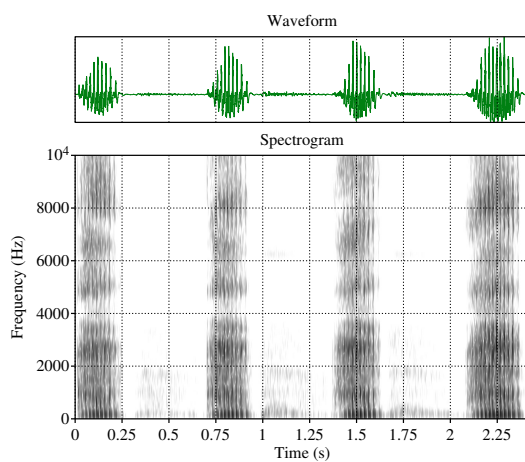
FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes du *Lips Roll Humming*  $\{\downarrow_{\mathbb{B}^1}\}$ . Attention, échelles de temps différentes.



(a) Sujet AI



(b) Sujet CJ



(c) Sujet GA



FIGURE : Comparaison intersujet des signaux acoustiques et des spectrogrammes du *Clic Roll* [ɹ̥]. (Pas de données pour le sujet VP). Attention, échelles de temps différentes.

# BIBLIOGRAPHIE

- Adamska-Sałaciak, A. (1998). Jan baudouin de courtenay's contribution to linguistic theory. *Historiographia Linguistica*, 25(1), 25-60. <https://doi.org/10.1075/hl.25.1-2.05ada>
- Alpatov, V. (2021). La place de Jan Baudouin de Courtenay dans l'histoire des sciences du langage. *Cahiers du Centre de Linguistique et des Sciences du Langage*, (65), 151-170. <https://doi.org/10.26034/la.cdcls.2021.1366>
- Ball, M. J., Howard, S. J. & Miller, K. (2018). Revisions to the extIPA chart. *Journal of the International Phonetic Association*, 48(2), 155-164. <https://doi.org/10.1017/S0025100317000147>
- Betillouloux, A. (2020, juin 15). *Ella Fitzgerald, pionnière du jazz* [France Musique]. <https://www.radiofrance.fr/francemusique/video-ella-fitzgerald-pionniere-du-jazz-1245807>
- Blaylock, R. (2022). *Beatboxing Phonology* (Doctoral Dissertation). University of Southern California.
- Blaylock, R., Patil, N., Greer, T. & Narayanan, S. S. (2017). Sounds of the human vocal tract. *Interspeech 2017*, 2287-2291. <https://doi.org/10.21437/Interspeech.2017-1631>
- Boersma, P. & Weenink, D. (2001). *Praat : doing phonetics by computer* (Version 6.2.06). <https://www.fon.hum.uva.nl/praat/>
- Bourdin, V. & Navion, A. (2013). *Mesure de l'efficacité vocale au sein d'une population de chanteur de Human Beatbox : Analyse acoustique, aérodynamique et observations comportementales* (Master's Thesis). Université Claude Bernard Lyon 1.
- Browman, C. P. & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, 6(2), 201-251. <https://doi.org/10.1017/S0952675700001019>
- Browman, C. P. & Goldstein, L. (1992). Articulatory phonology : an overview. *Phonetica*, 49(3), 155-180. <https://doi.org/10.1159/000261913>
- Byrd, D. & Tan, C. C. (1996). Saying consonant clusters quickly. *Journal of Phonetics*, 24(2), 263-282. <https://doi.org/10.1006/jpho.1996.0014>
- Castellengo, M. (2015). *Écoute musicale et acoustique : avec 420 sons et leurs sonagrammes décryptés*. Eyrolles.
- Catford, J. C. (1977). *Fundamental problems in phonetics*. Indiana University Press.
- Chitoran, I. & Cohn, A. C. (2009). Complexity in phonetics and phonology : gradience, categoriality, and naturalness. In F. Pellegrino, E. Marsico, I. Chitoran & C. Coupé (Éd.), *Approaches to phonological complexity* (pp. 21-46). Mouton de Gruyter. <https://doi.org/10.1515/9783110223958.19>
- Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. Harper & Row, Publishers.
- Dalla Bella, S. & Palmer, C. (2011). Rate effects on timing, key velocity, and finger kinematics in piano performance. *PLoS ONE*, 6(6), e20518. <https://doi.org/10.1371/journal.pone.0020518>
- De Torcy, T., Clouet, A., Pillot-Loiseau, C., Vaissière, J., Brasnu, D. & Crevier-Buchman, L. (2014). A video-fiberscopic study of laryngopharyngeal behaviour in the *human beat-*



- box. *Logopedics Phoniatrics Vocology*, 39(1), 38-48. <https://doi.org/10.3109/14015439.2013.784801>
- Dehais-Underdown, A., Buchman, L. & Demolin, D. (2019). Acoustico-Physiological coordination in the Human Beatbox : A pilot study on the beatboxed Classic Kick Drum. *19th International Congress of Phonetic Sciences*. <https://hal.science/hal-02284132>
- Dehais-Underdown, A., Vignes, P., Crevier-Buchman, L. & Demolin, D. (2021). In and out : production mechanisms in Human Beatboxing. *Proceedings of Meetings on Acoustics*, 45(1), 060005. <https://doi.org/10.1121/2.0001543>
- Demolin, D., Ngonga-Ke-Mbembe, H. & Soquet, A. (2002). Phonetic characteristics of an unexploded palatal implosive in hendo. *Journal of the International Phonetic Association*, 32(1), 1-15. <https://doi.org/10.1017/S0025100302000117>
- Dimitropoulos, K., Manitsaris, S., Tsalakanidou, F., Nikolopoulos, S., Denby, B., Kork, S. A., Crevier-Buchman, L., Pillot-Loiseau, C., Adda-Decker, M., Dupont, S., Tilmanne, J., Ott, M., Alivizatou, M., Yilmaz, E., Hadjileontiadis, L., Charisis, V., Deroo, O., Manitsaris, A., Kompatsiaris, I. & Grammalidis, N. (2014). Capturing the intangible an introduction to the i-Treasures project. *2014 International Conference on Computer Vision Theory and Applications (VISAPP)*, 2, 773-781. <https://doi.org/10.5220/0004871607730781>
- Donegan, P. J. & Stampe, D. (1979). The Study of Natural Phonology. In D. Dinnsen (Éd.), *Current Approaches to Phonological Theory*. Indiana University Press. <https://doi.org/10.2979/CurrentApproachestoP>
- 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. <https://doi.org/10.1017/S095267570600087X>
- Eklund, R. (2008). Pulmonic ingressive phonation : diachronic and synchronic characteristics, distribution and function in animal and human sound production and in human speech. *Journal of the International Phonetic Association*, 38(3), 235-324. <https://doi.org/10.1017/S0025100308003563>
- Esling, J. H., Moïsik, S. R., Benner, A. & Crevier-Buchman, L. (2019). *Voice Quality : The Laryngeal Articulator Model* (1<sup>re</sup> éd.). Cambridge University Press. <https://doi.org/10.1017/9781108696555>
- Fant, G. (1970). *Acoustic Theory of Speech Production : With Calculations based on X-Ray Studies of Russian Articulations*. Mouton.
- Garrigues, L. (2015). *Gestion des sons non pulmonaires et de la phonation inversée en Human Beatbox* (Master's Thesis). Université Pierre et Marie Curie.
- Gay, T. (1968). Effect of speaking rate on diphthong formant movements. *The Journal of the Acoustical Society of America*, 44(6), 1570-1573. <https://doi.org/10.1121/1.1911298>
- Gay, T. (1981). Mechanisms in the control of speech rate. *Phonetica*, 38(1), 148-158. <https://doi.org/10.1159/000260020>
- Gay, T., Ushijima, T., Hiroset, H. & Cooper, F. S. (1974). Effect of speaking rate on labial consonant-vowel articulation. *Journal of Phonetics*, 2(1), 47-63. [https://doi.org/10.1016/S0095-4470\(19\)31176-3](https://doi.org/10.1016/S0095-4470(19)31176-3)
- 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*, 55-58. <https://hal.science/hal-00142982>
- Gick, B., Wilson, I. & Derrick, D. (2013). *Articulatory phonetics*. Wiley-Blackwell.
- Goebel, W. & Palmer, C. (2008). Tactile feedback and timing accuracy in piano performance. *Experimental Brain Research*, 186(3), 471-479. <https://doi.org/10.1007/s00221-007-1252-1>



- Goebel, W. & Palmer, C. (2013). Temporal control and hand movement efficiency in skilled music performance. *PLoS ONE*, 8(1), e50901. <https://doi.org/10.1371/journal.pone.0050901>
- Gordon, M., Barthmaier, P. & Sands, K. (2002). A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association*, 32(2), 141-174. <https://doi.org/10.1017/S0025100302001020>
- Hermes, Z., Fu, M., Rose, S., Shosted, R. & Sutton, B. (2016). Representations of Place and Airstream Mechanism : A real-time MRI study of Tigrinya ejectives. *LabPhon15*.
- Hewlett, N. & Beck, J. M. (2006). *An introduction to the science of phonetics*. L. Erlbaum Associates.
- Hixon, T. (1966). Turbulent noise sources for speech. *Folia Phoniatica et Logopaedica*, 18(3), 168-182. <https://doi.org/10.1159/000263090>
- Hixon, T., Weismer, G. & Hoit, J. D. (2020). *Preclinical speech science : anatomy, physiology, acoustics, and perception* (Third edition). Plural Publishing Inc.
- Hofmann, A. & Goebel, W. (2014). Production and perception of legato, portato, and staccato articulation in saxophone playing. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.00690>
- Hume, E. (2006). Language Specific and Universal Markedness : An Information-theoretic Approach. *Colloquium on Information Theory and Phonology*, Linguistic Society of America.
- International Phonetic Association, I. (1999). *Handbook of the International Phonetic Association : a guide to the use of the international phonetic alphabet*. Cambridge University Press.
- Isshiki, N. (1965). Vocal intensity and air flow rate. *Folia Phoniatica et Logopaedica*, 17(2), 92-104. <https://doi.org/10.1159/000263031>
- Jackendoff, R. (2009). Parallels and nonparallels between language and music. *Music Perception*, 26(3), 195-204. <https://doi.org/10.1525/mp.2009.26.3.195>
- Jakobson, R., Fant, C. G. M. & Halle, M. (1951). *Preliminaries to speech analysis : the distinctive features and their correlates*. MIT Press.
- Jannedy, S., Fuchs, S. & Weirich, M. (2010). Articulation Beyond the Usual : Evaluating the Fastest German Speaker under Laboratory Conditions. In S. Fuchs, P. Hoole, C. Mooshammer & M. Zygis (Éd.), *Between the regular and the particular in speech and language* (p. 205-234). Peter Lang.
- Kingston, J. C. (1984). *The Phonetics and Phonology of the Timing of Oral and Glottal Events* (Doctoral Dissertation). University of California. Berkeley.
- Ladefoged, P. (1971). *Preliminaries to linguistic phonetics*. University of Chicago Press.
- Le Huche, F. & Allali, A. (2010). *La voix. Tome 1, Anatomie et physiologie des organes de la voix et de la parole* (4e édition). Elsevier Masson.
- Lehmann, C. (1974). Isomorphismus im sprachlichen Zeichen. In H. Seiler (Éd.), *Linguistic workshop II : Arbeiten des Kölner Universalienprojekts 1973/4* (p. 98-123). Fink.
- Leipp, E. (1971). *Acoustique et musique*. Eyrolles.
- Liberman, A. M. & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1), 1-36. [https://doi.org/10.1016/0010-0277\(85\)90021-6](https://doi.org/10.1016/0010-0277(85)90021-6)
- Lindblom, B. (1963). Spectrographic study of vowel reduction. *The Journal of the Acoustical Society of America*, 35(11), 1773-1781. <https://doi.org/10.1121/1.1918816>
- Lindblom, B. (1983). Economy of speech gestures. In P. F. MacNeilage (Éd.), *The production of speech* (p. 217-245). Springer. [https://doi.org/10.1007/978-1-4613-8202-7\\_10](https://doi.org/10.1007/978-1-4613-8202-7_10)

- Lindblom, B. (1990). On the notion of “possible speech sound”. *Journal of Phonetics*, 18(2), 135-152. [https://doi.org/10.1016/S0095-4470\(19\)30398-5](https://doi.org/10.1016/S0095-4470(19)30398-5)
- Lindblom, B. & Maddieson, I. (1988). Phonetic Universals in Consonant Systems Bjorn Lindblom, and Ian Maddieson. In V. Fromkin, L. M. Hyman & C. N. Li (Éd.), *Language, speech, and mind : studies in honour of Victoria A. Fromkin* (p. 62-78). Routledge, Chapman ; Hall.
- Maddieson. (1984). *Patterns of Sounds*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511753459>
- Malécot, A. (1969). The effect of syllabic rate and loudness on the force of articulation of american stops and fricatives. *Phonetica*, 19(4), 205-216. <https://doi.org/10.1159/000258628>
- Marchal, M. A. (2011). *Précis de physiologie de la production de la parole*. Solal.
- Martinet, A. (1960). *Éléments de linguistique générale* (5e éd). Armand Colin.
- Martino, R. (2009). *Le Human Beatbox et ses participants* (Master’s Thesis). Université Pierre-Mendès-France. Grenoble, France.
- Meyer, J. (2015). *Whistled languages : a worldwide inquiry on human whistled speech*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-45837-2>
- Miller, A. & Elsner, M. (2017, mars). Click reduction in fluent speech : a semi-automated analysis of Mangetti Dune !Xung. In A. Arppe, J. Good, M. Hulden, J. Lachler, A. Palmer & L. Schwartz (Éd.), *Proceedings of the 2nd Workshop on the Use of Computational Methods in the Study of Endangered Languages* (p. 107-115). Association for Computational Linguistics. <https://doi.org/10.18653/v1/W17-0115>
- Miller, A., Brugman, J., Sands, B., Namaseb, L., Exter, M. & Collins, C. (2009). Differences in airstream and posterior place of articulation among nluu clicks. *Journal of the International Phonetic Association*, 39(2), 129-161. <https://doi.org/10.1017/S0025100309003867>
- Munhall, K. & Löfqvist, A. (1992). Gestural aggregation in speech : laryngeal gestures. *Journal of Phonetics*, 20(1), 111-126. [https://doi.org/10.1016/S0095-4470\(19\)30242-6](https://doi.org/10.1016/S0095-4470(19)30242-6)
- Ohala, J. J. (1997). Aerodynamics of phonology. *Proc. 4th Seoul International Conference on Linguistics*, 92-97.
- Ostry, D. J. & Munhall, K. G. (1985). Control of rate and duration of speech movements. *The Journal of the Acoustical Society of America*, 77(2), 640-648. <https://doi.org/10.1121/1.391882>
- Ozerov, A., Vincent, E. & Bimbot, F. (2012). A General Flexible Framework for the Handling of Prior Information in Audio Source Separation. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(4), 1118-1133. <https://doi.org/10.1109/TASL.2011.2172425>
- Palmer, C., Koopmans, E., Loehr, J. D. & Carter, C. (2009). Movement-related feedback and temporal accuracy in clarinet performance. *Music Perception*, 26(5), 439-449. <https://doi.org/10.1525/mp.2009.26.5.439>
- Park, J., TyTe & Noise, W. (2005). *History of Beatbox : Old School* [Human Beatbox]. <https://www.humanbeatbox.com/articles/history-of-beatboxing-part-2/>
- Paroni, A. (2014). *How do beatboxers play with their tongue and lips ?* (Master’s Thesis). Università degli Studi di Padova.
- Paroni, A. (2022). *Human beatboxing : pushing the boundaries of human voice production* (Doctoral Dissertation). Université Grenoble Alpes. <https://theses.hal.science/tel-03992210>
- Paroni, A., Frère, J., Loevenbruck, H. & Henrich Bernardoni, N. (2022). Analyse électromyographique, acoustique et phonétique de la production de sons beatboxés. *Interaction*

- Musicien-Instrument : regards croisés entre sciences du mouvement humain et acoustique.* <https://hal.science/hal-03931058>
- Paroni, A., Henrich Bernardoni, N., Savariaux, C., Lœvenbruck, H., Calabrese, P., Pellegrini, T., Mouysset, S. & Gerber, S. (2021). 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. <https://doi.org/10.1121/10.0002921>
- Patel, A. D. (2008). *Music, language, and the brain*. Oxford University Press.
- Patil, N., Greer, T., Blaylock, R. & Narayanan, S. S. (2017). Comparison of basic beatboxing articulations between expert and novice artists using real-time magnetic resonance imaging. *Interspeech 2017*, 2277-2281. <https://doi.org/10.21437/Interspeech.2017-1190>
- Pillot-Loiseau, C., Garrigues, L., Demolin, D., Fux, T., Amelot, A. & Crevier-Buchman, L. (2020). Le human beatbox entre musique et parole : quelques indices acoustiques et physiologiques. *Volume!*, (16), 125-143. <https://doi.org/10.4000/volume.8121>
- 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. <https://doi.org/10.1121/1.4773865>
- Ratnovsky, A., Elad, D. & Halpern, P. (2008). Mechanics of respiratory muscles. *Respiratory Physiology & Neurobiology*, 163(1), 82-89. <https://doi.org/10.1016/j.resp.2008.04.019>
- Saltzman, E. & Kelso, J. A. (1987). Skilled actions : a task-dynamic approach. *Psychological Review*, 94(1), 84-106. <https://doi.org/10.1037/0033-295X.94.1.84>
- Sapthavee, A., Yi, P. & Sims, H. S. (2014). Functional endoscopic analysis of beatbox performers. *Journal of Voice*, 28(3), 328-331. <https://doi.org/10.1016/j.jvoice.2013.11.007>
- Saussure, F. d. (1916). *Cours de linguistique générale*. Éditions Payot & Rivages.
- Shadle, C. H. (2010). The aerodynamics of speech. In W. J. Hardcastle, J. Laver & F. E. Gibbon (Éd.), *The handbook of phonetic sciences* (p. 39-80). Wiley. <https://doi.org/10.1002/9781444317251.ch2>
- Shosted, R. (2006). Just put your lips together and blow ? The whistled fricatives of Southern Bantu. In H. C. Yehia, D. Demolin & R. M. Laboissiere (Éd.), *Proceedings of the 7th International Seminar on Speech Production*.
- Stankiewicz, E. (Éd.). (1972). *A Baudouin de Courtenay Anthology : The Beginnings of Structural Linguistics*. Indiana University Press. <https://muse.jhu.edu/book/84822>
- Stevens, K. N. (1971). Airflow and turbulence noise for fricative and stop consonants : static considerations. *The Journal of the Acoustical Society of America*, 50(4), 1180-1192. <https://doi.org/10.1121/1.1912751>
- Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17(1), 3-45. [https://doi.org/10.1016/S0095-4470\(19\)31520-7](https://doi.org/10.1016/S0095-4470(19)31520-7)
- Stone, M., Epstein, M. A. & Iskarous, K. (2004). Functional segments in tongue movement. *Clinical Linguistics & Phonetics*, 18(6), 507-521. <https://doi.org/10.1080/02699200410003583>
- Stowell, D. & Plumbley, M. D. (2008). *Characteristics of the beatboxing vocal style* (C4DM-TR-08-01). Centre for Digital Music, Dept. of Electronic Engineering, Queen Mary, University of London.
- Traill, A. (1985). *Phonetic and phonological studies of !Xóõ Bushman* (Doctoral Dissertation). Hamburg : Helmut Buske.
- Troubetzkoy, N. S. (1976). *Principes de phonologie*. Klincksieck.
- Tuhuse, B. II. & Traill, A. (1999). !Hán-!Hánsè, the desert Cisticola, implements an acoustic target. In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville & A. C. Bailey (Éd.), *Proceedings of the 14th International Congress of Phonetic Sciences* (p. 1041-1042). <https://doi.org/10.1017/S0025340900000000>

[//www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/p14\\_1041.html](http://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/p14_1041.html)

- TyTe & Splinter. (2014, septembre 18). *Standard beatbox notation (SBN)* [HUMAN BEATBOX]. <https://www.humanbeatbox.com/articles/standard-beatbox-notation-sbn/>
- Uecker, M., Hohage, T., Block, K. T. & Frahm, J. (2008). Image reconstruction by regularized nonlinear inversion—joint estimation of coil sensitivities and image content. *Magnetic Resonance in Medicine*, 60(3), 674-682. <https://doi.org/10.1002/mrm.21691>
- Uecker, M., Zhang, S., Voit, D., Karaus, A., Merboldt, K.-D. & Frahm, J. (2010). Real-time MRI at a resolution of 20 ms. *NMR in Biomedicine*, 23(8), 986-994. <https://doi.org/10.1002/nbm.1585>
- Van Son, R. & Pols, L. C. (1999). An acoustic description of consonant reduction. *Speech Communication*, 28(2), 125-140. [https://doi.org/10.1016/S0167-6393\(99\)00009-6](https://doi.org/10.1016/S0167-6393(99)00009-6)
- van Vugt, F. T., Furuya, S., Vauth, H., Jabusch, H.-C. & Altenmüller, E. (2014). Playing beautifully when you have to be fast : spatial and temporal symmetries of movement patterns in skilled piano performance at different tempi. *Experimental Brain Research*, 232(11), 3555-3567. <https://doi.org/10.1007/s00221-014-4036-4>
- Yeshoda, K. & Raveendran, R. (2021). Exploring the spectral and temporal characteristics of human beatbox sounds : a preliminary study. *Journal of Voice*, S0892199721003556. <https://doi.org/10.1016/j.jvoice.2021.10.011>
- Zipf, G. K. (1949). *Human Behavior And The Principle Of Least Effort*. Addison Wesley Press, Inc.

**Titre :** Étude phonétique de la production du Human Beatbox : approche articulatoire, aérodynamique et acoustique

**Résumé :**

Avant d'être un objet d'étude scientifique, le Human Beatbox (HBB) est avant tout une pratique musicale. Les beatboxeurs et les beatboxeuses utilisent leur conduit vocal pour produire des sons musicaux : sons de batterie, instruments à vent ou à corde, musique électronique ... Comment est produit le Human Beatbox ? Quelles sont les capacités du conduit vocal humain ? Existe-t-il un lien entre la production du Beatbox et de la parole ? Cette thèse tente d'apporter des éléments de discussion sur ces questions. À partir de données articulatoires, aérodynamiques et acoustiques, nous proposons une analyse de la production du Human Beatbox à différentes vitesses (90, 120, 150 battements par minute). Nos résultats montrent que les sujets utilisent des mécanismes de production similaires aux locuteurs des langues. Toutefois, ils combinent un plus grand nombre de mécanismes différents. Le Human Beatbox possède les propriétés d'un système combinatoire discret (Proctor et al., 2013). Les résultats sur les effets du tempo suggèrent que les stratégies individuelles de réorganisation temporelle des gestes sont similaires aux stratégies proposées par Byrd et Tan (1996). En effet, les sujets manipulent la durée des gestes et/ou des intervalles séparant les gestes. On a observé chez 2 sujets, des contraintes de coordination entre les gestes d'initiation et d'articulation. Le Human Beatbox jette un regard différent sur la phonétique en permettant d'actualiser son cadre théorique et de passer d'une approche purement linguistique à une approche anthropophonique (Catford, 1977 ; Lindblom, 1990) des phénomènes phonétiques.

Mots-clés : Human Beatbox, Phonétique, Articulation, Aérodynamique, Acoustique, Tempo

**Title :** A phonetic study of Human Beatboxing production : an articulatory, aerodynamic and acoustic approach

**Abstract :**

Before being a subject of scientific study, Human Beatboxing (HBB) is above all a musical practice. Beatboxers use their vocal tract to produce musical sounds : drum sounds, wind or string instruments, electronic music, and more. How is beatboxing produced? What are the capacities of the human vocal tract? Is there a connection between beatboxing and speech production? This thesis attempts to provide elements for discussion on these questions. Based on articulatory, aerodynamic, and acoustic data, we present an analysis of beatboxing production at different tempo (90, 120, 150 beats per minute). Our results show that subjects use similar production mechanisms to those found in linguistic systems. However, they combine a greater number of different mechanisms. Beatboxing possesses the properties of a discrete combinatorial system (Proctor et al., 2013). The results on tempo effects suggest that individual strategies for temporal reorganization of gestures are similar to the strategies proposed by Byrd et Tan (1996). Indeed, subjects modified the duration of gestures and the intervals between gestures. Coordination constraints between initiation and articulation gestures were observed for two subjects. Beatboxing offers a different perspective on phonetics, allowing for the update of its theoretical framework and a shift from a purely linguistic approach to an anthropophonic approach (Catford, 1977 ; Lindblom, 1990) of phonetic phenomena.

**Keywords :** Human Beatboxing, Phonetics, Articulation, Aerodynamics, Acoustics, Tempo