



HAL
open science

Variations de débit articulatoire sur un corpus conversationnel avec différents types d'interactions

Estelle Chardenon, C. Meunier, Cécile Fougeron

► To cite this version:

Estelle Chardenon, C. Meunier, Cécile Fougeron. Variations de débit articulatoire sur un corpus conversationnel avec différents types d'interactions. 34ème Journées d'Etudes sur la Parole, Jun 2022, Noirmoutier-en-l'Île, France. hal-03944868

HAL Id: hal-03944868

<https://hal.science/hal-03944868v1>

Submitted on 1 Feb 2023

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.

Variations de débit articulatoire sur un corpus conversationnel avec différents types d'interactions

Estelle Chardenon¹, Christine Meunier¹, Cécile Fougeron²

(1) Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France

(2) Laboratoire de Phonétique et Phonologie (CNRS & U.Sorbonne Nouvelle) 19 Rue des Bernardins, 75005 Paris, France

estelle.chardenon@univ-amu.fr, christine.meunier@univ-amu.fr,
cecile.fougeron@sorbonne-nouvelle.fr

RESUME

Des questions actuelles (robotiques, judiciaires, etc.) recourent au besoin de mieux comprendre les caractéristiques de la parole des individus. Les recherches, dans ce domaine se heurtent régulièrement aux questions de variabilité (intra et inter-locuteurs), tout particulièrement en situation de parole spontanée. Cette étude vise à mieux cerner les variations de débit des locuteurs dans un corpus de parole spontanée en prenant en compte différents facteurs pouvant influencer leur production tel que le type d'interaction. 16 locuteurs sont analysés dans un corpus de dialogues conversationnels. Nos résultats montrent que les locuteurs ont en moyenne un débit plus élevé en parole conversationnelle qu'en parole lue avec une variation inter-locuteurs assez forte. On observe malgré tout un important chevauchement des valeurs pour l'ensemble des locuteurs. Une observation plus fine des dialogues ne montre pas d'influence du type d'interaction mais confirme une adaptation du débit à l'interlocuteur.

ABSTRACT

Articulatory rate variations on a speech conversational corpus with different interaction types

Current questions (robotics, forensic, etc.) require a better understanding of individual speech properties. Research in this field is regularly confronted to variability (intra- and inter-speakers) issues, especially in spontaneous speech. This study aims to document rate variations in a spontaneous speech corpus, by taking into account various factors that can influence their production like type of interaction. 16 speakers are analyzed in a corpus of conversational dialogues. Our results show that speakers have an average higher speech rate in conversational speech than in read speech with a fairly strong inter-speaker variation. Nevertheless, we observe a large overlap between speaker's individual rates. A closer observation of a dialogues during the conversations doesn't show variation according to interaction types but confirm an adaptation to the rate of the interlocutor.

MOTS-CLES : Débit articulatoire, variation inter-locuteurs, variation intra-locuteurs.

KEYWORDS: Articulatory rate, inter-speaker variability, intra-speaker variability.

1 Introduction

La fréquence des études visant à comprendre les caractéristiques de parole des individus ne cesse de croître, notamment dans les domaines informatiques, robotiques et judiciaires (Campbell, 2009 ; Ajili, 2018). La plupart des études se concentrent sur des caractéristiques de la voix et la parole telles que la fréquence fondamentale ou les caractéristiques spectrales. Les valeurs de débit de chaque individu permettent également d'apporter des indices idiosyncratiques majeurs (Dellwo, 2008). Chardenon & al. (2020) ont montré que le débit est un paramètre qui varie entre les locuteurs mais qui s'avère robuste dans la variation intra-locuteur. Plusieurs facteurs peuvent provoquer ces variations : par exemple, Bona (2014) montre que les locuteurs âgés ont un débit plus lent que les locuteurs jeunes. Fougeron et al. (2021) observent aussi un ralentissement du débit qui décroît progressivement avec l'âge chez les adultes et que les femmes ont tendance à parler plus vite que les hommes. Quant au style de parole, les locuteurs auraient tendance à parler de façon plus lente lorsqu'ils sont en parole contrôlée comme le montrent Crystal & House (1988).

Un grand nombre d'études rapportant des caractéristiques individuelles de débit sont réalisées sur des corpus de répétition ou de parole contrôlée et lue (Pardo, 2010 ; Bell, 2003 ; Fougeron & al, 2021 ; Chardenon & al, 2020). Toutes obtiennent des résultats significatifs sur les variations inter-locuteurs en parole contrôlée. En revanche, encore peu de chercheurs se sont penchés sur les variations de débit dans un contexte de parole spontanée. Verhoeven (2004) montre, dans un corpus d'interaction en allemand, des différences de débit articulatoire en fonction du sexe, de l'âge mais également de la ville d'origine des locuteurs. En sus du fait que les études sur le spontané se font rares, une grande majorité de ces études observe exclusivement les variations inter-locuteurs.

En comparant des phrases similaires tirées d'une interview et d'une lecture de ces mêmes phrases, Dellwo (2015) montre des différences de complexité dans le choix des enchaînements vocaliques et consonantiques pouvant impacter les mesures rythmiques, principalement en parole spontanée. Il présente des différences inter-locuteurs sur des valeurs rythmiques qu'il assimile à des propriétés kynématiques propres à chaque locuteur. Cependant, il n'observe aucun lien rythmique majeur entre les différents styles de parole sur la variation inter-locuteurs. Alors, même si une majorité d'études concluent à une différence de débit entre les différents styles de parole, dans son étude de 2012, Leemann n'observe à nouveau pas de différence significative lorsqu'un locuteur change de style de parole. Des limites potentielles de ces études seraient respectivement la durée des enregistrements analysés (1 phrase) et le nombre de locuteurs enregistrés (7).

Entrer en interaction avec un individu provoque très souvent des adaptations de certains paramètres vocaux ou de parole. C'est notamment ce qu'observe De Looze (2012) sur l'adaptation de la fréquence fondamentale lors de conversations téléphoniques et Kousidis & al (2008) sur des adaptations d'intensité lors de conversations naturelles entre deux individus. Endlund (2009) observe également une adaptation des temps de pauses sur une conversation d'une trentaine de minutes entre deux individus. L'aspect temporel pourrait donc être influencé par l'interaction comme l'ont montrées les études de Cohen Priva & al. (2013) et Fuscone & al. (2018).

Des questions restent alors en suspens : plusieurs études montrent des variations de débit inter-locuteurs en parole contrôlée mais les conclusions de ces travaux ne vont pas toujours dans le même sens. Aussi, nous proposons d'observer ici le débit articulatoire de 16 locuteurs en interaction pendant une heure. Nous cherchons à caractériser le débit de la parole en conversation, mais aussi à évaluer la variabilité inter-locuteurs dans ces interactions. Nous cherchons également à évaluer d'éventuelles variations de débit en fonction des types d'interaction au sein de l'enregistrement (dialogue/passages

narratifs). Enfin, nous testons une potentielle adaptation du débit articulatoire entre les duos de locuteur sur l'heure d'enregistrement.

2 Méthodologie

2.1 Description du corpus et des locuteurs

Pour cette étude, nous avons utilisé les enregistrements du *Corpus of Interactional Data* (CID, Bertrand, 2008). Ce corpus comporte 8 heures d'enregistrement de dialogues en interaction avec 16 locuteurs (10 femmes et 6 hommes d'un âge moyen de 34 ans : min = 28 et max = 45). Les duos de locuteurs ont été choisis en fonction de leurs affinités (interactions personnelles fréquentes) (Bertrand, 2008) afin de créer un dialogue familier, le plus fluide et naturel possible. Un travail important de transcription et d'annotation automatiques a déjà été accompli sur ce corpus : segmentations en IPU (inter pausal units), en syllabes, en phones (Bertrand, 2008) et plus récemment en passages narratifs (Bertrand, 2017).

Nous utilisons des valeurs de débit obtenues lors d'une étude précédente en lecture de texte (Chardenon & al., 2020.) afin d'évaluer la différence éventuelle avec le débit caractéristique de la parole en interaction analysée dans cette étude. La moyenne de débit en lecture de texte (5 locuteurs) est de 12,5 phonèmes/seconde avec un minimum de 11,63 phonème/seconde et un maximum de 13,25 phonèmes/seconde.

2.2 Traitement des données

Extraction des données

A partir des enregistrements du CID et de la segmentation en IPU (Inter Pausal Unit), nous avons recueilli, à l'aide d'un script praat, le nombre de phonèmes par IPU ainsi que les temps de début et de fin pour chaque IPU. Nous avons également extrait les 'short pauses' (sp) ainsi que le nombre de sp par IPU afin de les exclure dans notre calcul de débit articulatoire. Pour les 16 locuteurs, nous avons en moyenne 59 sp allant de 11 à 152 sp par locuteur. Les IPU pouvant correspondre à des temps de silence aussi bien qu'à des temps de parole, nous avons fait le choix d'exclure tous les passages annotés 'gpF' (temps de silence) et ne conserver que ceux annotés 'gpd' (temps de parole). Enfin, nous avons calculé le débit articulatoire pour chaque IPU ainsi que son écart-type en procédant dans un premier temps par un calcul du temps total pour chaque IPU (temps de fin – temps de début) puis dans un second temps par un calcul de débit articulatoire (nombre de phonèmes/temps total en secondes). Tous les IPU contenant moins de 10 phonèmes, pouvant correspondre à des mots isolés, ont été exclus de notre analyse.

Analyses statistiques

Avec ces données, nous avons calculé un débit articulatoire moyen par locuteur (obtenu à partir du débit de chaque IPU). Nous avons également extrait la médiane par locuteur afin de s'assurer de la répartition homogène des valeurs des IPU par rapport à la moyenne. Nous avons retenu les valeurs extrêmes (maximum et minimum des valeurs de débit articulatoire sur tous les IPU) pour chaque locuteur avec son coefficient de variation, nous permettant d'observer la dispersion des moyennes de

débit articulatoire chez les locuteurs. Pour observer la significativité des différences nous avons utilisé des ANOVA ainsi qu'un test de Tukey nous permettant d'obtenir une comparaison multiple entre les locuteurs.

Analyses sur les passages narratifs

Nous avons ensuite sélectionné un autre type de dialogue pour procéder à des analyses plus détaillées. A partir de la segmentation en passages narratifs de Bertrand (2017), nous avons sélectionné les passages narratifs sur l'heure d'enregistrement que nous avons comparé aux passages non narratifs. Nous avons réutilisé les données de débit articulatoire pour ces différents passages. Nous avons en moyenne 9 passages narratifs par locuteur avec une forte dispersion selon les locuteurs. Pour LL nous avons recensé deux passages narratifs (comprenant 112 IPU) alors que MB présente 17 passages narratifs (comprenant 375 IPU). Le nombre d'IPU analysés sur les passages narratifs est de 46,71% alors que sur les passages non-narratifs il est de 53,29%.

A partir de ces données, nous avons comparés nos valeurs de débit articulatoire sur les passages narratifs pour les 16 locuteurs à l'aide d'une ANOVA et nous avons réalisé une comparaison multiple pour regarder la significativité des différences de débit entre les locuteurs. Enfin, nous avons comparé les valeurs de débit articulatoire de l'enregistrement avec celles des passages narratifs avec un t-test pour observer une potentielle différence entre ces types de dialogues.

Analyses sur les duos de locuteurs

Afin d'observer l'influence mutuelle des locuteurs dans leur duo de conversation, nous avons découpé notre heure d'enregistrement en 4 intervalles temporels : 1 intervalle toutes les 15 minutes. Nous avons ensuite calculé le débit articulatoire moyen (moyenne des débits par IPU) pour chaque intervalle. Enfin, nous avons obtenu la différence entre chaque locuteurs en soustrayant les débits articulatoire moyen des deux locuteurs pour chaque intervalle, ce qui permet d'évaluer une éventuelle convergence ou divergence des valeurs de débit par duo.

3 Résultats

3.1 Comparaison entre les 16 locuteurs

Le débit articulatoire moyen sur tous les locuteurs est de 12,6 phonèmes/seconde (P/S). Comme nous pouvons le voir dans la figure 1 ci-dessous, les 16 locuteurs analysés varient entre 11,15 (LL) et 14,41 (EB) phonèmes/seconde correspondant à une plage de variation maximale de 3,26 phonèmes/seconde. Nous pouvons remarquer dans la table 1 ci-dessous que tous les locuteurs ont des valeurs de débit dans un intervalle compris entre 6,25 phonèmes/seconde (min le plus élevé, ML) et 18,54 phonèmes/seconde (max le moins élevé, LJ). Cela signifie que dans cet intervalle plutôt large se trouvent la plus grande partie des productions de tous les locuteurs. Si les locuteurs présentent des moyennes de débit spécifiques, aucune valeur prise au hasard ne peut être attribuée à un locuteur spécifique.

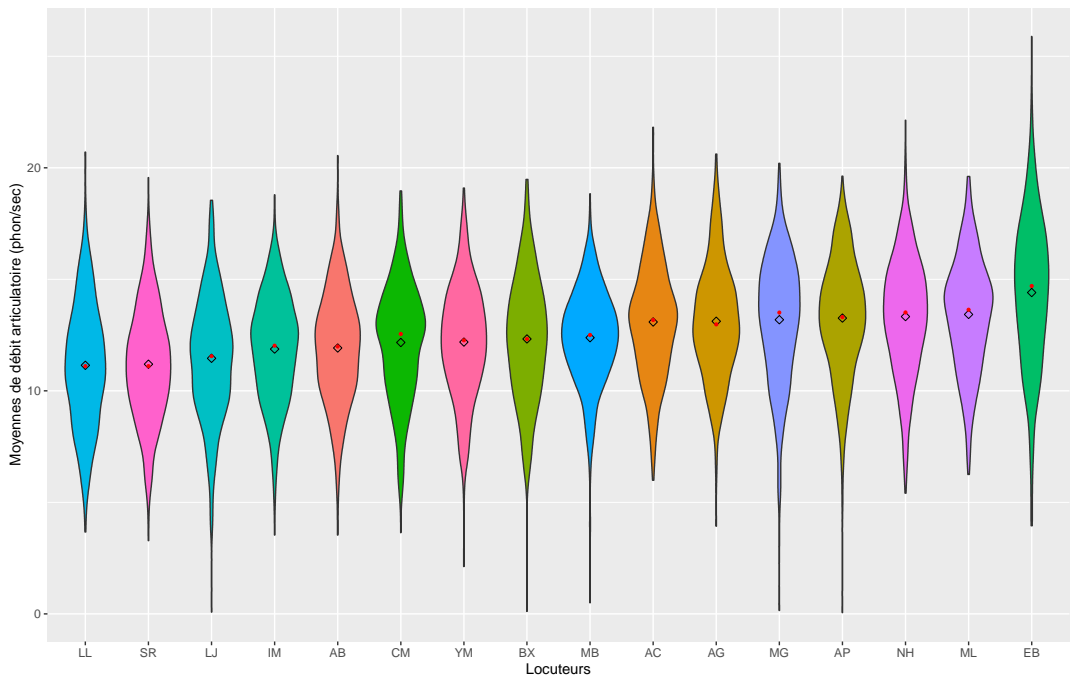


Figure 1. Distribution des valeurs de débit articulaire des IPU par locuteur et débit moyen par locuteur (losanges) et médiane (points rouges) pour les 16 locuteurs du CID.

	LL	SR	LJ	IM	AB	CM	YM	BX	MB	AC	AG	MG	AP	NH	ML	EB	Moy.
Min. débit	3.67	3.28	0.08	3.54	3.54	3.64	2.13	0.11	0.50	5.99	3.93	0.15	0.06	5.41	6.25	3.95	0.06
Max. débit	20.71	19.56	18.54	18.79	20.55	18.97	19.09	19.48	18.84	21.82	20.62	20.20	19.63	22.13	19.61	25.89	25.89
Coef. Var	7.64	6.37	7.36	5.60	6.14	6.56	6.47	7.65	4.68	6.43	6.53	7.90	6.73	6.93	6.24	11.18	7.59

Table 1. Valeurs minimales et maximales et coefficient de variation des débits articulaires par IPU pour les 16 locuteurs.

Concernant le coefficient de variation, nous remarquons qu'un locuteur varie son débit articulaire de façon plus importante : EB. Ce locuteur est également celui qui parle le plus vite. Cependant, les locuteurs qui varient le plus leur débit articulaire entre les IPU ne sont pas nécessairement les locuteurs qui parlent le plus vite ou le plus lentement (cf. Table 1).

On observe un effet significatif du locuteur sur les valeurs de débit ($F = 58,86$, $df = (15, 8072)$, $p < 0.001$). Les comparaisons des locuteurs deux à deux (test posthoc) ne sont pas toutes significativement différentes. Plus de la moitié des duos ont des moyennes significativement différentes. En effet, seuls les locuteurs parlant très vite (EB) ou très lentement (LL et SR) diffèrent significativement des autres. C'est notamment l'exemple des locuteurs LL et SR qui varient avec tous les locuteurs excepté LJ et SR ou LL et LJ qui sont les deux locuteurs ayant les moyennes les plus proches ou encore le locuteur EB qui varie son débit articulaire avec l'ensemble des locuteurs. Nous pouvons également noter que les locuteurs restants varient généralement avec des locuteurs assez éloignés.

3.2 Passages narratifs

Nous observons une moyenne de débit articulaire très légèrement plus rapide sur les passages narratifs. Nous n'observons pas de différence entre les deux types de dialogue (12,68 phonèmes/seconde pour les PN et 12,50 phonème/seconde pour les PNN). Pour autant, nous pouvons remarquer une organisation similaire entre les moyennes de débit articulaire pour les 16 locuteurs. En effet, le locuteur LL est toujours celui qui parle le plus lentement, quel que soit le type de dialogue tout comme le locuteur EB est toujours celui qui parle le plus rapidement (cf. figure 2).

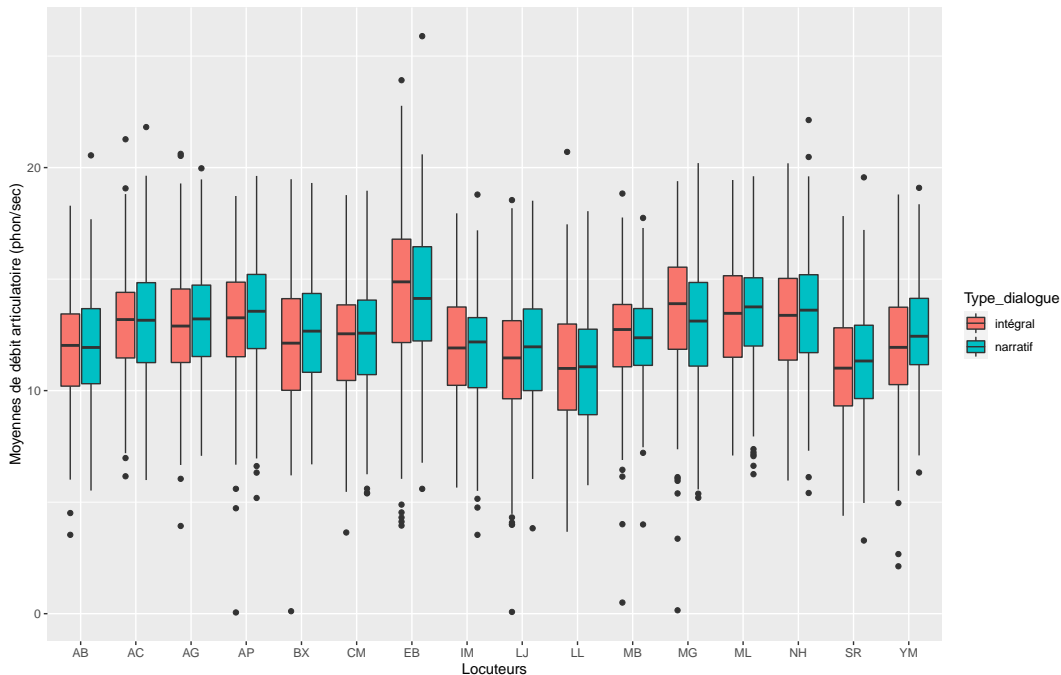


Figure 2. Dispersions des moyennes de débit articulaire des locuteurs du CID sur de l'enregistrement et sur les passages narratifs.

Lorsque nous nous concentrons exclusivement sur les résultats obtenus pour les passages narratifs, nous observons une différence significative des moyennes de débit articulaire ($F=15$, $df(3798)$, $p < 0,001$) entre les 16 locuteurs.

Ces différences significatives entre ces duos de locuteur sont similaires à celles observables dans la partie 3.1.

3.3 Adaptation du débit à l'interlocuteur

Les distances entre les débits des locuteurs d'un duo tendent vers un mouvement de convergence pour la majorité d'entre eux. Sur les 8 duos de locuteurs, 5 montrent une diminution de l'écart de leur débit articulaire de façon importante entre le premier et le dernier groupe (table 2, en rouge). 2 duos restent globalement stables entre le premier et le dernier groupe. Seul un duo ne converge pas entre

le début et la fin de l'enregistrement (duo LL-NH). En regardant plus finement les écarts entre les différents groupes et non pas seulement entre le début et la fin de l'heure d'enregistrement, nous remarquons que les locuteurs ne suivent pas toujours une même tendance mais que les écarts de débit articulaire varient au milieu de l'enregistrement. L'adaptation ne se fait pas toujours de façon linéaire : au milieu de l'enregistrement, les distances entre les membres d'un duo peuvent être élevées (duo AG-YM) ou proches (duo BX-MG). Cette affirmation est observable sur tous les duos sauf pour le duo IM-ML qui est en constante convergence et donc diminution de l'écart entre les moyennes de débit articulaire. Les duos IM-ML et EB-SR ont les écarts les plus importants de débit entre le début et la fin de l'enregistrement. Ces derniers ont également des moyennes de débit articulaire très éloignées (cf. figure 1) puisque le locuteur EB est le plus rapide des locuteurs alors que le locuteur SR fait partie des 3 locuteurs les plus lents. De même pour le locuteur ML qui est le locuteur le plus rapide après EB et qui reste très éloigné du locuteur IM qui est le 4^{ème} locuteur le plus lent.

	AB-CM	AG-YM	LL-NH	EB-SR	BX-MG	AP-LJ	AC-MB	IM-ML
Int1	0.55	0.81	2.09	3.63	0.82	2.38	1.07	2.14
Int2	0.52	1.13	2.10	3.21	1.63	1.29	0.59	1.59
Int3	0.88	0.80	1.99	3.39	0.02	1.50	0.61	1.37
Int4	0.19	0.92	2.64	2.80	0.98	2.14	0.53	1.19

Table 2. Écarts de débit articulaire entre les membres de chaque duo à 4 moments de l'enregistrement (Int1 à 4). En gras sont indiqués les locuteurs qui convergent et en rouge les convergences entre le début et la fin de l'enregistrement chez les locuteurs qui convergent.

4 Discussion et conclusion

Le but de notre étude était de regarder les potentielles variations de débit articulaire dans un corpus de parole interactionnelle en prenant en compte différents facteurs pouvant influencer cette variation.

L'analyse des moyennes de débit par IPU et par locuteur ont montré que le débit est légèrement plus rapide et variable en spontané que celui présenté dans la littérature pour de la parole contrôlée en français : environ 0,1 phonème par seconde de plus par rapport aux normes de Fougeron et al. (2021) et à celles de Chardenon & al (2020). De plus, les débits articulaires moyens varient d'un locuteur à un autre : les locuteurs ont globalement des débits différents. Le débit articulaire en parole conversationnelle chez un même locuteur reste cependant fluctuant : les plages de variation sont larges et se chevauchent. En prenant au hasard une moyenne de débit articulaire par IPU, il est possible de trouver des valeurs qui varient entre 6,5 et 18,54 phonèmes par seconde, ce qui ne permet pas d'attribuer une valeur de débit à un locuteur spécifique. On ne peut donc pas caractériser un locuteur à l'aide de son débit articulaire avec un tel chevauchement. De plus, les dispersions des débits articulaires sont globalement assez similaires même si certains locuteurs (MB vs EB par exemple) ont une variabilité intra-locuteur différente. Nous remarquons cependant que le débit articulaire moyen de chaque locuteur n'est pas directement lié avec les dispersions des valeurs de débit articulaire.

Les résultats sur les passages narratifs ont montré que le type de dialogue dans l'interaction n'exerçait pas d'influence sur la vitesse de parole des individus. En effet, les débits ne varient pas en fonction du style narratif vs interactionnel à l'intérieur d'une conversation. Nous observons une organisation quasi similaire dans les deux situations puisque les locuteurs les plus lents sur de l'enregistrement sont également les plus lents sur les passages narratifs et cette observation est également valable pour les locuteurs les plus rapides. Cette organisation similaire est aussi observée par Leemann (2012) lorsqu'il compare la parole spontanée à de la lecture.

Concernant la variation inter-locuteurs sur les passages narratifs, les variations étant similaires à celles observées dans l'intégralité de l'enregistrement, nous ne faisons pas ressortir de changement dans les caractéristiques de parole lié au type d'interaction. Les schémas étant pratiquement identiques dans les deux analyses, les locuteurs semblent donc être assez stables quel que soit le type de dialogue.

Cohen Parvi & al. (2017) ainsi que Fuscone & al. (2018) ont analysé des différences entre des situations interactionnelles et des productions en monologue et observent des différences entre les deux types de parole. On pourrait imaginer obtenir des différences de type de parole en comparant nos résultats du corpus interactionnel du CID avec des monologues en parole spontanée de ces mêmes locuteurs.

Nos observations concernant l'évolution du débit dans les duos de locuteurs mettent en évidence l'idée que certains locuteurs adaptent leur vitesse d'articulation lorsqu'ils sont en interaction avec un autre locuteur. En effet, nous observons chez plus de la moitié de nos duos de locuteurs une convergence de débit articuloire entre le début et la fin de l'heure d'enregistrement : l'écart entre les locuteurs diminue de façon souvent progressive. Cette adaptation semble bien plus importante lorsque les locuteurs ont des débits articuloires éloignés. Dellwo (2012) obtient des résultats présentant une adaptation rythmique à un autre individu puisqu'il montre une adaptation de la vitesse de parole lorsque des locuteurs parlent en synchronie. Pardo (2017) explique que ce phénomène de convergence serait involontaire et dû à la volonté de rendre une conversation plus fluide entre deux locuteurs avec pour but de renforcer les liens sociaux (Aubanel, 2020). L'observation que nous faisons sur la non-linéarité de la convergence peut être expliquée par l'intérêt des locuteurs pour les sujets de conversations qui varient au cours de l'heure d'enregistrement et qui peuvent avoir un impact sur la variation de débit (Fuscone & al., 2018) ou encore la fatigue vocale pouvant affecter la vitesse de parole et ses variations.

Pour conclure, les locuteurs montrent des valeurs de débit moyen qui leur sont spécifiques. Pour autant, les dispersions individuelles sont considérables. En conséquence, le débit d'un IPU pris au hasard ne peut, en aucun cas, être attribué à un locuteur plutôt qu'à un autre. Concernant la variation intra-locuteur, nous n'observons pas de variation en fonction du 'type' d'interaction dans un corpus conversationnel néanmoins, nous faisons ressortir un facteur de variation important dans les corpus interactionnels : l'influence mutuelle des locuteurs.

Remerciements

Ce travail a été financé dans le cadre d'un projet de thèse financé par Aix-Marseille Université et par l'ANR VoxCrim (ANR-17-CE39-0016). Nous remercions Gabriele CHIGNOLI ainsi que Benjamin O'BRIEN pour leur implication dans la rédaction des scripts Praat.

Références

- AJILI, M., ROSSATO, S., ZHANG, D., & BONASTRE, J. F. (2018, June). Impact of rhythm on forensic voice comparison reliability. In *Odyssey 2018 The Speaker and Language Recognition Workshop*.
- AUBANEL, V., & NGUYEN, N. (2020). Speaking to a common tune: Between-speaker convergence in voice fundamental frequency in a joint speech production task. *PloS one*, 15(5), e0232209. :
- BELL, L., GUSTAFSON, J., and HELDNER, M. (2003). Prosodic adaptation in human-computer interaction, in *ICPhS Barcelona, 2003*, 2453-2456.
- BERTRAND, R., BLACHE, P., ESPESSER, R., FERRE, G., MEUNIER, C., et al. (2008) Le CID - Corpus of Interactional Data - Annotation et Exploitation Multimodale de Parole Conversationnelle. *Revue TAL, ATALA (Association pour le Traitement Automatique des Langues)*, 49 (3), 105- 134. hal-00349893
- BERTRAND, R., ESPESSER, R. (2017). Co-narration in French conversational storytelling: A quantitative insight. *Journal of Pragmatics*, 111, 33-53. <https://doi.org/10.1016/j.pragma.2017.02.001>
- BONA, J. (2014). Temporal characteristics of speech: The effect of age and speech style. *The Journal of the Acoustical Society of America*, 136(2), 116-121.
- CAMPBELL, J. P., SHEN, W., CAMPBELL, W., M.I, SHCWARTZ, R. S, BONASTRE, J.F and MATROUF, D. (2009) Forensic speaker recognition. *IEEE Signal Processing Magazine*, 26 (2) 95-103, doi: 10.1109/MSP.2008.931100.
- CHARDENON, E., FOUGERON, C., AUDIBERT, N., & GENDROT, C. (2020). Dis-moi comment tu varies ton débit, je te dirai qui tu es. In *Actes de la 6e conférence conjointe Journées d'Études sur la Parole* 82-90.
- COHEN PRIVA, U., EDELIST, L., & GLEASON, E. (2017). Converging to the baseline: Corpus evidence for convergence in speech rate to interlocutor's baseline. *The Journal of the Acoustical Society of America*, 141(5), 2989-2996.
- CRYSTAL, T. H., & HOUSE, A. S. (1988). A note on the variability of timing control. *Journal of Speech and Hearing Research*, 31, 497-502. DOI : 10.1044/jshr.3103.497
- DE LOOZE, C., SCHERER, S., VAUGHAN, B., & CAMPBELL, N. (2012). Vers un mesure automatique de l'adaptation prosodique en interaction conversationnelle (Automatic measurement of prosodic accommodation in conversational interaction)[in French]. In *Proceedings of the Joint Conference JEP-TALN-RECITAL 2012, volume 1: JEP* 321-328.
- DELLWO, V., FRIEDRICHS, D. (2012). Variability of speech rhythm in synchronous speech. In: *Ma, Qiuyu; Ding, Hongwei; Hirst, Daniel. Proceedings of the 6th International Conference on Speech Prosody. Shanghai: Tongji University Press*, 539-542.

DELLWO, V., LEEMANN, A., KOLLY, M-J. (2015). Rhythmic variability between speakers: Articulatory, prosodic, and linguistic factors. *Journal of the Acoustical Society of America*, 137(3) 1513-1528.

DELLWO, V., & KOREMAN, J. (2008). How speaker idiosyncratic is measurable speech rhythm. *In Abstract presented at the annual IAFPA meeting.*

EDLUND, J., HIRSCHBERG, J. B., & HELDNER, M. (2009). Pause and gap length in face-to-face interaction.

FOUGERON, C., GUITARD-IVENT, F., & DELVAUX, V. (2021). Multi-dimensional variation in adult speech as a function of age. *Languages*, 6(4), 176.

FUSCONE, S., FAVRE, B., & PREVOT, L. (2018, May). Replicating speech rate convergence experiments on the switchboard corpus. *In Workshop on Replicability and Reproducibility of Research Results in Science and Technology of Language.*

KOUSIDIS, S., D. DORRAN, Y. WANG, B. VAUGHAN, C. CULLEN, D. CAMPBELL, C. McDONNELL and COYLE. E. (2008). Towards measuring continuous acoustic feature convergence in unconstrained spoken dialogues. *Interspeech*, 1692-1695.

LEEMANN, A., DELLWO, V., & KOLLY, M. J. (2012). Rhythmic differences between read and spontaneous speech: the case of% V. *Phonetik & Phonologie 8 Friedrich-Schiller-Universität Jena*, 16.

PARDO, J.S., JAY, I.C. & KRAUSS, R.M. (2010). Conversational role influences speech imitation. *Attention, Perception, & Psychophysics* 72, 2254–2264. <https://doi-org.lama.univ-amu.fr/10.3758/BF03196699>

ROUAS, J. L., FARINAS, J. (2004). Comparaison de méthodes de caractérisation du rythme des langues. *In Workshop MIDL 29*, 30-36.

VERHOEVEN, J., DE PAUW, G., KLOOTS, H. (2004). Speech rate in a pluricentric language: A comparison between Dutch in Belgium and the Netherlands. *Language and Speech* 47, 297–308.