



Étude exploratoire des événements articulatoires pendant la réalisation de pauses en parole spontanée

Ivana Didirková¹, Camille Fauth², Sébastien Le Maguer³

(1) F.R.S.-FNRS & VALIBEL, Université Catholique de Louvain, Belgique

(2) Université de Strasbourg, Institut de Phonétique, E.A. 1339 LiLPa, Strasbourg, France

(3) Saarland University, Allemagne

ivana.didirkova@gmail.com

RÉSUMÉ

Les disfluences en tant qu'événements arrêtant momentanément le déroulé du discours font partie intégrante de la production de la parole. De nombreuses études ont analysé leurs caractéristiques en se fondant principalement sur leurs propriétés acoustiques. L'objectif de cette étude est d'enrichir la description acoustique des disfluences produites par quatre locuteurs en parole spontanée en la corrélant à une description articulatoire, obtenue à partir d'un articulographe électromagnétique (EMA), synchronisée avec les données acoustiques. Pour ce faire, nous avons observé leur production à l'aide de deux différentes méthodes : une étude articulatoire effectuée de façon automatique à partir de la vitesse des mouvements des articulateurs et une observation experte du mouvement de ces mêmes articulateurs. Les résultats de cette étude exploratoire montrent l'existence de plusieurs configurations articulatoires pour un même type perçu de disfluence et inversement, plusieurs catégories acoustiques peuvent parfois être traduites par une même configuration articulatoire. De même, l'anticipation du geste articulatoire relatif au phone subséquent semble dépendre de l'activité articulatoire supra-glottique.

ABSTRACT

An exploratory study of articulatory events during pause realization in spontaneous speech.

Disfluencies, as a momentary stop event in a discourse, are considered as a constitutive part of speech production. Numerous studies analyzed their specificities by relying mainly on their acoustic properties. The goal of the presented study is to extend the acoustic analysis by including an articulatory description. This description is obtained using Electro-Magnetic Articulography (EMA) synchronized with the acoustic data. To achieve this study, we have observed the production of the disfluencies following two complementary methodologies: an automatic articulatory study based on the quantification of the velocity of the articulators and an expert subjective analysis of the movement of these articulators. The results of this exploratory study show the same perceived disfluency category can rely on multiple articulatory configurations. Multiple acoustic categories can also be realized following a same articulatory configuration. Furthermore, the anticipation of the subsequent gesture seems to be dependent on the supraglottic activity.

MOTS-CLES : Production de la parole, pauses, disfluences, description articulatoire, EMA

KEYWORDS: Speech production, pauses, disfluencies, articulatory description, EMA

1 Introduction

La communication ne se fonde pas exclusivement sur les mots employés ou sur les éléments prosodiques retenus par le locuteur. D'autres événements paralinguistiques, plus ou moins intentionnellement produits, participent à la construction du discours. Parmi ces éléments, notre attention se porte ici sur les disfluences (ou accidents de parole) qui viennent interrompre le discours et peuvent donc parfois entraver son intelligibilité (Mac Gregor *et al.*, 2009). Traditionnellement, deux types de disfluences sont observées (Fromkin & Ratner, 1998) : les pauses silencieuses (ou vides), qui sont une interruption du flux de parole se répercutant sur le signal acoustique par une amplitude nulle ou non-significative (Duez, 2003), et les pauses remplies ou pleines qui regroupent les répétitions, les faux-départs, les allongements de sons indépendants de l'allongement final de fin de syntagme (Duez, 2001), la réalisation d'un *schwa* (Maclay & Osgood, 1959 par ex.) ou l'utilisation d'un autre type de filler. De nombreuses études ont cherché à quantifier et à analyser ces disfluences, montrant que celles-ci sont d'autant plus fréquentes que le discours n'est pas préparé à l'avance (Corley & Stewart, 2008). De plus, l'endroit où se réalisent les disfluences ne serait pas anodin dans la mesure où elles ont plus de chance d'apparaître en début de prise de parole ou en début de phrase (Maclay & Osgood, 1959 ; Beattie & Bradbury, 1979) ou lorsque les énoncés à venir sont longs (Oviatt, 1995 ; Shriberg, 1996) ou enfin que le locuteur est peu familier du sujet discuté (Bortfeld *et al.*, 2001 ; Merlo & Mansur, 2004).

Toutefois, la plupart de ces études se fondent sur des données acoustiques et peu de recherches se sont portées sur les événements articulatoires pendant les disfluences. Signalons toutefois l'étude de Ramanarayanan *et al.*, (2009) qui ont observé le tractus vocal pendant la réalisation de pauses silencieuses grâce à l'IRM dynamique et montré que le comportement des articulateurs diffère en fonction de la nature de la pause. Ainsi, la vitesse des articulateurs décroît significativement lorsque la pause est planifiée alors que ce n'est pas le cas pour les pauses agrammaticales. Ces résultats sont à rapprocher de l'étude pilote de Lalain *et al.* (2016) qui a montré que les données articulatoires semblent pouvoir également contribuer à l'identification d'indices robustes pour caractériser les pauses respiratoires (abaissement systématique de la mandibule) et la déglutition (élévation de l'apex, du dos et du larynx).

Ainsi, ces travaux, généralement conduits à partir de données acoustico-perceptives, concluent qu'il existerait deux types de pauses : les pauses vides (qui peuvent-être syntaxiques ou non) et les pauses remplies. Notre *objectif*, avec cette étude, est d'observer si la dichotomie, fondée sur des indices acoustiques et perceptifs, entre pauses vides et pauses pleines correspond à l'activité articulatoire supra-glottique. En d'autres termes, est-ce que des patterns articulatoires spécifiques aux différents types pauses peuvent être observés ? Pour cela, nous nous baserons sur deux critères : présence / absence de mouvements articulatoires et anticipation de mouvement relatif au phone subséquent à la disfluence. Notre *hypothèse* est que, à l'instar des résultats présentés dans Didirková (2016) sur des données portant sur la production de la parole par des sujets qui bégaient, les catégories acoustico-perceptifs ne devraient pas être représentatives de l'activité articulatoire.

2 Méthodologie

2.1 Corpus et participants

Notre étude s'appuie sur des enregistrements articulatoires acquis à l'aide d'un articulographe électromagnétique de type Carstens AG501 3D au LORIA avec une fréquence d'échantillonnage de

250 Hz et une précision constructeur de 0,3 mm. Ces données ont été stockées sous format *.pos* et synchronisées avec des données acoustiques (format *.wav*, 44,1 kHz, 16 bits) obtenues à l'aide d'un microphone t.bone EM 9600. Les enregistrements se sont déroulés dans un endroit calme.

Quatre sujets ont pris part à cette étude, 2 femmes et 2 hommes de langue maternelle française, appariés en catégorie socio-professionnelle et en âge (âge moyen : 31,25 ans, ET : 5,12). Après la fixation des capteurs (1 au milieu de chaque lèvre, 1 sur la mandibule, 1 sur la pointe de la langue, 1 sur le pré-dos et 1 sur le dos de la langue, plus 3 capteurs servant à contrôler les mouvements de la tête dont 1 sur le front et 1 derrière chaque pavillon auriculaire), plusieurs tâches de production leur ont été demandées. Les analyses présentées dans le cadre de cette étude portent sur la parole spontanée (description d'une journée type, des passe-temps avec relance de l'expérimentateur si nécessaire) et ce, dans l'objectif d'obtenir des disfluences liées à la programmation linguistique. Les données ainsi acquises constituent un corpus d'une durée totale de 26 min 02 s, allant de 5 min 5 s à 8 min 14 s selon le sujet. À l'intérieur de ces laps de temps, les tours de parole des sujets (après exclusion des interventions de l'expérimentateur) représentent 21 min et 19 s au total.

2.2 Traitement de données

L'ensemble du corpus a été transcrit orthographiquement puis segmenté et annoté de manière semi-automatique à l'aide du logiciel Praat (Boersma & Weenink, 2017) et de l'outil d'alignement automatique et phonétique EasyAlign (Goldman, 2011). Une ligne d'annotation spécifique a été ajoutée pour y repérer les différents types de disfluences. Comme le remarque Candea (2000), ces éléments peuvent renvoyer, selon les auteurs, à des phénomènes différents. Pour ce travail, nous avons considéré les pauses selon les catégories suivantes : les prolongations correspondant aux allongements audibles d'un son (y compris celles des marqueurs d'hésitation de type « euh »), les répétitions de phonèmes et les pauses silencieuses (vides) syntaxiques ou non-syntaxiques (disfluentes). Afin de faciliter la lecture, toutes ces catégories vont être regroupées sous le terme de « disfluences », même si l'inclusion des pauses syntaxiques dans ce groupe peut faire débat. Ces catégories ont été choisies car, ne concernant que l'absence d'articulation ou un seul phone, elles permettent de minimiser l'influence de la coarticulation sur les analyses articulatoires. Les pauses vides ont été annotées sans seuil minimum.

Les analyses reposent ainsi sur une identification acoustico-perceptive des disfluences susmentionnées par deux auteurs. Les cas de désaccord ont été discutés. Les sujets ont produit 714 disfluences (voir TABLEAU 1 pour la distribution). Les pauses vides contenant des bruits de bouche (au nombre de 78) ont été exclues de la suite des analyses car possiblement provoqués par les capteurs collés dans la cavité buccale. Les analyses qui suivent couvrent donc 636 disfluences.

Locuteur	PV	SD	P	REPph	Total
F1	46	16	89	0	151
F2	33	17	57	0	107
M1	57	25	141	6	229
M2	95	13	40	1	149
Total	231	71	327	7	636

TABLEAU 1: Distribution des disfluences par locuteur et par type de disfluence. F - locuteur féminin, M - locuteur masculin. PV - pauses vides syntaxiques, SD - pauses vides non-syntaxiques, P - prolongations, REPph - répétitions de phonèmes.

Les TextGrids ont par la suite été exportés dans le logiciel Visartico (Ouni *et al.*, 2012) permettant une visualisation des données articulatoires dans différentes configurations, dont une vue en 2D avec coupe midsagittale ou en 3D. Ce logiciel rend également possible l'obtention d'un suivi de l'évolution spatio-temporelle des articulateurs synchronisé avec le signal de parole. Précisons que seules les *tiers* phone, words et disfluences ont été visualisées dans Visartico (FIGURE 1).

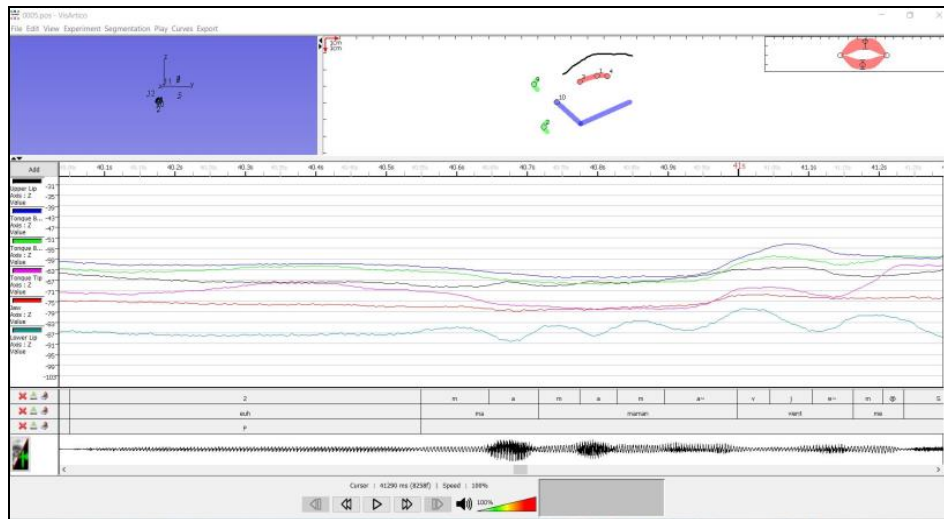


FIGURE 1 : Exemple d'annotation visualisé dans Visartico. En haut à gauche, une vue 3D des mouvements des capteurs. À droite, une vue sagittale médiane ainsi qu'une vue frontale des lèvres. Au centre, l'évolution spatio-temporelle des mouvements articulatoires verticaux. En bleu, les mouvements du dos de la langue ; en vert, les mouvements du prédos ; en noir, la lèvre supérieure ; en rose, l'apex ; en rouge, la mandibule et en gris, la lèvre inférieure.

Après l'importation des données, le logiciel Visartico a servi à la visualisation des mouvements articulatoires verticaux des articulateurs supra-glottiques mentionnés *supra* durant les disfluences. Ces mouvements ont par la suite été décrits puis classés dans différentes catégories. Enfin, l'extension de l'anticipation du mouvement lié au phone suivant immédiatement la disfluente a été calculée. Pour cette dernière mesure, les analyses portent systématiquement sur le mouvement de l'articulateur principal nécessaire à la production du phone suivant.

Signalons que l'analyse des signaux EMA a motivé l'exclusion de 12 occurrences de disfluences dans l'étude portant sur les *patterns* observés et ce, pour cause des anomalies sur le signal. En outre, l'anticipation n'a pas pu être analysée pour 90 disfluences puisqu'elles étaient suivies d'une pause silencieuse ou d'une fin de tour de parole.

2.3 Analyse automatique

Pour déterminer si une trame est en mouvement, nous avons choisi un critère basé sur la vitesse instantanée, détaillé dans Didirková *et al.* (2017). Tout d'abord, nous assimilons la vitesse instantanée à la dérivée première qui est calculée de manière discrète, telle qu'implémentée dans le système HTK¹. Ensuite, une trame est considérée comme étant en mouvement si sa dérivée dépasse un certain seuil. Ce seuil a été arbitrairement défini à 30% de la vitesse moyenne calculée sur l'ensemble des trames associées au locuteur analysé et ce, afin d'encourager la détection de

¹ Voir <http://htk.eng.cam.ac.uk> pour plus de détails.

mouvement. Cette étape effectuée, l'analyse est poursuivie à l'échelle du segment. L'objectif de notre analyse est de vérifier si un mouvement se produit dans un segment supposé statique. Ainsi, afin d'éviter les détections dues aux transitions, les premières trames ainsi que les dernières trames du segment concerné (soit 10% au début et à la fin du segment) ont été écartées.

L'analyse automatique est fondée sur deux mesures. La première consiste à calculer le ratio de trames en mouvement par rapport au nombre total de trames du segment. La seconde se focalise sur l'amorce du premier mouvement significatif. Pour cela, nous avons arbitrairement défini que si 10 trames (soit 50 ms) consécutives étaient en mouvement, alors nous étions en présence d'un mouvement significatif. Ainsi, nous pouvons calculer à quelle proportion du segment le premier mouvement débute.

3 Résultats

3.1 Pourcentage de trames en mouvement en fonction du type perceptif des disfluences

La première analyse automatique porte sur le pourcentage de disfluences considérées comme étant en mouvement (FIGURE 2). Cette première étude montre que le ratio de disfluences caractérisées par la présence d'un mouvement articulatoire supra-glottique atteint les 80% dans tous les types perceptifs des disfluences.

L'analyse a ensuite été affinée en s'intéressant au nombre de trames en mouvement par type de disfluente. Cette analyse plus détaillée est présentée dans la FIGURE 3. Même si globalement une grande majorité de disfluences présentent des mouvements, cela ne signifie pas pour autant que toutes les trames sont en mouvement suivant les critères définis *supra*. En effet, une tendance des silences non-syntaxiques à avoir davantage de trames en mouvement (80% en moyenne) que les pauses vides syntaxiques ou les prolongations peut être observée.

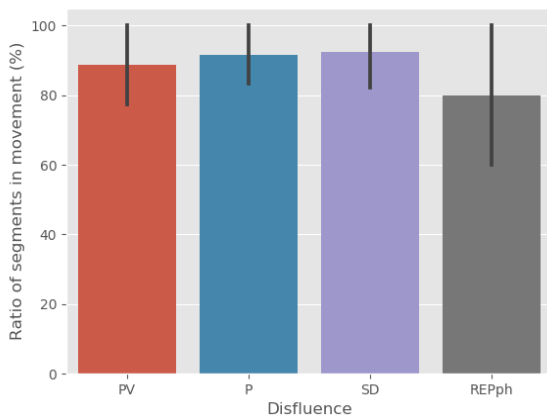


FIGURE 2 : Pourcentage de disfluences considérées comme étant en mouvement.

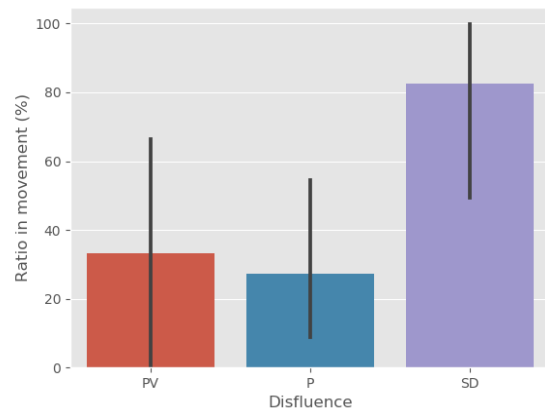


FIGURE 3 : Pourcentage de trames considérées comme étant en mouvement en fonction du type perceptif de la disfluente.

3.2 Patrons articulatoires

L'analyse experte manuelle des mouvements verticaux présents dans les disfluences a permis de faire ressortir l'existence de trois grands patrons articulatoires (test de χ^2 goodness-of-fit, $p < 0,01$). En effet, les disfluences analysées peuvent se traduire par le maintien global de la posture articulatoire avec ou sans sortie acoustique, la présence desdits mouvements tout au long de la disflue, ou une combinaison des deux premiers patrons et ce, dans un ordre pouvant varier.

Cela étant, il est à signaler que les sujets manifestent une nette préférence pour le pattern où des mouvements articulatoires sont présents tout au long de la disflue (89,94% de toutes les disfluences) par rapport à un maintien global de la posture articulatoire (6,16%) et à une combinaison des deux premiers patterns (3,9%). Cette observation confirme par ailleurs celle qui a été faite à partir de la détection automatique de mouvements, présentée *supra*.

Lorsque l'on s'intéresse à la distribution des patrons articulatoires susmentionnés par rapport au type perceptif de la disflue, on observe sans surprise que le patron caractérisé par la présence de mouvements peut être trouvé dans les quatre types analysés (FIGURE 4). En ce qui concerne le maintien global de la posture articulatoire, ce dernier peut principalement être observé dans les prolongations, alors que la combinaison des deux était surtout présente dans les pauses non-syntaxiques. Conformément au test de χ^2 pour l'indépendance, cette corrélation est significative ($p < 0,01$) mais l'effet n'est pas particulièrement fort (0,256).

L'analyse des durées des disfluences permet de montrer également une tendance du pattern « combinaison » à apparaître lorsque la disflue est plus longue (FIGURE 5). Ceci est vrai pour tous les types perceptifs des disfluences. Inversement, lorsque la durée de la disflue diminue, le comportement des types perceptifs des disfluences n'est pas le même. En effet, l'on observe notamment que les prolongations ainsi que les pauses vides non-syntaxiques les plus courtes sont caractérisées par la présence de mouvements articulatoires alors que dans les pauses vides syntaxiques les plus courtes, c'est un maintien global de la posture articulatoire qui est le plus représenté dans notre corpus.

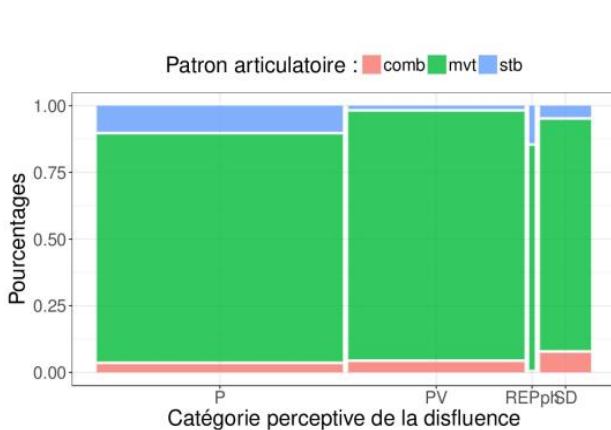


FIGURE 4 : Distribution des patrons articulatoires en fonction du type perceptif de la disflue. Comb = combinaison (en rouge), mvt = mouvements (en vert), stb = maintien global de la posture (en bleu). Sur l'ordonnée, la proportion observée (de 0 à 1). Sur l'abscisse, le type perceptif de la disflue.

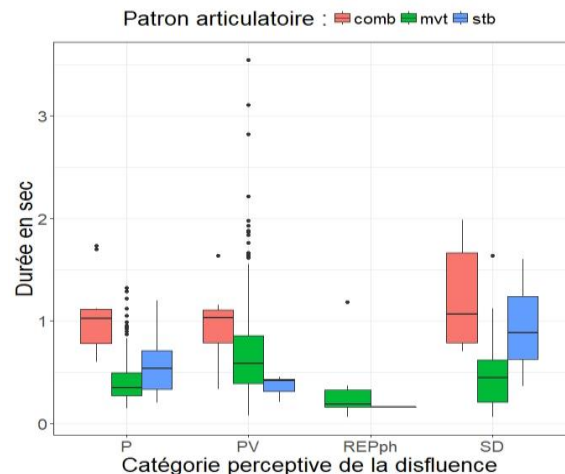


FIGURE 5 : Distribution des patrons articulatoires en fonction de la durée de la disflue. Comb = combinaison (en rouge), mvt = mouvements (en vert), stb = maintien global de la posture (en bleu). Sur l'ordonnée, la durée de la disflue (en secondes). Sur l'abscisse, le type perceptif de la disflue.

3.3 Anticipation dans les disfluences

La dernière analyse présentée dans cette étude concerne l'extension de l'anticipation dans les disfluences. Rappelons que l'objectif de cette analyse était de s'intéresser à la manière dont les phones suivant immédiatement la disfluence sont anticipés du point de vue temporel. Les résultats, présentés dans la FIGURE 6, montrent que l'extension de l'anticipation (en % sur la durée totale de la disfluence) est plus importante lorsque les articulateurs ont été en mouvement tout au long de la disfluence et ce, pour les prolongations et les pauses vides non-syntaxiques, laissant penser à une sorte de continuité dans les mouvements. Pour ces deux catégories perceptives, soulignons également que le maintien global de la position des articulateurs était le patron articulaire qui se prêtait le moins à une anticipation précoce. Dans ces cas-là, le démarrage du geste anticipatoire se fait à moins de 25% de la disfluence. En revanche, dans les pauses vides syntaxiques, la tendance semble inversée dans la mesure où c'est en situation de maintien de posture que l'anticipation vient le plus tôt. Cette différence pourrait s'expliquer par le fait que dans les pauses vides syntaxiques, le degré de préparation du message linguistique qui suit est plus important que dans d'autres types de disfluences. Toutefois, la variation est très importante dans tous les cas de figure. Signalons que 97 disfluences ont été exclues de cette analyse car suivies d'une pause vide.

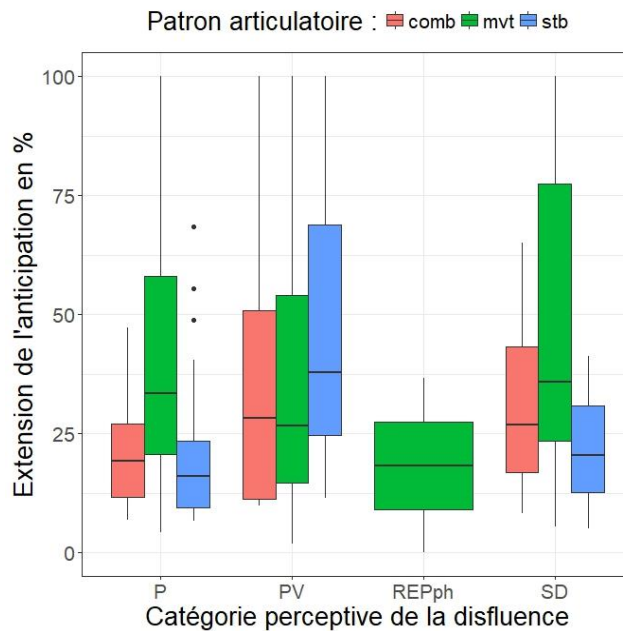


FIGURE 6: Extension de l'anticipation en fonction du type perceptif et du patron articulaire. Comb = combinaison (en rouge), mvt = mouvements (en vert), stb = maintien global (en bleu). Sur l'ordonnée, l'extension de l'anticipation (en %). Sur l'abscisse, le type perceptif de la disfluence.

4 Discussion et conclusion

Rappelons que l'objectif de cette étude était double. Il s'agissait (1) de vérifier si la catégorisation des disfluences en fonction de critères acoustiques et perceptifs correspond à l'activité articulaire supra-glottique et (2) de compléter la description de ces événements de la parole du point de vue de l'extension de l'anticipation du geste liée à la production du phone immédiatement subséquent à la disfluence.

Dans la première partie de l'étude, deux méthodes complémentaires ont été utilisées ; l'une basée sur des critères reposant sur la vitesse et permettant une analyse automatique, l'autre fondée sur une analyse manuelle des courbes de l'évolution spatio-temporelle des articulateurs supra-glottiques. Les deux méthodes ont permis de montrer que les disfluences présentes en parole spontanée se caractérisent, pour la plupart, par la présence d'une activité articulatoire supra-glottique. L'analyse manuelle a montré en outre que cette activité peut parfois être couplée avec une phase de maintien de la posture articulatoire au sein d'une même disfluence. Par ailleurs, un certain nombre de disfluences ne comportaient pas de mouvement détectable automatiquement ou manuellement. De fait, il peut être avancé que la dénomination des disfluences fondée sur les critères acoustiques et perceptifs ne reflète pas suffisamment la réalité articulatoire supra-glottique. À titre d'exemple, les prolongations ne sont pas systématiquement de simples « maintiens » de la posture articulatoire mais se caractérisent par la présence de mouvements articulatoires. De l'autre côté, les pauses ne sont pas exemptes de toute activité articulatoire, comme l'ont déjà souligné Lalain *et al.* (2016). Enfin, si les répétitions de phones n'étaient pas nombreuses dans le corpus, elles ont laissé voir que la répétition n'est pas toujours une réitération du mouvement articulatoire mais que parfois, les articulateurs peuvent être dans une position globalement stable. Dans ce cas, on peut supposer que la répétition sera due davantage à un acte phonatoire qu'articulatoire.

L'étude de l'extension de l'anticipation a également permis d'observer quelques tendances, montrant notamment que le patron articulatoire aurait une influence sur le degré d'anticipation dans la mesure où celui-ci était plus important lorsqu'une activité articulatoire était présente dans la disfluence. Cette tendance n'a cependant pas été confirmée pour les pauses syntaxiques, montrant que la préparation du message linguistique est un facteur à inclure dans les analyses à venir.

5 Perspectives

Si cette étude a permis d'observer certains phénomènes liés à la production des événements arrêtant momentanément le déroulé du discours, il est important de souligner que le nombre de locuteurs doit être augmenté afin de confirmer les tendances observées et éventuellement les compléter par d'autres patterns. En outre, il serait intéressant d'élargir l'analyse à d'autres critères que les mouvements verticaux utilisés pour les analyses manuelles et la vitesse employée pour celles automatiques. Par ailleurs, une analyse portant sur la vitesse avant et après les disfluences apporterait davantage de précisions sur la manière dont les planifications phonologique et articulatoire est faite en phase de disfluence en parole spontanée. L'étude portant sur le nombre de trames en mouvement détecté de manière automatique a aussi fait ressortir l'importance d'une analyse plus fine pour les disfluences considérées comme étant en mouvement. Nous avons notamment pu observer une tendance pour les silences non-syntaxiques à comporter un plus grand nombre de trames en mouvement que les prolongations et les pauses vides syntaxiques. Il semble ainsi intéressant d'affiner le pattern « mouvement » en fonction de plusieurs critères, tels que nombre d'articulateurs en mouvement ou encore la distinction entre l'articulateur principal (qui contribue directement à la production du phone en cours) et les autres. Enfin, les deux méthodes, automatique et manuelle, ayant démontré leur complémentarité, il serait profitable de développer, à terme, une analyse automatique permettant de guider l'analyse experte.

Références

- BEATTIE G. W., BRADBURY R. J. (1979). An experimental investigation of the modifiability of the temporal structure of spontaneous speech. *Journal of Psycholinguistic Research*, 8(3), 225-248.
- BORTFELD H., LEON S. D., BLOOM J. E., SCHOBBER M. F., BRENNAN S. E. (2001). Disfluency rates in conversation: effects of age, relationship, topic, role, and gender. *Language and Speech*, 44(Pt 2), 123-147.
- CORLEY M., STEWART O W. (2008). Hesitation Disfluencies in Spontaneous Speech: The Meaning of um. *Language and Linguistics Compass*, 2(4), 589-602.
- CANDEA M. (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes dits d' "hésitation" en français oral spontané*, Thèse de Doctorat, Paris 3 Sorbonne Nouvelle.
- DIDIRKOVÁ I. (2016). *Parole, langues et disfluences : une étude linguistique et phonétique du bégaiement*, Thèse de Doctorat, Université Paul-Valéry Montpellier 3.
- DIDIRKOVÁ I., LE MAGUER S., GBEDAHOU, D., HIRSCH F. (2017). What happens during stuttering-like disfluencies? An EMA study. *International Seminar on Speech production*, Tianjin, Chine.
- DUEZ D. (2001). Caractéristiques acoustiques et phonétiques des pauses remplies dans la conversation en français. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, 20, 31-48.
- DUEZ D. (2003). Le pouvoir du silence et le silence du pouvoir : comment interpréter le discours politique. *MediaMorphoses*, (8), 77-82.
- FROMKIN V. A., RATNER N. B. (1998). Speech production. In J. B. Gleason & N. B. Ratner (Éd.), *Psycholinguistics*. Harcourt Brace College Publishers.
- LALAIN M., LEGOU T., FAUTH C., HIRSCH F., DIDIRKOVÁ I. (2016). Que disent nos silences ? Apport des données acoustiques, articulatoires et physiologiques pour l'étude des pauses silencieuses. *JEP TALN RECITAL 2016*. Paris, France.
- MAC GREGOR L., CORLEY M., DONALDSON D. I. (2009). Not all disfluencies are equal: The effects of disfluent repetitions on language comprehension. *Brain and Language*, 1(111), 36-45.
- MACLAY H., OSGOOD C. E. (1959). Hesitation Phenomena in Spontaneous English Speech. *WORD*, 15(1), 19-44.
- MERLO S., MANSUR L L. (2004). Descriptive discourse: topic familiarity and disfluencies. *Journal of Communication Disorders*, 37(6), 489-503.
- OUNI, S., MANGEONJEAN L., STEINER I. (2012). VisArtico: a visualization tool for articulatory data, *Interspeech2012*, September 9-13, 2012, Portland, OR, USA.
- OVIATT S. (1995). Predicting and Managing Spoken Disfluencies During Human-Computer Interaction. *Computer Speech & Language*, 9(1), 19-35.
- RAMANARAYANAN V., BRESCH E., BYRD D., GOLDSTEIN L., NARAYANAN S. S. (2009). Analysis of pausing behavior in spontaneous speech using real-time magnetic resonance imaging of articulation. *The Journal of the Acoustical Society of America*, 126(5).
- SHRIBERG E. (1996). Disfluencies In Switchboard. *Proceedings International Conference on Spoken Langage Processing* (Vol. Addendum, p. 11-14). Philadelphie.