

Disfluences en parole continue en français : paramètres prosodiques des pauses pleines et des allongements vocaliques

Yaru Wu^{1,2}, Ivana Didirková³, Anne-Catherine Simon⁴

(1) EA 4255 CRISCO, Université de Caen Normandie, France

(2) UMR 7018 Laboratoire de Phonétique et Phonologie (CNRS Sorbonne-Nouvelle), France

(3) UR 1569 TransCrit, Université Paris 8 Vincennes - Saint-Denis, France

(4) Centre VALIBEL, Institut Langage et Communication,

Université catholique de Louvain, Belgique

yaru.wu@unicaen.fr, ivana.didirkova@univ-paris8.fr,

anne-catherine.simon@uclouvain.be

RÉSUMÉ

Cette étude examine les mesures prosodiques et le débit articulatoire de deux types de disfluences en français continu : les allongements vocaliques et les pauses pleines. Un sous-ensemble du corpus LOCAS-F ($\approx 2,5$ heures) a été utilisé. Les valeurs moyennes de f_0 et les intervalles mélodiques ont été analysés afin de comparer les corrélats prosodiques des disfluences par rapport aux syllabes environnantes fluides. Le débit articulatoire est également évalué. Les résultats montrent que la f_0 moyenne est plus basse dans les pauses pleines et les allongements vocaliques disfluents qu'en parole exempte de disfluences. Les pauses pleines sont produites avec une f_0 plus basse que les allongements vocaliques. Des intervalles mélodiques plus importants sont observés dans les transitions entre l'unité disfluente et la syllabe précédente/suivante de cette unité, par rapport à l'intervalle mélodique entre deux syllabes fluides. En outre, les disfluences liées aux pauses pleines et aux allongements vocaliques se produisent davantage dans les énoncés dont le débit articulatoire est plus bas.

ABSTRACT

Disfluencies in continuous speech in French : Prosodic parameters of filled pauses and lengthening.

This study examines prosodic parameters and speech rate in two disfluency types, namely vowel lengthenings and filled pauses, in continuous French. A subset of the LOCAS-F corpus, containing about 2.5 hours of French spoken data annotated for disfluences, was used in our analyses. Mean f_0 and melodic interval were analysed to compare prosodic correlates of disfluent syllables compared to their surrounding fluent syllables. Articulation rate is also studied. The results show that the average fundamental frequency is lower in filled pauses and disfluent lengthenings than in fluent speech. Filled pauses are produced with a lower f_0 than vowel lengthenings. Larger melodic intervals are observed in transitions between disfluent units and their preceding/following syllables, as compared to melodic intervals between fluent syllables. Disfluencies related to filled pauses and vowel lengthenings are shown to occur more in utterances with lower articulation rate.

MOTS-CLÉS : disfluences, allongements vocaliques, pauses pleines, fréquence fondamentale.

KEYWORDS: disfluencies, vowel lengthenings, filled pauses, pitch.

1 Introduction

L'objectif de cette étude est de décrire les propriétés acoustiques en contexte de deux marques de disfluence en français, les pauses pleines et les allongements vocaliques, d'une manière globale et selon les styles de parole ayant des degrés de préparation différents. Les disfluences sont des séquences d'une ou plusieurs syllabes qui interrompent le flux régulier de parole et devraient être effacées pour parvenir à la formulation visée par le locuteur (Shriberg, 1994). Si elles sont souvent présentées comme caractéristiques d'une parole spontanée (Moniz *et al.*, 2007), elles sont aussi régulièrement produites dans des styles formels ou préparés (Simon *et al.*, 2010). Les disfluences sont catégorisées comme pauses pleines (*euh* en français), allongements vocaliques, faux départs, répétitions ou productions verbales indiquant une correction. Les pauses silencieuses, lorsqu'elles sont précédées d'une marque du travail de formulation, sont aussi considérées comme disfluentes (Candea, 2000, 178). Les disfluences sont souvent interprétées comme le signal d'une difficulté dans le processus de production de la parole et leur taux baisse avec la possibilité de planifier ou de répéter son discours ; en contrepartie, les disfluences peuvent indiquer le caractère non prédictible du discours en train d'être produit (Goldman-Eisler, 1968, 42) et être exploitées perceptivement comme l'annonce d'un contenu nouveau ou complexe (Watanabe *et al.*, 2008; Grosman, 2018).

Dans cette étude, nous décrivons les corrélats prosodiques de deux marques de disfluence souvent traitées conjointement ou confondues (Duez, 1991, 71) : l'allongement vocalique disfluent et les pauses pleines. L'allongement vocalique disfluent se définit comme un allongement anormal de syllabe en position d'amorce ou de fin de mot et présentant un contour intonatif plat ou légèrement descendant (Candea, 2000, 26). Il affecte principalement les syllabes ouvertes (Grosjean & Deschamps, 1975) et sa durée varie entre 180 et 200 ms (Candea, 2000). L'allongement d'hésitation est complexe à identifier et à annoter, car il peut se confondre avec un allongement normal (prédictible) d'une syllabe qui se trouverait en fin de groupe accentuel ou intonatif (Di Cristo, 2000). Les pauses pleines, parfois appelées pauses remplies ou pauses sonores, correspondent à une marque explicite d'hésitation quasi-lexicalisée et généralement transcrite *euh*. Cette syllabe peut s'insérer en épenthèse en finale de mot ou se prononcer de manière indépendante, avant ou après un mot (Candea, 2000, 24).

Des études précédentes ont montré, pour tous types de disfluences confondus, une tendance pour la fréquence fondamentale à diminuer très localement sur la syllabe ou région disfluente (-0,9 DT en portugais, (Moniz *et al.*, 2007)) et à augmenter entre la zone de disfluence et la zone qui suit immédiatement (+0,8 DT). Sur un corpus de huit langues, Vasilescu *et al.* (Vasilescu *et al.*, 2004) ont montré que les pauses pleines (ou voyelles autonomes de support d'hésitation) présentent majoritairement une durée supérieure à 200 ms et qui culmine à 650 ms. Concernant la fréquence fondamentale, les autrices ne relèvent pas de différence significative entre les voyelles d'hésitation et les voyelles fluentes (schwa et eu fermé). Les voyelles d'hésitation présentent par contre souvent une qualité vocale irrégulière (*vocal fry*). En français, la valeur de fréquence fondamentale des pauses pleine est très similaire, pour un locuteur donné, à la valeur de l'onset de ses groupes de souffle (Duez, 2001, 3). Grosman (2018) étudie séparément les pauses pleines et les allongements vocaliques dans un large corpus en français et note que la durée moyenne des allongements vocaliques (347.25 ms) est significativement supérieure à la durée moyenne des pauses pleines (268.4 ms). La distribution des durées des allongements vocaliques est aussi plus étroite que celle des pauses pleines (Grosman, 2018, 191). Nous n'avons pas trouvé de résultats présentant la relation entre les syllabes disfluentes et le débit d'articulation environnant.

2 Méthodologie

2.1 Corpus

Nos analyses portent sur le corpus LOCAS-F (Louvain Corpus of Annotated Speech - French) (Martin *et al.*, 2014), un corpus oral multi-genres annoté pour les disfluences, les unités syntaxiques et les unités prosodiques. Le corpus contient 42 échantillons sonores représentant 14 genres de discours (ou activités communicatives) différents. Dans cette étude, les échantillons contenant des narrations conversationnelles ont dû être écartés en raison de nombreux chevauchements qui auraient pu fausser les relevés de f0. Nos analyses portent ainsi sur 2h38min35sec d'enregistrements (durée moyenne des enregistrements : 4min53).

2.2 Annotation manuelle des données

Les disfluences ont été annotées à deux niveaux selon un protocole adapté aux langues verbales et aux langues signées (Crible *et al.*, 2016) : disfluences simples, alignés sur une tokenisation de la transcription orthographique, et disfluences composées, regroupant plusieurs tokens contigus ou non contigus (Grosman, 2018, 429). Les éléments annotés sont nommés fluencèmes afin de rendre compte de leur contribution potentielle à la disfluence ou à la fluence du discours. Il s'agit des pauses silencieuses, des pauses pleines (euh), des allongements vocaliques, des marqueurs de discours (et, mais...), des termes d'édition (je veux dire), des faux départs, des troncations, des répétitions à l'identique ou avec modification, des substitutions et des insertions. L'annotation des pauses pleines et des allongements d'hésitation a été réalisée manuellement par une experte et évaluée sur une partie du corpus par le calcul d'un taux d'accord à partir des annotations des trois experts (Grosman, 2018, 123). Pour les pauses pleines, le score Kappa s'élève à 0,86 ($Z = 85,55$) et pour les allongements d'hésitation à 0,64 ($Z = 63,11$). Ces mesures sont considérées comme très satisfaisantes.

2.3 Sélection des données

Les analyses portent sur trois catégories cibles, dont deux liées à la disfluence (i.e. allongement vocalique (« AL ») et pause pleine « PP ») et qui sont comparées à de la parole exempte de toute disfluence (« Fluent »). Les allongements vocaliques finaux ne sont pas considérés comme disfluents en français et ne sont donc pas inclus dans les analyses. Les intervalles mélodiques sont mesurés dans Praat (Boersma, 2011) avec un script ad-hoc (1) en soustrayant la fréquence fondamentale de la dernière syllabe du mot précédent de la fréquence fondamentale de la séquence fluente, de la pause remplie ou de la voyelle allongée (= séquence en question (i.e. Fluent/AL/PP) – dernière syllabe du mot précédent) et (2) en soustrayant la fréquence fondamentale de la séquence analysée de la fréquence fondamentale de la première syllabe du mot suivant (= première syllabe du mot suivant – séquence en question (i.e. Fluent/AL/PP)). Le débit articulatoire, quant à lui, est mesuré pour chaque énoncé entre les pauses (unité interpausales) en syll/sec et exclut donc les pauses. Les données sont également regroupées en fonction du degré de préparation comme suit : (1) parole non préparée : comprend des interviews sociolinguistiques, télévisées et radio, ainsi que des récits conversationnels, (2) parole semi-préparée (performance radiophonique, débat politique, conversation formelle, conférence scientifique) et (3) parole préparée comprenant des journaux parlés, des discours politiques et des lectures.

3 Analyses et résultats

Dans cette section, nous présenterons dans un premier temps les comptages sur les allongements vocaliques, les pauses pleines, et les séquences fluides (section 3.1). La fréquence fondamentale moyenne et l'intervalle mélodique seront présentés dans les sections 3.2 et 3.3. Le débit articulatoire sera présenté en dernier dans la section 3.4. Précisons également que la fréquence fondamentale mesurée en Hz est convertie en demi-tons (DT) relatifs à 50 Hz.

Des modèles mixtes linéaires ont été utilisés pour analyser les mesures prosodiques (c'est-à-dire la fréquence fondamentale moyenne et l'intervalle mélodique) et le débit articulatoire en fonction de notre variable cible « F_DF » (comprenant trois niveaux : 1) fluent ; 2) allongement vocalique ; 3) pause pleine) en utilisant le packages *lme4* dans R (R Development Core Team, 2019). Quatre modèles distincts ont été réalisés étant donné que quatre jeux de données différents sont concernés. Pour tous les modèles, « F_DF » est inclus comme effet fixe. « Degré de préparation » et « Genre du locuteur » sont inclus comme variables de contrôle. En ce qui concerne les effets aléatoires, des intercepts ont été inclus pour le sujet et l'item pour les trois premiers modèles sur (1) la f0 moyenne (2) l'intervalle mélodique entre « F_DF » et la dernière syllabe du mot précédent (3) l'intervalle mélodique entre la première syllabe du mot suivant et « F_DF » . Dans le dernier modèle sur le débit articulatoire, l'intercept a été inclus pour le sujet. Pour les analyses statistiques, le débit articulatoire a été normalisé en fonction des sessions d'enregistrement.

3.1 Allongements vocaliques et pauses pleines dans le corpus LOCAS-F

Afin d'avoir une vue globale de la façon dont les données sont présentées en termes de disflue, le nombre (Occ.) et le pourcentage (PCT) de séquences fluides (« Fluent »), d'allongements vocaliques (« AL ») et de pauses pleines (« PP ») sont présentés en fonction de degrés de préparation, pour les hommes (à gauche) et les femmes (à droite) dans le tableau 1.

Catégories	Homme						Femme						Total	
	Non Préparé		Semi-préparé		Préparé		Non Préparé		Semi-préparé		Préparé			
	Occ.	PCT	Occ.	PCT	Occ.	PCT	Occ.	PCT	Occ.	PCT	Occ.	PCT	Occ.	PCT
Fluent	10160	94,7	9766	95,0	16843	98,1	12850	93,7	5951	95,2	3644	99,836	59214	95,8
AL	384	3,6	303	2,9	229	1,3	454	3,3	178	2,9	1	0,027	1549	2,5
PP	184	1,7	214	2,1	101	0,6	411	3,0	119	1,9	5	0,137	1034	1,7

TABLE 1 – Occurrence et pourcentage de séquences fluides (« Fluent »), d'allongement vocaliques (« AL »), de pauses pleines (« PP ») en fonction de degrés de préparation, pour les hommes (à gauche) et les femmes (à droite). « Occ. » correspond aux occurrences et « PCT » correspond aux pourcentages (%). Les PCT pour la parole préparée des femmes sont conservés avec 3 décimales (au lieu de 1) de sorte que le pourcentage total soit égal à 100 % précisément.

Le tableau 1 montre que, globalement, environ 96% des données concernent des séquences fluides, et que la disflue est observée dans un peu plus de 4% des données. Les disfluences sont plus nombreuses dans la parole non préparée que dans la parole préparée, avec la parole semi-préparée entre les deux, tant pour les hommes que pour les femmes. De manière générale, il y a légèrement plus de pauses pleines que d'allongements vocaliques dans nos données. Il est intéressant de noter que très peu de disfluences (< 0,2%) sont observées pour la parole préparée chez les femmes.

3.2 Fréquence fondamentale moyenne

Lorsque l’on s’intéresse aux valeurs moyennes de la fréquence fondamentale en séquences fluides en comparaison avec les séquences disfluides, les résultats montrent que la fréquence fondamentale moyenne est plus basse en allongement et en pause pleine, c’est-à-dire en séquence disfluente (moyenne = 18,04 DT, écart-type (ensuite ET) = 6,70 DT) qu’en parole hors disfluences (moyenne = 19,01 DT, ET = 5,97 DT). Cette observation se confirme pour les deux genres, avec les valeurs moyennes de 22,30 DT (ET = 4,96 DT) en disfluente et 23,66 DT (ET = 4,30 DT) en parole fluente pour les femmes et 15,42 DT (ET = 6,28 DT) en disfluente et 16,72 DT (ET = 5,32 DT) en parole fluente pour les hommes, ce qui représente systématiquement une baisse supérieure à 1 DT en séquence disfluente (Figure 1).

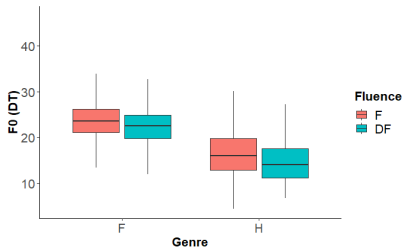


FIGURE 1 – F0 moyenne en séquences fluides vs disfluides et en fonction du genre. F = femmes, H = hommes

Si on décompose les disfluences en allongements vocaliques et en pauses pleines (cf. Figure 2), nous observons la même tendance avec une baisse de f0 plus prononcée dans les pauses pleines (moyenne = 17,90 DT, ET = 7,31 DT) que dans les allongements vocaliques (moyenne = 18,10 DT, ET = 6,43 DT) et que dans les séquences fluides (moyenne = 19,01 DT, ET = 5,97 DT). Cette différence est observée aussi bien chez les femmes avec une f0 moyenne de 21,54 DT (ET = 5,26 DT) dans les pauses pleines et de 22,71 DT (ET = 4,73 DT) dans les allongements vocaliques, que chez les hommes dont la f0 moyenne est de 14,97 DT (ET = 7,41 DT) dans les pauses pleines et de 15,59 DT (ET = 5,81 DT) dans les allongements vocaliques. De fait, la baisse de la f0 dans les pauses pleines atteint 1,75 DT chez les hommes et 2,12 DT chez les femmes par rapport à la parole fluente. Pendant les allongements, cette baisse est plus contenue, aux alentours de 1 DT (0,95 DT chez les femmes et 1,14 DT chez les hommes).

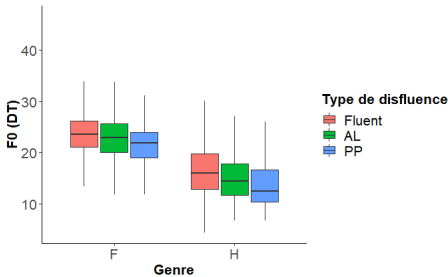


FIGURE 2 – F0 moyenne en séquences fluides vs disfluides et en fonction du genre. Séquences fluides en rouge (« Fluent »), allongement vocalique (« AL ») en vert et pause pleine (« PP ») en bleu. F = femmes, H = hommes.

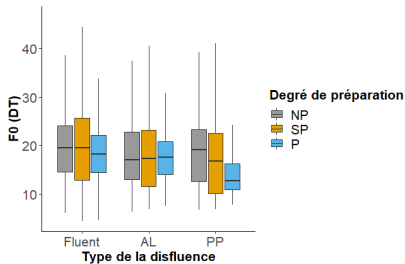


FIGURE 3 – F0 moyenne en fonction du degré de préparation et du type de la disfluente. En gris, prises de parole non préparées (NP), en bleu semi-préparées (SP) et en jaune préparées (P).

Enfin, en étudiant la f0 moyenne en fonction du degré de préparation, il apparaît que le comportement des pauses pleines en parole préparée est différent de celui des autres degrés de préparation et des allongements vocaliques, en ce que la f0 moyenne des pauses pleines baisse avec le degré de préparation : plus la parole est préparée, et plus la f0 moyenne de ces disfluences est basse. En effet, en parole non préparée, leur f0 moyenne est de 18,89 DT (ET = 7,11 DT). Elle est de 17,41 DT (ET = 7,91 DT) en parole semi-préparée et de 14,24 DT (ET = 5,01 DT) en parole préparée, soit une baisse de 4,66 DT entre la parole non préparée et la parole préparée. Il est à noter que les allongements vocaliques ne suivent pas cette tendance : avec des valeurs de respectivement 18,21 DT (ET = 6,42 DT) et 18,14 DT (ET = 5,60 DT), la différence de f0 moyenne entre la parole non-préparée et préparée n'est que de 0,07 DT dans ces disfluences. Enfin, en parole fluente, une baisse de 1,2 DT est observée entre la parole non préparée (moyenne = 19,68 DT, ET = 6,23 DT) et préparée (moyenne = 18,48 DT, ET = 5,29 DT), comme le montre la Figure 3. Les résultats du LMM confirment que la f0 moyenne d'« AL » [$\beta = -0,48307$; $t = -20,831$; $SE = 0,02319$] et de « PP » [$\beta = -1,56979$; $t = -17,372$; $SE = 0,09036$] est significativement inférieure à celle observée pour « Fluent ». Le test post-hoc basé sur le modèle montrent des différences significatives entre (i.e. « Fluent », « AL » et « PP », $p < 0,001$). Aucune différence significative n'est trouvée pour la valeur moyenne de f0 selon degré de préparation.

3.3 Intervalle mélodique

Les résultats sur les intervalles mélodiques sont présentés pour les changements mélodiques entre (1) la dernière syllabe du mot précédent et la séquence en question (i.e. la séquence fluente (« Fluent »), l'allongement vocalique (« AL ») ou la pause pleine (« PP »)) d'une part, et (2) la séquence en question (« Fluent », « AL » ou « PP ») et la première syllabe du mot suivant d'autre part (voir les détails sur les calculs dans la section 2.3). Les disfluences ne sont pas affichées dans la figure relative à la parole préparée chez les femmes, étant donné qu'il y a peu d'occurrences concernées.

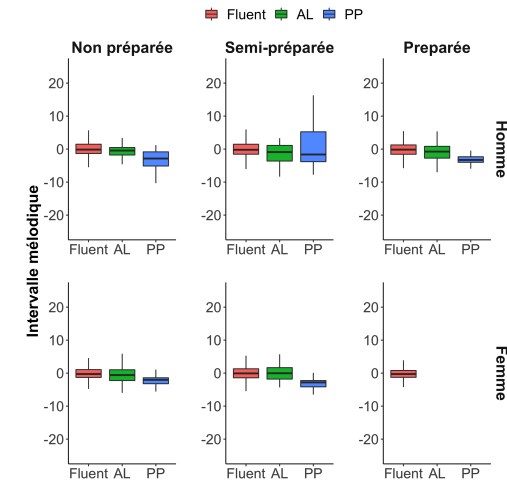


FIGURE 4 – Intervalle mélodique (en DT) entre la dernière syllabe du mot précédent et la séquence fluente (« Fluent », en rouge), l'allongement vocalique (« AL », en vert) ou la pause pleine (« PP », en bleu), en fonction du degré de préparation et du genre (de gauche à droite : parole non préparée, semi-préparée et préparée ; partie en haut pour les hommes, partie en bas pour les femmes).

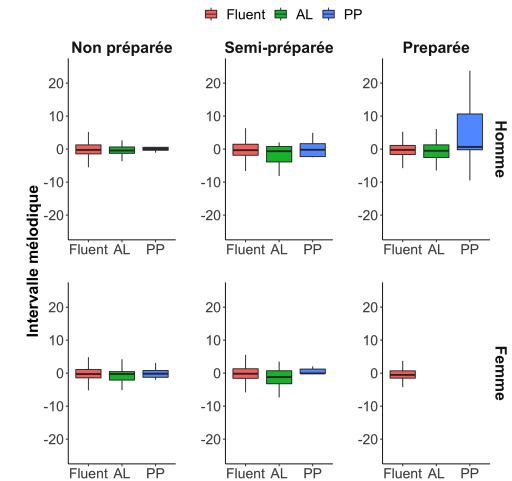


FIGURE 5 – Intervalle mélodique (en DT) entre la séquence fluente (« Fluent », en rouge), l'allongement vocalique (« AL », en vert) ou la pause pleine (« PP », en bleu) et la première syllabe du mot suivant, en fonction du degré de préparation et du genre (de gauche à droite : parole non préparée, semi-préparée et préparée ; partie en haut pour les hommes, partie en bas pour les femmes).

La Figure 4 montre que, pour l'intervalle mélodique entre la dernière syllabe du mot précédent et l'allongement vocalique (« AL »), la pause pleine (« PP ») ou la séquence fluente (« Fluent »), les voyelles allongées et les pauses pleines provoquent un changement mélodique négatif plus important que les séquences fluides de manière générale, à l'exception de la différence visuellement indétectable entre « Fluent » et « AL » pour la parole semi-préparée chez les femmes. Les résultats basés sur le modèle LMM correspondant confirment que « AL » [$\beta = -0,42520$; $t = -2,520$; $SE = 0,16871$] et « PP » [$\beta = -2,59138$; $t = -6,376$; $SE = 0,40645$] déclenchent davantage de changements mélodiques négatifs que « Fluent ». Le test post-hoc basé sur le modèle LMM montre que « Fluent », « AL » et « PP » sont significativement différents les uns des autres ($p < 0,001$ pour toutes les comparaisons).

La Figure 5 illustre l'intervalle mélodique entre la séquence fluente (« Fluent »), l'allongement vocalique (« AL ») ou la pause pleine (« PP ») et la première syllabe du mot suivant. « Fluent », « AL » et « PP » montrent un intervalle mélodique similaire dans toutes les figures, à l'exception de l'intervalle mélodique positif élevé observé pour « PP » dans la parole préparée chez les hommes. Aucune différence significative n'est trouvée pour l'intervalle mélodique entre la séquence en question (la séquence fluente, l'allongement vocalique ou la pause pleine) et la première syllabe du mot suivant.

3.4 Débit articulaire

Le débit articulaire des trois catégories cibles (de gauche à droite respectivement : (1) séquence fluente (« Fluent »), (2) allongement vocalique (« AL »), (3) pauses pleines (« PP ») est présenté en fonction de degrés de préparation chez les hommes et les femmes dans la figure 6. Les disfluences ne sont pas affichées dans la figure relative à la parole préparée chez les femmes, étant donné qu'il y a peu d'occurrences concernées.

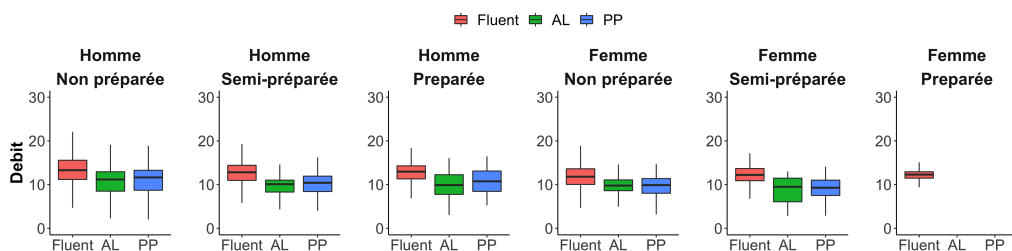


FIGURE 6 – Débit articulaire dans les séquences fluides par rapport aux séquences non fluides en fonction du degré de préparation et du genre (séquences fluides en rouge (« Fluent »), allongement vocalique (« AL ») en vert et pause pleine (« PP ») en bleu ; de gauche à droite : parole non préparée, semi-préparée et préparée ; partie en haut pour les hommes, partie en bas pour les femmes).

Un débit articulaire plus faible est observé pour les voyelles allongées (« AL ») et les pauses pleines (« PP »), par rapport aux « Fluent ». Cela vaut pour tous les degrés de préparation étudiés, chez les hommes et chez les femmes. Les résultats du LMM confirment que l'« AL » [$\beta = -0,952413$; $t = -28,401$; $SE = 0,033535$] et le « PP » [$\beta = -0,957557$; $t = -24,880$; $SE = 0,038487$] se produisent à un débit articulaire significativement plus lent que le « Fluent ». Les tests post-hoc basés sur le modèle LMM correspondant montrent une différence significative entre « Fluent » et « AL » ($p < 0,001$), ainsi qu'entre « Fluent » et « PP » ($p < 0,001$). Aucune différence significative n'est trouvée entre « AL » et « PP ».

4 Discussion, limites et perspectives

L'objectif de cette étude était de décrire le comportement mélodique des allongements et des pauses pleines en comparaison avec les séquences fluides (i.e., exemptes de toute disfluente) en français et ce, dans des styles de parole ayant des degrés de préparation différents. En effet, nous supposons que le fait d'avoir pu préparer son discours pourrait avoir une influence sur le comportement des disfluences. Plus spécifiquement, nous avons étudié la fréquence fondamentale moyenne en séquences disfluentes et fluides en fonction du genre du locuteur et du degré de préparation lié à la prise de parole, l'intervalle mélodique qui sépare les séquences disfluentes des séquences fluides, ainsi que le débit articulatoire dans ces deux séquences. Nos analyses ont porté sur plus de 1500 allongements et plus de 1000 pauses pleines, ainsi que sur plus de 59000 syllabes fluides. La disfluente concerne ainsi un peu plus de 4% des données, représentant une proportion assez faible dans la parole non pathologique.

Nos résultats ont montré une diminution de la fréquence fondamentale en séquence disfluente, rejoignant ainsi les résultats obtenus pour le portugais (Moniz *et al.*, 2007) et pour l'anglais américain (O'Shaughnessy, 1992). Il existe cependant une différence entre les allongements vocaliques et les pauses pleines, les premiers cités se détachant moins en termes de fréquence fondamentale. Cette différence, également observée en hongrois (Deme & Markó, 2013) pourrait s'expliquer par le caractère même de la disfluente, les pauses pleines formant des items autonomes. Par ailleurs, les pauses pleines se caractérisent également par une baisse de la f_0 plus importante lorsque la parole est préparée, ce qui serait en accord avec les nombreuses études laissant penser à une pluralité de catégories de pauses pleines (voir par ex. Kosmala & Crible, 2021). Un autre paramètre à prendre en compte serait également une habitude à travailler avec sa voix, notamment pour les journalistes des journaux parlés qui représentent un tiers du corpus de parole préparée.

Ce résultat est également appuyé par celui obtenu sur l'intervalle mélodique, où les pauses pleines et les allongements vocaliques se démarquent de la syllabe qui les précède par une baisse de la f_0 . De manière intéressante néanmoins, les pauses pleines font également baisser la fréquence fondamentale qui les suit, ce qui n'est pas le cas des allongements vocaliques où la modification de la f_0 n'est pas significative.

Enfin, un débit articulatoire plus faible est observé pour les pauses pleines (« PP ») et les allongements vocaliques (« AL »), comparé à celui relevé pour « Fluent », aussi bien chez les hommes que chez les femmes dans les trois degrés de préparation. Ces résultats suggèrent que, de manière générale, la disfluente est observée davantage dans les énoncés dont le débit articulatoire est plus bas. Ceci est cohérent avec nos attentes liées à la planification de la parole et à la production de la parole résultante. Les disfluences ont tendance à se produire lorsque la parole n'est pas entièrement planifiée. Dans ces situations, nous cherchons à parler plus lentement et à gagner du temps pour réfléchir aux énoncés que nous voulons produire. La production réalisée dans ces situations est donc accompagnée de disfluences dans certains cas.

Afin d'approfondir nos analyses, il serait intéressant de prendre en compte d'autres paramètres, notamment la localisation de la disfluente, et ce, aussi bien en termes grammaticaux (pause pleine placée entre deux unités syntaxiques distinctes ou à l'intérieur d'une unité) qu'en termes discursifs, tout en prenant en compte l'environnement prosodique, ce qui permettrait de mieux expliquer leur comportement spécifique par rapport aux allongements vocaliques. Par ailleurs, d'autres types de disfluences (notamment les répétitions et les faux départs) devraient être étudiés afin d'affiner nos conclusions.

Références

- BOERSMA P. (2011). Praat : doing phonetics by computer [computer program]. <http://www.praat.org/>.
- CANDEA M. (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes "d'hésitation" en français oral spontané. Etude sur un corpus de récits en classe de français*. PhD thesis, Université Paris III-La Sorbonne Nouvelle.
- CRIBLE L., DUMONT A., GROSMAN I. & NOTARRIGO (2016). *Annotation manual of fluency and disfluency markers in multilingual, multimodal, native and learner corpora. Version 2.0*. Rapport interne, Université catholique de Louvain.
- DEME A. & MARKÓ A. (2013). Lengthenings and filled pauses in hungarian adults' and children's speech. *Proceedings of Disfluency in Spontaneous Speech, DISS 2013*, **54**, 21.
- DI CRISTO A. (2000). Vers Une Modélisation De L'accentuation Du Français (seconde Partie). *Journal of French Language Studies*, **10**(01), 27–44.
- DUEZ D. (1991). *La pause dans la parole de l'homme politique*. Sons et Parole. Paris, editions du cnrs edition.
- DUEZ D. (2001). Acoustic-phonetic characteristics of filled pauses in spontaneous French speech : preliminary results. In *Proceedings of DiSS'01, Disfluency in Spontaneous Speech Workshop*, p. 41–44, Edinburg, UK : ISCA.
- GOLDMAN-EISLER F. (1968). *Psycholinguistics : experiments in spontaneous speech*. London, New York, academic press edition.
- GROSJEAN F. & DESCHAMPS A. (1975). Analyse contrastive des variables temporelles de l'anglais et du français : vitesse de parole et variables composantes : phénomènes d'hésitation. *Phonetica*, (31), 144–184.
- GROSMAN I. (2018). *Évaluation contextuelle de la (dis)fluence en production et perception : pratiques communicatives et formes prosodico-syntaxiques en français*. PhD thesis, UCL - Université Catholique de Louvain.
- KOSMALA L. & CRIBLE L. (2021). The dual status of filled pauses : Evidence from genre, proficiency and co-occurrence. *Language and Speech*.
- MARTIN L., DEGAND L. & SIMON A. C. (2014). LOCAS-F : un corpus oral multigenres annoté. In *CMLF 2014 - 4 ème Congrès Mondial de Linguistique Française*, p. 2613–2626, Berlin : EDP Sciences.
- MONIZ H., MATA A. I. & VIANA C. (2007). On filled-pauses and prolongations in european portuguese. In *Proceedings of Interspeech 2007*, p. 2645–2648, Anvers, Belgique : ISCA.
- O'SHAUGHNESSY D. (1992). Recognition of hesitations in spontaneous speech. In *[Proceedings] ICASSP-92 : 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 1, p. 521–524 vol.1.
- R DEVELOPMENT CORE TEAM (2019). *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- SHRIBERG E. E. (1994). *Preliminaries to a Theory of Speech Disfluencies*. PhD thesis, University of California, Berkeley.
- SIMON A.-C., AUCLIN A., AVANZI M. & GOLDMAN J.-P. (2010). Les phonostyles : une description prosodique des styles de parole en français. **2**, 71–88.

VASILESCU I., CANDEA M. & ADDA-DECKER M. (2004). Hésitations autonomes dans 8 langues : une étude acoustique et perceptive. In *Actes du colloque MIDL*, p. 25–30, Paris.

WATANABE M., HIROSE K., DEN Y. & MINEMATSU N. (2008). Filled pauses as cues to the complexity of upcoming phrases for native and non-native listeners. *Speech Communication*, **50**(2), 81–94.