

Iulia Grosman

Université catholique de Louvain & Institut Langage et Communication | Valibel – Discours et Variation

Anne Catherine Simon

Université catholique de Louvain & Institut Langage et Communication | Valibel – Discours et Variation

Liesbeth Degand

Université catholique de Louvain & Institut Langage et Communication | Valibel – Discours et Variation

Variation de la durée des pauses silencieuses : impact de la syntaxe, du style de parole et des disfluences

1. INTRODUCTION ¹

Les pauses silencieuses se définissent comme une interruption de la phonation. Elles ont fait l'objet d'études sur corpus visant à rendre compte de la fréquence des pauses ² et de leur durée en comparant les langues, les styles de parole ou les différentes localisations des pauses dans la chaîne parlée. Ces études débouchent généralement sur une catégorisation des pauses selon des critères formels (pauses brèves, moyennes ou longues) ou fonctionnels (pauses de respiration, grammaticales, d'hésitation, stylistiques, etc.). Les études de perception rendent compte des facteurs qui augmentent la saillance perceptive des pauses et mesurent si les pauses contribuent à identifier des événements prosodiques particuliers comme les proéminences ou les frontières.

À partir d'un corpus permettant de comparer 14 situations de parole en français (LOCAS-F ; Martin, Degand & Simon 2014), nous analysons la distribution et la durée de 4 969 pauses silencieuses relativement à la syntaxe et aux disfluences.

1. Cette recherche est effectuée dans cadre de l'Action de Recherche Concertée *A Multi-Modal Approach to Fluency and Disfluency Markers* avec le soutien de la fédération Wallonie-Bruxelles (convention n° 12/17-044). Anne Catherine Simon y a collaboré dans le cadre du Mandat d'Impulsion Scientifique PhraDiCo (*Prosodic Phrasing for Discourse Comprehension*) financé par le FSR-FNRS (convention n° F4520.16).

2. Nous utilisons « pauses » et « pauses silencieuses » comme des synonymes. Les pauses pleines (*filled pauses*), qui peuvent se combiner avec les pauses silencieuses, sont désignées par la forme *euh*.

Le premier objectif vise à établir quelles mesures sont les mieux appropriées pour calculer la durée des pauses et s'il est possible d'établir empiriquement des seuils de durée pour distinguer différentes catégories de pauses (brèves, moyennes, longues) (§ 5). Ensuite, nous évaluons si ces seuils sont généraux ou spécifiques à la situation de communication ou au locuteur (§ 6).

D'autre part, nous modélisons la durée des pauses en fonction de deux critères : leur cooccurrence avec des marques liées au travail de formulation (*euh*, allongements, répétitions, marqueurs de discours, etc.) (§ 7) et leur localisation syntaxique (§ 8). Des profils de pauses sont analysés en tenant compte des caractéristiques de la situation de parole (degré de préparation, degré d'orientation médiatique, etc.) (§ 9). En conclusion, nous revenons sur la distinction généralement admise entre pauses structurantes (à fonction de segmentation) et non structurantes (à fonction d'hésitation), afin de voir si celle-ci résiste à une analyse multifactorielle (§ 10).

2. DURÉE DES PAUSES

Cette section propose un état de l'art sur des questions méthodologiques relatives à la durée des pauses : les types de mesures (§ 2.1) et les seuils de durée (§ 2.2).

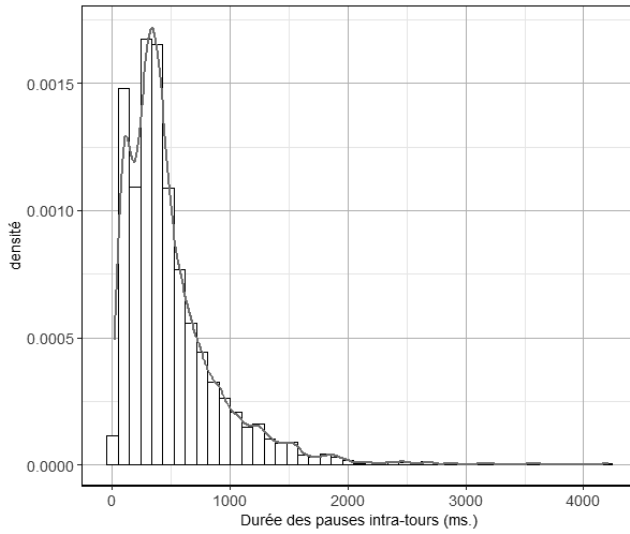
2.1. Types de mesure de durée

L'analyse de la durée des pauses est compliquée par le caractère typiquement asymétrique de leur distribution. La Figure 1 illustre la distribution des durées des 4 969 pauses internes aux tours de parole dans le corpus LOCAS-F³ (cf. § 5.1). Il s'agit d'une distribution non normale, où la moyenne de la durée des pauses ne peut pas servir de point de départ valide à une catégorisation des pauses (Oehmen, Kirsner & Fay 2010).

Certains chercheurs contournent cet écueil en appliquant une transformation logarithmique des données⁴ ou en appliquant des seuils minimum et maximum de durées (cf. § 2.2). Les études divergent cependant pour affirmer que la distribution des durées des pauses, même transformées selon une échelle logarithmique, forme une courbe normale. Si les données de N. H. De Jong et H. R. Bosker (2013 : 19) respectent les conditions de lognormalité, la majorité des distributions observées dans la littérature s'analysent plutôt comme une combinaison de deux, voire de trois, courbes normales (Rosen, 2005 : 411-412 ; Little

3. Cette figure rappelle celle publiée par Campione & Véronis (2002 : 200) à partir des durées de quelque 6 000 pauses extraites du corpus *Eurom*.

4. La transformation de la durée brute d'une pause peut s'opérer en relativisant celle des syllabes environnantes. Cette approche a été appliquée par Lacheret *et al.* (2013) pour la détection des syllabes proéminentes, et par Christodoulides *et al.* (2017) pour représenter la durée des pauses.



**Figure 1 : Distribution des durées pausales extraites de LOCAS-F
(durée absolue en ms)**

et al., 2013 : 135). Par conséquent, chaque courbe normale permettrait de distinguer une catégorie de pauses. Sur la base de l'analyse d'un corpus de parole en anglais, K. Kirsner, J. Dunn et K. Hird (2005 : 7-9) observent une distribution bimodale des durées pausales : ils considèrent que la première composante est formée de « pauses courtes », qui seraient liées aux processus d'articulation, tandis que les « pauses longues » seraient associées à des processus cognitifs de plus haut niveau, comme la segmentation discursive. Ces observations rejoignent celles de M. Demol, W. Verhelst et P. Verhoeve (2007). Ils soutiennent également que la durée logarithmique des pauses peut être approchée par deux distributions normales, que le débit de parole soit rapide ou lent, et dans toutes les langues étudiées (néerlandais, anglais, français, italien, roumain et espagnol).

D'autres auteurs analysent leurs données comme une combinaison de trois distributions. Dans un corpus de 5 h 30 min de parole lue et de parole spontanée dans cinq langues, E. Campione et J. Véronis (2002) distinguent les pauses brèves (inférieures à 200 ms), les pauses moyennes (de 200 ms à 1 000 ms) et les pauses longues (supérieures à 1 000 ms). Ces dernières ne sont toutefois observées que dans la parole spontanée. J.-P. Goldman *et al.* (2010 : 164) avancent que l'aspect unimodal, bimodal ou trimodal des distributions des durées pausales dépend des situations de parole observées. Sur la base d'un corpus en français de 40 minutes rassemblant quatre situations communicatives (lecture,

narration conversationnelle, journal parlé télévisé et conférence scientifique ⁵⁾, J.-P. Goldman *et al.* (*op. cit.*) observent que la distribution des durées pausales est majoritairement bimodale, à l'exception du journal parlé qui est le seul style de parole à faire apparaître des pauses brèves, moyennes et longues.

Ces travaux invalident les analyses basées sur une mesure de tendance centrale (p. ex. la durée moyenne des pauses) et suggèrent d'analyser de manière spécifique chaque catégorie de pauses (courtes, longues et parfois moyennes). Selon nous, cette catégorisation statistique des pauses à partir d'une transformation logarithmique doit également être motivée linguistiquement et pragmatiquement. Or, plusieurs objections surgissent.

Une première limite est inhérente à la méthode même de la linguistique de corpus : cette approche ne peut mesurer si une pause définie statistiquement comme brève ou longue sera perçue comme telle par un auditeur, en contexte. Quand on sait l'importance d'autres facteurs comme le débit, les modifications phonétiques ou les frontières d'unités syntaxiques sur la perception des pauses (Duez 1991 ; Simon & Christodoulides 2016), on ne peut que préconiser de combiner les études de production avec les études de perception.

En outre, l'observation d'une, de deux ou de trois catégories de pauses selon la situation de communication (lecture *vs* spontané, conférence *vs* journal parlé, etc.) pourrait être le résultat d'un artefact gommant la variabilité interindividuelle. Ainsi, chaque nouvelle partition d'un corpus en sous-corpus (selon la situation, selon le locuteur, etc.) pourrait produire une nouvelle distribution (cf. § 6). La détection de différentes composantes dans une distribution est donc à manipuler avec précaution.

Si l'on s'accorde sur le fait qu'il est motivé de distinguer différentes catégories de pauses, qu'en est-il de l'utilisation de seuils de durée ?

2.2. Utilisation de seuils de durée

Faut-il exclure certaines pauses parce qu'elles sont trop brèves ou parce qu'elles sont trop longues ? F. Grosjean et A. Deschamps (1975) considèrent une pause comme silencieuse à partir de 250 ms ; D. Duez (1991) établit un seuil adapté au débit du locuteur à partir de la durée moyenne des occlusives intervocaliques, celui-ci varie entre 180 et 250 ms ⁶ ; M. Candea (2000) retient les pauses à partir de 200 ms sans limite supérieure ; A. Lacheret-Dujour et B. Victorri (2002) utilisent un seuil minimum de 300 ms pour annoter les pauses dans un corpus radiophonique ; etc. La valeur du seuil minimal varie de 180 ms à 300 ms selon les études et E. Campione et J. Véronis (2002 : 202) montrent que la modification du seuil minimal de durée porte à conséquence sur les analyses subséquentes.

5. Cette étude se base sur une partie du corpus C-PROM (Avanzi *et al.* 2010). Les données conversationnelles, de journal télévisé et de conférences scientifiques sont reprises dans LOCAS-F.

6. Voir aussi de De Jong & Bosker (2013 : 18-19).

Ainsi montrent-ils que la durée moyenne des pauses est plus brève en parole spontanée qu'en parole lue si l'on n'utilise aucun seuil, que la proportion s'équilibre avec un seuil minimal de 200 ms et un seuil maximal de 2 000 ms, et que la durée moyenne des pauses devient plus brève en parole lue qu'en parole spontanée si l'on ne conserve que le seuil maximal de 2 000 ms. En d'autres mots, l'utilisation de seuils de durée modifie sensiblement un jeu de données et complexifie les comparaisons avec d'autres études (Hieke, Kowal & O'Connell, 1983 : 202).

En outre, ces seuils ne rendent pas nécessairement compte de la manière dont les pauses sont perçues. D. Duez, qui établit un seuil relatif au débit de chaque locuteur, a ainsi montré que la perception d'une pause peut même subvenir en l'absence de toute interruption de la phonation (cf. la notion de « pause subjective » dans Duez 1993). B. Zellner (1994 : 43) le confirme en observant que le seuil de perception d'une pause est modulé en fonction des segments la précédant. L'adoption d'un seuil minimum exprimé sous la forme d'une durée absolue, bien que séduisante, ne semble donc pas totalement motivée (cf. § 5.3 pour la solution adoptée dans cette étude).

3. LOCALISATION ET FONCTION DES PAUSES

La section précédente a permis de passer en revue les problèmes liés à une analyse de la fréquence des pauses en fonction de leur durée. Cette approche peut se combiner avec une analyse de la localisation séquentielle des pauses, *i.e.* de leur position relativement aux unités linguistiques de différents niveaux. F. Grosjean et A. Deschamps (1975) remettent en cause la possibilité d'assigner une fonction à une pause en fonction de sa localisation contextuelle :

Tout silence de plus de 0,25 sec [...] est considéré comme une pause silencieuse, quel que soit son rôle dans l'activité langagière, à savoir pause d'hésitation, pause de respiration, pause grammaticale (dans laquelle nous incluons les pauses utilisées pour l'emphase et celles permettant de lever certaines ambiguïtés). En effet, ces types de pauses se confondent très souvent ; un sujet profitera, par exemple, d'un arrêt grammatical pour respirer et pour hésiter, ou encore, se trouvant obligé d'hésiter à l'intérieur d'une phrase, en profitera pour respirer, etc. Essayer de déterminer rigoureusement la nature exacte de la pause silencieuse nous semble extrêmement délicat [...]. (Grosjean & Deschamps, 1975 : 148)

Malgré cette mise en garde, il nous a semblé pertinent de tester si la fréquence et la durée des pauses sont corrélées à leur position dans la chaîne parlée. Une distinction intéressante a été établie par M. Candea (2000) entre pauses structurantes et pauses non structurantes. Une pause structurante désigne « toute pause silencieuse qui est encadrée par des séquences sonores appartenant au même locuteur et qui n'est pas immédiatement précédée par une marque de TdF [Travail de Formulation] » (*ibid.* : 145). Lorsqu'une pause est immédiatement précédée d'une marque de TdF (*euh*, allongement d'hésitation, répétition, faux-départ, autocorrection immédiate), elle est considérée comme non structurante

(*ibid.* : 178). Nous tenons compte de cette distinction pour présenter l'état de l'art sur la variation de la durée des pauses selon leur localisation syntaxique (§ 3.1) et selon leur cooccurrence avec des marques d'hésitation (§ 3.2).

3.1. Fréquence et durée des pauses aux frontières syntaxiques

Plusieurs études montrent que la fréquence et la durée des pauses sont positivement corrélées avec le niveau de la frontière syntaxique où elles apparaissent, établissant une fonction segmentante et hiérarchisante de la pause. F. Grosjean et A. Deschamps (1975 : 167) montrent que 60,46 % des pauses se distribuent en fin de phrase, contre 39,54 % à l'intérieur d'une phrase (où elles se situent majoritairement dans le syntagme verbal ou entre le syntagme verbal et un autre constituant). Une différence de durée existe mais n'est pas significative (la médiane de la durée des pauses en fin de phrase s'élève à 500 ms contre 400 ms à l'intérieur des phrases). M. Candea (2000 : 154) rapporte que 77,66 % des pauses structurantes sont positionnées en fin d'énoncé ou de proposition, contre 19,05 % en fin de constituant, et seulement 3,30 % à l'intérieur d'un constituant. La durée moyenne des pauses est significativement plus longue en fin d'énoncé (881 ms) ou de proposition (608 ms) qu'en fin de constituant (486 ms)⁷.

Ces tendances sont confirmées par J.-P. Goldman *et al.* (2010 : 163) qui remarquent que la variable syntaxique⁸ la mieux corrélée avec la durée d'une pause est la position en fin d'unité de rection – l'unité syntaxique maximale dans leur modèle, qui équivaut globalement à la fin d'énoncé ou de proposition chez M. Candea (2000). Utilisant une procédure expérimentale de production contrôlée, J. Krivokapic (2006) met également en évidence, outre le niveau de frontière syntaxique, l'impact de la longueur des constituants qui précèdent et suivent la pause sur la longueur de celle-ci.

3.2. Fréquence et durée des pauses cooccurrentes à des marques d'hésitation

Selon M. Candea (2000), les pauses non structurantes, qui suivent immédiatement une marque de TdF, ont « un fonctionnement très différent des pauses structurantes » (*ibid.* : 146). Premièrement, elles n'occasionnent presque jamais une alternance de tour de parole (*ibid.* : 147) et elles « ne sont pas perçues par les auditeurs comme des indices de coupe » (*ibid.* : 148). Deuxièmement, ces pauses non structurantes ne semblent pas perçues par les auditeurs, que ce soit dans l'étude de perception mise en place par M. Candea (2000, chap. 3) ou dans une

7. Cette tendance est également observée en anglais (Goldman-Eisler 1972 ; Grosjean & Deschamps 1975).

8. Toutes variables confondues, c'est la présence d'une prise de souffle qui constitue le meilleur prédicteur qu'une pause sera longue (Goldman *et al.*, 2010 : 163).

étude similaire de D. Duez (Duez 1995, citée par Candea, 2000 : 178). Par conséquent, il semble légitime de considérer les pauses non structurantes comme une catégorie distincte ⁹.

Globalement, les pauses non structurantes (ou d'hésitation) sont moins fréquentes que les pauses structurantes (ou syntaxiques). Dans l'étude de M. Candea (2000 : 179), elles constituent 17,5 % du nombre total de pauses. Il est difficile de comparer ces chiffres car peu d'études ont distingué ces deux catégories de pauses, soit parce que les auteurs estiment qu'une telle distinction était difficile à appliquer (v. la citation de Grosjean & Deschamps *supra*), soit parce qu'une telle distinction requiert un corpus complètement annoté.

Concernant la durée, M. Candea (2000 : 161, 180) rapporte que la durée moyenne des pauses non structurantes (566 ms) est significativement plus brève que celle des pauses structurantes (677 ms). La pause qui suit une marque d'hésitation serait un prolongement de celle-ci et l'ensemble « marque d'hésitation + pause non structurante » serait perçu comme une seule unité.

Nos données permettent de tester quels facteurs rendent le mieux compte de la fréquence et de la durée des pauses.

4. HYPOTHÈSES ET OBJECTIFS

L'état de l'art (§ 2-3) a permis de montrer que de nombreuses études empiriques existent sur la fréquence et la durée des pauses en français. Cependant, les études basées sur les corpus les plus grands (Campioni & Véronis 2002 ; Christodoulides *et al.* 2017) ne peuvent pas traiter la question de la localisation des pauses parce que les annotations font défaut ; tandis que les corpus richement annotés sont de taille plus réduite et illustrent souvent une seule situation de parole (interview radiophonique chez Grosjean & Deschamps (1975), parole en classe chez Candea (2000), discours politique chez Duez (1991)) ou un nombre réduit de celles-ci (parole lue *vs* spontanée chez Campioni & Véronis (2002), quatre situations de parole chez Goldman *et al.* (2010)).

Le corpus à la base de cette étude (cf. § 5.1) contient des échantillons représentant 14 situations de parole et a fait l'objet d'une annotation des unités syntaxiques et des disfluences. Il permet donc de tester une série d'hypothèses générales :

- les pauses aux frontières syntaxiques sont plus nombreuses que les pauses dans l'environnement de marques d'hésitation ;
- la durée des pauses aux frontières syntaxiques est supérieure à la durée des pauses environnant les marques d'hésitation ;

9. Cela ne s'applique pas aux pauses qui sont suivies d'une marque de TdF (p. ex. « pause silencieuse + *eah* »), qui se comportent comme des pauses structurantes (Candea 2000).

- la durée d’une pause augmente proportionnellement au degré de frontière syntaxique où elle est localisée.

En outre :

- ces tendances pourront être nuancées ou précisées en fonction des situations de parole ;
- on cherchera à voir quels facteurs expliquent la variation de la durée des pauses environnant les disfluences ;
- finalement, on proposera une modélisation de la durée pausale intégrant les critères situationnels, syntaxiques et liés aux disfluences.

5. CORPUS D’ÉTUDE ET ANNOTATIONS

5.1. Corpus

LOCAS-F, acronyme de *LOuvain Corpus of Annotated Speech-French* (Martin, Degand & Simon 2014) compte 48 échantillons sonores produits entre 2005 et 2015, représentant 14 pratiques communicatives différentes : conférence scientifique, débat politique, discours académique, discours politique, homélie, interaction formelle (asymétrique), interaction informelle (symétrique), interview radiophonique, interview sociolinguistique, interview télévisée, journal parlé radiophonique, lecture ¹⁰, performance radiophonique à micro-ouvert ¹¹ et récit conversationnel en face-à-face.

D’une durée de 3 h 38 min, le corpus comprend environ 42 000 tokens, produits par 16 locuteurs en situation de monologue, 41 locuteurs en situation de dialogue et 19 locuteurs en situation de multilogue. Les locuteurs sont des francophones natifs, parmi lesquels 37 belges, 35 français et 5 suisses. Enfin, le corpus est relativement équilibré au niveau du sexe : 43 % des locuteurs sont féminins et 57 % masculins.

5.2. Annotations

Des études antérieures sur LOCAS-F (Degand & Simon 2005 ; Martin, Degand & Simon 2014 ; Simon & Degand 2011) ont assuré le développement de plusieurs annotations mobilisées dans la présente recherche. L’ensemble de ces annotations est disponible sous la forme de fichiers Textgrid (logiciel PRAAT, Boersma & Weenink 2015). Nous décrivons les annotations utilisées dans cette étude.

- Le corpus est transcrit phonétiquement et aligné au niveau du phonème, de la syllabe et du token (*via* le plug-in EasyAlign ; Goldman 2011).
- L’annotation en unités syntaxiques est linéaire et manuelle. Elle se fonde sur l’approche micro- et macrosyntaxique de l’école aixoise (Blanche-Benveniste

10. Il s’agit d’écrivains qui lisent un extrait de leur roman à la radio.

11. Il s’agit de journalistes improvisant une forme de chronique pour la chaîne Arte Radio.

et al. 1990 ; le protocole d'annotation complet est décrit par Tanguy *et al.* 2012)¹². Au niveau maximal, les échantillons ont été segmentés en unités de rection (UR) composées d'un verbe (le plus souvent) accompagné de ses dépendants ; les marqueurs de discours ou les éléments associés, qui peuvent entretenir une relation macrosyntaxique avec les constructions verbales, sont également annotés à ce niveau. Chaque unité de rection se décompose en séquences fonctionnelles (séquence sujet, séquence verbe, séquence objet, séquence régie, etc.) selon la proposition de M. Bilger et E. Campione (2002). La granularité de notre annotation syntaxique n'est pas aussi fine que dans certaines études mentionnées dans l'état de l'art (§ 3.2), car nous nous limitons à deux niveaux d'unités syntaxiques. Cela est partiellement compensé par l'annotation automatique en groupes accentuables (Mertens, 2008 : 100-102), définis par la combinaison d'un mot lexical et des mots grammaticaux qui en dépendent. Les groupes accentuables (ci-après « ga ») constituent dans notre annotation les constituants syntaxiques minimaux.

- L'annotation manuelle des disfluences suit le protocole développé dans le cadre d'un projet de recherche sur la fluence et la disfluence dans différentes langues et modalités (Crible *et al.* 2016)¹³. Les phénomènes suivants sont annotés : pauses, *eah*, allongement à valeur d'hésitation, faux-départs, répétitions, substitutions, troncations, insertions lexicales et parenthétiques (cf. Tab. 1).

Tableau 1 : Liste des disfluences annotées (d'après Crible *et al.* 2016)

Étiquette	Disfluence
LG	Allongements à valeur d'hésitation
UP	Pause silencieuse
FP	Pause pleine
DM	Marqueur de discours
ET	Terme d'édition
FS	Faux-départ
TR	Troncation
RI	Répétition à l'identique
RM	Répétition avec modification
SP	Substitution propositionnelle
SM	Substitution morphologique
IL	Insertion lexicale
IP	Insertion parenthétique

12. Le guide d'annotation syntaxique est téléchargeable [hdl.handle.net/2078/119280].

13. Le guide d'annotation des disfluences est téléchargeable [hdl.handle.net/2078.1/181221].

5.3. Extraction des pauses

Nous retenons pour l’analyse les pauses internes aux productions verbales des locuteurs, en excluant les pauses occasionnant un changement de locuteur (entre deux tours de parole ou à l’occasion d’un chevauchement de parole). Sur un total de 8 759 pauses, les pauses internes aux tours de parole s’élèvent à 4 969.

On observe, dans le diagramme en boîte à moustaches présenté à la Figure 2, la distribution des durées des pauses, en indiquant les seuils couramment utilisés. Ainsi, 50 % des pauses ont une durée comprise entre 250 et 600 ms (rectangle intérieur du diagramme). Si l’on exclut les pauses d’une durée supérieure à 1 000 ms, cela concerne 9,98 % des pauses ; si l’on exclut les pauses d’une durée supérieure à 2 000 ms, cela concerne 1,06 % des pauses. À l’autre extrémité du diagramme, on observe que 28,33 % des pauses ont une durée inférieure à 250 ms et que 17,36 % ont une durée inférieure à 150 ms.

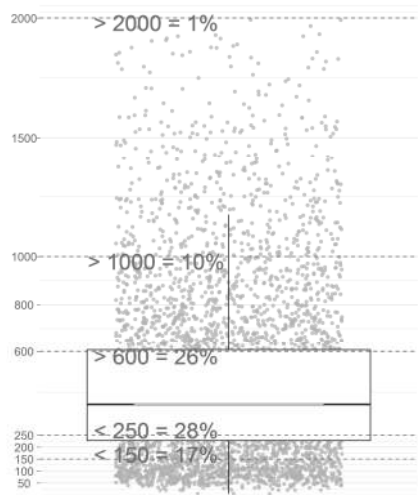


Figure 2 : Diagramme en boîte à moustaches de la distribution des durées pausales avec indications de la proportion des données en fonction des seuils communément utilisés (ms)

Ces données sont exprimées de manière plus détaillée dans le Tableau 2, qui présente la fréquence cumulée des pauses les plus brèves (colonne gauche) et les plus longues (colonne droite).

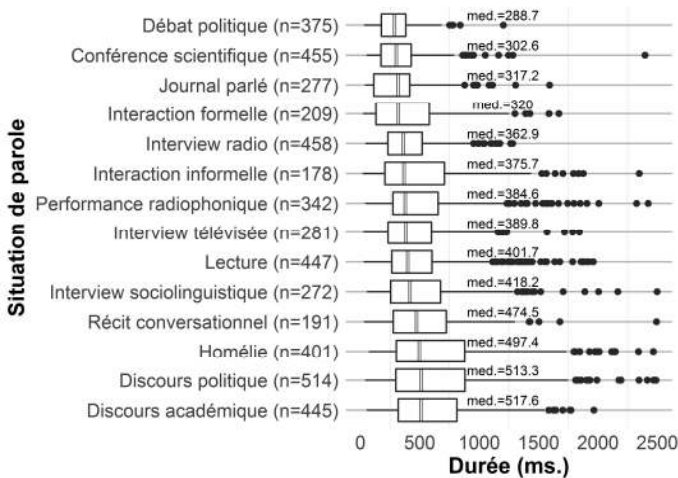
Tableau 2 : Fréquence cumulée des pauses en fonction des seuils inférieurs et supérieurs

Seuil inférieur (ms)	Fréquence	Proportion cumulée (%)	Seuil supérieur (ms)	Fréquence	Proportion cumulée (%)
50	116	2,24	600	1 353	26,12
100	522	10,07	800	816	15,76
150	899	17,36	1 000	517	9,98
200	1 163	22,46	2 000	55	1,062
250	1 467	28,33	3 000	17	0,33

Sur la base de ces données, nous excluons les pauses correspondant aux critères suivants :

- durée absolue : 7 pauses (0,4 % des données) supérieures à 500 ms ou inférieures à 20 ms ont été exclues, car elles étaient attribuables à des erreurs d'annotation ;
- logarithme de base 10 : 117 pauses ont été exclues, car leur valeur était 1,5 fois inférieure ou supérieure à l'écart interquartile (ces pauses représentent respectivement 1,79 % et 3,4 % des données ; v. Tukey 1977) ¹⁴.

La durée médiane de 4 845 pauses analysées dans la suite de cet article varie de 289 ms à 518 ms. L'observation succincte de la Figure 3 indique que la durée pausale varie en fonction des situations de parole, mais également à l'intérieur de celles-ci (cf. § 6).

**Figure 3 : Nombre d'occurrences et médiane de la durée pausale en fonction de la situation de parole**

14. L'exclusion sur base des valeurs transformées est motivée par la volonté d'exclure sur base de la distribution normale (ou quasi-normale) (cf. § 6).

5.4. Extraction des chaînes autour des pauses

Chaque pause intra-tour est catégorisée en fonction de son environnement immédiat. Les pauses dites « seules » sont précédées et suivies de tokens qui n'ont pas été annotés comme des disfluences potentielles (ex. 1) ; les pauses peuvent être précédées et/ou suivies d'un marqueur de discours (ex. 2), d'un *euh* (ex. 3) ou d'une autre marque (comme une troncation ou un faux-départ) (ex. 4). Enfin, les pauses en *interregnum* sont incluses entre un élément à corriger et sa correction (ex. 5).

Exemples de chaînes de pause

- (1) des millions et des millions de personnes viennent ici <pause> à la grotte de Massabielle (hom_2)
- (2) et <pause> selon le New-York Times (jpa_2)
- (3) on est tous montés dans les taxis londoniens euh <pause> quatre par quatre (rcv_3)
- (4) avant de pr/ <pause> et en prenant euh dans chaque maison une personne (rcv_3)
- (5) je ne rencontre pas une <pause> un individu à gauche qui ne me parle pas de cette question (pol_4)
- (6) voilà <pause> euh donc on est parti au mois de juillet (rcv_3)

Les pauses peuvent également s'insérer dans des chaînes de trois (ou plus) éléments (ex. 6). Ces chaînes peuvent contenir plusieurs pauses (p. ex. pause + marqueur + pause + marqueur). Nous avons fait le choix de les regrouper sous l'étiquette « mixte » ($n \geq 3$).

Dans les sections suivantes, nous présentons les analyses de fréquence et de durée des pauses par situation de parole (§ 6), selon les disfluences cooccurentes (§ 7), selon la localisation syntaxique (§ 8) et en prenant en compte tous ces facteurs simultanément.

6. FRÉQUENCE ET DURÉE DES PAUSES EN FONCTION DES SITUATIONS DE PAROLE

La Figure 4 représente la distribution des 4 845 pauses extraites de LOCAS-F, en fonction de leurs durées (durées brutes et durées transformées selon une échelle logarithmique¹⁵). La distribution logarithmique semble bimodale, *i.e.* elle pourrait être formée de deux composantes : un groupe réduit de pauses « brèves » et un groupe plus important de pauses « longues ». Cependant, ce type de distribution bimodale ne se retrouve pas systématiquement si l'on considère

15. Une contribution critique de l'utilisation des transformations logarithmiques a été présentée dans Christodoulides *et al.* (2017).

les durées des pauses produites par chaque locuteur individuellement ¹⁶ ou dans les échantillons de parole correspondant à une situation communicative spécifique.

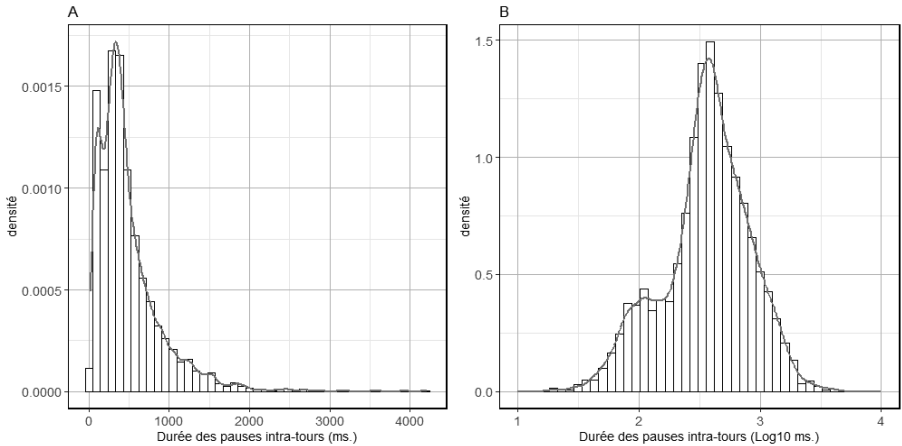


Figure 4 : Distribution des durées pausales sur le corpus LOCAS-F (en durée absolue et logarithmique)

Nous présentons les distributions pour quatre situations de parole représentatives. Les échantillons de discours politiques (Fig. 5a) montrent une distribution bimodale chez 3 locuteurs sur 5 et laissent entrevoir que les durées moyennes des pauses varient d'un locuteur à l'autre. Dans les journaux parlés radiophoniques (Fig. 5b), les distributions de chaque locuteur sont plus homogènes et également bimodales ; ce genre paraît donc relativement standardisé ¹⁷. Par contre, les récits conversationnels sont caractérisés par une distribution plutôt unimodale et relativement homogène entre les locuteurs (Fig. 5c). Enfin, les homélies sont intéressantes, car les pauses de l'ensemble des locuteurs se décrivent plutôt selon une distribution unimodale (Fig. 5d).

Nous en concluons que le caractère unimodal, bimodal, voire multimodal d'une distribution devrait être établi au minimum par situation de parole et parfois par locuteur.

16. Les données par locuteur ont été extraites pour les enregistrements présentant plus de 20 pauses par locuteurs.

17. Dans ce corpus, c'est aussi le cas pour les différents échantillons de conférence scientifique.

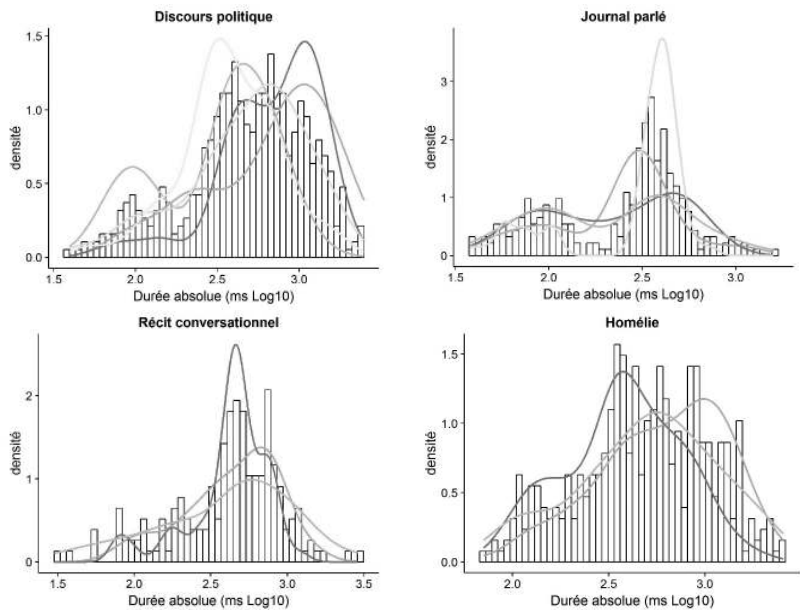


Figure 5 : Distribution des durées (logarithme de base 10) par situation de parole (les lignes représentent les distributions individuelles par locuteur)

7. PAUSES ET DISFLUENCES

Nous nous tournons maintenant vers la fréquence des pauses en fonction de leur cooccurrence avec différentes disfluences. L’annotation (présentée en § 5.2) a été exploitée de manière à extraire des chaînes des éléments qui précèdent et suivent immédiatement les pauses. Au total, 3 518 chaînes ont été extraites (certaines contiennent plus d’une pause) et le Tableau 3 présente leur fréquence.

Tableau 3 : Fréquence des pauses en fonction des cooccurrences les plus fréquentes (pause = pause silencieuse ; md = marqueur de discours ; euh = pause pleine)

Type de chaîne	Chaîne (description)			Fréq. (n)	Prop. (%)
Pause seule				2 593	53,5
Marqueur de discours (md)	pause	+	md	249	5,1
	md	+	pause	684	14,1
Pause pleine (euh)	pause	+	euh	146	3,0
	euh	+	pause	211	4,4

Tableau 3 : (suite)

Disfluence (dis)	pause	+	dis	115	2,4
	dis	+	pause	104	2,1
Pause <i>interregnum</i>				368	7,6
Chaînes $n \geq 3$	(pause + md + euh, pause + euh + pause, etc.)			375	7,7
Total				4 845	100

Les pauses intervenant entre de simples tokens sont les plus fréquentes (53,5 %), suivies par les pauses après un md (14,1 %). Si l'on y ajoute les pauses précédant un md (5,1 %), un *euh* (3 %) ou une disfluence (2,4 %), on obtient un total de 78 % de pauses qui ne suivent pas une disfluence et pourraient être considérées comme structurantes ou grammaticales. Ce total confirme les observations de M. Candea (2000, chap. 3) qui compte dans ses données 82,5 % de pauses structurantes ¹⁸.

Viennent ensuite les pauses internes aux disfluences (7,6 % et 7,7 %) et celles qui suivent un *euh* (4,4 %) ou une autre disfluence (2,1 %), soit un total de 22 % de pauses d'hésitation.

L'analyse de la distribution des chaînes de pauses en fonction du degré de préparation de la parole montre qu'il existe des associations fortes entre certains types de chaînes et les situations de parole préparées, semi-préparées ou spontanées (Fig. 6).

En ordonnée se trouvent les types de chaînes incluant des pauses et en abscisse les situations de parole regroupées selon le trait « degré de préparation ». La taille des boîtes est proportionnelle aux nombre d'occurrences et la taille de l'effet (corrélation positive ou négative) est indiquée par la mesure de résidu.

L'observation la plus évidente est que les pauses isolées sont prédominantes dans les situations de parole préparées (journaux parlés, homélies, discours politiques, etc.) et plus rares dans les styles non préparés (conversation informelle, narration conversationnelle, etc.). Les configurations les plus typiques des situations spontanées sont les séquences de plus de deux éléments, puis les pauses en *interregnum*, les *euh* suivis d'une pause et les pauses précédant un *euh*. En termes de fréquence, l'on n'observe pas de différence entre « *euh* + pause » et « pause + *euh* ». Enfin, les pauses qui suivent ou précèdent un md sont, de manière moins saillante, moins fréquentes dans les situations préparées.

18. En effet, selon Candea (2000), les pauses silencieuses suivies d'une marque de TdF forment une unité séparée de ce qui précède, la pause ayant alors une valeur de segmentation.

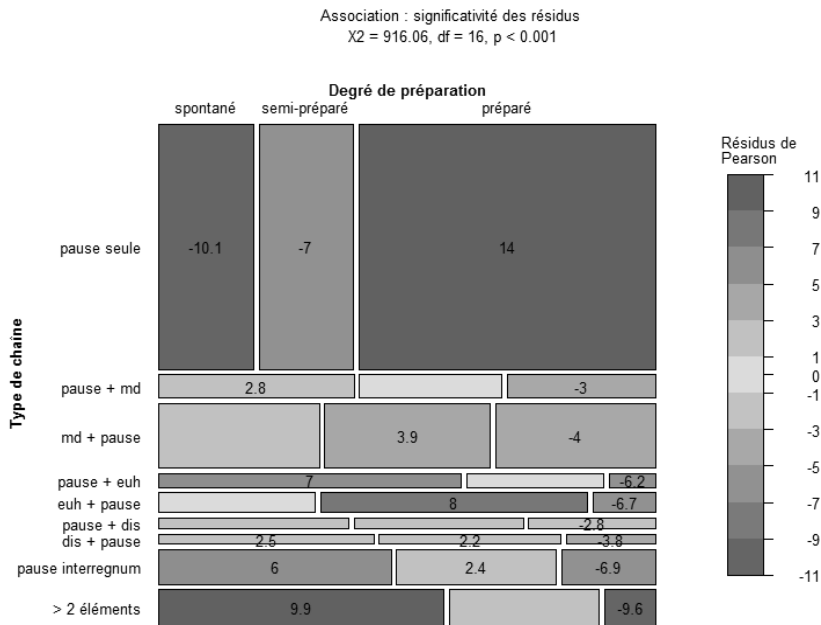


Figure 6 : Tableau d'association du degré de préparation et des cooccurrences

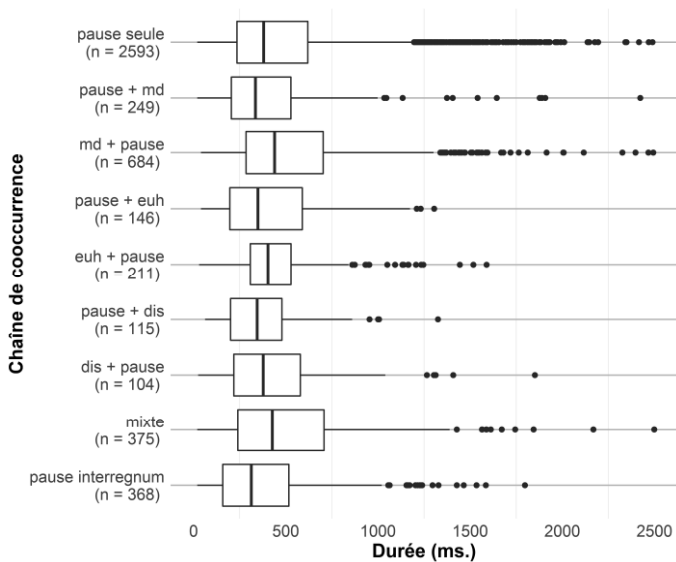


Figure 7 : Durées pauses en fonction des cooccurrences

La Figure 7 illustre les différences de durées en fonction des marques en cooccurrence avec la pause. De manière générale, les pauses précédées d'un *euh* semblent être les plus courtes. La tendance au raccourcissement relatif de la pause suivant une marque ne se retrouve pas dans la configuration « md + pause » ni dans « dis + pause », qui semblent au contraire plus longues.

Peu de différences émergent en fonction du degré de préparation du discours (Fig. 8). L'ordre des durées pausales est préservé entre les différents degrés de préparation, bien que ces durées soient globalement plus longues en situation préparée.

Dans la section suivante, nous centrons l'attention sur les pauses isolées en rendant compte du rôle des facteurs syntaxiques pour expliquer la distribution des pauses.

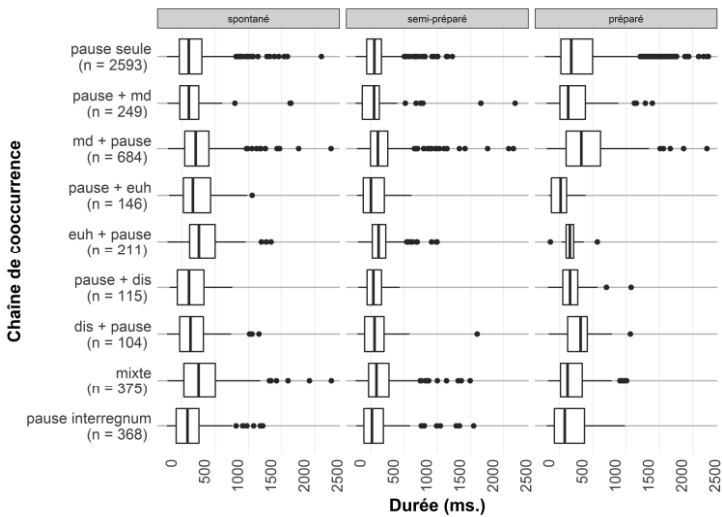


Figure 8 : Durées pausales en fonction des cooccurrences et du degré de préparation

8. PAUSES ET LOCALISATION SYNTAXIQUE

Les pauses se distribuent dans 44,1 % des cas (n = 2 138) entre les constituants syntaxiques minimaux (groupes accentuables ou ga), dans 35,8 % des cas (n = 1 735) à la fin des unités syntaxiques maximales (unités de rection), dans 10,79 % des cas (n = 518) entre les séquences syntaxiques (unités syntaxiques intermédiaires), et enfin dans 9,4 % des cas (n = 454) à l'intérieur des ga. Cette distribution (Fig. 9) est influencée par le degré de préparation. Comme l'on pouvait s'y attendre, les pauses internes aux constituants syntaxiques minimaux, qui interrompent une unité, sont moins fréquentes en parole préparée qu'en parole

semi-préparée ou spontanée. Les pauses qui caractérisent le mieux la parole préparée sont les pauses de niveau intermédiaire, entre les séquences syntaxiques, ce qui peut s’expliquer par le fait que les séquences syntaxiques sont plus longues et plus complexes en parole préparée qu’en parole spontanée ¹⁹.

Les pauses entre les unités syntaxiques maximales (de rection) sont les plus fréquentes en parole spontanée – ce qui peut s’expliquer par la brièveté des unités maximales dans ce style – et les moins fréquentes en parole semi-spontanée – cette dernière tendance étant plus difficile à interpréter.

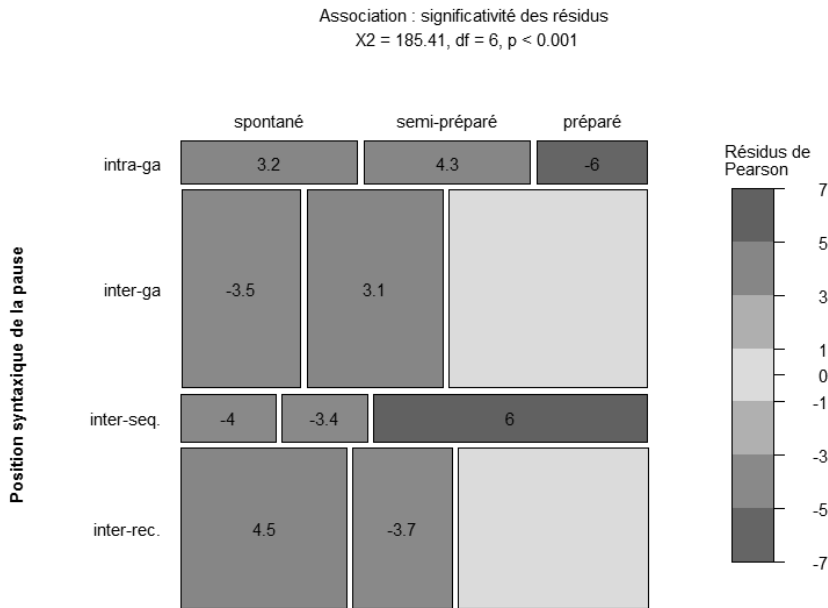


Figure 9 : Tableau d’association du degré de préparation et de la position syntaxique de la pause

La Figure 10 illustre que la durée moyenne d’une pause augmente en fonction du degré de frontière syntaxique et confirme les observations des études existantes (Grosjean & Deschamps 1975 ; Candea 2000 ; Goldman *et al.* 2010). On constate cependant que l’écart entre les pauses inter-séquences et inter-rections

19. LOCAS-F compte 26 607 ga, 16 979 séquences syntaxiques et 12 099 unités syntaxiques maximales de rection. Ces unités présentent des durées plus courtes en parole spontanée et plus longues en parole préparée (pour chaque unité, l’on donne la durée moyenne et la médiane) :

- ga en spontané, moyenne = 841 ms et médiane = 450 ms ; en situation préparée, moyenne = 623 ms et médiane = 560 ms ;
- séquences en spontané, moyenne = 1 110 ms et médiane = 465 ms ; en préparé, moyenne = 1 333 ms et médiane = 717 ms ;
- unités de rection en spontané, moyenne = 1 422 ms et médiane = 508 ms ; en préparé, moyenne = 2 506 ms et médiane = 1 063 ms.

est nettement moins marqué qu'entre pauses intra-ga, inter-ga et inter-séquences. La durée des pauses inter-rections est significativement plus longue dans le style préparé que dans le style semi-préparé ou spontané (Fig. 10-11), différence qui ne se marque pas aux niveaux inférieurs.

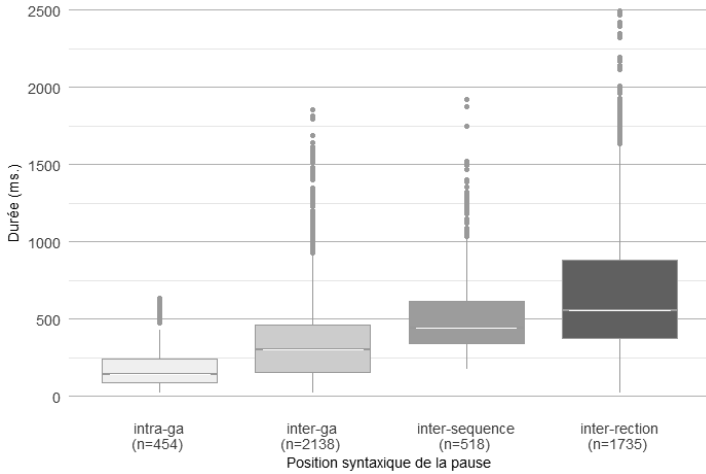


Figure 10 : Durées pausales en fonction de la position syntaxique de la pause
(Kruskal-Wallis chi-squared = 1457.2, dl = 3, $p < 0.001$)

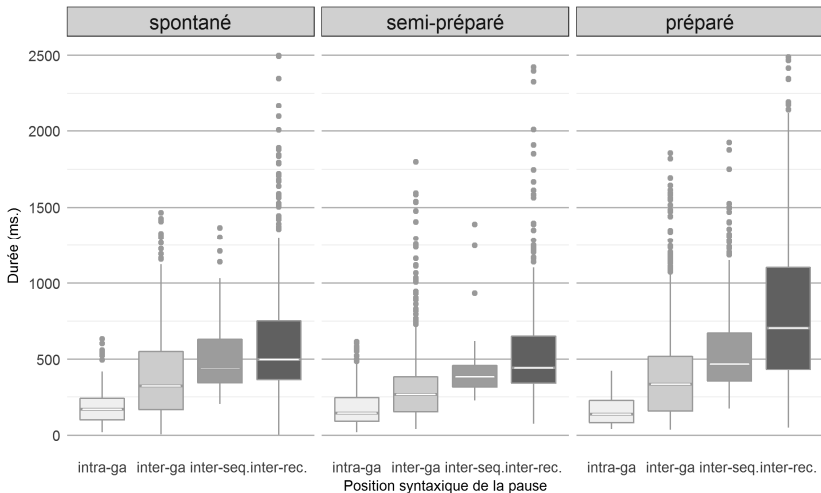


Figure 11 : Durées pausales en fonction de la position syntaxique et du degré de préparation
(K , spontané = 402.5 ; K , semi-préparé = 394.8 ; K , préparé = 643 ; dl = 3, $p < 0.001$)

Si l'on considère que les pauses internes aux *ga*, qui interviennent dans un constituant syntaxique minimal, sont disfluentes, ce schéma confirme que les pauses disfluentes sont plus brèves que les pauses structurantes (Candea, 2000 : 270).

9. ANALYSE MULTIFACTORIELLE ET MODÉLISATION DES DURÉES PAUSALES

Cette dernière section analyse l'interaction entre la position syntaxique et la combinaison avec des disfluences sur la durée pausale, les deux annotations ayant été faites indépendamment (cf. § 5.2). La Figure 12²⁰ permet de montrer que c'est la localisation syntaxique qui explique le mieux la durée d'une pause.

Dans presque tous les types de chaînes, la durée de la pause augmente avec le degré de frontière syntaxique. C'est le cas, en particulier, pour les pauses seules ou en cooccurrence avec des *md*, que nous qualifions de « pauses structurantes ou grammaticales ». En effet, les trois premières lignes de la Figure 9 montrent que les pauses sont les plus longues en position syntaxique inter-rection (frontière syntaxique maximale) et diminuent progressivement avec l'affaiblissement de la force de frontière syntaxique : les pauses sont moins longues aux frontières inter-séquences, encore moins longues en position inter-*ga* et plus courtes en position intra-*ga*.

Deux cas font exception. Premièrement, la durée des « pauses silencieuses + *eu*h » (*n* = 152) semble influencée par le degré de frontière syntaxique, sauf dans les cas intra-*ga* où la pause est beaucoup plus brève. C'est aussi le cas, dans une moindre mesure, pour « *eu*h + pause » et les pauses dans les séquences de plus de deux éléments. On pourrait en conclure que « pause + *eu*h » s'apparente davantage à un environnement disfluent que « *eu*h + pause ».

Nous entreprenons une modélisation des durées pausales à partir du paramètre « situation de parole » (les 14 situations représentées) et des paramètres linguistiques caractérisant l'environnement des pauses (position par rapport aux frontières des unités syntaxiques et cooccurrence avec des disfluences). L'application d'un algorithme de forêt d'arbres décisionnels de classification et de régression permet de sélectionner les variables à partir de leur contribution informationnelle. La Figure 13 présente la contribution de chaque variable au modèle. Nous concluons que c'est la syntaxe qui prédit le mieux la durée d'une pause, suivie de la situation de parole et de la cooccurrence avec les disfluences.

20. Par exemple, la ligne « *dis* + pause » représente les cas où une pause suit une disfluence, que ce soit au début d'une unité de rection, d'une séquence, d'un *ga* ou dans un *ga*.

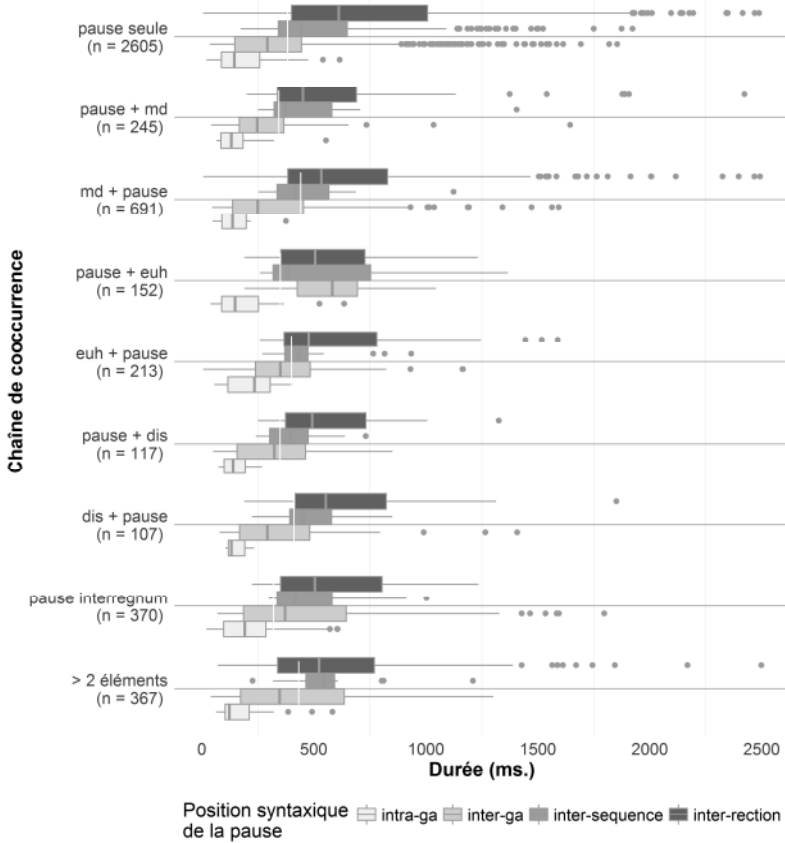


Figure 12 : Diagrammes en boîte à moustaches de la distribution des durées pausales en fonction de leurs cooccurrences et de leur localisation syntaxique

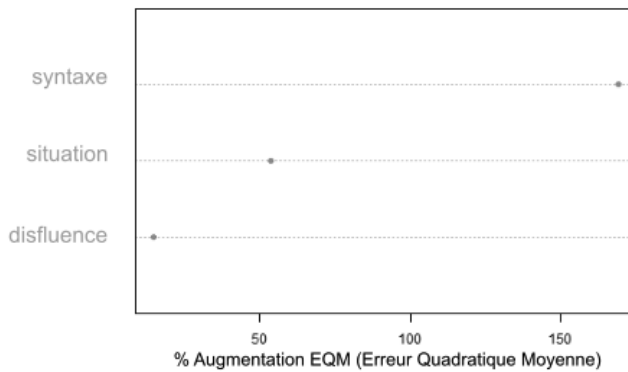


Figure 13 : Contribution des variables (ajustement à la forêt d'arbres décisionnels)

Ainsi, une modélisation à partir des ces trois variables permet d'expliquer 38,78 % de la variance des durées pausales. La permutation de la variable « situation de parole » avec celle « degré de préparation » diminue le pouvoir prédictif du modèle à 31,18 %. Enfin, la considération de seulement deux paramètres (« cooccurrence avec une disfluente » et « position syntaxique ») permet également de maintenir un taux honorable de 31,13 % d'explication de la variabilité des durées pausales. Sur la base de ces taux, nous avons générés les trois arbres de décision²¹. Seuls les deux derniers ont été représentés ici pour des raisons de lisibilité graphique.

L'arbre de décision illustré par la Figure 14 (Annexe) montre que les pauses les plus longues sont situées entre les unités de rection dans la situation de parole préparée. Cet effet de la syntaxe est moins marqué en parole semi-préparée et spontanée, avec pour conséquence que les pauses aux frontières de rection en parole semi-préparée ou spontanée ont la même durée que celles aux frontières de séquence en parole préparée. D'autre part, les pauses les plus brèves se situent à l'intérieur des unités syntaxiques minimales (ga), quel que soit le degré de préparation de la situation de parole. Entre ces deux extrêmes, se trouvent les pauses qui coexistent avec des disfluents. Les nœuds 11 et 12 séparent ainsi respectivement le groupe {*eu*h + pause, pause + *eu*h, pause en *interregnum* et combinaisons mixtes (n>3)} du groupe {pause + dis, dis + pause, pause + md, md + pause, pause seule}, ce dernier ensemble caractérisant des pauses plus brèves.

Les pauses seules se trouvent aux deux extrémités de l'arbre (nœuds 4 et 13). Les plus longues sont situées entre les unités syntaxiques maximales (inter-rection), tandis que les plus courtes sont situées entre ou dans les unités syntaxiques minimales (inter-ga et intra-ga).

Sans perdre en précision, nous avons décidé de représenter ces données selon deux variables uniquement (position syntaxique et cooccurrence avec des disfluents) (Fig. 15 en Annexe). Cette dernière option montre que, à position syntaxique égale, les pauses avant ou après une disfluente sont plus brèves que les pauses seules. Cela pourrait s'expliquer par le fait que la disfluente en elle-même offre déjà un temps de planification. Cependant, cette figure ne met pas systématiquement en évidence un comportement différent des pauses avant ou après disfluente (p. ex. « *eu*h + pause » vs « pause + *eu*h »), mais montre plutôt que les pauses dans un environnement de *eu*h ou en *interregnum* tendent à être un peu plus longues que celles dans l'environnement d'un md ou d'une disfluente (dans la position inter-ga).

21. Les arbres ont été générés sous R (R Core Team 2017) avec le package « party » (Hothorn, Hornik & Zeileis 2006). Les paramètres suivants ont été fixés : profondeur maximale de l'arbre = 4 ; significativité minimale = 0,95 ; nombre minimal d'observations par branche et par nœud terminal = 200 ; nombre d'arbres = 1 000.

10.CONCLUSION

Au terme de cette étude, nous proposons une modélisation de la distribution et de la durée des pauses en fonction des unités syntaxiques et des disfluences, sur la base de l'analyse de 4 845 pauses internes à des tours de parole.

Les résultats montrent que le caractère unimodal, bimodal ou multimodal des distributions de durée varie selon les situations de parole ou selon les locuteurs, et devrait donc être établi au cas par cas, sans qu'il soit possible de définir des durées normalisées générales pour le français parlé. Si la variabilité intra-genre est faible, l'on pourra s'accorder sur une représentation uni-, bi- ou tri-modale par genre. Si la variabilité entre les locuteurs est élevée au sein d'un genre, il faudra considérer une distribution par locuteur.

En raison de la variabilité des débits et des durées pausales, nous ne préconisons pas d'appliquer des seuils de durée fixes pour exclure certaines données (p.ex. comme on retrouve souvent dans la littérature, un seuil inférieur à 200 ms ou supérieur à 500 ms). Les seules pauses exclues dans nos analyses le sont sur la base d'un raisonnement statistique et linguistique (exclusions des données aberrantes liées au traitement des corpus). Cette méthodologie permet d'exclure un nombre réduit de données (moins de 4 % des pauses).

Ensuite, l'analyse de la distribution des pauses cooccurentes avec des contextes dits « fluents » ou « disfluents » fait apparaître que 78 % des pauses peuvent être qualifiées de structurantes ou grammaticales, alors que 22 % seraient des pauses d'hésitation. Plus spécifiquement, il existe des associations fortes entre certaines cooccurrences et le caractère préparé, semi-préparé ou spontané des situations de parole. Ainsi, les pauses structurantes sont prédominantes dans les situations de parole préparées (journaux parlés, homélies, discours politiques, etc.) et plus rares dans les styles non préparés (conversation informelle, narration conversationnelle, etc.). Ces derniers se caractérisent plutôt par des configurations de pauses dans des séquences mixtes (combinant différentes disfluences) ou dans une séquence réparatrice.

Par ailleurs, le degré de préparation de la situation de parole influence également la localisation syntaxique des pauses. Les pauses internes aux constituants syntaxiques minimaux, qui interrompent une unité, sont moins fréquentes en parole préparée qu'en parole semi-préparée ou spontanée. Les pauses qui caractérisent le mieux la parole préparée sont les pauses situées entre les unités syntaxiques intermédiaires (séquences), alors que les pauses entre unités maximales (de rection) sont les plus fréquentes en parole spontanée. Cette différence s'explique principalement par le fait que les unités syntaxiques sont plus longues et plus complexes en parole préparée et plus brèves dans les situations spontanées. Il pourrait donc s'agir d'un besoin d'insérer une pause à intervalle temporel régulier. Enfin, la durée moyenne d'une pause augmente en fonction du degré de frontière syntaxique, quelle que soit la situation de parole ou la séquence environnante.

L'analyse multifactorielle à partir de trois types de paramètres – situationnels, syntaxiques et liés aux disfluences – nous permet de conclure que c'est la syntaxe qui prédit le mieux la durée d'une pause, suivie de la situation de parole et de la cooccurrence avec les disfluences.

Certaines de ces observations empiriques ont été confirmées par des études de perception. A. C. Simon et G. Christodoulides (2016) ont montré que les pauses, même longues, ne sont pas perçues comme des indices de frontière lorsqu'elles sont situées à l'intérieur d'un groupe syntaxique, à cause de l'incomplétude syntaxique qui inhibe la perception de la pause. Il serait intéressant de tester plus spécifiquement la perception des pauses environnant les disfluences (Duez 1995) pour comprendre si elles augmentent ou diminuent la perception de ces marques et interrompent le traitement du flux discursif.

ANNEXES (FIGURES 14-15)

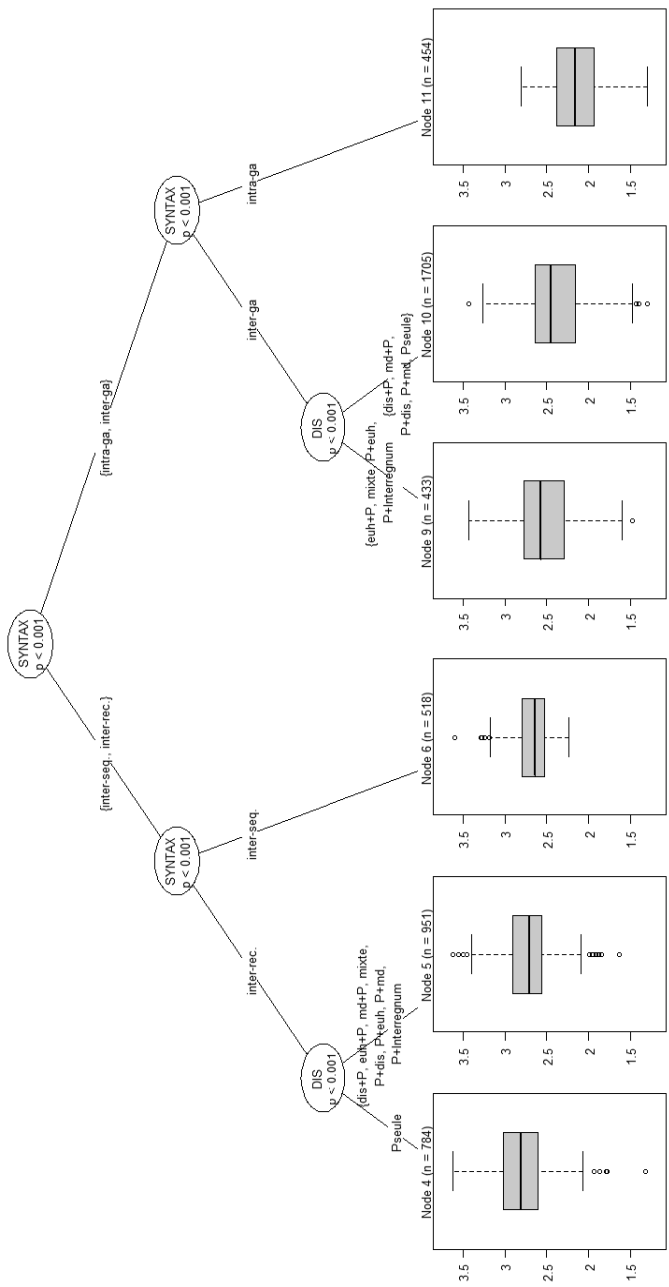


Figure 15 : Arbre de décision des durées pausesles (log10 ms) en fonction de la position syntaxique et de la cooccurrence avec des disfluences

Références

- AVANZI M. *et al.* (2010), "C-PROM. An annotated corpus for French prominence studies", *Proceedings of Prosodic Prominence: Perceptual and Automatic Identification, Proceedings of Speech Prosody 2010 Satellite Workshop*, Chicago (Illinois, USA). [speechprosody2010.illinois.edu/papers/102005.pdf]
- BILGER M. & CAMPIONE E. (2002), « Proposition pour un étiquetage en <séquences fonctionnelles> », *Recherches sur le français parlé* 17, 117-136.
- BLANCHE-BENVENISTE C. *et al.* (1990), *Le français parlé : études grammaticales*, Paris, Éditions du CNRS.
- BOERSMA P. & WEENINK D. (2015), *Praat: Doing Phonetics by Computer*, version 6.0.21. [www.praat.org]
- CAMPIONE E. & VÉRONIS J. (2002), "A large-scale multilingual study of silent pause duration", *Proceedings of the Speech Prosody 2002 Conference*, Aix-en-Provence (France), 199-202.
- CANDEA M. (2000), *Contribution à l'étude des pauses silencieuses et des phénomènes dits « d'hésitation » en français oral spontané : étude sur un corpus de récits en classe de français*, Thèse de l'Université de la Sorbonne nouvelle – Paris 3.
- CHRISTODOULIDES G. *et al.* (2017), "Methodologies for modelling silent pause length: Insights on individual and situational variation from a large-scale corpus study", Conference paper, *International Conference on Fluency and Disfluency across Languages and Language Varieties*, Louvain-la-Neuve (Belgium).
- CRIBLE L. *et al.* (2016), *Annotation Manual of Fluency and Disfluency Markers in Multilingual, Multimodal, Native and Learner Corpora*, version 2.0, rapport technique, Université catholique de Louvain & Université de Namur.
- DE JONG N. H. & BOSKER H. R. (2013), "Choosing a threshold for silent pauses to measure second language fluency", *Proceedings of the 6th Workshop on Disfluency in Spontaneous Speech – DiSS 2013*, Stockholm (Sweden), 17-20.
- DEGAND L. & SIMON A. C. (2005), "Minimal discourse units: Can we define them, and why should we", *Proceedings of SEM-05. Connectors, Discourse Framing and Discourse Structure: From Corpus-Based and Experimental Analyses to Discourse Theories*, Biarritz (France), 65-74.
- DEMOL M., VERHELST W. & VERHOEVE P. (2007), "The duration of speech pauses in a multilingual environment", *Proceedings of the 8th Annual Conference of the International Speech Communication Association – Interspeech 2007*, Antwerp (Belgium), 990-993.
- DUEZ D. (1991), *La pause dans la parole de l'homme politique*, Paris, Éditions du CNRS.
- DUEZ D. (1993), "Acoustic correlates of subjective pauses", *Journal of Psycholinguistic Research* 22 (1), 21-39.
- DUEZ D. (1995), "Perception of hesitations in spontaneous French speech", *Proceedings of International Congress of Phonetic Sciences – ICPhS*, Stockholm (Sweden), vol. 2, 498-501.
- GOLDMAN J.-P. (2011), "EasyAlign: An automatic phonetic alignment tool under Praat", *Proceedings of 12th Annual Conference of the International Speech Communication Association – Interspeech'11*, Firenze (Italy) (s.p.).
- GOLDMAN J.-P. *et al.* (2010), « Étude statistique de la durée pausale dans différents styles de parole », *Actes des XXVIII^{es} Journées d'Étude sur la Parole – JEP 2010*, Mons (Belgique), 161-164.
- GOLDMAN-EISLER F. (1972), "Pauses, clauses, sentences", *Language and Speech* 15 (2), 103-113.

- 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 (3/4), 144-184.
- HIEKE A. E., KOWAL S. & O'CONNELL D. C. (1983), "The trouble with 'articulatory' pauses", *Language and Speech* 26 (3), 203-214.
- HOTHORN T., HORNIK K. & ZEILEIS A. (2006), "Unbiased recursive partitioning: A conditional inference framework", *Journal of Computational and Graphical Statistics* 15 (3), 651-674.
- KIRSNER K., DUNN J. & HIRD K. (2005), "Language production: A complex dynamic system with a chronometric footprint", *Proceedings of the 7th International Conference on Cognitive Systems*, New Delhi (India), 1-11.
- KRIVOKAPIC J. (2006), "Prosodic planning: Effects of phrasal length and complexity on pause duration", *Journal of Phonetics* 35 (2), 162-179.
- LACHERET A. *et al.* (2013), "Prominence perception and accent detection in French: From phonetic processing to grammatical analysis", *Language Sciences* 39, 95-106.
- LACHERET-DUJOUR A. & VICTORRI B. (2002), « La période intonative comme unité d'analyse pour l'étude du français parlé : modélisation prosodique et enjeux linguistiques », *Verbum* XXIV (1-2), 55-72.
- LITTLE D. R. *et al.* (2013), "Fluency profiling system: An automated system for analyzing the temporal properties of speech", *Behavior Research Methods* 45 (1), 191-202.
- MARTIN L. J., DEGAND L. & SIMON A. C. (2014), « Forme et fonction de la périphérie gauche dans un corpus oral multigenres annoté », *Corpus* 13, 243-265.
- MERTENS P. (2008), « Syntaxe, prosodie et structure informationnelle : une approche prédictive pour l'analyse de l'intonation dans le discours », *Travaux de linguistique* 56, 97-124.
- OEHMEN R., KIRSNER K. & FAY N. (2010), "Reliability of the manual segmentation of pauses in natural speech", *Advances in Natural Language Processing* 6253, 263-268.
- R CORE TEAM (2017), *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna (Austria). [<http://www.R-project.org>]
- ROSEN K. M. (2005), "Analysis of speech segment duration with the lognormal distribution: A basis for unification and comparison", *Journal of Phonetics* 33 (4), 411-426.
- SIMON A. C. & CHRISTODOULIDES G. (2016), « Frontières prosodiques perçues : corrélats acoustiques et indices syntaxiques », *Langue française* 191, 83-106.
- SIMON A. C. & DEGAND L. (2011), « L'analyse en <unités discursives de base> : pourquoi et comment ? », *Langue française* 170, 45-59.
- TANGUY N. *et al.* (2012), *Projet FRFC « Périphérie gauche des unités de discours » – Protocole de codage syntaxique*, Document de travail. [hdl.handle.net/2078/119280]
- TUKEY J. W. (1977), *Exploratory Data Analysis*, Reading (MA), Pearson.
- ZELLNER B. (1994), "Pauses and the temporal structure of speech", in E. Keller (ed.), *Fundamentals of Speech Synthesis and Speech Recognition*, Chichester, John Wiley, 41-62.