

Fréquence et typologie des disfluences normales : différence(s) et similitude(s) entre locuteurs qui bégayaient et qui ne bégayaient pas

Laudy Arida¹

UMR8163 STL-Savoirs, Textes, Langage, Université de Lille & CNRS, France

UMR5267 Praxiling Université de Montpellier Paul-Valéry & CNRS, France

Résumé. L'objectif de la présente étude est de décrire les caractéristiques des disfluences « normales » émises par des locuteurs francophones adultes qui bégayaient et qui ne bégayaient pas. C'est dans ce contexte que la parole semi-spontanée de 18 locuteurs francophones adultes qui bégayaient et qui ne bégayaient pas a été analysée. La fréquence d'apparition des disfluences ainsi que leur répartition ont été étudiées afin d'identifier si ces deux groupes de locuteurs présentent des schémas comparables de disfluences. Les résultats obtenus ont montré que les personnes qui bégayaient (PQB) étaient plus disfluentes que les personnes qui ne bégayaient pas (PQBP). Les locuteurs des deux groupes ont émis davantage de disfluences sur les mots de fonction que sur les mots de contenu. De plus, les pauses remplies étaient le type de disfluences le plus récurrent dans les deux groupes. En outre, ils n'ont manifesté de différence significative qu'au niveau des marqueurs de discours qui étaient plus nombreux chez les PQB que chez les PQBP.

Mots-clés (Fr) : disfluences, fréquence d'apparition, répartition, locuteurs francophones qui bégayaient et qui ne bégayaient pas, bégaiement.

1 Introduction

De manière générale, un individu parlant ne peut pas être absolument fluent. Toute émission orale est marquée par des disfluences telles que les répétitions, les allongements, les pauses et autres. La présence de ces disfluences « normales » ne relève pas d'un trouble, mais s'avère essentielle vu qu'elle donne au locuteur du temps supplémentaire pour organiser ou rectifier son discours aussi bien au niveau de la forme que du contenu. Ainsi, la production de disfluences « normales » concerne tout locuteur et apparaît fréquemment lorsque la planification du discours rencontre une difficulté momentanée ([1], [2]), ou encore lorsque les ressources du locuteur se révèlent transitoirement insuffisantes au regard des exigences de la tâche langagière ([3] cité par [4]).

Au fil des années, les disfluences ont fait l'objet de nombreuses études auprès de différents groupes de locuteurs (locuteurs qui bégayaient et qui ne bégayaient pas, apprenants de

¹ laudy.arida15@gmail.com

langues secondes, adultes, enfants, ...), et ce, dans différents langues et contextes de productions (tâche de lecture, de parole semi-spontanée...).

Bon nombre de ces études ont visé à quantifier la fréquence d'occurrences des disfluences. [5] a montré que les locuteurs adultes hongrois présentaient 6,8 disfluences pour 100 mots produits dans une tâche de parole semi-spontanée, tandis que [6] a indiqué un taux de 3,6 disfluences pour 100 mots (in [7]). [8] a évoqué des taux de 7,23 pour 100 mots dans des corpus de parole spontanée en japonais (in [7]). Chez des adultes anglophones âgés de 20 à 51 ans, [9] ont trouvé que les taux moyens des disfluences étaient compris entre 6,4 et 7,8 disfluences pour 100 syllabes émises dans des monologues. Plus récemment, [10] se sont intéressées à l'effet de la nature de la tâche proposée (tâches de lecture, de description d'images et de parole spontanée) sur le taux d'apparition des disfluences chez des adultes francophones. Après avoir analysé les allongements, les modifications, les répétitions, les pauses remplies et les pauses silencieuses, les auteures n'ont pas trouvé de différences significatives entre le taux de disfluences produites dans une tâche de description d'images et de parole spontanée, alors que la lecture entraînait, sans surprise, le moins de disfluences. Des résultats comparables ont été observés lorsque les pauses silencieuses ont été exclues de l'analyse. En outre, [11] se sont intéressées entre autres, aux taux de disfluences « normales » (répétitions de mots multisyllabiques, répétitions de phrases, interjections, révisions, énoncés interrompus) produites à la fois par des locuteurs portugais qui bégayaient et qui ne bégayaient pas, de tout âge. Ces chercheuses ont observé que les deux groupes de locuteurs présentaient des taux comparables de disfluences, et ce, quelle que soit leur tranche d'âge.

Au-delà de l'identification du taux de disfluences émises, des chercheurs se sont intéressés à déterminer les types de disfluences les plus fréquemment observés. [5] a relevé que, chez des locuteurs hongrois, les pauses silencieuses – dont l'inclusion n'est pas systématique en fonction des études – représentaient les disfluences les plus fréquentes, suivies par les hésitations, les pauses remplies, les répétitions, les allongements et les erreurs grammaticales, et ce, indépendamment de l'âge. Contrairement à [5], [9] ont trouvé que les interjections et les révisions étaient les types de disfluences les plus récurrents suivies des répétitions de mots monosyllabiques, tandis que les blocages et les répétitions de mots multisyllabiques étaient plutôt rares. Ces résultats concordent avec ceux de l'étude de [12] qui ont montré qu'en parole semi-spontanée, les interjections étaient le type de disfluences le plus fréquent, suivies par les révisions et les répétitions de mots et de phrases. De même, [11] ont confirmé que les interjections constituaient les disfluences les plus récurrentes dans une tâche de parole spontanée, aussi bien chez les locuteurs qui bégayaient que chez ceux qui ne bégayaient pas, indépendamment de la tranche d'âge des locuteurs, alors que [13], ont trouvé, chez des locuteurs indiens parlant anglais, lors d'une présentation orale en classe, que les pauses remplies constituaient le type de disfluences le plus récurrent. Chez des locuteurs turcs, [14] ont signalé que ce sont les allongements, les pauses remplies et les hésitations qui représentent les disfluences les plus fréquentes, quelle que soit la tâche proposée (parole spontanée ou préparée). [10] ont conclu une forte corrélation entre la nature de la tâche et le type de disfluences le plus fréquemment produit. Plus précisément, elles ont trouvé que (1) la parole spontanée se caractérise principalement par des allongements des voyelles et par des pauses silencieuses et remplies ; (2) la description d'images se distingue par des pauses silencieuses et des allongements des voyelles et (3) la lecture se caractérise principalement par des pauses silencieuses.

Les études visant les disfluences « normales » ne se limitent pas à l'identification de leur taux d'occurrences ou au type le plus fréquemment produit, mais explorent un large éventail d'analyses. Ces dernières concernent, à titre d'exemple, des aspects phonétiques comme par exemple la fréquence fondamentale, la durée et le débit. Ainsi, [15], dans leur étude sur un

corpus oral multigenre français, ont trouvé que les pauses pleines et les allongements se caractérisent par une baisse au niveau de la fréquence fondamentale (F0). En outre, [16] ont complété ces observations en montrant que la prosodie des répétitions est relativement stable dans un corpus multigenre français et [17] ont trouvé que les marqueurs de discours étaient produits avec un débit plus rapide que les autres unités lexicales, puisqu'ils sont hypoarticulés.

Parallèlement à ces recherches portant sur l'aspect phonétique et articulatoire des disfluences « normales », d'autres se sont intéressées à l'aspect linguistique, comme par exemple celles de [18] et [19] qui ont étudié l'effet que peut avoir la classe grammaticale des unités lexicales sur l'apparition des disfluences chez des locuteurs adultes normo-fluents. Ces chercheurs ont trouvé que les mots de fonction étaient davantage disfluents en comparaison aux mots de contenu, dans des corpus de parole spontanée anglais et portugais brésilien. [20] ont rapporté que les mots de fonction étaient davantage touchés par des répétitions que les mots de contenu et que les allongements concernaient davantage les mots de fonction monosyllabiques que les mots de contenu pour lesquels l'allongement se portait sur la dernière syllabe. Un résultat similaire a été observé par [21] en suédois et par [22] en mandarin qui ont trouvé que les allongements surviennent plus fréquemment sur les mots de fonction que sur les mots de contenu. En revanche, dans une étude plus récente, [23] n'ont pas trouvé de différence significative entre le taux d'allongements qui touchent les mots de fonction et les mots de contenu. En outre, [22] ont trouvé que plus que la moitié des allongements était observée sur les unités lexicales monosyllabiques, qu'il s'agisse de mots de fonction ou de mots de contenu. Ces auteurs ont également analysé la répartition des allongements selon la catégorie grammaticale (adverbe, déterminant...) en mandarin. Leurs résultats ont montré que les verbes transitifs (18,96%), les adverbes (14,38%), les noms (11,32%) et les particules (16,67%) étaient les catégories les plus touchées. D'autres chercheurs se sont intéressés à la localisation des disfluences « normales » dont [24], qui a trouvé que la fréquence de disfluences était plus élevée en position initiale en raison (1) de la planification accrue que doit effectuer le locuteur à ce moment et (2) du comportement qu'il adopte pour prendre ou conserver son tour de parole.

Ainsi, nous remarquons une certaine hétérogénéité dans les résultats des différentes études. Alors que certains semblent se généraliser dans différentes langues (par exemple, les interjections constituaient les disfluences les plus récurrentes chez des locuteurs anglophones [9] et portugais [11]), alors que d'autres non (par exemple, dans un corpus de parole spontanée, les pauses silencieuses étaient le type de disfluences le plus fréquemment produit par des locuteurs adultes hongrois [5], alors que chez les locuteurs portugais, c'était les interjections [11] et chez des locuteurs turcs, c'était les allongements [14]). C'est dans ce contexte qu'il nous semble important de nous intéresser et à la fréquence et à la répartition des disfluences émises par des locuteurs francophones adultes qui ne bégaiement pas. En outre, il nous semble important d'élargir nos analyses en examinant aussi les disfluences « normales » produites par des locuteurs adultes francophones qui bégaiement. Même si le bégaiement, défini comme un trouble neurodéveloppemental de la fluence de la parole [25], principalement caractérisé par une incidence élevée de disfluences typiques du bégaiement (telles que des prolongations, des blocages, des répétitions de sons, de syllabes ou de mots monosyllabiques, ou encore des évitements de certains sons voire mots donnant lieu à des substitutions [26]), il faut toutefois préciser que ces locuteurs ne produisent pas exclusivement des disfluences « pathologiques », mais manifestent également des disfluences « normales », similaires à celles observées dans la parole de tout sujet. L'élaboration d'une comparaison inter-groupe permet d'identifier si les comportements et les stratégies utilisées par les locuteurs varient en fonction de la présence/absence du trouble de la fluence.

2 Objectifs et hypothèses

La présente étude a pour objectif de décrire les caractéristiques des disfluences « normales » – plus précisément leur fréquence et leur répartition –, chez des locuteurs adultes francophones qui bégaièrent et qui ne bégaièrent pas.

Étant donné que les causes d'apparition des disfluences « normales » sont similaires dans les deux groupes de locuteurs (problème de planification linguistique [2]), nous émettons l'hypothèse que les deux groupes présenteront des résultats comparables pour les disfluences (en termes de fréquence et de répartition). Nous nous attendons également à trouver que les traits lexicaux influenceront de façon similaire le taux d'apparition des disfluences au sein des deux groupes de locuteurs.

3 Méthodologie

3.1 Participants

Dans le cadre de cette étude, la parole de 18 locuteurs francophones adultes (recrutés dans le cadre du projet BENEPHIDIRE (ANR-18-CE36-0008)²) : 9 qui bégaièrent (PQB) et 9 qui ne bégaièrent pas (PQBP) a été analysée. Les deux groupes de locuteurs ont été appariés selon l'âge (l'âge moyen était de 26 ans et 11 mois (ET = 7,99 ; âge minimal = 19 ans ; âge maximal = 45 ans)), le genre (5 hommes et 4 femmes dans chaque groupe) et la catégorie socio-professionnelle (Tableau 1).

Pour l'ensemble des participants – qu'ils présentent ou non un bégaiement –, les critères d'inclusion étaient les suivants : les participants devaient être âgés entre 18 et 50 ans et avoir le français comme langue maternelle. Ils ne devaient présenter aucun (1) antécédent neurologique tel qu'un traumatisme crânien, un accident vasculaire cérébral, une tumeur cérébrale ; (2) (autre) trouble neurodéveloppemental (que le bégaiement pour les PQB) ; (3) antécédent psychiatrique. En outre, ils ne devaient pas prendre un médicament de type antidépresseur ou anxiolytique.

En ce qui concerne le groupe de locuteurs qui bégaièrent, ils devaient présenter un bégaiement développemental persistant diagnostiqué. Quant aux locuteurs qui ne bégaièrent pas, ils ne devaient/doivent présenter aucun problème de fluence, ou avoir un antécédent familial de bégaiement.

² ANR Benephidire (le bégaiement : la neurologie, la phonétique et l'informatique pour son diagnostic et sa rééducation), porteur : Fabrice Hirsch.

Tableau 1 : Présentation des participants qui bégaièrent et qui ne bégaièrent pas.

Participant	Âge	Sexe	Niveau socio-éducatif ³	Degré de sévérité du bégaiement (SSI-4 [27])	Participant	Âge	Sexe	Niveau socio-éducatif
B13	21	Homme	14	Léger à modéré	C15	22	Homme	15
B2	24	Homme	12	Modéré	C4	23	Homme	17
B7	25	Femme	17	Modéré	C9	25	Femme	17
B10	21	Femme	14	Modéré	C2	19	Femme	13
B17	19	Femme	13	Modéré	C10	20	Femme	14
B22	34	Homme	17	Modéré	C3	30	Homme	17
B20	20	Homme	13	Modéré à sévère	C14	21	Homme	14
B25	45	Homme	11	Modéré à sévère	C1	45	Homme	16
B18	27	Femme	11	Très sévère	C6	29	Femme	17

3.2 Corpus

Afin de mener cette étude, des entretiens dirigés ont été réalisés dans une salle calme. Les locuteurs ont été enregistrés à l’aide d’un micro-cravate Sennheiser ME-2 II et d’une caméra Panasonic Full HD HC-V750, en train de répondre à une série de questions ouvertes portant sur plusieurs thèmes comme par exemple décrire une journée type, parler du confinement, des vacances et autres. La durée moyenne des productions de parole semi-spontanée était de 4,03 minutes par locuteur (ET = 1,22 ; min = 2,41 min ; max = 6,35 min).

3.3 Identification des disfluences

Pour identifier les disfluences « normales », nous nous sommes basée sur le modèle d’annotation « ANODIS » [28].

Ces disfluences regroupent :

- Les répétitions « normales » qui correspondent à toute réitération à l’identique qui touche les phonèmes (ex : a-après), les syllabes (ex : ma-magasin), les unités lexicales (ex : pantalon-pantalon), ou les groupes des unités lexicales (ex : j’ai acheté un- j’ai acheté un pantalon).
- Les allongements « normaux » qui s’appliquent à tout allongement non phonologique de la durée d’un son (voyelle ou consonne) perçu comme anormal (ex : aprèèès). Pour l’identification de cette disfluence, nous n’avons pas eu recours à des seuils de durée.

³ Pour calculer ce score, un point a été attribué à chaque année d’étude en partant du cours préparatoire.

- Les pauses remplies qui renvoient à une vocalisation non lexicale telle que : /ø/.
- Les marqueurs de discours correspondant à une expression syntaxique qui n'apporte pas (ou très peu) de contenu propositionnel à l'énoncé tels que : ben, tu vois.

Notons que l'absence de tension, détectée lors du visionnement des vidéos et/ou à l'écoute des enregistrements, était le critère principal permettant de distinguer les répétitions et les allongements « normaux » – objets de la présente étude – des répétitions et des allongements « pathologiques » qui correspondent aux disfluences typiques du bégaiement.

Un type supplémentaire, appelé « disfluence composée », a été créé pour les cas où plusieurs disfluences apparaissaient simultanément sur une même unité lexicale (par exemple : unité lexicale à la fois répétée et allongée).

3.4 Analyse des données

L'ensemble des transcriptions orthographiques du corpus (y compris les disfluences) étaient importées dans le logiciel TreeTagger [29] afin d'identifier la classe grammaticale des unités lexicales émises par les locuteurs. Des vérifications et des réajustements manuels ont été effectués sur l'identification des unités lexicales proposée par le logiciel, tout en se basant sur le système d'annotation morpho-syntaxique du français [30]. Plus spécifiquement, les années, les nombres, les titres des films et des livres, les noms de ville et de pays, les sigles, les noms et les prénoms composés, les noms de famille composés avec ou sans particule, les noms de journaux avec ou sans article ainsi que les pronoms composés, démonstratifs composés avec -ci et -là ont été considérés comme une seule unité lexicale. En revanche, les cas d'élision ont été considérés comme plusieurs unités lexicales distinctes. Ensuite, les unités lexicales étaient classées en mots de contenu et mots de fonction. La première catégorie regroupait les noms communs ou propres, les verbes conjugués ou à l'infinitif (à l'exception des verbes modaux), les adverbes et les adjectifs qualificatifs, alors que la deuxième renfermait les déterminants, les pronoms, les conjonctions, les prépositions, les auxiliaires, les verbes modaux, les adjectifs non qualificatifs et les interjections.

Les analyses statistiques ont été menées dans le logiciel statistique IBM SPSS Statistics (version 27 [31]).

4 Résultats

4.1 Nombre de disfluences observées

Parmi les 4484 unités lexicales produites par le groupe des PQBP lors d'une tâche de parole semi-spontanée, 741 unités lexicales disfluentes ont été relevées (16,52%, ET = 0,74), alors que, chez les PQB, 929 unités lexicales disfluentes ont été observées sur un total de 4936 unités lexicales émises (18,82%, ET = 0,78). L'application du test de χ^2 montre que la différence entre ces taux de disfluences est significative ($\chi^2(1) = 8,488$; $p = 0,004$) (Figure 1).

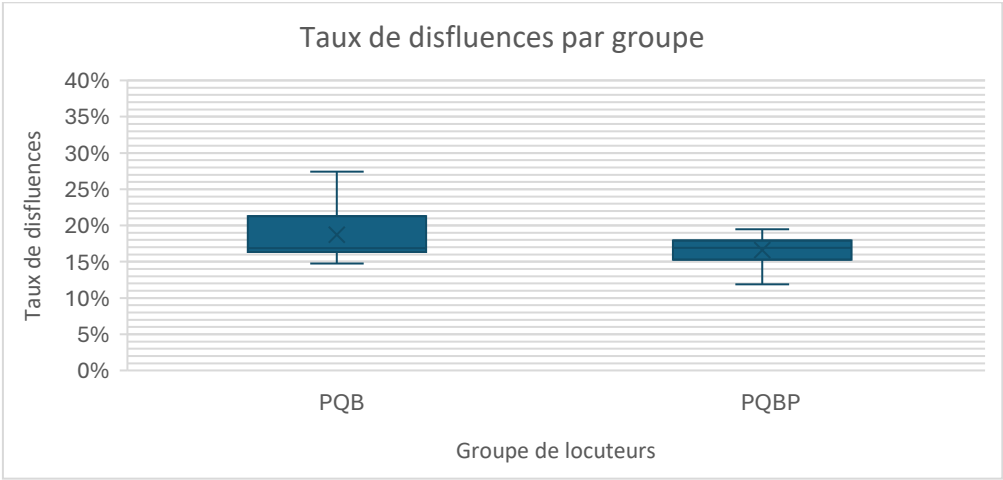


Figure 1 : Taux de disfluences par groupe.

4.2 Type de disfluences observées

Dans le groupe des PQBP, 741 disfluences ont été observées. Les pauses remplies présentaient le taux le plus élevé (n = 292 ; 6,51%), suivies des allongements (n = 161 ; 3,59%), des marqueurs de discours (n = 149 ; 3,32%), des répétitions de phrases (n = 64 ; 1,43%), des répétitions de mots (n = 36 ; 0,80%), des disfluences composées (n = 29 ; 0,65%), des répétitions de sons (n = 8 ; 0,18%), et des répétitions de syllabes (n = 2 ; 0,04%).

Dans le corpus semi-spontané des PQB, 929 disfluences ont été notées. Les pauses remplies étaient les disfluences les plus fréquentes (n = 332 ; 6,73%), suivies des allongements (n = 196 ; 3,97%), des marqueurs de discours (n = 190 ; 3,85%), des répétitions de phrases (n = 99 ; 2,01%), des disfluences composées (n = 67 ; 1,36%), des répétitions de mots (n = 36 ; 0,73%), des répétitions de sons (n = 7 ; 0,14%) et des répétitions de syllabes (n = 2 ; 0,04%).

Les comparaisons entre les proportions des divers types de disfluences, effectuées à l’aide du test de χ^2 avec correction de continuité, font apparaître des différences significatives pour toutes les paires comparées à l’exception des paires suivantes : allongements - marqueurs de discours, répétitions de sons - répétitions de syllabes et répétitions de mots - disfluences composées chez les PQBP et allongements - marqueurs de discours, répétitions de sons - répétitions de syllabes chez les PQB. Dans les deux groupes de locuteurs, les pauses remplies étaient significativement plus fréquentes que tous les autres types de disfluences (Figure 2 et Figure 3).

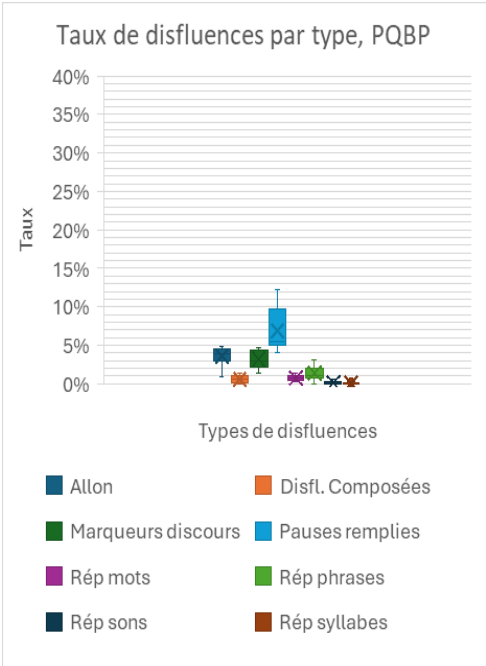


Figure 2 : Taux de disfluences par type, PQBP.

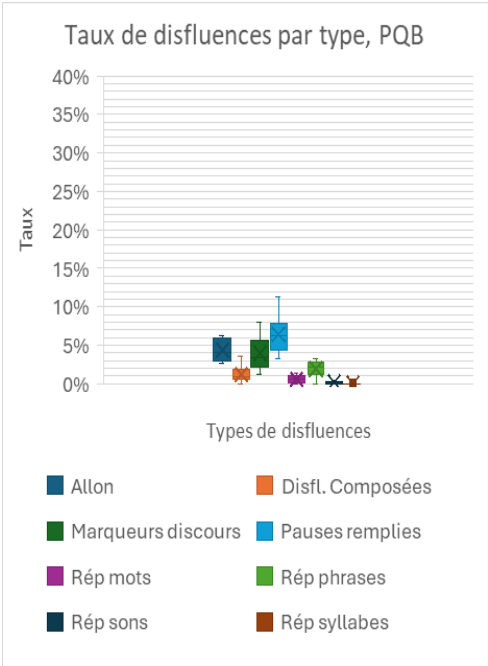


Figure 3 : Taux de disfluences par type, PQB.

Les résultats de la comparaison intergroupe, obtenus à l’aide du test de Mann-Whitney, montrent qu’aucune différence significative n’a été observée entre les deux groupes de locuteurs concernant le taux de l’ensemble des types de disfluences étudiés (allongements, disfluences composées, marqueurs de discours, pauses remplies, répétitions de mots, de phrases, de sons, et de syllabes) (Tableau 2).

Tableau 2 : Comparaison intergroupe entre le taux des différents types de disfluences émises par les PQB et les PQBP.

	Types de disfluences							
	Allon	Disfl. Composées	Marqueurs discours	Pauses remplies	Rép mots	Rép phrases	Rép sons	Rép syllabes
PQBP	3,59%	0,64%	3,32%	6,51%	0,80%	1,42%	0,17%	0,04%
PQB	3,97%	1,35%	3,84%	6,72%	0,72%	2,00%	0,14%	0,04%
Sig. asymptotique (bilatérale)	,402	,309	,566	,965	,453	,185	,771	,808
Sig. exacte [2*(sig. unilatérale)]	,436	,340	,605	1,000	,489	,190	,796	,863

4.3 Nombre de disfluences observées en fonction de la classe grammaticale

Lors de la tâche de parole semi-spontanée, les PQBP ont produit 4464 unités lexicales (2012 mots de contenu et 2452 mots de fonction). Au total, 741 disfluences ont été observées (176 sur les mots de contenu (8,74%, ET = 0,57) et 565 sur les mots de fonction (23,04%, ET= 0,84)). D’après le test de χ^2 , ces locuteurs sont significativement plus disfluents sur les mots de fonction que sur les mots de contenu ($\chi^2(1) = 163,126$; $p = 0,000$) (Figure 4).

De même, dans le groupe des PQB, le taux de disfluences sur les mots de fonction (24,02%, ET = 0,85 ; 665 disfluences sur un total de 2768) était significativement plus important que celui des mots de contenu (12,37%, ET = 0,66 ; 264 disfluences sur un total de 2134) ($\chi^2(1) = 106,569$; $p = 0,000$) (Figure 5).

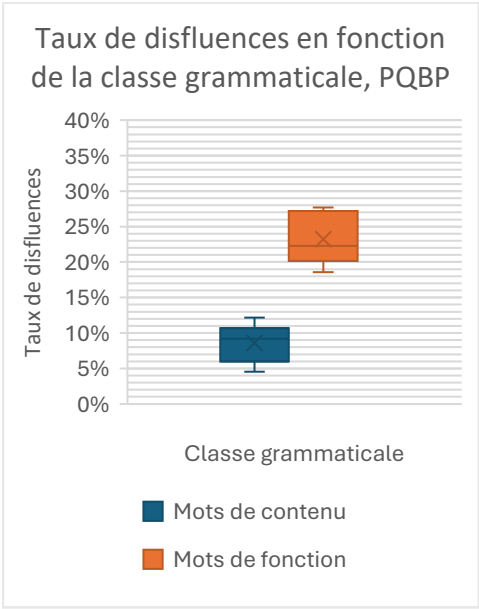


Figure 4 : Taux de disfluences en fonction de la classe grammaticale, PQBP.

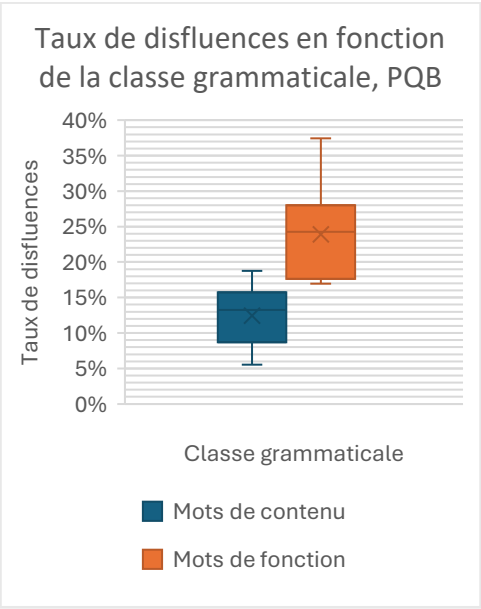


Figure 5 : Taux de disfluences en fonction de la classe grammaticale, PQB.

4.4 Type de disfluences observées en fonction de la classe grammaticale

Dans le groupe des PQBP, les mots de contenu étaient les plus fréquemment touchés par des allongements ($n = 91$; 4,52%), suivis des marqueurs de discours ($n = 44$; 2,19%), des répétitions de phrases ($n = 24$; 1,19%), des disfluences composées ($n = 6$; 0,30%), des répétitions de sons ($n = 5$; 0,25%), des répétitions de mots ($n = 4$; 0,20%) et des répétitions de syllabes ($n = 2$; 0,10%). Les résultats des comparaisons par paires effectuées à l’aide du test de χ^2 avec correction de continuité ont montré des différences statistiquement significatives entre la fréquence des allongements et l’ensemble des autres types de disfluences (répétitions de phrases ($p < 0,001$), répétitions de mots ($p < 0,001$), répétitions de syllabes ($p < 0,001$), répétitions de sons ($p < 0,001$), disfluences composées ($p < 0,001$) et marqueurs de discours ($p < 0,001$)). Les marqueurs de discours étaient significativement plus fréquents que les répétitions de mots, de syllabes, de sons et les disfluences composées ($p < 0,001$ pour l’ensemble des comparaisons). Aussi, les répétitions de phrases se

distinguaient significativement des répétitions de syllabes ($p < 0,001$), de mots ($p < 0,001$), de sons ($p < 0,001$) et des disfluences composées ($p = 0,001$) (Figure 6).

Pour les mots de fonction, le type de disfluences le plus fréquent était les pauses remplies ($n = 292$; 11,91%) – classées parmi les mots de fonction du fait qu’elles réfèrent à des interjections –, suivi par celui des marqueurs de discours ($n = 105$; 4,28%), des allongements ($n = 70$; 2,85%), des répétitions de phrases ($n = 40$; 1,63%), des répétitions de mots ($n = 32$; 1,31%), des disfluences composées ($n = 23$; 0,94%), et des répétitions de sons ($n = 3$; 0,12%). Les comparaisons par paires ont mis en évidence la prédominance des pauses remplies sur tous les autres types de disfluences ($p < 0,001$ pour toutes les comparaisons). Les marqueurs de discours survenaient plus fréquemment que les répétitions de phrases, de mots, de sons, et les disfluences composées ($p < 0,001$ pour l’ensemble des comparaisons). Les allongements étaient significativement plus récurrents que les disfluences composées ($p < 0,001$), les répétitions de sons ($p < 0,001$), de mots ($p < 0,001$) et de phrases ($p = 0,004$). En outre, des différences significatives ont été observées entre les répétitions de phrases et les disfluences composées ($p = 0,042$), les répétitions de phrases et de sons ($p < 0,001$), les répétitions de mots et de sons ($p < 0,001$), les disfluences composées et les répétitions de sons ($p < 0,001$) (Figure 7).

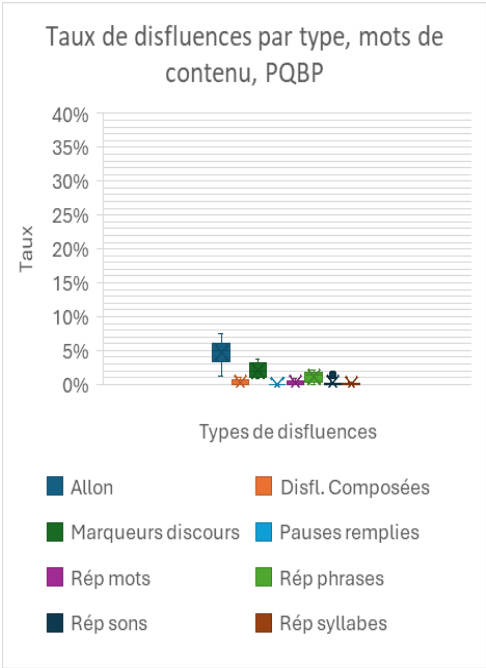


Figure 6 : Taux de disfluences par type, mots de contenu, PQBP.

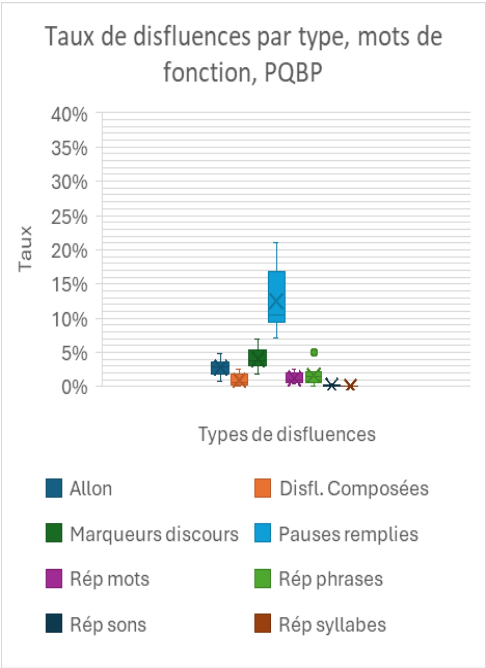


Figure 7 : Taux de disfluences par type, mots de fonction, PQBP.

L’analyse des types de disfluences touchant les mots de contenu chez les PQB montre que les marqueurs de discours présentaient la fréquence d’occurrences la plus élevée ($n = 108$; 5,06 %), suivis par les allongements ($n = 96$; 4,50 %), les répétitions de phrases ($n = 37$; 1,73 %), les disfluences composées ($n = 11$; 0,52 %), les répétitions de sons ($n = 6$; 0,28 %), les répétitions de mots ($n = 4$; 0,19 %) et les répétitions de syllabes ($n = 2$; 0,09 %). Les comparaisons par paires ont montré que les marqueurs de discours étaient significativement plus fréquents que les répétitions de phrases, de syllabes, de mots, de sons,

et les disfluences composées ($p < 0,001$ pour toutes les comparaisons). Les allongements étaient significativement plus récurrents que les répétitions de phrases, de syllabes, de mots, de sons et des disfluences composées ($p < 0,001$ pour l'ensemble des comparaisons). La fréquence des répétitions de phrases se présentait significativement plus importante que celles des répétitions de mots, de syllabes, de sons et des disfluences composées ($p < 0,001$). Une différence significative a été également observée entre les disfluences composées et les répétitions de syllabes ($p = 0,034$) (Figure 8).

En ce qui concerne les mots de fonction, les pauses remplies étaient les plus récurrentes ($n = 332$; 11,99%), suivies des allongements ($n = 100$; 3,61%), des marqueurs de discours ($n = 82$; 2,96%), des répétitions de phrases ($n = 62$; 2,24%), des disfluences composées ($n = 56$; 2,02%), des répétitions de mots ($n = 32$; 1,16%) et des répétitions de sons ($n = 1$; 0,04%). Les comparaisons par paires indiquent que les pauses remplies étaient significativement plus fréquentes que tous les autres types de disfluences ($p < 0,001$ pour l'ensemble des comparaisons). Les allongements, les marqueurs de discours, les répétitions de phrases et les disfluences composées présentaient tous une fréquence d'occurrences significativement plus importante que celle des répétitions de sons ($p < 0,001$). Les allongements étaient significativement plus importants que les répétitions de mots ($p < 0,001$), les disfluences composées ($p < 0,001$) et les répétitions de phrases ($p = 0,003$). Des différences statistiquement significatives ont également été observées entre les marqueurs de discours et les répétitions de mots ($p < 0,001$), les marqueurs de discours et les disfluences composées ($p = 0,029$), les répétitions de phrases et les répétitions de mots ($p = 0,003$), ainsi qu'entre les disfluences composées et les répétitions de mots ($p = 0,014$) (Figure 9).

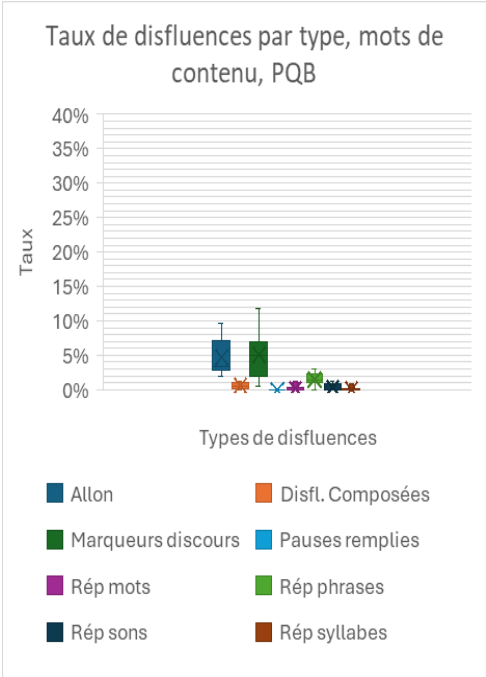


Figure 8 : Taux de disfluences par type, mots de contenu, PQB.

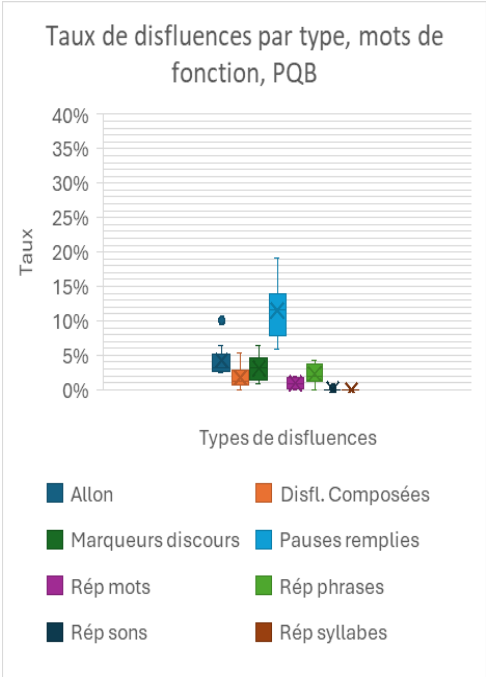


Figure 9 : Taux de disfluences par type, mots de fonction, PQB.

La comparaison intergroupe ne révèle aucune différence significative entre les deux groupes de locuteurs dans la fréquence des différents types de disfluences touchant d'une

part les mots de contenu et d'autre part les mots de fonction. Cependant, une différence marginalement significative a été observée pour les marqueurs de discours. Ces derniers sont légèrement plus fréquents dans les mots de contenu chez les PQBP que chez les PQB (Tableau 3 et Tableau 4).

Tableau 3 : Comparaison intergroupe entre le taux des différents types de disfluences, touchant les mots de contenu, émises par les PQB et les PQBP.

	Types de disfluences ^a						
	Allon- contenu	Disfl. composées- contenu	Marqueurs discours- contenu	Rép mots- contenu	Rép phrases- contenu	Rép sons- contenu	Rép syllabes- contenu
PQBP	4,52%	0,30%	2,19%	0,20%	1,19%	0,25%	0,10%
PQB	4,50%	0,52%	5,06%	0,19%	1,73%	0,28%	0,09%
Sig. asymptotique (bilatérale)	,627	,106	,047	,793	,269	,400	,903
Sig. exacte [2*(sig. unilatérale)]	,666 ^b	,136 ^b	,050 ^b	,863 ^b	,297 ^b	,489 ^b	,931 ^b
a. Variable de regroupement : groupes							
b. Non corrigé pour les ex aequo.							

Tableau 4 : Comparaison intergroupe entre le taux des différents types de disfluences, touchant les mots de fonction, émises par les PQB et les PQBP.

	Types de disfluences ^a						
	Allon- fonction	Disfl. composées- fonction	Marqueurs discours- fonction	Pauses remplies- fonction	Rép mots- fonction	Rép phrases- fonction	Rép sons- fonction
PQBP	2,85%	0,94%	4,28%	11,91%	1,31%	1,63%	0,12%
PQB	3,61%	2,02%	2,96%	11,99%	1,16%	2,24%	0,04%
Sig. asymptotique (bilatérale)	,233	,353	,270	,895	,627	,427	,203
Sig. exacte [2*(sig. unilatérale)]	,258 ^b	,387 ^b	,297 ^b	,931 ^b	,666 ^b	,436 ^b	,387 ^b
a. Variable de regroupement : groupes							
b. Non corrigé pour les ex aequo.							

5 Discussion

L'objectif de la présente étude était de décrire les caractéristiques des disfluences « normales » produites par des locuteurs adultes francophones qui bégayaient et qui ne bégayaient pas. Nous nous sommes plus particulièrement intéressée à identifier le taux de disfluences ainsi que le type de disfluences le plus fréquemment émis. Nous supposons que les deux groupes de locuteurs présenteraient des résultats comparables en termes de fréquence et de répartition des disfluences. Nous nous attendions également à ce que les traits lexicaux influencent le taux ainsi que la répartition des disfluences de façon comparable dans les deux groupes de locuteurs.

Contrairement à notre hypothèse postulant une similarité intergroupe, les PQB ont présenté significativement plus de disfluences en comparaison avec les PQBP. Ce résultat peut être expliqué par l'utilisation de stratégies individuelles de compensation réalisées par les PQB. Autrement dit, ces locuteurs pourraient par exemple produire des pauses silencieuses, des marqueurs de discours ou d'autres types de disfluences « normales », leur permettant de gagner du temps pour finaliser la planification motrice de l'unité à produire et, par conséquent éviter la production des disfluences « pathologiques ». Ce résultat est en discordance avec celui de l'étude de [11] qui ont rapporté que les groupes de locuteurs portugais qui bégayaient et qui ne bégayaient pas présentaient des taux comparables de disfluences. La différence entre les résultats des deux études pourrait être due à des différences méthodologiques, notamment en termes de types de disfluences analysés (répétitions de mots multisyllabiques, répétitions de phrases, interjections, révisions, énoncés interrompus dans l'étude de [11] vs. allongements, pauses remplies, marqueurs de discours, répétitions de sons, de syllabes, de mots et de phrases dans la présente étude).

Conformément aux travaux antérieurs portant sur l'anglais et le portugais brésilien ([18], [19]), les deux groupes de locuteurs ont manifesté plus de disfluences sur les mots de fonction que sur les mots de contenu. Cette tendance commune aux deux groupes de locuteurs peut être expliquée par le fait que la cause de production de ces disfluences « normales » est similaire dans les deux groupes (un retard au niveau de la planification du contenu linguistique). Ainsi, leur occurrence sur les mots de fonction permettrait au locuteur de maintenir son tour de parole pendant que la suite du message est en cours de planification.

L'identification du type de disfluences le plus récurrent chez les PQB et le PQBP a relevé une prédominance des pauses remplies dans les deux groupes de locuteurs. Ces résultats concordent avec ceux de [11] qui ont trouvé que les interjections (incluant les pauses remplies telles que « um »), étaient le type de disfluences le plus récurrent, dans des corpus de parole semi-spontanée des locuteurs qui bégayaient et qui ne bégayaient pas. Des résultats similaires ont été obtenus lorsque notre intérêt a porté sur l'identification du type de disfluence le plus fréquemment émis sur les mots de fonction (les pauses remplies étaient les plus récurrentes chez les PQB et les PQBP). En ce qui concerne les mots de contenu, les résultats obtenus chez les PQBP montrent que ce sont les allongements qui prédominent, alors que chez les PQB, ce sont les allongements et les marqueurs de discours.

L'analyse comparative intergroupe révèle une absence de différence significative pour les disfluences qui touchent : (1) les unités lexicales (toutes classes grammaticales confondues), et (2) les mots de fonction. En revanche, pour les disfluences touchant les mots de contenu, la comparaison intergroupe indique que seuls les marqueurs de discours présentaient une différence significative entre les groupes. Plus précisément, leur taux d'apparition est plus important chez les PQB que chez les PQBP. Ce résultat soutient l'idée

que les PQB recourent à des stratégies de compensation pour éviter la production des disfluences « pathologiques ».

6 Conclusion

La présente étude avait pour objectif de décrire les caractéristiques des disfluences « normales » produites par des locuteurs adultes francophones qui bégaièrent et qui ne bégaièrent pas. Bien que les PQB aient présenté un taux de disfluences significativement plus élevé que les PQBP, les analyses ont révélé des schémas de disfluences globalement comparables entre les deux groupes, avec des taux de disfluences plus importants sur les mots de fonction que sur les mots de contenu, ainsi que la prédominance des pauses remplies. Les seules différences significatives résident dans l'usage des marqueurs de discours. En outre, ces résultats soulignent la nécessité d'intégrer les stratégies individuelles et les variabilités inter-individuelles dans l'analyse. Les recherches futures devront donc explorer ces variabilités et l'influence des différents contextes de communication pour affiner notre compréhension des disfluences.

7 Références bibliographiques

- [1] Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84(1), 73-111. [https://doi.org/10.1016/S0010-0277\(02\)00017-3](https://doi.org/10.1016/S0010-0277(02)00017-3)
- [2] Clark, H. H. (2002). Speaking in time. *Speech Communication*, 36(1-2), 5-13. [https://doi.org/10.1016/S0167-6393\(01\)00022-X](https://doi.org/10.1016/S0167-6393(01)00022-X)
- [3] Starkweather, C.W. (1987). *Fluency and Stuttering*. N.J.: Prentice-Hall Inc, Englewood Cliffs.
- [4] Didirková, I. (2016). *Parole, langues et disfluences : Une étude linguistique et phonétique du bégaiement*. [Thèse de doctorat]. Université de Montpellier Paul-Valéry.
- [5] Menyhárt, K. (2003). Age-dependent types and frequency of disfluencies. *Proceedings of DiSS'03, Disfluency in Spontaneous Speech Workshop*, Sweden, 45-48.
- [6] Eklund, R. (2004). *Disfluency in Swedish human-human and human-machine travel booking dialogues*. [Thèse de doctorat]. Linköping University.
- [7] Lickley, R. (2017). Disfluency in typical and stuttered speech. In *Fattori sociali e biologici nella variazione fonetica (Social and biological factors in speech variation)* (Chiara Bertini, Chiara Celata, Giovanna Lenoci, Chiara Meluzzi, Irene Ricci, p. 373-387). Studi AISV.
- [8] Maekawa, K. (2004). Design, Compilation, and Some Preliminary Analyses of the Corpus of Spontaneous Japanese. *Spontaneous speech: Data and analysis*, 3, 87-108.
- [9] Roberts, P. M., Meltzer, A., & Wilding, J. (2009). Disfluencies in non-stuttering adults across sample lengths and topics. *Journal of Communication Disorders*, 42(6), 414-427. <https://doi.org/10.1016/j.jcomdis.2009.06.001>
- [10] Arida, L., & Didirková, I. (2025). The Impact of Speaking Task on Disfluency distribution in French : Reading Aloud, Picture Description, and Spontaneous Speech. *Proceedings of the 12th Workshop on Disfluency in Spontaneous Speech (DiSS 2025)*, Lisbon, Portugal.
- [11] Staróbole Juste, F., & Furquim de Andrade, C. R. (2010). Speech Disfluency Types of Fluent and Stuttering Individuals : Age Effects. *Folia Phoniatrica et Logopaedica*, 63(2), 57-64. <https://doi.org/10.1159/000319913>
- [12] Lutz, K. C., & Mallard, A. R. (1986). Disfluencies and rate of speech in young adult nonstutterers. *Journal of Fluency Disorders*, 11(4), 307-316. [https://doi.org/10.1016/0094-730X\(86\)90018-5](https://doi.org/10.1016/0094-730X(86)90018-5)

- [13] Kumaraswamy, S., Joseph, J., & S, R. (2022). Speech Disfluencies Among Adults During Class Presentation. *International Journal for Research in Applied Science and Engineering Technology*, 10(10), 241-248. <https://doi.org/10.22214/ijraset.2022.46972>
- [14] Altıparmak, A., & Kuruoğlu, G. (2018). Gender and Speech Disfluency Production : A Psycholinguistic Analysis on Turkish Speakers. *Psycholinguistics*, 24(2), 114-143. <https://doi.org/10.31470/2309-1797-2018-24-1-114-143>
- [15] Wu, Y., Didirková, I., & Simon, A.-C. (2022). Disfluences en parole continue en français: Paramètres prosodiques des pauses pleines et des allongements vocaliques. *XXXIVe Journées d'Études sur la Parole (JEP 2022)*. 900–909. <https://doi.org/10.21437/JEP.2022-95>.
- [16] Didirková, I., Wu, Y., & Simon, A.-C. (2024). Disfluences en parole continue en français : Paramètres prosodiques des répétitions. *JEP-TALN-RECITAL*, 33-41.
- [17] Prévot, L., Bertrand, R., & Rauzy, S. (2021). Investigating Disfluencies Contribution to Discourse-Prosody Mismatches in French Conversations. *Proceedings of the 10th Workshop on Disfluency in Spontaneous Speech (DiSS 2021)*, Saint-Denis, France.
- [18] Howell, P., Au-Yeung, J., & Sackin, S. (1999). Exchange of Stuttering From Function Words to Content Words With Age. *Journal of Speech, Language, and Hearing Research*, 42(2), 345-354. <https://doi.org/10.1044/jslhr.4202.345>
- [19] Juste, F. S., Sassi, F. C., & de Andrade, C. R. F. (2012). Exchange of disfluency with age from function to content words in Brazilian Portuguese speakers who do and do not stutter. *Clinical Linguistics & Phonetics*, 26(11-12), 946-961. <https://doi.org/10.3109/02699206.2012.728278>
- [20] 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.
- [21] Eklund, R. (2001). *Prolongations : A dark horse in the disfluency stable*. Workshop on Disfluency in Spontaneous Speech.
- [22] Lee, T.-L., He, Y.-F., Huang, Y.-J., Tseng, S.-C., & Eklund, R. (2004). *Prolongation in spontaneous Mandarin*. 2181-2184. <https://doi.org/10.21437/Interspeech.2004-239>
- [23] Schettino, L., & Eklund, R. (2023). Prolongation in Italian. *Proceedings of the Disfluency in Spontaneous Speech (DiSS 2023)*, Bielefeld, Germany, 81-85. <https://doi.org/10.21437/DiSS.2023-17>
- [24] Shriberg, E. (2001). To « errrr » is human : Ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, 31(1), 153-169.
- [25] American Psychiatric Association (ed) (2013). *Diagnostic and statistical manual of mental disorders* (5th Edition). Washington: American Psychiatric Publishing.
- [26] World Health Organization. (2024). *Clinical descriptions and diagnostic requirements for ICD-11 mental, behavioural and neurodevelopmental disorders (CDDR)*. <https://www.who.int/publications-detail-redirect/9789240077263>
- [27] Riley, G. (2009). *Stuttering Severity Instrument for Children and Adults*, Fourth Edition. Pro-Ed.
- [28] Crible, L., Didirková, I., Dodane, C., & Kosmala, L. (2022). Towards an inclusive system for the annotation of (dis)fluency in typical and atypical speech. *Clinical Linguistics & Phonetics*, 1-18. <https://doi.org/10.1080/02699206.2022.2126330>
- [29] Schmid H. (1995), Improvements in Part-of-Speech Tagging with an Application to German, *Proceedings of the ACL SIGDAT-Workshop*, Dublin, Ireland.
- [30] Abeillé, A., & Clément, L. (2015). *Annotation morpho-syntaxique—Les mots simples, les mots composés—Corpus Le Monde*. Université Paris 7.
- [31] IBM Corp. (2020). IBM SPSS Statistics for Windows (Version 27.0) [Computer software]. IBM Corp.