

UNIVERSITÉ SORBONNE NOUVELLE - PARIS III

Institut de Linguistique et de Phonétique Générales et Appliquées

Laboratoire de Phonétique et Phonologie (LPP) - UMR 7018



MÉMOIRE DE MASTER 2 RECHERCHE

Sciences du langage, spécialité « Phonétique et phonologie »

Présenté par François MOURET

Mesures rythmiques pour l'étude du bégaiement en français

Sous la direction de
Claire PILLOT-LOISEAU

Jury :

Nicolas AUDIBERT

Lise CREVIER-BUCHMAN

Cécile FOUGERON

Jacqueline VAISSIÈRE

ANNÉE UNIVERSITAIRE 2019-2020

DÉCLARATION SUR L'HONNEUR

Je soussigné François MOURET, né le 2 avril 1979, déclare avoir rédigé ce travail sans aides extérieures ni sources autres que celles qui sont citées. Toutes les utilisations de textes préexistants, publiés ou non, y compris en version électronique, sont signalées comme telles. Ce travail n'a été soumis à aucun autre jury d'examen sous une forme identique ou similaire, que ce soit en France ou à l'étranger, à l'université ou dans une autre institution, par moi-même ou par autrui.

Date : 31 août 2020

Signature :

François MOURET
signature électronique

RÉSUMÉ LONG

Objectifs

La présente étude vise à explorer au moyen de méthodes acoustiques les perturbations du rythme de la parole dans le bégaiement. Les recherches dans ce secteur ont surtout porté jusqu'à présent sur l'anglais et des langues germaniques apparentées, dites à « isochronie accentuelle » ou « fort degré d'accentuation » (Bergmann 1986). De plus, les travaux existants se limitent pour la plupart aux manifestations du rythme de la parole à l'échelle locale du segment ou de la syllabe, et ce en l'absence de disfluences (Maruthy *et al.* 2017). Par contraste, nous étudions les données du français, langue qui relève du type opposé à « isochronie syllabique » ou « faible degré d'accentuation ». Par ailleurs, nous considérons non seulement les séquences de parole perceptivement fluides, mais aussi celles qui ne le sont pas. Enfin et surtout, nous examinons le rythme dans ses manifestations locales, mais aussi supra-segmentales.

Méthodes

Notre étude repose sur les enregistrements de 5 sujets adultes présentant un bégaiement développemental persistant modéré à sévère lors d'une tâche de lecture d'un texte court, tous extraits du corpus « *Paroles disfluentes* » (Praxiling-UMR 5267 2017), que nous avons complétés par les enregistrements de 5 sujets contrôle appariés en âge et sexe. Nous utilisons deux types de mesures acoustiques, élaborées initialement pour contraster les langues selon leur degré d'accentuation ou « classe rythmique » :

- des mesures segmentales, fondées sur la durée des intervalles vocaliques et consonantiques (Ramus *et al.* 1999, Grabe & Low 2002) ;
- des mesures supra-segmentales, fondées sur la durée des groupes accentuels (Asu & Nolan 2006, Obin *et al.* 2012, Cichocki 2015).

Résultats

Sur le plan supra-segmental, les mesures considérées confortent l'hypothèse de Monfrais-Pfauwadel (2014) suivant laquelle les disfluences caractéristiques du bégaiement perturbent le phrasé en groupes accentuels ou « *chunking* » de la parole. Ainsi, l'eurythmie est typiquement rompue, se traduisant par une variabilité temporelle excessive des groupes accentuels, mais aussi par des violations des contraintes rythmiques qui gouvernent l'alternance entre battements faibles et battements forts (Martin 2014, 2017, 2018).

Sur le plan segmental, la parole des sujets qui bégaièrent se caractérise par une variabilité temporelle plus marquée des intervalles vocaliques et consonantiques, en accord avec les

résultats de Maruthy *et al.* (2017), et plus généralement avec les théories du bégaiement fondées sur l'hypothèse d'une instabilité motrice d'origine neurologique centrale (Zimmermann 1980, Packman & Onslow 1996). Les résultats diffèrent toutefois de ceux obtenus dans le domaine supra-segmental par plusieurs aspects. En premier lieu, la variabilité excessive constatée ne peut être directement imputée aux disfluences, comme c'est le cas au niveau des groupes accentuels : elle se maintient en effet en leur absence. D'autre part, le poids des mesures vocaliques et consonantiques n'est pas le même : alors que la variabilité des durées vocaliques distingue de manière générale les sujets qui bégaiement des sujets contrôle, la variabilité des durées consonantiques ne joue un rôle que chez certains sujets. Enfin et surtout, on peut se demander dans quelle mesure cette variabilité reflète véritablement une différence de nature rythmique entre les deux groupes étudiés.

REMERCIEMENTS

Ce travail n'aurait pu voir le jour sans les encouragements et le suivi minutieux de Claire Pillot-Loiseau : qu'elle soit ici remerciée chaleureusement. Il va de soi que toute erreur ou imprécision résiduelle dans le texte m'est entièrement imputable.

Merci par ailleurs aux membres du LPP de m'avoir accueilli « à distance » pendant deux ans, et tout particulièrement parmi eux à Nicolas Audibert, Lise Crevier-Buchman, Cécile Fougeron et Jacqueline Vaissière de me faire en outre l'honneur de participer au jury de soutenance. Jacqueline Vaissière y tient une place particulière pour m'avoir initié dès 2007 à la lecture de spectrogrammes, et avoir suivi avec un enthousiasme constant par la suite mon parcours de la syntaxe formelle à la phonétique expérimentale en passant par l'orthophonie.

Un grand merci à Fabrice Hirsch et Ivana Didirková qui ont rendu cette étude possible en me donnant accès au corpus « Paroles disfluentes », ainsi qu'à mes cinq « sujets contrôle » qui se reconnaîtront. Je remercie également Émilie Desportes et Marie Meyer, ainsi que Véronique Aumont-Boucand et Juliette de Chassey, pour m'avoir transmis en amont une quinzaine d'enregistrements, qui n'ont malheureusement pas pu être exploités.

Merci à Philippe Martin pour les discussions passionnantes que nous avons pu avoir sur l'accentuation et les ondes cérébrales, mais aussi pour la clarté de ses cours, qui m'ont permis de comprendre dans le détail le fonctionnement de la transformée de Fourier ainsi que divers aspects techniques du traitement du signal qui m'échappaient jusqu'alors.

Merci également à mes collègues linguistes de longue date Anne Abeillé, Olivier Bonami, Danièle Godard et Jean-Marie Marandin ; les échanges que nous avons pu avoir au fil des années, et notamment ceux des derniers mois dans le contexte de finalisation de la *Grande Grammaire du Français* ont sans nul doute façonné l'approche du rythme présentée dans cette étude.

Enfin, mes pensées vont comme toujours à ma femme, mes enfants et la famille étendue.

TABLE DES MATIÈRES

DÉCLARATION SUR L'HONNEUR	i
RÉSUMÉ LONG	ii
REMERCIEMENTS	iv
TABLE DES MATIÈRES	v
LISTE DES FIGURES ET TABLEAUX	vii
INTRODUCTION	1
CHAPITRE 1 : REPÈRES THÉORIQUES	2
1. Accentuation, rythme et tempo de la parole	2
1.1 L'accentuation.....	2
1.2 Le rythme.....	3
1.2.1 Dimensions du rythme dans les langues.....	3
1.2.2 Contraintes rythmiques liées à la structure accentuelle.....	5
1.2.3 Contraintes rythmiques liées à la structure syllabique	5
1.3 Le tempo	6
2. Généralités sur le bégaiement	7
2.1 Tableau clinique.....	7
2.1.1 Formes développementales.....	7
2.1.2 Formes acquises	11
2.2 Troubles apparentés	11
2.2.1 Le bredouillement ou « cluttering ».....	11
2.2.2 Les dysarthries et l'apraxie de la parole.....	12
2.3 Hypothèses causales et modélisation	13
2.4 Traitements	15
3. Revue de littérature.....	17
3.1 Bégaiement et perturbations segmentales	17
3.2 Bégaiement et perturbations supra-segmentales	17
CHAPITRE 2 : OBJECTIFS ET PROCÉDURE EXPÉRIMENTALE.....	19
1. Objectifs de l'étude.....	19
2. Procédure expérimentale.....	20
2.1 Participants	20

2.1.1 Présentation générale	20
2.1.2 Description des sujets pathologiques.....	21
2.1.3 Description des sujets contrôle.....	21
2.2 Matériel et protocole de passation	22
2.3 Mesures rythmiques	22
2.3.1 Mesures segmentales.....	22
2.3.2 Mesures supra-segmentales	25
2.4 Traitement des fichiers sonores.....	26
2.4.1 Analyse segmentale.....	26
2.4.2 Analyse supra-segmentale	29
2.4.3 Annotations complémentaires	31
2.5. Méthodes de calcul et d'analyse statistique	32
CHAPITRE 3 : RÉSULTATS ET DISCUSSION	33
1. Résultats segmentaux	33
1.1. Résultats avant élimination des disfluences	33
1.2. Résultats après élimination des disfluences	35
1.3. Comparaison des résultats par groupe de sujets.....	37
1.4 Discussion	38
2. Résultats supra-segmentaux	40
2.1. Résultats avant élimination des disfluences	40
2.2. Résultats après élimination des disfluences	42
2.3. Comparaison des résultats par groupe de sujets.....	44
2.4. Discussion	44
CONCLUSION	46
BIBLIOGRAPHIE.....	48
ANNEXES	61
1. Exemples d'analyses segmentales	61
2. Exemples d'analyses supra-segmentales	64
RÉSUMÉ COURT	70

LISTE DES FIGURES ET TABLEAUX

Fig. 1. Taxinomie des signes cliniques du bégaiement suivant Packman & Onslow (1998). ..	8
Fig. 2. Spectrogramme d'un blocage avant relâchement de la consonne occlusive /k/, sujet P5.	9
Fig. 3. Spectrogramme d'une prolongation sur la consonne fricative /v/, sujet P5.	9
Fig. 4. Spectrogramme de répétitions syllabiques de type bègue, affectant d'abord l'attaque consonantique de la syllabe /kɔ/, puis le noyau vocalique de la syllabe /oz/ sans la coda, sujet P2.	9
Fig. 5. Modèle à trois facteurs de Packman & Attanasio, extrait de Packman (2012 : 227)..	15
Fig. 6. Classifications des langues selon différentes mesures rythmiques d'après Ramus (2002).	24
Fig. 7. Segmentation maximale de la séquence « com- comment oses-tu me réveiller », sujet P5.	32
Fig. 8. Boîtes à moustaches illustrant la distribution des valeurs pour chaque mesure rythmique segmentale avant élimination des disfluences chez les sujets pathologiques et les sujets contrôle.	34
Fig. 9. Boîtes à moustaches illustrant la distribution des valeurs pour chaque mesure rythmique segmentale après élimination des disfluences chez les sujets pathologiques et les sujets contrôle.	36
Fig. 10. Classifications des sujets pathologiques et contrôle selon différentes mesures rythmiques segmentales appliquées à la parole avant élimination des disfluences.	39
Fig. 11. Classifications des sujets pathologiques et contrôle selon différentes mesures rythmiques segmentales appliquées à la parole après élimination des disfluences.	39
Fig. 12. Boîtes à moustaches illustrant la distribution des valeurs pour chaque mesure rythmique supra-segmentale avant élimination des disfluences chez les sujets pathologiques et les sujets contrôle.	41
Fig. 13. Boîtes à moustaches illustrant la distribution des valeurs pour chaque mesure rythmique supra-segmentale après élimination des disfluences chez les sujets pathologiques et les sujets contrôle.	43
Fig. 14. Oscillogramme, spectrogramme à bandes larges et courbes d'intensité (en jaune) et de F0 (en bleu) associés à la séquence « tu tu tu / tu tu me fait trop rire », produite par le sujet P2.	45

Tableau 1. Classification des dysarthries d'après Darley <i>et al.</i> (1969a, 1969b).....	13
Tableau 2. Critères d'inclusion et d'exclusion des participants de l'étude.....	20
Tableau 3. Échantillon de sujets pathologiques.....	21
Tableau 4. Échantillon de sujets contrôle.....	22
Tableau 5. Mesures rythmiques segmentales.....	23
Tableau 6. Mesures rythmiques supra-segmentales.....	25
Tableau 7. Critères acoustiques pour la segmentation en intervalles vocaliques et consonantiques.....	27
Tableau 8. Critères acoustiques pour la segmentation en groupes accentuels.....	29
Tableau 9. Valeurs des mesures rythmiques segmentales avant élimination des disfluences chez les sujets pathologiques (P), exprimées en millisecondes (ΔV , ΔC , r-PVI-V, r-PVI- C) ou en pourcentages (%V, VarcoV, VarcoC, n-PVI-V, n-PVI-C) arrondis à la deuxième décimale.....	33
Tableau 10. Valeurs des mesures rythmiques segmentales avant élimination des disfluences chez les sujets contrôle (C), exprimées en millisecondes (ΔV , ΔC , r-PVI-V, r-PVI-C) ou en pourcentages (%V, VarcoV, VarcoC, n-PVI-V, n-PVI-C) arrondis à la deuxième décimale.....	33
Tableau 11. Comparaison des valeurs obtenues par les sujets pathologiques et les sujets contrôle pour chaque mesure rythmique segmentale avant élimination des disfluences au moyen du test U bilatéral de Mann-Whitney.....	35
Tableau 12. Valeurs des mesures rythmiques segmentales après élimination des disfluences chez les sujets pathologiques (P), exprimées en millisecondes (ΔV , ΔC , r-PVI-V, r-PVI- C) ou en pourcentages (%V, VarcoV, VarcoC, n-PVI-V, n-PVI-C) arrondis à la deuxième décimale.....	35
Tableau 13. Valeurs des mesures rythmiques segmentales après élimination des disfluences chez les sujets contrôle (C), exprimées en millisecondes (ΔV , ΔC , r-PVI-V, r-PVI-C) ou en pourcentages (%V, VarcoV, VarcoC, n-PVI-V, n-PVI-C) arrondis à la deuxième décimale.....	36
Tableau 14. Comparaison des valeurs obtenues par les sujets pathologiques et les sujets contrôle pour chaque mesure rythmique segmentale après élimination des disfluences au moyen du test U bilatéral de Mann-Whitney.....	37
Tableau 15. Comparaison des valeurs obtenues par les sujets pathologiques pour chaque mesure rythmique segmentale avant et après élimination des disfluences au moyen du test T des rangs signés de Wilcoxon bilatéral pour échantillons appariés.....	37

Tableau 16. Valeurs des mesures rythmiques supra-segmentales avant élimination des disfluences chez les sujets pathologiques (P), exprimées en millisecondes (Δ GA, r-PVI-GA, Min(GA), Max(GA)) ou en pourcentages (VarcoGA, n-PVI-GA) arrondis à la deuxième décimale.....	40
Tableau 17. Valeurs des mesures rythmiques supra-segmentales avant élimination des disfluences chez les sujets contrôle (C), exprimées en millisecondes (Δ GA, r-PVI-GA, Min(GA), Max(GA)) ou en pourcentages (VarcoGA, n-PVI-GA) arrondis à la deuxième décimale.	40
Tableau 18. Comparaison des valeurs obtenues par les sujets pathologiques et les sujets contrôle pour chaque mesure rythmique supra-segmentale avant élimination des disfluences au moyen du test U bilatéral de Mann-Whitney.....	41
Tableau 19. Valeurs des mesures rythmiques supra-segmentales après élimination des disfluences chez les sujets pathologiques (P), exprimées en millisecondes (Δ GA, r-PVI-GA, Min(GA), Max(GA)) ou en pourcentages (VarcoGA, n-PVI-GA) arrondis à la deuxième décimale.....	42
Tableau 20. Valeurs des mesures rythmiques supra-segmentales après élimination des disfluences chez les sujets contrôle (C), exprimées en millisecondes (Δ GA, r-PVI-GA, Min(GA), Max(GA)) ou en pourcentages (VarcoGA, n-PVI-GA) arrondis à la deuxième décimale.	43
Tableau 21. Comparaison des valeurs obtenues par les sujets pathologiques et les sujets contrôle pour chaque mesure rythmique supra-segmentale après élimination des disfluences au moyen du test U bilatéral de Mann-Whitney.....	44
Tableau 22. Comparaison des valeurs obtenues par les sujets pathologiques pour chaque mesure rythmique supra-segmentale avant et après élimination des disfluences au moyen du test T des rangs signés de Wilcoxon bilatéral pour échantillons appariés. ...	44

INTRODUCTION

Le bégaiement est communément défini comme un trouble du rythme de la parole (World Health Organization 1977). Il se reconnaît en effet en premier lieu à la survenue de disfluences caractéristiques, accompagnées le cas échéant de comportements parasites, qui perturbent le déroulement normal de la parole dans le temps. De plus, il est bien connu qu'il disparaît ou du moins s'atténue sous certaines conditions qui font intervenir une modification du rythme ordinaire, comme le chant, ou encore la parole prolongée ou rythmée. Curieusement cependant, les recherches en phonétique ou phonologie qui s'attachent à décrire spécifiquement les perturbations rythmiques sont rares (Maruthy *et al.* 2017), les travaux ayant surtout porté jusqu'à présent sur la dimension segmentale (voir notamment Hirsch 2007, Pendelieu-Verdurand 2014 et Didirková 2016 pour le français).

La présente étude vise à explorer certains aspects du rythme de la parole dans le bégaiement au moyen de méthodes acoustiques. Plus précisément, nous souhaitons tester l'hypothèse selon laquelle les disfluences de type bégue perturbent non seulement la structure interne des syllabes, mais aussi le phrasé ordinaire en groupes accentuels ou « chunking » (Monfrais-Pfauwadel 2014). Nous utilisons pour ce faire différentes mesures rythmiques introduites initialement au niveau segmental pour contraster les langues selon leur degré d'accentuation (Ramus *et al.* 1999, Grabe & Low 2002), que nous étendons au niveau supra-segmental (Asu & Nolan 2006, Obin *et al.* 2012, Cichocki 2015).

Nous commençons par effectuer dans un premier chapitre une mise au point théorique sur les notions corrélées d'accentuation, de rythme et de tempo, puis le bégaiement de manière générale, avant de procéder à un brève revue de littérature sur les aspects phonétiques et phonologiques pertinents. Nous présentons ensuite plus en détail dans un deuxième chapitre nos objectifs et la procédure expérimentale retenue. Enfin, les résultats obtenus sont décrits et discutés dans le dernier chapitre avant de conclure.

CHAPITRE 1 : REPÈRES THÉORIQUES

1. Accentuation, rythme et tempo de la parole

1.1 L'accentuation

On parle dans le domaine prosodique d'accentuation, ou plus simplement d'accent, pour désigner non pas une façon de parler ou encore un signe graphique, mais la mise en valeur ou proéminence de certaines syllabes, perçues comme plus saillantes que d'autres dans la chaîne de parole (Delais-Roussarie & Di Cristo, à paraître). Suivant notamment Di Cristo (1999, 2011), on distingue entre les accents métriques, fixés par la phonologie, et les accents emphatiques, susceptibles de se superposer aux premiers avec une fonction sémantico-pragmatique (focus ou topique informationnel, soulignement contrastif ou marque émotive). Les accents métriques se présentent en outre sous deux formes : l'accent dit « primaire », obligatoire dans un domaine donné qu'il définit appelé ici « groupe accentuel », et l'accent « secondaire », qui n'est pas toujours réalisé.

Parmi les divers systèmes attestés dans les langues (Van der Hulst *et al.* 2010), le français se reconnaît à plusieurs caractéristiques générales. En premier lieu, l'accent métrique primaire a une position fixe prévisible : il frappe naturellement la dernière syllabe (ou alors la pénultième en cas d'insertion d'un schwa final) des mots de catégories ouvertes (noms, verbes, adjectifs, adverbes) ; plus difficilement celle, en général unique, des mots de catégories fermées (déterminants, prépositions, conjonctions, auxiliaires et pronoms clitiques), à moins que ceux-ci figurent immédiatement avant une rupture imposée par la structure, par exemple devant une incise (*Il a, paraît-il, réussi*). Par ailleurs, il n'est pas obligatoire (et parfois même impossible) d'accentuer chacun de ces mots dans un énoncé effectif ; on dit que toutes les syllabes accentuables ne sont pas nécessairement accentuées. Comme l'illustre Martin (2018 : 277), plusieurs découpages en groupes accentuels sont ainsi disponibles pour une séquence comme « la petite armoire violette » selon qu'on accentue les trois syllabes finales (/lapə'titaʁ'mwaʁvjɔ'let/), la première et la dernière (/lapə'titaʁmwaʁvjɔ'let/), la dernière et l'avant-dernière (/lapətitɑʁ'mwaʁvjɔ'let/) ou encore uniquement la dernière (/lapətitɑʁmwaʁvjɔ'let/). Enfin, le groupe accentuel peut sous certaines conditions être renforcé par un accent secondaire, qui tombe en général sur la syllabe initiale du premier mot lexical (Vaissière 1974, Fónagy 1980, Padeloup 1990, Astésano 2001).

Divers indices acoustiques contribuent selon les langues à l'identification des accents métriques ainsi décrits, parmi lesquels on peut citer la durée, mais aussi les variations de F0 ou encore l'intensité, auxquels s'ajoutent d'autres paramètres comme le timbre, ou encore la réalisation et la durée des pauses. En français, l'accent primaire du groupe accentuel est typiquement marqué par un allongement de la rime vocalique (Delattre 1966), et l'accent secondaire par un pic local d'intensité et/ou de F0 (Vaissière 1974, Fónagy 1980), suivant en cela la loi dite « iambique-trochaïque » qui associe la mise en valeur d'un battement par la durée à une frontière droite, et le recours aux variations tonales ou intensives à une frontière gauche (Fraisse 1974, Vaissière 1991). D'autres paramètres peuvent toutefois intervenir selon les réalisations (Astésano 2001). Du reste, la coïncidence avec les données perceptives demeure imparfaite. Comme le souligne Martin (2005, 2018), une syllabe peut être perçue comme accentuée alors qu'aucun indice acoustique ne le corrobore et inversement, ce que montrent bien les tests de perception sur des langues inconnues des sujets auxquels il est demandé d'identifier les proéminences, le taux d'erreur étant élevé dans ce cas (Mettouchi *et al.* 2007).

1.2 Le rythme

La notion de rythme est notoirement difficile à caractériser au-delà de « l'ordre dans les mouvements » décrit par Platon. De manière générale, on considère que la structuration naît d'une forme de périodicité, c'est-à-dire du retour à intervalles réguliers d'un événement distingué dans la succession (Fraisse 1956). Étant admis que la parole constitue un phénomène rythmique, on peut se demander à la fois quelle unité sonore sert d'événement distingué, et quelle forme de périodicité est mise en jeu. Nous allons rappeler brièvement les débats sur le sujet, avant de présenter les contraintes rythmiques communément admises dans la parole.

1.2.1 Dimensions du rythme dans les langues

Historiquement, la question du rythme de la parole a d'abord été considérée sous l'angle de la comparaison des langues. À partir des observations de Llyod James (1940) qui distinguait entre deux modèles rythmiques, l'un comparable à un code morse (« *morse-code rhythm* ») dans des langues comme l'anglais, l'autre comparable aux tirs d'une mitraillette (« *machine-gun rhythm* ») dans des langues comme l'espagnol, on a pu penser que certaines langues favorisaient le retour à intervalles réguliers de l'accent, la durée des syllabes au sein de chaque groupe accentuel pouvant, elle, varier, tandis que d'autres favorisaient le retour à intervalles réguliers de la syllabe. Pike (1945), à qui l'on doit les formulations « *stress timing* » ou « isochronie accentuelle » et « *syllable timing* » ou « isochronie syllabique »,

concevait ces deux modèles comme des tendances, susceptibles de coexister au sein d'une même langue, comme le rappelle Cumming (2010). Il faut attendre la monographie d'Abercrombie (1967) pour qu'émerge l'hypothèse d'une typologie plus rigide ; des langues comme l'anglais, l'allemand ou encore le russe illustrant ainsi l'isochronie accentuelle, des langues comme l'espagnol ou le français, l'isochronie syllabique, et des langues comme le japonais, une troisième classe fondée sur la périodicité de la more, unité inférieure à la syllabe (Ladefoged 1975, Port et al. 1987).

Dès lors, diverses études se sont attachées à vérifier sans succès la réalité acoustique de ces distinctions. Dès les années 60, Bolinger montrait, suivant en cela des travaux plus anciens de Classe (1939), que la durée des groupes accentuels dépend du nombre et du type des syllabes qu'ils contiennent dans une langue à isochronie accentuelle supposée comme l'anglais (Bolinger 1965). Parallèlement, Delattre observait que le contraste de durée entre syllabes non accentuées et syllabes accentuées en position non finale d'énoncé est moins marqué dans cette même langue qu'il ne l'est en français, langue supposée illustrer l'isochronie syllabique (Delattre 1966). De nombreux articles sont parus par la suite, parmi lesquels on peut citer ceux de Roach (1982) et Dauer (1983) sur diverses langues, ou encore, pour ce qui concerne plus spécifiquement le français, ceux de Wenk & Wioland (1982) et Duez & Nishinuma (1985), conduisant à abandonner à la fois l'hypothèse d'une unité rythmique de base différente selon les langues (la more, la syllabe ou le groupe accentuel), et celle d'une périodicité acoustique stricte.

Les recherches prennent une nouvelle direction à partir des travaux de Dasher & Bolinger (1982), Dauer (1983, 1987) puis Fant *et al.* (1991), qui proposent de dériver les contrastes observés de l'interaction entre deux phénomènes :

- une périodicité accentuelle de nature acoustico-perceptive, qui serait commune à toutes les langues, rejoignant ainsi la conception du rythme prônée par la phonologie métrique (Liberman & Prince 1977, Hayes 1984), mais aussi les études menées en psychoacoustique (Fraisse 1956, 1974) ;
- la combinaison variable selon les langues de certaines propriétés phonéto-phonologiques comme la complexité syllabique, la qualité vocalique ou encore certains corrélats acoustiques de l'accent (variations de F0, durée, intensité), contribuant à mettre en valeur le groupe accentuel sur le plan perceptif, ou au contraire à l'atténuer au profit de la more ou de la syllabe.

On distingue ainsi aujourd'hui entre deux dimensions du rythme dans la parole, l'une, de nature supra-segmentale, fondée sur la structure accentuelle, l'autre, de nature plus locale, fondée sur la structure interne de la syllabe, que nous allons à présent développer.

1.2.2 Contraintes rythmiques liées à la structure accentuelle

Deux contraintes générales sont communément admises sur le plan accentuel, visant à favoriser l'eurythmie, c'est l'alternance optimale entre battements faibles et battements forts (Di Cristo 2003). D'une part, la distance entre deux accents métriques ne doit être ni trop longue (« *no lapse* »), ni trop courte (« *no clash* ») (Selkirk 1984, Dell 1984, Nespor & Vogel 1989), ce qui découle en fait, comme l'a montré récemment Martin (2014, 2017, 2018), de la gamme de variation de certaines ondes cérébrales. Ainsi, un groupe accentuel ne peut être inférieur à 250 ms, ni supérieur à 1250 ms, ce qui coïncide avec les périodes minimale et maximale des ondes Delta. On explique ainsi les phénomènes de « déplacement » ou de « suppression » d'accent dans les configurations de collision potentielle où deux syllabes accentuables se suivent. Une séquence comme « un joli lit » ne peut recevoir la représentation /ẽʒo'li'li/ que dans le cas où la distance entre les deux syllabes finales est supérieure à 250 ms; dans le cas contraire, le premier accent ne sera pas réalisé (/ẽʒoli'li/) ou sera « déplacé », c'est-à-dire en réalité remplacé par un accent secondaire sur la syllabe initiale de l'adjectif : /ẽ,ʒoli'li/. On comprend par ailleurs que la fameuse « règle des 7 syllabes » de Meigret (1550), supposée fixer la taille maximale du groupe accentuel, ne soit correcte qu'à débit articulatoire moyen. À débit rapide en effet, le nombre de syllabes peut être plus élevé, la durée étant alors comprimée de façon à ne pas excéder 1250 ms.

D'autre part, les groupes accentuels successifs doivent être de durées comparables. Il est vrai, comme on l'a vu précédemment, que le retour de l'accent dans la chaîne de parole n'est pas strictement périodique sur le plan acoustique. Toutefois, les différences de durée constatées entre groupes accentuels successifs sont le plus souvent inférieures au seuil de perception, de l'ordre de 100 ms (Lehiste 1977). De plus, une durée moyenne comparable se retrouve dans la plupart des langues étudiées, avoisinant la période de référence de l'écoute musicale, qui est d'environ 600 ms (Di Cristo 2003).

1.2.3 Contraintes rythmiques liées à la structure syllabique

Comme le notent Nespor *et al.* (2011), l'approche métrique est insuffisante pour caractériser les classes rythmiques proposées par les linguistes, et effectivement discriminées par les nouveau-nés (du moins, pour ce qui concerne des langues prototypiques, Mehler *et al.* 1988, Nazzi *et al.* 1998). Il faut pour cela considérer la structure interne de la syllabe. Parmi les propriétés paramétriques recensées par Dauer (1987), ce sont surtout la complexité syllabique et la qualité vocalique qui ont retenu l'attention des chercheurs à partir des travaux de Ramus *et al.* (1999) et de Grabe & Low (2002), conduisant à distinguer trois types de combinatoires :

- les langues à fort degré d'accentuation, caractérisées par un riche inventaire de structures syllabiques, et une tendance à la réduction (par centralisation articulaire et compression de la durée) des voyelles non accentuées, comme l'anglais ou l'allemand ;
- les langues à faible degré d'accentuation, caractérisées par des structures syllabiques moins variées (comme en français), voire extrêmement simples (comme en japonais), et l'absence de réduction vocalique ;
- les langues intermédiaires (non prévues par les classifications initiales de Pike 1945, et Abercrombie 1967), caractérisées par des syllabes simples, mais l'existence de réduction vocalique en position non accentuée, comme le catalan, ou à l'inverse des structures syllabiques plus riches, mais une qualité vocalique stable, comme le polonais (Nespor 1990).

Sans reprendre ici tous les débats qui ont suivi (voir Mairano 2010 pour une revue), on retiendra que cette approche se heurte à plusieurs problèmes (Arvaniti 2009, 2012), parmi lesquels le manque d'études perceptives susceptibles d'étayer les distinctions proposées, l'insuffisance des critères retenus pour caractériser les langues dans leur diversité, ou encore le recours exclusif au paramètre acoustique de la durée dans les modélisations proposées (voir chapitre 2, section 2.3.1, et les remarques à ce sujet en conclusion de ce mémoire). Elle reste néanmoins communément admise dans le paysage théorique actuel, avec des résultats cohérents d'une étude à l'autre pour ce qui concerne une langue comme le français. De plus, elle a donné lieu récemment à plusieurs applications dans le champ clinique des dysarthries (Liss *et al.* 2009, Lowit *et al.* 2018), de l'apraxie de la parole (Ballard *et al.* 2016), mais aussi et surtout du bégaiement qui nous intéresse ici (Maruthy *et al.* 2017), ce qui justifie d'en explorer malgré tout le potentiel.

1.3 Le tempo

Le tempo correspond en musique à la vitesse d'exécution du rythme de la partition, exprimé en battements par minute. Dans le domaine de la parole, la notion recouvre celle de débit, qui est nettement plus complexe. En premier lieu, il faut distinguer différents types de mesures. Classiquement, on contraste le débit locutoire exprimé en syllabes par seconde, qui inclut les pauses silencieuses, et le débit articulaire, qui les exclut (Grosjean & Deschamps 1975). La plupart des études sur la parole disfluente opposent par ailleurs un débit articulaire brut, qui inclut les disfluences, et un débit articulaire « normalisé » sans celles-ci (Yaruss *et al.* 1994, van Zaalen 2009). D'autre part, la variation est importante d'une langue à l'autre, mais aussi au sein d'une même langue selon les locuteurs et les situations

(Miller *et al.* 1984). Enfin et surtout, la nature du débit a un effet direct sur le rythme, contrairement à ce qui se passe en musique : plus le débit est rapide, moins il y a d'accents (Zellner 1998, Padeloup 2004).

Nous procédons dans ce qui suit à un certain nombre de rappels sur le bégaiement, avant de considérer les perturbations phonétiques et phonologiques établies dans la littérature, et plus particulièrement celles affectant les paramètres décrits dans cette section.

2. Généralités sur le bégaiement

L'OMS définit le bégaiement comme « un trouble du rythme de la parole, dans lequel le sujet sait exactement ce qu'il veut dire, mais ne peut pas le dire sur le moment en raison de la répétition, du blocage ou de la prolongation involontaire d'un son donné » (World Health Organization 1977: 202, notre traduction¹). Comme l'ont noté de nombreux auteurs, cette caractérisation est insuffisante. D'une part, le trouble peut être développemental ou acquis, avec des propriétés sensiblement différentes. Par ailleurs, il ne se limite pas aux manifestations évoquées. Nous allons reprendre en détail le tableau clinique communément admis, avant d'évoquer plus brièvement les troubles apparentés, les hypothèses causales et leur modélisation, et enfin les traitements disponibles.

2.1 Tableau clinique

2.1.1 Formes développementales

Le bégaiement développemental survient de manière soudaine ou progressive chez environ 5% des enfants à un âge compris entre 2 et 6 ans en moyenne. La prévalence, toutefois, est estimée à 1% de la population mondiale (adultes inclus, 3 hommes pour 1 femme), ce qui tient au fait que le trouble régresse avec ou sans prise en charge chez environ 75% des enfants qui en sont affectés (Bloodstein & Bernstein-Ratner 2008). En l'état actuel des connaissances, rien ne permet de distinguer entre les cas transitoires et les cas persistants dans la tranche d'âge d'apparition.

Signes cliniques

Suivant la terminologie de Packman & Onslow (1998), le trouble se reconnaît à la combinaison variable de trois types de manifestations lors de la production de la parole : (i)

¹ « *Disorders in the speech rhythm, in which the individual knows precisely what he wishes to say, but at the time is unable to say it because of an involuntary repetition, prolongation or cessation of a sound.* »

des postures fixes, (ii) des mouvements répétés et (iii) des comportements superflus (figure 1).

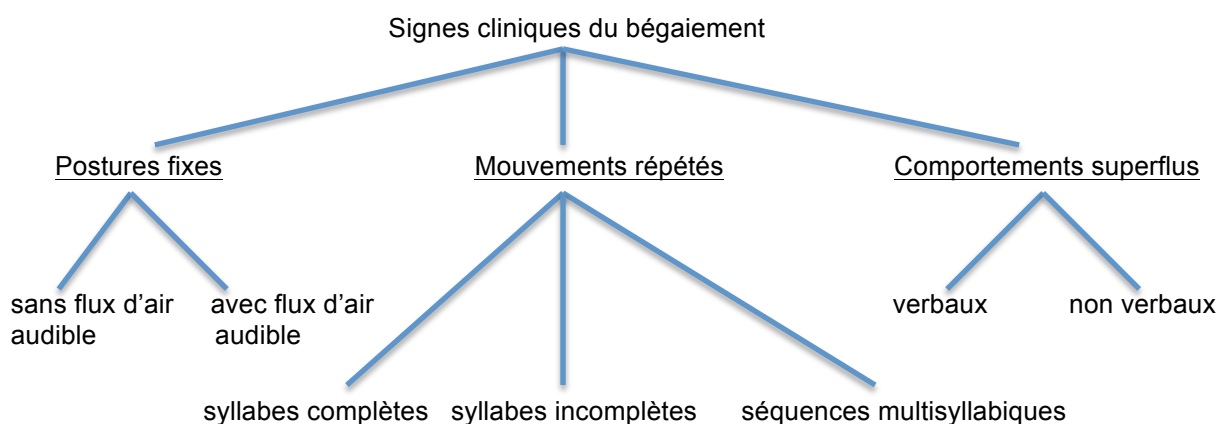


Fig. 1. Taxinomie des signes cliniques du bégaiement suivant Packman & Onslow (1998).

Les postures fixes et mouvements répétés recouvrent ce que l'on appelle plus communément les disfluences de type bègue ou « bégayages » (« *stuttering-like disfluencies* ») sur un continuum à l'opposé duquel on trouve les disfluences dites « normales » que constituent dans la parole spontanée les interjections, les allongements d'hésitation, les corrections et les répétitions de mots complets ou de séquences de mots. Il peut s'agir de blocages avant une voyelle ou une consonne occlusive (postures fixes sans flux d'air audible), de prolongations de voyelles ou de consonnes fricatives ou nasales (postures fixes avec flux d'air audible), ou encore de répétitions syllabiques. Parmi ces dernières, celles affectant des syllabes incomplètes sont parfois considérées comme pathognomoniques du bégaiement ; elles peuvent en effet générer des ruptures inhabituelles, notamment entre l'attaque et la rime (Wingate 1988, Monfrais-Pfauwadel & Teitler 1996). Les répétitions de syllabes complètes ou de séquences multisyllabiques, en revanche, peuvent se confondre avec des répétitions ordinaires de mots. On suppose généralement qu'il s'agit de bégayages lorsqu'une tension articulatoire excessive leur est associée, ou encore que le nombre d'éléments répétés est supérieur ou égal à 3 (Campbell & Hill 1994). Nous illustrons dans les figures 2, 3 et 4 les disfluences de type bègue les plus communes à partir d'extraits du corpus soumis à l'étude (voir chapitre 2, section 2).

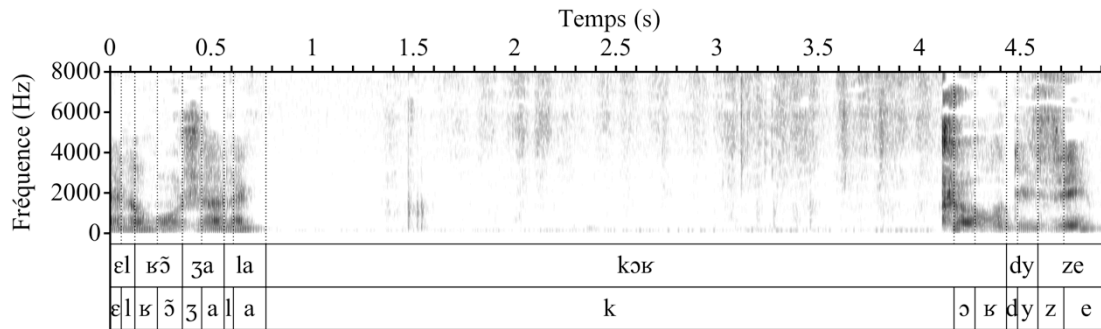


Fig. 2. Spectrogramme d'un blocage avant relâchement de la consonne occlusive /k/, sujet P5.

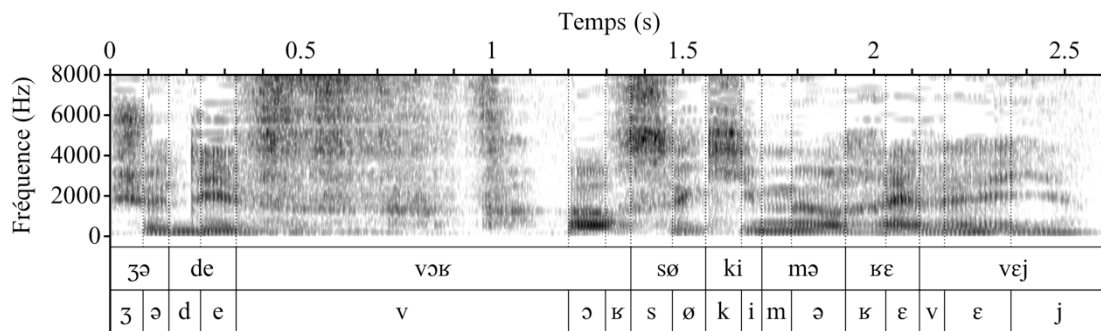


Fig. 3. Spectrogramme d'une prolongation sur la consonne fricative /v/, sujet P5.

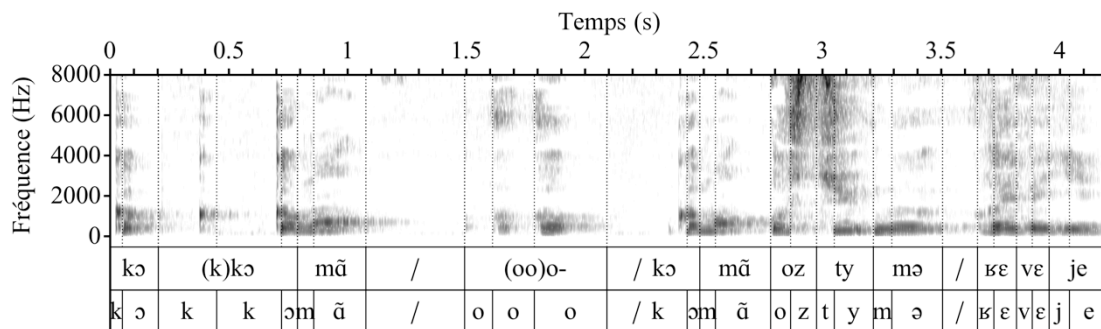


Fig. 4. Spectrogramme de répétitions syllabiques de type bègue, affectant d'abord l'attaque consonantique de la syllabe /kɔ̃/, puis le noyau vocalique de la syllabe /oz/ sans la coda, sujet P2.

Les comportements superflus correspondent par ailleurs à des manifestations qui parasitent l'énoncé en cours de production. Il peut s'agir d'interjections (« euh », « mmm », etc.) et expressions d'appui apparentées (« ben », « fin », etc.), qui constituent des disfluences normales, parfois appelées « pauses remplies », ou alors de comportements non verbaux variés, interprétables selon les cas comme des diffusions de tensions ou des manœuvres de désengagement : mouvements de la tête ou des membres, crispations de la face, clonies palpébrales, dilatation des narines, etc. À la différence des tics, ces manifestations motrices

se rencontrent uniquement lors de la production de la parole. Nous renvoyons le lecteur aux nombreux témoignages de personnes qui bégayaient disponibles en ligne sur Internet pour des illustrations.

Crucialement, ces signes cliniques ne surviennent pas systématiquement dans la parole du sujet qui bégaye. D'une part, le trouble est fluctuant selon divers facteurs comme le matériel linguistique et phonétique, la situation de parole ou encore l'état du sujet (voir Onslow 2020, chapitres 1 et 2 pour une revue fouillée). On parle ainsi de « moments de bégaiement » en dehors desquels la parole est jugée normale par les auditeurs, même si des anomalies peuvent être détectées sur les plans acoustique et articulatoire (voir *infra*, section 3.1). D'autre part, il est bien connu que certaines conditions inhibent temporairement le bégaiement, ou du moins l'atténuent, parmi lesquelles (i) le chant, (ii) la parole à l'unisson avec un autre locuteur ou sa propre voix par feedback auditif, phénomène connu sous le nom d'effet choral, (iii) la parole produite en écho, phénomène connu sous le nom d'effet de masque (« *shadowing* »), (iv) la parole prolongée, ou encore (v) la parole rythmée (Wingate 1969, Andrews *et al.* 1982). Clairement, ces conditions font toutes intervenir la dimension rythmique, par renforcement de l'alternance entre temps forts et temps faibles, comme c'est le cas dans le chant (Falk *et al.* 2016), et sans doute également dans la parole à l'unisson ou en écho ; ou au contraire par atténuation des contrastes, comme c'est le cas dans la parole prolongée ou rythmée à l'image d'une cadence, où les durées syllabiques tendent à être égalisées (Packman *et al.* 1994, 1996).

Symptômes associés

Suivant l'usage médical, on parle de symptômes pour désigner ce qui est ressenti par les sujets qui bégayaient. De manière générale, on distingue entre des symptômes somatiques comme des sensations de tensions musculaires susceptibles d'affecter tous les organes de la parole, notamment le diaphragme, le larynx, les organes mobiles de la cavité orale ou les muscles de la face d'une part, et divers symptômes psychiques typiques de l'anxiété sociale d'autre part, parmi lesquels la peur de bégayer, avec des conduites d'évitement de mots ou de situations associées, ou encore une faible estime de soi en lien avec des pensées et croyances négatives. Les symptômes somatiques surviennent essentiellement lors des moments de bégaiement, et ce quel que soit l'âge (Bloodstein & Bernstein-Ratner 2008), tandis que les symptômes psychiques se développent progressivement durant l'enfance et l'adolescence lorsque le trouble perdure, avec un impact sur la qualité de vie qui peut être majeur (Craig 2010).

2.1.2 Formes acquises

Le bégaiement acquis se présente sous trois formes (Van Borsel 2014), qui ont été globalement moins étudiées :

- le bégaiement pharmaco-induit, lié à la prise de certains médicaments, parmi lesquels divers antipsychotiques, antidépresseurs et anticonvulsivants ou encore certains broncho-dilatateurs (Brady 1998, Béné *et al.* 2014) ;
- le bégaiement neurogène, lié à des affections neurologiques variées comme l'épilepsie, un accident vasculaire cérébral, un traumatisme crânien, une tumeur ou une infection cérébrale, ou encore une maladie neuro-dégénérative ou génétique (Cruz *et al.* 2018) ;
- le bégaiement psychogène (Almada *et al.* 2016).

Contrairement à ce que l'on observe dans le bégaiement développemental, les disfluences se maintiendraient davantage dans les conditions qui les inhibent en principe comme le chant ou la parole à l'unisson. Elles seraient également moins fluctuantes selon les situations, moins sélectives quant au matériel linguistique et phonétique qu'elles affectent, et rarement accompagnées de comportements superflus moteurs, ou encore de symptômes anxieux relatifs à la parole.

2.2 Troubles apparentés

Le bégaiement dans ses différentes formes est communément classé parmi les troubles de la « fluence » ou « fluidité » (Starkweather 1987), notion particulièrement large si l'on entend par là « *l'aspect de la production de la parole qui renvoie à la continuité, la douceur, le débit et l'effort* » (ASHA 2015, notre traduction²). À ce titre, il s'apparente à d'autres pathologies, que nous allons présenter brièvement.

2.2.1 Le bredouillement ou « cluttering »

Le bredouillement constitue un trouble encore mal connu, notamment en français où les recherches se limitent pour l'heure à quelques mémoires de fin d'études en orthophonie (Dhuisme 2007, Desportes & Meyer 2014). Comme le note Aumont-Boucand (2012), le terme anglais « *cluttering* » (encombrement, désordre) illustre mieux ses caractéristiques, qui associent selon la définition consensuelle de St Louis & Schulte (2011), un débit perçu comme trop rapide, trop irrégulier ou les deux par moments à au moins un des signes suivants : (i) un nombre excessif de disfluences dites « normales » par opposition aux

² « Fluency is the aspect of speech production that refers to continuity, smoothness, rate and effort ».

disfluences de type bégue (voir *supra*, section 2.1.1), (ii) un nombre excessif de réductions et/ou d'élisions de syllabes dans les mots et expressions multisyllabiques, (iii) des perturbations dans l'accentuation, la distribution des pauses et/ou le rythme de la parole.

À la différence du bégaiement auquel il est souvent associé, le bredouillement est généralement mal identifié par les personnes qui en sont affectées. Par ailleurs, il se manifeste plus fréquemment dans les situations où le sujet exerce peu de contrôle sur son discours comme les conversations de la sphère privée, que dans les situations plus formelles où l'attention est centrée sur la parole et où le contenu linguistique est programmé d'avance, comme les discours publics, ou encore la lecture à voix haute.

Typiquement, le trouble est développemental : il survient au cours de l'acquisition du langage, *a priori* plus tardivement que le bégaiement, avec une composante génétique et un probable soubassement neurologique (van Zaalen 2009, Ward *et al.* 2015). Plus rarement, il est acquis dans le contexte d'une pathologie neurologique (Thacker & De Nil 1996).

2.2.2 Les dysarthries et l'apraxie de la parole

Les dysarthries constituent des troubles acquis de la production de la parole consécutifs à une atteinte du système nerveux central et/ou périphérique. On distingue classiquement entre 6 formes, recensées dans le tableau 1 (Darley *et al.* 1969a, 1969b). Parmi celles-ci, ce sont surtout les dysarthries résultant d'une atteinte des noyaux gris centraux ou de la région cérébelleuse qui peuvent s'apparenter au bégaiement. Dans certains cas, comme celui des syndromes parkinsoniens, la distinction est relativement aisée : les disfluences se présentent essentiellement sous la forme de répétitions de syllabes, notamment en cas de palilalies, sans tensions particulières ni ruptures entre l'attaque et la rime (Monfrais-Pfauwadel 2014). La frontière entre bégaiement acquis neurogène et dysarthrie peut en revanche être plus floue dans les cas de dystonies ou d'atteintes cérébelleuses.

Types	Localisations des lésions	Exemples de causes
Dysarthrie spastique	Premier motoneurone Faisceau pyramidal	AVC avec atteinte pseudo-bulbaire
Dysarthrie flasque	Atteinte des nerfs périphériques, des jonctions neuro-musculaires ou des muscles effecteurs	AVC avec atteinte bulbaire, myasthénie
Dysarthrie hypokinétique	Noyaux gris centraux	Maladie de Parkinson idiopathique et syndromes parkinsoniens apparentés
Dysarthrie hyperkinétique		Dystonies laryngées, Chorée de Huntington
Dysarthrie ataxique	Cervelet et voies cérébelleuses	AVC cérébelleux, maladie de Friedreich
Dysarthrie mixte	Lésions disséminées	Sclérose latérale amyotrophique, sclérose en plaque, traumatismes crâniens

Tableau 1. Classification des dysarthries d'après Darley *et al.* (1969a, 1969b).

L'apraxie de la parole ou « anarthrie », constitue également un trouble de la parole d'origine neurologique susceptible de perturber la fluence (McNeil *et al.* 2009). Elle se distingue en principe des dysarthries et du bégaiement en ce qu'elle affecte non pas le système effecteur ou la programmation, mais la conceptualisation des gestes articulatoires en amont, d'où une dissociation dite « automatico-volontaire » entre le discours spontané, qui est perturbé, et les formules (comptage, jours de la semaine, etc.), où la parole est en général mieux préservée. En pratique, le diagnostic différentiel peut toutefois être délicat, d'autant que le trouble est fréquemment associé à une aphasie.

2.3 Hypothèses causales et modélisation

À l'hypothèse d'une cause unique du bégaiement de nature psychologique ou encore organique explorée sans succès par le passé s'est substituée celle d'une étiologie multifactorielle, dont le soubassement serait à la fois génétique et neurologique. On sait d'abord qu'il existe une prédisposition héréditaire au bégaiement. Il est clair toutefois, ainsi que le rappelle Monfrais-Pfauwadel (2014 : 10), qu'il s'agit d'un mode de transmission complexe : le bégaiement est certes plus fréquent chez les jumeaux monozygotes que chez les jumeaux dizygotes, mais pas systématique, l'un pouvant présenter le trouble et l'autre non. Plus récemment, des mutations ont été identifiées sur quatre gènes, GNPTAB, GNPTG, NAGPA et AP4E1 (Kraft & Yairi 2012), qui interviendraient également dans d'autres pathologies qui peuvent affecter la fluence de la parole comme la maladie de Parkinson ou la maladie de Huntington. Comme le notent cependant Frigerio-Domingues & Drayna (2017), ces anomalies ne sont retrouvées que chez 20% des sujets adultes présentant un bégaiement développemental persistant.

L'existence de cas de bégaiement acquis neurogène ou pharmaco-induit implique par ailleurs un rôle essentiel de la composante neurologique. De nombreuses études de neuro-imagerie ont ainsi mis en évidence des spécificités cérébrales structurelles et fonctionnelles dans le bégaiement développemental, parmi lesquelles les suivantes chez l'adulte, selon la synthèse récente d'Etchell *et al.* (2017) fondée sur la revue d'une centaine d'articles parus entre 1995 et 2016 :

- une excitabilité réduite avant production de la parole dans les aires motrices ;
- une activation neuronale atypique dans les régions du gyrus frontal inférieur gauche et des aires auditives droites durant la parole ;
- des anomalies structurelles au niveau du gyrus frontal inférieur gauche et de ses connexions avec l'aire motrice supplémentaire, la région temporale et le gyrus frontal inférieur droit ;
- un gyrus temporal supérieur plus développé à droite qu'à gauche, et l'absence d'asymétrie des régions auditives droites et gauches en faveur de celles de l'hémisphère dominant.

Il reste toutefois difficile en l'état actuel des connaissances de distinguer entre les particularités qui favorisent le trouble et celles qui en découlent, qu'il s'agisse de dommages ou de modifications compensatoires. C'est pourquoi les recherches contemporaines dans ce secteur se concentrent depuis une dizaine d'années sur la population pédiatrique (Chang *et al.* 2008, 2015).

Du reste, ni les anomalies neurologiques, ni la composante génétique en amont ne suffisent à elles seules pour expliquer ce qui génère à un temps t les disfluences et comportements superflus associés, c'est-à-dire un moment de bégaiement dans les formes développementales. Il faut pour cela combiner plusieurs paramètres, comme le proposent Packman et Attanasio dans le cadre de leur modèle à trois facteurs (Packman & Attanasio 2004, Packman 2012), le plus abouti à l'heure actuelle pour rendre compte des traits essentiels du bégaiement développemental.

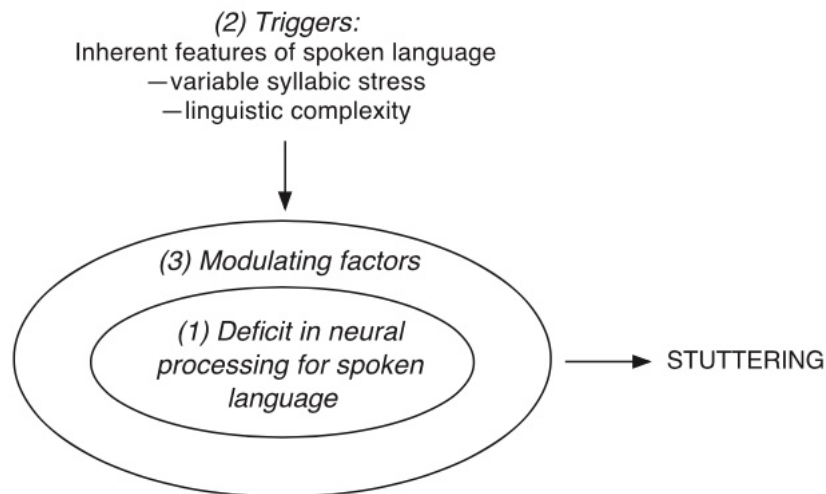


Fig. 5. Modèle à trois facteurs de Packman & Attanasio, extrait de Packman (2012 : 227).

Comme l'illustre la figure 5, le facteur central (1) correspond à un déficit neural dans le traitement du langage oral, qui fragiliserait le système de production de la parole. Le facteur secondaire (2) correspond par ailleurs à deux propriétés linguistiques : la nécessité d'alterner entre syllabes faibles et syllabes fortes ou variabilité de l'accentuation syllabique, et la complexité linguistique, qui fonctionneraient comme des déclencheurs, en élevant le niveau d'exigences motrices sur le système de production de la parole. Enfin, un ensemble de facteurs de modulations (3), propre à chaque individu, fixerait le seuil à partir duquel s'activent les déclencheurs et survient par conséquent un moment de bégaiement, comme l'excitation physiologique (« *physiological arousal* »), ou encore la disponibilité des ressources cognitives ou « charge cognitive » dans une situation donnée. On explique ainsi (i) l'âge de survenue des premiers moments de bégaiement chez l'enfant, qui coïncide avec la période entre 2 et 6 ans où se développent les structures syntaxiques et prosodiques, c'est-à-dire la complexité linguistique et la variabilité accentuelle, (ii) l'effet inhibiteur de diverses conditions comme le chant, la parole rythmée ou prolongée ou encore la parole à l'unisson qui font intervenir la dimension accentuelle, plausiblement interprétables comme des stabilisateurs du système moteur, (iii) la variabilité des moments de bégaiement, liée aux facteurs de modulations. Enfin, le présent modèle permet une analyse cohérente des divers traitements selon le facteur ciblé, comme nous allons le voir à présent.

2.4 Traitements

Les traitements du bégaiement pour lesquels il existe des essais cliniques sont communément regroupés en trois familles : (i) les thérapies orthophoniques directes, qui reposent sur une modification de la parole, (ii) les thérapies orthophoniques indirectes, qui

portent sur les facteurs environnementaux et individuels, et (iii) les traitements autres, basés sur la médication, ou encore une thérapie psychologique. Suivant Packman (2012), on peut considérer que les thérapies directes ciblent les déclencheurs du bégaiement, tandis que les autres ciblent davantage les facteurs de modulation, sauf peut-être certains traitements médicamenteux en lien avec le facteur neurologique central.

Parmi les approches directes récentes, on peut citer les programmes australiens Camperdown, Westmead et Lidcombe basés respectivement sur la parole prolongée, la parole rythmée et le recours au conditionnement opérant *via* des commentaires sur la parole (appelés « contingences verbales »), où les recherches sont les plus actives actuellement (voir Onslow 2020, chapitres 6 à 9 pour une revue) ; mais aussi le programme MPI (« *Modifying Phonation Intervals* ») développé aux États-Unis par l'équipe d'Ingham, qui repose sur la réduction des intervalles voisés brefs dans la parole *via* un système de bio-feedback (Ingham *et al.* 2015), ou encore les approches exploitant le feedback auditif (Hudock & Kalinowski 2014), apporté le cas échéant par un dispositif comparable à des prothèses auditives tel que *Speech Easy*®.

Les approches indirectes découlent, quant à elles, du modèle des demandes et capacités développé à partir de la fin des années 80 par Starkweather, suivant lequel le bégaiement surviendrait chez l'enfant lorsque la fluence nécessitée par l'environnement ou des représentations internes du sujet excède ses capacités motrices, cognitives ou psychosociales du moment (Starkweather & Givens-Ackerman 1997). La thérapie consiste dès lors à remodeler les interactions entre l'enfant et son entourage, comme cela est mis en œuvre dans des programmes comme le PCI (« *Parents-Child Interactions* ») élaboré au centre londonien « Michael Palin » (Kelman & Nicholas 2008, Milliard *et al.* 2008), ou encore le programme néerlandais RESTART, qui incorpore également des éléments de thérapie directe (de Sonnevile-Koedoot *et al.* 2015).

Parmi les traitements alternatifs, l'abord pharmacologique a été particulièrement étudié, avec des résultats souvent décevants (voir notamment Brady 1991, Bothe *et al.* 2006, Boyd *et al.* 2010 pour des revues). Suivant Monfrais-Pfauwadel (2014), les antagonistes de la dopamine seraient les plus efficaces sur les signes cliniques du bégaiement, mais avec de nombreux effets secondaires et des contre-indications, ce qui pose la question du rapport bénéfice-risque. Enfin, diverses thérapies psychologiques ont été explorées, notamment celles de type cognitivo-comportemental pour lesquelles on dispose de plusieurs essais cliniques dans le secteur du bégaiement (voir notamment Menzies *et al.* 2008), avec un effet positif établi en complément de la rééducation orthophonique lorsque les symptômes anxieux prédominent.

3. Revue de littérature

Que sait-on des caractéristiques phonétiques et phonologiques du bégaiement ? Nous rappelons brièvement les généralisations établies dans la littérature sur le plan segmental, suivant en cela la revue récente de Didirková *et al.* (2020), avant de considérer la dimension supra-segmentale, globalement moins étudiée.

3.1 Bégaiement et perturbations segmentales

Il a été montré que même lorsqu'aucune disfluence n'est perçue, la parole des sujets qui bégaiement présente certaines anomalies segmentales. D'une part, les délais d'établissement et de fins de voisement dans les contextes CVC tendent à être plus longs que chez les sujets contrôle (Agnello 1975, Hirsch 2007), ce qui concorde avec les observations nasofibroscopiques (Hirsch 2007) et cinéradiographiques (Zimmermann 1980), qui montrent respectivement un temps de réaction laryngée et une immobilisation des articulateurs supra-glottiques plus longs. D'autre part, l'espace vocalique des sujets qui bégaiement est typiquement plus réduit par centralisation (Blomgren *et al.* 1998, Hirsch 2007). Par ailleurs, plusieurs propriétés distinguent les disfluences de type bègue des disfluences normales, parmi lesquelles une affinité avec certaines catégories de sons (les consonnes sourdes, occlusives, et/ou associées au sein de syllabes complexes (Blomgren *et al.* 2012, Pendelieu-Verdurand 2014, Didirková 2016), des mouvements laryngés anarchiques (Conture *et al.* 1977, Monfrais-Pfauwadel *et al.* 2005, Hirsch 2007), une tendance à rompre la syllabe dans la zone de transition formantique ou « ligne de faille » (*fault lign*) entre l'attaque et la rime (Wingate 1988, Monfrais-Pfauwadel & Teitler 1996), ou encore un comportement anormal des articulateurs supra-glottiques sur les segments adjacents (Didirková & Hirsch 2019, Didirková *et al.* 2020). Enfin, la question de la coarticulation a suscité de nombreuses recherches tant lors des disfluences qu'en leur absence, avec des résultats qui restent pour l'heure contradictoires, certains travaux retrouvant des anomalies (voir par exemple Howell & Vause 1986, Pendelieu-Verdurand 2014), d'autres non (voir par exemple Sussman *et al.* 2011).

3.2 Bégaiement et perturbations supra-segmentales

De manière générale, l'intonation, ou mélodie de la parole principalement (mais pas uniquement) liée aux variations de F0, ne semble pas constituer un paramètre affecté dans le bégaiement (Schmitt & Cooper 1978, Bergmann 1986). Reprenant des observations de

Brown (1938), Wingate a insisté en revanche à diverses reprises sur l'importance de la dimension rythmique (voir notamment Wingate 1979, 1984, 1988). D'une part, il est admis que le *tempo* ou débit de parole est généralement plus lent chez les sujets qui bégaiement, y compris en l'absence de disfluences (Bloodstein 1944, Amster & Starkweather 1987, Schäfersküpfer & Dames 1987), ce qui constitue un critère de diagnostic différentiel avec le bredouillage (van Zaalen 2009). D'autre part, les disfluences se rencontrent plus fréquemment sur les syllabes accentuées que sur les syllabes inaccentuées dans les langues germaniques (Bergmann 1986, Prins *et al.* 1991, Arbisi-Kelm 2006, 2010). Comme l'a établi cependant Hubbard (1998), le facteur explicatif n'est pas tant l'accent *per se* que la position initiale de groupe prosodique, connue par ailleurs pour être un lieu de renforcement articulatoire (Fougeron & Keating 1997), ce que tendent à masquer les langues de type trochaïque comme l'anglais ou l'allemand où les deux facteurs se conjuguent typiquement. Plusieurs travaux sur ces mêmes langues ont mis en évidence par ailleurs une variabilité temporelle plus importante chez les sujets qui bégaiement sur des empanns tels que la phrase et le paragraphe (Cooper & Allen 1977), les intervalles inter-accentuels (Bergmann 1986), les syllabes (Schäfersküpfer & Dames 1987) ou encore les segments vocaliques (Onslow *et al.* 1992, Packman *et al.* 1994). À notre connaissance, l'étude de Maruthy *et al.* (2017) constitue cependant la première à rattacher explicitement le phénomène à la problématique du rythme, montrant au moyen de diverses mesures segmentales (voir chapitre 2, section 2.3.1) qu'en kannada, langue dravidienne à isochronie moraïque, les sujets qui bégaiement produisent des intervalles vocaliques et consonantiques de durées anormalement variables, ce qui rapproche leur parole du modèle des langues à isochronie accentuelle. Enfin, le bégaiement serait associé selon Wieland *et al.* (2015) à des capacités de discrimination rythmique moindres sur le versant perceptif ; l'étude se limite toutefois à la population pédiatrique et à des stimuli non verbaux.

CHAPITRE 2 : OBJECTIFS ET PROCÉDURE EXPÉRIMENTALE

1. Objectifs de l'étude

La présente étude vise à explorer les perturbations du rythme de la parole dans le bégaiement au moyen de méthodes acoustiques de type temporel.

Comme on l'a vu, les recherches dans ce secteur ont surtout porté jusqu'à présent sur l'anglais et des langues germaniques apparentées, dites à « isochronie accentuelle » ou fort degré d'accentuation. De plus, les travaux existants se limitent pour la plupart aux manifestations du rythme de la parole à l'échelle locale du segment ou de la syllabe, et ce en l'absence de disfluences. Par contraste, nous étudions les données du français, langue qui relève du type opposé à « isochronie syllabique » ou faible degré d'accentuation. Par ailleurs, nous considérons non seulement les séquences de parole perceptivement fluides, mais aussi celles qui ne le sont pas. Enfin et surtout, nous examinons le rythme dans ses manifestations locales, mais aussi supra-segmentales.

Plus précisément, nous proposons de tester l'hypothèse avancée par Monfrais-Pfauwadel (2014) suivant laquelle les disfluences de type bégue perturbent non seulement la structure interne des syllabes, mais aussi le « *chunking* » de la parole (Gilbert 2012), c'est-à-dire le regroupement des syllabes au sein d'unités temporelles supérieures, phénomène que nous assimilons ici sans discuter au phrasé en groupes accentuels, suivant en cela Martin (2014, 2017, 2018).

Pour ce faire, nous utilisons deux types de mesures acoustiques, élaborées initialement pour contraster les langues selon leur degré d'accentuation :

- des mesures segmentales, fondées sur la durée des intervalles vocaliques et consonantiques (Ramus *et al.* 1999, Grabe & Low 2002) ;
- des mesures supra-segmentales, fondées sur la durée des groupes accentuels (Asu & Nolan 2006, Obin *et al.* 2012, Cichocki 2015).

En accord avec les résultats de Maruthy *et al.* (2017), nous nous attendons à retrouver une variabilité temporelle anormale sur le plan segmental chez les sujets qui bégaièrent, rapprochant leur parole en français du modèle des langues à isochronie accentuelle. Par ailleurs, la tendance à l'eurythmie devrait être rompue à l'échelle supra-segmentale si le *chunking* est effectivement perturbé, se traduisant par des violations des contraintes de

durées minimale et maximale des groupes accentuels et/ou une variabilité temporelle plus marquée.

2. Procédure expérimentale

2.1 Participants

2.1.1 Présentation générale

Notre étude repose sur les enregistrements de 5 sujets pathologiques, extraits du corpus « *Paroles disfluentes* » (Praxiling-UMR 5267 2017), que nous avons complétés par les enregistrements de 5 sujets contrôle appariés en âge et sexe. Les critères d'inclusion et d'exclusion sont recensés dans le tableau 2.

	Critères d'inclusion	Critères d'exclusion
Sujets pathologiques	Bégaïement développemental modéré à sévère Âge $\geq$ 15 ans Monolinguisme français	Tout autre trouble connu susceptible d'affecter la parole
Sujets contrôle	Monolinguisme français	Tout trouble connu susceptible d'affecter la parole

Tableau 2. Critères d'inclusion et d'exclusion des participants de l'étude.

À l'instar des travaux antérieurs de Hirsch (2007), Pendelieu-Verdurand (2014) ou encore Didirková (2016) sur le français, nous nous limitons au bégaïement développemental, qui pose peu de problème de diagnostic différentiel, contrairement au bégaïement acquis. Le choix de travailler sur une population d'adultes présente par ailleurs deux intérêts. D'une part, les anomalies observées ne peuvent être imputées à un développement psychomoteur inachevé. D'autre part, la question du statut transitoire ou non du trouble ne se pose pas : il est par définition persistant. En contrepartie, la distinction entre anomalies structurelles et modifications de nature compensatoire est plus délicate. Enfin, le degré de sévérité du trouble retenu ici, *a minima* modéré, assure en principe que des perturbations de la parole seront présentes dans les échantillons recueillis, contrairement à un bégaïement léger.

Nous n'avons pas cherché en revanche à contrôler l'effet d'une prise en charge passée ou en cours. Il est en effet peu fréquent, du moins en France, de rencontrer des adultes qui bégaient n'ayant jamais bénéficié d'un suivi à un moment ou à un autre de leur vie. Du reste, la durée des effets d'une rééducation n'est pas connue et se combine plausiblement à des stratégies de compensation personnelles, comme le note Pendelieu-Verdurand (2014 : 137). De fait, exclure les sujets bénéficiant ou ayant bénéficié dans les 2 à 5 dernières années d'un suivi comme le font de nombreuses études n'assure en rien l'absence de biais.

Enfin, nous voudrions insister pour conclure sur l'intérêt de recourir à un corpus préexistant de données sonores pour des études exploratoires comme la nôtre. Il est vrai que ce choix implique de s'adapter à un matériel linguistique et des conditions d'enregistrement définis au départ pour d'autres objectifs. Il permet en revanche de contourner les obstacles auxquels se heurte le recrutement de sujets dans le champ clinique, parmi lesquels des dispositions législatives plus contraignantes, un éparpillement géographique fréquent des participants potentiels, ou encore une disponibilité moindre de ces derniers.

2.1.2 Description des sujets pathologiques

Notre échantillon de sujets pathologiques (tableau 3) se compose de 4 hommes et 1 femme dont les enregistrements, extraits du corpus précité, ont été réalisés en amont par Didirková pour une étude indépendante sur les disfluences (Didirková 2016). L'âge moyen est de 31,6 ans (médiane 30 ans, écart-type 11,8 ans, min. 15 ans, max. 43 ans), et les sujets appartiennent tous à la catégorie des cadres, excepté le sujet féminin qui était encore au lycée au moment des enregistrements. Le lieu d'habitation n'est pas renseigné, mais aucun sujet ne présente d'accent régional identifiable. Enfin, le bégaiement était jugé *a minima* modéré selon un thérapeute externe au moment du recrutement, et a été réévalué sur la base du pourcentage de mots bégayés dans les enregistrements recueillis, le trouble étant considéré modéré de 5 à 8%, modéré à sévère de 8 à 12%, et sévère de 12 à 25% (Didirková 2016 : 120-121).

Sujets	Sexe	Âge	Situation professionnelle	Degré de sévérité du bégaiement
P1	Masculin	30 ans	Cadre	Sévère
P2	Masculin	43 ans	Cadre	Modéré
P3	Féminin	15 ans	Étudiante	Modéré
P4	Masculin	27 ans	Cadre	Sévère
P5	Masculin	43 ans	Cadre	Modéré à sévère

Tableau 3. Échantillon de sujets pathologiques.

2.1.3 Description des sujets contrôle

Afin de disposer d'une référence à laquelle comparer les productions des sujets pathologiques, nous avons recruté dans notre entourage 5 sujets appariés en âge (± 2 ans) et sexe ne présentant aucun trouble de la parole (tableau 4). Comme les sujets pathologiques, aucun ne présente d'accent régional identifiable. Les professions représentées sont en revanche plus variées.

Sujets	Sexe	Âge	Situation professionnelle
C1	Masculin	30 ans	Commercial
C2	Masculin	43 ans	Professeur
C3	Féminin	15 ans	Étudiante
C4	Masculin	28 ans	Kinésithérapeute
C5	Masculin	45 ans	Professeur

Tableau 4. Échantillon de sujets contrôle.

2.2 Matériel et protocole de passation

Notre étude se limite à la parole lue, dont les propriétés rythmiques présentent moins de variation que la parole spontanée (Martin 2018). Afin de permettre une comparaison fiable, le protocole initial de Didirková (2016) a été reproduit à l'identique auprès de nos sujets contrôle. Tous les sujets ont ainsi été enregistrés en mono au format .wav 16 bits/44100 Hz, dans un milieu calme, assis face à l'expérimentateur, au moyen du même enregistreur PCM linéaire Tascam-DR100MKII®, placé au centre de la table séparant les deux individus. Après avoir pris connaissance des objectifs généraux de l'étude et signé (ainsi que leurs tuteurs légaux le cas échéant) un formulaire de consentement éclairé, les sujets étaient invités à lire silencieusement, puis à voix haute le texte dans l'encadré ci-dessous, choisi au départ par Didirková pour sa composition variée en mots simples et complexes sur les plans syllabique et morphologique. À l'instar de l'auteur précitée, nous n'avons volontairement pas demandé de relire à nouveau le texte lorsque la production comportait des disfluences. Un seul sujet est toutefois concerné, le plus jeune, et aucune des disfluences retrouvées, peu nombreuses au demeurant, n'est de type bégue : il s'agit soit de répétitions, soit de corrections liées à des erreurs de déchiffrement.

Par un long après-midi d'été, le lion dormait à l'ombre d'un arbre. Soudain, tout doucement, une petite queue lui toucha le bout du museau. Le lion se réveille en sursaut. «Comment oses-tu me réveiller ?» gronda-t-il. «Je dévore ceux qui me réveillent.» «Par pitié, ne me mange pas. Peut-être qu'un jour je te sauverai la vie.» «Toi ? Sauve-toi, tu me fais trop rire, sauve-toi.» Le lendemain, le lion se prit dans le filet d'un chasseur. Il se débattait, mais le filet se resserrait. La petite souris se précipita à son secours. Elle rongea la corde usée. «Tu m'as sauvé !» Même les petits animaux font de grands amis.

2.3 Mesures rythmiques

2.3.1 Mesures segmentales

Nous considérons les mesures segmentales les plus courantes, développées à partir de la fin des années 90 par Ramus *et al.* (1999), Grabe & Low (2002) puis divers chercheurs pour contraster les langues selon leur degré d'accentuation ou « classe rythmique » (tableau 5).

Mesure	Formule	Description
%V [Ramus <i>et al.</i> 1999]	$100 \times \left(\sum_{i=1}^n X_i / \sum_{i=1}^n (X_i + Y_i) \right)$	Proportion des durées vocaliques sur le total des durées vocaliques et consonantiques, exprimée en pourcentages
ΔV [Ramus <i>et al.</i> 1999]	$\sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}}$	Écart-type de la durée des intervalles vocaliques, défini comme la racine carrée de la variance, exprimé en ms
ΔC [Ramus <i>et al.</i> 1999]	$\sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n - 1}}$	Écart-type de la durée des intervalles consonantiques, défini comme la racine carrée de la variance, exprimé en ms
VarcoV [White & Matthys 2007]	$100 \times (\Delta X / \bar{X})$	Coefficient de variation de la durée des intervalles vocaliques, ou écart-type normalisé pour le débit, exprimé en pourcentages
VarcoC [Dellwo & Wagner 2003]	$100 \times (\Delta Y / \bar{Y})$	Coefficient de variation de la durée des intervalles consonantiques, ou écart-type normalisé pour le débit, exprimé en pourcentages
r-PVI-V [Grabe & Low 2002]	$\sum_{i=1}^{n-1} \frac{ X_i - X_{i+1} }{n - 1}$	Indice de variabilité par paires brut (<i>Raw Pairwise Variability Index</i>) de la durée des intervalles vocaliques, exprimé en ms
r-PVI-C [Grabe & Low 2002]	$\sum_{i=1}^{n-1} \frac{ Y_i - Y_{i+1} }{n - 1}$	Indice de variabilité par paires brut (<i>Raw Pairwise Variability Index</i>) de la durée des intervalles consonantiques, exprimé en ms
n-PVI-V [Grabe & Low 2002]	$\sum_{i=1}^{n-1} \frac{ X_i - X_{i+1} / (\frac{ X_i + X_{i+1} }{2})}{n - 1}$	Indice de variabilité par paires normalisé pour le débit (<i>Normalized Pairwise Variability Index</i>) de la durée des intervalles vocaliques, exprimé en pourcentages
n-PVI-C [Grabe & Low 2002]	$\sum_{i=1}^{n-1} \frac{ Y_i - Y_{i+1} / (\frac{ Y_i + Y_{i+1} }{2})}{n - 1}$	Indice de variabilité par paires normalisé pour le débit (<i>Normalized Pairwise Variability Index</i>) de la durée des intervalles consonantiques, exprimé en pourcentages

Tableau 5. Mesures rythmiques segmentales.

Légende : X = durée d'un intervalle vocalique ; Y = durée d'un intervalle consonantique.

Destinées à quantifier les manifestations dans le signal acoustique des deux paramètres supposés responsables des classes rythmiques, la réduction vocalique et la complexité syllabique (voir *supra* chapitre 1, section 1.2.3), ces mesures se veulent affranchies de toute hypothèse quant à la structure phonologique. Elles se fondent ainsi sur l'alternance acoustique la plus élémentaire entre les intervalles vocaliques (V), marqués par une concentration d'énergie associée à une ou plusieurs voyelles, et le matériel résiduel à l'exclusion des pauses silencieuses, c'est-à-dire les intervalles consonantiques composés d'une ou plusieurs consonnes (C). De fait, les frontières de phonèmes ou de syllabes ne sont pas toujours respectées, comme l'illustre la structure en intervalles de la phrase « *Comment oses-tu me réveiller ?* » : / k | ɔ | m | ão | zt | y | m | ə | ʁ | ε | v | ε | j | e / (voir *infra*, section 2.4.1). En outre, ces mesures reposent exclusivement sur la durée, guidées par l'hypothèse

que réduction vocalique et complexité syllabique se reflètent dans le signal par des variations temporelles. On s'attend ainsi à observer une variabilité plus élevée de la durée des intervalles vocaliques dans les langues présentant une forte réduction vocalique (ΔV , VarcoV, r-PVI-V, n-PVI-V), une variabilité plus élevée de la durée des intervalles consonantiques dans les langues présentant une riche inventaire de syllabes (ΔC , VarcoC, r-PVI-C, n-PVI-C), et une proportion de durées vocaliques plus faible dans les langues qui combinent l'un et l'autre des deux phénomènes (%V). Il est clair toutefois que d'autres facteurs interviennent dans les variations temporelles observées. C'est pourquoi toutes les mesures, à l'exclusion de %V, se déclinent sous une forme brute (ΔV , ΔC , r-PVI-V, r-PVI-C), mais aussi sous une forme normalisée pour le débit (VarcoV, VarcoC, n-PVI-V, n-PVI-C), destinée à neutraliser le cas échéant des variations non pertinentes pour l'analyse, comme certaines oppositions de durées intrinsèques (par exemple entre voyelles orales et voyelles nasales) ou contextuelles (par exemple entre voyelles non finales et voyelles finales). En outre, deux méthodes de calcul de la variabilité sont considérées : l'une fondée sur la variance (ΔV , ΔC , VarcoV, VarcoC), l'autre, sur la comparaison par paires des intervalles successifs (r-PVI-C, r-PVI-V, n-PVI-C, n-PVI-V). Enfin et surtout, aucune mesure ne suffit à elle seule pour obtenir une classification cohérente des langues. Il faut pour cela les combiner, comme l'illustrent les deux figures en 6, empruntées à Ramus (2002).

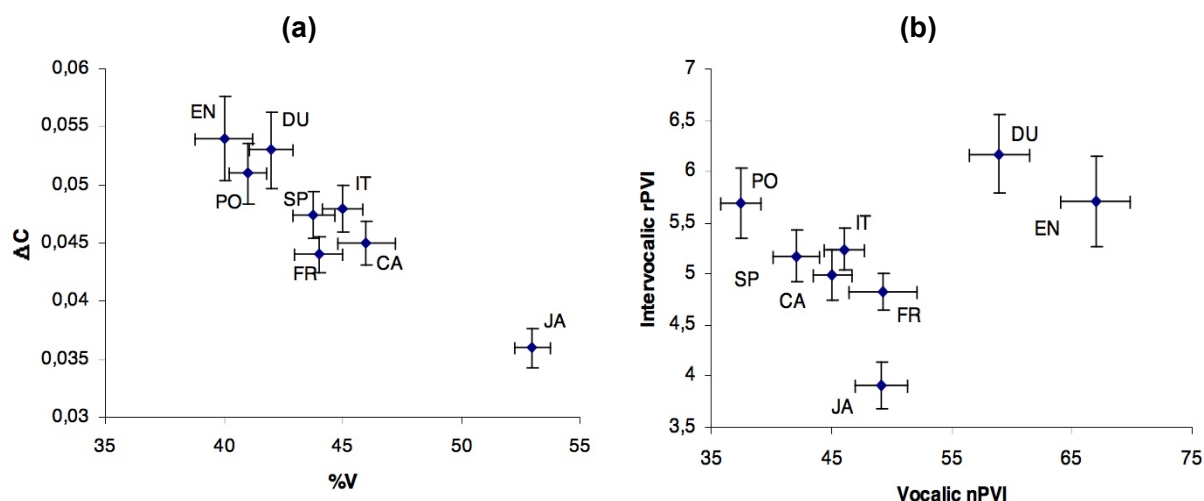


Fig. 6. Classifications des langues selon différentes mesures rythmiques d'après Ramus (2002).

Légende : en (a) l'espace obtenu en combinant la proportion de durées vocaliques (x) et l'écart-type de la durée des intervalles consonantiques (y), suivant en cela les propositions de Ramus *et al.* (1999) ; en (b) l'espace obtenu en combinant l'indice de variabilité par paires normalisé pour le débit de la durée des intervalles vocaliques (x) et l'indice de variabilité par paires brut de la durée des intervalles intervocaliques (c'est-à-dire consonantiques) (y), suivant en cela les propositions de Grabe & Low (2002). CA : catalan, DU : hollandais, EN : anglais, FR : français, IT : italien, JA : japonais, SP : espagnol, PO : polonais.

Comme on le voit, les langues dites « prototypiques » s'organisent de manière attendue dans les deux représentations. Ainsi les langues à fort degré d'accentuation ou « isochronie accentuelle » comme l'anglais et le hollandais se situent à l'opposé d'une langue moraïque comme le japonais, elle-même distante du cluster formé par les langues à faible degré d'accentuation ou « isochronie syllabique » comme l'espagnol, l'italien et le français. Comme le note Arvaniti (2009), l'analyse des langues dite « intermédiaires » est en revanche plus problématique. Ainsi le polonais, qui présente des syllabes complexes mais pas de réduction vocalique, s'apparente aux langues germaniques dans la représentation (a), tandis qu'il apparaît à l'opposé dans la représentation (b), à proximité des langues romanes. Le catalan, par ailleurs, qui présente la situation inverse, c'est-à-dire un phénomène de réduction vocalique mais des syllabes relativement simples, ne se distingue clairement des autres langues romanes dans aucune des deux représentations.

2.3.2 Mesures supra-segmentales

Afin d'étudier les contraintes d'eurythmie liées à la structure accentuelle, nous étendons les mesures de variabilité temporelle au niveau supra-segmental des groupes accentuels (GA), suivant en cela plusieurs études antérieures (Asu & Nolan 2006, Obin *et al.* 2012, Cichocki 2015). En outre, nous considérons deux mesures complémentaires empruntées à Martin (2014), Min(GA) et Max(GA), étroitement liées aux contraintes « *no clash* » et « *no lapse* » communément admises en phonologie métrique (tableau 6).

Mesure	Formule	Description
ΔGA	$\sqrt{\frac{\sum_{i=1}^n (Z_i - \bar{Z})^2}{n - 1}}$	Écart-type de la durée des groupes accentuels, défini comme la racine carrée de la variance, exprimé en ms
VarcoGA	$100 \times (\Delta Z / \bar{Z})$	Coefficient de variation de la durée des groupes accentuels, ou écart-type normalisé pour le débit, exprimé en pourcentages
r-PVI-GA	$\sum_{i=1}^{n-1} \frac{ Z_i - Z_{i+1} }{n - 1}$	Indice de variabilité par paires brut (<i>Raw Pairwise Variability Index</i>) de la durée des groupes accentuels, exprimé en ms
n-PVI-GA	$\sum_{i=1}^{n-1} \frac{ Z_i - Z_{i+1} / (\frac{ Z_i + Z_{i+1} }{2})}{n - 1}$	Indice de variabilité par paires normalisé pour le débit (<i>Normalized Pairwise Variability Index</i>) de la durée des groupes accentuels, exprimé en pourcentages
Min(GA)	$\min (Z_i)_{i=1, \dots, n}$	Durée minimale d'un groupe accentuel, exprimée en ms
Max(GA)	$\max (Z_i)_{i=1, \dots, n}$	Durée maximale d'un groupe accentuel, exprimée en ms

Tableau 6. Mesures rythmiques supra-segmentales.

Légende : Z = durée d'un groupe accentuel.

À la différence des intervalles vocaliques et consonantiques, les groupes accentuels ne se déduisent pas de manière simple des propriétés du signal acoustique. Nous développons plus loin, section 2.4.2, la méthode d'analyse retenue.

2.4 Traitement des fichiers sonores

Afin de procéder aux mesures rythmiques décrites, les fichiers sonores ont d'abord été segmentés manuellement en intervalles vocaliques et consonantiques, puis en groupes accentuels au moyen du logiciel libre *Praat* (version 6.0.29, Boersma & Weenink 2017). Nous détaillons dans ce qui suit les critères d'analyse, puis les annotations complémentaires réalisées.

2.4.1 Analyse segmentale

Pour déterminer précisément chaque frontière d'intervalle consonantique ou vocalique, nous nous sommes conjointement appuyé sur l'écoute attentive des données sonores et les représentations fournies par l'oscillogramme, le spectrogramme à bandes larges (fenêtre d'analyse gaussienne de 3,5 à 5ms, dynamique de 40 à 60 dB, échelle fréquentielle de 0 à 5500 Hz pour les voyelles, 0 à 8000 Hz pour les consonnes), et les courbes d'intensité et de F0 (échelle variable selon la voix). Les critères acoustiques retenus sont recensés dans le tableau 7 *infra* ; ils sont relativement standard (voir Peterson & Lehiste 1960). Dans un certain nombre de situations toutefois, ils s'avèrent insuffisants. Nous présentons les cas problématiques rencontrés et les solutions retenues. De manière générale, nous avons donné priorité aux indices fournis par l'oscillogramme et l'analyse perceptive lorsque la segmentation posait problème.

Pauses et tenues des occlusives non voisées

Pour identifier les pauses, nous avons retenu le seuil standard (mais arbitraire) de 200 ms de silence, entrecoupé le cas échéant de bruits sans valeur linguistique comme les reprises inspiratoires. En-deçà de cette durée, le silence a ainsi été intégré à l'intervalle précédent (contexte C_V), suivant (contexte V_C) ou englobant (contextes V_V et C_C). Un problème bien connu se pose lorsqu'est réalisée à la suite une consonne occlusive intrinsèquement ou contextuellement non voisée, dont le temps de tenue est lui aussi silencieux. À l'instar de Grabe & Low (2002) ou encore White & Matthys (2007), nous avons intégré l'ensemble de la consonne à la pause qui précédait dans les enregistrements des sujets contrôle, ce qui revient à l'exclure de l'analyse (voir *supra*, section 2.3.1). Il nous a semblé en revanche essentiel de ne pas perdre cette information chez les sujets qui bégayaient lorsque l'ensemble est clairement perçu comme un blocage pré-phonatoire. C'est pourquoi nous avons opté

pour la solution inverse consistant, dans les cas manifestes de blocage et seulement ceux-là, à intégrer dans un même intervalle consonantique le relâchement de l'occlusive et le silence précédent (annexe 1a), à l'exclusion toutefois de la portion initiale lorsque celle-ci se terminait par un bruit inspiratoire évocateur d'une pause (annexe 1b).

	Spectrogramme	Oscillogramme	Courbes de F0 et d'intensité
Début d'un intervalle vocalique	Début de F1, voire F2, si voyelle précédée d'un silence Début de F2, voire F3, si voyelle précédée d'une consonne sans formants Mouvements et intensité de F1, F2, voire F3, si voyelle précédée d'une consonne sonante	Point zéro avant le premier pic d'amplitude initiant un nouveau motif d'amplitude croissante	Montée d'intensité Apparition d'une courbe de F0 si voyelle précédée d'une consonne non voisée ou d'un silence
Fin d'un intervalle vocalique	Fin de F1, voire F2, si voyelle suivie d'un silence (tenue des occlusives non voisées incluse) Fin de F2, voire F3 si voyelle suivie d'une consonne sans formants Mouvements et intensité de F1, F2, voire F3, si voyelle suivie d'une consonne sonante	Point zéro après le dernier creux d'amplitude si voyelle suivie d'un silence (tenue des occlusives non voisées incluse) Point zéro avant le premier pic d'amplitude initiant un nouveau motif, d'amplitude moindre, si voyelle suivie d'une consonne (occlusives non voisées exclues)	Chute d'intensité Disparition de F0 si voyelle suivie d'une consonne non voisée ou d'un silence
Début d'un intervalle consonantique après silence (occlusives non voisées exclues)	Début du bruit de friction pour les fricatives non voisées Début de la barre de voisement pour les fricatives et les occlusives voisées Début de la barre de voisement et/ou des formants visibles pour les sonantes	Point zéro avant le premier pic d'amplitude initiant un nouveau motif	Montée d'intensité Apparition d'une courbe de F0 si consonne voisée
Fin d'un intervalle consonantique avant silence	Fin du bruit impulsif au relâchement pour les occlusives non voisées Fin du bruit de friction pour les fricatives non voisées Fin de la barre de voisement et/ou des formants pour les autres consonnes	Point zéro après le dernier creux d'amplitude	Chute d'intensité Disparition de F0 si consonne voisée

Tableau 7. Critères acoustiques pour la segmentation en intervalles vocaliques et consonantiques.

Légende : F1 = premier formant ; F2 = deuxième formant ; F3 = troisième formant.

Bruits parasites

Certains enregistrements comportent un bruit de fond assez proche du premier formant (F1) ou d'une barre de voisement et/ou un bruit d'écho qui génère des traces de formants de faible amplitude comparables à la phase d'amortissement d'une voyelle (« *voice decay*

time »). Nous nous sommes efforcé d'exclure systématiquement ces éléments de l'analyse en nous appuyant sur la forme de l'onde de l'oscillogramme, l'amplitude s'élevant anormalement après le déclin attendu en cas de bruit parasite (annexe 1c).

Voyelles

Avant ou après un silence, le recours au deuxième formant (F2) peut être nécessaire pour identifier le début ou la fin de la voyelle lorsque l'enregistrement comporte un bruit de fond important qui se confond avec le premier formant (F1). Avant ou après un autre segment, le recours au troisième formant (F3) est par ailleurs utile pour les voyelles antérieures. Enfin, les trois premiers formants doivent être pris en compte (et combinés dans tous les cas à l'analyse perceptive) lorsqu'il s'agit de séparer les voyelles des consonnes sonantes et semi-voyelles, tâche de loin la plus délicate. Certaines voyelles fermées, notamment /i/, peuvent par ailleurs se présenter sous une forme chuchotée et/ou fricatisée. La délimitation s'est surtout basée dans ce cas sur l'analyse perceptive, les frontières acoustiques étant particulièrement difficiles à établir (annexe 1d). En cas d'association à une consonne où le bruit de friction prédomine, la segmentation peut être impossible, et la séquence a alors été exclue de l'analyse (c'est-à-dire traitée comme une pause, annexe 1e). Dans d'autres cas, enfin, la voyelle est glottalisée à l'initiale ou en finale. Cette glottalisation est généralement graduelle, c'est pourquoi nous l'avons systématiquement intégrée à l'intervalle vocalique (annexe 1f). De fait, les voyelles en hiatus ne sont jamais séparées : elles forment toujours un seul et même intervalle (annexe 1g).

Consonnes et semi-voyelles

Nous avons fait le choix dans cette étude d'analyser uniformément les semi-voyelles /j/, /ɥ/ et /w/ comme des consonnes sonantes, à l'instar des nasales /m/, /n/ et des liquides /l/ et /ʁ/. Nous les avons séparées des intervalles vocaliques en combinant en général plusieurs critères :

- la forme du signal (d'amplitude moindre) ;
- les mouvements et l'intensité des formants F1, F2 et F3 ;
- les variations des courbes d'intensité et de F0 ;
- le traitement perceptif.

De même que certaines voyelles peuvent être fricatisées, les consonnes fricatives, et parfois également les occlusives voisées, peuvent être réalisées comme des approximantes, le bruit attendu étant remplacé par des traces de formants (annexe 1h). Nous les avons analysées malgré tout comme des consonnes.

Disfluences

Les disfluences comme les corrections ou les répétitions de séquences syllabiques ne posent pas de problèmes particuliers pour la segmentation en intervalles vocaliques et consonantiques. On peut en revanche se demander comment analyser celles de type bague affectant des unités inférieures à la syllabe. Nous avons déjà précisé plus haut comment sont intégrés les blocages pré-phonatoires. Les segments faisant l'objet d'une répétition n'ont été séparés par ailleurs par un intervalle de silence que dans le cas où celui-ci était supérieur ou égal à 200 ms, sans éléments perceptifs évocateurs d'un blocage (annexe 1i). Enfin, nous avons systématiquement traité comme des intervalles uniques les prolongations, ce qui recouvre les cas bien connus sur des fricatives ou des voyelles, mais aussi des formes plus atypiques générées par le relâchement intermittent d'une occlusion, avec ou sans voisement, qui combinent en fait des traits des trois types de disfluences bague (annexe 1j).

2.4.2 Analyse supra-segmentale

L'identification des prééminences qui fondent le phrasé en groupes accentuels constitue une tâche notoirement délicate dans les langues à faible degré d'accentuation comme le français (Rossi 1979, Vaissière 1991, 2010, Martin 2005). Suivant les recommandations d'Avanzi *et al.* (2010), nous avons combiné l'analyse perceptive sur une fenêtre relativement courte (l'unité inter-pausale) à l'inspection des indices acoustiques fournis de manière variable par les courbes d'intensité et de F0 (échelle variable selon les sujets), le spectrogramme à bandes étroites (fenêtre d'analyse gaussienne de 35 à 50 ms, dynamique de 40 à 60 dB, échelle fréquentielle fonction de l'étendue vocale), et dans une moindre mesure l'oscillogramme (tableau 8).

	Spectrogramme	Oscillogramme	Courbes de F0 et d'intensité
Fin d'un groupe accentuel	Mouvement montant ou descendant des harmoniques Allongement de la rime Silence subséquent	Silence subséquent	Mouvement montant ou descendant de F0 Pic d'intensité
Début d'un groupe accentuel	Mouvement montant des harmoniques (en cas d'accent initial) Allongement de l'attaque consonantique Silence précédent	Silence précédent	Pic local d'intensité et/ou mouvement montant de F0 (en cas d'accent initial) Rehaussement de la ligne de base (par comparaison avec le groupe précédent)

Tableau 8. Critères acoustiques pour la segmentation en groupes accentuels.

Nous illustrons dans l'encadré ci-dessous le phrasé le plus couramment rencontré dans les productions étudiées en l'absence de disfluences. Les groupes accentuels sont représentés par le parenthésage sans ponctuation ni majuscules, les accents primaires par les caractères gras, les accents secondaires (rares, et sans rôle évident de régulation rythmique dans les productions considérées, sauf peut-être dans la dernière phrase) par le soulignement, et les pauses supérieures ou égales à 200 ms par le symbole #. Bien entendu, d'autres réalisations sont possibles et attestées dans nos enregistrements. S'agissant de parole lue, la variation est toutefois limitée : elle concerne essentiellement les noms composés (*après-midi*) et groupes nominaux comprenant plusieurs mots lexicaux (*une petite queue, le bout du museau, la petite souris, la corde usée*), et quelques séquences analysables comme des phrases (*comment oses-tu me réveiller, tu me fais trop rire*) ou des groupes verbaux (*font de grands amis*). Dans les cas les plus simples, le groupe accentuel coïncide avec une unité prosodique supérieure, telle que le groupe intonatif (Delais-Roussarie & Di Cristo à paraître), ou encore l'unité inter-pausale. Il se caractérise alors par un mouvement mélodique ample montant ou descendant sur la syllabe finale (ou la pénultième en cas de schwa dit « extramétrique ») accompagné d'un allongement notable de la rime (annexe 2a), ou encore par des pauses encadrantes (annexe 2b). Dans d'autres cas, il s'insère dans une structure plus complexe et l'on peut parfois hésiter sur le découpage, les mouvements de F0 étant moins marqués et la durée peu exploitée. Nous renvoyons le lecteur aux figures annexées en 2c, 2d, 2e, 2f et 2g pour quelques exemples d'analyse.

(par un **long**) (après-midi) (d'été) # (le **lion**) (dormait) (à l'ombre) (d'un **arbre**) # (soudain) # (tout doucement) # (une petite **queue**) (lui toucha) (le **bout**) (du **museau**) # (le **lion**) (se réveille) (en sursaut) # (comment) (oses-tu) (me réveiller) # (gronda-t-il) # (je dévore) (ceux qui me réveillent) # (par pitié) # (ne me mange **pas**) # (peut-être) (qu'un **jour**) (je te sauverai) (la **vie**) # (**toi**) # (sauve-toi) # (tu me **fais**) (trop **rire**) # (sauve-toi) # (le lendemain) # (le **lion**) (se **prit**) (dans le **filet**) (d'un **chasseur**) # (il se débattait) (mais le **filet**) (se **resserrait**) # (la petite **souris**) (se précipita) (à son **secours**) # (elle rongeait) (la corde **usée**) # (tu m'as sauvé) # (même les petits **animaux**) (font de **grands amis**)

À quelques exceptions près où une borne a dû être ajoutée pour respecter la structure syllabique, les frontières de groupes accentuels ont été alignées sur celles des intervalles vocaliques ou consonantiques identifiées en amont. En accord avec nos choix segmentaux, les groupes accentuels introduits par une consonne occlusive non voisée ont par ailleurs été systématiquement exclus après une pause, c'est-à-dire traités eux-mêmes comme des pauses. L'analyse des disfluences pose de loin le plus de problèmes. Nous détaillons les cas

de figure rencontrés et les solutions retenues, qui concernent pour l'essentiel les sujets pathologiques (voir *supra*, section 2.2).

Blocages, prolongations et répétitions de segments

Blocages, prolongations liées ou non à une hésitation et répétitions de segments ont été intégrés de manière générale au groupe accentuel englobant (annexe 2h). Dans certains cas, toutefois, le segment cible est à la fois précédé d'une frontière (pause ou syllabe accentuée) et suivi d'une pause ; nous l'avons traité alors comme un groupe accentuel à lui seul, y compris en l'absence de voyelle (annexe 2i).

Répétitions de syllabes

Les répétitions syllabiques de type bègue se caractérisent par des tensions audibles liées, entre autres, à des élévations de l'intensité, ce que l'on peut interpréter comme des accentuations inappropriées (Packman *et al.* 1996). C'est pourquoi nous les avons analysées comme des groupes accentuels à part entière (annexe 2j), excepté la dernière syllabe, que nous avons intégrée au groupe accentuel subséquent sur le modèle des répétitions ordinaires, typiquement négligées par l'auditeur (annexe 2k).

Autres disfluences

Le groupe accentuel est parfois interrompu en cours de production avant la voyelle accentuée attendue, puis repris dans son intégralité, avec ou sans correction (annexes 2l et 2m). Nous avons considéré qu'il y avait alors des groupes accentuels distincts, le ou les premiers étant incomplets. Aucun enregistrement ne comporte en revanche d'interjections (*euuh*, *mmm*) ou de corrections impliquant un matériel linguistique complètement différent de la séquence corrigée, ce qui s'explique naturellement, du moins pour ce qui concerne les sujets contrôle, par le fait qu'il s'agit de parole lue.

2.4.3 Annotations complémentaires

Afin de comparer les productions avec et sans disfluences, les tires segmentale et supra-segmentale ont été dédoublées lorsque cela était nécessaire, celles sans disfluences étant obtenues en traitant comme une pause chaque syllabe où intervenait une disfluence au niveau segmental (y compris la dernière syllabe en cas de répétition), et les groupes accentuels englobants au niveau supra-segmental. Par ailleurs, une tire spécifique a été ajoutée pour coder les voyelles accentuées, ce qui permet d'identifier aisément les groupes accentuels atypiques qui en sont dépourvus (voir *supra*, section 2.4.2) ; elle ne joue cependant aucun rôle dans les calculs. Nous illustrons dans la figure 7 un exemple de segmentation maximale avec et sans disfluences. Au total, l'analyse complète d'une

production d'environ 45 secondes a nécessité en moyenne 30 heures par sujet pathologique, contre 15 heures par sujet contrôle.

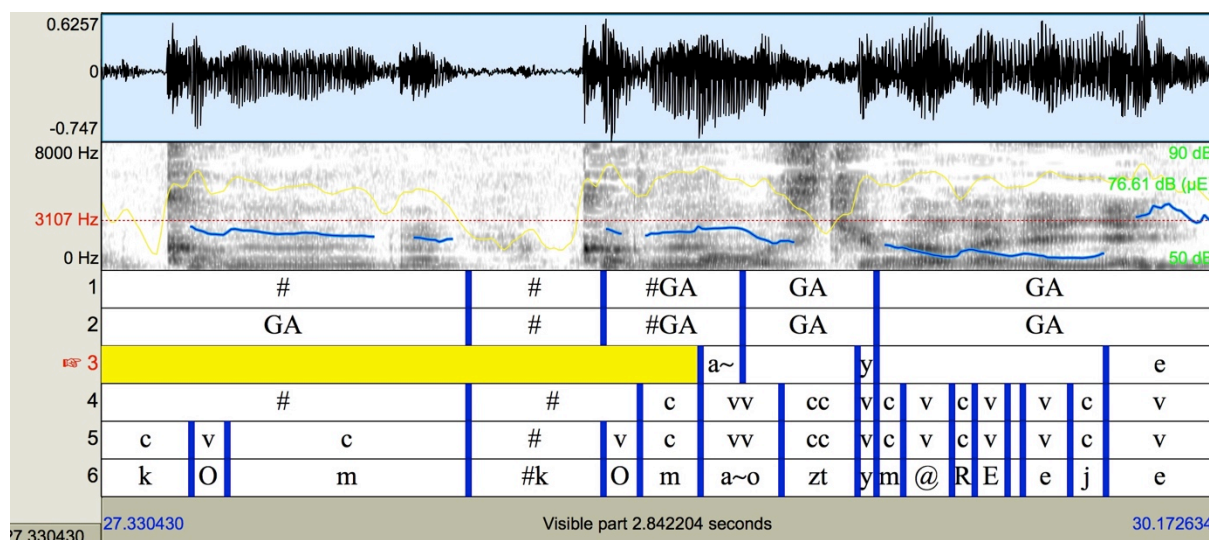


Fig. 7. Segmentation maximale de la séquence « com- comment oses-tu me réveiller », sujet P5.

Légende : la tire 1 correspond à la segmentation en groupes accentuels une fois éliminés les disfluences (notées #) et les groupes introduits par une consonnes occlusive non voisée après une pause (noté #GA), la tire 2 à la même segmentation avec les disfluences, la tire 3 à la segmentation en voyelles accentuées, la tire 4 à la segmentation en intervalles vocaliques et consonantiques une fois éliminées les disfluences et les consonnes occlusives non voisées introduites par une pause (notées #), la tire 5 à la segmentation des mêmes intervalles avec les disfluences et la tire 6 à la transcription Sampa phonème par phonème associée. La courbe d'intensité est représentée en jaune, la courbe de F0 en bleu, et le spectrogramme est réglé ici en bandes larges (fenêtre d'analyse gaussienne de 3,5 ms, échelle fréquentielle de 0 à 8000 Hz, dynamique à 50 dB).

2.5. Méthodes de calcul et d'analyse statistique

Les mesures rythmiques ont toutes été calculées au moyen du logiciel libre *Correlatore* (Mairano & Romano 2010), puis traitées statistiquement *via* l'extension *Real Statistics Resource Pack* (version 3.8, ©Charles Zaiontz), exécutable sous Microsoft Excel 2011©. Compte tenu du caractère exploratoire de notre étude et des faibles effectifs considérés, nous nous sommes limité pour les comparaisons pertinentes aux méthodes inférentielles les plus simples, de type non paramétrique.

CHAPITRE 3 : RÉSULTATS ET DISCUSSION

1. Résultats segmentaux

1.1. Résultats avant élimination des disfluences

Considérons d'abord la parole de manière globale, avant élimination des disfluences. Les valeurs obtenues par les sujets pathologiques et les sujets contrôle pour chaque mesure rythmique sont rassemblées dans les tableaux 9 et 10, et résumées dans la figure 8 sous la forme de boîtes à moustaches ou « *boxplots* » (Tuckey 1977).

	%V	ΔV	ΔC	VarcoV	VarcoC	r-PVI-V	r-PVI-C	n-PVI-V	n-PVI-C
P1	51,59	83,22	59,68	78,21	57,39	70,06	64,55	55,26	61,33
P2	50,57	56,61	59,20	58,91	59,34	58,54	56,97	54,15	53,62
P3	50,76	56,24	54,17	64,70	63,00	58,39	59,03	60,99	67,21
P4	37,37	65,14	319,61	55,63	158,62	71,60	186,39	56,17	66,06
P5	32,37	67,53	647,46	58,51	263,38	69,38	288,78	54,21	72,89
Médiane	50,57	65,14	59,68	58,91	63,00	69,38	64,55	55,26	66,06
Moyenne	44,53	65,75	228,03	63,19	120,35	65,60	131,14	56,16	64,22
Écart-type	9,00	10,98	260,47	9,02	90,69	6,56	103,73	2,83	7,21

Tableau 9. Valeurs des mesures rythmiques segmentales avant élimination des disfluences chez les sujets pathologiques (P), exprimées en millisecondes (ΔV , ΔC , r-PVI-V, r-PVI-C) ou en pourcentages (%V, VarcoV, VarcoC, n-PVI-V, n-PVI-C) arrondis à la deuxième décimale.

	%V	ΔV	ΔC	VarcoV	VarcoC	r-PVI-V	r-PVI-C	n-PVI-V	n-PVI-C
C1	46,67	53,87	55,90	64,59	57,41	50,80	57,13	51,55	56,68
C2	47,00	43,47	49,60	52,05	50,50	47,20	56,31	52,73	55,38
C3	53,14	37,34	51,48	40,38	61,07	41,15	52,69	43,76	60,11
C4	48,24	43,89	41,68	60,71	51,84	41,89	44,69	50,82	52,85
C5	44,94	23,01	42,68	33,25	49,25	23,75	47,84	34,42	55,42
Médiane	47,00	43,47	49,60	52,05	51,84	41,89	52,69	50,82	55,42
Moyenne	48,00	40,32	48,27	50,20	54,01	40,96	51,73	46,66	56,09
Écart-type	3,11	11,34	6,02	13,27	5,03	10,41	5,38	7,69	2,64

Tableau 10. Valeurs des mesures rythmiques segmentales avant élimination des disfluences chez les sujets contrôle (C), exprimées en millisecondes (ΔV , ΔC , r-PVI-V, r-PVI-C) ou en pourcentages (%V, VarcoV, VarcoC, n-PVI-V, n-PVI-C) arrondis à la deuxième décimale.

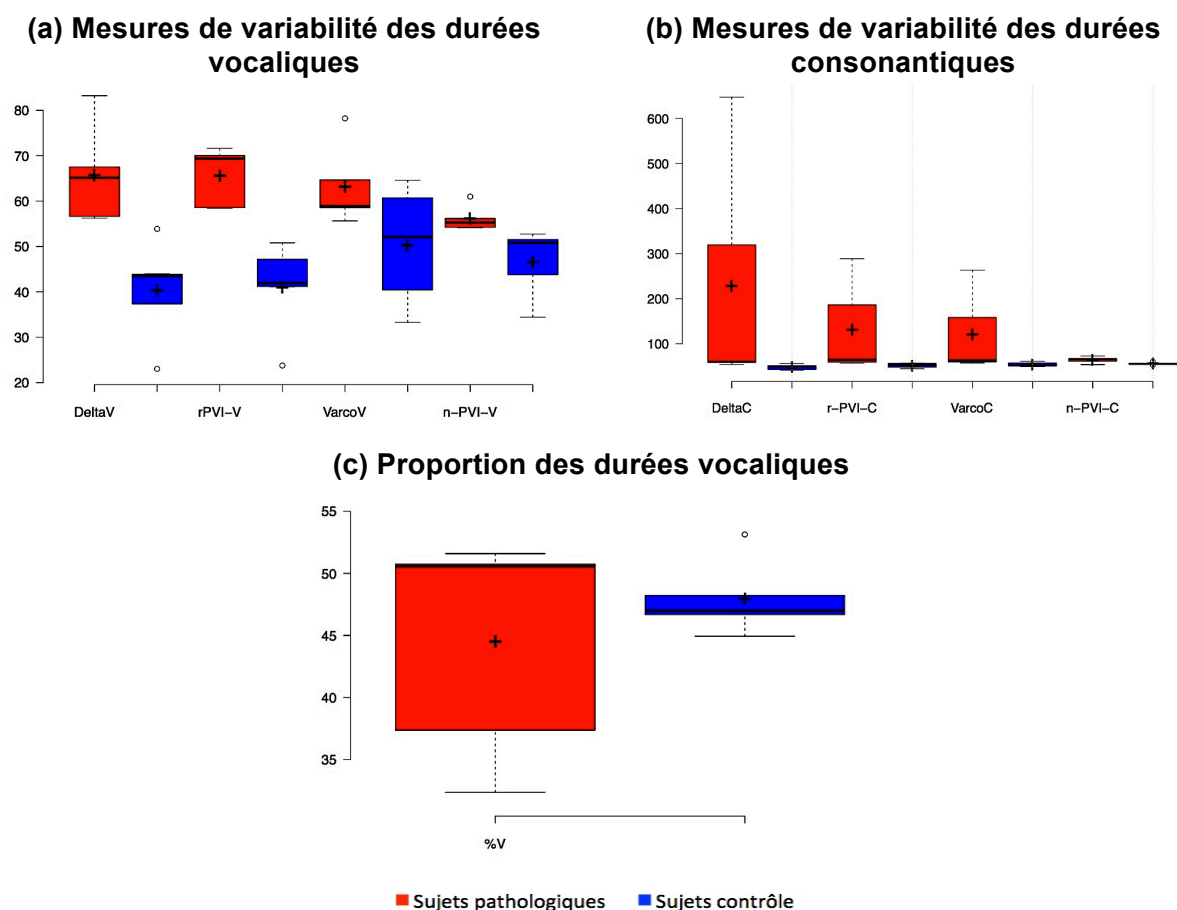


Fig. 8. Boîtes à moustaches illustrant la distribution des valeurs pour chaque mesure rythmique segmentale avant élimination des disfluences chez les sujets pathologiques et les sujets contrôle.

Légende : en (a) les mesures de variabilité des durées vocaliques, en (b) les mesures de variabilité des durées consonantiques, en (c) la mesure de proportion des durées vocaliques. Le rectangle principal de chaque boîte à moustaches correspond aux 50% des effectifs les plus proches de la médiane de l'échantillon considéré, c'est-à-dire à l'espace entre le premier quartile (Q1) délimité par le trait inférieur du rectangle, et le troisième quartile (Q3) délimité par le trait supérieur. Les « moustaches » haute et basse, quant à elles, correspondent aux 50% restants à l'exclusion des valeurs dites « atypiques » supérieures ($>Q3+1,5*(Q3-Q1)$) et inférieures ($<Q1-1,5*(Q3-Q1)$), représentées par le symbole °. La valeur moyenne est indiquée par le symbole + et la médiane par la barre noire horizontale.

Que l'on s'appuie sur la variance (ΔV , ΔC) ou sur la comparaison par paires (r-PVI-V, r-PVI-C), on constate que la variabilité temporelle brute est plus élevée chez les sujets qui bégayaient que chez les sujets contrôle. En outre, les valeurs médianes et moyennes maximales sont associées aux mesures consonantiques, ce qui s'explique plausiblement par le fait que blocages et prolongations affectent davantage les consonnes que les voyelles. Par ailleurs, la normalisation pour le débit (VarcoV, VarcoC, n-PVI-V, n-PVI-C) ne modifie que partiellement la situation : les valeurs sont nettement moins dispersées, notamment celles des mesures consonantiques, mais restent plus élevées chez les sujets qui bégayaient. L'analyse des résultats concernant la mesure %V est plus délicate. La moyenne est plus faible chez les sujets qui bégayaient que chez les sujets contrôle, mais la dispersion est importante et à la différence des autres mesures, la médiane, moins sensible aux valeurs

extrêmes, est au contraire plus élevée. Clairement, il faut distinguer entre deux profils parmi les sujets pathologiques : l'un caractérisé par une variabilité des durées consonantiques du même ordre de grandeur que la variabilité des durées vocaliques (P1, P2, P3), ce qui rapproche la valeur de %V de celle des sujets contrôle ; l'autre par une variabilité des durées consonantiques nettement supérieure à la variabilité des durées vocaliques, ce qui abaisse massivement la valeur de %V (P4, P5). En accord avec ces observations, la différence entre les deux groupes de sujets n'est pas statistiquement significative pour cette mesure, tandis qu'elle l'est pour les autres ($p < 0,05$), à l'exception de VarcoV et n-PVI-C (tableau 11).

	%V	ΔV	ΔC	VarcoV	VarcoC	r-PVI-V	r-PVI-C	n-PVI-V	n-PVI-C
U	12	0	1	6	3	0	1	0	4
z	-0,10	-2,61	-2,4	-1,35	-1,98	-2,61	-2,4	-2,61	-1,77
p	0,91	0,009*	0,01*	0,17	0,04*	0,009*	0,01*	0,009*	0,07

Tableau 11. Comparaison des valeurs obtenues par les sujets pathologiques et les sujets contrôle pour chaque mesure rythmique segmentale avant élimination des disfluences au moyen du test U bilatéral de Mann-Whitney.

1.2. Résultats après élimination des disfluences

Considérons à présent la parole dite « fluide », après élimination des disfluences. Les valeurs obtenues par les sujets pathologiques et les sujets contrôle pour chaque mesure rythmique sont rassemblées dans les tableaux 12 et 13, et résumées à nouveau dans la figure 9 par des boîtes à moustaches.

	%V	ΔV	ΔC	VarcoV	VarcoC	r-PVI-V	r-PVI-C	n-PVI-V	n-PVI-C
P1	50,13	58,50	59,08	60,47	57,96	55,92	63,81	51,83	62,26
P2	50,33	53,38	51,21	57,85	54,71	54,06	53,25	52,09	53,33
P3	50,20	56,20	49,44	66,21	58,32	56,48	56,26	59,50	67,20
P4	49,86	71,55	65,82	60,32	54,35	78,55	72,34	59,46	58,50
P5	47,97	61,35	57,89	55,74	50,06	63,55	67,32	52,97	57,91
Médiane	50,13	58,50	57,89	60,32	54,71	56,48	63,81	52,97	58,50
Moyenne	49,70	60,20	56,69	60,12	55,08	61,71	62,60	55,17	59,84
Écart-type	0,98	6,99	6,58	3,93	3,34	10,08	7,84	3,96	5,19

Tableau 12. Valeurs des mesures rythmiques segmentales après élimination des disfluences chez les sujets pathologiques (P), exprimées en millisecondes (ΔV , ΔC , r-PVI-V, r-PVI-C) ou en pourcentages (%V, VarcoV, VarcoC, n-PVI-V, n-PVI-C) arrondis à la deuxième décimale.

	%V	ΔV	ΔC	VarcoV	VarcoC	r-PVI-V	r-PVI-C	n-PVI-V	n-PVI-C
C1	46,67	53,87	55,90	64,59	57,41	50,80	57,13	51,55	56,68
C2	47,00	43,47	49,60	52,05	50,50	47,20	56,31	52,73	55,38
C3	50,22	37,84	60,59	41,68	68,31	41,61	58,56	44,97	62,28
C4	48,24	43,89	41,68	60,71	51,84	41,89	44,69	50,82	52,85
C5	44,94	23,01	42,68	33,25	49,25	23,75	47,84	34,42	55,42
Médiane	47,00	43,47	49,60	52,05	51,84	41,89	56,31	50,82	55,42
Moyenne	47,41	40,42	50,09	50,46	55,46	41,05	52,91	46,90	56,52
Écart-type	1,96	11,31	8,21	13,04	7,83	10,41	6,22	7,59	3,51

Tableau 13. Valeurs des mesures rythmiques segmentales après élimination des disfluences chez les sujets contrôle (C), exprimées en millisecondes (ΔV , ΔC , r-PVI-V, r-PVI-C) ou en pourcentages (%V, VarcoV, VarcoC, n-PVI-V, n-PVI-C) arrondis à la deuxième décimale.

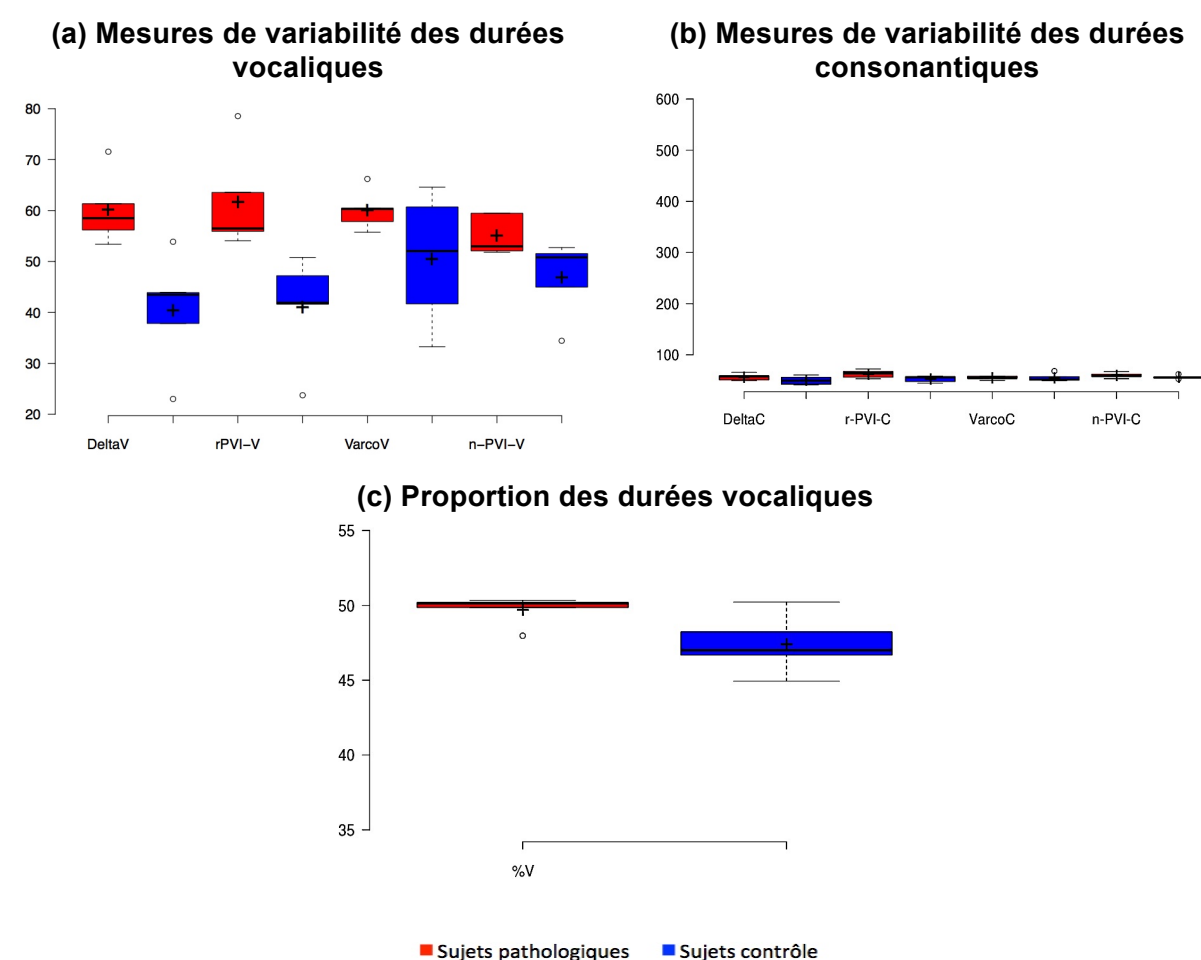


Fig. 9. Boîtes à moustaches illustrant la distribution des valeurs pour chaque mesure rythmique segmentale après élimination des disfluences chez les sujets pathologiques et les sujets contrôle.

Légende : en (a) les mesures de variabilité des durées vocaliques, en (b) les mesures de variabilité des durées consonantiques, en (c) la mesure de proportion des durées vocaliques. Le rectangle principal de chaque boîte à moustaches correspond aux 50% des effectifs les plus proches de la médiane de l'échantillon considéré, c'est-à-dire à l'espace entre le premier quartile (Q1) délimité par le trait inférieur du rectangle, et le troisième quartile (Q3) délimité par le trait supérieur. Les « moustaches » haute et basse, quant à elles, correspondent aux 50% restants à l'exclusion des valeurs « atypiques » supérieures ($>Q3+1,5*(Q3-Q1)$) et inférieures ($<Q1-1,5*(Q3-Q1)$), représentées par le symbole °. La valeur moyenne est indiquée par le symbole + et la médiane par la barre noire horizontale.

Comme on peut le voir, l'élimination des disfluences ne modifie pas radicalement le tableau d'ensemble pour ce qui concerne les mesures de variabilité des durées vocaliques : les valeurs médianes et moyennes restent plus élevées chez les sujets pathologiques que chez les sujets contrôle et la normalisation pour le débit se limite à réduire la dispersion. Par contraste, la variabilité des durées consonantiques s'approche chez les sujets pathologiques de celle observée chez les sujets contrôle, ce que confirme le test *U* de Mann-Whitney : la différence n'est statistiquement significative que pour les mesures vocaliques ($p < 0,05$), à l'exclusion de VarcoV (tableau 14). Par ailleurs, la tendance diffère s'agissant de %V par rapport aux observations relatives à la parole avec disfluences, avec cette fois une moyenne et une médiane plus élevées chez les sujets qui bégayaient, et une dispersion plus marquée des valeurs chez les sujets contrôle ; la différence n'est toutefois pas significative.

	%V	ΔV	ΔC	VarcoV	VarcoC	r-PVI-V	r-PVI-C	n-PVI-V	n-PVI-C
<i>U</i>	5	1	7	8	10	0	6	2	7
<i>z</i>	-1,56	-2,4	-1,14	-0,94	-0,52	-2,61	-1,35	-2,19	-1,14
<i>p</i>	0,11	0,01*	0,25	0,34	0,6	0,009*	0,17	0,02*	0,25

Tableau 14. Comparaison des valeurs obtenues par les sujets pathologiques et les sujets contrôle pour chaque mesure rythmique segmentale après élimination des disfluences au moyen du test *U* bilatéral de Mann-Whitney.

1.3. Comparaison des résultats par groupe de sujets

Considérons enfin les résultats par groupe de sujets. Comme on peut le voir dans les tableaux 12 et 13 *supra*, les valeurs médianes des différentes mesures sont identiques avant et après élimination des disfluences dans le groupe des sujets contrôle, ce qui exclut la comparaison au moyen du test des rangs signés de Wilcoxon. Ce résultat est attendu, seul l'un des sujets ayant produit des disfluences durant la tâche de lecture, et ce de manière marginale (voir *supra*, chapitre 2, section 2.2). En accord avec les observations de la section précédente, la différence n'est par ailleurs significative dans le groupe des sujets pathologiques que pour les mesures de variabilité des durées consonantiques sans normalisation pour le débit ($p < 0,05$, tableau 15).

	%V	ΔV	ΔC	VarcoV	VarcoC	r-PVI-V	r-PVI-C	n-PVI-V	n-PVI-C
<i>T</i>	6	4	0	6	1	4	0	4	3
<i>z</i>	-0,4	-0,94	-2,02	-0,04	-1,75	-0,94	-2,02	-0,94	-1,21
<i>p</i>	0,68	0,34	0,04*	0,68	0,07	0,34	0,04*	0,34	0,22

Tableau 15. Comparaison des valeurs obtenues par les sujets pathologiques pour chaque mesure rythmique segmentale avant et après élimination des disfluences au moyen du test *T* des rangs signés de Wilcoxon bilatéral pour échantillons appariés.

1.4 Discussion

Quel bilan dresser des résultats présentés ? Conformément à nos hypothèses, la parole des sujets qui bégaièrent se caractérise de manière générale par une variabilité des durées vocaliques et consonantiques plus marquée que celle des sujets contrôle. En outre, cette variabilité excessive tend à se maintenir en l'absence de disfluences, en accord avec les résultats de Maruthy *et al.* (2017), et plus généralement avec les théories du bégaiement fondées sur l'hypothèse d'une instabilité motrice d'origine neurologique centrale (Zimmermann 1980, Packman & Onslow 1996). Plusieurs précisions s'imposent toutefois. En premier lieu, les statistiques inférentielles utilisées sont à considérer avec prudence au regard des faibles effectifs considérés et surtout de l'hétérogénéité des profils au sein de l'échantillon de sujets pathologiques. Deuxièmement, la normalisation pour le débit semble peu utile : elle réduit la dispersion, mais ne gomme pas le contraste essentiel entre les deux groupes. Enfin, variabilité vocalique et variabilité consonantique n'ont pas le même poids, ce qui explique sans doute la faible sensibilité de la mesure %V. Comme l'illustrent bien les figures 10 et 11 ci-dessous, la variabilité des durées vocaliques discrimine bien les sujets qui bégaièrent des sujets contrôle, ce qui concorde avec les observations antérieures de Onslow *et al.* (1992) et Packman *et al.* (1994) sur le bégaiement en anglais. Par contraste, la variabilité des durées consonantiques n'est clairement distinctive que pour certains d'entre eux, les sujets P4 et P5, sans corrélation évidente avec le degré de sévérité, du moins dans l'échantillon considéré. Ainsi, le sujet P1 dont la parole comporte une variabilité des durées consonantiques nettement moindre présente un bégaiement tout aussi sévère (voir *supra*, chapitre 2, section 2.1.2, tableau 3). L'explication est sans doute à rechercher dans la nature des disfluences : le sujet P1 produit essentiellement des répétitions de segments ou de syllabes, tandis que les sujets P4 et P5 produisent essentiellement des prolongations et des blocages. Corrélativement, seuls ces derniers sujets présentent une parole véritablement proche du modèle rythmique des langues dites « à fort degré d'accentuation » sur le plan quantitatif (voir *supra*, chapitre 2, section 2.3.1, figure 6b), un point sur lequel nous revenons en conclusion de ce mémoire.

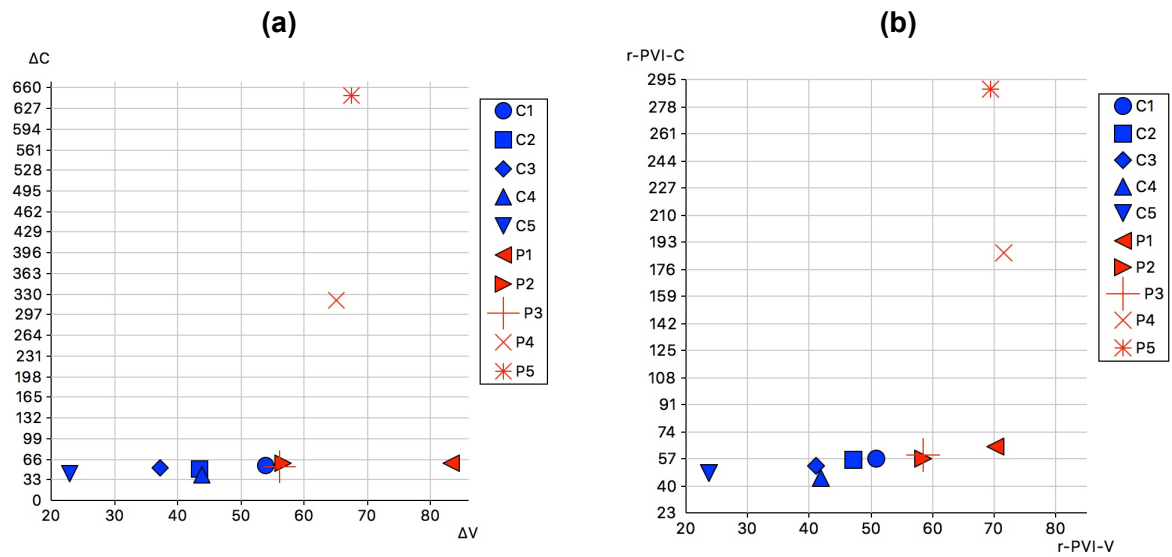


Fig. 10. Classifications des sujets pathologiques et contrôle selon différentes mesures rythmiques segmentales appliquées à la parole avant élimination des disfluences.

Légende : en (a) l'espace obtenu en combinant les écarts-types de durées des intervalles vocaliques (ΔV , x) et consonantiques (ΔC , y); en (b) l'espace obtenu en combinant les indices de variabilité par paires bruts de durées des intervalles vocaliques ($r\text{-PVI-V}$, x) et consonantiques ($r\text{-PVI-C}$, y).

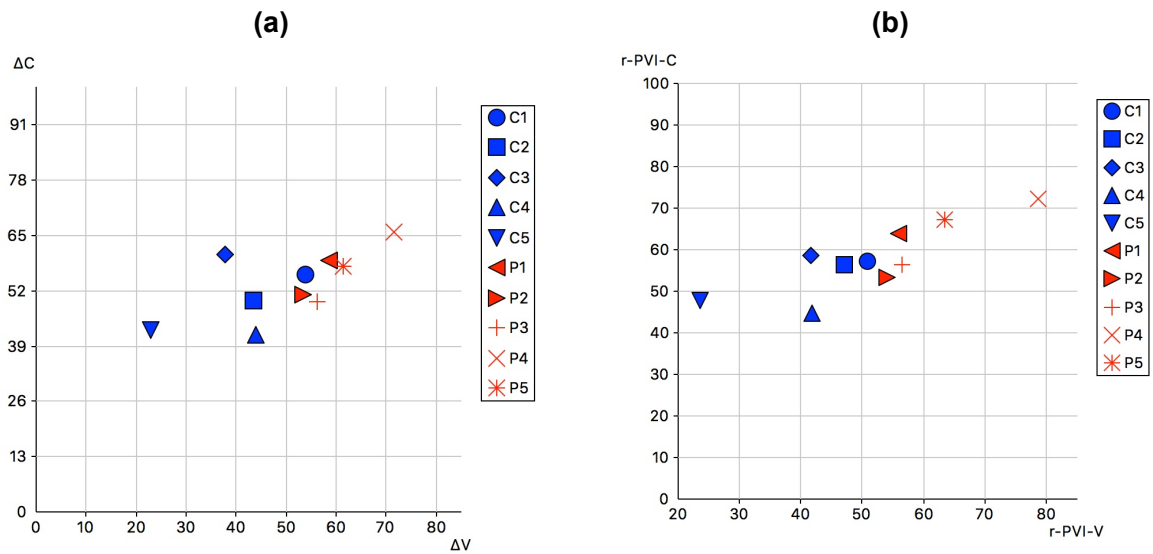


Fig. 11. Classifications des sujets pathologiques et contrôle selon différentes mesures rythmiques segmentales appliquées à la parole après élimination des disfluences.

Légende : en (a) l'espace obtenu en combinant les écarts-types de durées des intervalles vocaliques (ΔV , x) et consonantiques (ΔC , y); en (b) l'espace obtenu en combinant les indices de variabilité par paires bruts de durées des intervalles vocaliques ($r\text{-PVI-V}$, x) et consonantiques ($r\text{-PVI-C}$, y).

2. Résultats supra-segmentaux

2.1. Résultats avant élimination des disfluences

Sur le modèle de la section précédente, nous considérons d'abord la parole de manière globale, avant élimination des disfluences. Les valeurs obtenues par les sujets pathologiques et les sujets contrôle pour chaque mesure rythmique supra-segmentale sont rassemblées dans les tableaux 16 et 17, et résumées dans la figure 12 sous la forme de boîtes à moustaches.

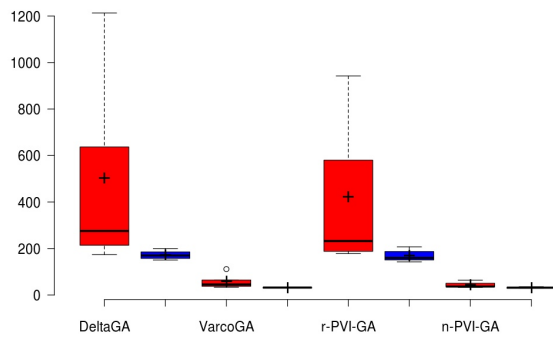
	Δ GA	VarcoGA	r-PVI-GA	n-PVI-GA	Min(GA)	Max(GA)
P1	275,45	45,23	232,03	35,99	282,71	1585,48
P2	173,60	33,07	178,04	36,39	127,29	930,19
P3	214,03	37,55	187,85	33,02	220,22	1061,04
P4	636,85	64,37	580,16	50,34	285,30	3015,49
P5	1213,11	111,54	942,06	63,69	256,77	7593,67
Médiane	275,45	45,23	232,03	36,39	256,77	1585,48
Moyenne	502,61	58,35	424,03	43,88	234,46	2837,17
Écart-type	437,60	32,05	333,87	12,94	65,38	2784,51

Tableau 16. Valeurs des mesures rythmiques supra-segmentales avant élimination des disfluences chez les sujets pathologiques (P), exprimées en millisecondes (Δ GA, r-PVI-GA, Min(GA), Max(GA)) ou en pourcentages (VarcoGA, n-PVI-GA) arrondis à la deuxième décimale.

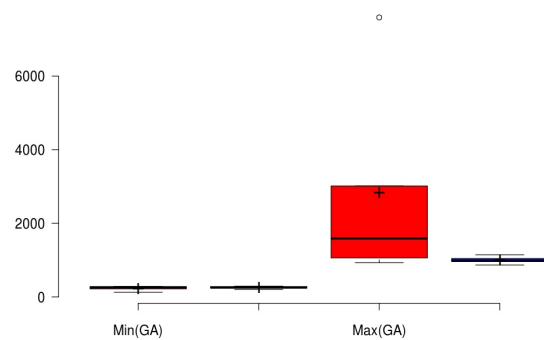
	Δ GA	VarcoGA	r-PVI-GA	n-PVI-GA	Min(GA)	Max(GA)
C1	199,92	33,87	207,10	35,22	240,67	1042,96
C2	185,33	32,43	186,91	32,69	281,95	1144,77
C3	170,06	30,29	159,48	29,04	294,10	979,59
C4	150,42	31,97	142,32	31,61	202,42	868,47
C5	157,48	29,76	150,88	29,09	256,07	967,68
Médiane	170,06	31,97	159,48	31,61	256,07	979,59
Moyenne	172,64	31,66	169,34	31,53	255,04	1000,70
Écart-type	20,22	1,66	26,94	2,60	36,15	101,93

Tableau 17. Valeurs des mesures rythmiques supra-segmentales avant élimination des disfluences chez les sujets contrôle (C), exprimées en millisecondes (Δ GA, r-PVI-GA, Min(GA), Max(GA)) ou en pourcentages (VarcoGA, n-PVI-GA) arrondis à la deuxième décimale.

(a) Mesures de variabilité temporelle des groupes accentuels



(b) Durées minimale et maximale des groupes accentuels



■ Sujets pathologiques ■ Sujets contrôle

Fig. 12. Boîtes à moustaches illustrant la distribution des valeurs pour chaque mesure rythmique supra-segmentale avant élimination des disfluences chez les sujets pathologiques et les sujets contrôle.

Légende : en (a) les mesures de variabilité temporelle des groupes accentuels, en (b) les mesures de durées minimale et maximale des groupes accentuels. Le rectangle principal de chaque boîte à moustaches correspond aux 50% des effectifs les plus proches de la médiane de l'échantillon considéré, c'est-à-dire à l'espace entre le premier quartile (Q1) délimité par le trait inférieur du rectangle, et le troisième quartile (Q3) délimité par le trait supérieur. Les « moustaches » haute et basse, quant à elles, correspondent aux 50% restants à l'exclusion des valeurs « atypiques » supérieures ($>Q3+1,5*(Q3-Q1)$) et inférieures ($<Q1-1,5*(Q3-Q1)$), représentées par le symbole °. La valeur moyenne est indiquée par le symbole + et la médiane par la barre noire horizontale.

Que l'on s'appuie à nouveau sur la variance (Δ GA, VarcoGA) ou sur la comparaison par paires (r-PVI-GA, n-PVI-GA), on retrouve au niveau supra-segmental du groupe accentuel les mêmes tendances qu'au niveau segmental. D'une part, la variabilité temporelle est significativement plus élevée chez les sujets pathologiques que chez les sujets contrôle ($p<0,05$), ce que ne gomme pas fondamentalement la normalisation pour le débit (tableau 18). D'autre part, la même partition est observable au sein du groupe pathologique entre les sujets P4 et P5 d'un côté, où les valeurs sont maximalelement élevées, et les sujets P1, P2 et P3, où elles sont plus modérées.

	Δ GA	VarcoGA	r-PVI-GA	n-PVI-GA	Min(GA)	Max(GA)
U	2	1	3	1	12	5
z	-2,19	-2,4	-1,98	-2,4	-0,1	-1,56
p	0,02*	0,01*	0,04*	0,01*	0,91	0,11

Tableau 18. Comparaison des valeurs obtenues par les sujets pathologiques et les sujets contrôle pour chaque mesure rythmique supra-segmentale avant élimination des disfluences au moyen du test U bilatéral de Mann-Whitney.

Considérons par ailleurs les mesures Min(GA) et Max(GA). Il est vrai que l'analyse statistique ne montre pas de différence significative entre les deux groupes (tableau 18). Il n'est pas clair toutefois que ce type de comparaisons basées sur les valeurs médianes fasse

sens lorsque l'on considère des contraintes temporelles d'allure catégorique comme les durées minimale et maximale des groupes accentuels (voir chapitre 1, section 1.2.2) ; il faut dans ce cas regarder les données dans le détail. À la lecture du tableau 17, on constate que les valeurs obtenues par les sujets contrôle respectent toutes (à 50ms près dans un cas isolé) les seuils inférieur et supérieur de 250 et 1250 ms établis par Martin (2014) sur la base de grands corpus de parole. Dans le tableau 16, par contraste, le seuil supérieur est largement dépassé par les sujets pathologiques P1, P4 et P5, c'est-à-dire ceux dont le bégaiement est le plus sévère (voir *supra*, section 1.4), ce qui élève les valeurs médiane et moyenne de Max(GA) et crée de la dispersion dans la figure 12b. En revanche, le seuil inférieur est, lui, respecté, à une exception près (P2), un point sur lequel nous revenons plus loin (voir *infra*, section 2.4).

2.2. Résultats après élimination des disfluences

Considérons à présent la parole après élimination des disfluences. Les valeurs obtenues par les sujets pathologiques et les sujets contrôle pour chaque mesure rythmique supra-segmentale sont rassemblées dans les tableaux 19 et 20, et résumées par les boîtes à moustaches de la figure 13.

	Δ GA	VarcoGA	r-PVI-GA	n-PVI-GA	Min(GA)	Max(GA)
P1	162,54	30,08	177,61	33,65	282,71	950,45
P2	158,43	29,25	162,45	30,31	298,88	885,88
P3	168,08	31,78	142,17	26,35	303,93	956,42
P4	232,70	32,42	220,86	32,61	285,30	1211,87
P5	246,58	37,64	257,29	38,94	341,01	1300,07
Médiane	168,08	31,78	177,61	32,61	298,88	956,42
Moyenne	193,66	32,23	192,08	32,37	302,36	1060,94
Écart-type	42,39	3,28	46,53	4,62	23,37	182,85

Tableau 19. Valeurs des mesures rythmiques supra-segmentales après élimination des disfluences chez les sujets pathologiques (P), exprimées en millisecondes (Δ GA, r-PVI-GA, Min(GA), Max(GA)) ou en pourcentages (VarcoGA, n-PVI-GA) arrondis à la deuxième décimale.

	Δ GA	VarcoGA	r-PVI-GA	n-PVI-GA	Min(GA)	Max(GA)
C1	199,92	33,87	207,10	35,22	240,67	1042,96
C2	185,33	32,43	186,91	32,69	281,95	1144,77
C3	159,99	30,29	152,60	29,57	294,10	910,31
C4	150,42	31,97	142,32	31,61	202,42	868,47
C5	157,48	29,76	150,88	29,09	256,07	967,68
Médiane	159,99	31,97	152,60	31,61	256,07	967,68
Moyenne	170,63	31,66	167,96	31,63	255,04	986,84
Écart-type	21,02	1,66	27,73	2,48	36,15	109,91

Tableau 20. Valeurs des mesures rythmiques supra-segmentales après élimination des disfluences chez les sujets contrôle (C), exprimées en millisecondes (Δ GA, r-PVI-GA, Min(GA), Max(GA)) ou en pourcentages (VarcoGA, n-PVI-GA) arrondis à la deuxième décimale.

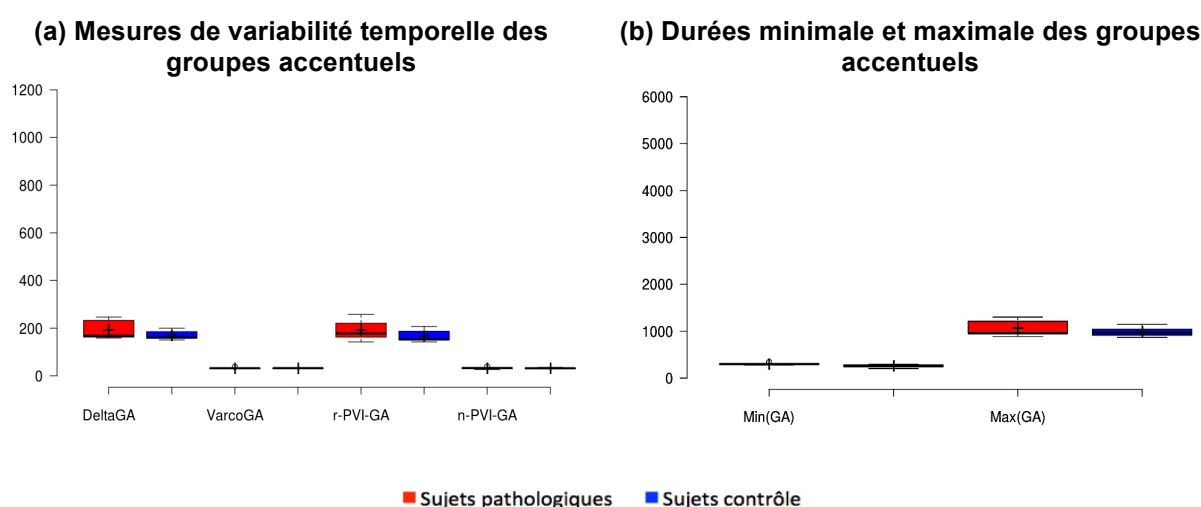


Fig. 13. Boîtes à moustaches illustrant la distribution des valeurs pour chaque mesure rythmique supra-segmentale après élimination des disfluences chez les sujets pathologiques et les sujets contrôle.

Légende : en (a) les mesures de variabilité temporelle des groupes accentuels, en (b) les mesures de durées minimale et maximale des groupes accentuels. Le rectangle principal de chaque boîte à moustaches correspond aux 50% des effectifs les plus proches de la médiane de l'échantillon considéré, c'est-à-dire à l'espace entre le premier quartile (Q1) délimité par le trait inférieur du rectangle, et le troisième quartile (Q3) délimité par le trait supérieur. Les « moustaches » haute et basse, quant à elles, correspondent aux 50% restants à l'exclusion des valeurs « atypiques » supérieures ($>Q3 + 1,5 \cdot (Q3 - Q1)$) et inférieures ($< Q1 - 1,5 \cdot (Q3 - Q1)$), représentées par le symbole °. La valeur moyenne est indiquée par le symbole + et la médiane par la barre noire horizontale.

Alors que la variabilité temporelle brute reste élevée chez les sujets P4 et P5, elle diminue chez les autres sujets pathologiques pour atteindre des valeurs comparables à celles des sujets contrôle, ce que ne représentent pas bien ici les boîtes à moustaches. En outre, la normalisation pour le débit gomme complètement le contraste, et aucune comparaison statistique ne donne de résultat significatif sur le plan statistique (tableau 21). Enfin, les seuils de durée minimale et maximale des groupes accentuels sont respectés (à 50 ms près dans deux cas) par tous les sujets, et ce dans les deux groupes. On remarque seulement que les valeurs de Min(GA) sont significativement plus élevées chez les sujets qui bégayaient

que chez les sujets contrôle, ce qui ne reçoit pas d'explication évidente, les valeurs de Max(GA) étant, elles, du même ordre dans les deux groupes.

	Δ GA	VarcoGA	r-PVI-GA	n-PVI-GA	Min(GA)	Max(GA)
U	7	11	9	11	2	10
z	-1,14	-0,31	-0,73	-0,31	-2,19	-0,52
p	0,25	0,75	0,46	0,75	0,02*	0,6

Tableau 21. Comparaison des valeurs obtenues par les sujets pathologiques et les sujets contrôle pour chaque mesure rythmique supra-segmentale après élimination des disfluences au moyen du test U bilatéral de Mann-Whitney.

2.3. Comparaison des résultats par groupe de sujets

La comparaison des résultats avant et après élimination des disfluences n'est pas réalisable au sein du groupe de sujets contrôle au moyen du test de Wilcoxon ; les valeurs médianes sont trop proches. La différence est en revanche significative au sein du groupe de sujets pathologiques pour toutes les mesures, excepté Min(GA).

	Δ GA	VarcoGA	r-PVI-GA	n-PVI-GA	Min(GA)	Max(GA)
T	0	0	0	0	1	0
z	-2,02	-2,02	-2,02	-2,02	-1,46	-2,02
p	0,04*	0,04*	0,04*	0,04*	0,14	0,04*

Tableau 22. Comparaison des valeurs obtenues par les sujets pathologiques pour chaque mesure rythmique supra-segmentale avant et après élimination des disfluences au moyen du test T des rangs signés de Wilcoxon bilatéral pour échantillons appariés.

2.4. Discussion

À la différence des mesures segmentales dont l'interprétation est parfois délicate dans le détail, les mesures supra-segmentales considérées donnent des résultats univoques. D'une part, le phrasé des sujets pathologiques se caractérise par une variabilité temporelle significativement plus marquée que celle observée chez les sujets contrôle en présence de disfluences, ce qui concorde avec les observations antérieures de Bergmann (1986) sur le bégaiement en allemand. D'autre part, la durée des groupes accentuels peut largement dépasser le seuil de 1250 ms quand interviennent des répétitions de segments, des blocages ou encore des prolongations, ce qui constitue une violation manifeste de la contrainte « *no lapse* » communément admise en phonologie métrique. Par contraste, la contrainte « *no clash* » dans la formulation temporelle qu'en propose Martin (2014, 2017, 2018) semble plus robuste, ce qui demanderait à être confirmé sur un échantillon de sujets plus important. Un seul exemple dans notre corpus y contrevient chez le sujet P2, avec un

groupe accentuel présentant une durée aberrante de 127,29 ms à la suite d'un premier groupe de 223,90 ms dans le contexte d'une répétition de syllabes (figure 14). Le découpage en groupes accentuels retenu peut cependant être discuté (voir *supra*, chapitre 2, section 2.4.2). Enfin, contrairement à ce que nous avons observé dans le domaine segmental, aucune anomalie ne se retrouve de manière franche en l'absence de disfluences à partir du moment où le débit est normalisé : en accord avec notre analyse perceptive, le phrasé en groupes accentuels des sujets qui bégayaient est alors comparable à celui des sujets contrôle.

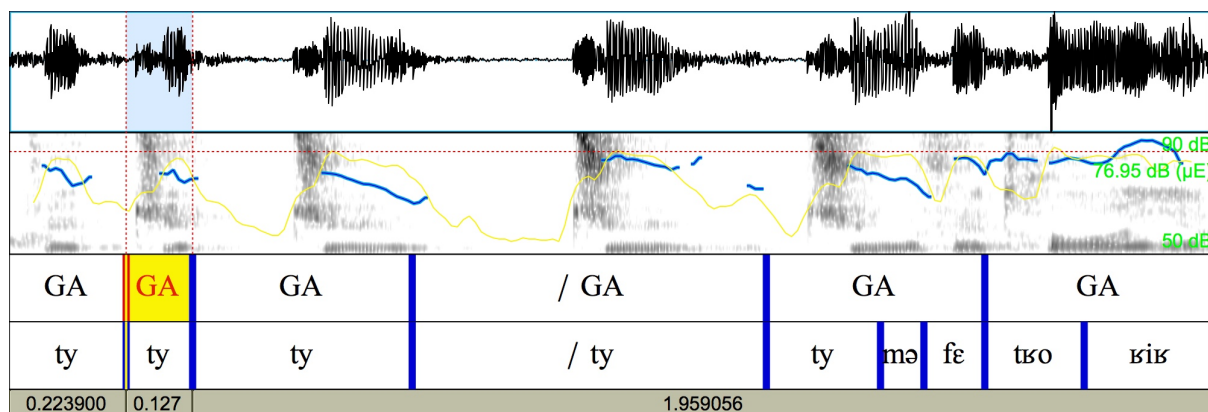


Fig. 14. Oscillogramme, spectrogramme à bandes larges et courbes d'intensité (en jaune) et de F0 (en bleu) associés à la séquence « tu tu tu / tu tu me fait trop rire », produite par le sujet P2.

CONCLUSION

La présente étude constitue à notre connaissance la première à explorer au moyen de méthodes quantitatives les perturbations du rythme de la parole chez les sujets qui bégaiement en français. Les résultats obtenus sont de deux ordres.

Sur le plan supra-segmental d'abord, les mesures considérées confortent l'hypothèse de Monfrais-Pfauwadel (2014) suivant laquelle les disfluences caractéristiques du bégaiement perturbent le phrasé en groupes accentuels ou « *chunking* » de la parole. Ainsi, l'eurythmie est typiquement rompue, se traduisant par une variabilité temporelle excessive des groupes accentuels, mais aussi par des violations des contraintes rythmiques qui gouvernent l'alternance entre battements faibles et battements forts. Les disfluences peuvent clairement éloigner deux accents métriques successifs au-delà de la distance maximale de 1250 ms caractéristique de la parole normale (Martin 2014, 2017, 2018). De manière marginale, elles semblent également pouvoir parfois rapprocher deux accents en-deçà du seuil de 250 ms ; une étude complémentaire sur un échantillon plus large de sujets serait nécessaire pour le confirmer.

Sur le plan segmental par ailleurs, la parole des sujets qui bégaiement se caractérise par une variabilité temporelle plus marquée des intervalles vocaliques et consonantiques, en accord avec les théories du bégaiement fondées sur l'hypothèse d'une instabilité motrice d'origine neurologique centrale (Zimmermann 1980, Packman & Onslow 1996). Les résultats diffèrent toutefois de ceux obtenus dans le domaine supra-segmental par plusieurs aspects. En premier lieu, la variabilité excessive constatée ne peut être directement imputée aux disfluences, comme c'est le cas au niveau des groupes accentuels : elle se maintient en effet en leur absence, à l'image d'autres propriétés segmentales distinctives comme le « *Voice Onset Time* » ou délai d'établissement du voisement. D'autre part, le poids des mesures vocaliques et consonantiques n'est pas le même : alors que la variabilité temporelle des intervalles vocaliques distingue de manière générale les sujets qui bégaiement des sujets contrôle, la variabilité temporelle des intervalles consonantiques ne joue un rôle que chez certains sujets. Enfin et surtout, on peut se demander dans quelle mesure cette variabilité reflète véritablement une différence de nature rythmique entre les deux groupes étudiés. Comme on l'a vu, seule la parole de certains sujets qui bégaiement, ceux qui produisent des blocages et/ou des prolongations de consonnes, s'apparente sur le plan quantitatif au modèle des langues dites « à fort degré d'accentuation » ou « isochronie accentuelle ». Une étude perceptive serait toutefois nécessaire pour le confirmer. Comme le souligne Arvaniti

(2009, 2012), les variations de durées vocaliques et consonantiques dépendent en effet de multiples facteurs, et pas uniquement de la complexité syllabique et de la réduction vocalique supposées fixer les classes rythmiques : d'autres indices acoustiques doivent être pris en compte (F0, intensité, timbre). Du reste, il n'est pas clair, comme le note ce même auteur, que ces deux facteurs suffisent pour rendre compte des partitions pertinentes.

Pour conclure, nous voudrions préciser quelques-uns des développements qu'appelle ce travail. Il serait intéressant d'abord de consolider nos résultats sur le corpus étudié au moyen de méthodes statistiques bayésiennes plus adaptées aux petits effectifs, après vérification de la fiabilité des segmentations réalisées. Sur le plan supra-segmental, par ailleurs, un prolongement naturel consisterait à explorer les contraintes d'eurythmie chez les sujets qui bégayaient en couplant l'approche acoustique à une analyse électro-encéphalographique (EEG), à la recherche d'anomalies dans les variations de période des ondes Delta. Enfin, les mesures de variabilité temporelle des intervalles vocaliques présentent un attrait évident sur le plan clinique à partir du moment où elles discriminent aussi bien les sujets qui bégayaient en présence de disfluences qu'en leur absence, ce qui demanderait à être confirmé, corrélé au degré de sévérité, puis normé sur de plus vastes échantillons.

BIBLIOGRAPHIE

Abercrombie, D. 1967. Elements of general phonetics. Edimbourg : Edinburgh University Press.

Agnello, J. 1975. Voice onset and voice termination features of stutterers. In Webster, L. M. & L. C. Furst (eds), Vocal tract dynamics and dysfluency. New York : Speech and Hearing Institute.

Almada, A., Simoes, R., Constante, M., Casquinha, P. & M. J. Heitor. 2016. Psychogenic stuttering : a case and review. European Psychiatry 33, S386-S387.

Amster, B. J. & C. W. Starkweather. 1987. Articulatory rate, stuttering and speech motor control. In Peters, H. F. M. & W. Hulstijn (eds), Speech Motor Dynamics in Stuttering. Vienne : Springer, 317-328.

Andrews, G., Howie, P., Dosza, M. & B. J. Guitar. 1982. Stuttering : speech pattern characteristics under fluency inducing conditions. Journal of Speech and Hearing Research 25, 208-216.

Arbisi-Kelm, T. R. 2006. An intonational analysis of disfluency patterns in stuttering. Thèse de doctorat, Université de Californie - Los Angeles.

Arbisi-Kelm, T. R. 2010. Intonation structure and disfluency detection in stuttering. Papers in Laboratory Phonology 10, 405–432.

Arvaniti, A. 2009. Rhythm, timing and the timing of rhythm. *Phonetica* 66:1-2, 46-63.

Arvaniti, A. 2012. The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics* 40, 351-373.

ASHA. 2015. Childhood fluency disorders. Rockville, MD : American Speech-Language-Hearing Association. Article disponible à l'adresse : <https://www.asha.org/Practice-Portal/Clinical-Topics/Childhood-Fluency-Disorders/>. [Page consultée le 26/03/2020]

Astésano, C. 2001. Rythme et accentuation en français : invariance et variabilité stylistique. Paris : L'Harmattan.

Asu, E. L. & F. Nolan. 2006. Estonian and English rhythm : a two-dimensional quantification based on syllables and feet. In *Proceedings of Speech Prosody 2006*, Dresden, 2-5 mai 2006, 249-252.

Aumont-Boucand, V. 2012. Le bredouillement, savoir l'identifier pour pouvoir le soigner. Les entretiens de Bichat 2012, 5-8.

Avanzi, M., Simon, A.C., Goldman, J.-P. & A. Auchlin. 2010. C-PROM. An annotated corpus for French prominence studies. In *Proceedings of of Prosodic Prominence : Perceptual and Automatic Identification (Speech Prosody Workshop)*, Chicago, 10 mai 2010.

Ballard, A. P., Azizi, L., Duffy, J. R., McNeil, M. R., Halaki, M., O'Dwyer, N., Scholl, D., Vogel, A. P. & D. A. Robin. 2016. A predictive model for diagnosing stroke-related apraxia of speech. *Neuropsychologia* 81, 129–139.

Béné, J., Auffret, M., Fedrizzi, S., Valnet-Rabier, M. B., Caron, J., Gautier, S., & the French Association of Regional Pharmacovigilance Centres. 2014. Drug-induced stuttering : a review of the French pharmacovigilance database. Article disponible en ligne à l'adresse : <http://www.pharmacovigilance-ndpc.fr>. [Page consultée le 25/03/2020]

Bergmann, G. 1986. Studies in stuttering as a prosodic disturbance. *Journal of Speech, Language and Hearing Research* 29:3, 290-300.

Blomgren, M., Robb, M. P., & Y. Chen. 1998. A note on vowel centralization in stuttering and nonstuttering individuals. *Journal of Speech, Language, and Hearing Research* 41:5, 1042–1051.

Blomgren, M., Alqhazo, M., & E. Metzger. 2012. Do speech sound characteristics really influence stuttering frequency ? In *Proceedings of the 7th World Congress of Fluency Disorders*, Tours, 2-5 juillet 2012.

Bloodstein, O. 1944. Studies in the psychology of stuttering : XIX. The relationship between oral reading and severity of stuttering. *Journal of Speech Disorders* 9:2, 161–173.

Bloodstein, O. & N. Bernstein-Ratner. 2008. A handbook of stuttering (6th ed.). New York : Delmar.

Boersma, P. & D. Weenink. 2017. Praat : doing phonetics by computer [Computer program]. Version 6.0.29. Logiciel disponible en téléchargement en ligne à l'adresse suivante : https://www.fon.hum.uva.nl/praat/download_mac.html [Page consultée le 06/07/2020]

Bolinger, D. 1965. Pitch accent and sentence rhythm. In Abe, I. & T. Kanekiyo (eds), *Forms of English : accent, morpheme, order*. Harvard : Harvard University Press, 139-180.

Brady, J. P. 1998. Drug-induced stuttering : a review of the literature. *Journal of Clinical Psychopharmacology* 18:1, 50-54.

Brown, S. F. 1938. Stuttering with relation to word accent and word position. *Journal of Abnormal and Social Psychology* 33:1, 112-120.

Campbell, J. G. & D. G. Hill. 1994. *Systematic disfluency analysis*. Evanston : Northwestern University.

Chang, S. E., Erikson, K. I., Ambrose, N. G., Hasegawa-Johnson, M. A. & C. L. Ludlow. 2008. Brain anatomy differences in childhood stuttering. *Neuroimage* 39, 1333-1344.

Chang, S. E., Zhu, D. C., Choo, A. L. & M. Angstadt. 2015. White matter neuroanatomical differences in young children who stutter. *Brain* 138, 694-711.

Cichocki, W. 2015. The timing of accentual phrases in read and spontaneous speech : data from acadian French. *Canadian Acoustics* 43:3, 109-110.

Classe, A. 1939. *The rhythm of English prose*. Oxford : Blackwell.

Conture, E. G., McCall, G. N., & D. W. Brewer. 1977. Laryngeal behavior during stuttering. *Journal of Speech and Hearing Research* 20:4, 661-668.

Cooper, M. H. & G. D. Allen. 1977. Timing control accuracy in normal speakers and stutterers. *Journal of Speech and Hearing Research* 20:1, 55-71.

Craig, A. (ed). 2010. The influence of fluency disorders on quality of life. *Journal of Fluency Disorders* 35:3, 159-332.

Cruz, C., Amorim, H., Beca, G. & R. Nunes. 2018. Neurogenic stuttering : a review of the literature. *Revue neurologique* 66:2, 59-64.

Cumming, R. E. 2010. Speech rhythm : the language specific integration of pitch and duration. Thèse de doctorat, Université de Cambridge.

Darley, F. L., Aronson, A. E. & J. R. Brown. 1969a. Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research* 12, 246-269.

Darley, F. L., Aronson, A. E. & J. R. Brown. 1969b. Clusters of deviant speech dimensions in the dysarthrias. *Journal of Speech and Hearing Research* 12, 462-496.

Dasher, R. & D. Bolinger. 1982. On pre-accentual lengthening. *Journal of the International Phonetic Association* 12, 58-69.

Dauer, R. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11, 51-62.

Dauer, R. 1987. Phonetic and phonological components of language rhythm. In *Proceedings of the 11th International Congress of Phonetic Sciences*, Tallinn, 1-7 Août 1987, 447-449.

De Sonnevile-Koedoot, C., Stolk, E., Rietveld, T. & M. C. Franken. 2015. Direct versus indirect treatment for preschool children who stutter : the RESTART randomized trial. *PloS One* 10.

Delattre, P. 1966. A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics* 4, 183-198.

Delais-Roussarie, E. & A. Di Cristo. À paraître. L'accentuation. In Abeillé, A. & D. Godard (eds), *Grande Grammaire du Français*, chapitre XIX. Arles : Actes Sud / Imprimerie Nationale.

Dell, F. 1984. L'accentuation dans les phrases en français. In Dell, F., Hirst, D. & J.-R. Vergnaud (eds), *Forme sonore du langage*. Paris : Hermann, 65–112.

Dellwo, V. & P. Wagner. 2003. Relations between language rhythm and speech rate. In *Proceedings of the 15th International Congress of Phonetics Sciences*, Barcelone, 3-9 août 2003, 471-474.

Desportes, E. & M. Meyer. 2014. Validation d'un test prédictif et d'une batterie d'évaluation du bredouillement. Mémoire pour l'obtention du Certificat de Capacité d'Orthophoniste, Université Claude Bernard - Lyon 1.

Dhuisme, F. 2007. Pour essayer de connaître le bredouillement. Mémoire de DU « Bégaiements et autres troubles de la fluence », Université Pierre et Marie Curie - Paris 6.

Di Cristo, A. 1999. Le cadre accentuel du français contemporain : essai de modélisation. *Langues* 2:3, 184–205 & *Langues* 2:4, 258–267.

Di Cristo, A. 2003. De la métrique et du rythme de la parole ordinaire : l'exemple du français. *Semen* 16. Article disponible en téléchargement en ligne à l'adresse suivante : <https://journals.openedition.org/semn/2944> [Page consultée le 06/07/2020]

Di Cristo, A. 2011. Une approche intégrative des relations de l'accentuation au phrasé prosodique du français. *Journal of French Language Studies* 21:1, 73-95.

Didirková, I. 2016. Parole, langues et disfluences : une étude linguistique et phonétique du bégaiement. Thèse de doctorat, Université Paul-Valéry - Montpellier 3.

Didirková, I. & F. Hirsch. 2019. A two-case study of coarticulation in stuttered speech. An articulatory approach. *Clinical Linguistics and Phonetics* 34:6, 517-535.

Didirková, I., Le Maguer, S. & F. Hirsch. 2020. An articulatory study of differences and similarities between stuttered disfluencies and non-pathological disfluencies. *Clinical Linguistics and Phonetics* 35:1, 1-21.

Duez, D. & Y. Nishinuma. 1985. Le rythme en français : alternance de durées syllabiques. *Travaux de l'institut de phonétique d'Aix* 10, 151-169.

Etchell, A. C., Civier, O., Ballard, K. J. & P. F. Sowman. 2017. A systematic literature review of neuroimaging research on developmental stuttering between 1995 and 2016. *Journal of Fluency Disorders* 55, 6-45.

Grabe, E. & E. L. Low. 2002. Durational variability in speech and the rhythm class hypothesis. In Gussenhover, C. & N. Warner (eds), *Papers in Laboratory Phonology 7*. Berlin : Mouton de Gruyter, 515-546.

Falk, S., Maslow, E., G. Thum & P. Hoole. 2016. Temporal variability in sung production of adolescents who stutter. *Journal of Communication Disorders* 62:4, 101-114.

Fant, G., Kruckenberg, A. & L. Nord. 1991. Durational correlates of stress in Swedish, French and English. *Journal of Phonetics* 19, 351-365.

Fónagy, I. 1980. L'accent français : accent probabilitaire. In Fónagy, I. & P. R Léon (eds), *L'Accent en français contemporain*, *Studia Phonetica* 15. Paris : Didier, 123–233.

Fougeron, C. & P. Keating. 1997. Articulation strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101:6, 3728-3740.

Fraisse, P. 1956. *Les structures rythmiques*. Louvain : Éditions Universitaires.

Fraisse, P. 1974. *Psychologie du rythme*. Paris : PUF.

Frigerio-Domingues, C. & D. Drayna. 2017. Genetic contributions to stuttering : the current evidence. *Molecular Genetics and Genomic Medicine* 5, 95-102.

Gilbert, A. 2012. *Le chunking perceptif de la parole : sur la nature du groupement temporel et son effet sur la mémoire immédiate*. Thèse de doctorat, Université de Montréal.

Grosjean, F. & A. Deschamps. 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.

Hayes, B. 1984. The phonology of rhythm in English. *Linguistic Inquiry* 14, 33-74.

Hirsch, F. 2007. Le bégaiement : perturbations de l'organisation temporelle de la parole. Thèse de doctorat, Université Marc Bloch - Strasbourg 2.

Hubbard, C. P. 1998. Stuttering, stressed syllables and word onsets. *Journal of Speech, Language and Hearing Research* 41:4, 802-808.

Hudock, D. & J. Kalinowski. 2014. Stuttering inhibition via auditory feedback during scripted telephone conversations. *International Journal of Language and Communication Disorders* 49, 139-147.

Howell, P. & L. Vause, L. 1986. Acoustic analysis and perception of vowels in stuttered speech. *The Journal of the Acoustical Society of America* 79:5, 1571–9.

Ingham, R. J., Ingham, J. C., Bothe, A. K., Wang Y. & M. Kilgo. 2015. Efficacy of the Modifying Phonation intervals (MPI) stuttering program with adults who stutter. *American Journal of Speech-Language Pathology* 24:2, 256-271.

Kelman, E. & A. Nicholas. 2008. Practical intervention for early childhood stammering. Milton Keynes : Speechmark Publishing Ltd.

Kraft, S. J. & E. Yairi. 2012. Genetic bases of stuttering : the state of the art. *Folia Phoniatica et Logopaedica* 64:1, 34-47.

Ladefoged, P. 1975. A course in phonetics, New York : Harcourt Brace Jovanovich.

Lehiste, I. 1977. Isochrony reconsidered. *Journal of Phonetics* 5:3, 253-263.

Liberman, M. Y. & A. Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8:2, 249-336.

Liss, J. M., White L., Mattys S. L., Lansford K., Lotto A. J. & S. M. Spitzer. 2009. Quantifying speech rhythm abnormalities in the dysarthrias. *Journal of Speech, Language, and Hearing Research* 52:5, 1334–1352.

Llyod James, A. 1940. Speech signals in telephony. Londres : Sir Isaac Pitman & Sons.

Lowit, A., Marchetti, A., Corson, S. & A. Kuschmann. 2018. Rhythmic performance in hypokinetic dysarthria : relationship between reading, spontaneous speech and diadochokinetic tasks. *Journal of Communication Disorders* 72, 26–39.

Mairano, P. 2010. Rhythm typology : acoustic and perceptive studies. Thèse de doctorat, Université de Turin.

Mairano, P. & A. Romano. 2010. Un confronto tra diverse metriche ritmiche usando Correlatore. In Schmid, S., Schwarzenbach, M. & D. Studer (eds), *La dimensione temporale del parlato, Proceedings of the V National AISV Congress*, Zurich, 4-6 février 2009, 79-100.

Martin, P. 2005. La transcription des proéminences accentuelles : mission impossible ? *Bulletin PFC* 6, 81-87.

Martin, P. 2014. Spontaneous speech corpus data validates prosodic constraints. In Campbell, N., Gibbon, D. & D. Hirst (eds), *Proceedings of Speech Prosody 2014*, Shanghai, 22-25 mai 2014, 525-529.

Martin 2017. Ondes cérébrales et accent syllabique. In Vičková-Mejvaldová (ed), *Komunikace řeči, komunikace lidí*. Prague : Presses de l'université Charles de Prague, 24-36.

Martin, P. 2018. Intonation, structure prosodique et ondes cérébrales. London : ISTE.

Maruthy, S., Venugopal, S. & P. Parakh. 2017. Speech rhythm in Kannada speaking adults who stutter. *International Journal of Speech-Language Pathology*, 19:5, 529-537.

McNeil, M. R., Robin, D. A. & R. A. Schmidt. 2009. Apraxia of speech : definition and differential diagnosis. In McNeil, M. R. (ed), *Clinical management of sensorimotor speech disorders*. New York : Thieme, 249-268.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J. & C. Amiel-Tison. 1988. A precursor of language acquisition in young infants. *Cognition* 29, 143-178.

Meigret, L. 1550. *Le tretté de la grammere francoëze*. Genève : Slatkine.

Mettouchi, A., Lacheret-Dujour, A. & V. Silber-Varod. 2007. Only prosody : perception of speech segmentation in Kabyle and Hebrew. *Cahiers de linguistique française* 28, 207-218.

Miller, J. L., Grosjean, F., Lomanto, C. 1984. Articulation rate and its variability in spontaneous speech : a reanalysis and its implications. *Phonetica* 41, 215-225.

Milliard, S. K., Nicholas, A. & F. M. Cook. 2008. Is parent-child interaction therapy effective in reducing stuttering ? *Journal of Speech, Language and Hearing Research* 51:3, 636-650.

Monfrais-Pfauwadel, M.-C. 2014. Bégaiement, bégaiements : un manuel clinique et thérapeutique. Paris : De Boeck-Solal.

Monfrais-Pfauwadel, M.-C. & N. Teitler. 1996. La décomposition de la syllabe chez le bégue : évidence phonétique, interprétation phonologique et perspectives de remédiation. *Revue française de linguistique appliquée* 1:1, 65-80.

Monfrais-Pfauwadel, M.-C., Tromelin, O., Mougin, A. L. & Y. Ormezzano. 2005. Utilisation des explorations multimédias synchrones dans l'objectivation des événements laryngés lors des bégayages, *Revue de Laryngologie, Otologie, Rhinologie* 126:5, 341-345

Nazzi, T., Bertoncini, J. & J. Mehler. 1998. Language discrimination by newborns : towards an understanding of the role of rhythm. *Journal of Experimental Psychology* 24:3, 756-766.

Nespor, M. 1990. On the rhythm parameter in phonology. In Roca, I. (ed), *The logical problem of language acquisition*. Dordrecht : Foris, 157–175.

Nespor, M., Shukla, M. & J. Mehler. 2011. Stress-timed vs. syllable-timed languages. In Van Oostendorp, M., Ewen, C.J., Hume, E. & K. Rice (eds), *The Blackwell companion to phonology*. Sussex : Wiley-Blackwell, 1147-1159.

Nespor, M. & M. Vogel. 1989. On clashes and lapses. *Phonology* 6, 69-116.

Obin, N., Avanzi, M., Bortal, G., Bardiaux, A. 2012. À la recherche des temps perdus : variations sur le rythme en français. In *Actes de la conférence conjointe JEP-TALN-RECITAL 2012*, volume 1, Grenoble, 4-8 juin 2012, 65-72.

Onslow, M. 2020. Stuttering and its treatment : eleven lectures. Monographie en ligne à l'adresse : <https://www.uts.edu.au/research-and-teaching/our-research/australian-stuttering-research-centre/resources/resources> [Page consultée le 25/03/2020].

Onslow, M., Van Doorn, J. & D. Newman. 1992. Variability of acoustic segment durations after prolonged-speech treatment for stuttering. *Journal of Speech and Hearing Research* 35, 529-536.

Packman, A. 2012. Theory and therapy in stuttering : a complex relationship. *Journal of Fluency Disorders* 37, 225-233.

Packman, A., Onslow, M. & J. Van Doorn. 1994. Prolonged speech and modification of stuttering : perceptual, acoustic, and electroglottographic data. *Journal of Speech and Hearing Research* 37, 724-734.

Packman, A., Onslow, M. & J. Van Doorn. 1996. Syllabic stress and variability : a model of stuttering. *Clinical Linguistics and Phonetics* 10:3, 235-263.

Packman, A. & M. Onslow. 1998. The behavioral data language of stuttering. In Cordes, A. & R. J. Ingham (eds), *Treatment efficacy in stuttering : a search for empirical basis*. San Diego : Singular Publishing Group, 27-50.

Packman, A. & J. S. Attanasio. 2004. *Theoretical issue in stuttering*. New York : Psychology Press.

Pasdeloup, V. 1990. *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*. Thèse de doctorat, Université de Provence - Aix-Marseille 1.

Pasdeloup, V. 2004. Le rythme n'est pas élastique : étude préliminaire de l'influence du débit de parole sur la structuration temporelle. In *Actes des XXVe Journées d'Études sur la Parole*, Fès, 19-22 avril.

Pendelieu-Verdurand, M. 2014. *Parole disfluente : aspects phonétiques et phonologiques*. Thèse de doctorat, Université de Grenoble.

Peterson, G. E. & I. Lehiste. 1960. Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America* 32:6, 693-703.

Pike, K. 1945. The intonation of american English. Ann Arbor : University of Michigan Press.

Port, R. F., Dalby, J. & M. O'Dell. 1987. Evidence for mora-timing in Japanese. *Journal of the Acoustical Society of America* 81:5, 1574-1585.

Praxiling - UMR 5267. 2017. Paroles disfluentes [Corpus]. ORTOLANG, disponible à l'adresse : <https://www.ortolang.fr/market/corpora/paroles-disfluentes/v1>. [Page consultée le 21/10/2019)

Prins, D., Hubbard, C. P. & M. Krause. 1991. Syllabic stress and the occurrence of stuttering. *Journal of Speech and Hearing Research* 34:5, 1011-1016.

Ramus, F. 2002. Acoustic correlates of linguistic rhythm : Perspectives. In *Proceeding of Speech Prosody 2002*, Aix-en- Provence, 11-13 avril 2002, 115–120.

Ramus, F., Nespor, M. & J. Melher. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition*, 73:3, 265-292.

Roach, P. 1982. On the distinction between « stress-timed » and « syllable-timed » languages. In Crystal, D. (ed), *Linguistic controversies : essays in honour of F. R. Palmer*. Londres : Edward Arnold, 73-79.

Rossi, M. 1979. Le français, langue sans accent ? *Studia Phonetica* 15, 13-51.

Schmitt, L. S. & E. B. Cooper. 1978. Fundamental frequencies in the oral reading behavior of stuttering and nonstuttering male children. *Journal of Communication Disorders* 11, 17-23.

Selkirk, E. 1984. *Phonology and syntax. The Relation between sound and structure*. Cambridge : MIT Press.

Schäfersküpper P., Dames M. 1987. Speech rate and syllable durations in stutterers and nonstutterers. In Peters, H. F. M. & W. Hulstijn (eds), *Speech Motor Dynamics in Stuttering*. Vienne : Springer, 329-336.

St Louis, K., & K. Schulte. 2011. Defining cluttering : the lowest common denominator. In Ward, D. & K. Scaler Scott (eds), *Cluttering : a handbook of research, intervention and education*. East Sussex : Psychology Press, 233–253.

Starkweather, C. W. 1987. *Fluency and stuttering*. Englewood Cliffs : Prentice-Hall.

Starkweather, C. W. & J. Givens-Ackerman. 1997. *Stuttering*. Austin : Pro-Ed.

Sussman M., Byrd, C. T. & B. Guitar. 2011. The integrity of anticipatory coarticulation in fluent and non-fluent tokens of adults who stutter. *Clinical Linguistics & Phonetics* 25:3, 169–86.

Thacker R C. & L. De Nil. 1996. Neurogenic cluttering. *Journal of Fluency Disorders* 21, 227-238.

Tuckey, J. W. 1977. *Exploratory data analysis*. Reading MA : Addison-Wesley.

Vaissière, J. 1974. On French prosody. *Quarterly Progress Report* 114, Research Laboratory of Electronics, MIT, 212-223.

Vaissière, J. 1991. Rhythm, accentuation and final lengthening in French. In Sundberg, J., Nord, L. & R. Carlson (eds), *Music, Language, Speech and Brain*. Stockholm : Macmillan Press, 108-120.

Vaissière, J. 2010. Le français, langue à frontières par excellence. In Delomier, D. & M.-A. Morel (eds), *Frontières, du linguistique au sémiotique*. Limoges : Lambert-Lucas, 39-62.

Van Borsel, J. 2014. Acquired stuttering : a note on terminology. *Journal of Neurolinguistics* 27, 41-49.

Van der Hulst, H., Goedemans, R. & E. van Zanten (eds). 2010. *A survey of word accentual patterns in the languages of the world*. New York : Mouton de Gruyter.

Van Zaalen, Y. 2009. *Cluttering identified. Differential diagnostics between cluttering, stuttering and learning disability*. Thèse de doctorat, Université d'Utrecht.

Ward, D., Connally, E. L., Pliatsikas, C., Bretherton-Furness, J. & J. Watkins. 2015. The neurological underpinnings of cluttering : some initial findings. *Journal of Fluency Disorders* 43, 1-16.

Wenk, B., & F. Wioland. 1982. Is French really syllable-timed ? *Journal of Phonetics* 10, 193-216.

White, L. & Matthys, S. L. 2007. Calibrating rhythm : first language and second language studies. *Journal of Phonetics* 35:4, 501-522.

Wieland, E. A., McAuley, J. D., Dilley, L. C. & S-E. Chang. 2015. Evidence for a rhythm perception deficit in children who stutter. *Brain & Language* 144, 26-34.

Wingate, M. E. 1969. Sound and pattern in « artificial » fluency. *Journal of Speech and Hearing Research* 12:4, 677-686.

Wingate, M. E. 1979. The loci of stuttering : grammar or prosody ? *Journal of Communication Disorders* 12:4, 283-290.

Wingate, M. E. 1984. Stuttering as a prosodic disorder. In Curlee, R. F. & W. H. Perkins (eds), *Nature and Treatment of Stuttering : New directions*. San Diego, CA : College-Hill.

Wingate, M. E. 1988. *The structure of stuttering : a psycholinguistic analysis*. New York : Springer-Verlag.

World Health Organization. 1977. *Manual of the international statistical classification of diseases, injuries and causes of death*. Volume 1. Genève : World Health Organization.

Yaruss, S., Logan, K., & E. Conture, E. 1994. Speaking rate and diadochokinetic abilities of children who stutter. *Journal of Fluency Disorders* 19, 221-222.

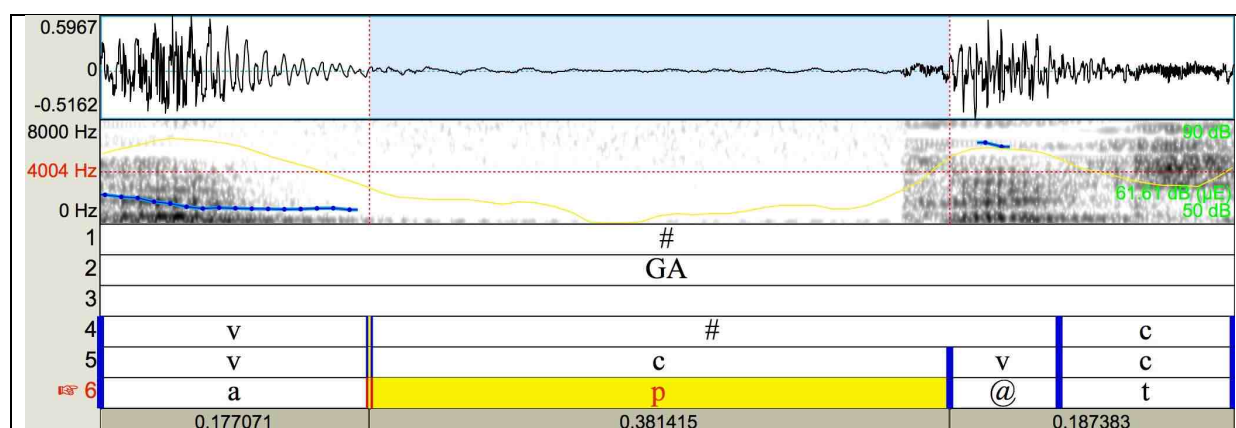
Zellner, B. 1998. *Caractérisation et prédiction du débit de parole en français. Une étude de cas*. Thèse de doctorat, Université de Lausanne.

Zimmermann, G. 1980. Articulatory dynamics of fluent utterances of stutterers and nonstutterers. *Journal of Speech Language and Hearing Research* 23:1, 95-107.

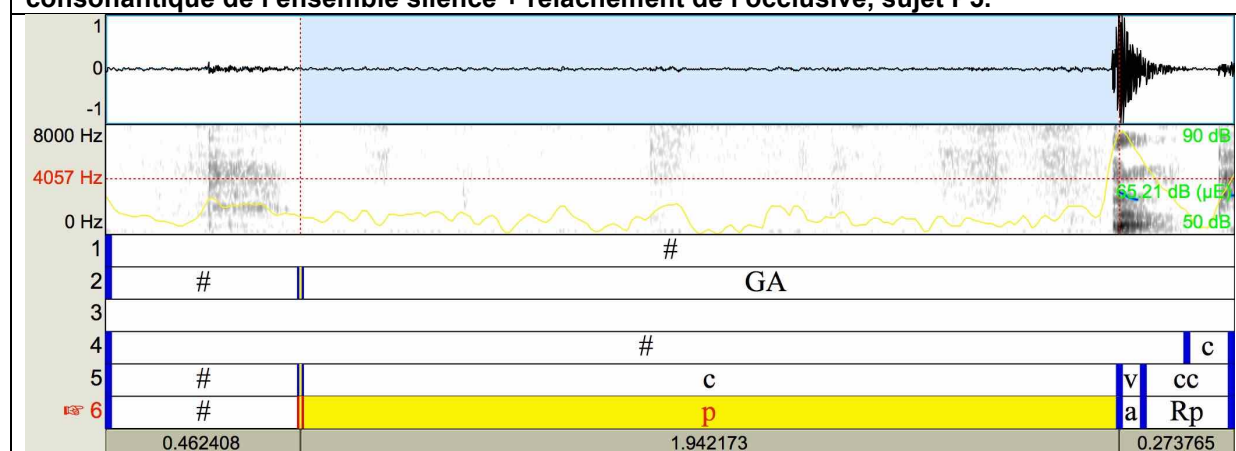
ANNEXES

1. Exemples d'analyses segmentales

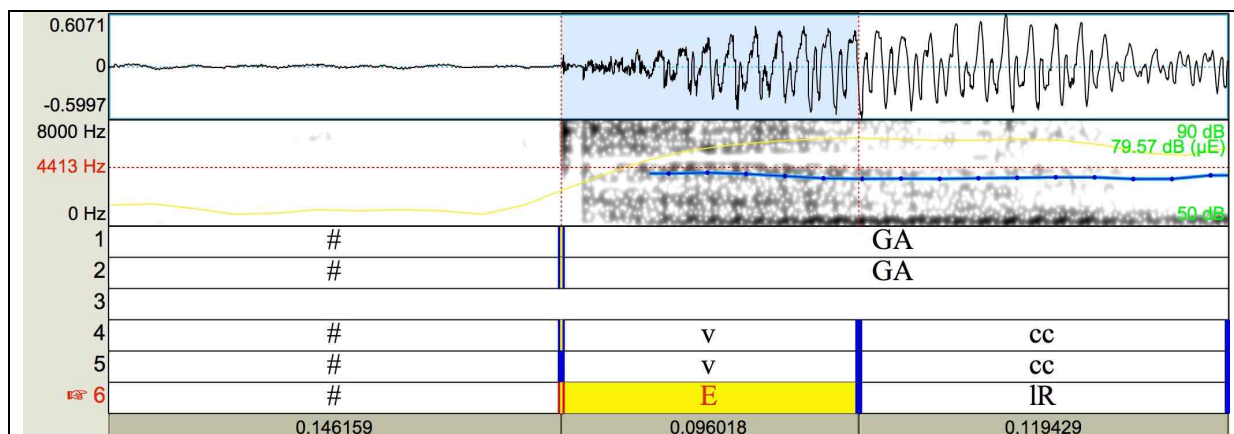
Nous illustrons les principales difficultés d'analyse segmentale (chapitre 2, section 2.4.1). L'intervalle pertinent est représenté en jaune. La tire inférieure correspond à la segmentation brute en intervalles vocaliques et consonantiques transcrits au moyen de l'alphabet Sampa, la tire immédiatement au-dessus à la même segmentation transcrite au moyen des seuls symboles c (intervalle consonantique), v (intervalle vocalique) et # (pause), et la tire supérieure, lorsqu'elle est présente, à la segmentation obtenue après exclusion des disfluences par syllabes, traitées comme des pauses (#). Enfin, les trois premières tires correspondent à l'analyse supra-segmentale, non pertinentes ici (voir annexe 2). Le spectrogramme est réglé en bandes larges (fenêtre d'analyse gaussienne de 3,5 ms, dynamique de 40 dB, échelle fréquentielle de 0 à 8000 Hz), la courbe d'intensité est représentée en jaune et la courbe de F0 en bleu (échelle variable selon les sujets).



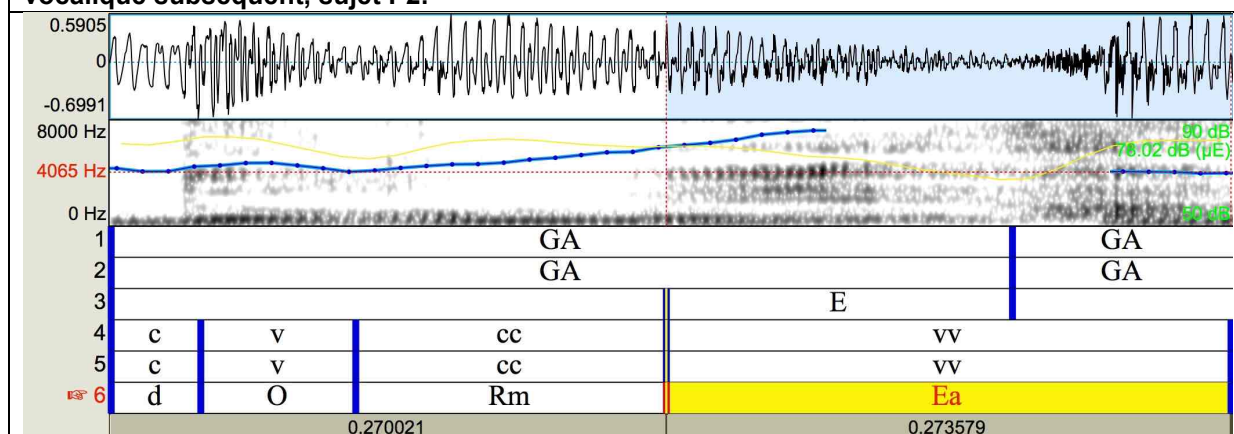
Annexe 1a. Blocage à l'initiale du mot « petite » avec intégration au sein d'un même intervalle consonantique de l'ensemble silence + relâchement de l'occlusive, sujet P5.



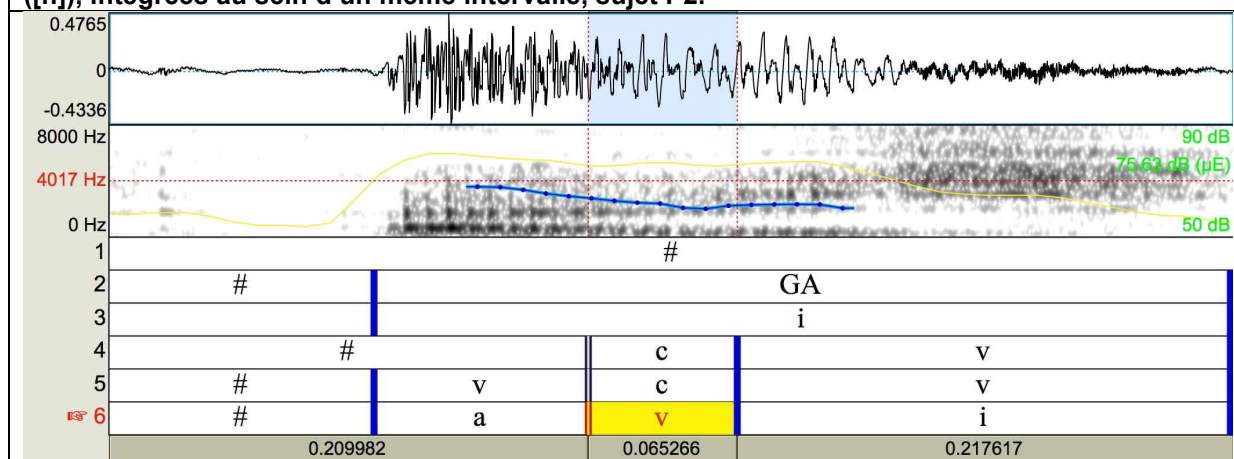
Annexe 1b. Blocage à l'initiale du groupe prépositionnel « par pitié », avec intégration au sein d'un même intervalle consonantique de l'ensemble silence + relâchement de l'occlusive à l'exclusion de la portion initiale terminée par un bruit inspiratoire, traitée comme une pause, sujet P5.



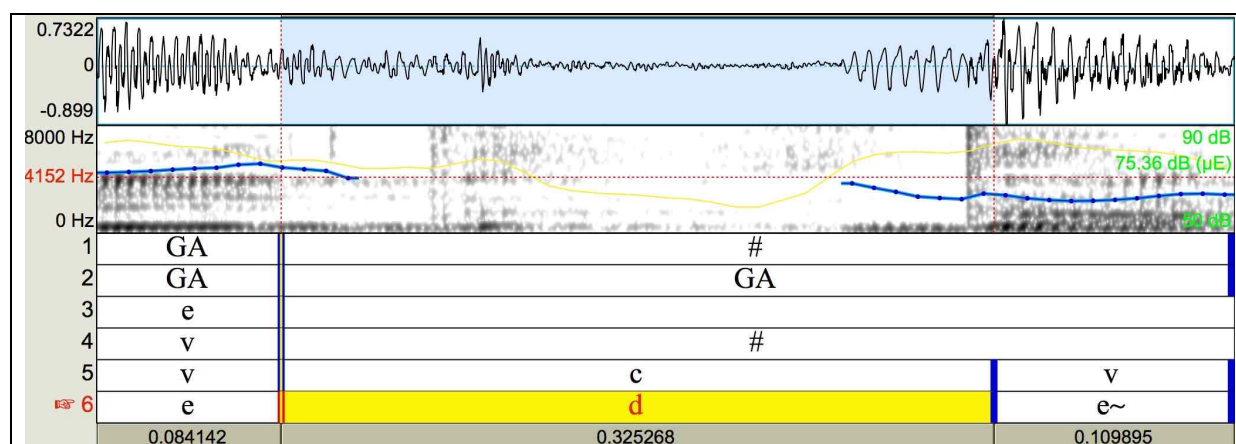
Annexe 1f. Glottalisation occlusive ([ʔ]) à l'initiale du mot "elle", intégrée à l'intervalle vocalique subséquent, sujet P2.



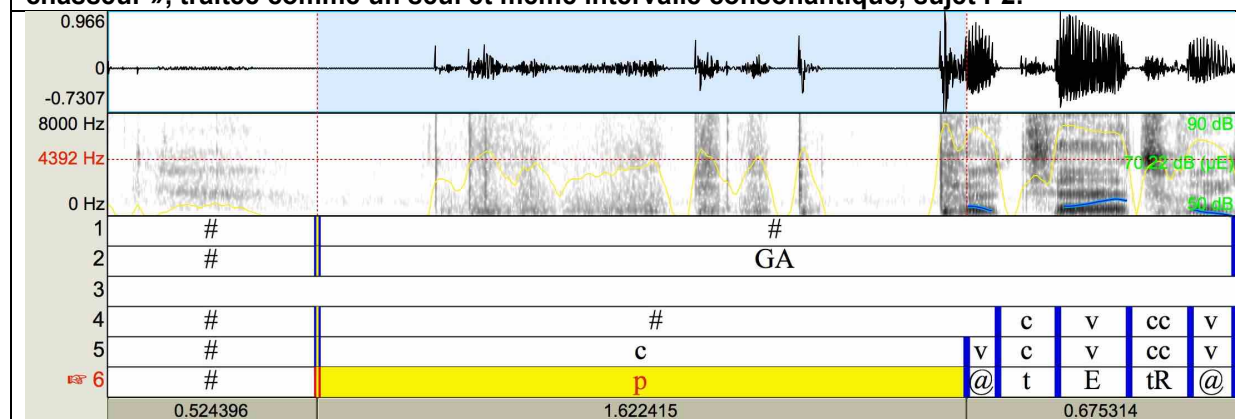
Annexe 1g. Voyelles en hiatus (« dormait à l'ombre ») avec glottalisation intercalaire fricative ([h]), intégrées au sein d'un même intervalle, sujet P2.



Annexe 1h. Consonne fricative /v/ réalisée comme une approximante à l'initiale du mot « vie », sujet P5.



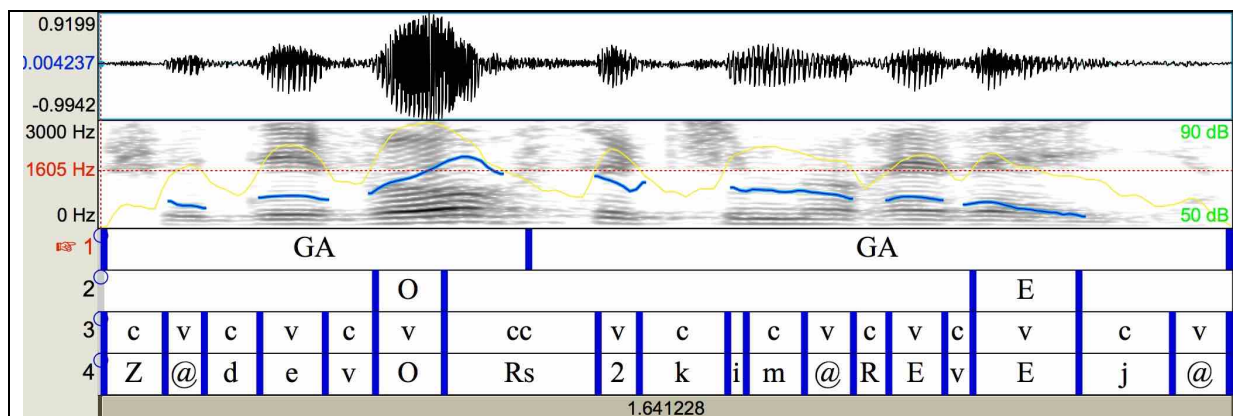
Annexe 1i. Répétition de la consonne /d/ à l'initiale du groupe prépositionnel « d'un chasseur », traitée comme un seul et même intervalle consonantique, sujet P2.



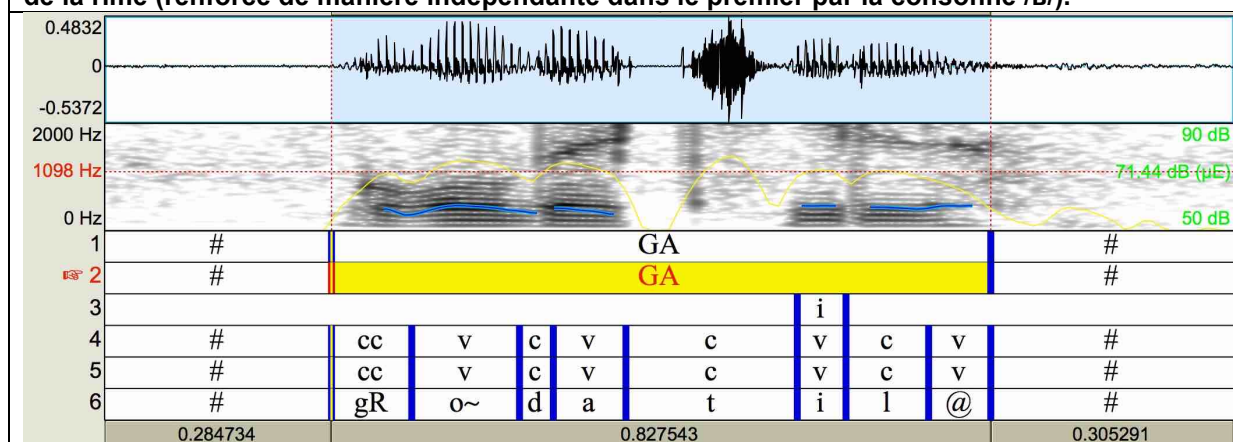
Annexe 1j. Prolongation par relâchement intermittent de l'occlusive /p/ à l'initiale du mot « peut-être », sujet P4.

2. Exemples d'analyses supra-segmentales

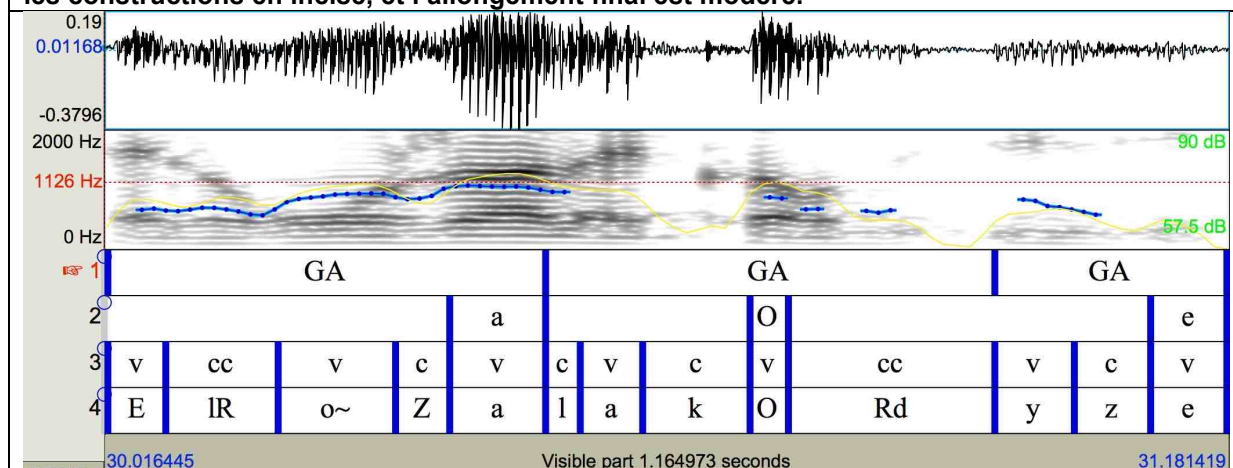
Nous présentons quelques exemples d'analyses supra-segmentales (chapitre 2, section 2.4.2). La tire supérieure correspond à la segmentation en groupes accentuels à l'exclusion de ceux présentant des disfluences, traités comme des pauses (#), la tire immédiatement inférieure, lorsqu'elle est présente, à la segmentation en groupes accentuels avec disfluences, la tire suivante aux voyelles accentuées, et les tires inférieures à l'analyse segmentale (voir annexe 1). La courbe d'intensité est représentée en jaune et la courbe de F0 en bleu. Le spectrogramme est réglé en bandes étroites (fenêtre d'analyse gaussienne de 35 ms, dynamique de 40 dB), sauf lorsqu'il s'agit d'illustrer un phénomène segmental (fenêtre d'analyse gaussienne de 3,5 ms), et l'échelle de fréquences varie suivant les sujets.



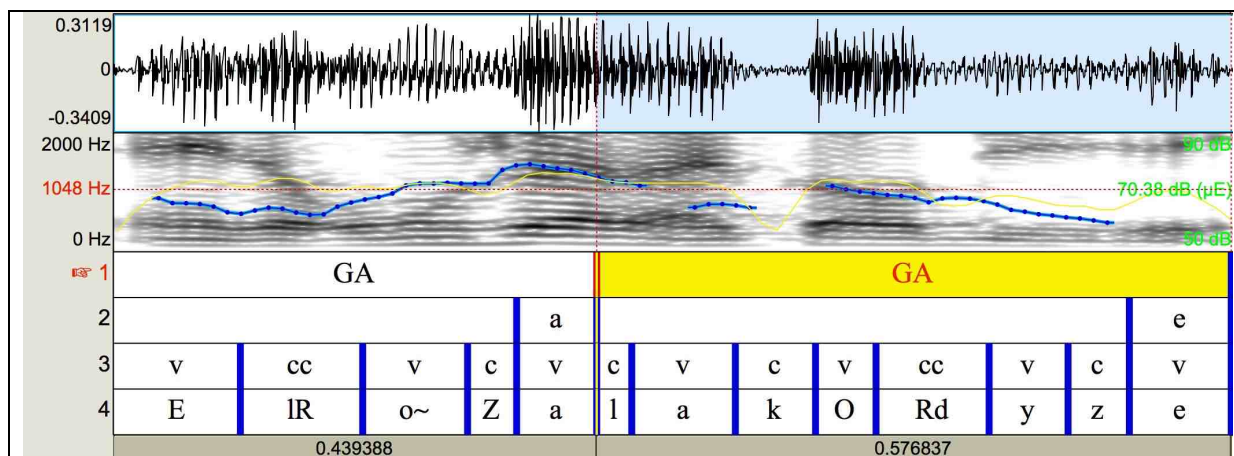
Annexe 2a. Groupes accentuels coïncidant avec des groupes intonatifs, sujet C2. Le premier « je dévore » se caractérise par une montée mélodique ample sur la syllabe finale, le second (« ceux qui me réveillent ») par une descente et chacun des deux par un allongement notable de la rime (renforcé de manière indépendante dans le premier par la consonne /ʁ/).



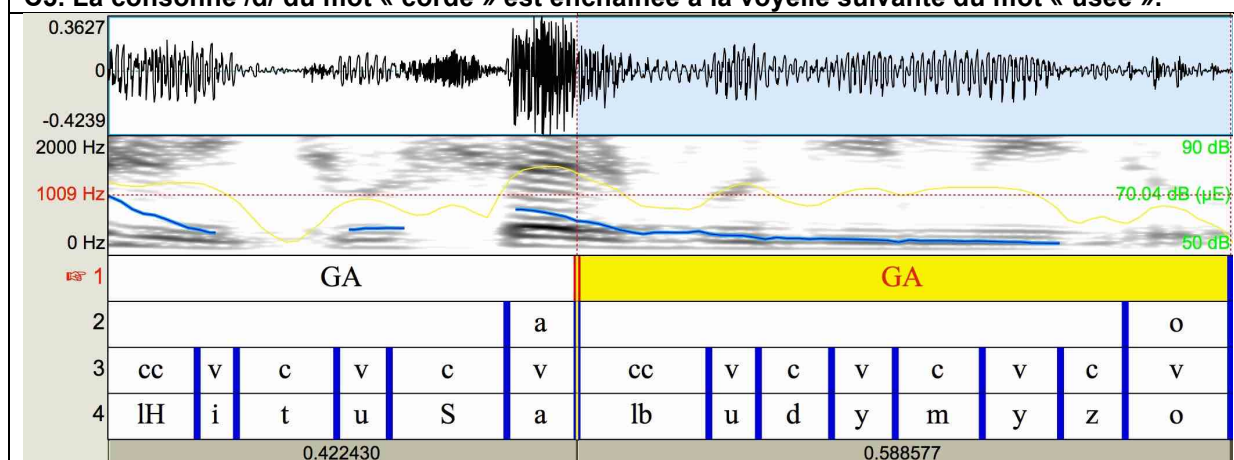
Annexe 2b. Groupe accentuel coïncidant à la fois avec un groupe intonatif et une unité inter-pausale (« gronda-t-il »), sujet P4. Le contour intonatif est plat, comme cela est typique dans les constructions en incise, et l'allongement final est modéré.



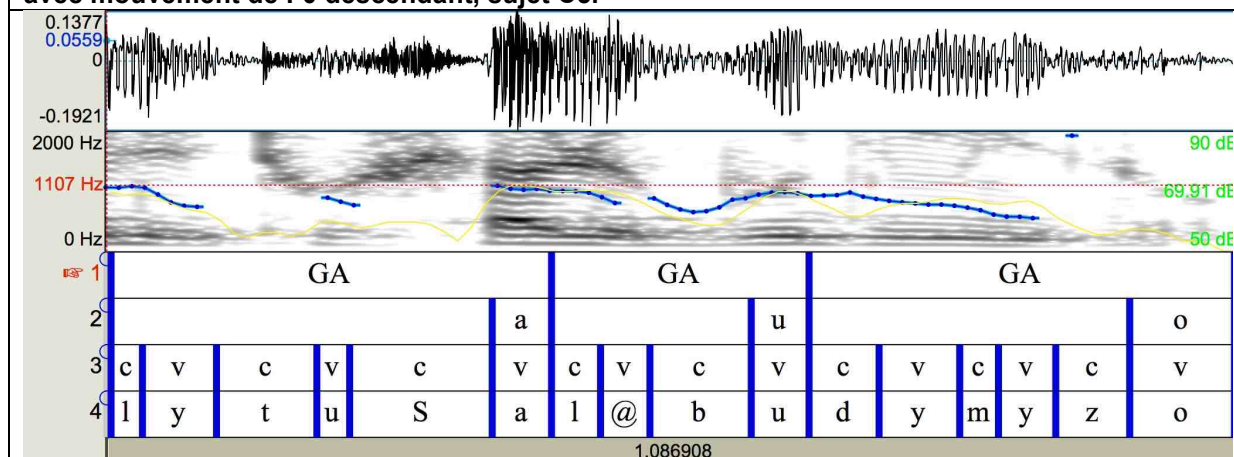
Annexe 2c. Réalisation du groupe nominal « la corde usée » en deux groupes accentuels, séparés par une pause inférieure à 200 ms (intégrée à l'intervalle consonantique dans les tires segmentales), sujet C4. La consonne /d/ du mot « corde » n'est pas enchaînée à la voyelle suivante du mot « usée ».



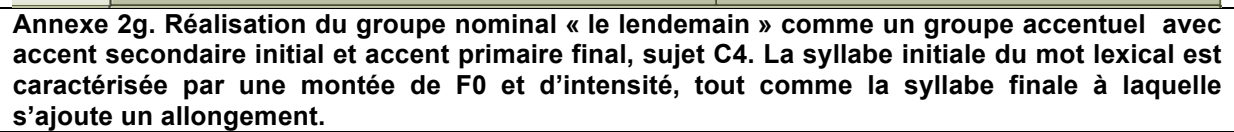
Annexe 2d. Réalisation du groupe nominal « la corde usée » en un seul groupe accentuel, sujet C5. La consonne /d/ du mot « corde » est enchaînée à la voyelle suivante du mot « usée ».

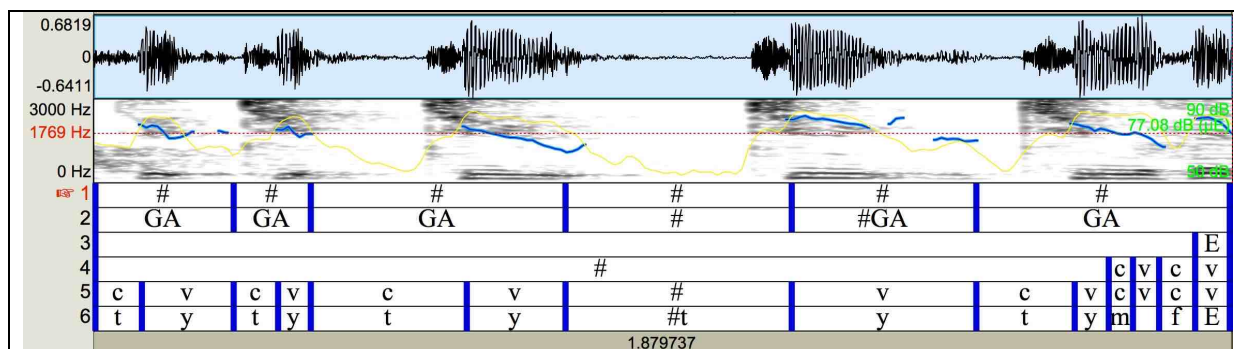


Annexe 2e. Réalisation du groupe nominal « le bout du museau » en un seul groupe accentuel avec mouvement de F0 descendant, sujet C5.

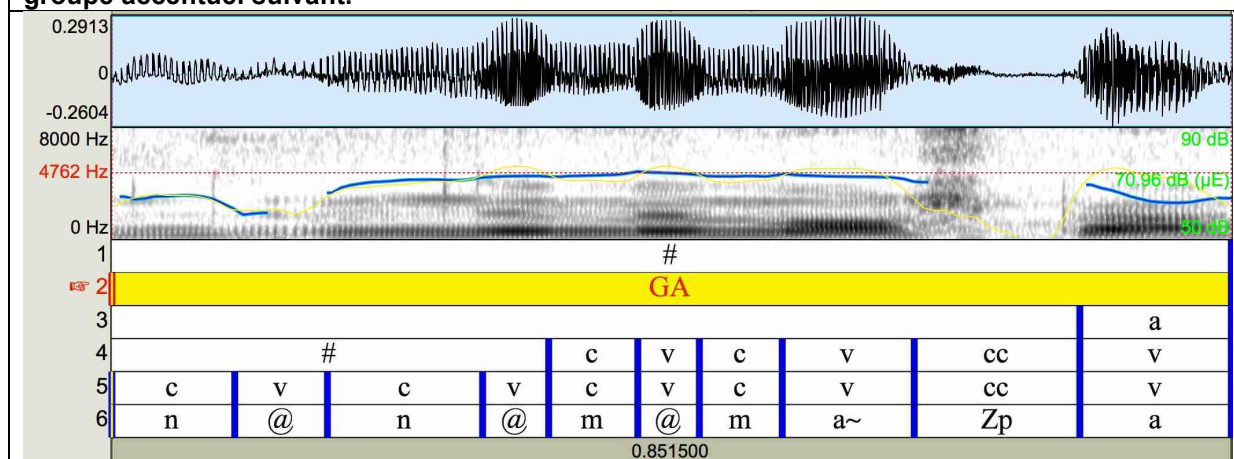


Annexe 2f. Réalisation du groupe nominal « le bout du museau » en deux groupes accentuels, le premier étant associé à un mouvement montant de F0 et le second à un mouvement descendant, sujet C4.

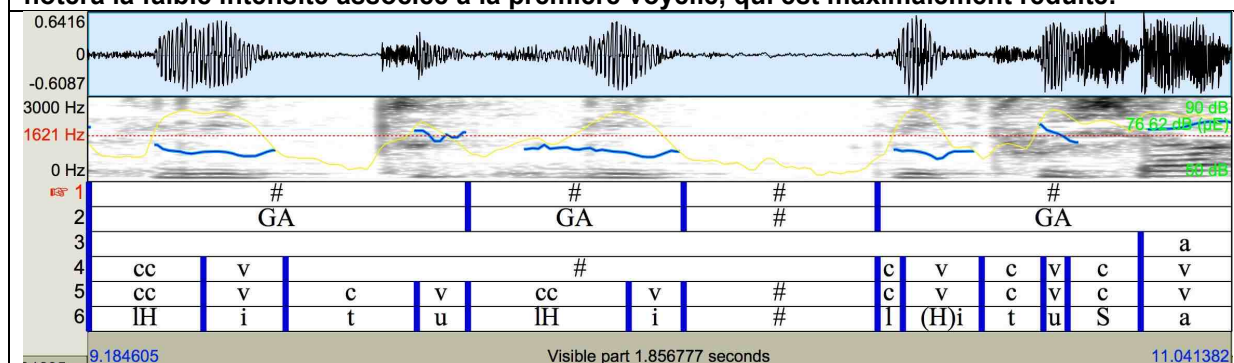




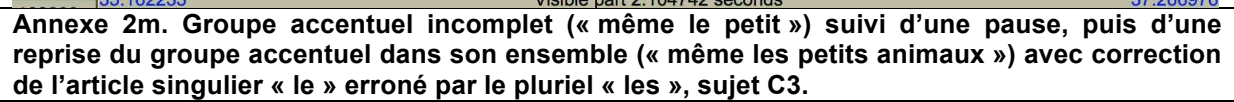
Annexe 2j. Répétition tendue de syllabes, affectant le pronom « tu », sujet P2. Les trois premières syllabes sont traitées comme des groupes accentuels à part entière, la troisième, séparée des précédentes par une pause, est exclue de l'analyse, et la dernière est intégrée au groupe accentuel suivant.



Annexe 2k. Répétition de syllabes sans tension particulière, affectant la particule « ne », sujet C3. Les deux syllabes forment avec le matériel qui suit un seul et unique groupe accentuel. On notera la faible intensité associée à la première voyelle, qui est maximale réduite.



Annexe 2l. Suite de deux groupes accentuels incomplets (« lui tou- », puis « lui ») suivie d'une pause, puis du groupe accentuel complet (« lui toucha »), sujet P2.



RÉSUMÉ COURT

La présente étude vise à explorer les perturbations du rythme de la parole dans le bégaiement en français au moyen de méthodes acoustiques. Nous utilisons pour ce faire différentes mesures rythmiques introduites initialement au niveau segmental pour contraster les langues selon leur degré d'accentuation (Ramus *et al.* 1999, Grabe & Low 2002), que nous étendons au niveau supra-segmental (Asu & Nolan 2006, Obin *et al.* 2012, Cichocki 2015) sur un corpus de parole lue composé des enregistrements de 5 locuteurs adultes présentant un bégaiement développemental modéré à sévère et 5 locuteurs contrôle appariés en âge et sexe. Les résultats obtenus sont de deux ordres. Sur le plan supra-segmental, les mesures considérées confortent l'hypothèse de Monfrais-Pfauwadel (2014) suivant laquelle les disfluences caractéristiques du bégaiement perturbent le phrasé en groupes accentuels ou « *chunking* » de la parole : l'eurythmie est typiquement rompue. Sur le plan segmental, la parole des sujets qui bégaiant se caractérise par une variabilité temporelle plus marquée des intervalles vocaliques et consonantiques, en accord avec les résultats récents de Maruthy *et al.* (2017), et plus généralement les théories du bégaiement fondées sur l'hypothèse d'une instabilité motrice d'origine neurologique centrale (Zimmermann 1980, Packman & Onslow 1996).