

Comportements d'autocorrection et d'hésitation manifestés par les apprenants de FLE au cours de conversations orales spontanées

Hugues PÉTERS

UNSW, Australia

School of Humanities and Languages, Sydney NSW 2052, Australie

h.peters@unsw.edu.au

There is a need for more precise descriptions of disfluency markers in the actual oral dialogic productions of learners of French as a foreign language to inform language pedagogy. This research aims at documenting communicative strategies used by learners of French when confronted with gaps in communication, namely phenomena of hesitation and repairs (repetition, self-correction and false start) occurring during SLA learners' spontaneous speech at different stages of proficiency. It is based on a longitudinal oral learner corpus of actual spontaneous speech by Jamaican learners of French. We will also compare the observed results with the descriptors available in the Common European Framework describing a conscious communicative strategy of self-correction and suggest pedagogical ways to improve oral communication.

Keywords:

CEFR, false start, fillers, French learner corpus, pause, repetition, SLA, self-repair

1. Introduction

Cette recherche vise à décrire de manière empirique les comportements d'hésitation et les stratégies discursives d'autocorrection, de répétitions et de faux départs, utilisées par des apprenants de français langue étrangère (FLE) au cours de conversations orales spontanées¹. Elle se base sur l'analyse fine de données de trois apprenants de FLE, extraites d'un corpus oral et longitudinal de productions d'apprenants jamaïcains en contexte guidé. Le caractère longitudinal d'un tel corpus permet de décrire l'évolution des capacités d'expression orale relevant de stades différenciés de développement et de comparer entre eux des apprenants relativement moins avancés à d'autres plus avancés, et s'intéresse donc aux processus de différenciation des pratiques langagières en contexte d'enseignement. Ces comparaisons d'apprenants à apprenants nous permettront de déceler une évolution des comportements discursifs de reprise en fonction du niveau de compétence.

Dans une perspective didactique, il sera aussi utile de comparer nos résultats avec les descripteurs du Cadre Européen Commun de Référence (CECR 2001), qui identifient, entre le niveau B1 où l'apprenant peut "corriger les

¹ Je remercie les audiences de Genève et de Kingston, le Dr. Tabensky et deux évaluateurs pour leurs commentaires avisés.

confusions de temps qui ont conduit à un malentendu" et le niveau C1 où il peut "revenir sur une difficulté et reformuler ce qu'il veut dire sans interrompre complètement le fil du discours" (CECR 2001: 54), un niveau B2 où l'apprenant atteint un nouveau degré de conscience de la langue: il peut ainsi "corriger les fautes qui ont débouché sur des malentendus; prendre note des "fautes préférées" et contrôler consciemment le discours pour les traquer." (CECR 2001: 33). C'est sur la base d'un état des lieux plus précis de tels phénomènes d'autocorrection que l'on pourra éventuellement élaborer des applications pédagogiques pour améliorer les compétences conversationnelles en L2.

2. Recherche sur la parole spontanée

Les études de corpus ont montré que la parole spontanée est ponctuée de phénomènes dits d'hésitation qui passent normalement inaperçus: il s'agit de pauses silencieuses ou remplies (*euh*), d'allongements vocaliques, ainsi que, dans des proportions moindres et en mesure variable selon le locuteur et la langue, d'amorces ou de répétitions de mots ou de syntagmes, d'autocorrections phonétiques, lexicales ou grammaticales, et de faux départs impliquant l'abandon complet d'un énoncé, toutes traces du travail de formulation (TdF) d'un énoncé en temps réel (Candéa 2000). De tels phénomènes se retrouvent naturellement dans la parole spontanée des apprenants de langues étrangères ou secondes. Cependant, comme nous allons le voir, ils trahissent le caractère non natif de ces derniers, sont perçus par les interlocuteurs, et peuvent même gêner la compréhension.

Il existe un consensus parmi les chercheurs pour analyser ces phénomènes dans le cadre du modèle de production de Levelt (1989) adapté à l'acquisition d'une L2 (Kormos 2006; Segalowitz 2010), et pour en attribuer la cause à un manque d'automatisation des processus de recherche lexicale ou de programmation syntaxique, surchargeant la mémoire de travail des apprenants lors de l'encodage linguistique. Le locuteur doit en effet penser à ce qu'il va dire, planifier l'organisation de son discours, chercher ses mots, construire ses phrases et contrôler ce qu'il dit au fur et à mesure avec une mémoire de travail limitée. L'effort cognitif nécessaire à ce TdF donne lieu à des "hésitations" presque à chaque mot.

Temple (2000) démontre que les apprenants de FLE manifestent ces comportements à des taux significativement plus élevés que les natifs, comme le montre l'exemple suivant extrait de notre corpus:

- (1) et <il y a un> [-] &hum le gagnant &hum (.) yeah@s &euh peut [//] va &euh (.) faire un &hum (.) cd avec leurs chansons, quelque chose comme ça, &*INV:oui, &euh oui.
[0831_Ex]²

² Dans l'encodage des exemples, les deux premiers chiffres (08) représentent le code de l'apprenant, les deux suivants le niveau (21 = deuxième année, premier semestre), les lettres représentent la situation (Ex = Examen). Sur la ligne orthographique, le & indique des éléments

Pour exprimer une proposition simple (*le gagnant va faire un cd avec [ses] chansons*), le locuteur produit trois pauses brèves (.), des pauses remplies de type anglophone (&hum) ou francophone (&euh), des auto-évaluations (*oui, yeah@s*) avec une interférence de la L1, un faux départ (<il y a un> [/-/]), en l'occurrence une amorce d'énoncé incomplet, et une autocorrection (*peut [//] va*) qui substitue un auxiliaire d'aspect à un auxiliaire de mode: en bref, une accumulation de comportements discursifs pour résoudre des problèmes d'encodage, avec une fréquence telle que la compréhension en est rendue ardue, d'autant plus que le résultat final contient une erreur de choix de déterminant possessif, typique pour un anglophone.

En outre, Temple (2000) suggère une différence qualitative entre apprenants et natifs quant aux types de réparations effectuées. Ainsi note-t-elle que si, chez un natif, réparations et faux départs sont surtout des reformulations lexicales destinées à mieux exprimer sa pensée ou en raison du registre de langue, les autocorrections grammaticales, normalement absentes chez ce dernier, sauf lapsus ou accident, pullulent chez les apprenants. Ces observations sont raffinées par Hilton (2014) qui propose une analyse détaillée des phénomènes de pauses et des types de reprises sur la base d'un corpus symétrique d'apprenants/natifs de français/anglais L1-L2 en une tâche monologique de résumé à différents stades de compétence.

Dans la recherche en acquisition des langues secondes ou étrangères, les études de 'pausologie' sont le plus souvent liées à la question de la perception de l'aisance de parole (*fluency*) analysée de manière objective et quantifiable par des mesures temporelles de productions orales. Ainsi, dans son étude séminale, Lennon (1990) avait-il démontré que si le nombre moyen de mots entre deux pauses silencieuses ou la vitesse de parole élaguée (nombre de mots, sans les reprises, par unité de temps) correspondent à une impression d'aisance de parole par des examinateurs, aucune des stratégies discursives impliquant une modification de l'énoncé (répétition, autocorrection ou faux départ), prise isolément, n'est liée de manière directe à une telle perception (voir Freed (2000) et Préfontaine & Kormos (2016) pour des études, respectivement quantitative et qualitative, dans cette double perspective de production et perception en FLE). Selon Tavakoli & Skehan (2005), il faut distinguer trois aspects: la vitesse (*speed fluency*), les pauses silencieuses et remplies qui interrompent le flux (*breakdown fluency*) et les reprises et hésitations utilisées pour réparer le discours (*repair fluency*). De plus, parmi ces dernières, Olynyk et al. (1990) établissent une distinction entre reprises 'progressives' ou répétitions (auxquelles s'ajoutent les pauses remplies) qui, pour les natifs et les apprenants avancés, mais non pour les apprenants moins avancés, peuvent

transcrits phonétiquement ou non encodés sur la ligne morphologique. "&*INV:oui" marque un encouragement en incise par l'investigateur et "&euh", une pause remplie. Les italiques ajoutés par l'auteur dans les exemples soulignent le phénomène illustré.

avoir un impact positif sur l'aisance de parole, et reprises 'régressives' (autocorrections et faux départs) qui ont normalement un impact négatif car elles contribuent au ralentissement du débit. Selon Witton-Davies (2010), il est clair qu'un taux excessif de reprises influe négativement sur la perception de facilité de parole, mais qu'une certaine quantité de répétitions pourrait en fait aider le locuteur à maintenir son rythme de parole en lui permettant notamment d'éviter des pauses à l'intérieur des énoncés. Enfin, même les autocorrections ne sont pas toujours jugées négativement, car les natifs semblent apprécier les efforts d'autocorrection d'apprenants de FLE (Préfontaine & Kormos 2016).

À quelques exceptions près telles que Michel (2011) ou Witton-Davies (2014) notamment, la recherche sur l'aisance de parole a été conduite sur des monologues. Le mode dialogique présente en effet des difficultés accrues en raison des tours de parole, interruptions et chevauchements divers, pour ce qui est du calcul du débit de parole. Tavakoli (2016) résume la spécificité des dialogues par rapport aux monologues: débit plus rapide, pauses plus courtes à l'intérieur des énoncés (mais en nombre identique), augmentation du nombre de pauses remplies, diminution du nombre de reprises, et impression globale d'aisance de parole accrue. Ces caractéristiques manifestent chez un locuteur l'avantage d'utiliser le temps de parole de son interlocuteur pour planifier sa propre intervention ainsi que la prise en compte des besoins de cet interlocuteur. Une meilleure connaissance de la parole en interaction, le mode de communication le plus fréquent et le plus naturel, est de première importance pour l'enseignement de la communication orale.

La présente recherche, par l'étude d'un corpus de productions orales de FLE en interaction, se penche sur la description des stratégies discursives d'hésitations et d'autocorrections employées au cours de conversations spontanées, liant cette analyse à des considérations d'ordre pédagogique.

3. Méthode

3.1 Corpus de FLE

Les données de cette étude préliminaire proviennent d'un corpus oral et longitudinal de productions de dix apprenants jamaïcains de FLE enregistrés six ou sept fois chacun sur une période de 18 mois au cours de leurs études à l'Université des Indes Occidentales, Mona, en Jamaïque (Péters 2009)³. Les entretiens ont été transcrits, annotés et encodés selon les procédures de CHILDES avec une ligne de décomposition morphologique en dessous de chaque ligne de transcription orthographique (MacWhinney 2000). De plus,

³ Ce corpus a initialement été constitué pour étudier l'influence des L1 en contexte diglossique sur la morphosyntaxe des productions orales en L2. Faute d'espace, je réfère le lecteur à Péters (2014). D'emblée, néanmoins, s'est posée la question de la transcription des pauses et reprises, inhérente au système de CHILDES.

chaque ligne de transcription a été liée à la séquence sonore correspondante. Ce corpus contient 33.527 mots d'apprenants.

La présente étude se limite à l'analyse de deux séries d'entretiens de trois apprenants (L08, L18 et L20), en tête à tête avec l'investigateur, à un an d'intervalle, à la fin du premier semestre de la deuxième année (II1) de leur bachelor et à la fin du premier semestre de leur troisième année (III1), pour un total de 4,299 mots d'apprenants.

3.2 *Entretiens*

Les sujets de conversations au niveau II1 étaient une discussion impromptue autour d'une personnalité, suivie d'une discussion sur le déroulement des études, les activités de loisir, les vacances et les projets d'avenir. L'entretien au niveau III1 était constitué d'une conversation similaire, suivie d'activités guidées: narration d'une bande dessinée courte, obtention de formes interrogatives par un jeu de devinette, description d'une personnalité à l'aide d'affirmations et de négations, jeu de rôles sur une réalité sociale familiale.

3.3 *Participants*

Les participants constituent un groupe homogène de locuteurs jamaïcains caractérisés par une diglossie anglais-créole jamaïcains, tous apprenants de français en contexte guidé, n'ayant jamais visité de pays francophone au moment des entretiens, et inscrits dans les mêmes cours de français à l'université. Ces apprenants avaient néanmoins des expériences distinctes d'étude de FLE avant de commencer leur programme universitaire: L08, ayant quelques notions d'espagnol, s'était mis à l'étude du français à l'université par un cours intensif d'un an de 250 heures, L18 avait étudié le français au lycée pendant quatre ans avec un nombre d'heures de contact estimé à 250 heures et L20 avait étudié le français et l'espagnol ensemble pendant 6 ans.

Ces trois apprenants ont été sélectionnés en fonction de leur niveau de compétence bien distinct, que je qualifierai d'élémentaire (L18), d'intermédiaire (L08) et d'avancé (L20), établi sur la base de la mesure de la diversité lexicale D calculée sur la ligne morphologique élaguée, c'est-à-dire sans hésitations ni reprises⁴. Nous avons en outre éliminé de ce calcul les mots inintelligibles, les emprunts à la L1 et les répétitions immédiates au mot à mot de l'investigateur.

Plus le résultat de VOCD est élevé, plus l'étendue du vocabulaire actif du locuteur est importante. Le premier locuteur (VOCD = 33.36/55.98) est en dessous du deuxième (48.83/62.90) et ce dernier en dessous du troisième (57.12/77.53), même si nous observons des chevauchements entre la mesure

⁴ Le programme VOCD (McKee et al. 2000) intégré au système CHILDES donne une mesure de diversité lexicale D moins sensible à la longueur du texte considéré que la simple mesure du rapport Type-Occurrence (*Type Token Ratio TTR*). Pour cette dernière, en effet, plus le texte est long, plus le TTR tend à diminuer, comme pour L08, entre le niveau II1 et le niveau III1.

en troisième année des moins avancés et la mesure en deuxième année des plus avancés. Ce chevauchement est naturel vu que les sujets de conversations au niveau III1 sont plus diversifiés et que la compétence des étudiants s'est améliorée en un an.

	L18	L08	L20
II1	33.36	48.83	57.12
TTR	$138 \div 543 = 0.254144$	$151 \div 520 = 0.290385$	$152 \div 458 = 0.331878$
III1	55.98	62.90	77.53
TTR	$201 \div 667 = 0.301349$	$237 \div 832 = 0.284856$	$233 \div 595 = 0.391597$

Fig.1: Mesure de VOCD sur la ligne morphologique.

4. Catégorisation des reprises

Le corpus de production orale nous permettra d'explorer les facteurs quantitatifs et qualitatifs de trois types de reprises effectuées dans les six entretiens sélectionnés: répétitions, autocorrections et faux départs.

4.1 Répétition

Par la catégorie des répétitions, codées [I], nous entendons les répétitions identiques et contiguës de mots ou de syntagmes. Les répétitions sémantiquement signifiantes (*très très*) ou d'usage courant (*oui oui*), ainsi que les amorces de mot, transcrites phonétiquement sur la ligne principale en code SAMPA⁵ et précédées d'un & (&Ze en 2d), ne sont pas incluses.

- (2) a. *ils [I] ils n' ont pas la passion pour [I] pour des [I] des [I] des sujets [1821]*
 b. *<c'est> [I] (.) c'est ennuyeux [I] très ennuyeux [1831]*
 c. *il y a <beaucoup de personnes> [I] 0 [=! click of the tongue] beaucoup de personnes gentilles [2031]*
 d. *&Ze j'étudie [...] et <j'aime> [I] (.) &eh (.) j'aime le politique [1821]*

La séquence d'origine répétée est généralement d'un ou deux, voire de trois mots (*beaucoup de personnes*). Des pauses, des "&eh", des phénomènes non linguistiques (toux, clic), voire des interférences ponctuelles de l'interlocuteur, peuvent intervenir entre les deux parties de la répétition.

Les répétitions de (séquences de) mots outils (MO) (en 2a-b): déterminants, pronoms, prépositions, conjonctions, auxiliaires ou "pro. + aux.", "prép. + dét.", etc. sont de loin les plus fréquentes (207/245, 84.49%) par rapport à celles de (syntagmes contenant des) mots pleins (MP) (en 2c-d): catégories lexicales ouvertes (verbes, noms, adjectifs) (38/245, 15.51%). Cela indique la difficulté du travail cognitif de recherche lexicale des MP⁶.

⁵ Le code SAMPA (Speech Assessment Method Phonetic Alphabet) est un alphabet phonétique international lisible à la machine (voir MacWhinney 2000).

⁶ Les appellations 'mots outils' ou 'mots pleins' sont utilisées à la suite de Candéa (2000) par facilité didactique.

En outre, les répétitions de MO peuvent être doubles, voire triples:

- (3) a. ils *ne* [/] *ne* [/] *ne* sont pas *son* [/] *son* [/] *son* sujet [1821]
 b. *les* [/] *les* [/] *les* [/] *les* services sociales [1831]

Chez les natifs, ces répétitions multiples sont d'habitude suivies d'un mot générique (*chose, mec, faire, etc.*) (Candéa 2000: 329), mais chez les apprenants, elles dénotent simplement une difficulté accrue dans la recherche de vocabulaire. Ces répétitions multiples sont d'ailleurs le fait de l'apprenant le plus faible, L18, qui en produit 20 (dont 2 triples) sur les 22 instances notées.

Lorsqu'il y a répétition de syntagme incluant un MP, un locuteur natif répètera les éléments dépendants précédant la tête du syntagme: notamment le pronom avec le verbe. En revanche, chez les apprenants se remarque fréquemment l'oubli du pronom clitique devant le verbe:

- (4) ils &hum (.) *enseignant* [/] *enseignant* &hum créole Jamaïcain [0831]

L08 et L18 oublie systématiquement le pronom personnel de troisième personne lors des trois cas de répétition d'un syntagme à verbe plein qu'ils produisent. Leur attention complète semble accaparée par le choix des verbes (*enseigner, améliorer, réaliser*), pourtant relativement fréquents, et de leur complément. C'est une indication de la faiblesse d'analyse morphosyntaxique des clitiques, assez typique chez des locuteurs anglophones qui ré-analysent les clitiques du français comme des pronoms forts. Néanmoins, ces mêmes apprenants n'omettent normalement pas les pronoms devant auxiliaires, ni à la première personne devant des verbes pleins⁷.

4.2 Autocorrection

Lors d'autocorrections immédiates, codées [/] (ou [///] pour des reformulations syntaxiques ou lexicales), l'apprenant se reprend immédiatement après avoir perçu une erreur. Comme chez les natifs, il peut s'agir de modifications liées au registre (5a) (1 occurrence de ce type), au choix du genre du déterminant (5b) (11 occurrences), à une élision (5c) (16 occurrences), à une correction phonétique (5d) (4 occurrences), à l'ajout d'un modifieur pour préciser sa pensée (5e) (6 occurrences) ou à une substitution lexicale (5f) (14 occurrences):

- (5) a. <tu sais> [/] *vous savez* qu'est saramaccan ? [0831]
 b. c'est *le* [/] *la* chanteuse principale de la groupe [2021]
 c. et je [/] (.) *je* [/] *j'aime* son ténacité ? [1821]
 d. il (.) ne peut pas <&z amuser> [/] &*INV:hm *s'amuser* [1831]
 e. c'est *ennuyeux* [/] *très ennuyeux* [1831]
 f. <j'aime> [///] <je voudrais> [/] <je voudrais> [///] *j'aimerais* lire [0821]

⁷ Il y a cependant trois cas d'omission devant auxiliaire en contexte interrogatif ou négatif sur les 21 exemples relevés. Cela se confirme dans les autres entretiens de ces apprenants.

Mais, à la différence des natifs, les apprenants font des corrections de traits morphosyntaxiques de genre (6a) (6 occurrences de ce type), de temps (6b) (11 occurrences), d'accord sujet verbe (6c) (2 occurrences), ou des reformulations syntaxiques par ajout ou suppression de mots, changement de structure avec ou sans substitution lexicale (6d) (16 occurrences) et d'autres modifications: choix de déterminant (8 occurrences) ou de préposition (7 occurrences), de nombre (7 occurrences), d'auxiliaire (4 occurrences), de polarité (4 occurrences), de pronom relatif ou interrogatif (2 occurrences), de personne (1 occurrence) ou de forme du verbe (6e) (1 occurrence), toutes corrections que les natifs ne font d'habitude pas, soit qu'ils ne fassent pas ce genre d'erreurs, soit qu'elles passent inaperçues dans le flux de la parole.

- (6) a. quelle est <son nationalité> [/] sa &na nationalité ? [0831]
 b. <j'assiste> [/] j'ai assisté [2021]
 c. tous les enfants <doit aller> [/] doivent aller à l'école [1831]
 d. <elle était un [/] un film> [///] elle était dans un film [0821]
 e. j'ai apprend [/] appris [0831]

La séquence d'origine autocorrigée par l'apprenant est généralement d'un ou deux mots, mais peut aller jusqu'à cinq mots (voir (7a) ci-dessous).

Notons que les apprenants ne perçoivent qu'une partie des erreurs qu'ils produisent. De plus, un trait particulièrement saillant quoique minoritaire (21/122, 17.21%) de ces reprises que les apprenants font afin de corriger une erreur qu'ils ont perçue est que certaines introduisent une erreur autre, voire rendent incorrecte une forme qui était correcte (ou partiellement correcte): par exemple l'accord sujet verbe en (7a) ou le pronom relatif sujet en (7b):

- (7) a. j' <ai toujours allé à Carib> [/] &*INV:et: va [/] va toujours à Carib [1831]
 b. un communauté <qui s'appelle Name_Of_Community> [/] &*INV:hmhm &euh que s'appelle Name_Of_Community [1831]

On assiste parfois à des errements dans la recherche de grammaticalité ou de précision lexicale, l'apprenant semblant tester ses connaissances en passant d'une forme à l'autre:

- (8) a. je n'aime pas coûter [///] couper [///] coûter d'argent [1821]
 b. son [/] <sa père> [/] &hum son père [2021]

Le résultat final de la correction dépend de l'état du moniteur grammatical⁸.

Ces autocorrections, selon le CECR (2001), constituent une stratégie discursive consciente et intentionnelle. Ceci est confirmé par un comportement typique des apprenants jamaïcains de s'excuser d'avoir fait une erreur:

⁸ Le Moniteur grammatical est, selon Krashen (1982), un module d'autocontrôle conscient, constitué de règles explicites apprises par l'apprenant. Comme noté par un évaluateur, cela pose la question de la norme enseignée aux apprenants.

- (9) les <lire que j'aime> [/] &hum livres, *désolée*, que j'aime &hum sont &hum fantastiques [0821]

Tout cela ouvre une fenêtre aussi bien sur l'état de la compétence lexicogrammaticale inconsciente que sur l'état du moniteur grammatical conscient.

4.3 Faux départ

Un faux départ, codé [/ -], représente la situation d'un énoncé interrompu et laissé inachevé, suivi d'un redémarrage avec un nouvel énoncé lexicalement et syntaxiquement distinct.

- (10) a. <j'aime> [/ -] elle me plaît [0811]
 b. <mais c'est> [/ -] je ne sais pas &hum pourquoi, *mais* &hum c'est probablement à cause du fait que Beyoncé, <elle est> [/] &hum mm elle est dans les films maintenant [2021]

Blanche-Benveniste (2010: 53) identifie des "productions par bribes" au cours desquelles, souvent après une incise ou une modalisation, le locuteur reprend plus loin la construction qu'il semblait avoir abandonnée (comme en 10b).

Chez des apprenants, les faux départs qui se terminent sur un MO sont souvent liés à un problème insoluble de recherche lexicale ("alcoolique" en 11a), qui peut déboucher sur une impasse (en 11b). En (12), le locuteur gagne du temps en se posant des questions en aparté, comme indiqué par la baisse d'intonation (du formant F0), jusqu'à trouver le mot recherché.

- (11) a. <il est un> [/ -] il boive trop de vin [0831]
 b. mm *je* [/ -] par exemple, quand les autobus passent sur la route, *je* +... [0831]

(12) Transcription: Extrait de 0831_Conv

- 01 *STU: <c'est> [/ -], comment [/] comment dit on?
 02 *STU: quel est le nom de &sE langue [/], cette langue?
 03 *STU: <c'est un> [/ -] (..) saramaccan !

5. Synthèse

Le tableau récapitulatif qui suit présente, par apprenant et par niveau, le nombre de répétitions et de reprises régressives (autocorrections et faux départs mis ensemble) et, entre parenthèses, le nombre de mots inclus dans les séquences d'origine du phénomène considéré, ainsi que, sur une deuxième ligne, la fréquence en mots des différents phénomènes, c'est-à-dire le nombre moyen de mots qui apparaissent entre deux occurrences d'un type de reprise (nombre total de mots prononcés divisé par le nombre d'occurrences + 1 de reprises) et, entre parenthèses, le pourcentage de mots inclus dans les séquences d'origine de reprises par rapport au nombre total de mots d'apprenants produits (voir Witton-Davies (2010) pour l'utilité de distinguer ces deux mesures).

Pour les répétitions, il n'y a aucune évolution individuelle entre les niveaux II1 et III1: les trois locuteurs ont des fréquences stables entre ces niveaux. Les répétitions sont néanmoins plus longues en III1 comme le montre le

pourcentage de mots de segments originaux (ainsi, pour L20, les mots répétés représentent 5.44% du nombre total de mots produits en II1 et 7.37% en III1). L'accroissement proportionnel du nombre de mots repris est dû aux répétitions de séquences plus longues (de MO ou de MP).

	L18	L08	L20
II1: # total de mots produits	689	596	551
# de répétitions (# de mots)	71 (79)	16 (21)	23 (30)
Fréquence (% de mots)	9.56 (11.46%)	35.05 (3.52%)	22.95 (5.44%)
# de repr. régressives (# de mots)	25 (37)	17 (34)	26 (47)
Fréquence (% de mots)	26.5 (5.37%)	33.11 (5.70%)	20.40 (8.52%)
III1: # total de mots produits	832	953	678
# de répétitions (# de mots)	79 (102)	27 (40)	29 (50)
Fréquence (% de mots)	10.4 (12.25%)	34.03 (4.19%)	22.6 (7.37%)
# de repr. régressives (# de mots)	36 (68)	29 (55)	13 (24)
Fréquence (% de mots)	22.48 (8.17%)	31.76 (5.77%)	48.42 (3.53%)

Fig.2: Fréquence des reprises et pourcentage de mots repris.

En outre, un des locuteurs, L18, produit plus de la moitié des répétitions (150/245) avec une fréquence moyenne d'une répétition tous les 10.07 mots au total, deux fois plus que L20 et trois fois plus que L08. Cela pourrait être dû au niveau de compétence: les étudiants les plus faibles répétant plus que les étudiants les plus avancés. Mais cette tendance ne se confirme pas entre L08 et L20, puisque L20 pourtant plus avancé que L08 produit proportionnellement plus de répétitions que ce dernier.

Il est utile ici de prendre en compte la fréquence des pauses remplies (*euh*, *hum*) en parallèle avec le nombre de répétitions, car ces diverses marques d'hésitation remplissent de fait la même fonction (Witton-Davies 2014): faire avancer le discours en évitant les pauses silencieuses, en même temps qu'on en planifie la suite.

	L18	L08	L20
II1: Fréq. (# de euh/hum/total)	11.29 (56/4/60)	6.54 (18/72/90)	9.83 (15/40/55)
III1: Fréq. (# de euh/hum/total)	10.02 (78/4/82)	10.95 (29/57/86)	12.32 (22/32/54)

Fig.3: Fréquence des pauses remplies.

Le locuteur L08 en produit beaucoup plus que les deux autres en II1 (une tous les 6.54 mots en moyenne) et plus que L20 en III1. Ce qui suggère que le choix de l'une ou l'autre marque peut faire partie du style d'un locuteur indépendamment de sa compétence en L2.

Quant aux reprises régressives, seul L08 a une fréquence relativement stable entre les deux niveaux. L20 produit deux fois moins de ces reprises en III1 qu'en II1 (passant d'une reprise régressive tous les 20.40 mots à une tous les 48.42 mots) et L18 en produit un peu plus (fréquence de 26.5 à 22.48) et de plus longues (de 5.37% à 8.17% des mots produits). En outre, L08 aux deux niveaux et L20 en II1 produisent plus de reprises régressives que de répétitions alors

que les natifs en produisent normalement moins (Candéa 2000, Hilton 2014). Cela pourrait correspondre chez les apprenants à une étape de prise de conscience accrue des erreurs suite au développement de connaissances plus sûres et donc à un accroissement des autocorrections. L20 produisant moins d'autocorrections au niveau III1 indiquerait alors une intégration plus poussée de ses connaissances menant à une diminution du nombre d'erreurs.

6. Conclusion

Dans cette recherche basée sur une analyse de corpus longitudinal de productions orales d'apprenants de FLE, nous avons décrit les reprises effectuées en conversations spontanées et observé que la fréquence de répétitions et de pauses remplies reflèterait un style personnel stable du locuteur, que les autocorrections conscientes sont surtout utilisées pour corriger des traits morphosyntaxiques, la fréquence de ce comportement étant vraisemblablement liée au développement du système grammatical, qu'un faux départ s'avère souvent causé par une recherche lexicale infructueuse, et que c'est bien un faisceau de facteurs (reprises et pauses remplies ensemble) plutôt qu'un facteur séparé qui peut être corrélé à la maîtrise langagière et à la perception d'une difficulté d'expression.

Ces observations permettent de mieux cerner la nature des stratégies intentionnelles de contrôle et d'autocorrection explicitement relevées par le CECR reposant sur des règles explicites contenues dans le moniteur grammatical, dont une étape est la prise de conscience accrue par l'apprenant de (certaines de) ses erreurs et dont l'étape ultime est la capacité de l'utilisateur à contrôler consciemment sa production et à se corriger sans heurt. Pour qu'elles passent inaperçues, l'apprenant devra produire moins d'autocorrections de traits morphosyntaxiques et plus de nature lexicosémantique. Cette évolution passe par le développement du système lexicogrammatical sous-jacent qui lui permettra de ne plus faire d'erreurs morphosyntaxiques en premier lieu et de se concentrer sur la tâche de préciser sa pensée. Néanmoins, il est utile avant ce stade de veiller à lire les réactions de l'interlocuteur, afin de ne pas multiplier les autocorrections qui ne sont nécessaires que lors d'un malentendu (et aussi, inutile de s'excuser) afin de ne pas interrompre le fil du discours.

Si les répétitions, surtout de MP, voire les pauses remplies de type francophone (*euh* plutôt que *hum*), utilisées avec modération, ne sont pas toujours néfastes à l'aisance de parole lorsqu'elles remplacent des pauses silencieuses à l'intérieur d'un énoncé, les apprenants peuvent être amenés à prendre conscience de ces phénomènes infra-conscients dans leurs comportements langagiers (par exemple, par l'analyse de leurs propres productions spontanées). Des recommandations pédagogiques pourraient aussi encourager l'adoption de stratégies réduisant l'usage des hésitations et reprises

lorsqu'il s'agit de gérer des difficultés dans l'activation du vocabulaire ou de la construction adéquate. Outre les bénéfices démontrés du temps de préparation ou de la répétition de tâche sur l'aisance de parole (de Jong & Perfetti 2011), ces stratégies pourraient faire l'objet d'un entraînement spécifique en classe de langue, notamment quant à l'emploi d'un répertoire de locutions toutes faites (*comment dit-on?, je ne sais pas pourquoi mais...*) ou de connecteurs (*eh bien, bon, okay*) (Dörnyei 1995; Tavakoli et al. 2016). En outre, le traitement de l'autocorrection soulève la question de la place de la norme dans la conception de l'enseignement et oblige l'enseignant(e) à choisir l'aspect privilégié de la performance à développer à un moment donné: la correction grammaticale (quelle norme et pour qui ?), la complexité syntaxique ou, plus à propos, l'aisance de parole.

Finalement, l'examen de l'ensemble du corpus devrait nous aider à affiner nos observations notamment en comparant les situations de communication informelles à celles en contexte d'examen. Le stress de cette dernière situation devrait normalement accroître le nombre de marques de TdF.

BIBLIOGRAPHIE

- Blanche-Benveniste, C. (2010). *Approches de la langue parlée en français*. Paris: Ophrys.
- Candéa, M. (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes dits "d'hésitation" en français oral spontané*. Thèse de doctorat, Université de Paris III.
- Conseil de l'Europe (2001). *Cadre européen commun de référence pour les langues*. Paris: Didier.
- de Jong, N. & Perfetti, C. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning* 61(2), 533-568.
- Dörnyei, Z. (1995). On the teachability of communication strategies. *TESOL Quarterly*, 29, 55-85.
- Freed, B. (2000). Is fluency, like beauty, in the eyes and ears of the beholder? In H. Riggenbach (éd.), *Perspectives on fluency* (pp. 243-265). Ann Arbor: University of Michigan Press.
- Hilton, H. (2014). Oral fluency and spoken proficiency: considerations for research and testing. In P. Leclercq, A. Edmands & H. Hilton (éds.), *Measuring L2 proficiency: Perspectives from SLA* (pp. 27-51). Bristol: Multilingual Matters.
- Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah: Lawrence Erlbaum.
- Préfontaine, Y. & Kormos, J. (2016). A qualitative analysis of perceptions of fluency in second language French. *International Review of Applied Linguistics*, 54(2), 151-169.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Lennon, P. (1990). Investigating fluency in EFL. *Language Learning*, 40(3), 387-417.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum Associates.
- McKee, G., Malvern, D. & Richards, B. (2000). Measuring vocabulary diversity using dedicated software. *Literary and Linguistic Computing*, 15(3), 323-337.
- Michel, M. (2011). Effects of task complexity and interaction on L2 performance. In P. Robinson (éd.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 141-173). Amsterdam; Philadelphia: John Benjamins.

- Olynyk, M., D'Anglejan, A. & Sankoff, D. (1990). A quantitative and qualitative analysis of speech markers in the native and second language speech of bilinguals. In R. Scarcella, R. Anderson & S. Krashen (éds.), *Developing communicative competence in a second language* (pp. 139-155). Rowley, MA: Newbury House.
- Péters, H. (2009). Développement d'un corpus oral d'apprenants. *UPR Working Papers in Linguistics*, 2(2), 21-32.
- Péters H. (2014). Étude sur la complémentation infinitive des verbes de déplacement sur la base d'un corpus de productions orales d'apprenants jamaïcains de FLE et suggestions pédagogiques. In L. de Serres, F. Ghillebaert, P-A. Mather & A. Bosch (éds.), *Aspects culturels, linguistiques et didactiques de l'enseignement-apprentissage du français à un public non-francophone* (pp.62-84). Québec: Association Internationale des Études Québécoises.
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge.
- Tavakoli, P. (2016). Fluency in monologic and dialogic task performance: Challenges in defining and measuring L2 fluency. *International Review of Applied Linguistics*, 54(2), 133-150.
- Tavakoli, P., Campbell, C. & McCormack, J. (2016). Development of speech fluency over a short period of time: Effects of pedagogic intervention. *TESOL Quarterly*, 50(2), 447-471.
- Tavakoli, P. & Skehan, P. (2005). Strategic planning, task structure and performance testing. In R. Ellis (éd.), *Planning and task performance in a second language* (pp. 239-277). Amsterdam: John Benjamins.
- Temple, L. (2000). Second language learner speech production. *Studia Linguistica*, 54(2), 288-297.
- Witton-Davies, G. (2010). The role of repair in oral fluency and disfluency. *Selected papers from the 19th international symposium on English teaching* (pp. 119-129). Taipei: Crane Publishing.
- Witton-Davies (2014). Oral fluency development over four years at a Taiwanese university. *Selected Papers from the 23rd international symposium on English teaching* (pp. 134-145). Taipei: Crane Publishing.

Annexes

Conventions de transcription

Nous suivons les conventions de CHAT du système CHILDES (MacWhinney 2000)

Notation de la parole

&	transcription phonétique	(.) (..)	micro-pauses
&=laugh	évènement non linguistique	xx	segment inaudible
< >	délimitation des phénomènes	:	allongement vocalique
[/]	répétition	[//] [///]	autocorrection
[-]	faux départ	@s	emprunt à la L1