

L'analyse des corpus d'énoncés de locuteurs allophones : une démarche pour former les enseignants à la langue et à la didactique de la langue

CAMUSSI-NI Marie-Armelle CELLAM EA 3206

Université Rennes 2

Pour citer cet article : Camussi-Ni Marie-Armelle (2024). L'analyse des corpus d'énoncés de locuteurs allophones : une démarche pour former les enseignants à la langue et à la didactique de la langue. *Scolagram* n°11. Novembre 2024. Enjeux des corpus de grammaire en didactique de la langue pour la formation des enseignants. En ligne : https://scolagram.u-cergy.fr/index.php/content_page/item/1354-l-analyse-des-corpus-d-enonces-de-locuteurs-allophones-une-demarche-pour-former-les-enseignants-a-la-langue-et-a-la-didactique-de-la-langue

Résumé : Cet article met en perspective une démarche de formation de futurs enseignants de Français Langue Étrangère et Seconde (FLES) à partir d'extraits de corpus d'apprenants allophones avec les recherches sur corpus menées en FLES et en acquisition des langues étrangères, en particulier sur corpus d'apprenants, articulés aux différents points de vue sur l'erreur qui ont pu avoir cours dans ce domaine. Il s'agit, dans ce cadre, de témoigner d'une démarche originale d'utilisation de corpus pour la formation d'enseignants. En effet, les corpus d'apprenants allophones permettent d'extraire des écarts à la langue-cible récurrents et une méthode d'analyse permet de rechercher les représentations de la langue qui leur sont sous-jacentes, à la façon des travaux des acquisitionnistes qui évaluent les productions langagières dans les micro-systèmes que peuvent se construire logiquement les apprenants selon leur degré de maîtrise de la langue-cible. Les énoncés erronés organisés dans une progressivité de dévoilement des structures langagières auxquels ils font défaut peuvent constituer ainsi le support d'une formation de futurs enseignants qui conjugue linguistique et didactique de la langue. Cette démarche s'appuie sur une théorie linguistique dont la portée euristique soutient la démarche, celle de la théorie de la médiation.

Mots clés : corpus d'apprenants, corpus d'apprentissage, formation des enseignants, grammaire, obstacles d'apprentissage, didactique du FLES, théorie de la médiation.

Les dernières décennies ont vu croître les propositions didactiques en Français Langue Étrangère et Seconde¹ (FLES) appuyées sur des corpus de langue. Il s'agissait tout d'abord de mesurer la distance entre les énoncés fabriqués dans les manuels de langue et la réalité de l'emploi des natifs (par exemple, Delahaie, 2008) puis de proposer directement les corpus comme support d'apprentissage dans la mouvance du *data-driven learning* (Johns, 1986) ou de l'apprentissage sur corpus (Boulton et Tyne, 2014), voire même les métadonnées des corpus pour repérer les contextes des variantes (André, 2020)

Les corpus d'apprenants, quant à eux, voient un développement plus tardif et sont moins sollicités pour la didactique de la langue, en particulier en Français Langue Étrangère. Parallèlement, les chercheurs en acquisition du FLE (Selinker, 1972 ; Corder, 1980 ; Perdue, 1980 ; Py, 1980 ; Véronique *et al.*, 2017) ont élaboré des corpus qualitatifs, longitudinaux ou transversaux, dans une démarche empirique de mise au jour de stades d'acquisition. Dans cette perspective, les écarts à la langue cible sont les témoins d'une réflexion sur la langue qui se construit progressivement. Quoiqu'il en soit, les écarts à la langue cible fréquents ont sans doute toujours été pris en compte, empiriquement, par les enseignants de langue étrangère. Aussi nous semble-t-il important d'interroger la place que ces corpus d'énoncés allophones peuvent prendre dans la formation des enseignants.

Cependant, les objectifs de ces recherches et de la didactique se disjoignent comme le souligne Marquillo Larruy (2003) et la question de la manière dont elles peuvent venir innover la formation des enseignants est trop rarement posée (à l'exception de Véronique, 2005 et 2017). Perdue estime que les erreurs sont le résultat d'« une activité cognitive qui consiste à faire des hypothèses sur la structure de la langue cible » (Perdue, 1980, 93). Aussi la réflexion sur les erreurs pourrait-elle consister à examiner de quel échantillon de la langue a pu disposer l'apprenant pour parvenir à telle ou telle hypothèse et de quel échantillon il aurait besoin pour déplacer cette hypothèse. Ce qui revient à dire qu'au corpus d'énoncés erronés répond un corpus d'énoncés auquel on pourrait exposer l'apprenant pour tenter d'anticiper l'erreur.

Nous voudrions donner une idée de la logique d'un enseignement destiné à de futurs enseignants FLES qui se base sur ces principes et qui s'appuie sur un corpus d'énoncés d'allophones analysés et classés que nous utilisons depuis une quinzaine d'années dans le cadre d'un enseignement universitaire. Ces énoncés sous-tendent l'enseignement à partir de deux entrées : la méconnaissance de ce qui structure la langue (structuration des unités nominales ou verbales, contraintes de la syntaxe) ; la méconnaissance de ce qui vient brouiller ce repérage des éléments structurants de la langue (homophonie, allomorphisme, fragmentation de morphèmes, amalgame,

¹ Le terme FLES sera privilégié dans cet article. Il faut le comprendre comme englobant les situations très différentes dans lesquelles peut se trouver un apprenant allophone. Le terme FLE est réservé à des recherches ciblées sur des apprenants en contexte non francophone.

absence significative). Nous verrons que l'analyse de ces énoncés permet de mettre en évidence la logique de la langue à enseigner, les moyens d'analyse dont dispose l'enseignant et l'intelligence des erreurs commises par les locuteurs allophones.

Mais avant cela, nous ébaucherons un tour d'horizon des différentes perspectives autour de la « langue de l'apprenant » et de ses écarts à la langue cible, basée sur l'étude de corpus.

1. Les corpus dans la recherche en FLES et la place de l'erreur

Le statut de l'erreur en FLES est spécifique car l'erreur est incontournable dans ce domaine tant elle est massive et généralisée, particulièrement aux niveaux débutants. Aussi la réflexion sur ces erreurs et leur correction a-t-elle toujours fait partie des questionnements sur l'enseignement du Français Langue Étrangère (FLE) et on peut en trouver des exemples très anciens, comme le soulignent Cuq *et al.* (2003, p. 86), citant les premiers travaux de Bibbesworth en 1296, de John Palsgrave en 1532 ou de Claude Mauger, fin 17^{ème}, début 18^{ème}². La question est toujours régulièrement l'objet de publications de didacticiens du FLES (voir, par exemple, Pourquoi, 1977, Reichler-Béguelin *et al.*, 1998, De Salins, 1996, Cuq *et al.*, 2003, Marquilló Larruy, 2003, Cuq et Gruca, 2003...).

Cependant, ce sont différentes perspectives sur l'erreur qui ont été développées, visant des objectifs différents, somme toute complémentaires, comme nous allons le voir.

1.1 Le repérage systématique des erreurs en FLES

1.1.1 Une démarche « naturelle »

La démarche qui consiste à répertorier les erreurs les plus fréquentes liées à un calque de la langue première pour les corriger ou les anticiper en priorité paraît « naturelle » et est souvent revendiquée. De fait, les enseignants chevronnés face à un public monolingue repèrent aisément ces interférences et adoptent spontanément cette démarche. Aussi celle-ci a-t-elle été fréquemment un objet de recherches ou de classements à partir de corpus constitués *in situ* (pour un inventaire, Noyau et Véronique, 1986) et le reste encore (Debrenne, 2011 ; Dubois *et al.*, 2014, Vergne, 2020).

Dans la formation des enseignants, l'analyse de calques de langue peut avoir sa place pour des enseignants souhaitant exercer dans un milieu unilingue, par exemple pour des étudiants souhaitant devenir professeur de FLE dans un pays anglophone. C'est moins vrai dans le cadre de la formation d'enseignants susceptibles de rencontrer un public international, aux langues sources multiples, comme dans les

² Voir aussi le panorama historique de l'enseignement des langues de Germain (1993).

cours de FLE des universités françaises ou dans les dispositifs d'accueil des élèves allophones de l'éducation nationale en France.

Cependant, ce recueil d'erreurs a pris une autre perspective dans le cadre du développement des corpus électroniques.

1.1.2 Les corpus électroniques dans l'enseignement du FLES

1.1.2.1 Les corpus de langue cible

Soulignons tout d'abord que cette pratique renouvelée s'inscrit dans un cadre plus large correspondant à une démarche ambitieuse d'apprentissage sur corpus (*data driven learning*) des apprenants FLES (Boulton *et al.*, 2013 ; Boulton et Tyne, 2014). Il s'agit de former les apprenants à utiliser des corpus de la langue cible pour leur permettre d'avoir accès à des exemples de langue française attestés (citons par exemple, CLAPI-FLE développé par l'équipe d'ICAR à Lyon), mais aussi de répondre précisément à leurs besoins, dans les contextes où eux-mêmes évoluent, en tenant compte des contextes de production qui peuvent être analysés grâce aux métadonnées, comme dans le corpus Fleuron (André, 2020) et en définitive, d'être autonomes. Il s'agit alors de faire des étudiants des « chercheurs en linguistique ou en sociolinguistique de corpus ». V. André souligne que cela pose cependant la question de la formation des enseignants pour qu'ils sachent guider les premiers pas des apprenants dans la découverte du corpus³ (André, 2020).

1.1.2.2 Les corpus d'apprenants

Les corpus d'apprenants allophones, apparus à la fin des années 1980 dans leur forme de recueil électronique (Granger, 2004), sont plus rarement utilisés directement auprès des apprenants allophones, mais davantage destinés à des études acquisitionnistes. En ce qui concerne les corpus qui visent une analyse quantitative des erreurs, ce sont surtout des corpus d'apprenants en langue anglaise qui ont été développés et outillés à ce jour, dont certains cherchent à distinguer parmi les erreurs récurrentes celles qui sont spécifiques à un public d'une langue 1, rendant toujours actuelle la question des interférences. À leur suite, des corpus d'apprenants de différentes L2 ont commencé à être développés, dont des corpus de FLE, qui restent relativement fragmentaires, même si certains projets permettent de les réunir comme le projet *Clarín*, une infrastructure européenne qui met à disposition des ressources sur les langues dont des corpus numériques d'apprenants⁴, la plateforme *Ortolang*⁵ qui réunit plusieurs corpus d'allophones, ou encore la *Base de données de Louvain* qui regroupent des corpus d'apprenants de plusieurs langues cibles⁶.

³ Taous *et al.*, dans ce numéro, esquissent quelques pistes pour former les enseignants de FLES à cette nouvelle posture et proposent un modèle de didactisation conduisant progressivement à l'autonomie des apprenants.

⁴ <https://www.clarin.eu/resource-families/L2-corpora>

⁵ <https://www.ortolang.fr/market/corpora>

⁶ <https://uclouvain.be/fr/instituts-recherche/ilc/corpus-linguistiques-et-mediatiqes.html>

Ces corpus d'apprenants en FLES sont encore peu nombreux et peu accessibles. De fait, à cause de leur distance à la norme, ce type de corpus pose de grandes difficultés d'annotations automatiques. Benazzo et Watorek (2021) montrent aussi que les ambiguïtés posées par les corpus oraux posent des problèmes de fiabilité de leur transcription sous la forme orthographique et donc des dilemmes entre accessibilité et justesse de la transcription.

En outre, on peut s'interroger sur la destination des corpus d'apprenants. Il peut sembler périlleux d'exposer des apprenants allophones à des énoncés erronés, le risque étant qu'ils mémorisent ces formes ou même, pour les plus débutants, qu'ils n'en comprennent pas le statut erroné. De fait, les ouvrages qui ont tiré parti des recherches sur corpus d'apprenants de la langue anglaise (ALE) sont destinés à des apprenants de niveau avancé. De Cock et Tyne (2014) en donnent trois exemples : deux dictionnaires enrichis de notes d'erreurs, le Macmillan English Dictionary for Advanced Learners (MED2 ; Rundell, 2007), le Louvain English for Academic Purposes Dictionary (Lead) ("dictionary-cum-writing aid" – Paquot, 2012), et un programme de remédiation en ligne pour apprenants sinophones étudiant dans des universités américaines, GrammarTalk (Chuang et Nesi 2006). Les résultats des recherches sur corpus d'apprenants servent alors un objectif de remédiation, lequel pourrait se faire en autonomie ou dans un usage encadré par des enseignants.

Cependant, sur le plan méthodologique, De Cock et Tyne (2014) suggèrent qu'il est plus efficace, en situation d'apprentissage, de travailler sur des corpus locaux, « sur l'observation de ses propres productions ou des productions que l'on aurait pu produire ». Or, poursuivent-ils :

Ce genre d'approche présuppose que les enseignants ne soient pas complètement étrangers à certains aspects de la linguistique de corpus grâce à des programmes de formation continue ou de publications (comme Millar & Lehtinen, 2008, par exemple). Cependant, il n'est pas évident des enseignants, qui n'ont du reste pas tous un goût prononcé pour la linguistique, qu'ils deviennent des linguistes sur corpus. Voilà pourquoi pour certains une approche plus "soft" (cf. Gabrielatos, 2005), ou cherchant simplement à mettre en avant un certain "esprit de corpus" (Boulton & Tyne, 2014 : 81) serait plus favorable (du moins dans un premier temps).

Nous rallions ce point de vue, l'écart est grand entre les possibles développés par les chercheurs et la réalité des usages didactiques sur le terrain. C'est pourquoi nous estimons que les recherches sur les corpus d'allophones présentent leur plus grand intérêt quand ils sont destinés à la formation des enseignants qui pourront reconnaître les types d'erreurs chez leurs apprenants et transposer les analyses. Par le biais des enseignants, les publics débutants peuvent alors bénéficier des recherches sur corpus d'apprenants. C'est ainsi que ces corpus autorisent à viser une anticipation et non pas seulement une remédiation des erreurs. Enfin, ces compétences développées chez les enseignants permettent aux apprenants d'avoir une interface réactive au plus près de leurs besoins. Nous estimons, en effet, que l'échelon « enseignant » est une pièce maîtresse du système de l'enseignement-

apprentissage et qu'il est important de ne pas le négliger dans les efforts de progrès de la formation.

Cependant, d'autres types de travaux sur corpus d'apprenants allophones peuvent être utiles à la formation des enseignants de FLES : ceux sur l'acquisition de la langue.

1.2 Les perspectives acquisitionniste et sociolinguistique sur l'erreur

1.2.1 La disparition de l'erreur

Une autre perspective sur l'erreur, parallèle, considère que les erreurs des apprenants allophone, somme toute, n'en sont pas mais sont plutôt des constructions transitoires qui témoignent d'une appropriation en cours d'une langue étrangère. Depuis les années 1970, le concept d'interlangue défini tout d'abord par Selinker (Selinker, 1972, Selinker *et al.*, 1975) est devenu central pour traiter de la question de l'écart à la langue cible (Py, 1980 ; Rattunde, 1980 ; Vogel, 1995 ; Galligani, 2003 ; Bartning et Schlyter, 2004 ; Véronique *et al.*, 2017). Il s'agit de la perspective acquisitionniste qui vise à définir des stades dans l'acquisition, des séquences de développement de notions et à étudier la mosaïque de micro-systèmes (Py, 1980) qui se mettent en place et se reconfigurent au cours de l'acquisition. Ces analyses qualitatives sont conduites à partir de corpus longitudinaux, quasi longitudinaux ou transversaux (Véronique *et al.*, 2017, p. 39) qui ont été élaborés dans le cadre de grands programmes depuis les années 1970 : par exemple, l'ESF, le Corpus Lund, InterFra, FLLOC⁷. Il s'agit de suivre sur un temps long l'évolution d'apprenants (parfois en nombre très réduit), qu'ils soient en apprentissage guidé ou spontané, pour repérer les étapes par lesquels ils passent, les systèmes provisoires qu'ils construisent, le rythme de leur apprentissage. La focale est posée sur les réussites, leurs prémisses et leur consolidation, plutôt que sur l'erreur.

Ces travaux rencontrent une perspective sociolinguistique qui, mettant en avant l'existence de variantes et récusant l'idée d'une norme unique, reconnaît la langue de l'apprenant comme une variante non native du français (Galligani, 2003, Cadet et Guérin, 2012...).

Ces perspectives ont toute leur place dans la formation des enseignants de FLES : l'approche acquisitionniste peut permettre⁸ de poser des jalons dans l'apprentissage, d'en penser la progressivité, d'analyser les micro-systèmes provisoires des apprenants et de dédramatiser les écarts par rapport à la langue-cible ; la perspective sociolinguistique en permettant de mettre à distance le concept de « norme » et d'introduire la notion de variantes, légitime à son tour l'existence des écarts à une norme ou à des normes aux frontières bien moins rigides, de même que

⁷ Pour en connaître les détails et les publications qui en ont découlé, voir Véronique *et al.* (2017).

⁸ Soulignons que la prise en compte des travaux des acquisitionnistes du FLE en didactique de la langue reste un vœu pieu selon Véronique *et al.* (2017) qui déplorent le cloisonnement entre disciplines de recherche.

le plurilinguisme qui valorise les « parlars bilingues » (Auger, 2005 ; Galligani, 2022). La formation des enseignants peut en bénéficier : ces points de vue sont utiles pour penser une progression dans l'apprentissage de la langue qui intègre des objectifs réalisables à chaque stade d'apprentissage, pour tolérer des écarts à la langue-cible correspondant à un niveau d'acquisition, pour repérer les variantes à l'œuvre dans les productions des allophones et leur permettre de valider leurs contextes d'exploitation et *in fine*, considérer la langue comme un ensemble de possibles plutôt qu'un idéal normé.

1.2.2 La visée didactique du terme de l'apprentissage

Pour autant, même si l'enseignant révisé son appréciation des erreurs dans un parcours d'apprentissage et relativise la norme, son objectif dans la perspective d'un enseignement du FLES est bien de permettre à un apprenant de s'approcher au plus près de la langue-cible comme le rappelle Marquilló Larruy (2003), d'en saisir la norme ou les normes et leurs contextes d'application, quitte ensuite à jouer de ces variantes et de l'atypie qu'offre le bilinguisme comme le suggèrent Devaux (2005), Suchet (2014) ou encore Auger et Le Pichon-Vorstman (2021).

Cette exigence demande donc à compléter la perspective descriptive des stades d'acquisition en élaborant la question du terme de l'apprentissage, en la mettant en regard avec la notion de la fossilisation, en examinant l'influence des situations d'apprentissage sur les microsystèmes en place. En effet, il arrive que des microsystèmes construits à un stade de l'apprentissage perdurent dans la mesure où l'apprenant peut choisir d'arrêter son apprentissage de façon précoce mais aussi parce que peut se produire une fossilisation, c'est-à-dire une situation où l'apprenant automatise une forme éloignée de la langue-cible et ne parvient plus à se corriger. Dans le cadre du concept d'interlangue et à propos d'acquisition spontanée du FLE, Galligani (2003) s'interroge : comment nommer la langue des apprenants qui ont décidé d'arrêter d'apprendre alors que leur usage témoigne encore d'écarts importants à la langue ? Elle propose le terme de « parler bilingue ». Cependant, dans le cadre de l'enseignement de la langue, s'il s'agit de respecter la langue de l'apprenant, l'objectif est de mettre en œuvre des stratégies pour lui éviter ces fossilisations.

En effet, ces fossilisations qui se produisent en cas d'arrêt précoce de l'apprentissage, font vivre au locuteur un conflit cognitif. Que ce dernier ait seulement conscience d'une anomalie du microsystème employé sans pouvoir se donner les moyens de le déconstruire ou qu'il connaisse la forme à employer sans pouvoir se défaire de son usage inapproprié (Galligani, 2003), l'apprenant ressent un sentiment d'empêchement lié à la langue. En outre, certes, il communique et se fait entendre mais jusqu'à un certain seuil de tolérance qui correspond aux capacités de concentration de son interlocuteur. À cela s'ajoute le risque de stigmatisation lié à la langue dans la société française. Même si l'enseignant parvient à se dégager de tout jugement, il doit pour autant en avoir conscience et en tenir compte dans son désir d'enseigner le français à un apprenant allophone.

1.3. Faire dialoguer les perspectives dans la formation des enseignants

La constitution de corpus de productions orales ou écrites d'apprenants allophones, qu'ils soient quantitatifs ou qualitatifs présente un grand intérêt pour repérer les représentations de la langue récurrentes dans l'apprentissage et utiliser cette connaissance comme un des socles de la didactique de la langue. Aussi importe-t-il de clarifier les perspectives en jeu⁹ qui, en dialoguant, sont susceptibles de favoriser l'enseignement-apprentissage et d'éviter la fossilisation d'erreurs. Cependant, ce dialogue entre sous-disciplines différentes de la linguistique appliquée au FLES reste encore un horizon désirable comme le souligne Véronique (2017) qui appelle de ses vœux l'exploitation des travaux acquisitionnistes en didactique de la langue, ce qui pourrait irriguer la formation des enseignants de FLES. Comme l'indique Weber (2021), « les manifestations hétérogènes [de l'interlangue] sont d'ailleurs déroutantes pour l'enseignant, s'il n'est pas sensibilisé pendant sa formation à toutes ces caractéristiques ». En tout état de cause, former les enseignants à prendre en compte les représentations que l'apprenant se fait de la langue-cible apparaît comme un élément central de la formation. Ce que nous proposons ci-dessous est très embryonnaire par rapport à ces perspectives de croisement des recherches pour la didactique du FLES.

Cependant, nous verrons que cette proposition pour la formation des enseignants de FLES, en rapportant les erreurs à la langue-cible et à ce qu'elle laisse paraître de sa logique en fonction des échantillons de langue auquel l'apprenant a été exposé, permet d'outiller des enseignants s'adressant à des publics internationaux, que ce soit dans le supérieur ou dans des classes de l'éducation nationale, tout en leur proposant une approche de l'intelligence de l'erreur qui évite un point de vue normatif et stigmatisant.

Nous voudrions montrer dans la suite de cet article comment on peut structurer une formation en didactique de la grammaire pour de futurs enseignants de FLES à partir d'extraits de corpus de production d'apprenants allophones. Le corpus utilisé a été constitué, de façon empirique, dans les années 2000, à partir des tests d'admission des étudiants d'un centre universitaire de langue (le CIREFE, Rennes 2) auxquels se sont ajoutées des rédactions d'étudiants de niveau B1-B2¹⁰ du même centre ainsi que des productions orales et écrites d'apprenants en milieu non institutionnel de niveau A0-A1 (mères d'élèves allophones dans le cadre d'un projet d'école à l'école maternelle de Cleunay, Rennes). L'analyse du corpus a porté sur les erreurs récurrentes en morphosyntaxe ainsi qu'en sémantique pour extraire des énoncés regroupés en fonction de ce qui fait obstacle à l'accès à la langue-cible. Dans le cadre de cet article sont présentés quelques énoncés utilisés pour argumenter et exemplifier l'usage d'un corpus analysé dans le cadre d'un enseignement de futurs professeurs de FLES. Le principe retenu est en effet que c'est à travers ce qui fait obstacle au

⁹ Cette clarification des différents points de vue est aussi l'objet de *L'interprétation de l'erreur* (Marquillo Larruy 2003).

¹⁰ Niveau intermédiaire selon le CECRL.

repérage du système de la langue que se met au jour ce système. Il est clair que cette démarche qualitative d'analyse de corpus pourrait être affinée par la mise en perspective de ces extraits de corpus avec les séquences d'acquisition mises au jour par les recherches sur l'acquisition d'une langue étrangère ainsi que par l'usage statistique de grands corpus d'allophones comme ceux du ALE (Anglais Langue Étrangère).

2. Un corpus d'énoncés d'allophones pour enseigner la langue

Il existe un mythe selon lequel il suffirait de parler une langue pour pouvoir l'enseigner. Or l'usage ne donne pas accès automatiquement à la structure et la confrontation à des apprenants allophones en fait nécessairement la démonstration. Le locuteur simple usager de sa langue peut dire qu'un énoncé n'est pas correct mais souvent ne saura pas expliquer pourquoi, ce qui laisse l'apprenant allophone dans une forme de frustration mais surtout lui interdit de réinvestir la correction proposée qui n'est pas généralisable tant que l'erreur n'est pas analysée.

Aussi est-il légitime d'estimer que des enseignants de langue étrangère doivent être un minimum « linguistes ». Former de futurs enseignants à partir des erreurs d'apprenants allophones leur permet déjà de comprendre cette différence entre « correction » et « analyse d'erreurs », entre « usage de la langue » et accès aux régularités de la langue qui correspondent à sa structure.

Nous l'avons vu, un certain nombre d'erreurs des apprenants allophones sont intralinguales (Besse et Porquier, 1984), celles-ci ne peuvent pas être comprises en prenant en compte la langue première mais bien la langue cible. L'analyse de ces erreurs demande d'adopter le point de vue sur le français d'un apprenant qui n'en a pas l'usage, en imaginant les représentations qu'il peut s'en faire à partir de ce que le français laisse à voir, voire la grammaire personnelle, provisoire ou fossilisée, qu'il se construit. Ces hypothèses à partir de la langue cible concerneront nécessairement des erreurs interlinguales pour l'une ou l'autre des langues sources. L'enseignant qui se retrouvera dans un milieu monolingue pourra alors compléter sa première hypothèse d'une analyse appuyée sur la langue source de l'apprenant. L'analyse à partir de la langue cible sera l'unique ressource quand l'enseignant aura à charge un public immigré en France, dans des centres de langue ou dans le cadre de l'Éducation Nationale – en UPE2A ou dans des classes ordinaires, du fait de la multiplicité des langues-sources en présence.

En outre, un corpus d'énoncés erronés peut constituer un moyen heuristique de comprendre le système de la langue *a contrario*, tout en mettant en contact de futurs enseignants avec des éléments authentiques du terrain et en leur donnant les moyens d'anticiper certaines de ces erreurs. Ainsi, former des étudiants, futurs enseignants de FLES, à partir d'énoncés allophones choisis pour leur aspect récurrent et systématisable sert une dialectique complexe : il s'agit de convaincre les futurs enseignants de FLES que la maîtrise du système de la langue est incontournable pour

un apprenant allophone tout en repérant ce qui vient brouiller son aspect systémique. Dès lors, se donner ces compétences d'analyse de la langue apparaît nécessaire pour l'enseignant qui doit répondre aux impasses de l'apprenant allophone mais aussi les limiter en les anticipant.

2.1. Comprendre le système de la langue

Comprendre le système de la langue, c'est-à-dire ce qui peut être appréhendé par les apprenants allophones pour soutenir leurs compétences langagières, c'est se décentrer de sa compétence liée à l'usage pour adopter le point de vue d'un apprenant allophone qui ne dispose pas (encore) de l'usage et qui, ne pouvant pas se contenter de l'immersion éventuelle dans la langue étrangère pour l'apprendre, doit s'appuyer sur les éléments de régularité qui structurent le langage, sur une compréhension explicite de la structuration de la langue.

Le premier objectif qu'on peut se donner auprès de futurs enseignants de FLES, c'est de leur montrer l'existence d'une structuration entre les mots qui ne sont pas simplement juxtaposés comme pourrait le laisser croire les blancs de la graphie.

Un corpus d'apprenants allophones débutants (niveau A0 ou A1 du CECRL) ou intermédiaires (B1) peut permettre d'extraire des énoncés mettant en jeu la structure de base de la langue, c'est-à-dire celle des unités de la langue. Il s'agit d'énoncés¹¹ comme :

- (1) *Frère Paris demain.* (oral)¹²
- (2) *Mais salade est frais et bon pour **vous** santé.*
- (3) ***Moi aller** de faire du ski.*
- (4) *Il joue du tennis avec **je**.*
- (5) *Je **étudié** à l'alliance française Rio de Janeiro.*
- (6) *Une infirmière emmène **lui** dans la salle d'opération.*
- (7) *Donc quelqu'un **a leur demander** de s'échapper.*

L'énoncé (1)¹³ permet de prendre conscience de l'agrammaticalité d'un énoncé constitué de juxtapositions de lexèmes. En l'occurrence, cette agrammaticalité empêche la construction d'un sens autre qu'approximatif (mon frère va à Paris demain ? Ton frère vient de Paris demain ? Son frère travaille à Paris demain ? etc...). Connaître du lexique apparaît bien insuffisant pour communiquer, la langue est

¹¹ Les énoncés présentés sont pour la plupart des énoncés écrits. Quand ils sont issus d'un corpus oral transcrit, le terme « oral » le spécifie à la suite. Les mots en gras indiquent ici les éléments sur lesquels est attirée l'attention des étudiants.

¹² Cet énoncé qui correspond à la variété pré-basique selon les études acquisitionnistes (Véronique *et al.*, 2017, p. 373) a été réalisé par une mère de famille résident depuis dix ans en France, ce qui témoigne du fait que la fossilisation de cette variété peut se produire chez des allophones qui bénéficient de très peu de contacts avec des locuteurs de la langue cible.

¹³ Notons que la démarche d'analyse des erreurs se double utilisation de l'analyse des acquis, ce qui ne pourra pas être déployé dans le cadre de cet article. On peut se référer aux analyses de Camussi-Ni et Coatével (2013, p. 11-12) pour en trouver un exemple.

structurée par des interdépendances. L'absence de verbe, de déterminant et de préposition pourra être repérée ici.

L'énoncé (2) rend compte de la structure des unités nominales par son écart à la norme. L'unité nominale nécessite que l'élément entre la préposition et le lexème, s'il figure, appartienne à une liste de déterminants. Le sens de *vous* ne peut suffire à assurer la grammaticalité de l'énoncé et à terme, la fluidité de la communication.

Les énoncés (3), (4), (5) rendent compte de la construction de l'unité verbale avec la présence d'un préfixe de personne (*je*) selon la terminologie adoptée par Creissels (1995) et par Camussi-Ni et Coatéval (2013)¹⁴.

La comparaison des énoncés (3) et (4) montre l'interdépendance du lexème avec le préfixe de personne (*je*) dans l'unité verbale : *je* ne peut pas être utilisé comme un pronom personnel autonome ; il nécessite d'être suivi d'un verbe. Et inversement, le verbe de l'énoncé (3) ne peut s'accommoder d'un pronom personnel autonome (*moi*), il nécessite un préfixe de personne comme *je*. Le choix des éléments devant le verbe est contraint. Il y a bien un schéma d'organisation.

L'énoncé (5) souligne le lien nécessaire entre le suffixe personnel du verbe et celui du préfixe de personne. Avec l'énoncé (3), il permet de mettre au jour la flexion des suffixes verbaux (en temps, mode, personne) qui échappe souvent aux apprenants débutants.

Enfin, les énoncés (6) et (7) rendent compte d'une contrainte de disposition des unités verbales, ce qui participe de leur structuration.

Une première étape de formation des enseignants passe donc par la reconnaissance d'une structuration basique de la langue qui pose des difficultés de communication quand elle n'est pas respectée. Cette structuration basique s'appuie sur le repérage de l'interdépendance de certains éléments qui, dès lors, se regroupent dans des ensembles autonomes, soumis à la variation. Ce repérage fait émerger de véritables schémas qui structurent la langue : ceux de l'unité nominale et de l'unité verbale présentés ci-dessous :

Préposition	déterminant	lexème	Suffixe de nombre
pour	votre	santé	Ø
sans	un	regard	Ø
avec	ces	idée	s

Schéma de l'unité nominale substantive

¹⁴ Cette terminologie rend compte à la fois de l'interdépendance de ce morphème avec le lexème mais aussi, dans notre acception, de sa corrélation avec le suffixe de personne.

prépo- sition	Con- jonction	Préfixe de personne	Préfixe d'asser- tion	Morphèmes Complé- mentaires	lexème	Suffixe mode/ temps	Suffixe de personne	Suffixe d'assertion
pour	qu'	ils	∅	le	dis	∅	ent	∅
dès	que	tu	∅	∅	part	ira	s	∅
∅	quand	je	n'	en	veu	∅	x	pas
∅	si	vous	ne	vous	lev	∅	ez	plus

Schéma de l'unité verbale déclarative

Ces schémas d'unité sont issus d'une théorie linguistique, la théorie de la médiation, qui les a mis au jour à partir d'une analyse d'une compétence glossologique qui opère par segmentation et par opposition¹⁵. Ils rendent compte de l'interdépendance de leurs éléments, de leur disposition contrainte ainsi que de la spécificité de certains d'entre eux en fonction de la catégorie. Ce sont donc des outils d'analyse qui permettent de s'abstraire de la présentation orthographique et des exemples en jeu. Dans ce cadre théorique, l'unité morphologique est un élément minimal de signification. Comme l'indiquent Bonnet et Barreau qui la définissent comme « la plus petite unité sémiologique sur l'axe du texte (axe horizontal des combinaisons) », « je ne puis, en parlant, dire moins » (Bonnet et Barreau, 1974, p. 66).

L'analyse qualitative du corpus s'appuyant sur cette théorie met nécessairement l'accent sur des distorsions par rapport à ces schémas. L'hypothèse théorique est que cette description linguistique est pertinente pour former des apprenants et leurs enseignants.

À terme, les analyses d'énoncés erronés mettent l'accent sur l'interdépendance et sur la disposition des éléments dans ces schémas structurants, ce qui permet aux futurs enseignants de les repérer. L'appréhension des schémas de l'unité nominale, de l'unité verbale et de leurs variantes permet de porter un autre regard sur la langue et de donner des éléments de réponse aux difficultés des apprenants FLE.

2.2. Repérer les obstacles à l'appréhension du système de la langue

Cependant, les éléments de structuration de la langue qui facilitent, par leur régularité, son apprentissage vont de pair, comme nous allons le voir, avec des phénomènes de brouillage qui, eux aussi, participent de la description de la langue.

Il s'agit donc pour les futurs enseignants de comprendre ce qui fait obstacle à l'appréhension de la langue, de façon interne à la langue et parfois externe, là aussi avec une visée systématique pour avoir accès à des microsystèmes transitoires parfois personnels, parfois partagés dans une communauté ou par tout allophone, mais qui risquent de se fossiliser. C'est pourquoi leur connaissance peut avoir pour

¹⁵ Pour plus de précisions sur cette théorie et sur les variantes de ces schémas, voir Bonnet et Barreau (1974), Pergnier (1986), Jongen (1993), Urien (1999), Camussi-Ni (2020).

objectif de remédier les erreurs mais surtout de les anticiper, la didactique pouvant être orientée par ces obstacles comme le suggère Astolfi (1992 et 1997).

Aussi proposons-nous de former les enseignants à quelques phénomènes qui peuvent fausser les représentations des apprenants dans le domaine de la morpho-syntaxe : l'homophonie, l'allomorphisme, l'amalgame, la fragmentation des morphèmes et l'absence significative. Nous allons détailler quelques exemples d'énoncés d'allophones permettant de mettre au jour ces phénomènes et des microsystèmes qui peuvent leur être associés.

2.2.1. L'homophonie

L'homophonie est un phénomène qui peut perturber de nombreux lieux d'appréhension de la langue. Nous la définirons par l'existence de segments phoniques identiques qui appartiennent à des listes de commutation différentes¹⁶. Ces homophones peuvent être homographes¹⁷.

Certaines homophonies s'exposent directement dans des énoncés allophones permettant ainsi de la définir auprès des futurs enseignants, comme *mais*, homophone de *mes*, dans l'énoncé (2) :

(2) ***Mais*** *salade est frais et bon pour vous santé.*

D'autres se repèrent facilement du fait qu'elles déclenchent des accords par la confusion d'un morphème complémentaire avec un déterminant (*les*) ou un préfixe de personne (*nous*) :

(8) ***Il les*** *changes.*

(9) *Il a répondu que ça **nous** permettions de prendre des week-ends de quatre jours.*

Ces homophonies viennent perturber alors le repérage de l'interdépendance des éléments de l'unité verbale, isolant *que*, par exemple, du reste de l'unité.

D'autres énoncés présentant des perturbations de la structuration morphologique demandent d'opérer des commutations pour repérer l'homophonie originale qui en a été la cause. C'est le cas des énoncés (5) et (6) :

(5) *Il joue du tennis avec **je**.*

(6) *Une infirmière emmène **lui** dans la salle d'opération.*

Dans l'énoncé (5), le préfixe de personne *je* ne présente aucune homophonie mais commute avec le préfixe de personne *elle* qui, lui, a un homophone homographe, pronom personnel autonome :

¹⁶ Nous excluons de cette définition les lexèmes dans la mesure où nous estimons qu'il n'existe qu'une seule liste de lexèmes. De ce fait, leur commutation ne permet pas de distinguer entre plusieurs listes. On distingue donc les lexèmes homophones par leur sens radicalement différent (par exemple : *une porte/je porte*).

¹⁷ Les homophones hétérographes constituent non seulement un obstacle orthographique pour les allophones mais aussi de compréhension à l'oral au même titre que les homophones homographes. Aussi ne mettrons-nous pas forcément l'accent sur cette différence entre les homophones.

(5bis) *Il joue du tennis avec **elle**.*

C'est cette homophonie qui a conduit l'apprenant à penser que les préfixes de personne peuvent être utilisés comme des pronoms personnels autonomes.

L'homophonie qui sous-tend l'erreur de l'énoncé (6), celle de *lui*, pronom personnel autonome, commutant avec *moi, toi ... eux, elles*, et de *lui*, morphème complémentaire commutant avec (*me, te, ... leur*) est susceptible d'être à la source de représentations erronées comme en témoignent ces énoncés :

(10) *Je **lui** discute.*

(11) *J'allais chez ma grand-mère. Avec **lui**, je jouais beaucoup.*

(12) *Malheureusement, on ne peut pas **eux** demander, ils sont morts comme la moitié de la population.*

(13) *Ils apportaient le grand poisson chez **leur**.*

L'énoncé (10) souligne que cette homophonie peut rendre caduque la notion de disposition puisque l'un des *lui*, le morphème complémentaire est antéposé au lexème de l'unité verbale quand l'autre *lui*, le pronom autonome, lui est postposé. L'apprenant allophone peut avoir l'illusion d'une seule liste de morphèmes qu'on peut placer à son gré avant ou après le lexème.

L'énoncé (11) montre que cette illusion peut conduire les apprenants à neutraliser le genre du pronom personnel autonome en le confondant avec le morphème complémentaire.

Enfin, les énoncés (11) et (12) témoignent de logiques opposées dépendant des échantillons de langue rencontrés par les apprenants. Ces logiques permutent des éléments des deux listes dans lesquelles commutent les deux homophones *lui*.

Aussi, l'analyse des sous-bassements de ces énoncés témoigne au futur enseignant de la nécessité de maîtriser l'outil de commutation, de penser de façon abstraite à l'ensemble du système de la langue pour comprendre en quoi l'erreur de l'apprenant à laquelle il est confronté est systémique et donc reproductible. Elle met aussi l'accent sur une terminologie qui ne distingue pas suffisamment des éléments qui entrent dans des interdépendances bien différentes. Le terme « pronom personnel » est ainsi trop générique et demande à être décliné par différents termes si on ne veut pas conduire un apprenant à une impasse.

Un autre phénomène d'homophonie sous-jacent, celui bien connu des suffixes /e/ peut être à la base de la construction de microsystèmes distincts de la langue cible mais aussi des microsystèmes construits par des francophones.

C'est le cas de l'énoncé (4) qui peut être relié à d'autres énoncés mettant en jeu la même homophonie et amenant à construire le même microsystème :

(4) *Je **etudié** à l'alliance française Rio de Janeiro.*

(14) ***Jai** rentré à la maison demain.*

(15) ***Jai** apprendre des chansons.*

L'apprenant a systématisé la présence d'un suffixe /e/ en fin de verbe et a constitué un système où la non-flexion du verbe peut faire souvent illusion à l'oral (car correspondant à ces formes : *j'ai étudié, vous étudiez, je vais étudier voire à j'étudiais*) et qui s'appuie alors sur le principe d'une variante phonologique du préfixe de la personne 1 *je/jai*. L'homophonie des suffixes *ez, er, é, ai* empêche de distinguer les formes composées et les formes simples. De son côté, l'interlocuteur francophone ne perçoit pas l'absence généralisée de flexion de l'apprenant, trompé lui aussi par l'homophonie des suffixes /e/, même si le microsystème n'est évidemment pas sans contradiction. Ainsi surgissent des indices comme la position de l'adverbe différente entre la forme composée et la forme simple (16), celle de la négation (17) ou des morphèmes complémentaires (18). Ces erreurs ne permettront pas de comprendre qu'un système sans flexion verbale s'est mis en place. Elles seront mises au compte de la disposition et de coquilles peu graves. Si un énoncé témoigne clairement de l'absence de flexion (16), il est noyé dans un ensemble de verbes apparemment bien fléchis. Quant à la discordance éventuelle entre un repère temporel du futur et un verbe apparemment conjugué au passé composé (19), c'est un indice fort mais il faut déjà connaître l'existence de ce microsystème pour pouvoir l'y relier¹⁸.

(16) *Les enfants toujours pleurer.* (oral)

(17) *Jai ne pleurer pas.* (oral)

(18) *Jai lui parler.* (oral)

(19) *J'ai préparé demain.* (oral)

Ces quelques exemples témoignent de l'intérêt de conduire une analyse des homophonies pour de futurs enseignants. Au-delà d'un contact avec l'authenticité des réalisations, il s'agit de reconnaître un obstacle majeur de l'apprentissage du FLE, l'homophonie, et de construire une partie son enseignement, au moins, pour déjouer les obstacles qu'elle produit.

2.2.2. L'allomorphisme

L'allomorphisme (ou existence de variantes) est un deuxième obstacle important du repérage de la structuration de la langue, d'autant plus quand l'écriture n'en atteste pas ou qu'il vient se combiner avec des phénomènes d'homophonie. Les erreurs d'apprenants allophones liées aux variantes mettent au jour le fait qu'ils n'ont pas conscience de l'existence de deux variantes (ou plus) pour un même morphème ou lexème. Des erreurs de production seront alors facilement repérées où l'apprenant généralise la variante la plus usuelle ou la seule enseignée :

(20) *Tout le monde voyait **ce** enfant et il ne faisait rien.*

(21) *J'ai raté **le** examen.*

(22) *Maintenant, j'ai un **nouveau** ami très gentil.*

¹⁸ Précisons que, pour notre part, c'est l'insistance d'une apprenante à interroger « C'est bizarre, vous les français, des fois, vous dites *je*, des fois vous dites *jai* » qui nous a permis d'accéder à ce microsystème pour ensuite en confirmer l'existence chez d'autres apprenants évoluant dans des milieux d'apprentissage non institutionnels. Cet éclairage est susceptible de modifier d'ailleurs certaines analyses sur les séquences de développement du verbe chez les acquisitionnistes.

(23) /*ilarivəRɔ̃ãrətɑR*/ (Il arriveront en retard) (oral)

(24) *Ils ne saient pas.*

L'un des enjeux de la présentation de ces énoncés à de futurs enseignants est de repérer que l'allomorphisme ne se présente pas comme un phénomène facilement appréhendable. Il ne touche pas qu'une seule catégorie de la langue (les déterminants, par exemple, comme en (20) et (21)) mais peut affecter aussi bien un lexème (22) et (24) qu'un préfixe de personne (23) ou tout autre morphème. Les variantes se repèrent parfois à l'oral comme l'absence de liaison (23) et demandent donc aux enseignants de se doter d'une analyse orale de la langue. En outre, ces variantes correspondent à une variété de phénomènes (élision, liaison ou forme différente) et s'organisent dans une variété de configurations logiques (c'est le cas des différents déterminants ou des lexèmes dans les conjugaisons). Ces variétés peuvent conduire à un autre type d'erreur, celle qui correspond à un choix erroné parmi des variantes possibles, soit par création d'une variante selon une des logiques possibles (25, 26), soit par usage d'une variante existante dans une autre configuration que celle de l'usage (27, 28) :

(25) *Est-ce qu'il y a **d'**invités ?*

(26) *À **s'**arrivée, chacune personne s'est levée.*

(27) ***Cet** bâtiment, on aperçoit beaucoup des fenêtres.*

(28) *Ils **veulaient** savoir si on était là.*

Les apprenants peuvent aussi ne pas tenir compte que les variantes ont des limites liées à des configurations syntaxiques pour (29), les variantes s'établissant à l'intérieur des unités¹⁹, ou à la voyelle élidée pour (30) :

(29) *J'ai un projet **nouvel** à exécuter.*

(30) ***S'**elle est d'accord, moi aussi.*

La voyelle i est rarement élidée de même que la voyelle u. Pour autant, il s'agit de distinguer la production de **s'** (30) de celle de **t'** pour **tu** ou de **qu'** pour **qui** qui sont des variantes attestées chez les francophones (*t'arrives ? C'est lui qu'a réussi*). Il ne s'agit donc pas d'analyser ces dernières comme des erreurs mais de les replacer dans des contextes d'emploi. Se limiter à une norme empêche de repérer ces variantes que l'apprenant allophone au contact de francophones rencontrera très probablement.

Or, ce type de variantes pose des difficultés supplémentaires en constituant des homophones, conduisant à un double brouillage. C'est le cas, également, des déterminants féminins devant initiale vocalique, *mon, ton, son* :

(31) *À **s'**arrivée, chacune personne s'est levée.*

(32) *En fait, elle est face à une vie qui est le contraire de **sa** envie.*

(33) *Il m'a raconté **son** histoire.*

Ces déterminants peuvent être évités comme dans les énoncés (31) et (32) ou encore conduire à « masculiniser » un nom féminin comme dans (33) où le mot « histoire » a perdu son « e » dans un essai de rationalisation.

¹⁹ À l'exception notable de l'adjectif épithète antéposé au nom.

En outre, les variantes des déterminants peuvent être à la source de découpages erronés, quand l'initiale d'un mot est identique à un déterminant ou à sa finale, par exemple :

- (34) *Je veux **m**abouret.*
- (35) *Où est **la** petite vabo ?*
- (36) *Prêtes-moi **ta** pareil-photo !*
- (37) *Il fait cet **l**exercice.*

À condition de se décentrer pour prendre le point de vue de l'apprenant, ce type d'erreurs, parfois fossilisées, permet de comprendre ce que signifie le brouillage de la langue par le jeu des allomorphismes avec l'homophonie et la nécessité de prendre en compte ces phénomènes très tôt dans l'apprentissage. Elles peuvent également faire l'objet par de futurs enseignants d'élaboration de corpus d'anticipation (cf. les propositions dans Camussi-Ni et Coatéval, 2013, p. 11-12).

2.2.3. L'amalgame

L'amalgame peut être défini comme la compression de deux morphèmes en un seul segment phonique. Les erreurs liées à l'absence d'amalgame sont, pour celles qui concernent les prépositions *à* et *de* et les articles définis *le* ou *les*, faciles à repérer. C'est une bonne occasion de se décentrer et de comprendre que l'existence de ces amalgames et leur périmètre ne peuvent pas être devinés si on n'est pas francophone. De fait, ces erreurs sont très fréquentes et susceptibles de fossilisation :

- (38) *J'ai demandé **à le** directeur pour changer de groupe.*
- (39) *Il y a un grand contrôle **de les** naissances.*

Inversement, certains apprenants étendent ces amalgames aux articles définis *la* et *l'* ou aux morphèmes complémentaires *le* ou *les* dans le schéma de l'unité verbale (41) ou encore ne repèrent pas l'amalgame (42) :

- (40) *Ils se battent parce que l'un des deux a volé une photo **du** autre.*
- (41) *Ils demandent **des** aider.*
- (42) *Le reporter a demandé même chose **aux** plusieurs personnes.*

La théorie de la médiation utilisée ici permet d'analyser aussi comme défaut d'amalgame certaines erreurs concernant les relatifs. En effet, elle pose que les relatifs sont l'amalgame de la conjonction *que* avec un préfixe de personne (*qui*), un morphème complémentaire 1 ou direct (*que*) ou un morphème complémentaire 3 ou indirect (*dont*) (Bonnet et Barreau, 1974, p. 246-247). Ainsi, pour l'énoncé (43) peut-on considérer que l'apprenant ne connaît pas l'amalgame de la conjonction *que* avec le préfixe de personne *ils* qui donne *qui* ; pour l'énoncé (44), l'amalgame aurait dû opérer entre la conjonction *que* et le morphème complémentaire *les* pour donner un homophone de la conjonction *que*, cette homophonie complexifiant l'analyse ; enfin, dans l'énoncé 45, c'est l'amalgame entre la conjonction *que* et le morphème complémentaire *en* qui n'a pas été réalisé.

- (43) *C'était son oncle et son beau-père **qu'ils** lui ont donné son premier et 2^{ème} travail.*

(44) *Les repas **que** je **les** mange ici ne sont pas très bons.*

(45) *C'est le vieux homme **que** je t'**en** ai parlé.*

Ces défauts d'amalgame sont très fréquents chez les apprenants allophones, mais ils apparaissent aussi chez les locuteurs francophones (Frei, 1929, Gadet, 1989).

2.2.4. La fragmentation de morphèmes

Les erreurs liées à la fragmentation de morphèmes correspondent au marquage de la personne, du genre ou du nombre à deux endroits d'une unité. Nous considérons qu'il s'agit d'un seul morphème marqué de façon discontinue. Ce double marquage pose des difficultés évidentes qui peuvent être liées à un sentiment de redondance chez des apprenants qui ne connaissent pas cette fragmentation dans leur langue maternelle. Ici aussi, ils ne peuvent pas deviner des marques qui apparaissent parfois uniquement à l'écrit mais au-delà ils peuvent avoir le sentiment d'une forme d'absurdité. Quelques exemples permettent d'illustrer la difficulté auprès de futurs enseignants, sans que cette notion demande beaucoup d'analyse :

(46) *On peut trouver toutes les choses que l'**étudiants** n'ont plus besoin.*

(47) *Il y a **des** image **qui** parle.*

(23) /ilaRiVƏRɔ̃ɑ̃RətɑR/ (**Il arriveront** en retard) (oral)

(48) **Nous** oublie partir.

Les exemples (46) et (47) présentent des cas où le préfixe et le suffixe du morphème de nombre sont discordants dans le cadre d'une unité nominale. Le préfixe de nombre, interne au déterminant, est singulier quand un suffixe du pluriel est postposé au lexème (46), ou inversement (47).

Cette non-reconnaissance d'un morphème fragmenté en deux éléments apparaît aussi dans des unités verbales. L'énoncé (23) concerne la discordance entre le suffixe de personne 6 audible au futur dans l'amalgame /ɔ̃/ et le préfixe de personne 3 indiqué par l'absence du phonème /z/ qui est nécessairement sonorisé dans le cadre de la liaison obligatoire. Dans l'énoncé (47), il s'agit de la même discordance pour la seconde erreur, **qui** amalgamant **elles** du fait de sa position syntaxique, le suffixe de personne devrait également être à la personne 6 **-ent**. Enfin, l'exemple (48) témoigne de la même discordance entre un préfixe de personne 4 (**nous**) qui devrait imposer son pendant, le suffixe de personne 4 (**-ons**).

2.2.5. L'absence significative

Le dernier obstacle au repérage des unités morphologiques concerne l'absence significative. Il s'agit des cas où une différence de sens est marquée par une opposition entre un segment phonologique et une absence, un silence.

L'explication précédente concernant l'énoncé (48) peut être reliée également à la notion d'absence significative. L'apprenant n'entendant pas les suffixes des personnes 1 à 3, peut ne pas avoir conscience de ces suffixes qui existent par leur opposition aux suffixes **-ons** et **-ez**. Dès lors, il n'a pas conscience de l'existence de la flexion verbale.

Il est possible aussi d'observer des erreurs dans le choix d'un suffixe de genre lié au fait que le suffixe masculin est marqué très souvent par une absence significative (49). Dès lors, un enseignement limité aux formes masculines de l'adjectif, comme l'ont repéré les acquisitionnistes (Véronique *et al.*, 2017, p. 101-102) pose d'évidentes difficultés.

(49) *Elles sont toutes remplites.*

Dans de nombreux cas comme dans (50), l'apprenant n'ayant eu affaire qu'à la forme masculine à l'oral (*rempli*) se retrouve en situation de deviner le suffixe oral qui permet de réaliser la forme orale et teste différentes possibilités /t/ dans notre exemple (49) ou /s/, /z/, /d/ ou encore un zéro, parfois moins évident pour qui veut marquer le genre.

La présence d'un déterminant devant le nom peut aussi être surgénéralisée, l'apprenant ne connaissant pas son absence significative dans le cas de la négation (50) et (51) ou après la préposition *de* (52) :

(50) *Cet homme n'a pas **des** parents.*

(51) *Le sol de la chambre sans **de** moquette est brillant.*

(52) *Cet bâtiment, on aperçoit beaucoup **des** fenêtres.*

Dans l'énoncé (50), nous analysons *des* comme un amalgame de l'article partitif *de*, indiquant une prise globale et de l'article défini *les* indiquant l'ensemble dans lequel se fait la prise. En contexte négatif, cet ensemble est vide, ce que dit le zéro attendu après le partitif *de*. L'énoncé (51) peut faire penser que l'apprenant utilise l'article partitif *de*, de façon automatique en contexte négatif. Or, *de* n'est pas une simple variante de *des* en contexte négatif. L'absence de déterminant après la préposition *sans* peut signifier l'ensemble vide. Le cas d'absence significative de déterminant est encore différent dans l'énoncé (52) où l'apprenant a voulu amalgamer la préposition *de* et le déterminant *des*. Or, ce déterminant s'efface après la préposition *de*, selon un phénomène appelé haplogogie par Wilmet (1997).

L'enseignant gagnera à connaître la notion d'absence significative, car il pourra facilement anticiper le risque de la négliger à l'aide de corpus où elle côtoie sa commutation.

Ce type d'analyse se poursuivra pour développer d'autres compétences, syntaxiques ou sémantiques des enseignants de FLES autour des obstacles que rencontrent les apprenants allophones. Nul doute également que ce type de démarche pourrait être amélioré en croisant les données de corpus numériques d'apprenants allophones et les séquences d'acquisition mises au jour par les acquisitionnistes.

2.3. Élaborer des corpus d'apprentissage anticipant les obstacles repérés

L'analyse outillée des erreurs d'énoncés d'allophones est un levier pour la formation des enseignants. L'acquisition de cette compétence permet au futur professeur d'apporter rapidement une réponse à une difficulté de grammaire repérée

chez un apprenant, une réponse qui lui permette de systématiser au lieu de corriger simplement. De plus, elle doit lui permettre d'adopter une position d'anticipation des erreurs, susceptible d'éviter leur fossilisation.

La formation de futurs enseignants de FLES peut faire déboucher l'exploration des phénomènes qui font obstacle à l'appréhension de la langue sur la constitution de corpus d'apprentissage. En effet, à l'analyse de ces énoncés erronés répond le questionnement : à quels énoncés l'apprenant aurait-il eu besoin d'être exposé pour éviter cette représentation faussée de la langue ? Ces énoncés peuvent donc être le déclencheur d'une activité en miroir, la fabrication de corpus d'apprentissage.

Pour Cuq et Gruca (2003) « le problème didactique est de recréer artificiellement mais efficacement des conditions d'appropriation linguistique, qui, hors système didactique ne sont pas ou très peu contraintes par l'espace et le temps : la classe est donc le lieu où se focalise la procédure de compression. » (Cuq & Gruca, 2003 : 121). Nous comprenons les corpus d'apprentissage proposés aux apprenants comme le résultat de cette compression. Il s'agit d'échantillons qui condensent les corpus en permettant d'en extraire rapidement les caractéristiques. Cette compression doit permettre d'accélérer l'apprentissage et de réduire les obstacles pour les apprenants qui arrêteraient précocement ou pour leur éviter la fossilisation, ce qui est un enjeu majeur selon nous.

Ainsi est-il possible, une fois le phénomène d'homophonie et ses différents impacts repérés de conduire les futurs enseignants à sélectionner une difficulté et à créer un mini-corpus d'apprentissage qui permette de dévoiler l'homophonie et d'anticiper les erreurs qui en découleraient. On peut en donner un exemple avec cet extrait de *Comprendre la grammaire* (Camussi-Ni et Coatéval, 2013, p. 21) :

- Elle fait un peu de baby-sitting avec **une** petite fille de 4 ans. C'est **la** petite fille de la voisine.

- Ah ! c'est **cette** petite fille-là ?

- Oui, elle **la** conduit à l'école et elle **me** conduit ensuite à la fac.

- Elle **te** conduit tous les jours ?

- Non, quand même pas !

Mise en évidence de l'homophonie du déterminant la et du morphème la

Il en est de même pour l'allomorphisme. Le futur enseignant peut choisir de cibler une difficulté et un niveau pour l'anticiper. L'extrait ci-dessous cible le niveau débutant A1, (Camussi-Ni et Coatéval, p. 26) :

Je comprends **le** français et **j'**apprends **l'**anglais.

Est-ce que **vous** parlez français ? Est-ce que **vous** écrivez en français ?

Est-ce qu'**on** va en cours ? oui, **on** y va.

Initiation à la notion de variantes liées à l'environnement phonétique

Dans ces exemples, les éléments surlignés demandent à être comparés pour mettre au jour homophonie ou allomorphisme. C'est leur objectif minimal qui permet de déconstruire la difficulté et qui sera à faire comprendre aux futurs enseignants.

Cette déconstruction permet, en outre, un grand nombre de sujets de création de corpus pour répondre aux erreurs au plus près des besoins des apprenants.

Conclusion

Au terme de ce parcours, il est apparu que les énoncés erronés d'apprenants pouvaient permettre de construire une compétence linguistique des enseignants, en même temps que la compétence didactique qui permet de les déjouer ou de les anticiper. Nous voudrions insister sur deux points : nous nous proposons d'appuyer les analyses d'erreurs sur une analyse proprement linguistique et non sur des procédés généralement évoqués comme la simplification ou la surgénéralisation ; nous accordons une grande attention à la formation des enseignants à une époque où l'accent sur les outils numériques pourraient amener à les négliger. En effet, l'usage des corpus pourrait rester affaire de spécialistes ou être destiné uniquement aux apprenants avancés. Or, nous l'avons repéré, c'est dès leur entrée dans la langue que l'enseignant peut proposer aux apprenants de déjouer des obstacles, ce qui est un enjeu majeur pour les apprenants qui n'atteindront pas le niveau avancé. Encore faut-il qu'il soit formé pour cela ! Astolfi s'interroge :

Il est permis de se demander si les enseignants n'ont pas comme une intuition de l'ampleur des difficultés qu'ils rencontreraient s'ils prétendaient s'attaquer sérieusement aux problèmes conceptuels qui se posent aux élèves et si, inconsciemment, ils ne préfèrent pas la prudence. Peut-être pressentent-ils qu'il leur faudrait disposer d'outils intellectuels qu'ils n'ont pas, comme par exemple celui d'objectif-obstacle. (Astolfi, 1992, p. 107)

Nous pensons qu'offrir à un enseignant la capacité d'analyser des énoncés erronés permet de disposer de ces outils intellectuels mais aussi conduit à une nouvelle posture : il ne s'agit plus de corriger une erreur mais de la comprendre en profondeur pour donner les moyens à l'apprenant de réviser la représentation qui est à la base de toutes sortes d'énoncés erronés équivalents. Il s'agit donc de rechercher ce qui fait l'intelligence de l'erreur et non de la fustiger. La valorisation de l'erreur, par son ouverture à la bienveillance, à la reconnaissance de l'autre, redouble son efficacité didactique. En somme, il s'agit pour l'enseignant d'assumer pleinement et dans tous les sens du terme, l'erreur de l'apprenant.

Bibliographie

- André, V. (2020). Faire de la linguistique de corpus avec des apprenants de français langue étrangère, in P. Larrivée & F. Lefeuvre. *La didactisation du français vernaculaire*. Caen : Presses Universitaires de Caen. 37-66.
- Astolfi, J.P. (1992). Apprendre par franchissement d'obstacles ? *Repères* 5, 103-116.
- Astolfi, J.-P. (1997). *L'erreur, un outil pour enseigner*, Paris, ESF éditeur.

- Auger, N. (2005). *Comparons nos langues. Démarches d'apprentissage du français auprès d'enfants nouvellement arrivés*. Montpellier : CRDP Languedoc-Roussillon (DVD 26 mn).
- Auger, N. et Le Pichon-Vorstman, E. (2021). *Défis et richesses des classes multilingues. Construire des ponts entre les cultures*. Paris : E.S.F. Sciences humaines.
- Bartning, I. et Schlyter, S. (2004). Itinéraires acquisitionnels et stades de développement en français L2. *Journal of French Language Studies* 14/1, 281-299.
- Benazzo, S. et Watorek, M. (2021). Transcription de corpus oraux d'apprenants débutants en français L2 : quelques enjeux théoriques. In Lorenzo Spreafico, Giuliano Bernini, Ada Valentini & Jacopo Saturno (éds.). *Superare l'evanescenza del parlato. Un vademecum per il trattamento digitale di dati linguistici*. Bergamo: Sestante. 127-165.
- Besse, H. et Porquier, R. (1984). *Grammaire et Didactique des langues*. Paris: Hatier.
- Bonnet, J. et Barreau, J. (1974). *L'esprit des mots. Traité de linguistique française*. L'école.
- Boulton, A. et Tyne, H. (2014). *Des documents authentiques aux corpus : démarches pour l'apprentissage des langues*. Paris : Didier.
- Boulton, A., Canut, E., Guerin, E., Parisse, C. et Tyne, H. (2013). Corpus et appropriation de L1 et L2. *Linx* 68-69. 9-32.
- Cadet, L. et Guérin, E. (2012). FLM, FLS, FLE. Au-delà des catégories. *Le français aujourd'hui* 176. 3-8.
- Camussi-Ni, M.-A. et Coatéval, A. (2013). *Comprendre la grammaire. Une grammaire à l'épreuve de la didactique du FLE*. PUG.
- Camussi-Ni, M.-A. (2020). Le Mot. Critique théorique et implications didactiques. *Tétralogiques* 25. 105-136.
- Chuang et Nesi (2006). Chuang, F.-Y. et Nesi, H. (2006). An analysis of formal errors in a corpus of L2 English produced by Chinese students. *Corpora*, vol. 1, n° 2. 251-271.
- Corder, S.P. (1980a). Que signifient les erreurs des apprenants ? *Langages* 57. 9-15.
- Corder, S.P. (1980b). La sollicitation de données d'interlangue. *Langages* 57. 29-38.
- Creissels, D. (1995). *Éléments de syntaxe générale*, PUF.
- Cuq, J.-P. (dir.) (2003). *Dictionnaire de didactique du français langue étrangère et seconde*. Clé International.
- Cuq, J.-P. et Gruca, I. (2003). *Cours de didactique du français langue étrangère et seconde*. PUG. collection FLE.
- Debrenne, M. (2011). *Les erreurs de français russophones*. Lambert-Lucas.
- De Cock, S. et Tyne, H. (2014). Corpus d'apprenants et acquisition des langues. *Recherches en didactique des langues et des cultures* 11-1. En ligne : 10.4000/rdlc.1716. DOI : 10.4000/rdlc.1716.

- Delahaie, J. (2008). *Français parlés et français enseignés. Analyses linguistiques et didactiques de discours de natifs, de non-natifs et d'enseignants*. Thèse de doctorat. Université de Nanterre - Paris X.
- De Salins, G. (1996). *Grammaire pour l'enseignement/apprentissage du FLE*. Coll. Didactique du français. Didier/Hatier.
- Devaux, A. (2005). L'asymptote du français avancé : les difficultés résistantes. *ELA* 138. 223-241.
- Dubois, M., Kamper, A. et Skupien Dekens, C. (2014). Tenir compte des langues premières dans l'enseignement/apprentissage des langues vivantes. *Mélanges Crapel* 35. 5-11.
- Frei, H. (1929). *La grammaire des fautes*. Bellegarde (Ain) : Société anonyme des arts graphiques de France.
- Gabrielatos, C. (2005). Corpora and language teaching: just a fling or wedding bells?. *TESL-EJ*, vol. 8, n° 4.
- Gadet Françoise. La relative non standard saisie par les grammaires. *Linx* 20, 1989. Analyse grammaticale des corpus oraux. 37-49.
- Galligani, S. (2003). Réflexion autour du concept d'interlangue pour décrire des variétés non natives avancées en français. *Linx*, 49. 141-152.
- Galligani, S. (2022). L'interlangue et son rôle dans la construction de la compétence plurilingue d'élèves allophones en contexte scolaire. *Le français aujourd'hui*, 217. 47-55.
- Germain, C. (1993). *Évolution de l'enseignement des langues : 5 000 ans d'histoire*. Montréal : Hurtubise HMH.
- Granger, S. (2004). Computer learner corpus research: current status and future prospects. In Connor, U. & Upton, T. (dir.). *Applied corpus linguistics: a multidimensional perspective*. Amsterdam/Atlanta: Rodopi. 123-145.
- Jongen, R. (1993). *Quand dire c'est dire. Initiation à une linguistique glossologique et à l'anthropologie clinique*. De Boeck Université. Collection Raisonances.
- Johns, T. (1986). Micro-concord: a language learner's research tool. *System* 14(2). 151-162.
- Marquilló Larruy, M. (2003). *L'interprétation de l'erreur*, CLE International.
- Millar, N. & Lehtinen, B. (2008). DIY local learner corpora: bridging gaps between theory and practice. *JALT CALL Journal*, vol. 4, n° 2. 61-72.
- Noyau C. et Véronique D. (1986). Survey Article. *Studies in Second Language Acquisition* 8(3). 245-263.
- Paquot, M. (2012). The LEAD dictionary-cum-writing aid: an integrated dictionary and corpus tool. In Granger, S. & Paquot, M. (dir.). *Electronic lexicography*. Oxford: Oxford University Press. 163-187.

- Perdue, C. (1980). L'analyse des erreurs : un bilan pratique. *Langage* 57. 87- 94.
- Pergnier, M. (1986). *Le Mot*, PUF. Collection Linguistique nouvelle.
- Porquier, R. (1977). L'analyse des erreurs : problèmes et perspectives. *Études de Linguistique Appliquée* 25. 23-43.
- Py, B. (1980). Quelques réflexions sur la notion d'interlangue. *Travaux Neuchâtelois de Linguistique* 1. 31-55.
- Rattunde, E. (1980). Stabilité vs. instabilité des connaissances intermédiaires. *Encrages*, 46-53.
- Rundell, M. (dir.). (2007). *Macmillan English dictionary for advanced learners*, second edition. Oxford: Macmillan Education.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics* X(3). 209-231.
- Selinker, L., Swain, M. & Dumas, G. (1975). The interlanguage hypothesis extended to children. *Language Learning* 25(1). 139-152.
- Suchet, M. (2014). *L'Imaginaire hétérologue. Ce que nous apprennent les textes à la croisée des langues*. Classiques Garnier.
- Urien, J.-Y., 1999, « Le critère du grammatical ». In Giot J. et Schotte J. (dir.). *Langage, clinique, Épistémologie. Achever le programme saussurien*. De Boeck-Supérieur. 29-71.
- Vergne, M. (2020). Analyser et classer les fausses analogies entre l'italien et le français : les (re)connaître, les enseigner. *Dorif.it*.
<https://doi.org/10.1051/shsconf/202419107002>
- Véronique, D. (2005). Les interrelations entre la recherche sur l'acquisition du français langue étrangère et la didactique du français langue étrangère. *Acquisition et interaction en langue étrangère* 23. 9-41.
- Véronique, D., Carlo, C., Granget, C., Kim, J.-O. et Prodeau, M. (2017). *L'acquisition de la grammaire du français langue étrangère*. Didier.
- Véronique, G.-D. (2017). La grammaire en français langue étrangère : questions d'acquisition et d'intervention. *Lidil* 56. <https://doi.org/10.4000/lidil.4734>
- Vogel, K. (1995). *L'interlangue : la langue de l'apprenant*. Toulouse : Presses universitaires du Mirail.
- Weber, C. (2021). Chapitre 19. Acquisition d'une langue étrangère ou comment l'adulte apprend une langue étrangère ? In Chiss J.-L. (éd.). *Le FLE et la francophonie dans le monde*. Paris : Armand Colin, 219-226.
- Wilmet, M. (1997). *Grammaire critique du français*. Hachette Supérieur. Duculot.