



Sociolinguistique des interactions verbales et exploitations didactiques

Virginie André

► To cite this version:

Virginie André. Sociolinguistique des interactions verbales et exploitations didactiques. Linguistique. Université Lumière Lyon 2, 2019. tel-03231366

HAL Id: tel-03231366

<https://theses.hal.science/tel-03231366>

Submitted on 20 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Lumière Lyon 2 – Laboratoire ICAR (UMR 5191)

Dossier présenté pour
l'Habilitation à Diriger des Recherches

Sociolinguistique des interactions verbales et exploitations didactiques

Virginie André

Université de Lorraine
ATILF (UMR 7118 - CNRS)
(Équipe Didactique des langues et sociolinguistique)

22 novembre 2019

Conseillère scientifique

Véronique Traverso, Directrice de Recherche au CNRS

Jury

Hervé Adami, Professeur à l'Université de Lorraine

Alex Boulton, Professeur à l'Université de Lorraine

Isabelle Racine, Professeure à l'Université de Genève (Suisse)

Élisabeth Richard, Professeure à l'Université de Rennes 2

Anita Thomas, Professeure à l'Université de Fribourg (Suisse)

Volume 3

Note de synthèse

Remerciements

La recherche ne se fait jamais seule. Ce présent travail doit à beaucoup de monde.

Je remercie Véronique Traverso pour avoir accepté d'être ma conseillère scientifique dans cette aventure ainsi que pour ses précieux conseils. Je la remercie pour avoir posé les bonnes questions, celles qui m'ont permis d'avancer scientifiquement et d'en rendre compte dans ce travail de synthèse.

Je remercie Hervé Adami, Alex Boulton, Isabelle Racine, Élisabeth Richard et Anita Thomas pour avoir accepté de participer à ce jury et pour m'avoir chacun·e communiqué son enthousiasme et ses encouragements. Je les remercie aussi pour nos riches collaborations passées et à venir.

Je remercie tous les collègues auprès desquels j'ai grandi et je me suis nourrie tout au long de ces dernières années : mes collègues nancéens – de l'ATILF, du « CRAPEL » (les anciens et les nouveaux ou les crapéliens et crapUL), du DÉFLE et du groupe LTF – et d'ailleurs.

Je remercie tous les étudiants qui ont fait le choix de me suivre dans mes travaux, dans certaines de mes expérimentations et qui se sont lancés dans les leurs...

Enfin, je remercie tous mes proches qui se reconnaîtront parce qu'ils m'ont accompagnée, encouragée et aidée... à commencer et à finir ce manuscrit... avec une mention spéciale pour un soutien tout aussi... spécial...

Sommaire

Introduction	5
Parcours de recherche	5
Présentation du dossier	14
Chapitre 1. Un cadre pour une sociolinguistique des interactions verbales.....	17
1.1. Une schématisation de l'analyse sociolinguistique des interactions verbales	18
1.1.1. Démonstration et illustration du schéma-cadre	26
1.1.2. Exploitation du cadre dans d'autres études.....	31
1.2. Pourquoi une approche sociolinguistique des interactions verbales ?	38
1.3. Les apports de l'analyse des situations de travail.....	43
Conclusion du chapitre.....	48
Chapitre 2 : Des données à analyser, des corpus à constituer.....	49
2.1. Un corpus, des corpus, quels corpus ?	51
2.2. TCOF – des interactions verbales authentiques.....	53
2.3. FLEURON – un corpus du domaine universitaire	56
2.4. ILO – des locuteurs en insécurité langagière	60
2.5. Des corpus pour quoi faire ?	64
Conclusion du chapitre.....	70
Chapitre 3. Des analyses sociolinguistiques d'interactions verbales – vers une interface sociolinguistique et didactique	73
3.1. Des pratiques interactionnelles sous influence	73
3.1.1. Des énonciations conjointes	75
3.1.2. Des questions	82
3.2. Identification des compétences socio-interactionnelles	90
3.2.1. Des outils d'évaluation de compétences socio-interactionnelles	91
3.2.2. Écrire, un prétexte pour acquérir des compétences socio-interactionnelles....	94
3.2.3. Interactions verbales avec des locuteurs en insécurité langagière	97
3.3. Interface sociolinguistique et didactique	103
Conclusion du chapitre.....	109
Chapitre 4. L'analyse sociolinguistique des interactions verbales au service de l'enseignement/apprentissage de la langue	111
4.1. Exposition des apprenants à des interactions verbales authentiques	113
4.1.1. PFC-EF, FLORALE et CLAPI-FLE	117
♦ PFC-EF	118
♦ FLORALE	119
♦ CLAPI-FLE	122

4.1.2. FLEURON	124
4.2. Écart entre <i>langue prescrite</i> et <i>langue réelle</i> . Des pratiques interactionnelles non décrites pour le FLE	127
4.2.1. <i>le truc c'est que</i> dans CLAPI-FLE	130
4.2.2. <i>genre</i> et sa fiche descriptive	132
4.2.3. La duplication du sujet avec un pronom disloqué - <i>moi je</i>	138
4.3. Apprendre à interagir avec les ressources de FLEURON.....	142
4.3.1. Avec des étudiants étrangers.....	143
4.3.2. Avec des adultes migrants peu scolarisés.....	150
Conclusion du chapitre.....	155
Chapitre 5. L'apprentissage sur corpus : un concordancier au service de l'enseignement et de l'apprentissage de la langue	157
5.1. <i>Data-driven learning</i> (DDL) ou apprentissage sur corpus (ASC)	158
5.2. De l'analyse de lignes de concordances à l'analyse sociolinguistique des données en interaction	166
5.3. L'apprentissage sur corpus avec FLEURON	171
5.3.1. Le concordancier de FLEURON.....	171
5.3.2. Les étudiants-apprenants-utilisateurs de FLEURON : des Sherlock Holmes ...	175
5.3.2.1. Les consignes	176
5.3.2.3. Le processus d'analyse : l'apprenant devient un détective.....	179
5.3.4. Bilan des recherches d'occurrences.....	194
5.4. Des usages détournés de FLEURON	202
5.4.1. Avec des adultes migrants peu scolarisés.....	202
5.4.2. Avec des adultes francophones natifs en insécurité langagière.....	214
5.4.2.1. Une transcription pour tester le degré d'illettrisme	215
5.4.2.2. Une séquence pédagogique à partir de corpus.....	217
5.4.2.3. Le choix de FLEURON	220
Conclusion du chapitre.....	228
Conclusion.....	234
Bibliographie.....	242
Table des illustrations (figures et tableaux)	262

Introduction

Parcours de recherche

Ce travail de synthèse repose sur un parcours de recherche fait d'événements de différentes natures : des rencontres scientifiques et humaines, des croisements de regards, des confrontations conceptuelles et des expériences de terrain. Toutes ces occasions m'ont nourrie, sont entrées en résonance avec mes préoccupations scientifiques et sociales, m'ont fait réfléchir et avancer. Elles ont finalement, directement ou indirectement, participé au présent travail. Pour saisir le sens de celui-ci, je vais tenter de résumer mon cheminement scientifique. Ce résumé va inévitablement être lacunaire, tributaire des manquements de ma mémoire et, parfois, injustement sélectif. Il ne sera pas toujours chronologique, parce que certaines rencontres et certaines recherches se font en parallèle et parce que leurs enchevêtrements se prêtent difficilement à la linéarité qu'impose l'écriture. Il fera donc des allers-retours sur l'axe temporel de ma vie de chercheuse.

Tout d'abord, je me suis inscrite en Sciences du langage en 1997, à Nancy, à la suite d'une discussion avec une amie qui avait une amie qui avait suivi les trois premières années de cette formation universitaire. De cette courte discussion, je n'ai retenu que l'expression : « étude scientifique du langage ». Elle a immédiatement réveillé en moi une curiosité qui s'était animée pour la première fois dans mes cours de philosophie, consacrés au langage. Pour être honnête, c'est donc avec très peu d'informations sur la discipline que j'ai commencé à m'y intéresser... Mes premiers cours à l'université ont été déterminants. Ils ont confirmé mon choix et mon intérêt pour l'étude de la langue. Les enseignements passionnants de sociolinguistique et d'ethnolinguistique, dispensés par Philip Riley, ont éclairé les premiers pas de mon chemin scientifique. Je me souviens avoir dit à mes camarades de promotion que c'était « ça » que je voulais faire, sans trop savoir à l'époque ce que cela pouvait signifier. Ce n'est donc pas sans émotion et fierté que presque dix années plus tard, je dispensais moi-même ces enseignements après le départ en retraite de Philip Riley, d'abord en tant qu'ATER en 2006 puis en tant que maître de conférences en 2007.

Après ma Licence de Sciences du langage pendant laquelle j'ai pu aborder, avec exaltation, tous les aspects de la langue, ma maîtrise et mon DEA ont été consacrés à l'analyse des interactions en situation de travail. Les prémices d'une « analyse

sociolinguistique des interactions verbales »... Lors de ces deux années, j'ai découvert ce que l'on appelle « le terrain », ainsi que les difficultés pour y entrer, s'y faire accepter, s'y familiariser. J'ai découvert aussi les joies et les peines du recueil de données et de la transcription... à cette époque, avec des cassettes pour enregistrer et un traitement de texte pour transcrire... Les premières organisations dans lesquelles j'ai enquêté ne m'ont pas toujours accueillies avec affabilité mais certaines ont été extrêmement bienveillantes, sans réellement comprendre les raisons de ma présence... Mes premières analyses des interactions verbales en réunion de travail ont confirmé les voies scientifiques dans lesquelles j'avais choisi de m'engager et surtout mon appétence pour ce type de travail.

A la suite de la soutenance de mon mémoire DEA sur l'étude de l'influence du statut du locuteur dans les interactions verbales en entreprise, je souhaite poursuivre mes recherches. Philip Riley me propose de m'engager dans un doctorat sous sa direction. Très honorée de cette proposition, je m'inscris en septembre 2002 avec une bourse du Ministère de l'Enseignement Supérieur et de la Recherche. Ma thèse sera consacrée à l'analyse de la construction collaborative du discours au sein de réunions de travail en entreprise (André 2006a). Je me suis immergée dans une entreprise de droit public, j'ai recueilli quarante heures d'enregistrement, transcrit 331 000 mots et commencé à élaborer un cadre théorique et méthodologique, affiné d'année en année et dont l'aboutissement est présenté dans ce travail. Depuis 2002, je participe également à la vie de mon équipe, mes recherches se font en parallèle et parfois de façon conjointe aux siennes.

L'été qui précède mon inscription en doctorat va être mon premier bain dans un bouillonnement scientifique et mes premiers pas au Centre de Recherches et d'Applications Pédagogiques en Langue (CRAPEL), devenu aujourd'hui l'équipe « Didactique des langues et sociolinguistique » de l'ATILF. J'ai été recrutée en tant que membre du comité d'organisation d'une école d'été sur l'enseignement des langues vivantes dans le premier degré. Ce travail de petites mains m'a permis d'observer en coulisse la recherche en train de se faire. Bien que le domaine ne m'ait pas été familier, ce n'est pas sans plaisir que je me suis investie dans la tâche que l'on m'avait confiée. J'en ai profité pour côtoyer de plus près certains de mes enseignants qui avaient participé au début de mon parcours universitaire et qui deviendront mes collègues. Ces derniers m'ont en réalité un peu testée lors de cette école d'été, je m'en suis rendue compte bien des années plus tard.

Ce test a dû être concluant parce qu'en septembre 2002 je fus « cooptée » au CRAPEL. C'est ici que commence l'aventure. Dans cette équipe, la tradition était que tous

les nouveaux arrivants assistent aux réunions de tous les groupes de recherche. A mon arrivée, ils y en avaient trois, très actifs, qui s'intéressaient : à l'apprentissage autodirigé et intégré de deux langues, aux nouvelles technologies et au multimédia, à la sociolinguistique. D'autres groupes se formaient temporairement au gré des questions scientifiques des membres de l'équipe et des projets en cours (pour une présentation des thématiques de recherche du CRAPEL, voir Holec 2000). Ma participation au sein de tous les groupes fut particulièrement enrichissante mais j'ai dû choisir (ayant également ma thèse en cours) et j'ai rejoint le Groupe de réflexion sociolinguistique (GREFSOC) dirigé par Philip Riley. De sociolinguistique et d'interactions verbales, il était question, mais également de didactique des langues. N'ayant pas eu de formation universitaire dans ce domaine, j'ai en quelque sorte « appris sur le terrain » mais accompagnée et guidée par des chercheurs expérimentés, brillants et généreux. Ainsi, parallèlement à mon travail de thèse, j'ai fait mes premiers pas dans la recherche en analysant notamment des relations de service et leurs aspects sociaux et culturels. Ces analyses ont mené le groupe à définir une notion qui sera reprise en didactique, celle d'« étranger compétent » (Adami, André *et al.* 2003 ; André, Castillo 2005). Ces travaux sur les analyses des relations de service et leurs exploitations didactiques ont été discutés lors d'un colloque organisé à Nancy en 2005 et ont fait l'objet d'un numéro spécial de la revue *Mélanges* (André 2008). Avec ces travaux et les membres du GREFSOC, j'entrevois les premiers liens entre sociolinguistique et didactique des langues.

Mon intégration au sein du CRAPEL me plonge dans les recherches pionnières et innovantes en didactique des langues notamment au sujet de l'utilisation des documents authentiques comme supports pédagogiques (Holec 1990). Cette rencontre, entre l'analyse des interactions verbales et l'exposition des apprenants à une langue *authentique*, paraît évidente et naturelle. D'ailleurs, elle s'est faite spontanément et simplement car la conception de la langue était commune. Le cadre scientifique que j'étais petit à petit en train de me construire s'est trouvé en parfaite résonnance avec le cadre didactique de l'équipe. J'ai assisté, d'abord pour apprendre, aux réflexions liées à la préparation de formations de formateurs et d'enseignants de langue étrangère, notamment de français langue étrangère (désormais FLE), qui venaient du monde entier pour se former auprès des chercheurs du CRAPEL. J'ai ensuite assisté aux formations puis dispensé ces formations avec les membres de l'équipe. En allant former des enseignants de FLE dans leur pays, notamment en Tunisie avec Florence Poncet, lors de la mise en place de la réforme de l'enseignement du français confiée au CRAPEL, j'ai appris la socio-didactique. En d'autres

termes, les formations à l'enseignement de la langue doivent être contextualisées et prendre en compte les situations dans lesquelles elles s'inscrivent. Voici encore des liens évidents entre sociolinguistique et didactique...

Ces connexions ont été mises à l'épreuve de la formation linguistique des plus faiblement dotés en capital symbolique (Bourdieu 1994), et notamment des migrants, lors de la création du groupe « Langage, Travail et Formation » (désormais LTF) au sein de l'équipe Didactique des langues et sociolinguistique de l'ATILF. Depuis 2006, avec Hervé Adami, nous animons ce groupe de recherche. Il est né d'une demande sociale de la branche professionnelle des métiers de la propreté, relayée par Hervé Adami qui avait des liens étroits avec un organisme de formation pour adultes migrants. Cette demande consistait à créer des outils d'évaluation des compétences langagières des salariés de la propreté. Elle a permis de réunir Hervé Adami, sociolinguiste et spécialiste de la formation linguistique des migrants, Philip Riley, socio-ethno-linguistique et didacticien, Florence Poncet, didacticienne et enseignante de FLE, et moi-même, sociolinguiste des interactions (et déjà intéressée par ses exploitations didactiques). Plusieurs membres du CRAPEL sont également intervenus ponctuellement dans nos réflexions. Depuis 2006, nous avons mené de nombreux projets articulant les trois item *langage*, *travail* et *formation*. Nous avons créé un réseau international¹ du même nom, dans la lignée du réseau *Langage et Travail* (Borzeix, Fraenkel 2001), et une revue en ligne (diffusée sur le site du réseau). Le premier numéro de la revue paru en 2015 présente les objectifs du réseau (Adami, André, Ciekanski 2015). Toutes les activités de recherche menées au sein du groupe LTF s'inscrivent dans différents cadres disciplinaires, tels que la sociolinguistique, la sociologie, l'ergonomie, et réinterrogent les concepts de la didactique. Dans ce cadre, plusieurs actions de valorisation de la recherche ont été réalisées par le groupe. Elles permettent de mettre en application les cadres scientifiques que nous abordons. Les réalités sociales que nous analysons et que parfois nous transformons ou proposons de transformer bénéficient légitimement des résultats de la recherche publique.

Dans cette perspective, ma collaboration avec Hervé Adami est riche, fructueuse et, parfois, à contre-courant. Nous nous aventurons sur des terrains délaissés de la didactique, ceux de la formation linguistique des faiblement qualifiés et peu scolarisés, natifs et migrants (Adami, André 2012a ; 2012b). Nous démontrons les enjeux du français langue d'intégration (désormais FLI) avec la nécessaire triple intégration – sociale, économique et

¹ Voir son site : <https://reseaultf.atilf.fr/>

citoyenne – des migrants et nous introduisons les corpus dans les formations FLI (Adami, André 2013). Nous proposons ainsi une méthodologie adaptée au public migrant, accueilli en France, et à ses enjeux d'émancipation. Nous avons abouti à la définition de la notion d'« insécurité langagière » (Adami, André 2014). Nous considérons que des locuteurs ou scripteurs sont en insécurité langagière lorsqu'ils ne réussissent pas à interagir (à l'écrit et/ou à l'oral) de façon efficace et satisfaisante. Nous avons dénoncé la doxa du plurilinguisme et coordonné un ouvrage qui révèle cette im-posture scientifique (Adami, André 2015). Nous réfléchissons à des innovations didactiques pour des natifs en insécurité langagière et coordonnons, avec Valérie Langbach qui a rejoint LTF en 2014, un ouvrage sur les enjeux sociaux et didactiques des insécurités langagières (Adami, André, Langbach à paraître). Je ne peux pas mentionner tous les projets du groupe LTF, tant ils sont nombreux.

Parallèlement aux travaux du groupe LTF, je mène plusieurs activités scientifiques qui vont également enrichir mon parcours. Je ne peux pas toutes les citer, certaines seront développées dans le corps de ce travail. Je vais donc les sélectionner dans cette présentation de mon cheminement. Cela ne signifie pas que celles qui ne sont pas décrites n'ont pas eu d'importance. Je pense que chaque pas, chaque rencontre, chaque collaboration, a alimenté, consciemment ou non, ma réflexion scientifique. Par exemple, je coordonne la partie « interactions entre adultes » du corpus TCOF depuis 2007, d'abord avec ma collègue Jeanne-Marie Debaisieux (une rencontre marquante...), jusqu'à son départ de Nancy en 2011. Ce corpus nancéen a pour ambition de collecter un grand échantillon de données orales en français parlé et d'être diffusé librement et gratuitement. Les choix effectués pour cette diffusion ont été longuement discutés (André, Canut 2010) et ont occasionnés des rencontres avec différents spécialistes en informatique, traitement du signal, traitement automatique des langues, linguistique outillée... Mes premières collaborations avec le laboratoire lorrain de recherches en informatiques et ses applications (LORIA) voient le jour avec les réflexions autour de TCOF.

En 2006, une fois ma soutenance de thèse terminée, Catherine Kerbrat-Orecchioni m'a proposé de venir travailler avec elle à Lyon. Sans aucune hésitation et avec une grande exaltation, j'ai rejoint son groupe de travail sur les formes nominales d'adresse. Cette première collaboration lyonnaise fut à la fois riche, intense et généreuse. Les longues séances de travail d'analyse des données de chacun des membres du groupe étaient l'occasion d'échanges scientifiques passionnants. Les ouvrages sur les formes nominales d'adresse, en français et dans d'autres langues, ont mis au jour des éléments interactionnels

peu étudiés alors que leurs rôles, leurs valeurs et les effets pragmatiques sont non négligeables (André 2010a ; Kerbrat-Orecchioni 2010, 2014).

En 2011, dans le cadre de l'ANR Corpus intitulée ORFÉO (outils et recherches sur le français écrit et oral) coordonnée par Jeanne-Marie Debaisieux, mes collaborations avec l'équipe LIS (Langues, Interactions, Situations) du laboratoire ICAR (UMR 5191) de Lyon se sont construites autour des études pilotes du projet. Nous nous sommes plus précisément intéressés à l'analyse interactionnelle des actualisations de *oui* et de ses variantes (*oui oui*, *ah oui*, etc.) et de l'analyse des émotions (Quignard *et al.* 2016). Ces travaux ont été menés à partir des corpus d'interactions lyonnais et nancéen, CLAPI et TCOF. Depuis, mes collaborations avec l'équipe LIS, et notamment avec Carole Etienne, ne se sont jamais interrompues. Toutes les deux, et avec les collègues qui nous entourent, nous nous intéressons à l'utilisation et à la didactisation des corpus oraux et multimodaux. Nos travaux sont complémentaires et ont abouti à la création de ressources interactionnelles pour la didactique : CLAPI-FLE à Lyon et FLEURON à Nancy.

Lorsque je suis arrivée au CRAPEL en 2002, je ne m'en suis pas rendue compte immédiatement mais il planait une idée qui semblait enthousiasmer les membres de l'équipe. Cette idée paraissait toutefois avoir des difficultés à se concrétiser... Elle était évoquée dans tous les groupes de recherche (notamment dans NTM : nouvelles technologies et multimédia) mais paraissait insaisissable. Cette idée, c'est celle de FLEURON qui se décline : français langue étrangère universitaire : ressources et outils numériques. En 2009, sous l'impulsion de Jeanne-Marie Debaisieux, avec Henry Tyne, nous reprenons l'idée et commençons à y réfléchir. Nous envisageons FLEURON comme un dispositif d'apprentissage du FLE destiné à des étudiants étrangers qui souhaitent se préparer à leur séjour universitaire en France. Le cœur de FLEURON sera un corpus d'interactions, notamment dans les services administratifs de l'université, et un concordancier qui permettra d'interroger le corpus. En 2011, avant son départ pour Paris, Jeanne-Marie Debaisieux m'a confié le projet. Henry Tyne étant parti à Perpignan, je me suis donc retrouvée responsable, je dois bien l'avouer, un peu malgré moi, du projet. A cette époque, à part quelques enregistrements de situations de communication authentiques, peu de choses sont abouties. Dans un premier temps, je me suis demandée si j'allais poursuivre les travaux ou si je n'allais pas tout simplement abandonner l'idée... Et puis, dans un deuxième temps, quelques idées – déjà de liens entre corpus d'interactions et didactique – m'ont poussée à persévérer. J'ai collaboré avec Yann Boniface (LORIA) pour développer, avec ses

étudiants en informatique, une première mouture d'un site Internet. D'année en année, au fur et à mesure de projets d'étudiants et de stages, un site a vu le jour (qui a été stabilisé en 2016, puis amélioré en 2019). Je me suis appropriée le projet, j'ai pris du temps pour enrichir le corpus ainsi que les fonctionnalités du site. Surtout, j'ai attribué au projet les objectifs scientifiques qui me correspondaient et me plaisaient. Et enfin, dans un troisième temps, le projet était de plus en plus enthousiasmant, j'ai accéléré les recherches et je me suis lancée dans les premières expérimentations du dispositif avec Florence Poncet, directrice du département de FLE à Nancy et membre de LTF, à mes côtés. Ce résumé de la naissance de FLEURON est extrêmement bref (pour en savoir plus, voir [André 2016](#)). Il ne rend pas hommage aux collègues et étudiants qui m'ont encouragée dans cette aventure. Il ne mentionne pas non plus, par exemple, les innombrables recherches de financement qui illustreraient le fonctionnement de la recherche publique.

Le site FLEURON² associe désormais un corpus d'interactions verbales authentiques et un outil d'interrogation des données (un concordancier). Il met à disposition librement et gratuitement ce corpus multimodal intégré à un dispositif d'apprentissage. D'un point de vue scientifique, le projet FLEURON puise ses origines dans les thématiques du CRAPEL : séparation des aptitudes, documents authentiques, apprentissage en autonomie. De plus, le dispositif FLEURON est conçu selon l'hypothèse didactique que le contexte d'interaction proposé par le recours au corpus soutiendrait le développement de compétences socio-interactionnelles. C'est en s'exposant à des situations de communication que l'on apprend à interagir. FLEURON permet une exposition à la langue de deux façons différentes : la première en ayant recours aux ressources multimodales (audios ou vidéos), une par une, et la seconde en explorant le corpus entier *via* le concordancier. Ces deux possibilités d'exploiter les données seront exposées respectivement dans les chapitres [4](#) et [5](#) de ce travail. Elles mettent au jour les liens entre sociolinguistique des interactions verbales et didactique de l'oral en FLE.

Enfin, parmi les rencontres étonnantes, celle avec le groupe Psyphine en 2015 le fut particulièrement. Ce groupe pluridisciplinaire s'intéresse aux interactions entre les hommes et les robots. Les regards croisés de l'informatique, des neurosciences, de l'intelligence artificielle, de la philosophie, de la psychologie, de l'anthropologie, de la robotique et de la sociolinguistique permettent de répondre à la question de l'attribution d'intentionnalité, d'intelligence, de cognition voire d'émotions, à des entités naturelles ou des dispositifs

² <https://fleuron.atilf.fr/>

artificiels. Il s'agit de comprendre et d'analyser les interactions qui prennent place entre un humain et ces dispositifs. Ce groupe est particulièrement actif, nous organisons régulièrement des expériences, avec une lampe robotisée, pour approcher ces interactions. Nous avons organisé plusieurs journées d'étude dont les principaux travaux sont en cours de publication (André, Collectif Psyphine à paraître 2019) et nous organisons une conférence sur ces mêmes thématiques, intitulée *Drôles d'objets*, en avril 2020 à La Rochelle. Dans ce cadre, l'analyse des interactions homme/robot a permis d'affiner mon propre cadre théorique.

Parallèlement à ces rencontres et travaux, mes enseignements au sein de l'UFR puis du département Sciences du langage de Nancy sont les premiers à bénéficier de mes recherches, en cours ou abouties. Depuis que je suis maître de conférences, j'enseigne une variété de cours en Licence et Master de Sciences du langage. Ils sont tous en lien avec mes recherches et concernent notamment la sociolinguistique, l'analyse des interactions verbales, le recueil et le traitement des données orales et le domaine langage et travail ainsi que ses applications en formation pour un public en insécurité langagière. Au premier semestre 2018, j'ai créé un nouvel enseignement intitulé « Corpus et didactique » pour les étudiants de deuxième année de Master FLE. Cette promotion est la première, au moins à Nancy, à être formée à l'exploitation de corpus à des fins didactiques. Plusieurs étudiants ont consacré leur stage de second semestre à mettre en œuvre cette méthodologie avec différents publics et dans différents pays. J'ai participé à mettre en place leurs expérimentations et certaines seront mentionnées dans ce travail.

Revenons à ce travail de synthèse. Le bain scientifique dans lequel je suis plongée depuis mes premiers pas en recherche a façonné mes propres travaux et orienté mes préoccupations, à la fois scientifiques et sociales. Ce travail est plus particulièrement consacré, comme son titre l'indique, à l'analyse sociolinguistique des interactions verbales et à ses exploitations didactiques. En d'autres termes, j'ai souhaité exposer de quelles façons les analyses de pratiques interactionnelles pouvaient être utiles à l'enseignement et à l'apprentissage du français parlé en interaction. Par conséquent, j'ai souhaité examiner différentes façons d'introduire des corpus d'interactions dans les salles de classe ou dans un dispositif d'apprentissage. Les expérimentations de FLEURON sont sans nul doute à l'origine de ce travail. Lorsqu'ils entrent sur le site Internet, les enseignants et les apprenants explorent d'abord les ressources multimodales. Ils examinent les transcriptions et les sous-

titres. Cette technique m'a amenée à proposer une méthodologie d'analyse des interactions appuyée par la transposition didactique des résultats de la recherche dans ce domaine. L'utilisation du concordancier est inspirée des travaux de Johns (1991) et du *data-driven learning*, qui consiste à appliquer l'utilisation outillée des corpus dans le processus de l'enseignement et de l'apprentissage des langues. En outre, les questions scientifiques que soulèvent l'utilisation de FLEURON sont liées à l'analyse sociolinguistique des interactions, c'est-à-dire à l'analyse de la variation des phénomènes interactionnels dans des situations de communication particulières. Cette analyse révèle les compétences socio-interactionnelles nécessaires pour participer aux échanges communicatifs. En d'autres termes, j'ai cherché à faire entrer dans le champ de la didactique de l'oral en FLE des notions et des méthodologies souvent ignorées ou délaissées.

Les apports et les applications de l'analyse des interactions et des analyses de corpus sont nombreux dans différents domaines, y compris en didactique (voir par exemple Avanzi *et al.* 2014 ; Tyne, André *et al.* 2014 ; Guinamard *et al.* 2015 ; Teston-Bonnard *et al.* 2017 ; Longhi 2017). L'introduction de corpus en situation d'apprentissage doit beaucoup aux réflexions sur les documents authentiques qui visent à exposer les apprenants à la langue telle qu'elle est rencontrée des situations de communication réelles. Cependant, les documents authentiques ne sont pas si fréquents et sont encore aujourd'hui beaucoup issus de la radio ou des médias. Ils sont aussi parfois des transcriptions sans leur support sonore ou visuel (Detey 2010). Le recours aux corpus oraux reste encore marginal (Vialleton, Lewis 2014). Le français « standard [qui] est avant tout une idéologie » reste le choix consensuel des manuels de langue qui construisent une « fiction linguistique » à enseigner (Candéa, Véron 2019 : 51). En outre, la variation n'apparaît guère dans les cours de FLE, « elle est souvent reléguée à un apprentissage informel, lors de séjour en immersion, ou réservée aux niveaux les plus avancés. (...) Le français ordinaire demeure un luxe. » (Paternostro 2017 : 281-281). Les recherches du CRAPEL ont montré qu'un document lui-même n'est pas complexe ou difficile, la difficulté réside dans les tâches définies par la consigne (Holec 1990). Par conséquent, les corpus devraient pouvoir être exploités avec des apprenants, quel que soit leur niveau.

Dans le monde anglo-saxon, les corpus sont exploités à des fins didactiques (voir notamment Aston 2001 ; Sinclair 2004 ; Breyer 2011 ; Leńko-Szymańska, Boulton 2015) même s'ils le sont essentiellement avec des apprenants avancés qui souhaitent améliorer leurs compétences rédactionnelles. En revanche, en France, les corpus tardent à arriver

dans le champ de la didactique. Quelques corpus, notamment d'anglais écrit, de spécialité sont utilisés (voir par exemple Landure 2011 ; Kübler, Hamilton 2018). Cependant, dans la 4^{ème} édition du Cours de didactique de Cuq et Gruca (2017) qui semble être une référence dans le domaine, il n'y a aucune mention de l'utilisation des corpus. Dans le cadre européen commun de référence en langues (CECRL), il est bien question d'interactions mais il semble que, dans ce référentiel, les résultats de la recherche ne soient que peu pris en compte également. Par ailleurs, l'analyse du CECRL montre qu'il a été élaboré (et révisé en 2018) sous l'influence des enjeux idéologiques de la construction européenne et de la mondialisation (Maurer 2011), c'est pourquoi il est peu évoqué ou qu'il est critiqué dans mes travaux.

Enfin, l'objectif de ce travail est de créer des liens entre la sociolinguistique des interactions verbales, l'exploitation des corpus et l'enseignement/apprentissage du français. Ces liens peuvent prendre différentes formes et peuvent être synthétisés dans différentes actions : exposer des apprenants au français parlé en interaction, donner accès aux corpus d'interactions aux enseignants et aux apprenants, les aider à comprendre le fonctionnement des interactions en proposant des transpositions didactiques des résultats de la recherche, utiliser les outils d'exploration des corpus (concordancier), proposer des méthodologies d'observation et d'analyse des données. Ce sont ces différentes actions qui sont traitées, qui ont été expérimentées et qui sont présentées dans ce travail. Celui-ci a permis de mettre en connexion et en synergie l'ensemble de mes activités de recherche et d'en tester le résultat. Ce travail donc est à la fois un aboutissement et le début d'une nouvelle étape. Elle consistera à déployer plus largement les résultats qui ont émergé dans cette synthèse.

Présentation du dossier

Ce dossier comporte trois volumes :

1. Un **curriculum vitae**, détaillant l'ensemble de mes activités de recherche et d'enseignements ; mes responsabilités scientifiques, pédagogiques et administratives ; et mes publications.
2. Une **sélection de mes publications** scientifiques, présentées dans un ordre antichronologique, les écrits les plus récents paraissant en premier, dont deux chapitres d'ouvrage et un article de revue à paraître. Cette sélection a été

opérée en cohérence avec les thématiques de la note de synthèse, ainsi qu'ainsi en fonction de l'importance qu'elles ont eu dans mon cheminement scientifique et dans mes collaborations.

Les publications citées dans la note de synthèse et présentes dans ce volume sont référencées en couleur, de cette façon : (André 2015).

3. La présente **note de synthèse**, intitulée « Sociolinguistique des interactions verbales et exploitations didactiques ».

Cette note synthèse comporte cinq chapitres. Le **premier chapitre** est consacré à mon approche sociolinguistique des interactions verbales. Il précise l'ancrage scientifique de ce travail en mentionnant les outils et les concepts qui sont convoqués pour analyser des interactions. Il montre les évolutions de ce cadre ainsi que ses utilisations dans d'autres études. Il précise également l'apport des sciences du travail et les concepts transférables et utiles à l'analyse des interactions. Le **deuxième chapitre** tente de définir la notion de corpus, il donne des exemples et propose des utilisations possibles de corpus. Il s'intéresse aux données recueillies et constituées en corpus pour analyser les interactions et présente les trois corpus que je coordonne et que j'enrichis. Ce chapitre montrera également que, parmi la chaîne de traitement des données, il est indispensable de documenter ou de donner des informations situationnelles sur les corpus pour qu'ils puissent être analysés d'un point de vue sociolinguistique et exploités en didactique. Le **troisième chapitre** expose des analyses de pratiques interactionnelles, dans le cadre défini au sein du premier chapitre. Ce cadre et ces analyses permettent d'identifier les compétences socio-interactionnelles donnant la possibilité aux locuteurs de participer de façon appropriée aux situations de communication qu'ils rencontrent. Ce chapitre central présente les avantages voire la nécessité d'une interface entre sociolinguistique et didactique. Le **quatrième chapitre** propose une première exploitation des corpus à des fins didactiques. Il insiste, tout d'abord, sur l'indispensable exposition des apprenants à la langue cible. Il propose des transpositions didactiques de certains phénomènes spécifiques à l'oral. Enfin, il rend compte des expérimentations de l'exploitation du corpus FLEURON avec des apprenants souhaitant apprendre à interagir. Le **cinquième et dernier chapitre** met à l'épreuve l'exploitation de corpus avec un concordancier à des fins d'enseignement et d'apprentissage de la langue. L'objectif est d'étudier la façon dont se comportent les enseignants et les apprenants dans

le cadre du *Data-Driven Learning* (DDL) ou, en français de l'Apprentissage Sur Corpus (ASC). Il sera notamment analysé si les apprenants, plus ou moins accompagnés par les enseignants, développent une conscience langagière grâce aux activités métalangagières et cognitives que demande la méthodologie didactique de l'ASC.

Je conclurai ce travail en insistant sur la nécessité de mettre les résultats de la recherche en sociolinguistique des interactions verbales à la disposition des enseignants et des apprenants de langue. Ces derniers ne sont pas des spécialistes de la langue mais il est légitime qu'ils puissent s'approprier les connaissances qui leurs sont utiles. Cette mise à disposition n'est pas toujours évidente, elle demande un travail de transposition à co-construire entre les chercheurs et les praticiens. Cette collaboration est une des conditions de réussite de l'exploitation des corpus à des fins didactiques. Une autre collaboration, entre chercheurs, est également évoquée dans les perspectives de ce travail. Elle permettra, à terme, d'offrir un panel de ressources langagières et interactionnelles ainsi que des propositions d'exploitations qui contribueront à l'enseignement et à l'apprentissage du français parlé en interaction.

Chapitre 1. Un cadre pour une sociolinguistique des interactions verbales

J'ai choisi une nouvelle appellation – « sociolinguistique des interactions verbales » – pour faire référence à une nouvelle approche. Même si cette dernière fait appel à différents courants, disciplines et concepts existants, elle les met en synergie afin de saisir au mieux les interactions verbales, leur fonctionnement, leur dynamique, leurs enjeux, et leurs intrications à la société. Ce premier chapitre présente l'aboutissement de ce que j'avais appelé « une approche multidimensionnelle », il y a un peu plus de dix ans maintenant (André 2006a). Cette approche s'inspirait déjà des travaux éclairants de Kerbrat-Orecchioni qui, elle, qualifie son approche d'« éclectique » (2005 : 5). Je me retrouve parfaitement dans les principes d'une telle approche qui convoque les outils adéquats pour analyser les données plutôt que de les observer au prisme d'une école de pensée ou d'un courant d'analyse spécifique. Ces approches éclectique ou multidimensionnelle répondent aux besoins d'une analyse de la complexité des interactions verbales. En d'autres termes, ce sont les données à analyser, leur nature et leur forme, qui conduisent le chercheur à recourir à différents outils.

Si l'« approche multidimensionnelle » de ma recherche est devenue aujourd'hui une « approche sociolinguistique » des interactions verbales, ces deux expressions renvoient à des approches qui ne sont ni contradictoires ni exclusives. Néanmoins la seconde est plus précise et définit davantage les enjeux des interactions, leurs influences et leurs intrications à la situation ou à la société. De plus, cette approche sociolinguistique est nécessairement multidimensionnelle dans la mesure où elle fait appel à de nombreuses dimensions de la langue à la fois individuelle, interactionnelle, collaborative, praxéologique, sociale ou encore située.

Ce chapitre sera donc consacré à la présentation de cette approche sociolinguistique des interactions verbales, sans m'aventurer trop en avant (ou trop en arrière) dans des questions épistémologiques mais suffisamment pour justifier ce point de vue scientifique. J'aborderai cette approche tout d'abord en présentant un schéma synthétique qui peut être utilisé comme point de départ pour exposer les spécificités du cadre théorique, affiné dans le cadre de cette synthèse ([1.1.](#)). Ensuite, je montrerai que cette approche a été et est exploitée au sein de différentes études qui appréhendent les interactions verbales dans

différents contextes. J'insisterai sur les contributions qu'elle apporte et, réflexivement, sur les apports de ces études. J'expliquerai pourquoi j'ai préféré *sociolinguistique* à *multidimensionnelle* pour qualifier cette approche et, plus généralement, pourquoi il me semble indispensable d'adopter une approche sociolinguistique de la langue (1.2.). Enfin, je pointerai les apports et les implications de l'analyse des situations de travail à cette approche (1.3.).

1.1. Une schématisation de l'analyse sociolinguistique des interactions verbales

En plusieurs années, mon approche sociolinguistique des interactions verbales s'est affinée, ainsi que le schéma qui me permet de la représenter (André 2014, 2015a) – peut-être plus clairement. Le schéma ci-dessous est l'un des aboutissements de mon cheminement scientifique et va constituer le point de départ de cette synthèse théorique. Il peut être représenté de la façon suivante :

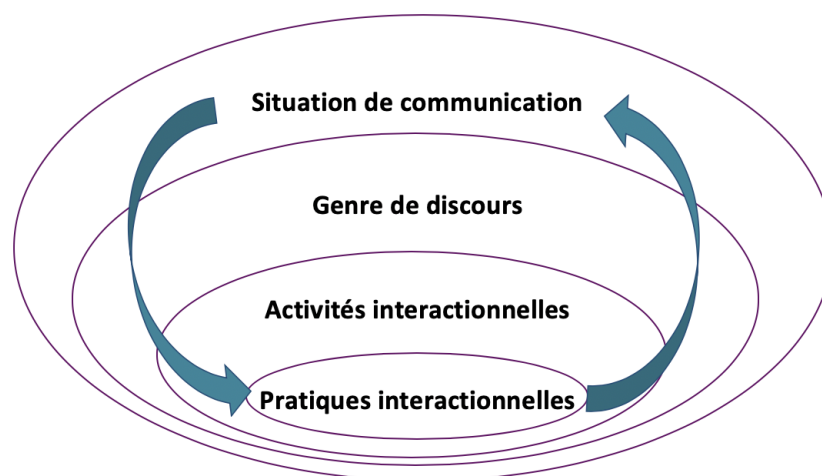


Figure 1 : Schéma - Intrications et influences dans les interactions verbales

Il est légèrement différent des schémas que j'ai conçus auparavant. Les pratiques et les activités *langagières* (André 2014) sont devenues des pratiques et des activités *interactionnelles*, notamment parce que les pratiques langagières des locuteurs sont en réalité toujours interactionnelles (sauf cas pathologiques). Avant de commenter ce schéma, je vais apporter quelques précisions terminologiques :

- Les **pratiques interactionnelles** sont les ressources langagières et non langagières (verbales et non verbales) mises en œuvre par les locuteurs pour interagir. Ces ressources sont variées et concernent par exemple le lexique, les constructions syntaxiques, les répétitions, les reformulations, les marqueurs de discours, les ponctuations spécifiques à l'oral, la prosodie, les interjections, les bruits-sons (pff, mh, tss), les gestes et les postures, les regards et les expressions faciales. Elles sont tournées vers le ou les interactants et fonctionnent en interaction. Elles prennent sens dans la temporalité de l'interaction et configurent l'interaction. Comme toutes les pratiques sociales, elles s'inscrivent dans une situation. Les pratiques accomplissent des activités sociales et interactionnelles.
- Les **activités interactionnelles** désignent ce que font ou ce qu'accomplissent les locuteurs avec leurs pratiques interactionnelles, les buts qu'elles atteignent (leurs valeurs et effets pragmatiques) dans et grâce à l'interaction. Les activités interactionnelles sont de différentes natures.
 - Elles correspondent à des activités telles que poser des questions, s'excuser, expliquer, raconter une histoire, complimenter, informer, reprocher, convaincre, saluer, remercier, inviter, insulter, *etc*³.
 - Ces mêmes activités peuvent être accomplies avec plusieurs activités différentes. Par exemple, convaincre son interlocuteur est une activité interactionnelle mais elle peut être accomplie avec plusieurs autres activités telles qu'argumenter, expliquer, illustrer, *etc*. Dans ce cas, une macro-activité est accomplie grâce à plusieurs micro-activités (Kerbrat-Orecchioni 2001a). Les différentes activités interactionnelles peuvent donc être entremêlées pour accomplir les objectifs de la communication⁴.
 - Les activités interactionnelles peuvent être complexes. Elles peuvent avoir plusieurs valeurs pragmatiques. C'est par exemple le cas de *poser une question* qui peut cumuler les valeurs pragmatiques *obtenir des informations*

³ La traditionnelle notion d'actes de parole telle qu'elle a été décrite par Austin (1962) puis reprise par Searle (1969) et en pragmatique est largement revisitée dans le cadre de cette approche (voir également Kerbrat-Orecchioni 2001a).

⁴ J'ai choisi de ne pas faire de différences entre *actions langagières* et *activités interactionnelles* (par exemple comme Ravazzolo *et al.* (2015 : 223) le font), d'abord parce que le terme *activité* correspond davantage à une dimension praxéologique du langage, que j'expliquerai en 1.3., et ensuite parce que toutes les activités se réalisent en interaction et qu'elles sont orientées vers et co-construites avec les interlocuteurs, d'où le terme *interactionnelles*. Les activités interactionnelles, telles que je les définis, ont donc plusieurs dimensions.

et *faire un reproche* (*Ibid.* : 96). Elles peuvent également avoir des valeurs pragmatiques non prévues par les locuteurs. C'est le cas d'un compliment sincère interprété comme un reproche.

- Les activités interactionnelles structurent également les interactions lorsqu'elles ouvrent les échanges, permettent de changer de thème, de prendre la parole, d'interrompre son interlocuteur, de le compléter ou encore de clore les échanges (Traverso 2016).
 - Ces activités peuvent être négociées, acceptées, refusées, co-construites ou réalisées conjointement par les interlocuteurs.
 - Les activités interactionnelles ont des effets pragmatiques, c'est-à-dire des incidences ou des conséquences sur les interactions et sur la situation de communication. Ces effets ne sont pas toujours ceux escomptés par les locuteurs.
- Le **genre de discours** correspond au type d'interaction, au *speech event* (Hymes 1972) ou encore au G1 qui est « un ensemble discursif plus ou moins institutionnalisé dans une société donnée » selon Kerbrat-Orecchioni et Traverso (2004 : 42). Ils sont définis par des critères extralinguistiques, tels que le cadre situationnel, le cadre participatif, le degré de formalité ou l'objectif de l'échange. Les genres de discours sont par exemple : des conversations, des réunions de travail, des entretiens, des débats, des conférences, des cours, des relations transactionnelles, etc. Ils sont constitués d'activités interactionnelles. En d'autres termes, ces « G1 sont composés de G2 », selon la terminologie de Kerbrat-Orecchioni et Traverso (2004 : 44), c'est-à-dire composés d'activités interactionnelles.
 - La **situation de communication** est constituée de l'ensemble des éléments extralinguistiques qui conditionnent le déroulement de l'interaction. Elle peut être, en partie, décrite à l'aide du modèle SPEAKING de Hymes (1972) qui permet de mémoriser les principaux éléments de la situation et qui se décline en : *setting*, *participants*, *ends*, *act sequences*, *keys*, *instrumentalities*, *norms*, *genres*. Il s'agit de décrire de façon ethnographique la situation dans laquelle prend place une interaction. La situation est composée de tous les éléments qui préexistent à l'interaction. Par exemple, la notion de *participants* comprend de multiples éléments, tels que le cadre participatif (Goffman 1987), le nombre de locuteurs,

l'identité des locuteurs, leurs statuts, fonctions, rôles sociaux et conversationnels, leurs relations interpersonnelles et hiérarchiques, mais également leur histoire conversationnelle (Golopentia 1988). Les situations de communication ont une épaisseur historique (Boutet 1994), c'est-à-dire qu'elles s'inscrivent dans une histoire sociale, économique, culturelle particulière. La description des éléments de la situation de communication permet de s'interroger sur leur influence en interaction.

Ce schéma a pour objectif de représenter **les multiples intrications et influences en jeu dans les interactions verbales**. Les pratiques interactionnelles sont les productions verbales et non verbales des participants aux interactions. Elles permettent de réaliser des activités interactionnelles, telles que raconter, expliquer, informer, dans un genre de discours particulier (conversation, cours, conférence, relation de service) et dans une situation de communication particulière, telle qu'elle peut être définie à partir du modèle descriptif SPEAKING de Hymes (1972). Les cercles représentent ces intrications, chacun des éléments du schéma est lié aux autres. Les flèches symbolisent les influences traversant ces quatre cercles. Plus précisément, les flèches illustrent la réciprocité des influences, de la situation de communication vers les pratiques interactionnelles et des pratiques vers la situation. Ainsi, les pratiques interactionnelles sont toujours situées, influencées par les éléments de la situation et, de façon réflexive (Garfinkel 1967), transforment certains de ces éléments.

Il me semble que ce schéma peut représenter, au moins de façon mnémotechnique ou synthétique, les principales influences et intrications qui entrent en jeu lors de l'analyse d'interactions verbales. Chaque élément du schéma renvoie à un ensemble de choix théoriques, conceptuels, d'outils et de courants d'analyse ou encore d'éléments transdisciplinaires, c'est-à-dire traversant les disciplines scientifiques et apportant des éclairages sur mon objet d'étude. Ce sont ces choix qui vont être expliqués ici dans la mesure où ils ont contribué à la construction d'un cadre d'analyse qui me paraît adéquat pour appréhender tous types d'interactions verbales. Les choix et leurs justifications sont présentés de façon linéaire. Néanmoins, ils sont liés les uns aux autres et n'ont pas forcément été réalisés selon la chronologie rédactionnelle qui suit.

Tout d'abord, ce schéma est largement influencé par les travaux en socio-ethno-anthropo-linguistique et par la nécessité d'une approche qui lie les faits langagiers et les

faits sociaux. Cette approche m'a été présentée pour la première fois par Philip Riley dès ma première année à l'université. Elle a d'abord provoqué chez moi une sorte d'évidence sociale avant de devenir une évidence scientifique. Les études des pratiques communicatives dans des groupes sociaux (voir notamment Hymes 1964, Gumperz et Hymes 1972 ou encore Saville-Troike 2002) ont particulièrement attiré mon attention et ont contribué à la construction de mon cadre scientifique. L'étude de la parole chez les Dogon de Calame-Griaule (1965)⁵ a été quasiment le point de départ de mon intérêt pour la sociolinguistique. L'analyse méticuleuse, réalisée par Geneviève Calame-Griaule, de la conception de la parole des Dogon ainsi que de l'importance qu'ils lui accordent est fascinante. En outre, elle présente une approche de la parole et de son fonctionnement dans un contexte social et culturel particulier à un moment où la sociolinguistique américaine et l'ethnographie de la communication n'ont pas encore traversé l'océan Atlantique et où, en France, les travaux de Cohen, défendant une sociologie du langage (1956), sont encore ignorés. L'aspect presque révolutionnaire de cette approche, dans mon parcours universitaire consacré essentiellement à la linguistique formelle, a accentué ma curiosité pour cette étude. Ainsi, cet intérêt pour l'ethno-sociolinguistique, qui étudie le sens et le rôle de la parole dans une société, a largement influencé mes orientations scientifiques ainsi que mes choix conceptuels.

Naturellement, je me suis intéressée à la langue parlée et à ses variations. L'approche pronominale et les travaux pionniers sur le français parlé impulsés par le Groupe Aixois de Recherche en Syntaxe (Blanche-Benveniste, Jeanjean 1987 ; Blanche-Benveniste 1990a) ont rapidement orienté mes intérêts vers la description de l'élaboration du discours à l'oral (Richard 2018). Les difficultés à la fois politiques et scientifiques que rencontrent les chercheurs sur le français parlé dans les années 1990 sont, en partie, à l'origine de mon engagement dans certaines études. Ces difficultés sont liées au poids de la langue écrite en France, au mépris des études sur l'oral et à l'absence de financement de programmes de recherche consacrés à l'oral (jusqu'à la fin des années 2000). En 1999, le discours sur l'état de la langue prononcé par Maurice Druon à l'Académie Française explique à propos du français oral vernaculaire ou ordinaire :

Ainsi ramasse-t-on dans le ruisseau tous les détritiques tombés des poubelles langagières. Mais comme ces horreurs sont « dans un dictionnaire » chacun peut se croire encouragé à les

⁵ En 2009, les éditions Lambert-Lucas ont publié une troisième édition revue et corrigée (après celle de 1987 publiée par l'Institut d'Ethnologie de Paris).

utiliser. « Viens, ma cocotte, on va se zoner ». Voilà le parler de la génération du PACS, acronyme qui était déjà entré dans l'usage avant même que la loi qu'il désigne ait été votée. Voilà ce que l'on nous affirme être « le vrai français »

La langue verte a de la saveur certes ; mais quand elle verdit trop, c'est qu'elle pourrit.

Cette volonté délibérée de « faire populaire », cette démagogie lexicale, ce gauchisme du vocabulaire sont, comme tous les gauchismes, un péché de l'esprit.⁶

Sans accorder trop d'importance aux discours de cette institution, ses opinions se diffusent, y compris dans les sphères les plus hautes de la recherche scientifique. Même si mes premiers pas en recherche sur l'oral, au début des années 2000, sont facilités par mes prédécesseurs, « il reste encore beaucoup à faire pour décrire les différents usages du français parlé » (Blanche-Benveniste 2000 : 3). Les freins politiques, liés au poids et au prestige de l'écrit, sont une forte motivation pour poursuivre le travail entamé. De plus, la langue parlée, abordée d'un point de vue sociolinguistique, est sujette à la variation ou aux variations (Labov 1969/1978 ; Gadet 2003a). Ce sont ces variations et leurs origines extralinguistiques qui vont motiver mes travaux.

La linguistique interactionnelle, ou les études du français parlé *en interaction*, (Kerbrat-Orecchioni 1990-1994 ; Vion 1992 ; Gülich, Mondada 2001 ; Mondada 2002 ; Traverso 1996, 1999, 2016) a été naturellement une grande inspiration pour ce modèle d'analyse. Une attention et une description fine des ressources, langagières et non langagières, utilisées par les locuteurs pour élaborer leur discours, au fur et à mesure de l'interaction, sont à l'origine de l'analyse sociolinguistique des interactions. Dès le début de mon cheminement scientifique, la lecture de l'article fondateur de Sacks, Schegolff et Jefferson (1974) a été l'occasion de découvrir et d'entrer pleinement dans un cadre d'analyse minutieux et méthodique des interactions. Ainsi, plusieurs des principes de l'approche interactionnelle sont adoptés. Le premier est que « la parole en interaction est une parole toujours en train d'être produite et toujours en train d'être reçue » (Traverso 2016 : 17). Le second principe, nous y reviendrons, met en évidence la co-construction du discours : « il s'agit d'une parole produite en interaction, c'est-à-dire *pour* et *avec* ceux à qui elle est adressée⁷ » (*Ibid.*). Le caractère situé de l'interaction est également fondamental dans l'approche interactionnelle même si différentes positions sont adoptées en ce qui concerne la préexistence du contexte. L'analyse conversationnelle d'inspiration

⁶ Disponible en ligne : <http://www.academie-francaise.fr/discours-sur-letat-de-la-langue-seance-publique-annuelle-3>

⁷ En italique dans le texte.

ethnométhodologique appréhende le contexte comme un élément qui se construit uniquement en interaction, ou plus précisément, « la parole contribue à accomplir le contexte d'une part, et (...) constitue elle-même le contexte d'autre part » (Mondada 2006 : 119). Mon approche n'adopte pas cette position, elle envisage le contexte à la fois comme préexistant à l'interaction et comme défini ou redéfini par et dans l'interaction. Certains des éléments de la situation sont bien présents avant même l'ouverture de l'interaction mais ils peuvent tout à fait être modifiés par l'interaction (cette conception de la communication est représentée par les flèches du schéma). Les éléments de la situation sont pris en compte lorsqu'ils semblent pertinents pour les participants et qu'ils influencent leurs pratiques interactionnelles.

Ensuite, ce schéma a vu le jour avec l'influence des travaux de recherche effectués dans le domaine « Langage et Travail » (Borzeix, Fraenkel 2001). J'ai participé aux derniers séminaires du réseau « Langage et travail » fondé en 1985, de 2001 à 2007. Ces rencontres ont été éclairantes pour moi parce qu'elles associaient ethnométhodologie, ethnographie de la communication et théories de l'activité. Il s'agissait de saisir les ressources langagières mobilisées par les salariés pour effectuer leur travail. Les situations de travail sont typiquement des objets d'étude de la sociolinguistique puisqu'il s'agit d'analyser le langage dans des situations sociales particulières. La construction de mon cadre d'analyse s'est poursuivie avec mes propres travaux au sein du groupe de recherche « Langage, Travail et Formation » (LTF), que je co-dirige avec Hervé Adami depuis sa création en 2006. En 2011, Josiane Boutet, une des fondatrices du réseau « Langage et travail » nous a passé le relais pour créer un nouveau réseau international du même nom - « Langage, Travail et Formation » - qui intègre donc désormais une dimension didactique⁸. Je reviendrai sur cette partie de mon activité à plusieurs moments dans ce travail de synthèse. La conception de la langue prend alors, dans ce cadre d'analyse, une dimension éminemment sociale et praxéologique (Filliettaz 2002, Bronckart 2008). Cette dernière transparait dans les termes « pratiques » et « activités » interactionnelles du schéma. L'analyse de l'activité et l'analyse du travail ont participé à l'élaboration de mon cadre d'analyse, j'y reviendrai dans la dernière partie de ce chapitre.

L'analyse des interactions verbales dans différentes situations, et notamment au sein de réunions de travail (André 2006a), m'a amenée à donner une place importante au genre de discours. Le genre fait partie de la liste des éléments de la situation de

⁸ Voir le site du réseau LTF : <https://reseaultf.atilf.fr/>.

communication telle que la définit Hymes (1972) mais mérite une place particulière tant son influence sur les interactions verbales est prégnante (Bakhtine/Volochinov 1977/1929 ; Bakhtine 1984 ; Haliday 1985 ; Biber 1988). Comme je l'ai expliqué dans les précisions terminologiques au début de ce point, le genre correspond à ce que Kerbrat-Orecchioni et Traverso (2004) ont appelé le G1, c'est-à-dire l'évènement de communication ou, selon l'ethnographie de la communication, le *speech event* (Hymes 1974). La notion de genre de discours occupe une place centrale dans mon approche. C'est pourquoi elle est représentée schématiquement par un cercle dans le schéma. Nous verrons qu'elle occupe également une place centrale dans l'exploitation de ce cadre à des fins didactiques. Bakhtine (1984 : 284) précisait déjà : « Si les genres de discours n'existaient pas et si nous n'en avions pas la maîtrise, et qu'il nous faille les créer pour la première fois dans le processus de la parole, qu'il nous faille construire chacun de nos énoncés, l'échange verbal serait quasiment impossible ». Ainsi, le genre de discours crée des attentes de la part des locuteurs, il conditionne l'apparition et l'interprétation des activités et des pratiques interactionnelles. L'analyse de Hymes (1984 : 61) rejoint cette position, il précise que « la notion de genre suppose qu'il est possible d'identifier des caractéristiques formelles qui sont traditionnellement reconnues ». Cette affirmation est liée à la notion de compétence de communication. Pour interagir de façon appropriée, il est nécessaire de saisir les attentes ouvertes par le genre de discours. Le genre est à l'origine de nombreuses variations langagières. C'est pourquoi il représente un élément fondamental dans l'analyse des interactions. L'analyse de différents corpus a montré l'importance des genres (Biber, Conrad 2009). Biber (1988) a montré qu'un faisceau de paramètres linguistiques permet d'identifier des genres de textes et de discours. Dans mon cadre théorique, comme l'explique Cappeau (2016 : 33), « c'est la mise au jour d'un genre qui permet d'apprécier le matériau linguistique employé ».

A ce schéma synthétique s'ajoute une conception collaborative de la construction du discours en interaction (André 2014). Les genres interactifs oraux sont sans cesse négociés et conjointement élaborés par les locuteurs dans leur situation de communication. Ainsi, comme l'avait déjà précisé Kerbrat-Orecchioni (1990 : 13) : « Tout discours est une construction collective ». Les pratiques interactionnelles sont donc analysées comme étant élaborées progressivement et collectivement par les interlocuteurs, au fur et à mesure des interactions ou encore dans leur séquentialité (Garfinkel 1967 ; Sacks, Schegolff, Jefferson 1974). Les locuteurs s'ajustent en permanence les uns aux autres, adaptent leurs pratiques

en fonction de celles qui sont déjà produites, ainsi qu'en fonction des éléments de la situation de communication et du genre de discours.

Ainsi, d'une part, les approches sociale, ethnographique, interactionnelle, conversationnelle, discursive, pragmatique, praxéologique, énonciative mais aussi pronominales et syntaxiques de la langue parlée alimentent le modèle d'analyse sociolinguistique des interactions verbales. D'autre part, ce modèle est réinterrogé sans cesse par les données qu'il permet d'analyser. La complexité des interactions demande à la fois souplesse et rigueur dans la façon d'appréhender les données. Mon attention se porte sur les traces laissées et les ressources employées par les interlocuteurs pour construire collaborativement un discours commun, en d'autres termes, et en utilisant l'image de la marche, pour « marcher ensemble » (Gilbert 2003). Ces traces et ces ressources sont parfois attendues, parce qu'elles sont fréquentes, récurrentes ou classiques dans les interactions, mais également parfois surprenantes, parce qu'elles sont inhabituelles et extraordinaires. Observer et analyser les interactions verbales offrent au chercheur une source inépuisable de surprises, d'étonnement et de fascination tant les façons d'interagir des locuteurs sont variées et diversifiées.

1.1.1. Démonstration et illustration du schéma-cadre

Après ces brèves descriptions et explications du cadre sociolinguistique des interactions verbales, je vais préciser, grâce au schéma, les applications et les implications de ce cadre. Pour commencer, je choisis volontairement un extrait court d'une interaction pour faciliter une mise en pratique du schéma et la compréhension de ses sous-jacents et fondations théoriques. Voici un extrait d'une interaction verbale entre deux locuteurs :

Exemple 1 : Une jeune femme interroge un ami qui vient de quitter l'armée française (TCOF – Armée_MER_15⁹).

1. L1 mais toi tu en penses quoi de l'armée franchement /
2. L2 à l'heure d'aujourd'hui ça vaut rien

⁹ Cette interaction fait partie du corpus nancéen TCOF présenté dans le chapitre 2, accessible à l'adresse suivante : <http://cnrtl.fr/corpus/tcof/> et sur la plateforme ORTOLANG : <https://www.ortolang.fr/>. En ce qui concerne les conventions de transcription dans cet extrait : la barre oblique (/) symbolise une intonation montante et le plus (+) une pause. L'intégralité des conventions de transcription est consultable sur le site du corpus TCOF.

- 3. L1 ***c'est-à-dire*** /
- 4. L2 ça vaut plus rien du tout
- 5. L1 ben explique /
- 6. L2 ben quand on voit euh comment + les les moyens euh déployés par euh l'armée française et les moyens déployés par euh l'armée américaine

Dans cet extrait, les deux premiers tours constituent une paire adjacente du type question-réponse. Ensuite, L1 produit *c'est-à-dire* (ligne 3). C'est ce tour que je vais analyser plus précisément en utilisant le schéma. Ce tour est tout à fait particulier, L1 utilise une locution, qui selon les grammaires traditionnelles ne fonctionne pas seule, afin de relancer son interlocuteur ou de l'inviter à préciser son discours. En d'autres termes et dans les termes du cadre proposé, L1 produit une pratique interactionnelle (*c'est-à-dire*) qui lui permet d'accomplir une activité interactionnelle (inviter à préciser ou à poursuivre). D'ailleurs, à la ligne 4, L2 saisit la place de transition laissée par son interlocuteur (Sacks, Schegloff, Jefferson 1974) et répond. Ces pratiques et activités interactionnelles apparaissent dans une conversation (Traverso 1996), un genre de discours qui contraint assez peu les interventions des locuteurs. Enfin, ces éléments prennent place dans une situation de communication particulière dans laquelle : les interlocuteurs sont jeunes (L1 a 20 ans et L2 a 24 ans) ; ce sont des amis de longue date ; ils sont nés en Lorraine ; ils se connaissent bien ; ils partagent un certain nombre de connaissances sur la vie de chacun ; ils ont une histoire conversationnelle (Golopentia 1988) ; ils sont installés dans le canapé de L1 (dans son appartement).

C'est en prenant en compte l'ensemble de ces éléments, même si certains ont plus d'influence que d'autres, qu'il est possible d'analyser les interactions entre ces locuteurs. Cet exemple est relativement simple mais permet de saisir les intrications et les influences proposées par le schéma. Afin d'être tout à fait claire, il est possible de reproduire le schéma avec cet exemple :

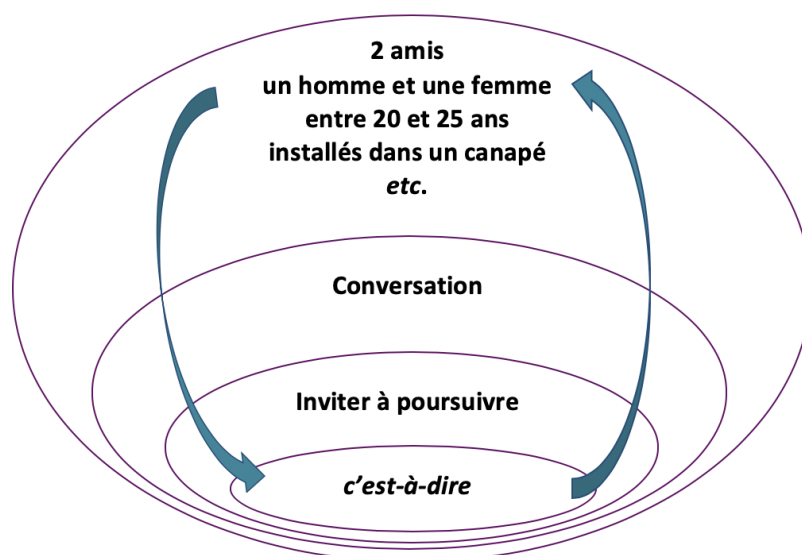


Figure 2 : Illustration du schéma avec l'analyse d'une pratique interactionnelle

Ainsi, l'analyse de l'actualisation de *c'est-à-dire* n'a pas de sens si elle se passe des autres éléments du schéma. En bref, c'est parce que les locuteurs sont ce qu'ils sont, qu'ils entretiennent la relation qu'ils entretiennent, qu'ils sont installés confortablement, qu'ils ont une conversation spontanée – non préparée, non formelle – que la production de *c'est-à-dire* est permise pour accomplir l'activité interactionnelle visée. De plus, cette pratique interactionnelle, légitimée par les éléments déjà mentionnés, permet d'obtenir l'effet pragmatique escompté, faire poursuivre ou relancer L2 (même si cette poursuite sera ensuite négociée avec la reformulation de L1 : *ben explique*, ligne 5). En outre, les flèches marquent les influences et les conséquences des éléments du schéma. La flèche descendante, comme je viens de l'expliquer, montre que les éléments du schéma conditionnent une telle pratique interactionnelle. Réflexivement, l'actualisation de cette pratique influence les interactions, puisqu'elle demande à l'interlocuteur d'accomplir une nouvelle activité interactionnelle ou de préciser celle qui vient d'être réalisée. L2 s'ajuste à son interlocuteur et coopère, il prend son tour en reformulant légèrement ses propos.

L'analyse de ce tour est réalisée de façon isolée pour comprendre le fonctionnement du schéma et pour analyser une pratique interactionnelle spécifique. Néanmoins, le tour de L1 analysé ici apparaît dans une temporalité particulière qui est prise en compte dans l'analyse. Les interlocuteurs s'ajustent afin de construire collaborativement un discours commun. En outre, la compréhension du discours est assurée par l'espace intersubjectif qui se crée entre les deux locuteurs. En effet, en entrant en interaction, deux subjectivités se

rencontrent pour créer un espace intersubjectif dans lequel une partie de chacun des mondes subjectifs des locuteurs se superpose (Riley 2001).

Pour continuer à illustrer mon cadre sociolinguistique des interactions verbales, je vais analyser un second exemple. Il s'agit d'une interaction entre deux étudiants.

Exemple 2 : Une étudiante (E1) va se présenter à son voisin (E2) (FLEURON – Se présenter à son voisin)¹⁰.

1. E1 bonjour
2. E2 salut
3. E1 je viens de visiter l'appart du dessus et du coup ben:: c'est bon il est pour moi donc je viens me présenter je suis la nouvelle voisine
4. E2 ok salut ben moi c'est Pierre
5. E1 Siam enchantée du coup ben je déménage d'ici la semaine prochaine
6. E2 < ok
7. E1 *** > voilà c'est tout c'était pour faire un petit coucou
8. E2 ouais ben c'est sympa merci ben si tu as besoin de quoi que ce soit < tu me dis
9. E1 ok et ben > ça marche pas de souci je retiens
10. E2 cool
11. E1 voilà bon ben bonne journée
12. E2 à toi aussi merci
13. E1 au revoir

Cette interaction et certaines des pratiques interactionnelles de cet extrait peuvent être analysées à partir des éléments suivants :

- En ce qui concerne la **situation de communication**, les éléments les plus pertinents (et accessibles à l'analyste) sont les suivants : les locuteurs sont des étudiants, âgés d'environ 20 ans, ils ne se connaissent pas, habitent dans une résidence réservée aux étudiants. L'interaction se déroule sur le palier de E2.
- Le **genre de discours** auquel les locuteurs participent semble être une conversation. On peut toutefois préciser qu'il s'agit d'une conversation de rencontre ou de visite.
- Les **activités interactionnelles** de cette interaction sont diverses malgré la brièveté de l'échange (32 secondes). On peut identifier une structure bien connue en ce qui concerne l'organisation de l'interaction : ouverture – corps – clôture. Les tours ligne 1 et 2 sont une paire adjacente de salutations, ils ouvrent l'interaction. Ligne 3, E1 explique la raison de sa visite. Ligne 4, E2 signifie qu'il a compris, renouvelle ses

¹⁰ Cette interaction fait partie du corpus multimodal FLEURON présenté dans le chapitre 2, accessible à l'adresse suivante : <https://fleuron.atilf.fr/>. Les conventions de transcription de cet extrait : un allongement *ben::*, une suite de syllabes inaudibles *****, des chevauchements de paroles entre <...>.

salutations et donne son prénom. En réponse, E1 donne également son prénom, ajoute une formule de politesse et précise une information quant à son déménagement. Ligne 6, E2 montre qu'il a compris et E1 prend la parole (ligne 7) en même temps. Elle garde le tour et réalise une pré-clôture de l'échange. En réponse, ligne 8, E2 formule une évaluation de l'activité de son interlocutrice et propose son aide. Avant qu'il n'ait terminé son tour E1 prend la parole (ligne 9) pour lui confirmer la réception de l'aide. A partir de la ligne 10, la clôture de l'échange se fait progressivement avec différentes formules de politesse. On peut également identifier une macro-activité interactionnelle – se présenter à son voisin – réalisée avec différentes micro-activités – saluer, donner son prénom, proposer de l'aide, remercier, etc.

Ces éléments influencent les **pratiques interactionnelles** des locuteurs (flèche descendante du schéma). Par exemple, le premier tour est une salutation réalisée avec *bonjour* et l'interlocuteur répond *salut*. Cette pratique est appropriée parce que les locuteurs sont des étudiants, dans une résidence universitaire. Dans la grande majorité des échanges entre locuteurs du même âge, les pratiques de salutations sont identiques. Le locuteur va répéter *salut* lors de son deuxième tour (ligne 4). On peut imaginer que cette répétition vise à mettre à l'aide son interlocutrice qui arrive dans la résidence. Cette pratique interactionnelle viserait donc plusieurs activités interactionnelles : saluer de nouveau et mettre à l'aise en guise de bienvenue. De façon réflexive, ces pratiques interactionnelles redéfinissent la situation de communication de départ, l'ambiance est plus détendue, les étudiants ont fait connaissance (flèche ascendante du schéma).

Par ailleurs, les locuteurs se tutoient (ligne 8). On peut tout de même faire l'hypothèse d'une stratégie d'évitement mise en œuvre par la nouvelle voisine. De plus, les échanges sont spontanés, les locuteurs n'ont pas de rôle hiérarchique ou conversationnel particulier. L'alternance des tours de parole respecte la règle du chacun son tour avec quelques « ratés » qui conduisent à des chevauchements de paroles. Plusieurs petits mots de l'oral ou marqueurs sont actualisés pour articuler les activités interactionnelles (*du coup*, *ben*, *ok*, *voilà*). En résumé, certains de ces phénomènes sont influencés par le genre de discours, notamment l'alternance des tours ou la symétrie entre les participants.

De façon plus globale, certains éléments de la situation de communication conditionnent l'existence même de cette interaction. En effet, cet échange existe parce que les locuteurs sont des étudiants et qu'ils habitent une résidence réservée aux étudiants. Il

est habituel de se présenter à ses voisins dans ce type de résidence mais ce n'est pas une pratique très courante, notamment dans un milieu urbain. Les présentations entre voisins se font mais plutôt au hasard des rencontres sur les paliers. Dans cette situation, l'étudiante sonne à la porte de son voisin (un étage plus bas) *pour faire un petit coucou* (ligne 7). En outre, même si l'étudiante qui arrive est un peu réservée au début, les pratiques et activités interactionnelles co-construites par les locuteurs contribuent à instaurer une ambiance sympathique et presque amicale entre des individus qui ne s'étaient jamais rencontrés auparavant. Pour finir, cette analyse est partielle et non exhaustive, d'autres éléments interactionnels pourraient être analysés au prisme de la situation.

Toutes les analyses mentionnées dans ce travail sont effectuées en suivant ce modèle. Tout comme le SPEAKING de Hymes (1972), deux traitements successifs et indissociables doivent être menés pour rendre ce modèle opérationnel. Tout d'abord, ce modèle est descriptif. Il présente un moyen visuel et mnémotechnique permettant de faire la liste des éléments à prendre en compte lors de l'analyse d'interactions. Ensuite, l'étape suivante consiste à rendre ce modèle analytique en s'interrogeant sur les interinfluences et les intrications de ses éléments. Il s'agit d'analyser comment les locuteurs interagissent, pourquoi ils le font et s'ils réussissent à le faire. En d'autres termes, il s'agit de questionner et d'analyser l'actualisation des traces et des ressources interactionnelles des locuteurs au fur et à mesure de leurs échanges dans la situation de communication à laquelle ils participent. Ce modèle est exploitable pour analyser tous types d'interactions verbales et non verbales.

1.1.2. Exploitation du cadre dans d'autres études

Ce cadre d'analyse sociolinguistique des interactions verbales a été utilisé et mis à l'épreuve dans différentes études que j'ai encadrées ou auxquelles j'ai participé. Par exemple, dans sa thèse, Valérie Langbach a analysé des interactions verbales dans lesquelles est engagé au moins un locuteur peu qualifié et en situation d'insertion (Langbach 2014). Elle s'est inscrite dans cette analyse sociolinguistique des interactions verbales, avec une attention particulière pour les travaux de Hymes (1972, 1984) et pour le débat entre Bernstein (1975) et Labov (1978). Ce cadre théorique a notamment permis de mettre au

jour et d'analyser les problèmes de maîtrise de la langue orale en interaction chez des locuteurs peu qualifiés. Elle remarque, par exemple, que certains problèmes sont liés au fait que les locuteurs ne prennent pas en compte les éléments de la situation de communication. Elle précise : « Pour qu'un échange se déroule de manière satisfaisante, les locuteurs doivent prendre en compte l'ensemble des composantes du modèle [SPEAKING de Hymes] et produire un discours approprié à l'ensemble de celles-ci. » (Langbach 2014 : 56). Ainsi, un locuteur qui ne présente pas de problème de maîtrise de la langue s'appuie, inconsciemment évidemment, sur les éléments de la situation de communication pour se comporter de façon appropriée. Cependant, certains locuteurs, souvent ceux qui sont peu habitués à interagir dans des situations qui ne sont pas des conversations de connivence entre proches, ne respectent pas les contraintes imposées par la situation de communication et le genre de discours auxquels ils participent. La schématisation du cadrage théorique, telle que je la propose, facilite la compréhension de ces phénomènes d'insécurité langagière, qui empêchent les interactions réussies (voir Adami, André 2014 pour une définition, j'y reviendrai dans le chapitre 2, en [2.4.](#)). Certaines pratiques interactionnelles peuvent se révéler inappropriées au regard des autres éléments du schéma. Ces pratiques peuvent ne pas réussir à accomplir les activités interactionnelles visées. Par exemple, Langbach (2014) montre que certains locuteurs ne réussissent pas à faire état de leur parcours professionnel de façon satisfaisante. Dans certains entretiens d'insertion, un genre de discours tout à fait particulier qui contraint le contenu thématique ainsi que l'apparition de certaines activités interactionnelles, des locuteurs évoquent des thèmes qui ne sont ni prévus ni attendus. Ces « ruptures thématiques » (*Ibid.* : 229) montrent que des locuteurs ne s'appuient pas sur les éléments de la situation et qu'ils ne saisissent pas les spécificités des genres de discours. Ces analyses montrent également que les multiples intrications et interinfluences de l'ensemble des éléments du cadre ne sont pas perçues, ce qui nuit à la production d'un discours approprié. Je reviendrai sur la poursuite de ces travaux dans le chapitre 3, en [3.2.](#)

Deux années plus tard, dans sa thèse, Guillaume Nassau, qui s'est intéressé à l'expression des émotions dans un dispositif d'apprentissage de langue, a analysé des entretiens de conseil en s'inspirant du schéma afin de ne pas réduire les interactions à la simple co-présence des interlocuteurs (Nassau 2016). Il a donc utilisé le cadre de l'analyse sociolinguistique des interactions « dans le contexte de l'entretien de conseil » (*Ibid.* : 53) afin de saisir la complexité et l'intrication des éléments qui s'influencent. Il a

notamment cherché à décrire la réalité sociale des situations de communication qu'il a analysées ; en pointant les aspects professionnels du genre de discours *entretien de conseil*, et inversement en pointant la part langagière de cette activité de travail, et en examinant le rôle de l'identité des participants, de leur relation ou encore de leur histoire conversationnelle. En résumé, l'approche sociolinguistique des interactions verbales a permis à Nassau (2016) de réaliser une analyse de l'expression des émotions et de mettre au jour des émotions appropriées ou non à la situation de communication ainsi que leurs effets sur l'apprentissage de la langue cible.

Anouchka Divoux, également dans le cadre de sa thèse, mène une analyse sociolinguistique de la question au sein de réunions de travail (Divoux 2017, 2018). Elle montre que les questions, selon les formes qu'elles prennent, accomplissent différentes activités interactionnelles, en plus de celle qui vise à obtenir des informations. Elle précise que ces procédés sont accentués, permis et légitimés par des éléments de la situation de communication et notamment par les caractéristiques spécifiques des situations de travail. Je pourrais schématiser une partie du travail de thèse d'Anouchka Divoux de la façon suivante :

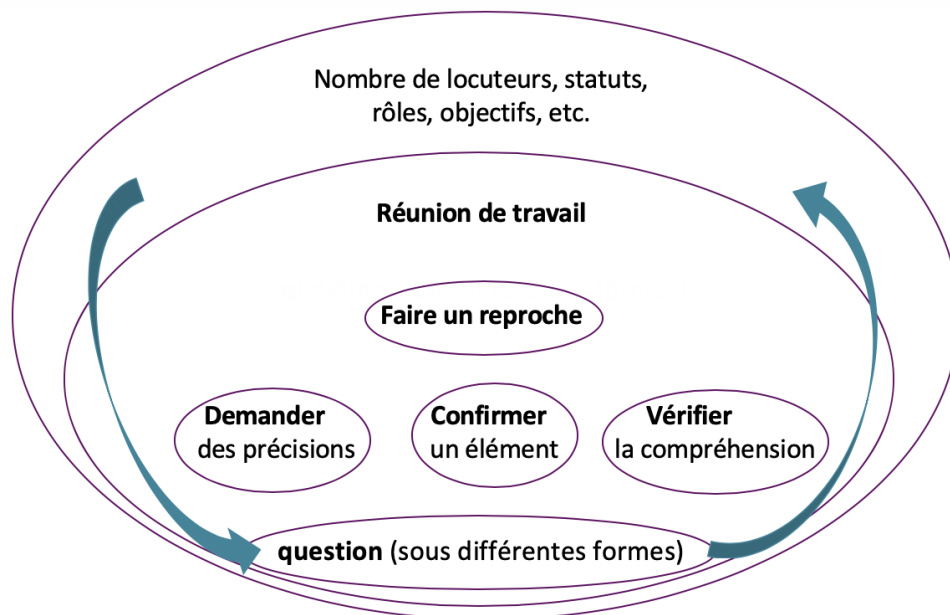


Figure 3 : Illustration du schéma avec une partie des analyses de la question

Ces analyses s'inscrivent dans le cadre de la sociolinguistique des interactions verbales. Ce schéma met en exergue plusieurs phénomènes, tels que :

- les pratiques interactionnelles utilisées comme ressources par les locuteurs pour poser des questions et obtenir des réponses de la part de leurs interlocuteurs sont variées et prennent différentes formes ;
- ces pratiques peuvent accomplir différentes activités interactionnelles et viser différentes valeurs pragmatiques en fonction de leurs formes ;
- une même pratique interactionnelle peut accomplir plusieurs activités interactionnelles en même temps, c'est-à-dire qu'une question, selon sa forme peut cumuler différentes activités (demander des précisions et vérifier la compréhension ou demander des précisions et confirmer un élément ou encore demander des précisions et faire un reproche) ;
- ces pratiques et les activités interactionnelles sont influencées par le genre de discours *réunion de travail* ;
- certaines pratiques accomplissent plusieurs activités en même temps parce qu'elles sont produites par des locuteurs ayant des statuts hiérarchiques et des rôles conversationnels particuliers.

Cette liste n'est pas exhaustive, en analysant les questions dans les interactions, le travail d'Anouchka Divoux exploite et met à l'épreuve le modèle sociolinguistique des interactions verbales. Elle identifie des phénomènes particulièrement intéressants.

Au-delà de ces travaux approfondis de thèse, plus ponctuellement, différentes études s'inscrivent également dans ce cadre théorique qui permet d'analyser les interactions verbales. Ce cadre sociolinguistique et le schéma qui l'illustre sont à la fois simples et complexes. Ils permettent de visualiser et de comprendre rapidement les enjeux des interactions verbales et leur forte intrication à tout un ensemble de facteurs extérieurs. Ainsi, ce cadre complexe, qui peut être abordé de façon schématique, est présenté, expliqué et utilisé dans mes différentes situations de travail et d'enseignement : des étudiants de première année de Licence de Sciences du langage aux étudiants en thèse en passant par des étudiants en orthophonie ou en sociologie et des enseignants de langue. Tous ces publics auxquels j'ai présenté mon cadre et ses applications ont pu, à leur niveau, l'exploiter pour saisir ce qu'il se passe lorsque des locuteurs se rencontrent et interagissent. Par exemple, plusieurs travaux de master s'inscrivent dans ce cadre et plusieurs enseignants de langue s'en inspirent pour décrire des interactions à des apprenants. En outre, l'objectif de ce travail est d'exploiter les résultats des recherches réalisées dans ce cadre à des fins

didactiques, j'y reviendrai à de nombreuses reprises et plus précisément dans les deux derniers chapitres.

Enfin, ce cadre d'analyse est déployé et mis à l'épreuve dans le projet Psyphine¹¹ qui cherche notamment à étudier les interactions entre les hommes et des objets robotisés. Ce projet a vu le jour en 2013 et ses membres m'ont demandé de l'intégrer en 2015. Ce groupe de recherche vise à confronter et articuler les apports de différentes disciplines (psychologie, philosophie, informatique, neurosciences, intelligence artificielle, robotique, sociolinguistique) à la question de l'attribution d'intentionnalité, d'intelligence, de cognition voire d'émotions, à des dispositifs artificiels (des robots, par exemple). Il s'agit d'appréhender les questions naturellement posées par l'interaction homme/objet, à savoir celles liées à l'interprétation du comportement de l'objet jusqu'à la confiance qui peut lui être ou non accordée. Plus précisément, le groupe se demande quels sont les types d'interactions que les hommes mettent en œuvre avec des objets robotisés non humanoïdes. Dans ce cadre, l'analyse sociolinguistique des interactions peut apporter une contribution non négligeable. L'objet robotisé qui permet au groupe de réfléchir à ces questions est une lampe, voici sa photo :

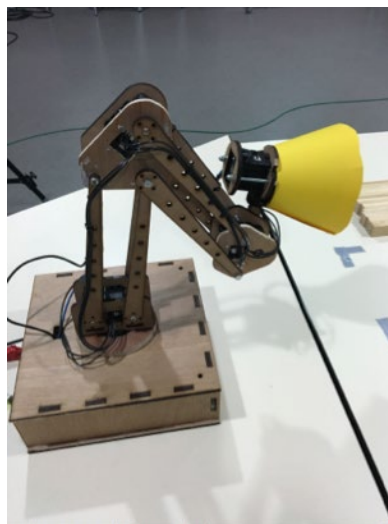


Figure 4 : La lampe robotisée Psyphine

Plusieurs expérimentations ont eu lieu afin de tenter de répondre aux problématiques du projet (André, Boniface à paraître 2019 ; Becker, André, Dutech 2019). La dernière expérimentation, intitulée DECIDE, permet d'analyser les interactions entre deux

¹¹ Voir la présentation sur le site de la MSH Lorraine : <http://www.msh-lorraine.fr/index.php?id=846>

participants face à la lampe, entre les participants et la lampe, ainsi que de recueillir des verbalisations sur les expérimentations, au moyen de questionnaires, d'entretiens et d'autoconfrontations (Clot *et al.* 2000). Au-delà d'un apport méthodologique (mise en place des expérimentations, réalisation des autoconfrontations, analyse qualitative des données), je tente d'éclaircir la question de la nature des interactions qui prennent place entre un humain et un objet robotisé. Pour cela, comme je l'ai précisé, mon cadre théorique et d'analyse, développé pour appréhender les interactions verbales, est mis à l'épreuve et tente de comprendre les interactions. Cette mise à l'épreuve peut paraître éloignée des préoccupations sociolinguistiques. Cependant, en cherchant à examiner les possibles similitudes entre les interactions entre humains et les interactions avec des machines, le cadre est nécessairement affiné et peaufiné puisque de nouvelles questions émergent.

Par exemple, lors de l'expérience DECIDE, j'ai pu analyser les pratiques interactionnelles utilisées par les participants face à la lampe (André, Boniface 2018). Ces dernières accomplissent des activités interactionnelles et nous renseignent sur l'attribution d'intentions à cet objet ou sur sa prise en compte en tant que partenaire communicatif. Les pratiques interactionnelles réalisées sont étroitement liées à la situation de communication, à la fois déterminée par le cadre expérimental et par les éléments non calibrés, tels que l'identité des participants, leur relation, le lieu de l'expérience ou encore les connaissances des participants quant aux technologies informatiques et robotiques. La prise en compte de ces éléments permet de saisir le comportement interactionnel des locuteurs et de tenter de définir probablement un nouveau genre de discours. Ces questionnements peuvent être schématisés. Voici une représentation possible pour analyser un énoncé produit par l'un des participants et adressé à la lampe : *tu fais la tête + pourquoi tu boudes + tu boudes*¹² :

¹² Les + symbolisent des pauses.

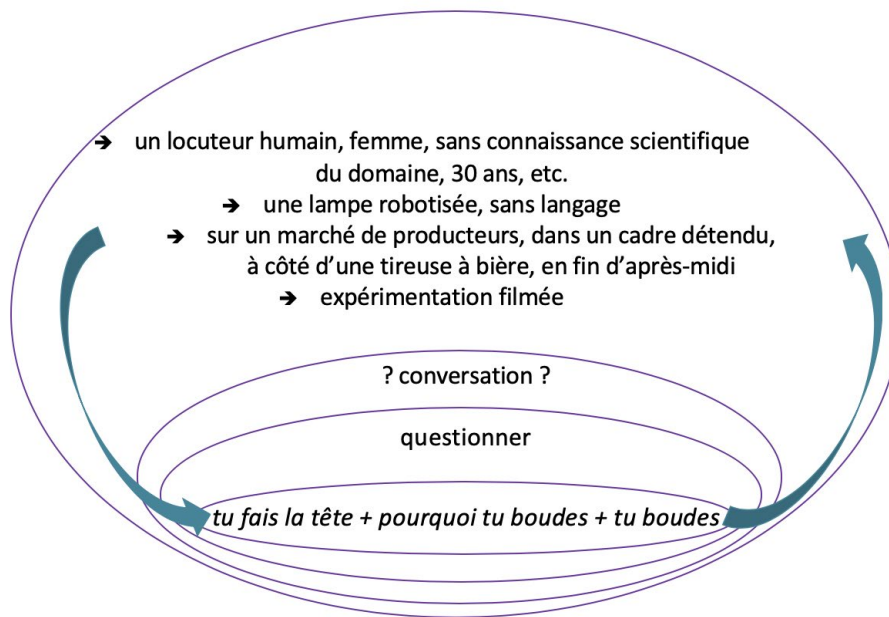


Figure 5 : Illustration du schéma avec l'analyse d'une interaction homme/lampe robotisée

La prise en compte des éléments de la situation de communication ainsi que les activités interactionnelles réalisées nous aident à identifier ce nouveau genre de discours et, par conséquent, la nature des interactions qui prennent place entre des humains et des machines. Les multiples influencent identifiées lors d'interactions verbales existent également lors de ces interactions spécifiques. Ce processus d'analyse se révèle utile pour saisir les conditions d'interaction, les types d'interaction, ainsi que l'acceptation des machines par les humains. De nombreuses études pluridisciplinaires sont en cours actuellement mais rares sont celles qui font appel à la sociolinguistique.

Par ailleurs, l'analyse de ces expérimentations révèle l'importance des deux derniers cercles du schéma, celui de la situation de communication mais surtout celui du genre de discours. Lors de la toute dernière série d'expérimentations, dans une médiathèque, j'ai pu identifier que les participants étaient à la recherche d'une situation et d'un genre connus. Comme la figure précédente le montre, le genre de discours n'est pas facilement identifiable, les participants sont alors déstabilisés et cherchent à se référer à un genre connu dans une situation connue. Lors des autoconfrontations, certains locuteurs nous ont expliqué qu'ils avaient comparé cette situation, face à l'objet robotisé, à une situation face à leur chat ou à d'autres animaux. En réalité, les interactions auxquelles les chercheurs leur demandent de participer ne correspondent à rien d'habituel ni de connu, malgré la présence de plus en plus importante d'objets robotisés ou de logiciels dans le quotidien. En d'autres

termes, il manque aux participants des éléments du schéma pour savoir comment se comporter dans la situation dans laquelle ils se trouvent. Dans toute situation de communication, les interlocuteurs connaissent généralement suffisamment d'éléments de la situation pour s'y ajuster et interagir de façon appropriée. De plus, ils savent dans quel genre de discours ils s'engagent et ont donc des attentes en ce qui concerne le déroulement de l'interaction, sur les mécanismes et l'alternance des tours de parole, sur l'objectif communicatif, sur les thèmes à aborder, etc. Dans les interactions hommes/robots, un certain nombre de ces éléments reste encore indéfini. Ces manques expliquent en partie pourquoi les humains ne savent pas encore tout à fait interagir avec des robots. Par la même occasion, ces analyses mettent en évidence l'importance des éléments contextuels et du genre de discours pour mettre en œuvre une véritable compétence de communication et pour interagir de façon satisfaisante.

1.2. Pourquoi une approche sociolinguistique des interactions verbales ?

L'importance donnée à la situation de communication donne son nom à mon approche « sociolinguistique » des interactions verbales. En effet, le caractère situé des interactions prime sur toutes approches uniquement formelles, descriptives et analytiques, sans les rejeter. En d'autres termes et selon mon point de vue, il n'existe pas de linguistique qui ne soit pas de la sociolinguistique. Cette affirmation a déjà été posée, discutée et interrogée. Dès le milieu du 20^{ème} siècle, même si ses écrits ont été ignorés, Cohen (1956) avait déjà une vision sociologique du langage, dans laquelle la sociolinguistique se reconnaîtra. Marcellesi, à la suite de Labov (1976), posait déjà cette question : « si la langue est chose éminemment sociale, n'est-on pas en droit d'estimer qu'il n'y a pas de véritable *linguistique* sans *sociolinguistique* et que de ce fait la *sociolinguistique* est la linguistique véritable » (1980 : 9). Décrire et analyser une langue sans tenir compte de la situation dans laquelle elle est actualisée n'aurait pas de sens. Il s'agirait de faire une abstraction d'un objet qui n'a pas d'existence en dehors de ses usagers, de leur subjectivité et de leur intentionnalité. Ce postulat fort amène à repenser le découplage des niveaux de description de la langue, tels qu'ils sont par exemple enseignés à l'université¹³. Je ne nie pas

¹³ Par exemple, l'offre de formation de la Licence Sciences du langage de l'Université de Lorraine propose en première année des enseignements intitulés : phonétique/phonologie, morphologie, lexique, grammaire, syntaxe.

l'utilité d'un tel cloisonnement afin de décrire les phénomènes langagiers et de comprendre la structure de la langue. Néanmoins, il est indispensable de (re)complexifier ces descriptions en les confrontant aux éléments de la situation de communication sans quoi aucune analyse n'est valable voire possible. Je m'associe donc à Guespin et Marcellesi (1986 : 26) : « De la phonologie à la stylistique, toute étude des pratiques langagières doit nécessairement aujourd'hui croiser paramètres langagiers et paramètres sociaux. ».

Je me place ici en opposition frontale avec les positions de Noam Chomsky, telles qu'elles ont été présentées encore assez récemment lors d'une conférence à Genève, en 2013¹⁴. A l'issue d'une conférence plutôt intéressante sur la grammaire générative, ses origines et ses évolutions, une personne du public a posé une question intrigante. Je la reproduis en substance : « M. Chomsky, vous qui vous intéressez tellement et tellement bien à la société dans vos écrits politiques¹⁵, comment pouvez-vous ne pas tenir compte de la société quand vous analysez la langue ? ». Cette question si évidente a reçu une réponse tout aussi évidente aux yeux du conférencier : « une langue est un outil, c'est comme un marteau, qu'il soit utilisé pour enfoncer le crane de quelqu'un ou enfoncer un clou dans un mur, c'est la même chose, rien ne change ». Ce bref échange a été à la fois surprenant et éclairant sur les objectifs d'une linguistique qui s'intéresse à la langue comme objet mental. Cet éclairage trouve facilement son opposition avec le célèbre exemple donné par Bloomfield (1933/1970) concernant notamment l'énoncé « j'ai faim ». Cet énoncé prononcé par une personne dans un environnement où rien n'est comestible et ce même énoncé prononcé par quelqu'un, accompagné d'une autre personne, suffisamment grande pour attraper une pomme qui se trouve dans un arbre à côté d'eux, n'ont rien de commun, mis à part leur forme linguistique. Cette forme est décrite de la façon suivante par Bloomfield (1933/1970 : 35) : « Les cordes vocales, la langue, les lèvres, etc. du locuteur contrarient l'émission d'air de telle sorte qu'elles produisent des ondes sonores ; ces ondes se propagent dans l'air et frappent les tympons de l'auditeur, qui alors vibrent à l'unisson ». Il souligne l'importance de l'utilisation du langage et précise que seules les conditions d'un événement de parole permettent de saisir les objectifs des locuteurs ainsi que l'organisation

¹⁴ Cette conférence intitulée « Qu'est-ce que le langage et en quoi est-ce important ? Un point de vue personnel » a été donnée à l'occasion du 19^e Congrès International des Linguistes (CIL) le 25 juillet 2013 à l'université de Genève.

¹⁵ Il est fait référence ici aux positions anarcho-libertaires de Noam Chomsky ainsi qu'à ses activités militantes menées depuis les années 1960.

de la société. Un exemple similaire est fourni par Riley (2001 : 91) à propos de l'identité sociale des locuteurs :

quand un vendeur au marché m'interpelle « Arrêtez, monsieur ! Venez par ici ! », j'ai le choix. Mais quand c'est un gendarme avec son radar au bord de la route, dont l'uniforme et l'arme de service symbolisent et manifestent son pouvoir considérable, il vaut mieux obtempérer.

La langue est bien un outil, comme le disait Chomsky à Genève, mais un outil utilisé, donc tout change. Voilà ici, selon moi, l'intérêt de la sociolinguistique, d'une linguistique qui met en corrélation l'analyse des formes et des aspects *sociaux*.

Les travaux de Dell Hymes ont largement contribué au développement de la sociolinguistique, en tant que discipline qui permet de comprendre et d'analyser la variation des pratiques communicatives qui prennent place nécessairement dans une situation de communication particulière, que l'on peut décrire de façon ethnographique avec le modèle SPEAKING (Hymes 1972). L'approche sociolinguistique proposée permet de saisir les influences des éléments situationnels sur les productions verbales des locuteurs. Ce cadre s'oppose lui aussi fermement à celui de Chomsky pour qui l'objet de la linguistique est le locuteur-auditeur idéal qui appartient à une communauté parfaitement homogène, qui maîtrise parfaitement sa langue, qui fait preuve d'une performance en lien avec cette maîtrise, sans influence socioculturelle. Hymes posera ainsi les bases d'une linguistique socialement ancrée, d'une sociolinguistique, sur laquelle mon approche s'appuie grandement.

En France, la crise de la linguistique est expliquée par Marcellesi :

C'est que la linguistique formelle, après avoir rendu compte d'un grand nombre de faits linguistiques, se trouvait incapable d'intégrer de manière satisfaisante la variation et en même temps ne répondait pas réellement à des questions qu'elle refusait de se poser mais que la vie lui posait, celles de la place et du rôle des phénomènes langagiers dans la société : le cumul de ces deux incapacités, objectées en même temps ou séparément, a conduit à la mise en cause de la linguistique formelle et au succès de ce que les optimistes appelleront une discipline et les pessimistes une étiquette : *la sociolinguistique*. (1980 : 4).

La linguistique ne souhaite pas s'intéresser aux éléments extralinguistiques, considérés comme des objets qui ne concernent pas leur discipline, mais se restreint à l'analyse du *système*. Cette position fut largement adoptée, renforcée par la grammaire générative de Chomsky, après que Saussure a pourtant posé le caractère éminemment social de la langue.

Il serait sans doute plus pertinent de trouver un préfixe encore plus englobant à « sociolinguistique ». Par commodité, le préfixe *socio-* est préféré mais il est chargé de bien plus que des éléments *uniquement* sociaux. Il comporte également des éléments psychologiques, économiques, politiques, philosophiques ou encore épistémiques. Ces éléments sont inextricables d'une situation de communication. En outre, la réflexion exposée par Renaud (2004) concernant les orthographes en débat, *socio-linguistique*, *(socio)linguistique* ou *sociolinguistique*, semble mettre en lumière des approches et des conceptions différentes de la langue et des pratiques langagières et interactionnelles, ainsi que de la façon de les analyser dans leur contexte social. Il précise que

l'analyse sociolinguistique, sans tiret, est celle d'une activité langagière considérée comme constitutive de pratiques qui la débordent tout en intégrant l'objet d'étude : elle contraint le linguiste à développer une linguistique de formes « situées », prenant sens dans la temporalité même et dans la logique d'activités qui contribuent à les accomplir. » (*Ibid* : 96-97).

Cette précision orthographe-terminologique permet d'identifier plusieurs orientations théoriques possibles et plusieurs façons d'envisager les liens entre la langue et son contexte. D'un côté, la socio-linguistique (avec un tiret) s'intéresserait à ce qui préexiste à l'interaction verbale et la sociolinguistique (sans tiret) analyserait ce qui se construit au fur et à mesure du déroulement de l'interaction, aux pratiques langagières mises en action par les participants. Les deux approches ne s'excluent pas. S'il semble que la sociolinguistique soit la voie d'accès privilégiée à la compréhension des pratiques interactionnelles mises en œuvre pour atteindre des activités interactionnelles dans des situations de communication, il semble également que des éléments extralinguistiques préexistent et font varier les pratiques interactionnelles (Gadet 2003a).

Ce vieux débat, qui concerne notamment l'objet d'étude que représente la langue, pourrait paraître ringard ou dépassé. Cependant, le quotidien des chercheurs en « Sciences du langage » (pour trouver un terme non discuté, quoi que...) est encore aujourd'hui fait de cette dichotomie entre linguistique et sociolinguistique. Encrevé (1977) dans sa présentation du numéro 34 de la revue *Langue française* intitulé *Linguistique et sociolinguistique*, soulignait déjà que ce débat pourrait être considéré comme dépassé alors que les travaux de l'époque montraient encore de quelle façon la sociolinguistique devait sans cesse « combattre pour être reçue à l'intérieur du champ, et réfuter les arguments avec lesquelles on la refoule, tout en maintenant ses propres critiques. » (p.4). Aujourd'hui

encore, il suffit d'analyser les discours des enseignants-chercheurs et chercheurs en Sciences du langage pour constater le jugement sévère des *linguistes*, qui ne s'intéressent qu'aux formes, et le mépris latent pour des recherches qui ne s'intéressent pas exclusivement à la langue mais aux autres facteurs qui pourraient l'influencer. Le reproche d'une absence de purisme scientifique est encore présent aujourd'hui. Là encore, Encrevé explique ces postures, qui perdurent plus de quarante ans plus tard :

La polémique et la différence de pratique portent d'abord en effet, ceci est bien connu, sur la question des *données* et sur la méthodologie de l'enquête. La sociolinguistique veut étudier la langue utilisée dans la communauté linguistique, la langue dans son usage quotidien, et sa méthode consiste à enregistrer les locuteurs dans les situations ordinaires des échanges linguistiques. (...) A l'opposé, la linguistique générative et transformationnelle récuse les données observables de la parole qu'elle tient pour dégradées, masquant les règles grammaticales par l'intervention d'erreurs de production ; les seules données sur lesquelles doit s'élaborer la grammaire, ce sont les jugements d'acceptabilité-grammaticalité portés par le locuteur sur des exemples construits, jugements dont on affirme que c'est l'intuition linguistique qui les rend, donnant ainsi un accès qualifié de direct à la compétence, la langue, du *sujet-juge*. (*Ibid.* : 4-5)¹⁶.

Cette polémique se perpétue dans le champ scientifique de l'analyse de la langue. Dans sa critique de la grammaire générative de Chomsky, Encrevé (*Ibid.* : 7) souligne que la langue que cette approche analyse est une langue « loin du réel aussi, ce qui est plus grave »¹⁷. En effet, cette analyse d'une langue construite par intuition, susceptible de ne pas exister, semble assez saugrenue.

Pourtant, ce qui ne relève pas de l'analyse du système peut encore aujourd'hui être considéré comme une sous discipline ou comme une application vulgaire, c'est-à-dire sans valeur scientifique. La sociolinguistique est encore regardée et décrite comme une « branche de la linguistique » alors qu'elle semble être la discipline capable d'analyser la langue. De plus, la sociolinguistique met en lien la langue et ses utilisateurs dans différents contextes, elle permet donc d'aborder différents objets d'étude ; de l'analyse des usages à l'enseignement/apprentissage de ces usages en passant par l'analyse de l'enjeu de ces usages, en situation de travail par exemple. Ainsi, analyser la langue, en tant que pratique sociale, non seulement permet de saisir le fonctionnement des interactions humaines mais offre également la possibilité d'appréhender d'autres pratiques sociales, comme les

¹⁶ Les italiques se trouvent dans le texte original.

¹⁷ Encrevé souligne juste avant cet extrait que la langue étudiée par Chomsky est « assez loin de la « langue » saussurienne » (*Ibid.* : 7), d'où la présence de « aussi » dans cette citation.

pratiques éducatives ou professionnelles. De plus, l'analyse sociolinguistique des interactions verbales gagne à se nourrir d'autres disciplines qui peuvent occasionner un regard complémentaire sur la langue. Cette démarche est à l'opposé de celle préconisée par les puristes de la langue, pourtant elle est riche d'un point de vue scientifique. Dans le point suivant, je vais analyser les apports conceptuels et analytiques des disciplines qui s'intéressent à l'activité de travail humaine.

1.3. Les apports de l'analyse des situations de travail

Les études qui s'intéressent au travail se sont naturellement penchées sur le langage pour appréhender les situations de travail. Dès les années 1960 aux États-Unis avec les débuts de l'ergonomie puis dans les années 1980 en France, la question du langage a été posée comme un objet d'étude nécessaire. L'évolution du travail et l'augmentation de sa part langagière (Boutet 1995, 2001 ; Borzeix, Fraenkel 2001) ont favorisé la prise en compte de la dimension communicative des situations de travail. La rencontre des disciplines qui analysent le travail et celles qui analysent les interactions verbales se révèle fructueuse à la fois pour les analyses du langage et pour les analyses du travail. De cette collaboration a notamment émergé un « virage discursif de l'analyse du travail » et un « virage actionnel des sciences du langage » (Fillettaz 2004 : 69). En France, le réseau « Langage et Travail » s'est d'ailleurs constitué afin de faire travailler ensemble différentes disciplines et quelques concepts à l'interface des sciences du langage et des sciences du travail (Borzeix 1998). Je m'attarderai peu sur les apports mutuels, désormais bien connus, de chacune des disciplines mais je vais examiner plus précisément l'apport des sciences du travail à mon cadre d'analyse sociolinguistique des interactions verbales.

Le cadre d'analyse présenté dans les points précédents m'a permis tout d'abord d'appréhender les phénomènes langagiers et interactionnels collaboratifs au sein de réunions de travail dans une entreprise chargée de la construction et de l'entretien des routes d'un département français (André 2006a). Les situations de communication au travail sont tout à fait particulières et méritent d'être problématisées dans la mesure où certaines influences et certaines contraintes sont exacerbées (Drew, Heritage 1992). La contextualisation des données pour analyser les interactions verbales est indispensable, quelle que soit la situation, mais elle paraît encore plus essentielle lorsque les interactions

prennent place dans des situations aussi complexes que celles du travail. En outre, les travaux menés avec l'objectif de montrer l'importance et les enjeux du langage au travail ont contribué à faire émerger une conception praxéologique du langage (Filliettaz 2002).

L'analyse des interactions verbales tire profit d'un détour par la compréhension du travail. En effet, comme interagir, travailler, c'est s'inscrire dans une histoire sociale et collective, et c'est également apporter une contribution à la société (Clot 1999). L'analyse du travail ne peut se faire sans une analyse plus globale de la société¹⁸. Travailler n'est pas simplement fournir une tâche ou un service contre une rémunération. La complexité de l'analyse du travail n'est pas sans rappeler la complexité de l'analyse des interactions. Les intrications dans lesquelles se trouvent les actions de travailler et d'interagir sont multiples. C'est pourquoi analyser les interactions en situation de travail pointe les caractéristiques complexes de leurs inscriptions dans une situation sociale particulière. Je ne reviendrai pas sur les enjeux du langage au travail (André 2006a) mais je renvoie tout de même aux travaux pionniers du réseau « Langage et travail » (Borzeix, Fraenkel 2001), déjà mentionnés précédemment. Ils ont permis de saisir le sens des dimensions actionnelles ou praxéologiques du langage et d'orienter mes analyses interactionnelles. Les recherches en sociolinguistique ont introduit la problématique de l'action : « les productions langagières attestées sur les lieux de travail exigent l'abandon d'une vision idéaliste de la communication et la prise en considération de la nature complexe des rapports qui se tissent entre les conduites humaines et les situations d'action dans lesquelles elles prennent sens » (Filliettaz 2004 : 10).

En outre, l'analyse de l'activité de travail rejoint nécessairement l'analyse de l'activité interactionnelle. Les théories de l'activité ainsi que l'analyse clinique de l'activité (Clot 1999) apportent un éclairage intéressant sur la conception de la langue. Celle-ci est une ressource pour travailler, elle est exploitée à des fins de production. Sa réalisation, au même titre que toutes les autres activités professionnelles, permet l'effectuation du travail (Lacoste 2001). En réalité, le cadrage de l'analyse sociolinguistique des interactions verbales est fait de nombreux allers-retours entre différentes conceptualisations et théorisations, du travail et de la langue. Ces allers-retours sont présentés brièvement ci-dessous dans un ordre qui

¹⁸ On comprend aisément pourquoi des philosophes ou des économistes se sont intéressés au travail, c'est notamment le cas de John Locke, Adam Smith, Emile Durkheim, Max Weber ; voir par exemple Marx (1972).

n'est ni chronologique ni hiérarchique. Chacun des apports a autant d'importance et a contribué tout au long de mes travaux de recherches dans ses différentes dimensions.

Parmi les apports de l'analyse du travail à l'analyse des interactions, la distinction entre *travail prescrit* et *travail réel* mérite d'être mentionnée. D'ailleurs, je la réinvestis en distinguant *langue prescrite* et *langue réelle* (voir chapitre 4, point [4.2.](#)) afin de saisir les enjeux d'une transposition didactique de l'analyse des interactions. L'analyse des situations de travail conduit à distinguer le travail prescrit – explicite, normatif, théorisé, défini et organisé à l'avance, par un concepteur, par la hiérarchie, par des procédures et des techniques – et le travail réel – effectif, réalisé par l'opérateur, tacite (c'est-à-dire non décrit), individualisé, personnalisé, adapté aux conditions de travail et aux contraintes de la situation (Leplat, Hoc 1983 ; Hubault 1995 ; Schwartz 2000 ; Lonchamp 2003 ; Clot 1995). Cette distinction implique nécessairement une réinterrogation du travail prescrit et un écart entre celui-ci et le travail réel. Si l'on prend en compte le facteur humain au travail (Dejours 1995), la subjectivité du travailleur ou encore ce que Dejours (2016) appelle « l'intelligence rusée », le travail tel qu'il est envisagé et prescrit par un tiers est nécessairement redéfini. La mise en situation de l'activité de travail est influencée par différents éléments, tels que l'identité des travailleurs, leurs relations, leurs connaissances partagées et distribuées au sein du collectif, l'organisation réelle du travail, la représentation de son travail au sein de la société, etc. Ainsi, le travail est une notion complexe qui prend place dans une situation aux multiples influences et qui est intriquée à une multitude d'éléments. Le salarié actualise sans cesse, dans le travail réel, ses connaissances, ses compétences et ses valeurs afin de réaliser une activité qui évolue au fur et à mesure du développement de son apprentissage professionnel, de son réseau de contacts, de sa compréhension de l'activité et de ses pratiques de travail, constamment redéfinies. Cette conception du travail et de son effectuation a des conséquences sur la définition des compétences.

Ainsi, l'analyse de l'activité de travail et plus précisément l'analyse de la compétence ou des compétences en situation de travail sont également utiles à l'appréhension des interactions verbales et à plus particulièrement à la définition d'une compétence de communication (décrite en termes de compétences socio-interactionnelles dans le chapitre 3, en [3.2.](#)). La notion de compétence est généralement mal définie dans la mesure où elle est surutilisée dans de nombreux domaines, aussi bien scientifiques qu'administratifs ou managériaux. Au travail, Durrieu (2016 : 9) explique que « Définir la compétence ne va pas de soi. Comme point de départ, nous dirons que la compétence s'entend comme *une*

*hypothèse*¹⁹ : une hypothèse sur la façon dont quelqu'un fera face à un problème (au sens d'une question à résoudre), dans une situation donnée ». La compétence est alors nécessairement mise en œuvre par un individu ou un collectif d'individus – en fonction de son identité, son expérience, son histoire – dans une situation particulière. Il ne s'agit donc pas ici – au moins pas exclusivement – d'une définition déshumanisée de la notion de compétence en termes de listes d'actions qu'un individu est censé connaître et être capable de faire, telle que cette définition apparaît aujourd'hui en entreprise mais aussi de l'école à l'université (Dejours 2006). La compétence serait alors la possibilité de se comporter de façon appropriée dans une situation, en répondant à la fois à ses attentes et en s'adaptant à son caractère situé. Cette approche permet de confirmer l'importance de la prise en compte des éléments de la situation de communication par les participants qui interagissent. Néanmoins, aucune activité de travail ne peut se réaliser sans une définition théorique préalable ou un cadre prédéfini. Ce cadre est alors redéfini par la situation, comme le travail prescrit est redéfini par le travail réel. Je mets fortement en lien cette analyse du travail avec l'analyse des interactions. Le genre de discours sert de cadre, créé des attentes et propose un canevas sans lequel il serait impossible d'entrer en interaction (Bakhtine 1984), mais comme l'activité de travail, il est redéfini par les éléments de la situation de communication, tels que les participants, leur identité, leur relation mais aussi par leurs pratiques interactionnelles.

Un locuteur qui interagit de façon appropriée met en œuvre des pratiques interactionnelles en s'appuyant sur l'ensemble de ses connaissances situationnelles. Il investit ses savoirs sociolangagiers en interaction et en fonction de la situation. La compétence au travail, comme la compétence socio-interactionnelle ou communicative, mettent en lien « ce que l'on *sait* et ce que l'on *fait*²⁰ » (Durrive 2016 : 17). En d'autres termes, les locuteurs, comme les travailleurs, actualisent leurs savoirs et savoir-faire en s'appuyant sur les conditions, les contraintes et les différents et nombreux éléments de la situation dans laquelle ils se trouvent. Clot et Faïta (2000), également dans le cadre de l'analyse clinique de l'activité, montrent que le genre professionnel est sans cesse retravaillé ou redéfini par le style, c'est-à-dire par la façon d'agir de l'individu ou du collectif. Ils expliquent que le genre est « réitérable dans chaque situation de travail, il ne prend sa forme achevée que dans les traits particuliers, contingents, uniques et non réitérables qui

¹⁹ En italique dans le texte.

²⁰ En italique dans le texte.

définissent chaque situation vécue » (*Ibid.* : 15). Cette redéfinition du genre avec les éléments de la situation, y compris ses participants, est tout à fait éclairante pour saisir les influences en situations de communication. Ces dernières sont mises en lumière avec les deux flèches du schéma présenté au début de ce chapitre (voir [1.1](#)).

Aux notions de travail prescrit, de travail réel et de compétences, s'ajoute la notion de collectif de travail dans laquelle il est également possible de puiser afin d'éclairer l'analyse des interactions verbales. En effet, l'analyse des collectifs de travail montre d'une part que les interactions en sont, naturellement, le socle et d'autre part que les collectifs construisent un espace d'intersubjectivité et de compréhension réciproque (Zarifian 1996). Cet espace donne une importance à l'agir ensemble ou à la force du collectif. Cette importance donnée à la coopération des travailleurs se rapproche de la coopération interactionnelle entre les locuteurs. Elle permet de mettre l'accent sur la construction collaborative du discours, sur les ajustements mutuels observés en interaction ou encore sur la recherche de la construction d'un discours intelligible par tous les participants. Les sciences cognitives apportent une contribution intéressante aux collectifs à la fois de travailleurs et d'interlocuteurs. Les travaux sur la cognition située et distribuée éclairent la compréhension des collectifs (Hutchins 1990 ; Lonchamp 2003). Ils permettent d'appréhender plus finement la définition des participants aux interactions dans la mesure où ils mettent l'accent sur leurs connaissances, partagées ou non, c'est-à-dire distribuées entre les différents interactants et sur lesquelles ces derniers peuvent s'appuyer pour collaborer et construire leur discours.

L'analyse des collectifs de travail permet d'affiner le regard du sociolinguiste sur le fonctionnement des interactions en général. Comme l'expliquent Borzeix *et al.* (2015 : 3) à propos de la participation au travail, « le monde du travail pose avec force, tel un miroir grossissant, le problème de l'asymétrie des acteurs et des inégalités dans la participation ». La situation de travail exacerbe des influences, notamment à cause des relations hiérarchiques, de la répartition des compétences ou de la cognition distribuée et de la contribution individuelle et collective à un objectif commun, c'est-à-dire à la production d'un bien ou d'un service. Ces situations s'intègrent dans des organisations ou des entreprises particulières, avec des modes de gestion allant de l'autogestion à l'autoritarisme en passant par les différentes formes modernes de management participatif. Ces organisations prennent, elles aussi, place dans une dimension politique et dans des situations allant du collectivisme au capitalisme. Ainsi, ces différents cercles d'influences

pèsent sur les relations de travail, sur les représentations que les travailleurs ont de leur métier, sur les objectifs fixés par l'organisation et par eux-mêmes et sur les interactions verbales qui prennent place au sein même des situations de travail. Ces questions de travail prescrit, de compétences et de mode de gestion des organisations ont été creusées lors d'une étude des conséquences de l'hégémonie de l'anglais en situation de travail (André 2015b). Je reviendrai sur cette étude afin de montrer l'intérêt d'une interface entre sociolinguistique et didactique des langues (chapitre 3, point 3.3.).

Conclusion du chapitre

L'objectif de ce chapitre était de mettre au jour et d'expliquer mon cadre d'analyse sociolinguistique des interactions verbales. Ce cadre est inspiré de différents cadres scientifiques, de différents travaux de recherches, de différentes études interactionnelles mais également de différentes expériences personnelles à la fois d'analyste d'interactions et d'implication sur le terrain, notamment du travail et de la formation.

En résumé, les interactions verbales sont appréhendées dans leurs intrications à la situation de communication. De multiples influences pèsent sur les pratiques interactionnelles des locuteurs lorsqu'ils réalisent leurs activités interactionnelles. Cette situation n'est jamais saisissable dans sa globalité à la fois pour l'analyste et pour les locuteurs. Cependant, lorsque des éléments échappent aux participants, ils peuvent les interroger ou les négocier. L'analyste fait avec ce qu'il a, c'est-à-dire avec les éléments qu'il connaît de la situation et les éléments qui se révèlent au fur et à mesure des interactions. Ces éléments influencent le déroulement des interactions, les activités et les pratiques des locuteurs. Réciproquement, les pratiques et les activités interactionnelles redéfinissent le genre de discours qui proposent un cadre ou un script dans lequel ces dernières s'inscrivent. La situation de communication se trouve elle-même redéfinie et modifiée au fur et à mesure de l'actualisation des pratiques et des activités. En outre, celles-ci sont forcément interactionnelles et co-construites, dans la mesure où parler c'est interagir (Gumperz 1982, Kerbrat-Orecchioni 1990). L'analyse des interactions ne peut se passer de données, que le chercheur doit soigneusement sélectionner, recueillir et traiter.

Chapitre 2 : Des données à analyser, des corpus à constituer

Une analyse sociolinguistique des interactions verbales implique nécessairement le chercheur dans une science de terrain pour recueillir les données qui permettront de mettre au jour le fonctionnement, la complexité et la dynamique des interactions. Le cadre théorique définit dans le premier chapitre impose la mise en place d'une méthodologie de recherche tout à fait particulière qui garantit la compréhension des pratiques interactionnelles intriquées à la situation de communication dans laquelle elles sont actualisées (André 2006b). Cette méthodologie doit être pensée pour être en adéquation avec les différents objets d'étude manipulés ainsi qu'avec la façon de les appréhender. Pour recueillir des données à analyser, il ne s'agit pas d'arriver quelque part, de poser un enregistreur ou une caméra, d'appuyer sur le bouton *on* pour commencer puis sur *off* pour terminer. Un certain nombre de précautions, de contraintes et de démarches sont indispensables. Ces dernières sont inspirées des pratiques d'enquête anthropologiques ou ethnographiques. En outre, elles sont le reflet de la conception que le chercheur a de la langue.

Pour illustrer certaines de ces différentes positions, je vais présenter trois corpus que je coordonne et auxquels je contribue. Je vais les présenter par ordre chronologique du début de leur constitution :

1. TCOF : Traitement de Corpus Oraux en Français
2. FLEURON : Français Langue Étrangère Universitaire – Ressources et Outils Numériques
3. ILO : Insécurités Langagières à l'Oral

Chaque terrain est singulier et demande des ajustements de la part du chercheur pour qu'il saisisse *ce qu'il fait, pourquoi il le fait et comment il le fait* (Mahmoudian 1998 ; Mondada 1998 ; Gadet 2003b). La façon dont les données de ces corpus ont été recueillies et traitées en permet une exploitation à la fois sociolinguistique et didactique. En d'autres termes, les données primaires (le son/le film) et secondaires (la transcription) sont complétées aussi précisément que possible par d'autres données secondaires (les métadonnées) précisant les aspects *sociaux* des interactions.

Chaque corpus a été constitué en tentant de respecter au maximum un ensemble de règles, de procédures et de normes juridiques qui régissent le recueil et le traitement des données sonores et visuelles²¹. Néanmoins, chaque corpus est différent. Si les réflexions menées pour la constitution de chacun d'entre eux bénéficient aux autres, certains aspects varient d'un corpus à l'autre. Aucune méthodologie n'est reproductible à l'identique, elle est nécessairement pensée et adaptée aux participants, aux conditions de recueil, aux situations de communication ainsi qu'aux analyses et aux exploitations que les données vont permettre d'accueillir. Par exemple, l'anonymisation est la première règle qui est annoncée aux participants et elle respectée dans la grande majorité des cas. Elle est considérée comme un impératif éthique. Néanmoins, elle ne résiste pas toujours à la contextualisation nécessaire des données et se trouve mise à mal avec la diffusion de données filmées destinées par exemple à un dispositif d'apprentissage du français, comme c'est le cas pour le corpus FLEURON (même si l'image des participants n'est pas associée à leur nom). La multimodalité, indispensable pour certaines exploitations des corpus, ne peut être conservée si les visages sont floutés. Le caractère situé des interactions verbales demande parfois un traitement particulier. Celui-ci peut rendre utile la levée de l'anonymisation de certaines métadonnées. Dans un tout autre contexte, celui d'une enquête sur le tourisme sexuel en Thaïlande, Roux (2010) explique qu'il a révélé le nom de l'ONG avec laquelle il a travaillé afin de rendre possible la contextualisation des observations et des analyses effectuées. Ainsi, l'intelligibilité et la compréhension de certains phénomènes sociaux et interactionnels demandent parfois de fournir des informations non anonymes, notamment sur les institutions ou organisations au sein desquelles les enquêtes sont menées. Je préciserai les dispositions prises pour chacun des corpus. Rien n'est jamais simple ni totalement reproductible lors de recueils et de traitements de données. Aucun des corpus n'est constitué de données recueillies à l'insu des locuteurs, quelles que soient les situations de communication enregistrées et les exploitations envisagées des données. En outre, tous les participants enregistrés peuvent à tout moment annuler leur accord de diffusion, conformément aux articles du code pénal relatifs au droit à la voix et à l'image et aux recommandations de la Commission National de l'Informatique et des Libertés (CNIL).

Avant de présenter les corpus que je coordonne, je vais rapidement examiner différentes définitions de la notion de « corpus ». Ces dernières varient encore aujourd'hui

²¹ Depuis que des corpus oraux et multimodaux sont constitués, de nombreuses réflexions ont été menées et les pratiques ont évolué, voir notamment Baude (2006).

même si certains points semblent faire consensus. Cette notion de corpus est centrale dans ce travail dans la mesure où elle est utilisée et manipulée lors de l'analyse sociolinguistique des interactions verbales et lors de son exploitation à des fins didactiques. Si sa définition précise est moins importante que ce l'on en fait, le corpus traverse aujourd'hui différentes approches de la langue (voir notamment O'Keeffe et McCarthy 2010 pour une synthèse).

2.1. Un corpus, des corpus, quels corpus ?

La question de la nature et de la définition des corpus a été abordée à de nombreuses reprises et donne lieu à des débats constants. Elle est à la fois simple et complexe. Wynne (2005 : 3) ouvre la préface d'un guide de bonnes pratiques en ce qui concerne la linguistique de corpus avec ces phrases :

A linguistic corpus is a collection of texts which have been selected and brought together so that language can be studied on the computer. Today, corpus linguistics offers some of the most powerful new procedures for the analysis of language, and the impact of this dynamic and expanding sub-discipline is making itself felt in many areas of language study.

Il semble qu'un consensus existe autour de la définition du corpus comme étant une grande collection informatisée de données langagières sélectionnées pour être représentative d'une langue (ou d'une variété de langue). Cette définition est également celle de Sinclair (2005) après avoir discuté de nombreuses notions ou critères associés aux corpus (telles que la représentativité, l'échantillonnage, la taille, *etc.*) :

A corpus is a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research.

Cependant, cette question, malgré les discussions, les débats, l'examen des nuances, les accords, les désaccords ou les compromis, reste à la fois simple et compliquée. Chaque critère et chaque notion abordés, par Sinclair (2005) ou par d'autres, peuvent être rediscutés et ont déjà été rediscutés dans de nombreux travaux (voir par exemple Biber *et al.* 1998, Habert 2000, Hunston 2002, McEnery *et al.* 2006). Je ne souhaite pas retracer ici ces discussions (voir McCarthy et O'Keeffe (2010) pour une perspective historique) mais je

vais simplement mentionner les critères importants dans le cadre de mon travail sur corpus. En outre, je m'intéresserai ici uniquement aux corpus oraux et multimodaux, puisque l'analyse des interactions verbales est mon objet de recherche et que j'essaie de transposer cette analyse pour la didactique de l'oral.

Parmi les nombreuses approches, Williams (2005 : 13) précise que : « La linguistique de corpus est un domaine qui s'intéresse aux textes, aux textes réels, c'est-à-dire produits pour des raisons de communication entre êtres humains et non des productions artificielles produites par introspection des linguistes ». Cette définition est un point de départ. Elle s'applique également aux productions orales, accompagnées de leur transcription. Elle me permet également de faire un lien entre corpus et didactique dans la mesure où l'approche didactique à laquelle je ferai référence plaide pour l'utilisation de textes réels, appelés « documents authentiques ». J'y reviendrai.

La linguistique de corpus a bénéficié des progrès en informatique, notamment en ce qui concerne la gestion, le stockage et la manipulation ou l'interrogation des données. Les corpus oraux ont bien évidemment été des bénéficiaires de taille étant donné les difficultés que représente le traitement des données sonores. Le traitement des corpus écrits est également devenu indéniablement plus aisé. La linguistique de corpus a donc pu prendre son essor avec l'arrivée des ordinateurs personnels performants, avec une capacité de stockage et de calcul permettant la manipulation de données. Plus récemment, les problématiques des corpus multimodaux sont venues s'ajouter à celles des corpus oraux. Pour traiter les données filmées, il faut encore plus d'espace de stockage et plus de puissance.

S'intéresser aux interactions verbales réelles, c'est donc faire de la sociolinguistique de corpus. En effet, les échantillons de langue analysés dans ce cadre sont extraits de situations de communication réelles, qu'elles aient été initiées ou non par un chercheur. Dans ce cadre, un corpus peut être défini par :

- un échantillon de langue ou un ensemble de données interactionnelles ;
- de taille importante ;
- extrait de situations de communication authentiques ;
- choisi ou sélectionné en fonction des besoins d'une étude ou pour accueillir certains types d'études ;
- traité informatiquement ;
- et documenté par des données situationnelles.

Pour simplifier davantage la notion de corpus en sociolinguistique et se focaliser sur l'essentiel, il est possible de ne conserver que le premier et le dernier point : un corpus est un ensemble de données interactionnelles documenté par des données situationnelles (André, Tyne 2014).

Bien entendu, chaque élément peut être discuté, calibré, redéfini. La taille, par exemple, est un sujet important souvent au cœur des préoccupations. Combien d'heures d'enregistrement et combien de mots transcrits doit comporter l'ensemble pour qu'il puisse représenter un corpus de travail ? Pour quoi faire ? Cette question interroge le choix et la sélection des données. Quelles variétés de langue ? Quels genres de discours ? Quels locuteurs ? Quels objectifs communicatifs ? *etc.* En réalité, il existe différents types de corpus, en fonction de la façon dont ils ont été recueillis et en fonction de leurs exploitations. Les corpus TCOF, FLEURON et ILO, ont en commun d'être trois corpus de français parlé en interaction. Ils se distinguent toutefois, notamment par les objectifs d'analyse visés lors de leur constitution et par leurs utilisations, par des chercheurs, des enseignants ou des apprenants. Les différences entre ces corpus m'ont amenée à les distinguer et à ne pas les réunir dans un seul et même corpus.

2.2. TCOF – des interactions verbales authentiques

Le corpus TCOF (Traitement de Corpus Oraux en Français) est constitué d'interactions verbales authentiques de différentes natures. Il se présente en deux parties, une première composée d'interactions entre adultes et enfants et une seconde d'interactions entre adultes. C'est cette seconde partie que je coordonne et c'est à celle-ci uniquement que je vais m'intéresser ici. Le corpus TCOF est librement accessible dans la partie *corpus* du CNRTL²² et dans la partie *corpus oraux* de la plateforme ORTOLANG²³. La partie *adultes* du corpus compte actuellement de 221 enregistrements d'une durée totale de 55 heures et évalué à 700 000 mots transcrits (109 enregistrements sont en cours de traitement).

Ce projet est né de la mise en commun de corpus recueillis par des chercheuses désireuses de partager leurs données (André, Canut 2010). Depuis sa naissance, le corpus

²² <http://cnrtl.fr/corpus/tcof/>

²³ <https://www.ortolang.fr/>

TCOF est donc alimenté par les données de certains membres du projet et par des chercheurs du laboratoire ATILF. Les étudiants en Licence et en Master de Sciences du langage de l'Université de Lorraine, sur le site de Nancy, participent également à l'enrichissement du corpus. Depuis quelques années, les étudiants en Master de Sociologie déposent également les corpus qu'ils constituent dans le cadre de leur recherche. Les données orales sont enregistrées dans des contextes variés. Elles se répartissent en deux catégories et présentent des genres de discours différents :

1. Des interactions naturelles authentiques : conversations (à bâtons rompus, autour d'un repas, d'un café, entre étudiants, en famille, entre amis, *etc.*), réunions de travail (au sein de différentes entreprises privées, de laboratoires de recherche, et d'associations), débats publics (assemblées générales, réunions publiques). Ces interactions auraient eu lieu même si elles n'avaient pas été enregistrées.
2. Des interactions sollicitées : entretiens (enquêtes sociologiques, pédagogiques, sur différentes pratiques sociales et culturelles), récits de vie ou d'expériences (narrations de faits marquants, de maladies, de guerres, de mariages, de naissances, de migrations, de voyages, *etc.*), explications professionnelles (pompiers, serveurs, enseignants, éducateurs, formateurs, ouvriers, *etc.*), explications techniques ou passions (escalade, cinéma, jeux de rôle, moto, boxe, *etc.*). Ces interactions sont déclenchées à l'initiative du chercheur ou de l'étudiant qui enregistre.

Au fur et à mesure de l'enrichissement du corpus, l'échantillonnage est affiné. Le corpus TCOF compte actuellement 330 enregistrements, de quinze minutes environ chacun, transcrits (dont 109 en cours de traitement). Chaque enregistrement est transcrit et aligné texte-son avec le logiciel libre *Transcriber*²⁴. Chaque enregistrement est également traité par deux réviseurs. Ces derniers sont des étudiants ou des ingénieurs d'études, recrutés au sein du projet. J'effectue moi-même une dernière vérification avant la diffusion des données.

Ce corpus est constitué afin d'accueillir des analyses de différentes natures. Tout d'abord, l'objectif de départ, lié aux intérêts scientifiques des membres du projet, était de

²⁴ *Transcriber* est disponible à l'adresse suivante : <http://trans.sourceforge.net/> et une version pour mac a été créée par Matthieu Quignard, disponible à l'adresse suivante : <http://perso.ens-lyon.fr/matthieu.quignard/Transcriber/> (cette version permet également de prendre en compte un nombre non limité de locuteurs qui parlent en même temps). *Transcriber* a été choisi parce qu'il est libre, gratuit et simple d'utilisation.

Faire une étude syntaxique et/ou interactionnelle des productions orales à partir d'un grand nombre de corpus, en particulier :

1. la description linguistique (aspects lexicaux, syntaxiques) et pragmatique des pratiques langagières ;
2. la comparaison (du point de vue de leurs caractéristiques linguistiques) entre des verbalisations d'adultes dans des situations de communication ordinaire avec d'autres adultes (récits, conversation, explication, etc.) et des verbalisations d'adultes s'adressant à de jeunes enfants (moins de 7 ans) ;
3. la comparaison entre les productions linguistiques des adultes et celles des jeunes enfants (mise en lien entre diversité du répertoire linguistique proposé et développement langagier) ;
4. la possibilité de faire des recherches transversales sur la langue, s'appuyant à la fois sur des corpus écrits et sur des corpus oraux. (André, Canut 2010 : 39).

Le corpus TCOF est conçu pour accueillir un grand nombre d'analyses. Les transcriptions sont assez « nues » (voir les conventions listées et expliquées sur le site mentionné en note 22), et respectent l'orthographe standard, sans trucage, selon les recommandations de Blanche-Benveniste et Jeanjean (1987), mais peuvent être enrichies selon les besoins. Les métadonnées sont riches, ce qui facilite les analyses sociolinguistiques. Les formats sont partageables, les fiches de métadonnées ainsi que les transcriptions sont normalisées selon la norme internationale *Text Encoding Initiative (TEI)*. En résumé, ce projet a mené des réflexions importantes en ce qui concerne les aspects méthodologiques, éthiques, juridiques, informatiques (liés à la pérennité recherchée des données et aux traitements des données) et aux modalités de partage des données. Ce projet est également intégré et participe à d'autres projets :

- des projets institutionnels et structurant plus larges ou plus englobants, tels que :
 - l'ANR Corpus ORFEO coordonnée par Jeanne-Marie Debaisieux²⁵
 - la plateforme ORTOLANG
- des projets de recherche centrés sur des aspects spécifiques, tels que :
 - l'alignement semi-automatique de corpus - le logiciel *Jtrans* a été réalisé en collaboration avec le LORIA (Cerisara, Mella, Fohr, 2009)
 - l'interrogation du corpus avec un concordancier aligné texte-son (le logiciel *Jconc* a été réalisé en collaboration avec le LORIA)
 - l'annotation automatique en macrosyntaxe (Benzitoun, Fort, Sagot 2012)
 - l'étude des émotions dans les interactions (Quignard *et al.* 2016)

²⁵ <https://www.projet-orfeo.fr/>

- l'exploitation didactique de corpus (André 2018a)

De plus, le corpus TCOF est utilisé dans le cadre d'études dont j'ignore l'existence mais que j'espère nombreuses. L'objectif de ce corpus est de permettre une description documentée et plus fine de la langue et de ses usages²⁶.

Ce corpus tente de respecter les critères consensuels qui définissent généralement les corpus et qui sont exposés dans le point précédent. Ainsi :

- il tente d'être représentatif du français parlé en interaction,
- il s'enrichit régulièrement et fréquemment afin d'augmenter sa taille,
- il est composé d'enregistrements réalisés dans des situations de communications authentiques,
- ces enregistrements sont sélectionnés pour qu'ils puissent accueillir différents types d'analyse,
- il est traité informatiquement,
- il est documenté par des métadonnées.

Grace à ces éléments, le corpus TCOF peut être utilisé pour une analyse sociolinguistique des interactions verbales ainsi que pour une exploitation à des fins didactiques.

2.3. FLEURON – un corpus du domaine universitaire

Le corpus FLEURON (Français Langue Étrangère Universitaire – Ressources et Outils Numériques) est un corpus multimodal et concerne le domaine universitaire. Il est intégré à un dispositif d'apprentissage du français²⁷ et présente des ressources audios et vidéos qui permettent aux étudiants étrangers qui souhaitent faire un séjour universitaire en France de se préparer, avant leur arrivée et pendant leur séjour, à interagir dans différentes situations (André 2016). Le corpus est constitué de situations de communication authentiques que les étudiants étrangers sont susceptibles de rencontrer dès leur arrivée pour gérer et organiser leur séjour. Le corpus a été élaboré (et ne cesse d'être enrichi) afin d'être un échantillon représentatif des interactions auxquelles les étudiants doivent ou

²⁶ Des extraits du corpus ont par exemple été utilisés pour illustrer *La Grande Grammaire du français* (Abeillé, Godard à paraître).

²⁷ Disponible librement et gratuitement à l'adresse suivante : <https://fleuron.atilf.fr/>.

peuvent participer lors de leur séjour universitaire. Il est également constitué de ressources explicatives sur le fonctionnement de l'université française et de la vie universitaire. Il compte actuellement 160 enregistrements dont la durée est comprise entre 30 secondes et 15 minutes (30 enregistrements sont en cours de traitement). Concrètement, ce corpus est composé de trois types de ressources²⁸ :

1. Des interactions naturelles authentiques

- entre étudiants et personnels administratifs de l'université, dans différents lieux : scolarité, secrétariat, bibliothèque, maison de l'étudiant, loge de l'appariteur, service des relations internationales, service délocalisé de la préfecture, centre de langues, service d'aide aux logements, service santé, service d'assistance sociale et médicale, service informatique, associations étudiantes, *etc.* ;
- entre étudiants et enseignants : choix de cours, constitution d'emplois du temps, *etc.* ;
- des conversations entre étudiants français et/ou étrangers ;
- des interactions de la vie quotidienne en dehors du campus universitaire : à la banque, à la gare, avec des voisins, à la piscine municipale, au cinéma, au musée, *etc.*

2. Des témoignages

- d'étudiants français et d'étudiants étrangers, des narrations (modalités et qualité de l'intégration, chocs culturels, *etc.*) ;
- de situations difficiles à filmer (le déroulement d'une visite médicale, d'un entretien à la préfecture, *etc.*).

3. Des explications sollicitées

- du fonctionnement de la vie universitaire : de la bibliothèque, de la carte du restaurant universitaire, des horaires d'ouverture, de la sécurité sociale, des cours informatiques, des aides au logement, des écoles doctorales, *etc.*
- du fonctionnement de certains aspects de la vie en dehors du campus mais utiles aux étudiants : la location des vélos de la ville, obtenir une carte de réduction dans des lieux culturels, comprendre les panneaux dans les gares, acheter un billet de train, *etc.*

²⁸ Il existe quelques ressources audios dans le corpus FLEURON. Néanmoins, toutes les nouvelles ressources sont des vidéos, lorsque les situations le permettent.

Les situations concernent le domaine universitaire mais sont exclus les cours et les contenus disciplinaires. D'autres projets de recherche constituent des corpus de cours à des fins didactiques (voir par exemple Mangiante, Parpette 2004, 2011). Le corpus FLEURON est utilisé à des fins d'apprentissage de compétences socio-interactionnelles dans des situations de communication en dehors des cours. Je présenterai ce dispositif et l'exploitation du corpus dans les chapitres 4 et 5.

Le corpus FLEURON est constitué d'interactions entre des locuteurs natifs mais également de quelques interactions exolingues avec ce que nous avons appelé des « étrangers compétents » (André, Castillo 2005). Cette notion désigne des locuteurs non natifs mais qui réussissent à interagir de façon appropriée à la situation de communication à laquelle ils participent. Cette notion a été développée au sein du GREFSOC (groupe de réflexion sociolinguistique) qui s'est notamment intéressé aux relations de service (André 2008) et aux vertus communicatives (Cameron 1995 ; Marui *et al.* 1996) attendues dans ces situations, très culturellement marquées. L'étranger compétent remplace également, dans le domaine de la didactique des langues, le traditionnel modèle du locuteur natif. L'objectif pour un apprenant n'est pas de se transformer en natif mais d'apprendre à interagir en réussissant à s'appuyer sur les éléments pertinents de la situation de communication.

Le corpus FLEURON a bénéficié des réflexions menées dans le projet TCOF, notamment en ce qui concerne les aspects éthiques et juridiques, les modalités de recueil des données, la transcription et le traitement de données. Néanmoins, ce corpus présente quelques spécificités. Tout d'abord, toutes les situations, dans la mesure du possible, sont filmées. Tous les participants donnent leur accord pour participer aux captations. Ils signent une autorisation de diffusion de leur voix et de leur image dans le cadre du dispositif FLEURON. La captation vidéo fait l'objet de réflexions à chaque nouvelle situation (emplacement de la caméra, présentation et informations sur le projet, droit de rétractation, *etc.*). Ensuite, la transcription du corpus est également « nue » et respecte l'orthographe standard. Toutefois, elle est simplifiée par rapport à celle de TCOF étant donné qu'elle est avant tout destinée à être lue par des apprenants de français. Toutes les traces de l'élaboration du discours oral sont conservées (hésitations, répétitions, pauses remplies, *etc.*) mais les chevauchements ne sont pas signalés graphiquement. Ce choix s'explique par l'alignement transcription-vidéo. En effet, toutes les vidéos sont sous-titrées. Les sous-titres apparaissent sous la vidéo, il est donc aisé de se rendre compte des parties du discours qui sont prononcées simultanément par plusieurs locuteurs. En outre, l'objectif

didactique prime ici sur le marquage des chevauchements qui pourrait perturber la lecture des sous-titres par des apprenants.

Le corpus FLEURON est accessible ressource par ressource (ainsi que par un concordancier, j’y reviendrai). Les ressources sont classées par catégorie, en voici la liste :

- Démarches administratives
- Démarches à la préfecture
- Étudiants Erasmus
- Aides sociales
- Questions pédagogiques
- Santé
- Explications du système universitaire
- Utiliser les transports
- Vie en dehors du campus
- Étudiants en doctorat

Ces catégories sont évolutives et sont créées en fonction des besoins des utilisateurs du corpus, en l’occurrence des apprenants et des enseignants de FLE. Par exemple, la dernière catégorie de la liste a été créée pour répondre à des demandes de la part de doctorants étrangers à l’Université de Lorraine désireux d’avoir accès à ce type de ressources. De nombreuses interactions du corpus ont été recueillies à la suite de suggestions d’étudiants étrangers, notamment de la part de ceux qui ont participé au projet. Le corpus est sans cesse enrichi, par moi-même ainsi que par les étudiants du Master FLE de Nancy qui choisissent de s’investir dans le projet et par les contractuels que j’encadre (vacataires, ingénieurs d’études et contrat postdoctoral) recrutés dans le projet en fonction de ses différents financements.

Le corpus est également interrogeable par un concordancier, directement intégré au site ²⁹. Ce concordancier unique est aligné texte-son-vidéo. Il permet de faire des requêtes dans les transcriptions, d’afficher les résultats sous la forme de lignes de concordances et d’accéder aux occurrences prononcées dans les audios et les vidéos correspondants. Ce dispositif est incontestablement un outil puissant pour l’apprentissage du FLE, il sera au cœur de ce travail de synthèse dans le dernier chapitre.

²⁹ Le concordancier de FLEURON est présenté dans le chapitre 5, en [5.3.1](#).

2.4. ILO – des locuteurs en insécurité langagière

Le corpus ILO (Insécurités Langagières à l'Oral) est composé d'interactions verbales auxquelles participe au moins un locuteur francophone natif en insécurité langagière. Avec Hervé Adami, nous définissons « l'insécurité langagière comme la difficulté pour un locuteur/scripteur de gérer de façon efficace les interactions verbales dans lesquelles il est engagé, d'un point de vue linguistique, interactionnel, pragmatique et social. » (Adami, André 2014 : 77). Une personne en insécurité langagière à l'écrit se trouve en situation d'illettrisme. Cette difficulté est bien identifiée, elle est même nommée. Elle est également l'origine de plans nationaux de lutte contre l'illettrisme, ce fut la grande cause nationale en 2013, elle a une agence, l'ANLCI (Agence nationale de lutte contre l'illettrisme), les médias publics réalisent des documentaires ou des fictions mettant en scène des personnes ou des salariés qui ne savent ni lire ni écrire³⁰. Si cette difficulté est prise en compte et en charge par différentes instances, ce n'est pas le cas avec les insécurités langagières à l'oral. Ce pendant de l'illettrisme à l'oral n'a pas de mot en français, et à ma connaissance dans aucune autre langue. Ce phénomène innomé est généralement un impensé scientifique, il fait l'objet de très peu d'études. Son existence est même parfois niée, rejetée ou renvoyée à un relativisme langagier et culturel consistant à tolérer les différences sans se préoccuper des conséquences. En d'autres termes, il n'est pas rare (même de la part de sociolinguistes) d'entendre que tout le monde sait parler (sauf cas pathologiques) et que toutes les variétés sont équivalentes. Cependant, ces difficultés peuvent représenter un véritable obstacle à une vie sociale et professionnelle. Je reviendrai sur les insécurités langagières à l'oral dans le chapitre suivant (le point [3.2.](#) est consacré à l'identification des compétences socio-interactionnelles).

A priori, il est difficile d'identifier une personne en insécurité langagière à l'oral avant d'avoir enregistré, transcrit et analysé ses paroles en interaction alors qu'identifier une personne en insécurité langagière à l'écrit, ou en situation d'illettrisme, est plus facile. Pour cela, il suffit de lui faire écrire ou lire quelques lignes et les difficultés sont rapidement observables (voir les photos de productions écrites de personnes en situation d'illettrisme en [5.4.2.1.](#)). Néanmoins, plusieurs indices permettent de faire une hypothèse forte sur les éventuelles difficultés de communication à l'oral. Ces difficultés supposées seront ensuite

³⁰ Je pense notamment au film de Jean-Pierre Améris, intitulé *Illétré* et diffusé sur France Télévision en 2018.

validées ou invalidées à la suite de l'analyse d'interactions verbales. Cela signifie qu'il est impossible d'identifier des difficultés tant que l'on n'a pas le matériel langagier nécessaire à analyser. Dans cette perspective, le corpus ILO est en cours de constitution, il compte actuellement une dizaine d'heures d'enregistrements. Les réflexions qui l'entourent sont nombreuses et très spécifiques. Différents éléments peuvent guider le recueil de données, à différents niveaux.

Tout d'abord, une expérience personnelle m'a orientée vers une certaine catégorie de la population. Cette expérience, qui ressemble plutôt à une anecdote, s'est déroulée il y a longtemps maintenant, alors que j'étais étudiante en thèse. J'ai rencontré un éducateur spécialisé qui s'occupait d'un centre pour adolescents et jeunes adultes en décrochage avec la société. Cet éducateur était parfois amené à rencontrer les parents de ces jeunes pour faire le point sur leur situation. Il m'a contactée pour m'expliquer que lorsqu'il rencontrait les parents, ils ne réussissaient pas à se comprendre. Ils partageaient pourtant la même langue première, le français. Ce témoignage a été, pour moi, la première prise de conscience d'un problème non-dit. De nombreuses questions ont émergé de cette rencontre et de ce constat. Parmi ces questions, une, d'ordre général, a primé : Quelles étaient ces familles avec lesquelles les interactions étaient si compliquées ? Ces familles appartenaient à ce que Schwartz (2011) et d'autres (Bourdieu 1983 ; Grignon, Passeron 1989) appellent les classes populaires, même si cette catégorie est difficile à circonscrire. Ensuite, une caractéristique importante commune à ces familles était leur faible niveau de qualification voire de scolarisation.

J'ai laissé de côté cette problématique pour me consacrer à mes travaux en cours jusqu'à ma rencontre avec Valérie Langbach en 2010, qui collaborait temporairement avec mon groupe de recherche – LTF – pour discuter de projets pédagogiques développés au sein de l'organisme de formation pour lequel elle travaillait. Valérie Langbach était chargée de projets et de développement mais également formatrice dans cet organisme de formation, pour migrants et natifs en situation d'insertion. Elle s'est également trouvée confrontée à des échanges avec des francophones natifs pendant lesquels de nombreux problèmes venaient nuire à l'intercompréhension et au bon déroulement des interactions verbales. Elle mentionnait également les mêmes constats de la part de ses collègues, qu'ils soient formateurs linguistiques ou conseillers en insertion. Fortement questionnée par ces problèmes, que nous qualifions aujourd'hui d'insécurité langagière, Valérie Langbach a entrepris une thèse, sous la direction d'Hervé Adami et moi-même, afin de mesurer et

d'analyser ces problèmes sociolangagiers et interactionnels (Langbach 2014). Une partie de son corpus de thèse est intégré au corpus ILO³¹.

Recueillir des interactions avec des personnes en insécurité langagière consiste en réalité à ouvrir un espace de parole quasiment inexistant. Ces personnes ne sont pas familières des prises de parole, en dehors des conversations quotidiennes, et encore moins lorsqu'elles sont sollicitées par un chercheur et enregistrées ou filmées. De plus, pour ces personnes, leur parole n'a généralement pas une grande valeur sur le « marché linguistique » (Bourdieu 1984). Les locuteurs enregistrés dans le cadre du corpus ILO sont généralement rencontrés dans des associations de quartier ou des organismes de formation au sein desquels ils viennent suivre des ateliers ou des formations, généralement de remise à niveau en français et/ou en mathématiques. Ces personnes sont conscientes de la précarité dans laquelle elles se trouvent. Cette précarité représente souvent un frein au recueil de données langagières dans la mesure où ces locuteurs estiment fréquemment leur parole illégitime. Le travail du chercheur est alors de rendre cette parole légitime et surtout de trouver les moyens de « faire parler » en montrant l'intérêt à s'exprimer, à (se) raconter ou à témoigner. Demazière et Zune (2019) font également état de ces difficultés dans leur recueil de la parole de chômeurs radiés des aides sociales et d'autres enquêtes sociologiques évoquent les difficultés d'enquêter auprès de ce que certains appellent les « acteurs faibles » (Payet *et al.* 2008). Cette catégorie sociologique « propose surtout une perspective analytique : elle postule que la faiblesse n'est pas un état mais un processus et qu'il convient dès lors de comprendre des dynamiques et des processus d'affaiblissement et de renforcement. » (Payet 2011 : 2). Dans le cas du corpus ILO, la faiblesse existe parce les personnes enregistrées sont peu scolarisées, peu dotées en capital culturel (Bourdieu, Passeron 1970) et se trouvent dans une précarité économique. Ces personnes rencontrent des difficultés à interagir de façon satisfaisante dans certaines situations de communication. La réalisation de certaines activités interactionnelles peut être problématique, dans certains contextes. Ces mêmes acteurs ne sont pas en difficulté dans toutes les situations qu'ils rencontrent ou pour accomplir toutes les activités qu'ils ont à réaliser. Les objectifs de la constitution de ce corpus s'inscrivent dans une double perspective : d'une part, il s'agit d'identifier et d'analyser les insécurités langagières et, d'autre part, de trouver des façons de les réduire en proposant des pistes didactiques et des actions de formation.

³¹ L'autre partie est constitué d'interactions entre des conseillers en insertion et des demandeurs d'emploi qualifiés, à des fins de comparaison.

Aussi, dans le cas du corpus ILO, l'intérêt pour les participants eux-mêmes n'est pas extradiscursif, c'est-à-dire que la parole recueillie ne profite pas *directement* aux interlocuteurs. Elle n'est pas utilisée dans le cadre d'une enquête qui pourrait, par exemple, améliorer l'accès à l'emploi ou au logement, mettre en lumière des invisibles auprès d'institutions, présenter des parcours professionnels à des employeurs ou encore faire émerger des solutions liées à des besoins de première nécessité, comme c'est le cas avec certaines enquêtes sociologiques. Les objectifs premiers tels qu'ils sont exposés aux participants relèvent plutôt du recueil et de la conservation de la parole. Dans certains cas, les objectifs secondaires sont liés à la formation et à l'amélioration de la formation linguistique ou à visée professionnelle.

Le corpus ILO compte actuellement trois types d'interactions verbales authentiques :

1. des entretiens d'insertion menés par un conseiller pôle emploi ou un formateur d'un organisme d'insertion ;
2. des interactions naturelles spontanées, recueillies lors de différentes expérimentations didactiques dans des associations et des organismes de formation ;
3. des interactions sollicitées par un chercheur ou par un étudiant ayant pour objectif de recueillir des récits de vie, des expériences marquantes, des passions, des opinions, *etc.*

La constitution de ce corpus est particulièrement complexe, comme je l'ai déjà précisé, dans la mesure où il tente de recueillir la parole d'individus peu habitués à interagir dans les situations de communication proposées. Le premier type d'interactions a l'avantage d'exister en dehors de tout protocole d'enquête. Les entretiens d'insertion ont lieu par nécessité dans le parcours des demandeurs d'emploi qui y participent. Le deuxième type d'interactions, quant à lui, est recueilli par un chercheur ou un formateur linguistique lors d'activités de formation. Le troisième se présente souvent sous la forme d'entretiens ou de conversations entre un chercheur ou un étudiant (du master FLE de l'université de Lorraine) et un locuteur choisi pour son faible niveau de qualification, pour son appartenance aux classes populaires ou parce qu'il est en situation d'illettrisme. Ce choix ne conduit bien évidemment pas nécessairement à des interactions verbales présentant des insécurités langagières, certains locuteurs en faisant preuve (Bourdieu, Passeron 1964) et d'autres non

(Lahire 1995). Des études en psychologie ont toutefois montré de forts liens entre les difficultés à l'écrit et la mise en mots à l'oral (voir par exemple Eme *et al.* 2009).

La façon de présenter l'enquête est importante et s'intègre parfois dans une démarche de formation plus globale pour réduire les effets de la formalité de l'enregistrement, audio ou vidéo. Dans tous les cas, les métadonnées concernant le niveau de scolarisation, le parcours scolaire et professionnel sont primordiales pour saisir les éventuelles insécurités langagières. L'analyse de ce corpus ne peut, elle non plus, se passer d'une approche sociolinguistique pour saisir à la fois les influences préexistantes à l'interaction et les pratiques interactionnelles qui y prennent place.

Comme le corpus FLEURON, le corpus ILO bénéficie des réflexions, notamment technologiques et juridiques, menées pour la constitution et le traitement du corpus TCOF. Les données du corpus ILO sont transcrites avec *Transcriber*, en respectant les mêmes conventions de transcription que celles mises au point dans TCOF. Si le corpus ILO est un corpus spécifique, il est intégralement en cours d'intégration dans TCOF afin de le rendre disponible librement. La recherche par métadonnées sur la plateforme du CNRTL permet de sélectionner des sous corpus et d'extraire le corpus ILO. Néanmoins, avec ce corpus, les spécificités et les difficultés du travail de terrain font davantage l'objet de réflexions que pour les deux corpus précédents.

2.5. Des corpus pour quoi faire ?

Depuis les années 1980, de grands corpus écrits et oraux de français sont constitués dans différents pays européens (notamment en Angleterre, au Québec, en Italie, au Portugal ou en Espagne). Le *British National Corpus* (BNC), avec ses quatre-vingt-dix millions de mots écrits et surtout ses dix millions de mots pour sa partie orale présentant un large échantillon de locuteurs variés interagissant dans des genres de discours différents, représente une sorte d'objectif à atteindre pour les autres langues. Ces nouvelles données ont orienté et orientent encore les méthodes de travail ainsi que les descriptions des langues et de leur fonctionnement en contexte. Je ne m'intéresserai ici uniquement qu'aux données interactionnelles orales et aux corpus oraux et multimodaux de français, tels que ceux présentés précédemment.

En France, le travail de recueil de données orales sous la forme de grands corpus est plus récent, il a notamment été impulsé par les travaux du Groupe Aixois de Recherche en Syntaxe (GARS) sur le français parlé. Aujourd'hui, la constitution de corpus fait partie des préoccupations de plusieurs équipes de recherche françaises et francophones, de grands projets et d'infrastructures. On peut citer, parmi les corpus librement accessibles, le corpus CLAPI³² de Lyon, ESLO³³ d'Orléans, le projet PFC³⁴ ou le corpus suisse OFROM³⁵. La plateforme ORTOLANG pointe vers ces corpus oraux et multimodaux, et bien d'autres. La très grande infrastructure HUMA-NUM facilite le traitement des données numériques. L'objectif principal et prioritaire est souvent la constitution de *grands* corpus échantillonnés afin d'essayer de tendre vers une certaine représentativité de la langue.

Les corpus de français parlé permettent d'aborder de nouveaux domaines d'études, avec de nouvelles méthodologies et de nouveaux outils (voir notamment Richard 2018). L'analyse des interactions et celle de la variation sociolinguistique à partir de ces types de corpus vont particulièrement bénéficier de ces nouvelles données. Néanmoins, le risque est aujourd'hui d'accumuler des données, sans objectif, uniquement *pour faire du chiffre*. En outre, l'utilisation de données constituées par d'autres pose un certain nombre de problèmes. Il n'est pas rare, et il serait presque aisé de le comprendre, que certains chercheurs se « rabattent » sur des corpus constitués par d'autres chercheurs, par manque de temps ou de goût pour les enquêtes de terrain. Cette pratique peut se comprendre pour certaines analyses. Par exemple, des mesures de fréquence d'apparition de certains mots ou l'apparition de certains phénomènes morpho-syntaxiques peuvent et doivent s'effectuer à partir de grands corpus (c'est certainement l'objectif premier d'une mutualisation des données, y compris celles de TCOF). Dans ces cas, les éléments de la situation de communication échappent souvent à l'analyste. Ce dernier peut également faire le choix de ne pas s'y intéresser.

Il en est nécessairement autrement pour l'étude de la variation et des usages de la langue. La compréhension de la complexité des facteurs extralinguistiques qui entrent en jeu dans les échanges langagiers est plus ardue lors de l'utilisation de banques de corpus. L'exemple désormais classique de la réalisation de la négation à l'oral illustre à la fois l'apport des grands corpus et leurs limites. Le caractère facultatif de *ne* dans *ne... pas* est

³² <http://clapi.ish-lyon.cnrs.fr/>

³³ <http://eslo.huma-num.fr/>

³⁴ <https://www.projet-pfc.net/>

³⁵ <http://www11.unine.ch/>

soutenu par des observations et des statistiques effectuées à partir de corpus de plusieurs millions de mots (Gadet 2000). Néanmoins, l'analyse de la présence ou de l'absence de *ne* est fortement liée, au moins en grande partie, au contexte sociolinguistique des interactions. Ainsi, les métadonnées des corpus utilisées pour faire ce type d'étude sont indispensables même si, même lorsqu'elles sont accessibles, les corrélations restent difficiles. Il faudrait analyser chaque occurrence de négation, corrélée chacune aux types de locuteurs, au genre de discours, aux thèmes abordés, aux objectifs de la communication, au degré de formalité ou encore aux activités interactionnelles réalisées, pour viser une analyse sociolinguistique. Toutefois, aussi fines puissent être ces analyses, elles ne permettent pas toujours de donner des explications des phénomènes langagiers. La masse de données dans laquelle les analystes puisent pour chiffrer les occurrences de phénomènes ne donne que peu la possibilité de se pencher sur les facteurs de variation. Et pourtant, cette question de la négation est soulevée régulièrement par des apprenants de français, j'y reviendrai.

La linguistique de corpus permet de renouveler ou « d'étendre » (Deulofeu 2000 : 220) l'analyse distributionnelle. Là encore, l'utilisation de ses propres données ou des données des autres influencent nécessairement au moins le type d'analyses réalisé, sinon leur qualité. Si les corpus, et leurs outillages, facilitent le recours à un co(n)texte large d'une pratique interactionnelle étudiée, ce contexte peut difficilement être saisi sans les connaissances nécessaires à la compréhension des interactions. Parfois, selon les analyses réalisées, les métadonnées donnent suffisamment d'informations pour saisir les liens entre certaines formes et certains genres de discours (Biber *et al* 1998). Cependant, en cohérence avec ce que j'ai présenté dans le chapitre précédent, toute analyse des faits de langue est nécessairement sociolinguistique. Cette position implique une connaissance des données secondaires (extralinguistiques) pour saisir les données primaires (le matériel langagier). Avec la priorité donnée au recueil de grands corpus, je m'interroge, tout comme Gadet (2015 : 33) : « Le sociolinguiste a dès lors à se demander si la pratique du *terrain* est compatible avec le recueil d'une importante masse de données, supposées exploitables par d'autres que ceux qui les ont recueillies. ».

Ces méfiances et ces précautions liées à l'utilisation de grands corpus sont nées ou plutôt ont été confirmées par un énoncé spontané : « ton corpus j'y comprends rien³⁶ ». Ces

³⁶ J'ai exposé cette problématique du partage des données lors du colloque de l'AFLS à Nancy en 2011 dans une communication intitulée : « *Ton corpus, j'y comprends rien !* » : *de la difficulté de partager des données orales*.

paroles ont été prononcées par une collègue spécialiste de macro-syntaxe à qui j'avais donné mon corpus de thèse (enregistrement sonore, transcription et métadonnées) pour que nous réalisions une étude mêlant analyse des interactions et macro-syntaxe. Cette exclamation a été révélatrice des difficultés d'exploiter des données que l'on n'a pas recueillies et surtout pour lesquelles on n'a pas assez de connaissances extralinguistiques. Mon corpus de thèse a été recueilli pendant une enquête de terrain qui a duré un an et qui a nécessité ensuite de nombreux allers-retours entre mes analyses et les informations fournies par les locuteurs pour que je comprenne les interactions que j'avais enregistrées et auxquelles j'avais participé. Le travail d'analyse que nous avons réalisé a pu voir le jour grâce aux longues explications que j'ai pu fournir et qui ne peuvent pas être consignées dans les métadonnées (André, Pira 2006). En effet, même lorsque la situation de communication est bien décrite et que les métadonnées sont précises, certains éléments contextuels sont, au cours de l'interaction, parfois pertinents et parfois pas (Traverso 2016). Il n'est pas exclu que selon les moments de l'interaction, les thèmes abordés, les locuteurs en place, ou d'autres éléments de la situation (Hymes 1972) soient plus saillants que d'autres. L'analyse des interactions doit pouvoir réussir à identifier les éléments pertinents qui entrent dans un jeu d'influences. Néanmoins, il n'est pas toujours aisé pour l'analyste de corpus de repérer ces éléments pertinents.

Toutefois, les corpus de français parlé en interaction permettent d'identifier les usages réels de la langue. Ils offrent la possibilité d'étudier tous les aspects du français – phonétiques, phonologiques, morphologiques, lexicaux, syntaxiques, sémantiques, pragmatiques, discursifs et interactionnels. Ces aspects sont corrélés aux éléments situationnels pour une sociolinguistique de corpus. L'idée n'est pas de défendre l'utilisation des corpus, telle une doxa (Bourdieu 1997), mais de souligner les apports et les précautions à prendre (Cappeau, Gadet 2007). Le sociolinguiste doit être conscient de ce qu'il fait avec quels matériaux interactionnels, comment il peut les utiliser, les comprendre, les interpréter et les analyser.

Enfin, les corpus peuvent être exploités à des fins didactiques. C'est l'objectif principal de ce travail qui cherche à montrer les possibilités et les avantages de l'exploitation des corpus oraux et multimodaux pour enseigner et apprendre une langue. Ces possibilités ne sont encore que rarement discutées et envisagées alors que différents corpus peuvent être utilisés et que l'approche sur corpus peut être mise en œuvre de différentes façons. Deux de ces façons sont facilement identifiables : l'utilisation de grands corpus, créés initialement

pour accueillir des analyses sociolinguistiques, ou l'utilisation de petits corpus, créés spécifiquement pour illustrer des phénomènes précis et atteindre des objectifs didactiques particuliers. Entre ces deux approches, il en existe d'autres qui peuvent être représentées sur un continuum :

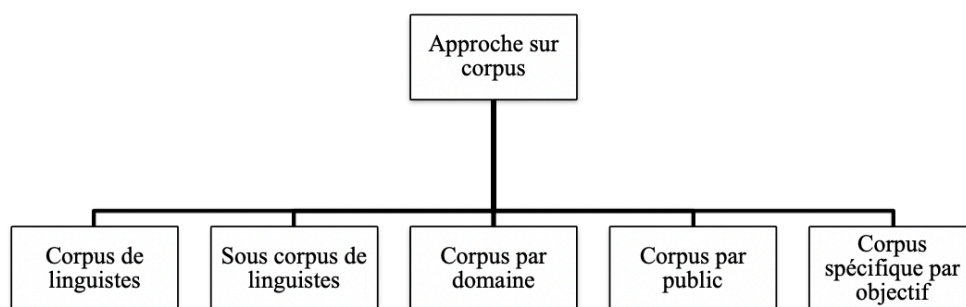


Figure 6 : Différents types de corpus dans l'approche sur corpus (André à paraître 2019)

Les **corpus de linguistes**, créés sans visée didactique, permettent aux enseignants et aux apprenants d'avoir accès à un grand échantillon de langue. Le *data-driven learning* (Johns 1991) a montré comment utiliser ces corpus, notamment de langue anglaise écrite, et leurs contributions dans l'appropriation de la langue. Par exemple, Bernardini (2000) utilise le *British National Corpus (BNC)* avec ses apprenants. J'expliquerai cette approche, pour la didactique de l'oral, dans le dernier chapitre.

Si la masse de données réunies dans ces corpus peut effrayer et déstabiliser les utilisateurs, des **sous-corpus de linguistes** peuvent être constitués. Par exemple, le *BNC Baby* est extrait du BNC. Il compte quatre millions de mots d'anglais écrit et parlé (Burnard 2004). Les critères pour choisir les données de ces sous-corpus sont variés : genre de discours, identité des locuteurs, nombre de locuteurs, thèmes abordés, activités interactionnelles, pratiques interactionnelles, etc. Ces critères sont choisis en fonction des objectifs d'apprentissage. Ces sous-corpus sont parfois didactisés. C'est le cas de CLAPI-FLE³⁷, le sous-corpus de CLAPI destinés aux enseignants de FLE, sélectionné notamment selon les activités langagières et les pratiques langagières présentes dans les enregistrements (audio et vidéo). Je présenterai ce sous-corpus dans le chapitre 4 ainsi que sa complémentarité avec le corpus FLEURON.

³⁷ <http://clapi.icar.cnrs.fr/FLE/accueil.php>

Le corpus FLEURON est un **corpus par domaine**, en l'occurrence universitaire. Il est donc représentatif des situations de communication d'un domaine particulier que les apprenants de la langue sont susceptibles de rencontrer. Johns (1997: 103) avait constitué un tel corpus : « to give a rough reflection of an overseas student's life in Birmingham, both on and off the campus ». Les formations en français langue professionnelle (Mourlhon-Dallies 2008) peuvent être également à l'origine de la constitution de ce type de corpus. Il s'agit par exemple des corpus des métiers de la propreté, du bâtiment, de l'hôtellerie mais aussi des chercheurs en sciences humaines, des négociateurs en vin, *etc.*

Les publics ou les profils d'apprenants contribuent à orienter les méthodologies didactiques. Des **corpus par public** peuvent ainsi être constitués afin de correspondre au mieux aux attentes de ces publics. Par exemple, la formation linguistique des migrants doit accorder une attention particulière au public, à son profil et à ses objectifs (Adami 2009a). Le français langue d'intégration (FLI) vise l'autonomie des apprenants qui doivent réussir à participer aux interactions qui leur permettent d'organiser leur vie dans la société d'accueil. Les formations en FLI s'appuient sur des corpus (Adami, André 2013) qui présentent des données généralement inexistantes dans les corpus précédemment cités : des interactions à l'école, à la mairie, chez le médecin, à la pharmacie, à la préfecture, à la gare, à la poste, *etc.* D'autres ressources peuvent être rassemblées dans ce type de corpus, telles que des annonces sonores (dans les gares, stations de métro, *etc.*), des boîtes vocales (de la CAF, des urgences, *etc.*) des documents administratifs (demande de carte de séjour, de mutuelle, des factures, *etc.*) ou encore des panneaux de signalisation. Ces corpus sont d'autant plus importants qu'ils représentent les interactions, orales et écrites, auxquelles les migrants doivent apprendre et réussir à participer pour assurer leur (sur)vie. Il est possible également de viser un public en fonction de son âge. Par exemple, le projet européen *SACODEYL* de l'université de Tübingen propose des corpus d'adolescents, âgés de 13 à 18 ans, filmés dans leur langue maternelle (allemand, anglais, espagnol, français, italien, lituanien, roumain) pour des apprenants du même âge.

Malgré la forte croissance du nombre de corpus, parfois les données manquent. Certains enseignants constituent alors eux-mêmes leurs corpus de travail malgré l'investissement que cela représente (Gilmore 2007). Ils sont parfois de petites tailles mais sont créés parce qu'ils représentent des **corpus spécifiques par objectif** didactique particulier ou ce que Braun (2007) appelle des *pedagogically relevant corpora* et Meunier et Gouverneur (2009) des *textbook corpora*. Par exemple, des enseignants recueillent des

corpus d'entretiens d'embauche, d'exposés scientifiques, d'interactions à l'accueil de bâtiments publics (Janot 2017), d'annonces sonores publiques, d'explications d'itinéraires, etc. Ces données permettent de travailler, avec des interactions authentiques, la compréhension et la production de pratiques interactionnelles spécifiques.

Conclusion du chapitre

Le travail de constitution de données interactionnelles est complexe. Quel que soit le corpus, plusieurs étapes sont indispensables et font partie d'une chaîne de traitement, schématisée ci-dessous :

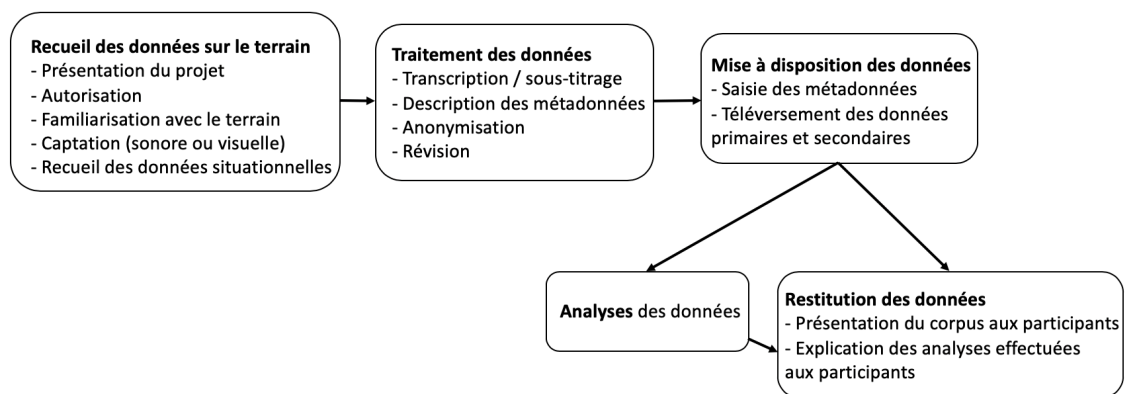


Figure 7 : Chaîne de traitement des corpus de données interactionnelles

Chaque élément de ce schéma peut être discuté, négocié et adapté au terrain, aux participants, aux situations de communication enregistrées, au type d'anonymisation, au mode de diffusion ou aux objectifs de la recherche.

Par exemple, comme je l'ai mentionné précédemment, mon corpus de réunions de travail (André 2006a, 2006b) a nécessité des allers-retours réguliers sur le terrain pendant un an, puis plus espacés, jusqu'à la restitution des données. Cette immersion sur le terrain est parfois nécessaire pour comprendre les données interactionnelles recueillies et pour pouvoir les analyser. Les contacts avec des informateurs permettent également de co-construire l'intelligibilité des données, les incompréhensions ou les doutes peuvent être évoqués avec les participants. Ainsi, à la manière des ethnologues ou des anthropologues, le sociolinguiste doit se familiariser et comprendre l'environnement au sein duquel sont recueillies les données. Dans ce cas, l'analyse des pratiques interactionnelles des

participants ne peut se faire sans connaissance des contextes économique, relationnel ou encore socioculturel. Enquêter devient nécessaire, il faut donc se faire accepter sur le terrain. A ce propos, Céfaï (2003 : 476) souligne que « mener une enquête de terrain, c'est accomplir des opérations très semblables à celle de l'immigrant, mais sur un mode expérimental. Il faut être capable de se mouvoir à la frontière de plusieurs mondes, comme un déclassé ou un déraciné professionnel ». Ainsi, ce travail ethnographique et investigateur permet de saisir plus aisément les interactions entre les locuteurs et, notamment, de saisir les liens entre les pratiques et les activités interactionnelles et la situation de communication. Dans cette perspective, le recueil des données situationnelles peut se faire grâce à un carnet de bord, tel qu'il a pu être défini par Winkin (2001) ou encore par Malinowski (1967) afin de documenter aussi précisément que possible le corpus et décrire les métadonnées. Toutefois, ces métadonnées ne sont pas toujours facilement descriptibles et compréhensibles par ceux qui n'ont pas constitué le corpus et enquêté sur le terrain. La diffusion des données se trouve donc parfois conséquemment problématique voire impossible.

Un autre exemple d'adaptation est l'étape délicate de la transcription qui est fortement liée à l'analyse des données. Les réflexions menées sur les façons de transcrire sont anciennes mais toujours aussi dynamiques et les débats sont toujours aussi vifs. En 2005, à Perpignan, lors de l'école d'été *Transcription de langue parlée. Aspects théoriques, pratiques et technologiques*, un des objectifs, après une semaine de discussions, était de clore la rencontre avec des conventions de transcription communes. Nous n'avons jamais atteint cet objectif, même en optant pour des conventions simples ou « nues », enrichissables selon les besoins analytiques de chacun. Ces conventions communes n'existent toujours pas quinze ans plus tard et l'objectif d'harmoniser les conventions s'est éloigné. D'ailleurs, des « moulinettes » sont créées pour passer de conventions à d'autres, à l'instar des conversions de format de données. Ces réticences sont compréhensibles, la transcription n'est pas un simple accessoire, elle est à la fois le reflet des réflexions théoriques du transcripteur (voir par exemple Ochs 1979 ; Bilger 1999, 2000 ; Mondada 2000, 2008 ; Bilger 2008) et la présentation de certains aspects des interactions verbales (Duranti 1997 ; Racine *et al.* 2011). Les analyses syntaxiques de corpus n'ont pas les mêmes besoins que les analyses interactionnelles. Néanmoins, si un compromis entre fidélité et lisibilité des données fait partie des principes qui régissent la transcription, sa granularité reste en question. Par souci de fidélité, tout ce qui est prononcé est transcrit et par souci de

lisibilité, tous les phénomènes ne sont pas transcrits. Dans les corpus présentés dans ce chapitre, dans la mesure où certains accueillent différents types d'analyse, un choix a été opéré pour tenter de garantir ce compris. La transcription orthographique de ces corpus est complétée par certains phénomènes (les pauses mais non chronométrées, les chevauchements, par exemple) mais les phénomènes prosodiques ne sont pas mentionnés. En outre, la transcription orthographique n'est pas « truquée » dans la mesure où tous les enregistrements sont alignés texte/son et que les faits de prononciation sont facilement repérables. Ce choix est également lié aux effets péjoratifs que peut induire une transcription qui ne respecte pas l'orthographe usuelle (Blanche-Benveniste, Jeanjean 1997). Le choix effectué est celui de la diffusion facilitée des données, chaque chercheur pourra enrichir la transcription fournie.

Ainsi, chacun des éléments de la chaîne de traitement des données peut et doit faire l'objet de réflexions à la fois théoriques et pratiques pour garantir l'analysabilité des données interactionnelles. Je n'ai mentionné que deux éléments mais les autres ont également toute leur importance (André *et al.* 2010). La façon de se présenter sur le terrain peut garantir le recueil des données souhaitées ou le compromettre. Dans ce cas, c'est l'investigation du chercheur sur son terrain d'enquête qui est également compromise. Les dispositifs de captation doivent être pensés pour atteindre un compromis entre qualité des données et discrétion. A chaque nouveau recueil de données, les éléments du schéma sont réinterrogés et se reconfigurent. C'est au prix de ces étapes, souvent longues et laborieuses, que l'analyse sociolinguistique des interactions verbales est possible.

Chapitre 3. Des analyses sociolinguistiques d'interactions verbales – vers une interface sociolinguistique et didactique

Le cadre théorique étant posé, les données étant décrites, je peux désormais illustrer les deux précédents chapitres avec des analyses d'interactions verbales. Le choix des phénomènes présentés et analysés ici est fortement lié à leurs applications didactiques. En effet, j'ai sélectionné des pratiques interactionnelles particulières parce qu'elles ont fait l'objet de demandes d'une description de leur fonctionnement de la part d'enseignants et/ou d'apprenants de français, lors d'expérimentations en séances de formation linguistique. Ces expérimentations seront présentées dans les chapitres suivants. Ce chapitre central représente une sorte de pivot qui pose les bases d'une interface entre sociolinguistique et didactique. Il s'agit ici de dépasser « l'idéologie descriptive » (Guespin, Marcellesi 1986 : 7) afin de contribuer à la mise au jour de l'utilité de la sociolinguistique.

Le premier point se concentre sur le point de départ de mes travaux : l'analyse sociolinguistique de pratiques interactionnelles ([3.1.](#)). J'analyserai deux phénomènes qui me paraissent intéressants à la fois pour montrer la dynamique des interactions et pour leurs exploitations didactiques possibles : les énonciations conjointes ([3.1.1.](#)) et les questions ([3.1.2.](#)). Le deuxième point montrera de quelle façon le cadre théorique que je propose permet de définir les compétences socio-interactionnelles des locuteurs, de les identifier et de les évaluer, qu'ils soient apprenants de français ou en situation d'insécurité langagière et qu'ils soient allophones ou francophones natifs ([3.2.](#)). Enfin, le dernier point de ce chapitre soulignera l'articulation nécessaire entre sociolinguistique et didactique en proposant une interface cohérente ([3.3.](#)).

3.1. Des pratiques interactionnelles sous influence

Les premières études interactionnelles que j'ai réalisées mettaient en évidence la notion de coopération ou, plus précisément ce que j'ai appelé la construction collaborative du discours (André 2006a). La coopération, notamment au sens de Marx (1872), pointe l'activité combinée de plusieurs travailleurs dans la réalisation d'une tâche. En interaction, cette coopération existe, elle met en scène plusieurs locuteurs qui construisent ensemble,

de manière collaborative, leur discours, dans l'interaction. Le cadre d'analyse élaboré, et présenté dans le premier chapitre, permet une analyse sociolinguistique des interactions verbales. Il permet de rendre compte de la construction collaborative du discours en analysant ses traces et ses ressources langagières et interactionnelles. Parmi celles-ci, certaines pratiques interactionnelles ont retenu mon attention.

Mes travaux de thèse (André 2006a) ont mis au jour quatre phénomènes participant à la construction collaborative du discours, analysés en prenant en compte les éléments de la situation de communication. Ainsi, j'ai analysé au sein de réunions de travail l'apparition du couple *oui non*, du marqueur *hein*, des reprises et des énonciations conjointes. Ces phénomènes mettent au jour l'élaboration du discours au sein d'un collectif de travail. Ensuite, j'ai pu analyser ces mêmes phénomènes au sein d'autres corpus afin de confirmer et d'affiner les valeurs et les effets pragmatiques déjà repérés en analysant les activités interactionnelles accomplies par les locuteurs (André 2014, 2011, 2010). Ces phénomènes ont tous été étudiés au prisme de la situation de communication dans laquelle les réunions d'une part et les autres genres de discours analysés d'autre part prenaient place. Ces pratiques interactionnelles sont en quelque sorte « sous influence » (André 2005), c'est-à-dire qu'elles sont conditionnées, légitimées ou influencées par des éléments extralinguistiques. Seule une analyse sociolinguistique des interactions verbales peut les appréhender dans la mesure où elle s'appuie, pour son analyse, sur les éléments de la situation de communication ainsi que sur la construction collaborative de l'interaction et sur les ajustements mutuels des locuteurs. Elle prend en compte les multiples influences qui pèsent sur le déroulement des interactions, sur leur forme et leur nature. Elle prend en compte également les intrications entre les pratiques, les activités interactionnelles, le genre de discours et la situation. Les pratiques interactionnelles sont analysées au fur et à mesure de leur production et de leur élaboration *on line* (Traverso 2016) dans l'interaction. Elles sont également analysées au fur et à mesure de la reconfiguration de la situation qu'elles engendrent. Ces intrications et influences sont identifiables, interprétables (Kerbrat-Orecchioni 1989) et analysables, au moins en partie, grâce à la façon dont les corpus sont constitués, comme je l'ai expliqué précédemment.

Afin de poursuivre l'analyse de certaines pratiques interactionnelles, et de mettre au jour les multiples influences qu'elles subissent, je vais en (ré)examiner quelques-unes. Je vais tout d'abord étudier des énonciations conjointes et montrer que ce phénomène interactionnel collaboratif influence et est influencé par, à la fois des pratiques

interactionnelles cotextuelles et des éléments de la situation de communication (3.1.1.). Ensuite, j'aborderai certaines formes de question en interaction ainsi que les influences réciproques qu'elles mettent en œuvre (3.1.2.). La description et l'analyse de ces pratiques langagières *sous influence* sont inévitablement utiles à la didactique de l'oral qui n'aborde presque jamais ces phénomènes présents en interaction.

3.1.1. Des énonciations conjointes

Tout d'abord, je vais revenir sur l'analyse des énonciations conjointes (André 2006a, 2011, 2014). Ce phénomène interactionnel, déjà repéré notamment par Sacks, Schegloff, Jefferson (1974), puis Lerner (1991, 1996), Sacks (1992) et Janneret (1999), illustre parfaitement le cadre d'analyse présenté dans le premier chapitre. L'énonciation conjointe est constituée de deux tours de parole, le second – la complétion – s'inscrivant dans la construction syntaxique du premier, comme c'est le cas dans l'exemple suivant.

Exemple 1 : L7 ouvre l'assemblée générale d'une association municipale (A1). (Corpus Ag_AEL_08 – TCOF)

- L7 ben moi je voudrais vous féliciter déjà vous êtes nombreux parce que euh hein tu disais tout à l'heure que finalement il y avait pas mais je crois bien que c'est la première fois à l'assemblée générale de l'A1 où il y a autant de monde hein ou alors euh ça remonte déjà à loin loin **je sais pas si c'est**
- L8 **l'effet Beaujolais**
- L7 j'ai pas dit que c'était l'effet Beaujolais hein + (rires)

Le tour de L8 s'inscrit dans la continuité syntaxique de celui de L7 pour constituer une énonciation conjointe (en gras). En d'autres termes, un même énoncé est construit par deux locuteurs différents. Au tour suivant, L7 invalide cette complétion de L8 avant de poursuivre son tour (après une pause, marquée par +). Les participants à l'assemblée générale rient de cette énonciation conjointe, y compris le locuteur complété et celui qui a complété. Cette connivence et ces rires sont liés au déroulement prévu de la réunion qui doit se terminer par un pot et la dégustation du Beaujolais nouveau, un vin de primeur qui est mis en vente dans le monde entier le troisième jeudi du mois de novembre. Cette énonciation conjointe est à la fois une trace et une ressource de la construction collaborative du discours entre les locuteurs. Elle est autorisée par le principe de coopération entre les participants qui interagissent et révèle cette coopération. En outre, cette pratique interactionnelle est

influencée par les éléments de la situation de communication, notamment par le fait que l'assemblée générale réunit, dans une atmosphère détendue, des personnes qui se connaissent bien et qui ont l'habitude d'interagir ensemble dans ce même cadre. L8 peut légitimement compléter les propos de L7, bien qu'il soit le maire de la commune, en blaguant, sans que cela nuise aux interactions. Au contraire, cette complétion, de façon réflexive, accentue le caractère informel et détendu de la situation. Cette micro-analyse permet de saisir l'approche sociolinguistique, avec les multiples intrications et influences (schématiser en [1.1.](#)) et avec une attention particulière à la négociation de la structure locale de l'interaction et à l'organisation séquentielle des tours de parole (Sacks, Schegloff, Jefferson 1974 ; Goodwin 1979 ; Lerner 1991 ; Sacks 1992 (vol 1) ; Schegloff 2007).

Il est possible de schématiser le processus interactionnel de l'énonciation conjointe. Le tableau ci-dessous montre qu'elle est constituée de deux tours de parole et que le troisième tour permet d'interpréter la réception de ce phénomène.

Énonciation conjointe ⇒ deux tours de parole	{ Premier locuteur : Tour 1 Second locuteur : Complétion /Tour 2
Tour suivant : Tour 3 ⇒ Validation/ratification de la complétion ⇒ Refus de la complétion ⇒ Ignore la complétion	

L'expression *énonciation conjointe* est utilisée pour désigner la paire de tours de parole qui construit conjointement une unité commune. J'ai choisi l'expression *énonciation conjointe* mais d'autres coexistent, par exemple *collaborative sentence* (Jefferson 1973), *collaboratively built sentence* (Sacks 1992), phrase en collaboration (Jeanneret 1999), locuteur collectif (Loufrany 1981) ou encore *joint construction of utterances* (Coates 1994). Cette liste n'est pas exhaustive mais elle donne un aperçu de la diversité des études s'intéressant à ce phénomène. Le terme *complétion* est également utilisé par André-Larochebouvy (1984). Ensuite, la deuxième ligne du tableau montre que le tour 3 qui suit l'énonciation conjointe est un ajustement, le locuteur complété peut réagir de différentes façons. Il peut ratifier la complétion, le plus souvent avec un *oui* (ou ses combinaisons) ou une hétéro-répétition (voir également Oloff 2018) ou la refuser, le plus souvent avec un *non* ou un commentaire sur la complétion (comme dans l'interaction ci-dessus). Ce tableau

représente un schéma interactionnel prototypique de l'énonciation conjointe et de la réaction qui suit.

L'analyse sociolinguistique des interactions verbales fait également appel aux études plus spécifiques du français parlé telles qu'elles ont pu être menées par le Groupe Aixois de Recherche en Syntaxe (GARS) par exemple (Blanche-Benveniste, Jeanjean 1987 ; Blanche-Benveniste 1990b). Certains énoncés, montrant les traces de leur élaboration, méritent d'être disposés en grille afin de pouvoir observer leur organisation. L'exemple suivant montre à la fois ces traces de l'élaboration du discours oral en interaction et la construction collaborative du discours, ou encore ce que la linguistique interactionnelle appelle des réparations.

Exemple 2 : Une mère et ses deux filles discutent de théâtre et des possibilités d'avenir dans de ce milieu, notamment en intégrant le cours Florent à Paris. (Corpus Ave_BAT_08 – TCOF).

1. L1 tu veux pas être comédienne
2. L2 c'est pas possible
3. L1 ah bah c'est sûr si on part défaitiste < ***
4. L2 non mais euh > dans la classe **il y en a déjà euh trois qui v- qui v-**
5. L1 **qui veulent < être**
6. L3 **qui veulent > descendre < à Paris**
7. L2 **oui** > qui pensent qu'à ça quoi enfin quand tu vois Charlotte elle son rêve c'est de devenir comédienne j-
8. L3 Charlotte c'est laquelle < la rouquine
9. L2 la rousse > ouais < ouais

Dans cet extrait, les trois locutrices coopèrent pour construire leur discours commun. La répétition de L2 à la ligne 4 (« qui v- qui v- » en gras dans l'extrait) est un indice pour les interlocutrices d'un moment opportun de prise de tour ou d'un point possible de transition, notamment pour aider la locutrice en place. L1 et L3 produisent chacune leur tour (ligne 5 et 6) une complétion qui sera validée par L2 (avec « oui », ligne 7). Les tours de L1 et L3 exploitent la structure syntaxique du tour de L2 et s'ajustent mutuellement. Il est possible de disposer ces énoncés en grille (Blanche-Benveniste 1990b) afin de percevoir leur élaboration :

L2	il y en a déjà euh trois	qui v- qui v-
L1		qui veulent être
L3		qui veulent descendre à Paris

Cette représentation graphique met au jour l'élaboration de la langue parlée en interaction. Les traces de cette élaboration ou les piétinements sont forcément visibles (ou plutôt audibles) à l'oral, alors qu'elles sont supprimées (effacées) à l'écrit. Si l'on imagine que ces traces peuvent être supprimées à l'oral, il suffit d'observer les dernières lignes de chaque colonne pour voir le « produit fini » du discours oral, de la même façon que nous ne voyons que le produit fini des textes écrits. Cet extrait associe une énonciation conjointe – à trois locuteurs – et une collaboration discursive avec une aide à l'élaboration de l'énoncé à la suite d'un piétinement du locuteur en place. Cette énonciation conjointe est actualisée dans une situation de communication où les locutrices se connaissent bien, elles sont de la même famille et participent à une conversation informelle dans la maison familiale.

L'analyse de cet extrait disposé en grille, montre, en outre, l'« éclectisme » de mon cadre théorique qui est orienté par les données. En effet, j'utilise l'approche aixoise du français parlé, qui s'intéresse seulement aux faits langagiers, mais j'ajoute la prise en compte d'autres éléments, influençant inévitablement les pratiques interactionnelles. Ainsi, j'ajoute la prise en compte notamment des locuteurs et de leurs ajustements mutuels pour réaliser leurs activités interactionnelles et co-construire leur discours, en situation. L'association de plusieurs outils d'analyse permet de tirer profit de chacun d'entre eux et, en les combinant, de saisir la complexité des interactions.

Les phénomènes d'énonciation conjointe apparaissent dans tout type d'interactions et quel que soit le genre de discours mais sont influencés par certains éléments de la situation de communication. Ils mettent en lumière la coordination des activités interactionnelles des locuteurs, leurs activités conjointes (Sacks 1992 (vol 1) ; Clark 1996) et reflètent une collaboration à la fois interactionnelle et praxéologique. Ils mettent en évidence certains éléments de la situation qui légitiment ou expliquent leurs apparitions. Enfin, ils produisent également divers effets pragmatiques comme la création d'alliances ou de coalitions entre les locuteurs dans la mesure où ils soulignent leur coopération. Je vais mentionner brièvement d'autres exemples d'énonciations conjointes rencontrés dans différentes situations de communication. Le tableau ci-dessous propose une présentation synthétique des extraits accompagnés de certaines métadonnées et d'une petite analyse.

Nom du corpus et extrait	Métadonnées et informations nécessaires à la compréhension	Analyse de l'énonciation conjointe
<i>Professeurdesecoles_ROL_16</i> L2 c'est une formation nécessaire parce que ça donne les bases mais je pense qu'il y a des choses qui pourraient être faites L1 < autrement L2 en plus > pour qu'il y ait davantage de concret et moins de théorie	Une femme (38 ans) présente à une étudiante (19 ans) les études qu'elle a suivies pour devenir professeur des écoles.	L1 complète par anticipation son interlocutrice, qui prend une respiration après <i>qui pourraient être faites</i>. L1 s'appuie sur le discours de L2 pour le compléter. Il n'y a pas d'indice dans le discours de L2 qui pourrait laisser penser qu'elle a besoin d'aide. L1 propose donc une complétion à la fois pour collaborer, pour montrer qu'elle suit et qu'elle est impliquée dans le discours commun. La complétion de L1 est produite en même temps que la reprise du discours de L2 avec <i>en plus</i>. Même si la complétion ne correspond pas au discours que L2 souhaitait produire, elle ne nuit absolument pas au discours et L2 poursuit son récit. L'énonciation conjointe est produit par des locutrices qui ne se connaissent pas.
<i>Professeurdesecoles_ROL_16</i> L2 alors il y a pas une méthode après ce que je disais tout à l'heure c'est que chaque année euh on quitte des enfants des élèves et puis l'année d'après L1 c'est plus les mêmes L2 c'est plus les mêmes donc euh	Cet exemple est extrait de la même situation de communication que le précédent. Les locutrices parlent des méthodes d'enseignement.	L1 complète également par anticipation. Elle montre ainsi qu'elle suit le discours de son interlocutrice. La complétion est ratifiée par L2 qui répète la complétion et poursuit son discours
<i>Aidesoignante_VIV_17</i> L2 ça serait peut-être la possibilité pour moi pour rebondir + et puis justement faire quelque chose + L1 qui te plaît /	L2 raconte à sa fille L1 son parcours professionnel et comment elle est devenue aide-soignante. Dans cet extrait, L2 explique que son	L2 fait une pause (+). L1 l'interprète comme un moment opportun de prise de tour et complète le tour de L2 avec une intonation montante.

L2 oui oui et puis justement faire quelque chose qui aboutisse à une qualification un diplôme	licenciement lui a donné l'opportunité de faire une formation pour obtenir un diplôme.	L2 ratifie la complétion avec <i>oui oui</i> et poursuit son discours. La complétion sert à la fois à poursuivre le discours de L2 et à l'interroger sur ses motivations.
<i>Dimanche_NAG_16</i> L1 René Ruth c'est un allemand mais là par contre j'ai eu un petit soucis c'est que L4 il parle que allemand L1 il parle que allemand bah il y avait Didier	Lors d'un dîner en famille, L1 (le père) raconte son week-end en Allemagne avec des amis. Il raconte sa rencontre avec d'autres personnes et notamment avec René Ruth, un allemand.	L4, la femme de L1, complète la proposition. L4 ratifie et valide cette complétion en la répétant puis poursuit son discours. Dans cette situation, il est possible de faire l'hypothèse que la complétion soit à la fois une trace de et une ressource marquant la complicité ou la proximité entre les locuteurs.
<i>Atable_HAM_16</i> L1 c'est de la de la maison euh de la bibliothèque royale de euh + L3 Autriche L1 je crois et donc ils sont ils sont sublimes	Lors d'une conversation, deux jeunes femmes échangent sur des livres achetées par l'une d'elles. Elles admirent les couvertures et parlent des éditeurs.	L3 complète L1 qui hésite. L1 produit une hésitation et une pause. Ce sont des indices qui peuvent donner lieu à une complétion. L3 complète L1 en apportant l'information recherchée. L1 n'invalide pas la complétion mais fait un commentaire sur celle-ci (<i>je crois</i>) et poursuit son discours.

Tableau 1 : Présentation synthétique d'exemples d'énonciations conjointes dans TCOF

Ainsi, les énonciations conjointes sont des traces et de ressources de la construction collaborative du discours. Ces pratiques interactionnelles permettent d'accomplir différentes activités interactionnelles, qui s'ajoutent à la collaboration. En complétant l'énoncé de son interlocuteur, celui qui produit la complétion collabore et coopère au discours commun. Simultanément, il montre qu'il suit le discours, qu'il s'implique dans son élaboration, qu'il vérifie sa compréhension, qu'il est proche de son interlocuteur, qu'il partage des connaissances avec lui, qu'il est d'accord avec lui ou encore qu'il possède des connaissances supplémentaires. En outre, elles ont différents effets pragmatiques dans les interactions. En d'autres termes, elles sont à la fois influencées par les éléments de la

situation de communication et les pratiques interactionnelles du locuteur en place et elles influencent cette même situation dans la mesure où elles permettent de créer des alliances, des accords ou des désaccords entre les locuteurs.

Au-delà de l'analyse sociolinguistique de ces pratiques interactionnelles sous influence, il est intéressant de saisir les implications, notamment didactiques, de la présence de ce type de phénomènes dans les interactions verbales. Tout d'abord, dans d'autres analyses d'interactions (Langbach 2014), il a été montré que la non maîtrise des adaptations nécessaires du discours au contexte est une des principales causes des insécurités langagières. Les compétences socio-interactionnelles sont effectives lorsque les pratiques interactionnelles sont utilisées de façon appropriée à la situation de communication. Je reviendrai plus précisément sur ces points dans ce même chapitre. Dans le cas des énonciations conjointes, comme je l'ai expliqué, leurs apparitions sont conditionnées aux éléments de la situation et à la coopération entre les interactants. Par conséquent, les locuteurs qui utilisent comme ressources interactionnelles ces pratiques, font preuve d'une compétence interactionnelle ou d'une « compétence d'interaction verbale » (Pekarek Doehler 2006 : 22). Ce constat a inévitablement des conséquences sur la façon d'enseigner et d'apprendre une langue, notamment à interagir, ainsi que sur la façon d'identifier et d'évaluer ces compétences.

Les énonciations conjointes mettent en évidence la compétence interactionnelle, celle qui ne relève pas exclusivement de la responsabilité du seul locuteur en cours. D'autres pratiques interactionnelles peuvent être intéressantes à décrire et à analyser avec une visée didactique, c'est-à-dire en pensant à l'acquisition de compétences socio-interactionnelles facilitant les échanges exolingues. En réalité, toutes les pratiques interactionnelles utilisées par les locuteurs sont intéressantes à analyser à des fins didactiques. Néanmoins, certaines pratiques sont bien décrites et facilement identifiables (par des enseignants de FLE par exemple) et d'autres échappent aux descriptions (des manuels de FLE par exemple). J'aborderai ces aspects dans le chapitre suivant toutefois je vais examiner d'autres exemples de pratiques interactionnelles fortement sous influence, c'est-à-dire fortement intriquées à la situation de communication, relevées dans différents corpus d'interactions.

3.1.2. Des questions

Parmi les pratiques interactionnelles aux prises avec de multiples influences, il est intéressant d'examiner les questions en interaction. L'objectif n'est pas de proposer une analyse exhaustive des différentes formes de questions (voir notamment Kerbrat-Orecchioni 1991 ; Coveney 1996 ; Rossari *et al.* 2005 ; Beguelin *et al.* 2018) mais d'identifier en quoi l'analyse sociolinguistique des interactions verbales permet de pointer des phénomènes socio-interactionnels spécifiques, aux prises avec de multiples influences. Mon intérêt pour les questions est né d'une collaboration avec Jeanne-Marie Debaisieux, lorsque nous préparions un cours commun sur l'utilité des corpus oraux, destiné aux étudiants de deuxième année de la Licence de Sciences du langage à Nancy. Ce travail collaboratif a été poursuivi et a donné lieu à une des rares études qui tentent d'exploiter l'analyse de corpus à des fins didactiques (Debaisieux, Boulton 2007). Nous avons proposé aux étudiants une présentation qui visait à montrer que les questions, qui sont définies dans les grammaires et les manuels de FLE, n'avaient pas toujours les formes canoniques décrites. Je reviendrai sur l'importance de cet écart entre ce que montrent ces documents et ce que montrent les corpus dans le chapitre 4 (en [4.2.](#)). Je vais simplement présenter les résultats chiffrés d'une petite étude que nous avons réalisée à partir d'un corpus de deux heures d'enregistrement (une heure de conversations et une heure d'interactions à la scolarité de l'université). Nous avons analysé toutes les questions, c'est-à-dire toutes les demandes d'informations (Kerbrat-Orecchioni 2001), de ce corpus, en fonction de ce qui est présenté dans les grammaires et les manuels. Nous avons constaté la répartition suivante, en ce qui concerne la forme des questions :

- Inversion sujet-verbe : 0,67%
- *est-ce que* : 4%
- Mot interrogatif : 30,6%
- Intonation : 34%

L'ordre des résultats est présenté en respectant l'ordre préférentiel des grammaires et des manuels (généralement sans distinguer écrit et oral, sauf pour l'intonation). Ainsi, dans ces documents de référence, la première façon de poser une question est d'inverser le sujet et le verbe mais elle représente moins de 1% des questions de notre corpus. L'intonation, qui est présentée en quatrième position, est rarement mise en valeur en tant ressource interactionnelle mais représente 34% des questions du corpus. En outre, il reste 31% des

questions dont la forme n'est pas présentée dans ces documents de référence. Ce pourcentage élevé, même s'il est issu d'une toute petite étude, est intrigant. Quelles sont ces ressources qui permettent de poser des questions et qui ne sont pas décrites dans les grammaires et les manuels ? Ce travail syntaxique et pragmatique, pour notre enseignement, m'a ensuite conduite à mener une analyse sociolinguistique des questions dans différents corpus. Cette analyse permet d'identifier au moins une partie des 31% de questions non décrites.

Tout d'abord, je vais reprendre l'exemple présenté précédemment, dans le point [1.1.1.](#), qui a permis d'éclaircir le cadre d'analyse sociolinguistique des interactions verbales et le fonctionnement du schéma associé. Il illustre parfaitement la variété des formes que peuvent prendre les questions ainsi que leur caractère situé. Dans cet exemple, la locution *c'est-à-dire* est utilisée pour demander à un interlocuteur de poursuivre son discours. Pour rappel, voici l'extrait analysé :

Exemple 1 : Une jeune femme interroge un ami qui vient de quitter l'armée française (TCOF – Armee_MER_15).

1. L1 mais toi tu en penses quoi de l'armée franchement /
2. L2 à l'heure d'aujourd'hui ça vaut rien
3. L1 **c'est-à-dire** /
4. L2 ça vaut plus rien du tout
5. L1 **ben explique** /
6. L2 ben quand on voit euh comment + les les moyens euh déployés par euh l'armée française et les moyens déployés par euh l'armée américaine

Les deux tours de paroles de L1 qui servent de relance (*c'est-à-dire*, ligne 3, et *ben explique*, ligne 5) visent à obtenir des informations. Ainsi, ils fonctionnent comme des questions, selon les définitions de Kerbrat-Orecchioni (2001 : 86) : « Nous considérons comme question *tout énoncé qui se présente comme ayant pour finalité principale d'obtenir de son destinataire un apport d'information* » ou de Ravazzolo *et al.* (2005 : 56) : « La question peut être définie comme un énoncé qui appelle généralement une réponse et qui permet au locuteur d'obtenir de son destinataire un apport d'informations ». Ces définitions interactionnelles et pragmatiques permettent d'appréhender différentes pratiques comme des questions, telles que les tours de L1 dans l'exemple précédent. En outre, l'activité interactionnelle qui consiste à questionner, interroger ou obtenir des informations peut être réalisée par des pratiques interactionnelles différentes et dont les formes ne sont pas prévisibles. Je reviendrai sur l'importance de transposer ces analyses en didactique.

En outre, certaines de ces formes sont conditionnées par le genre de discours et les éléments de la situation de communication, comme je l'ai expliqué dans le premier chapitre. Par exemple, le tour de L1 *ben explique* dans l'extrait précédent (ligne 5) fonctionne comme une question mais sa forme, un petit mot de l'oral suivi d'un impératif, est liée notamment au fait que les interlocuteurs sont des amis de longue date et qu'ils participent à une conversation informelle ou à bâtons rompus. Les ressources interactionnelles utilisées par les locuteurs pour obtenir des informations sont influencées par le genre de discours et la situation de communication. Dans ce tour, l'impératif perd sa valeur d'ordre (voir par exemple Arrivé *et al.* 1986 : 318 ou Riegel *et al.* 1994 : 330) au profit d'une valeur de sollicitation grâce aux conditions de l'interaction. Dans une autre situation de communication, l'utilisation de cette pratique avec ce même objectif pourrait être totalement inappropriée.

Je vais maintenant analyser le premier tour de L1 dans cet extrait : *mais toi tu en penses quoi de l'armée franchement* (ligne 1). Ce tour est une question. Il est composé de plusieurs unités dont certaines sont disloquées. Il peut être découpé de la façon suivante :

<i>mais toi</i>	<i>tu en penses quoi</i>	<i>de l'armée</i>	<i>franchement</i>
-----------------	--------------------------	-------------------	--------------------

L'organisation de ce tour semble répondre à l'intention communicative de L1. La locutrice s'adresse à son ami en posant un élément de cadrage qui permet de focaliser le discours sur lui (*mais toi*) et de changer de thème (De Fornel 1988). Ce tour apparaît après que L2 a expliqué qu'il existe des personnes avec des bracelets électroniques qui préfèrent entrer dans l'armée plutôt que d'aller en prison. Ensuite, L1 pose le cœur de sa question avec un mot interrogatif (*tu en penses quoi*) avant de produire une dislocation à droite du thème ou du topic (*de l'armée*, probablement par opposition au thème concernant les personnes avec des bracelets électroniques). Enfin, L1 termine son tour par l'adverbe *franchement*, accompagné d'une intonation descendante, pour préciser quelle attitude énonciative elle attend (Hermoso 2009). La construction progressive de ce tour permet à L2 de comprendre le discours de L1 et de prendre son tour en s'y ajustant. Par ailleurs, cet exemple montre de quelle façon le discours se construit à l'oral. Les informations sont posées progressivement, permettant à l'interlocuteur de prendre connaissances des éléments topicaux et de se remémorer ceux dont il est question. Le mode d'élaboration du discours oral, la progression

de la mémoire discursive et le rappel des connaissances partagées conduisent les locuteurs à poser successivement plusieurs unités.

Ainsi, la situation de communication peut permettre à des activités interactionnelles de se réaliser avec des pratiques interactionnelles particulières voire inattendues. Lors de l'étude des questions dans différents corpus, j'ai relevé, à la suite de Debaisieux et Boulton (2007), des énoncés, visant à obtenir des informations, prenant la forme de semi-assertions ou d'énoncés syntaxiquement inachevés. Ce sont à ces derniers que je vais m'intéresser ici. Une fois encore, les analyses présentées dans ce travail de synthèse sont choisies parce qu'elles ont été l'origine de discussions ou de suggestions lors de ma collaboration avec des enseignants de FLE ou lors de mes propres expériences d'enseignante de FLE ou de FL21, notamment lors d'expérimentations.

Parmi les énoncés syntaxiquement inachevés ou suspendus qui permettent de poser des questions, les énoncés alternatifs ne mentionnant qu'une seule des propositions sont relativement fréquents. Ils s'achèvent généralement par un allongement sur la conjonction de coordination *ou*, comme c'est le cas dans l'exemple qui suit.

Exemple 2 : Lors d'un entretien de type sociologique, L1 interroge L2 sur son anorexie (TCOF – Anorexie_BAN_17).

- L2 ça a duré vraiment on va dire médicalement parlant avec un suivi et un arrêt de travail pendant presque deux ans pendant deux ans je ne suis pas retournée travailler
- L1 et quand tu dis que tu as eu un suivi **tu as eu un suivi euh avec un médecin euh généraliste plutôt comme ça ou::**
- L2 au départ au départ c'était généraliste mais il a vite été dépassé par la situation

L'analyse sociolinguistique du tour de L2 permet de montrer que cette question reflète certains éléments de la situation de communication. Les locuteurs ne se connaissent pas et le sujet de l'échange est personnel et potentiellement délicat. L1 souhaite demander si L2 a été suivie par un médecin ou par un spécialiste, par exemple un psychologue. Cependant, afin de préserver la face de son interlocutrice (Goffman 1974), L1 choisit de poser une question alternative sans proposer d'alternative à « médecin généraliste » qui ne risque que peu de mettre mal à l'aise L2, puisqu'elle vient de dire qu'elle a eu un traitement et qu'elle a été suivie. L'allongement produit sur la conjonction *ou* (symbolisé par ::) peut être interprété comme un choix de ne pas proposer un achèvement syntaxique à cet énoncé et de ne pas proposer une ou plusieurs alternatives. Toutefois, cet énoncé est bien achevé

pragmatiquement, l'activité interactionnelle est réalisée, et L2 prend le tour pour répondre à la question posée (58 secondes après cette réponse, L2 dit qu'elle est allée voir un psychologue).

Dans d'autres situations, ce procédé apparaît lorsque le locuteur ne sait pas quelle alternative il pourrait proposer ou lorsque que l'alternative est la négation de la première proposition. Il est parfois difficile de se prononcer pour l'une ou l'autre des analyses, comme c'est le cas dans l'exemple suivant :

Exemple 3 : Quatre amies discutent de cirque, que pratique L1. L'extrait porte sur le nom donné à une école de cirque (TCOF – Cirque_HEL_17).

- L1 mais il aurait pu à mon sens ça aurait été plus juste d'appeler ça l'académie Gachet
 que de garder le nom Fratellini < parce que
- L3 hum hum > ça n'avait rien à voir
- L4 mais il l'a gardé
- L1 d'une certaine manière non mais ça avait à voir parce que c'était du cirque
- L4 **il l'a gardé par hommage ou:**
- L1 il l'a gardé parce que c'est Paul Fratellini qui est venu le chercher pour < essayer
- L2 ouais >
- L1 de sauver l'école d'Annie Fratellini et que la l'école d'Annie Fratellini est devenue
 l'académie Fratellini avec un tout autre projet dans d'autres locaux

Après que L1 a donné son avis sur le nom d'une école de cirque, L4 demande une précision sur les raisons de ce choix avec l'énoncé alternatif, syntaxiquement inachevé *il l'a gardé par hommage ou* (en gras). Il semble que L4 n'ait pas d'alternative à proposer, ce qui explique l'allongement sur la conjonction *ou*. L4 fait une première proposition et laisse une place de transition pour que L1 réponde, ce qui se produit immédiatement.

Ce type de question est présent également dans le corpus FLEURON. De la même façon, des éléments de la situation de communication expliquent l'usage d'un énoncé inachevé. Dans l'exemple ci-dessous, une étudiante interroge un agent administratif.

Exemple 4 : A la scolarité de l'université, l'agent (A) explique à l'étudiante (E) qu'elle devra utiliser l'adresse mail de l'université (FLEURON – Inscription complète Erasmus 3).

- A c'est important je j'écirai sur cette adresse l'uni- l'adresse de l'université Nancy 2
- E **et c'est moi qui dois euh le créer ou::**
- A non elle se fera automatiquement hein quand vous ferez la connexion à Internet
- E d'accord

L'étudiante demande si elle doit elle-même créer son adresse mail. L'alternative pourrait être que l'agent la crée ou qu'un service informatique s'en charge. Ces éventuelles deux possibilités n'ont pas besoin d'être prononcées, A répond à la question (*non*) à la suite de l'allongement (*ou::*). Le choix de la proposition ne semble pas être fait au hasard, l'étudiante demande d'abord si c'est elle qui doit faire la tâche avant de demander si quelqu'un d'autre s'en occupe. Le statut d'étudiant peut expliquer l'apparition de la seule proposition de réaliser la tâche soi-même.

Les énoncés inachevés ou suspendus permettent donc d'interroger son interlocuteur et d'obtenir des informations. Ils permettent de produire des énoncés stratégiques, soit en termes de préservation de la face de son interlocuteur ou de la sienne, soit en termes de méconnaissances du domaine abordé, soit les deux. De nombreux exemples apparaissent dans les corpus avec lesquels je travaille. Le tableau suivant en mentionne quelques-uns.

Nom du corpus et extrait	Métadonnées et informations nécessaires à la compréhension	Analyse de la question
<i>Professeurdesecoles_ROL_16</i> L1 donc euh les études que vous avez faites à l'ESPE vous avez été à l'ESPE L2 alors c'est pas l'ESPE nous c'était encore l'IUFM L1 c'est la même chose ou euh L2 c'est à peu près le même concept euh il y a euh deux années et euh	Une femme (38 ans) présente à une étudiante (19 ans) les études qu'elle a suivies pour devenir professeur des écoles. <i>Nb. L'ESPE (École Supérieure du Professorat et de l'Éducation) est le nouveau nom de l'IUFM (Institut Universitaire de Formation des Maîtres)</i>	L1 pose une question avec un énoncé alternatif inachevé. La conjonction <i>ou</i> est suivie d'une hésitation <i>euh</i>. Il n'est pas rare que ces deux éléments apparaissent successivement dans le cas de ces énoncés inachevés qui permettent de poser des questions.
<i>Professeurdesecoles_ROL_16</i> L2 donc là on a eu une mise en place de nouveaux programmes + et peut-être quand on changera de président voir de ministre peut-être qu'on aura d'autres programmes L1 c'est euh c'est dur de faire la transition entre les les programmes ou euh	Cet exemple est extrait de la même interaction que l'exemple précédent.	L1 pose une question avec un énoncé inachevé qui se termine en <i>ou euh</i> . On peut faire l'hypothèse que l'alternatif est la négation de la première partie. L2 répond immédiatement à la question.

L2 pas for- pas forcément difficile mais c'est euh + c'est c'est les approprié		
<i>Association_MAU_17</i> L2 en fait ce qu'ils veulent c'est juste passer un bon moment et oublier < un peu le reste L1 c'est ça > L2 quoi je pense < * L1 hum > ils sont suivis par des thérapeutes ou euh + L2 euh ben il y a une psychologue qui vient	Une jeune femme qui travaille dans une association pour migrants raconte son activité à une étudiante.	L1 pose une question avec un énoncé inachevé qui se termine en <i>ou euh</i> . L2 répond à la question, en prenant son tour après la pause. Cette question inachevée cherche à préserver la face de la locutrice qui la produit. Elle pose une question délicate sur le suivi psychologique (probablement qu'elle estime nécessaire) des migrants de l'association.
<i>Lyme_CRO_17</i> L1 alors ça fatigue physique euh psychique euh ou L2 euh euh physique d'autant plus que cet excès de fatigue je veux dire euh + des difficultés à faire un geste des douleurs articulaires euh je me sou- d'un seul coup je me sentais comme une personne qui avait pris euh je ne sais pas moi trente ans d'âge	Un père raconte à son fils les conséquences de sa maladie de Lyme.	L1 interroge L2 sur la fatigue qu'il ressent à cause de sa maladie en faisant deux propositions puis ouvre les propositions avec un <i>ou</i> , précédé d'une hésitation <i>euh</i> . Cet énoncé qui a la forme d'un énoncé inachevé permet d'interroger l'interlocuteur, de lui demander des explications et de poursuivre son discours.

Tableau 2 : Présentation synthétique d'exemples d'énoncés inachevés qui permettent de poser des questions.

Ces énoncés inachevés ou suspendus permettent de poser des questions pour obtenir des informations de la part de son interlocuteur. Ils apparaissent dans différents genres de discours (conversation, entretien, relation de service) et dans différentes situations de communication (formelles, informelles, amicales, familiales, professionnelles, avec des locuteurs qui se connaissent ou qui se rencontrent pour la première fois).

Dans un tout autre contexte, lors de la réalisation d'outils d'évaluation des compétences langagières des salariés des métiers de la propreté (présentés ci-dessous, dans le point [3.2.1.](#)), un corpus d'interactions a été constitué. Son analyse a permis d'identifier

les besoins langagiers des travailleurs. Cette analyse a montré que les questions ne prenaient quasiment jamais de formes canoniques. Elles sont ancrées dans la situation de communication et relèvent de ce que l'on peut appeler un langage « opératif » (Falzon 1991 : 236) en situation de travail. Les questions relevées sont chargées de déictiques, comme dans les quatre exemples suivants qui sont extraits d'échanges entre un cadre et des opérateurs :

Exemples : Une chef d'équipe (C) fait le point avec ses employées sur le travail effectué ou à effectuer. Le chantier de nettoyage est un Centre Hospitalier Universitaire (CHU).

C1 : là vous avez bien fait là

C2 : et là c'est bon ça

C3 : et là-haut tout est ok

C4 : demain c'est qui est-ce qui fait là (*montre les chambres et les couloirs*)

Ces quatre exemples suffisent pour rendre compte de la particularité des questions dans ces situations de travail. Ces énoncés sont accompagnés par des gestes de pointage du doigt, de la main, de la tête ou du buste du locuteur (Mondada, Dubois 2004). L'ancrage situationnel montre que l'économie du discours est importante, l'agent ne produit que ce qui est utile à l'échange. Dans la mesure où les interactions sont intriquées à la situation dans laquelle elles apparaissent, les pratiques interactionnelles sont situées et réduites à ce qui est nécessaire pour exécuter le travail.

Volontairement, je n'analyse ici que peu d'exemples, l'objectif étant de montrer que le cadre d'analyse sociolinguistique des interactions verbales permet de saisir des pratiques et des activités construites en interaction et interinfluencées par de multiples facteurs. Les intrications, représentées par le schéma illustrant le cadre d'analyse, des pratiques interactionnelles, des activités interactionnelles, du genre de discours et de la situation, sont à prendre en compte pour saisir la dynamique des interactions verbales. Ces analyses montrent également comment certaines ressources peuvent accomplir des activités interactionnelles et avoir les effets escomptés parce que certains éléments de la situation les légitiment ou les conditionnent. Enfin, ces analyses ont nécessairement des implications didactiques, notamment en ce qui concerne les objectifs à atteindre à les compétences à évaluer.

3.2. Identification des compétences socio-interactionnelles

La conception sociolinguistique de la langue, et de ses réalisations individuelles, collaboratives et interactionnelles, a nécessairement une implication sur l'identification des compétences communicatives des locuteurs. Conséquemment, cette conception et cette identification des compétences ont également des répercussions sur les objectifs d'enseignement et d'apprentissage en FLE. Ces derniers ne peuvent être définis, comme je l'ai montré précédemment, sans une analyse sociolinguistique de la langue, et sans s'intéresser à la mise en œuvre de la langue, par les individus qu'ils l'utilisent en situation.

Ce point est consacré à des travaux menés autour des compétences sociolangagières ou socio-interactionnelles, de leur identification, de leur amélioration et de leur évaluation. J'utilise socio-interactionnelles pour évoquer les compétences en langue qui sont nécessairement réalisées en interaction, même lorsqu'il s'agit de compétences scripturales. Toutes les ressources langagières utilisées par les locuteurs et les scripteurs sont mises en œuvre pour interagir. De plus, ces ressources, pour atteindre les objectifs de la communication, doivent être appropriées à la situation et doivent donc prendre en compte les aspects sociaux des interactions. C'est pourquoi j'ajoute le préfixe socio- à interactionnel(le).

La cohérence voudrait que ce soit dans un ordre logique – identification-évaluation-amélioration – que mes travaux concernant les compétences soient présentés, cependant je vais procéder autrement. Je vais présenter mes travaux dans l'ordre chronologique de leur réalisation, d'abord parce que cette synthèse de mon parcours de recherche peut y gagner en clarté, ensuite parce que chaque étude a nourri les suivantes. Les trois points présentés dans cette section ont contribué à l'aboutissement du cheminement exposé dans ce travail et ont permis de consolider les liens entre l'analyse des interactions et les préoccupations didactiques.

Avant de présenter ces trois études, je vais revenir brièvement sur la notion de compétence et tenter de la définir en m'inspirant de l'analyse des situations de travail, comme je l'ai expliqué dans le premier chapitre. Pour compléter cet apport, je vais me référer aux travaux de Hymes (1970, 1984) et à ce qu'il a nommé la « compétence de communication » ainsi qu'à la notion de « compétence élargie » de Bourdieu (1977). Élaborées en opposition à la notion de performance ou de compétence générative de Chomsky, ces compétences – de communication et élargies – sont proches du point de vue

de leur définition. Contrairement à la position de Chomsky qui se réfère à la notion de locuteur-auditeur idéal, la célèbre formule de Hymes, « There are no rules of use without which the rules of grammar would be useless » (1972b : 278), met en avant la mise en œuvre des savoirs en situation. Ainsi, comme les compétences en situation de travail (Durrive 2016), les compétences en langue sont mises à l'épreuve du réel des situations de communication.

3.2.1. Des outils d'évaluation de compétences socio-interactionnelles

Les premiers travaux du groupe Langage, Travail et Formation sont nés d'une demande sociale (André 2009a, 2015a). Ce groupe s'est d'ailleurs constitué, en 2006, avec cette demande, formulée par la branche professionnelle des métiers de la propreté. Je ne détaillerai pas précisément cette demande et les travaux qui ont été réalisés pour y répondre mais j'insisterai sur les aspects sociolinguistiques. L'Organisme Paritaire Collecteur Agréé (OPCA)³⁸ du secteur de la propreté a souhaité évaluer les compétences en langue de ses salariés, dans un premier temps, et de ses candidats à l'emploi, dans un second temps. Les origines et les objectifs de ces évaluations linguistiques sont multiples. Tout d'abord, dans le secteur de la propreté, 91% des salariés ont un niveau V ou infra V³⁹ et 31% sont de nationalité étrangère. Ce secteur, à l'instar de celui du bâtiment, de l'hôtellerie ou de la restauration, est un secteur en tension, c'est-à-dire qu'il rencontre des difficultés pour recruter et qu'il y a plus d'offres d'emploi que de candidats. Pendant longtemps, le secteur de la propreté a recruté des salariés sans s'intéresser à leurs compétences en français, à l'écrit comme à l'oral. Ce désintérêt s'explique en partie par la non qualification des emplois du secteur et par la quasi absence de communication dans les différentes situations de travail rencontrées par les salariés. Avec les outils d'évaluation demandés, il s'agit à la fois d'avoir une vision de la maîtrise ou de la non maîtrise du français par les salariés ou futurs

³⁸ En France, jusqu'en mars 2018, les OPCA recevaient et mutualisaient les contributions financières des entreprises adhérentes correspondant à un pourcentage de leur masse salariale annuelle. Ces organismes finançaient et élaboraient avec leurs partenaires des actions de formation pour les salariés de ces mêmes entreprises. Depuis le 1^{er} avril 2019, les OPCA sont remplacés par les OPCO (opérateurs de compétences), chargés également d'accompagner les formations.

³⁹ Pour rappel, l'INSEE établit un classement des niveaux de scolarisation. Le niveau V correspond aux personnes sorties de leur scolarisation après l'année terminale de CAP ou BEP ou sorties après un 2nd cycle général et technologique avant l'année terminale (seconde ou première). Pour voir les différents niveaux : <https://www.insee.fr/fr/metadonnees/definition/c1076>

salariés et de prévoir des formations en langue si besoin. Ces évaluations, suivies de formations, arrivent à un moment où le travail se transforme (André 2009b). La part langagière (Boutet 2001) augmente dans tous les secteurs et pour tous les métiers, quelles que soient les fonctions hiérarchiques occupées, et de nouveaux besoins langagiers apparaissent.

En outre, les métiers de la propreté ont fait l'objet d'une revalorisation avec la création de diplômes de branche. Pour faciliter l'accès à ces Certificats de Qualification Professionnelle (CQP), le secteur a mis en place un dispositif de Validation des Acquis de l'Expérience (VAE). Les salariés devaient constituer un dossier de VAE avec l'aide d'un accompagnateur. Néanmoins, avant même la constitution de ce dossier, l'OPCA souhaitait savoir si les salariés concernés avaient les compétences langagières nécessaires pour s'engager dans cette démarche qui impliquait une part de verbalisation du travail (Adami 2007). Le premier outil d'évaluation des compétences socio-interactionnelles des salariés permet de les positionner en donnant des indications précises quant à leur capacité à s'engager dans le processus de VAE. Il évalue également la capacité des salariés à participer à différents types d'interaction. Dans le cadre de la revalorisation des métiers de la propreté, les conditions de travail ont été modifiées et améliorées. Le travail en équipe apparaît. Les horaires décalés diminuent. Par conséquent, les interactions verbales augmentent, entre les collègues, avec la hiérarchie, avec les clients. Ces mutations du travail concernent tous les secteurs d'activité. Ces nouveaux besoins langagiers, identifiés également au niveau européen dans différents secteurs par les groupes de travail menés par Grünhage-Monetti *et al.* (2004, 2005), peuvent faire l'objet de diverses formations continues pour les salariés et doivent être pris en compte au niveau du recrutement des futurs salariés (Mourlhon-Dallies 2008). Les outils d'évaluation que nous avons créés pour le secteur de la propreté ont permis à la fois d'identifier les compétences socio-interactionnelles nécessaires à l'effectuation des activités professionnelles et de les évaluer chez les salariés et futurs salariés.

Cette demande de création d'outils de positionnement concentre presque tous les aspects de mes travaux – la sociolinguistique des interactions verbales, l'exploitation de corpus et la didactique du français langue d'intégration et d'insertion (FL2I, voir Adami, André 2012b). En outre, cette demande a permis de réunir un groupe de personnes dont les thématiques dominantes de recherche étaient complémentaires. En effet, comme je l'ai expliqué en introduction, le groupe, composé initialement de Hervé Adami, Philip Riley,

Florence Poncet et moi-même, a contribué à faire émerger une recherche capable d'articuler les trois domaines – langage, travail, formation – avec des préoccupations interactionnelles, sociales, professionnelles et didactiques. Ainsi, pour répondre à la demande de création d'outils d'évaluation des compétences langagières des métiers de la propreté, nous avons tout d'abord réalisé une analyse sociolinguistique des interactions, orales et écrites, auxquelles les salariés étaient susceptibles de participer. Une vaste enquête de terrain a été menée sur différents chantiers de nettoyage afin de recueillir un corpus d'interactions à partir duquel les compétences socio-interactionnelles nécessaires à la réalisation de l'activité de travail ont été identifiées. La démarche entreprise correspond à une conception sociale et praxéologique de la langue. Nous avons cherché à identifier les réels besoins interactionnels des salariés en situation de travail (même si nous avons tout de même accordé une place au travail prescrit et aux éventuels besoins langagiers liés aux prescriptions mais non observés sur les chantiers). Le corpus de travail est constitué d'enregistrements d'interactions – entre les salariés, avec les responsables d'équipe, avec les clients – sur différents types chantiers (collectivités, hôpitaux, bureaux, *etc.*), d'entretiens avec des professionnels (encadrement, opérateurs) et de documents écrits (fiches de poste, référentiels métiers, manuels, cahiers de liaison, modes d'emploi, plannings, petits mots laissés sur des bureaux, consignes de sécurités, *etc.*). L'analyse de ce corpus a permis d'identifier les différentes situations de communication (et leurs éléments), les genres de discours récurrents, les activités interactionnelles réalisées et les pratiques interactionnelles actualisées.

L'évaluation que nous avons proposée est nécessairement située et contextualisée. Conformément à mon approche de la notion de compétence, il n'est pas envisageable d'évaluer la production ou la compréhension de façon décontextualisée, c'est-à-dire sans qu'un individu puisse mettre en œuvre ses ressources interactionnelles dans une situation particulière. Il s'agit alors d'évaluer la mobilisation et l'adaptation des savoirs et savoir-faire ou encore des ressources interactionnelles disponibles en interaction dans une situation particulière. Comme l'explique également Pekarek Doehler (2006 : 41) à propos de l'évaluation des compétences en langue :

Il ne peut plus être question de prendre la production de l'apprenant comme un indicateur indépendant d'un progrès quelconque, mais il s'agit d'interroger l'apprentissage comme un forme située de comportement humain. L'observable n'est alors plus la production de l'apprenant prise isolément, mais relève de sa participation à l'organisation d'activités, à leurs enchaînements séquentiels, à leur coordination, et à la configuration même des contextes

d'action. L'analyse prend dès lors pour départ non pas des formes linguistiques définies dans l'abstrait, mais les manières dont les participants accomplissent des activités situées (...).

Par conséquent, une articulation doit être effectuée entre les données de terrain recueillies pour constituer un corpus de travail et les données issues des évaluations. Les premières permettent de saisir les interactions en situation et les secondes doivent permettre l'évaluation des compétences socio-interactionnelles, c'est-à-dire la capacité pour les salariés de se comporter de façon appropriée dans les situations identifiées et analysées dans le corpus de travail. L'analyse des premières données nourrit le barème de l'évaluation des secondes. En d'autres termes, les interactions auxquelles participent les salariés lors du test sont analysées et évaluées au regard des interactions produites en situation de travail réelle. Par exemple, l'analyse des questions du corpus de travail révèle l'utilisation d'un grand nombre de déictiques verbaux et gestuels (voir le point précédent [3.1.2.](#)), les questions produites par les salariés lors de l'évaluation dans une situation similaire à celle du corpus du travail pourront donc être évaluées positivement, si elles comportent ces mêmes déictiques et si elles sont appropriées à la situation. Dans d'autres situations, les déictiques ne seraient pas appropriés. En résumé, sans chercher à réduire les compétences des salariés de la propreté, leur évaluation est nécessairement située, relative à la situation, aux activités interactionnelles visées et au genre de discours dans lequel les pratiques interactionnelles apparaissent. Cette approche des compétences socio-interactionnelles et de la compétence de communication a également des conséquences sur la façon d'aborder les formations linguistiques.

3.2.2. Écrire, un prétexte pour acquérir des compétences socio-interactionnelles

En 2016, avec Valérie Langbach, nous avons mené une étude pilote au sein d'une association d'insertion d'une petite ville de Meurthe-et-Moselle à la périphérie de Nancy. Cette étude⁴⁰, intitulée « Mémoire sociale et écriture », avait plusieurs objectifs⁴¹, dont les deux suivants :

⁴⁰ Cette étude est née d'une idée originale d'Hervé Adami. Elle a été réalisée avec le soutien de la Direction Générale à la Langue Française et aux Langues de France (DGLFLF) et de la Direction Régionale des Affaires Culturelles (DRAC) Lorraine, dans le cadre de l'appel à projet « Action culturelle au service de la maîtrise de la langue ».

⁴¹ Cette étude est présentée de façon détaillée sur le site du réseau *Langage, Travail et Formation*, à l'adresse suivante : <https://reseaultf.atilf.fr/?p=803#more-803>. Le déroulé de la présente étude ainsi que

1. permettre aux participants de s'exprimer sur un sujet culturel ou de société ;
2. faciliter l'entrée dans l'écrit en proposant la rédaction aidée, individuelle ou collective, d'un texte (sur le principe de la dictée à l'expert).

Cette expérimentation est menée avec des personnes en insécurité langagière (Adami, André 2014) salariées de l'association d'insertion et qui suivent des cours de remise à niveau en français et mathématiques. Nous avons souhaité exploiter les possibilités offertes par la rédaction aidée individuelle ou collective d'un texte écrit pour à la fois faciliter l'entrée dans l'écrit et solliciter l'expression orale. De façon plus globale, l'objectif était de faire prendre conscience, à ces personnes en insécurité langagière, du fonctionnement du langage pour améliorer leurs compétences socio-interactionnelles (à l'écrit et à l'oral).

Cette recherche-action a pour objectif premier de saisir la façon dont la mémoire ouvrière se transmet dans une région de tradition industrielle. La première étape de cette étude a été de recueillir la parole ouvrière d'anciens salariés du bassin industriel de la région. Nous avons choisi d'interroger des anciens ouvriers de la tuilerie de Jeandelaincourt, petite commune située au nord de Nancy, détruite en 1985 mais très connue dans la région et au-delà. Les entretiens spontanés ont été filmés aux domiciles des participants. Ces films ont ensuite été visionnés par les salariés en insertion de l'association pendant les séances de formation linguistique en accord avec la formatrice. Le premier objectif de cette étude était de *faire parler* les salariés. Ces derniers ne sont pas habitués à parler, comme nous l'avons expliqué lors de la présentation du corpus ILO dans le chapitre 2 (en [2.4.](#)). Les thématiques liées au travail en usine, aux conditions difficiles, aux relations avec la hiérarchie, à la répartition des tâches ou encore aux modes de recrutement ont suscité un vif intérêt pour ces salariés en insertion. Certains avaient connu des situations proches et d'autres avaient entendu les récits de ces mêmes conditions faits par leurs parents ou grands-parents. Chacun des participants s'est exprimé, plus longuement que d'habitude selon la formatrice qui travaillait avec eux depuis plusieurs mois. Cet espace de discussion s'est réalisé en groupe, chacun a expliqué son point de vue, raconté des situations proches de celles des témoignages ou a confronté son avis avec ceux de ses collègues⁴².

certaines enregistrements l'illustrant sont disponibles. Un guide pratique qui permet de reproduire ce travail a également été élaboré et mis en ligne.

⁴² Toutes ces interactions ont été enregistrées, transcrites et versées au corpus ILO.

Ensuite, nous avons proposé aux participants de rédiger un texte dont les destinataires ont été choisis lors d'une discussion collective. Unanimement, ils ont décidé d'écrire un texte pour les anciens salariés de la tuilerie dont ils avaient visionné les témoignages. Ils ont souhaité leur faire part de leur admiration face aux conditions difficiles que leur imposait leur activité au sein de la tuilerie. Ces textes ont été dictés par les participants mais écrits par moi, ma collègue ou la formatrice. Chaque participant a collaboré à un texte dont la structure a été négociée lors des séances de travail. Ces dernières ont pointé ce qui pouvait s'écrire et quelles formes étaient appropriées. Je ne détaillerai pas précisément cette partie de l'expérimentation pour me concentrer sur les aspects oraux de l'étude. Je vais souligner les difficultés rencontrées lors de ce passage à l'écrit. La mise en mot d'un récit, destiné à être écrit, a été rendue ardue notamment à cause de l'absence de lien logique entre les propositions énoncées par les participants, des anaphores floues, de l'absence de concordance des temps, des implicites ou des inférences irréalisables par un lecteur. Toutefois, le fait de discuter, en tant que rédacteur ou « teneur de stylo », avec les participants accroît leurs compétences socio-interactionnelles, même si la discussion n'améliore pas nécessairement les compétences scripturales. Avant même de penser au texte, les participants ont verbalisé ce qu'ils souhaitaient exprimer. Cette phase de verbalisation s'est révélée compliquée parce qu'elle demandait une mise en mots décontextualisée destinée à un auditeur/lecteur « naïf », c'est-à-dire n'ayant pas d'information sur la situation relatée.

Le travail effectué avec les participants peut être analysé grâce au cadre sociolinguistique des interactions verbales. En effet, pour faciliter la verbalisation appropriée du récit ou l'utilisation de ressources interactionnelles adéquates, les éléments du schéma d'analyse présenté dans le chapitre 1 ont été sollicités et se sont révélés utiles. Ainsi, j'ai pointé (en collaboration et négociation avec les participants) les éléments saillants à prendre en compte et sur lesquels s'appuyer pour produire un récit intelligible. Le premier point identifié comme important, au cours des discussions pendant la séance, est le genre de discours. Ensuite, j'ai interrogé les participants sur les activités qu'ils souhaitaient réaliser en insistant sur celles qui peuvent apparaître dans le genre de discours choisi, ici une lettre, et celles qui en sont exclues. Les destinataires ont été identifiés, ainsi que les objectifs de la communication, les thèmes à aborder, *etc.* afin de produire des énoncés qui s'appuient sur tous ces éléments et qui y sont appropriés. C'est la compréhension de la nécessité de prendre en compte ces éléments qui a permis aux participants de construire un discours

intelligible. J'y reviendrai ultérieurement mais l'élément le plus compliqué à prendre en compte concerne les connaissances partagées entre le scripteur et le lecteur. Il est difficile pour des individus en insécurité langagière de saisir le besoin de contextualisation de leurs paroles ou de leurs écrits. Cette difficulté est souvent liée à des problèmes de décentration, c'est-à-dire que des éléments nécessaires à l'intercompréhension ne sont pas suffisamment explicités. En outre, il est possible de faire le lien entre les rapports à l'écrit de ces individus, en situation d'illettrisme, et leurs insécurités langagières à l'oral. Ces différentes expériences et ces différents résultats semblent rejoindre les travaux de Goody (1979, 2007) ou de Terrail (2009) sur l'impact de l'écriture et des représentations graphiques sur les compétences cognitives et sur le développement de la pensée. Les exercices que nous avons réalisés lors de cette expérimentation ont tenté de contribuer au développement d'« une mise en mots » et d'une capacité de décentration ou d'abstraction, qui manque généralement aux locuteurs en insécurité langagière et qui est souvent à l'origine des problèmes socio-interactionnels de ces derniers. Concrètement, ces locuteurs éprouvent parfois des difficultés à faire correspondre les informations qu'ils communiquent avec les connaissances de leurs interlocuteurs. Par exemple, des implicites difficiles à inférer peuvent être à l'origine des problèmes dans les interactions. C'est également ce que va révéler l'analyse présentée dans le point suivant.

3.2.3. Interactions verbales avec des locuteurs en insécurité langagière

Dans le cadre de l'expérimentation présentée ci-dessus, plusieurs étapes ont été mises en œuvre afin de tester des activités de remédiation en français avec un public en situation d'illettrisme. Dans le dernier temps de l'expérimentation, nous avons demandé aux participants de faire un témoignage, à leur tour et à l'instar des anciens salariés de la tuilerie dont ils avaient visionné les témoignages. Pour les aider, nous leur avons préalablement demandé la thématique qu'ils souhaitaient aborder dans ces témoignages. Les thèmes choisis étaient généralement leurs passions (musique, couture, défense des animaux, etc.). Nous avons tout d'abord réalisé des entretiens filmés avec des locuteurs qualifiés, scolarisés au minimum jusqu'à l'obtention du baccalauréat, sur ces mêmes thématiques (parfois des témoignages sujets à controverse, par exemple le témoignage d'un chasseur diffusé à quelqu'un qui défend la cause animale). Nous avons diffusé ces

entretiens aux participants puis échangé avec eux à propos des entretiens. A la fin de la diffusion et de la discussion autour de ces entretiens, nos participants se sont volontiers pliés à l'exercice ayant pour objectif affiché de recueillir de la parole de différents locuteurs pour enrichir un corpus oral. Ce sont ces récits, élaborés en interaction, qui ont attiré mon attention.

Ces récits oraux sont parfois difficilement compréhensibles parfois même complètement incompréhensibles alors qu'ils abordent des thématiques bien connues des participants. Les récits choisis par les locuteurs, co-construits avec l'interlocuteur (chercheuses, étudiante ou formatrice), sont parfois tellement ancrés dans les connaissances du locuteur qu'il est difficile de les comprendre, même en le questionnant. Pour illustrer ces problèmes de compréhension, je propose d'examiner l'exemple suivant, extrait d'une interaction avec un locuteur faiblement qualifié, de niveau 4^{ème} collège (infra V, selon les niveaux de l'INSEE) et en situation d'illettrisme. L1 s'apprête à clore l'interaction mais L2 souhaite ajouter quelque chose et relater une anecdote. L'extrait présente d'abord un long tour de parole de L2 (lignes 2 à 12) qui répond à L1 (ligne 1) puis un passage plus interactif entre les deux locuteurs (lignes 13 à 32).

Exemple : Interaction entre un chercheur (L1) et un locuteur de niveau 4^{ème} (L2).

- | | | |
|----|----|---|
| 1 | L1 | il y a quelque chose à ajouter *P* {rire} |
| 2 | L2 | euh ouais je veux bien parce que moi il y a une grande histoire dans ma |
| 3 | | famille c'est euh au niveau de la musique c'était pendant la guerre c'était |
| 4 | | euh le père de mon père + il était dans les camps de concentration euh sa |
| 5 | | famille ils ont été dans les camps de concentration et euh {ffff} les Allemands |
| 6 | | lui demandaient si il connaissait une musique mais il connaissait pas puis on |
| 7 | | lui demandait souvent si il connaissait des gens qui savaient jouer la musique |
| 8 | | là et puis on leur disait si vous savez la faire bien on vous libère + puis il(s) |
| 9 | | connaissai(en)t et puis plus tard euh c'est ma mère qui m'a expliqué bien sûr |
| 10 | | euh on les a renvoyés vers des personnes qui étaient des cousins de ma mère |
| 11 | | euh des lointains cousins ils ont joué la musique et puis ils ont été libérés |
| 12 | | parce qu'ils ont ils ont arrivé à jouer |
| 13 | L1 | c'était les Allemands qui voulaient une musique en particulière |
| 14 | L2 | oui une musique précise ouais + |
| 15 | L1 | avec euh votre du coup votre euh c'était votre |
| 16 | L2 | ouais de la famille |
| 17 | L1 | euh votre grand-père votre arrière-grand-père qui |
| 18 | L2 | c'est ça ben c'est des cousins de ma mère mais lointains ça j'en ouais ça fait |
| 19 | | quand même un certain temps ouais |
| 20 | L1 | tout le monde jouait déjà de la musique et euh |
| 21 | L2 | oui chez moi c'est assez c'est tout ce qui est guitare accordéon violon + |
| 22 | L1 | et vous savez c'était quoi la musique qu'ils voulaient qu'elle qui soit jouée |
| 23 | L2 | ah non j'aimerais bien savoir mais non + |
| 24 | L1 | ils ne veulent pas vous le dire ou ils ne s'en souviennent pas |
| 25 | L2 | non ils se ils le savent pas |

- | | | |
|----|----|--|
| 26 | L1 | ils connaissent juste le rythme euh |
| 27 | L2 | même pas c'est le c'est une histoire quand comme je l'ai racontée mais |
| 28 | | c'est tellement loin qu'on se souvient pas personne |
| 29 | L1 | et ils ont pu être libérés grâce à la musique |
| 30 | L2 | grâce à une musique + |
| 31 | L1 | ouais du coup c'est une belle histoire |
| 32 | L2 | hum |

Dans cet extrait, un certain nombre de constructions s'éloigne de la norme. Cependant, au-delà de ces écarts à la norme assez fréquents et qui passent souvent inaperçus en interaction, d'autres éléments nuisent à la compréhension détaillée du récit. En effet, si l'histoire se comprend dans sa globalité, les détails sont plus difficiles à saisir. Par exemple, l'absence de lien logique et temporel empêche de reconstituer précisément l'histoire racontée. Les protagonistes sont posés mais il est difficile pour l'interlocuteur de saisir exactement quels sont leurs liens et leur implication dans l'événement. Les anaphores floues sont également à l'origine de la confusion entre les différents acteurs. Les liens entre les personnages et le narrateur ainsi que les liens éventuels entre les Allemands et les cousins sont flous et difficilement identifiables. A partir de la ligne 13, L1 tente de négocier la compréhension des détails du récit en posant des questions et en demandant des précisions à L2. La locutrice L1 a bien saisi l'histoire générale et réussit à en restituer l'essentiel à la fin de l'interaction : *et ils ont pu être libérés grâce à la musique*. Néanmoins, à la suite du témoignage, L2 avouera qu'elle n'a pas compris la totalité des détails de l'histoire racontée par L1, bien qu'elle ait tenté d'avoir davantage d'informations.

Dans cette étude, une distinction entre compréhension globale et compréhension détaillée est opérée. L'histoire est globalement comprise, c'est-à-dire que les éléments essentiels ont été repérés (lieu, moment, action). A partir de ce type de compréhension, il est possible de faire un bref résumé de l'histoire, comme l'a fait L1 à la fin de l'interaction. La compréhension détaillée consiste ici à saisir l'identité des personnages, leurs relations et leur implication directe ou indirecte dans l'histoire. Les détails de l'histoire doivent permettre également à l'interlocuteur d'inférer ce qu'il ne connaît pas. Ces inférences s'appuient sur les paroles du locuteur et sur les connaissances partagées avec son interlocuteur. Dans cet extrait, ce sont les détails de l'histoire racontée par L2 qui sont parfois difficiles à saisir.

Pour être certaine que la compréhension détaillée de cette interaction ne soit pas seulement problématique pour le locuteur L1 et pour moi-même, j'ai réalisé un test de compréhension auprès de vingt personnes. Ce test s'est déroulé en deux parties. Lors de la

première, j'ai diffusé le film de l'interaction des lignes 1 à 12 et j'ai demandé aux participants de me dire s'ils comprenaient le récit, ce qu'ils en comprenaient et s'ils pouvaient identifier les acteurs de l'histoire, c'est-à-dire qui sont les personnes évoquées et ce qu'elles font. Aucun des vingt participants n'affirme avoir compris totalement le récit, dans tous ses détails. Voici trois extraits des interactions que j'ai enregistrées lors des tests avec les participants :

Exemples : Extraits des interactions avec les participants au test (P)

- VA alors qu'est-ce que tu comprends de l'histoire qu'il nous raconte
P1 je suis pas sûr d'avoir tout compris euh une musique a libéré des gens de sa famille des camps de concentration mais je comprends pas plus

- VA est-ce que tu la comprends son anecdote
P19 plus ou moins je comprends pas bien comment on fait pour contacter les cousins mais à part ça euh oui si il parle de ses grands-parents oui
VA de ses grands-parents
P19 qui étaient en camp de concentration
VA donc c'est ses grands-parents pour toi qui étaient en camp de concentration
P19 ben c'est ce qu'il a eu l'air de dire ou alors j'ai zappé quelque chose
VA non non (...)
P19 par contre le fait d'aller chercher les arrières-petits-cousins pour aller qui connaissaient la musique là je comprends rien

- VA est-ce que tu comprends qui est qui dans cette histoire
P20 absolument pas au début je crois qu'il parle de son grand-père si j'ai bien compris mais en fait non parce qu'il parle pas du grand-père il parle de la famille du grand-père et après je suis j'ai un peu lâché l'affaire avec les cousins

Ces trois extraits montrent que les participants à cette petite enquête comprennent globalement le sens de l'histoire. Dans l'extrait 1, la question porte sur le sens global de l'histoire et la réponse montre que P1 l'a comprise. Dans l'extrait 2, à une question globale P19 tente de répondre sur les détails de l'histoire. Dans ce cas, P19 expriment ses difficultés à saisir les détails. Dans l'extrait 3, la question porte explicitement sur les détails de l'histoire et sur ses personnages (*est-ce que tu comprends qui est qui dans cette histoire*). P20 répond négativement en essayant de comprendre les liens familiaux. Lors de la seconde partie du test, je diffuse l'interaction des lignes 13 à 32. Même avec les questions posées dans cette partie de l'interaction, les vingt participants ne réussissent pas totalement à reconstituer les détails de l'histoire racontée par L2. Les participants avouent ne pas comprendre plus de détails après cette diffusion. De plus, malgré les cinquante secondes que dure la seconde

partie de cette vidéo, les participants au test se découragent rapidement et ne font pas autant d'efforts de compréhension qu'ils ont pu faire lors de la première partie. La majorité des participants s'excuse de ne pas comprendre davantage.

Cet exemple est probablement analysable des différentes façons. Il est possible que l'histoire racontée par le locuteur ne soit pas claire parce qu'il n'en connaît pas tous les éléments ou peut-être parce que c'est une légende véhiculée dans sa famille (on peut s'interroger sur la libération de déportés en camps de concentration parce qu'ils savaient jouer de la musique, cette question a d'ailleurs été soulevée par un des participants aux tests de compréhension). Il est possible également que les souvenirs du locuteur soient confus et que, par conséquent, il rencontre des difficultés à les restituer. Toutefois, il a choisi de faire le récit de cette histoire pendant cette interaction. L'objectif est de s'interroger sur les ressources mises en œuvre pour réaliser cette activité.

Les difficultés à mettre en récit une histoire, une anecdote ou une expérience semblent être liées au faible niveau de scolarisation des participants à ces études. La compétence d'expression à l'oral est encore le parent pauvre de l'enseignement. Comme l'explique Bautier (2016 : 96) :

la familiarité avec l'oral pour apprendre, pour (s')interroger et problématiser et ainsi élaborer une compréhension à valeur générique et conceptuelle, non pour simplement retenir des faits et les restituer, mais pour comprendre des phénomènes, compréhension donc fondée dans le travail avec le langage ne fait que rarement objet d'un apprentissage scolaire, d'un enseignement.

Ainsi, non seulement l'expression orale, qui permet de mettre sa pensée en mot, n'est que peu travaillée à l'école, mais elle est révélatrice d'inégalités entre les locuteurs qui n'ont souvent pas d'autres occasions de la travailler. Les participants à cette étude ont eu un parcours scolaire court et chaotique, ce qui raréfie davantage les occasions d'un travail sur l'expression orale. De plus, ils vivent dans des conditions économiques et sociales précaires, ce qui leur laisse peu la possibilité d'interagir dans beaucoup d'autres situations que celles qu'ils partagent avec leur famille ou leurs proches, vivant dans les mêmes conditions.

Dans un travail qui s'inscrit également dans le cadre de cette étude, Brié (2018) a montré que les interactions avec des locuteurs peu qualifiés pouvaient poser des problèmes quant au nombre important d'inférences demandées à l'interlocuteur. Son corpus est constitué d'interactions entre des locuteurs peu qualifiés, engagés dans une formation linguistique au sein de structures d'insertion (en Meurthe et Moselle et dans l'Hérault) et

un interlocuteur fortement scolarisé et n'étant pas en insécurité langagière. L'analyse de ce corpus montre que cet interlocuteur est contraint de réaliser de nombreuses inférences, parfois erronées, parce que ses interlocuteurs (peu qualifiés) ne donnent pas suffisamment d'informations contextuelles pour garantir l'intercompréhension des échanges. Comme nous l'avions déjà constaté avec Langbach (2014), les interactions sont co-construites de façon particulière, elles reposent notamment sur des ajustements liés à la vérification des inférences effectuées. Brié (2018) explique qu'un tiers des inférences erronées ne font pas l'objet de réaction, c'est-à-dire qu'elles ne sont pas infirmées ou réorientées par les locuteurs en insertion, ce qui nuit à la compréhension du contenu des interactions. Cette analyse confirme le manque de décentration et de capacité d'abstraction de ces locuteurs et les difficultés à évaluer ce qui doit être explicité pour garantir l'intercompréhension en interaction.

Cette étude et les différentes expérimentations qui s'y rattachent montrent, néanmoins, que les problèmes de maîtrise de la langue rencontrés par ces individus sont souvent liés à une non prise en compte du contexte, c'est-à-dire des éléments de la situation de communication. Dans le cadre d'une méta-analyse, Eme *et al.* (2011) pointent des résultats intéressants. Ils décrivent plusieurs études qui montrent une analogie entre des difficultés en lecture (restitution d'un texte) et en expression orale. Les participants ont tendance à restituer des textes lus en s'appuyant sur leurs propres connaissances plutôt que sur le sens du texte. En expression orale, un constat proche peut être établi, les événements sont racontés sans véritable contextualisation pour l'interlocuteur, qui est censé savoir ce que le locuteur sait également (Eme *et al.* 2009). Ce procédé rend peu compréhensible la restitution des textes lus, comme celles d'événements racontés à l'oral. Cette absence de contextualisation semble effectivement être une trace de l'insécurité langagière. Par ailleurs, le corpus ILO en cours de constitution permettra de confirmer ou d'infirmier ce type d'analyse.

Ces expérimentations mettent au jour les liens entre les interactions et le contexte social dans lequel elles prennent place. Par la même occasion, elle révèle l'importance de la sociolinguistique et de ses implications. Les actions menées et les démarches proposées relèvent de ce que l'on peut appeler la sociolinguistique impliquée (Adami, André 2012a). En effet, cette démarche « intègre dans son processus de construction épistémologique, dans ses hypothèses de travail, l'éventualité d'une implication possible sur un terrain

d'expérience. Le travail théorique n'est donc plus conçu comme une fin en soi mais comme une référence capable de fournir un cadre en mesure de répondre à d'éventuelles demandes sociales » (*Ibid.* : 66). Les problèmes de maîtrise de la langue identifiés sur mes différents terrains de recherche représentent un « handicap linguistique global » (Bentolila 1996). Ces problèmes relèvent de la sociolinguistique dans la mesure où ils concernent les usages de la langue dans la cité et la possibilité d'être citoyen.

3.3. Interface sociolinguistique et didactique

L'analyse sociolinguistique des interactions verbales met en évidence la façon dont les locuteurs utilisent des ressources interactionnelles, langagières et non langagières, pour réussir à communiquer en coopérant avec des interlocuteurs dans des situations particulières. Elle montre que les ressources utilisées s'appuient sur différents éléments pour être appropriées à la situation et pour accomplir les activités interactionnelles visées. J'ai montré que certaines pratiques interactionnelles inattendues, c'est-à-dire non décrites par les grammaires, pouvaient accomplir certaines activités avec succès et obtenir les effets pragmatiques escomptés. Ainsi, les ressources mises en œuvre par les locuteurs pour arriver à leurs fins peuvent être infinies et non canoniques. Malgré ces analyses, la langue orale enseignée à des apprenants de français est assez éloignée de celle qui est utilisée par les locuteurs et, plus gênant, cette langue ne permet pas toujours aux apprenants d'interagir de façon satisfaisante dans toutes les situations de communication qu'ils rencontrent. Cette langue ne permet donc pas l'émancipation des locuteurs et leur autonomie interactionnelle.

Dans les années 1970, d'autres s'intéressaient à ce problème dans le cadre de l'enseignement du français langue maternelle (FLM). Ce problème est finalement récurrent, et encore actuel. C'est celui du rapport à la norme. En effet, le premier numéro des *Cahiers de Linguistique Sociale*, publié en 1976 les interventions d'une journée d'étude, s'intitule *La Norme*, et tente de saisir les liens entre norme et enseignement. Cette journée avait été organisée par le CALEF, le Centre d'Application de la Linguistique à l'Enseignement du Français, sous la direction de Christiane Marcellesi. Le nom de cette équipe résonne pour moi comme une évidence, même si mes travaux tentent de s'appliquer à la didactique de l'oral en FLE. Marcellesi (1976) explique parfaitement que les enseignants de français à

l'école primaire, en suivant les manuels, sont motivés par l'enseignement de la grammaire et surtout de l'orthographe. Cependant, elle pose cette alternative éclairante :

Or s'agit-il d'enseigner la grammaire uniquement pour apprendre l'orthographe à l'enfant, pour lui apprendre à « bien » écrire, et subsidiairement à « bien » parler, ou pour le doter d'un instrument qu'il aura appris à faire fonctionner, qui lui permettra de s'exprimer en toutes occasions, en toutes situations, instrument de libération pour un individu inséré dans les luttes qui, dans notre société, opposent les classes entre elles. Apprendre à parler, faire fonctionner la langue, c'est déjà se libérer (ou du moins c'est s'en donner le moyen, c'est se doter d'une arme). (*Ibid.* : 9).

Cette question est toujours d'actualité, plus de quarante ans plus tard. De plus, la langue proposée aux enfants francophones natifs ressemble à la langue proposée aux apprenants de FLE, encore aujourd'hui. Elle est orientée par la norme, la grammaire et l'orthographe. Cette question se pose également avec des adultes francophones natifs en insécurité langagière. Le principal problème, nous y reviendrons, est leur incapacité à « faire fonctionner la langue » dans toutes les situations de communication qu'ils rencontrent. Cette question de la norme rejoint celle des compétences socio-interactionnelles qui permettent aux locuteurs d'interagir de façon appropriée aux situations de communication auxquelles ils sont amenés à participer. Marcellesi cite Genouvrier et je m'y associe également :

Il s'agit de mettre précisément l'enfant en mesure de choisir sa langue en lui apportant ce que son milieu socio-culturel n'a pu lui donner. Car encore faut-il, avant de parler de niveaux de langue, que l'élève en ait plusieurs à sa disposition. Démocratiser l'enseignement du français, ce n'est, ni rejeter la langue soutenue, gage d'accès aux richesses culturelles (et j'ajouterai « d'accès à nombre d'emplois sur le marché du travail »), ni exclure les autres, dont celles de l'élève lui-même : mais faire en sorte que le plus grand nombre accède au privilège qui est le nôtre : savoir varier nos registres et apprécier en conséquence le discours d'autrui. (*Ibid.* : 10).

Je terminerai les références au FLM en faisant cette fois référence aux travaux de Jean-Baptiste Marcellesi qui souligne : « Enseigner le français c'est d'abord connaître, non pas seulement ni tellement dans sa forme standard surtout écrite, mais dans ses variations géographiques, sociales et stylistiques (au sens labovien du terme). » (Marcellesi 1980 : 14). Ces problématiques du FLM mettent en évidence l'importance de l'exploitation de la sociolinguistique à des fins didactiques, que ce soit pour des enfants et des adultes francophones natifs ou des adultes allophones qui apprennent le français. Elles soulignent

l'importance accordée à l'émancipation des locuteurs, quelle que soit la langue dans laquelle ils s'expriment.

Après ce détour par le FLM, je vais revenir à la didactique des langues. La question de l'émancipation des locuteurs est tout aussi importante. Être capable de faire face de façon satisfaisante, c'est savoir se comporter de façon appropriée à la situation de communication, c'est maîtriser les usages et les variantes de la langue. Pour illustrer l'indispensable connexion entre sociolinguistique et didactique, je vais évoquer rapidement quelques-uns de mes travaux qui m'ont conduite à un plaidoyer en la faveur d'une interface disciplinaire.

Tout d'abord, dans le cadre d'une étude de la place de l'anglais dans différentes situations de travail, j'ai également pu montrer l'importance d'établir des liens entre sociolinguistique et didactique et entre pratiques linguistiques et enjeux humains et sociaux. « La domination symbolique et pratique de l'anglais, langue officieuse de la mondialisation capitaliste, n'est pas un épiphénomène mais un élément majeur conséquences humaines et sociales importantes. » (André 2015b : 91). En effet, l'utilisation de l'anglais au travail place de nombreux salariés en insécurité langagière. Les situations de communication qu'ils rencontrent deviennent des obstacles à l'effectuation du travail. Les conséquences de l'imposition de l'anglais et de la valorisation de « plurilinguisme », telle une doxa, conduisent certains salariés à subir de nouvelles souffrances au travail (Dejours 1998). L'analyse sociolinguistique des situations de travail permet, dans ces situations où l'anglais semble naturellement hégémonique, d'identifier les insécurités langagières, les souffrances au travail et, plus largement, de proposer une alternative à une domination linguistique. Cette domination est légitimée par la didactique du plurilinguisme, telle qu'elle est défendue au Conseil de l'Europe (Adami 2015, Gohard-Radenkovic *et al.* 2015), puisqu'on assiste actuellement à une didactique mondiale de l'anglais, sans se préoccuper des conséquences. La sociolinguistique a toute sa place pour démontrer les intrications entre décisions politiques, didactique et enjeux sociaux.

Ensuite, toujours pour illustrer cette nécessaire interface entre sociolinguistique et didactique, je vais faire un détour par le français langue d'intégration (FLI). Les premières réflexions que j'ai menées sur cette interface ont émergé avec la naissance et les travaux du groupe *Langage, Travail et Formation*. Pour le numéro zéro de la revue du même nom, j'expliquais déjà mon cheminement allant de l'analyse sociolinguistique des interactions aux applications sociales, sur le terrain de la formation linguistique (André 2015a). Après la

création des outils d'évaluation que j'ai mentionnés précédemment, la branche professionnelle des métiers de la propreté a fait une nouvelle demande à LTF, la réalisation d'un lexique pour les métiers de la propreté. La création de ce lexique a été entièrement réalisée à partir de corpus de différentes natures : de documents techniques, de fiches de poste, de manuels techniques, d'interactions verbales entre les salariés, etc. L'analyse sociolinguistique des interactions verbales a permis de saisir quels étaient les termes les plus utilisés dans les situations de communication les plus fréquentes ainsi que les termes indispensables à la pratique du métier. Un corpus représentatif des situations d'interaction rencontrées par les salariés visés par le lexique a été constitué afin de saisir la liste des mots essentiels. Un écart entre les mots prescrits par les manuels techniques et par les formateurs, et ceux réellement utilisés a été constaté. Nous avons opté pour un compromis entre *langue prescrite* et *langue réelle*, dont les définitions seront présentées en 4.2., dans la mesure où les accidents, les pannes et les événements imprévus sont rares mais importants et peuvent solliciter des mots moins fréquents mais tout aussi essentiels dans les situations professionnelles. En bref, ce travail lexical a été effectué dans le cadre de l'analyse sociolinguistique des interactions (en ce qui concerne l'analyse du corpus d'interactions), avec une visée didactique puisque le lexique devait faciliter l'acquisition des mots et qu'il a été créé comme support didactique pour des formations techniques et professionnelles⁴³.

Plus globalement, le champ de la didactique qui s'intéresse à la formation linguistique pour les migrants ou de qu'on appelle désormais le français langue d'intégration (FLI) nécessite des liens étroits avec la sociolinguistique. Tout d'abord, les enjeux d'une telle formation sont contextuels et politiques, les nouvelles vagues migratoires en Europe entraînent une nouvelle réalité sociale dans le champ de la formation linguistique des migrants (Adami, André 2013). Ensuite, les vagues migratoires en France sont caractérisées par le faible niveau de scolarisation voire par l'analphabétisation de certains migrants. Enfin, les objectifs des formations doivent être situés et doivent permettre aux migrants d'interagir dans les situations de communication qu'ils vont rencontrer pour assurer leur vie, voire leur survie, dans le pays d'accueil. Les enjeux de ces formations visent l'intégration citoyenne, sociale et professionnelle des migrants. Cette intégration permet une autonomisation et une émancipation de ces derniers. En termes didactiques, pour atteindre

⁴³ Voir le lexique : <https://perso.atilf.fr/VirginieAndre/wp-content/uploads/sites/47/2019/07/Lexique-LesMotsClesdelaProprete.pdf>

ces objectifs, il faut apprendre à interagir de façon appropriée. D'un point de vue méthodologique, cela signifie que les formateurs soient capables d'analyser des interactions et d'identifier les compétences socio-interactionnelles nécessaires pour y participer.

En termes sociolinguistiques, cela signifie s'appuyer sur les éléments du modèle présenté dans le chapitre 1 et sur les analyses qu'il permet d'effectuer. Il s'agit alors de faire prendre conscience de la variation à des apprenants de langue pour qu'ils acquièrent des compétences socio-interactionnelles. L'objectif de proposer aux apprenants d'acquérir une langue qui leur permettra d'interagir de façon appropriée et de s'autonomiser d'un point de vue interactionnel semble tout à fait légitime. C'est pourquoi, on peut s'interroger sur les raisons pour lesquelles les manuels de langue, même les plus récents, fournissent aux apprenants une langue construite qui n'existe pas en dehors du manuel (Giroud, Surcouf 2016). Une des hypothèses, celle que je formule ici, est le manque d'interface ou de communication entre les chercheurs des domaines concernés, sociolinguistique et didactique, ainsi qu'entre les chercheurs et les enseignants (Gilmore 2007).

Enfin, lors d'une étude menée avec Henry Tyne, nous avons montré que la compétence sociolinguistique ne repose pas sur ce que préconisent habituellement les manuels et certains enseignants de langue (André, Tyne 2012). En comparant des interactions entre des natifs avec des interactions entre un natif et un apprenant de FLE, nous avons montré que plus les apprenants sont avancés, plus ils utilisent des pratiques interactionnelles généralement proscrites par les manuels et les grammaires mais utilisées par les natifs. Par exemple, comme je l'ai déjà mentionné précédemment, l'analyse sociolinguistique des interactions montrent que les locuteurs construisent leur discours de façon collaborative. Cette pratique qui vise à s'appuyer sur le discours de l'autre ou à solliciter l'aide de l'autre pour élaborer son discours est fréquente dans les interactions. Cependant, les apprenants de langue apprennent à ne pas hésiter, ne pas piétiner, ne pas montrer d'indices de demande d'aide à la construction du discours et donc à construire seuls leur discours, sans s'appuyer sur l'interlocuteur. A ce propos, Pekarek-Doelher (2009 : 21) propose de déconstruire le mythe du « *'lonely speaker'*, soit du locuteur isolé dont la production relève de seule responsabilité ». Les constructions collaboratives, telles que les énonciations conjointes (André 2014), ne sont pas évaluées comme relevant de la compétence interactionnelle des apprenants. Les hésitations, par exemple, qui peuvent être analysées comme des pauses sonores et des traces de la sollicitation de son interlocuteur, sont considérées comme des *dysfluences*, c'est-à-dire comme des dysfonctionnements ou

des manques de compétences. Ainsi, le manque de liens entre les analyses interactionnelles et la didactique des langues conduit à une conception erronée de la langue, de la part des enseignants. Le discours en interaction est co-construit quel que soit le statut – natif ou non natif – du locuteur. La prise en compte de l'interaction est indispensable, aucun locuteur ne produit un monologue continu, sans les traces de l'élaboration de son propre discours et de sa collaboration avec les interlocuteurs (Chambers 1997, Guillot 1999, McCarthy 2010). Ce qui est identifié comme une *dysfluence* est, si on se réfère à l'analyse des interactions, en réalité une marque de la compétence socio-interactionnelle de l'apprenant. C'est en établissant des liens entre les analyses d'interactions et la didactique qu'il est possible d'évaluer réellement les compétences des locuteurs-apprenants.

Certaines initiatives ont déjà été menées afin d'exploiter les analyses sur corpus à des fins didactiques. Parmi les projets les plus connus et les plus anciens, après le Français Fondamental (Gougenheim *et al.* 1964), la constitution du premier corpus ESLO et le dictionnaire COBUILD sont probablement les pionniers. Tout d'abord, dès 1968, des enseignants de français du Royaume-Uni ont mené les Enquêtes Sociolinguistiques à Orléans (ESLO) afin de constituer un corpus dans lequel il était possible de puiser les usages réels de la langue parlée à des fins didactiques, c'est-à-dire pour les enseigner à des apprenants britanniques de français. L'objectif était de profiter des nouvelles technologies (enregistrements, stockage, diffusion en centre de langues) pour moderniser les méthodes pédagogiques de l'enseignement du FLE. Que cette initiative vienne d'outre-manche n'est pas vraiment étonnant, l'oral ne bénéficie pas d'une mauvaise presse comme en France (Lotfi, Baude 2005). Ensuite, en 1987 un dictionnaire entièrement basé sur corpus voit le jour à l'université de Birmingham, c'est le *Collins COBUILD English Language Dictionary* sous la direction de John Sinclair. Il est composé d'un lexique de base destiné à des apprenants de la langue anglaise. Enfin, un peu plus récemment, une grammaire à partir de corpus a vu le jour avec le *Longman Grammar of Spoken and Written English* (Biber *et al.* 1999). Malgré ces initiatives pionnières, peu de travaux s'inscriront ensuite dans cette perspective. J'en aborderai quelques-uns dans les chapitres suivants avant de présenter mes propositions d'exploitations didactiques de la sociolinguistique des interactions verbales.

L'interface entre sociolinguistique et didactique peut se manifester concrètement par ce que l'on peut appeler une « transposition didactique ». L'analyse des interactions est nécessairement utile à l'enseignement et à l'apprentissage de la langue. Néanmoins, les apprenants et les enseignants ne sont pas des scientifiques de la langue. Par conséquent,

une transposition didactique permet une exploitation des recherches à des fins d'enseignement et d'apprentissage. A ce propos, Debaisieux (2000 : 148) s'excuse de la transposition didactique qu'elle propose : « Je suis consciente qu'elle pourra être décriée, à la fois par les linguistes, à juste titre, puisque les analyses y sont nécessairement « déformées » du fait même de la transposition, et par les didacticiens, au nom de l'autonomie de la discipline. ». Cependant, la didactique de l'oral a besoin d'oral et des analyses de l'oral. Cet oral est à la fois difficile à consulter et à manipuler, et à la fois complexe à manipuler à des fins d'enseignement et d'apprentissage de la langue. C'est pourquoi, les exploitations didactiques des corpus d'interactions ainsi que des résultats des recherches qui leur sont consacrés méritent d'être mises au jour.

Conclusion du chapitre

Créer des liens entre des disciplines, afin qu'elles puissent se nourrir mutuellement et qu'elles puissent être utiles à la société et répondre à des préoccupations ou des demandes sociales, n'est pas toujours bien considéré, à la fois par les chercheurs attachés à leur objet d'étude qu'ils estiment « pur » et par les institutions attachées à ce même purisme. Dans le domaine des Sciences du langage, la sociolinguistique, parce qu'elle s'intéresse à d'autres éléments que la langue subit parfois une légère déconsidération. Alors que la langue est une pratique sociale, finalisée, collective et située. Son étude demande de saisir à la fois les influences qu'elle reçoit et les influences qu'elle opère lors de sa mise en pratique en interaction.

Cette double articulation, entre influences subies et provoquées, confère aux interactions verbales une grande complexité. Les interactants mettent en œuvre une infinité de ressources pour participer à des interactions. Ces ressources sont *sous influence*, c'est-à-dire qu'elles sont entremêlées à la situation de communication, à d'autres pratiques sociales, finalisées et collectives. Elles permettent aux locuteurs d'être acteurs dans leur société, à condition qu'ils soient en *sécurité langagière*, c'est-à-dire qu'ils puissent interagir de façon efficace dans les situations auxquelles ils participent. En revanche, certains locuteurs se trouvent en *insécurité langagière* (Adami, André 2014), c'est-à-dire qu'ils ne réussissent pas à interagir de façon appropriée dans toutes les situations de communication qu'ils rencontrent. En d'autres termes, ces locuteurs produisent des pratiques et des

activités interactionnelles sans prendre suffisamment en compte le genre de discours auxquels ils participent et les éléments de la situation. La sociolinguistique sait décrire et analyser les interactions verbales, leur fonctionnement, leur dynamique, leur nature, leur forme et leurs enjeux. Ces descriptions et analyses peuvent être exploitées à des fins didactiques afin de mesurer et analyser les compétences socio-interactionnelles des individus puis d'essayer de les conforter ou de les améliorer.

J'ai mentionné le jugement sévère des linguistes dans le premier chapitre. Celui-ci s'adresse à l'égard de la sociolinguistique, parfois, comme je l'ai expliqué. Il est également parfois réservé un mauvais sort à la didactique des langues. Que l'on pense ou que l'on envisage, ou pas, la didactique comme une application, elle n'en demeure pas moins une discipline qui se dote d'une méthodologie scientifique. Je n'entrerai pas dans le débat qui consiste à déterminer si la didactique est une discipline autonome ou si elle est seulement le fruit de réflexions menées dans d'autres disciplines (psychologie, sociologie, philosophie, ergonomie, sociolinguistique, *etc.*). Néanmoins, elle comporte des objets d'étude qui peuvent être expérimentés, analysés, théorisés, au même titre que d'autres disciplines scientifiques. Dans le cadre d'un travail associant déjà sociolinguistique et didactique, avec Hervé Adami, nous proposons une approche sociolinguistique impliquée (Adami, André 2012a). Cette implication a tout à fait sa place dans une démarche scientifique rigoureuse. Cette approche m'a conduite à envisager des exploitations didactiques de l'analyse sociolinguistique des interactions verbales. Elles feront l'objet des deux prochains chapitres. Le chapitre 4 mettra au jour l'utilité de l'analyse, séquentielle ou tour par tour, des interactions pour l'enseignement et l'apprentissage de la langue. Le chapitre 5 exploitera un outil de la sociolinguistique de corpus, le concordancier, à des fins didactiques. Je montrerai que ces exploitations sont complémentaires et peuvent également faire appel à d'autres ressources interactionnelles pour une plus grande complémentarité et une meilleure exploitation de l'analyse des interactions en didactique du FLE.

Chapitre 4. L'analyse sociolinguistique des interactions verbales au service de l'enseignement/apprentissage de la langue

Ce travail de synthèse est une sorte de plaidoyer en faveur d'un renforcement, voire d'un établissement parfois, de liens entre sociolinguistique et didactiques des langues. Pour enseigner une langue, savoir la décrire est indispensable. Cependant, il ne suffit pas de faire une analyse formelle décontextualisée ou désincarnée. L'analyse des interactions telle que je l'ai présentée précédemment contribue nécessairement à l'enseignement et à l'apprentissage de la langue cible. Je reviendrai plusieurs fois sur le fait que certaines de mes analyses interactionnelles ont été guidées par des demandes de la part d'enseignants de FLE, et parfois d'apprenants. Les réponses que peuvent apporter les analyses sociolinguistiques sont essentielles aux formations en langue et à l'appropriation de la langue.

Tout d'abord, en s'inscrivant dans une approche sociolinguistique de la langue, il va de soi que l'approche didactique correspondante doit présenter aux apprenants la langue telle qu'elle est réellement utilisée par ses locuteurs, avec ses variations situationnelles. Nous verrons, même si ces constatations ont déjà été faites par ailleurs, que les descriptions faites par les grammaires et les manuels de langue sont encore en vigueur alors qu'elles sont souvent inappropriées et s'éloignent des usages effectifs de la langue. Par exemple, les approches communicative puis actionnelle, telles qu'elles sont proposées par le Cadre Commun Européen de Référence en Langue (CECRL 2011)⁴⁴, mettent l'accent sur l'enseignement et l'apprentissage de l'interaction. L'importance de la composante interactionnelle fait partie du référentiel, il est par exemple précisé, à propos de la perspective actionnelle :

Dans l'interaction, au moins deux acteurs participent à un échange oral et/ou écrit et alternent les moments de production et de réception qui peuvent même se chevaucher dans les échanges oraux. Non seulement deux interlocuteurs sont en mesure de se parler mais ils peuvent simultanément s'écouter. Même lorsque les tours de parole sont strictement respectés, l'auditeur est généralement en train d'anticiper sur la suite du message et de préparer une réponse. Ainsi, apprendre à interagir suppose plus que d'apprendre à recevoir et à produire des énoncés. On accorde généralement une grande importance à l'interaction dans l'usage et l'apprentissage de la langue étant donné le rôle central qu'elle joue dans la communication (CECRL : 18).

⁴⁴ Disponible en ligne : <https://rm.coe.int/16802fc3a8>

De plus, il est fait mention à plusieurs reprises de certains aspects sociolinguistiques de l'interaction tels que les marqueurs de relations sociales, les règles de politesse ou encore les accents régionaux. Cependant, au-delà de ces recommandations qui devraient permettre l'acquisition d'une compétence sociolinguistique, peu d'éléments pratiques sont proposés. D'autres initiatives plus concrètes sont entreprises autour du CECRL. Par exemple, certains travaux proposent de donner accès aux cadres théoriques et aux concepts qui devraient être mobilisés pour respecter le référentiel européen et ses niveaux de langue. Ainsi, dans Traverso (2004a et 2004b), le lecteur peut trouver une description détaillée des discours oraux interactifs et de la conversation à bâtons rompus. Néanmoins, les descripteurs du CECRL, même dans sa dernière version de 2018, ne donnent pas d'information précises sur le fonctionnement des interactions en situation. Les recherches sur les interactions verbales sont rarement pointées et encore moins exploitées. En outre, certaines pratiques interactionnelles ne sont parfois simplement pas décrites, ni dans les manuels ni par les enseignants en formation linguistique, j'en examinerai certaines.

Ce chapitre sera consacré à l'apport de l'analyse sociolinguistique des interactions verbales à l'enseignement et à l'apprentissage du français en tant que langue étrangère. J'évoquerai l'indispensable exposition à la langue pour un apprenant et certains travaux qui mettent en œuvre des méthodologies en accord avec ce principe (4.1.). Je mentionnerai quatre projets qui proposent une exploitation de corpus oraux et multimodaux à des fins didactiques. Ensuite, j'emprunterai à l'analyse du travail ses définitions de l'activité pour proposer une nouvelle distinction et certaines caractéristiques de la *langue prescrite* et de la *langue réelle* (4.2.). Dans cette perspective, je proposerai d'analyser certaines pratiques interactionnelles réelles peu ou pas évoquées dans les manuels ou dans les formations en FLE alors qu'elles apparaissent dans des corpus d'interactions. J'illustrerai ce constat avec des exemples de pratiques interactionnelles analysées en interaction et pouvant faire l'objet d'une transposition didactique. Enfin, je présenterai de quelle façon il est possible d'exploiter les analyses sociolinguistiques des interactions à des fins didactiques en examinant plusieurs expérimentations réalisées avec le corpus FLEURON (4.3.). Ces expérimentations montrent les avantages de ces analyses interactionnelles pour l'appropriation de la langue.

4.1. Exposition des apprenants à des interactions verbales authentiques

En arrivant au CRAPEL, comme je l'ai mentionné en introduction, je me suis rapidement trouvée immergée dans un cadre scientifique didactique en résonnance avec mes propres travaux sur l'analyse des interactions verbales. L'utilisation des documents authentiques dans les formations de formateurs en langue et la théorisation de cette pratique méthodologique par les membres de l'équipe (Duda, Esch, Laurens 1972 ; Abé, Carton, Cembalo, Régent, 1979 ; Holec 1990) m'ont paru comme aller de soi tellement il me paraissait évident de donner à voir et d'enseigner une langue, et ses usages, intriquée à son contexte social. Pour rappel, un document authentique est un « document produit à des fins autres que l'apprentissage d'une langue seconde » proposant « un énoncé produit dans une situation réelle de communication » (Abé *et al.* 1979 : 2). En justifiant l'utilisation des documents authentiques, Holec (1990 : 68) souligne : « si les discours auxquels l'apprenant est confronté ne sont pas sinon identiques tout au moins analogues à ceux auxquels il sera confronté dans les situations où il aura à utiliser ses savoir-faire acquis, il courra le risque d'être incapable de faire face de manière satisfaisante ». De plus, c'est en étant exposé à différentes situations de communication authentiques que les apprenants se sensibilisent à la variation.

Il m'a fallu quelques interventions didactiques auprès d'enseignants de langue pour me rendre compte que l'utilisation de documents authentiques n'était pas une pratique majoritaire. Les freins à l'utilisation des documents authentiques étaient à la fois individuels et institutionnels. La première formation à laquelle j'ai participé avec Sophie Bailly s'est déroulée dans les locaux de l'IUFM⁴⁵ et était destinée à des enseignants d'espagnol. Si nous avons réussi à les convaincre d'exposer leurs apprenants à des documents authentiques, ils nous ont rapidement fait comprendre que leur institution, l'éducation nationale, n'était pas prête à valider cette approche et qu'ils risquaient fortement d'être mal notés si, le jour de leur inspection, ils mettaient en œuvre nos recommandations. Ces échanges se sont révélés éclairants. Je pourrais également dire que ces échanges ont révélé ma naïveté quant aux pratiques pédagogiques des enseignants de langue. Ces mêmes constatations ont bien évidemment également été faites avec des enseignants de FLE, notamment lors de la formation des enseignants qui venaient en stage au sein de l'équipe ou encore lors des

⁴⁵ L'institut universitaire de formation des maîtres, devenu aujourd'hui l'école supérieure du professorat et de l'éducation (ESPE).

formations d'enseignants dans les centres de langues de Tunisie que j'ai réalisées avec Florence Poncet en 2008.

Si les travaux pionniers des collègues nancéens étaient et sont parfois encore (plusieurs dizaines d'années après leurs publications) novateurs, ils étaient loin d'être mis en pratique sur le terrain de la formation en langue. Évidemment, pour l'analyste des interactions verbales que je suis, cette réticence à l'utilisation de documents authentiques était intrigante et me posait questions. Pourquoi ne pas enseigner une *vraie* langue ? c'est-à-dire une langue qui varie en fonction de facteurs extérieurs, sociaux ? Pourquoi se limiter à une seule variété de langue ? Et aussi et surtout, pourquoi enseigner l'oral comme si c'était de l'écrit oralisé ? Ces questions m'ont poussée à m'intéresser à la didactique de l'oral et aux applications pratiques – souvent négligées – des recherches en sociolinguistique.

Il est encore fréquent aujourd'hui de rencontrer des apprenants de FLE qui sont surpris par la langue qu'ils entendent lorsqu'ils arrivent en France ou dans un pays francophone. Certains sont même surpris par la langue qu'ils entendent à la radio ou à la télévision. Lors de diverses séances d'expérimentation, dont certaines sont décrites dans les sections suivantes, j'ai moi-même entendu des apprenants avouer que la langue que je leur proposais d'observer, dans différents corpus, était inconnue pour eux. Certains faisaient le (triste) constat que durant leurs années d'apprentissage du français (plusieurs années parfois), ils n'avaient jamais été exposés à ce type de langue, autrement dit à de l'oral spontané, quotidien ou ordinaire. L'anecdote racontée par Durán et McCool (2003), et reprise avec autant d'étonnement par Surcouf et Ausoni (2018), illustre parfaitement cette aberration. Les premiers expliquent que l'un d'entre eux (Durán) a fait la triste expérience de ne pas comprendre la réponse, qu'ils transcrivent « Chépas », d'une vendeuse dans un magasin. L'auteur explique : « No instructor in any of my French courses had ever alluded to the transformation of the formal *je ne sais pas* to the colloquial *chépas* (*j'sais pas*). I had taken three years of college French and, despite my success in those courses, I found my language preparation highly inadequate in certain contexts. » (Durán, McCool 2003 : 288). Cette incompréhension n'est pas exceptionnelle. La langue à laquelle les apprenants sont exposés, qu'elle soit issue des manuels ou qu'elle soit celle des enseignants, est bien souvent éloignée de la langue telle qu'elle est utilisée par des locuteurs natifs en situation de communication authentique non pédagogique (Holec 1990 ; Vialleton, Lewis 2014 ; Ravazzolo *et al.* 2015). En comparant les dialogues de manuels à des interactions authentiques, Parpette (1998) relevait déjà que les caractéristiques de ces dernières

n'apparaissent pas dans les premiers. Elle souligne « l'absence de dialogues porteurs des caractéristiques les plus fréquentes en situations authentiques » (*Ibid.* : 128). Elle remarque que les dialogues de manuels, contrairement aux interactions authentiques, présentent une forte unité thématique, un déroulement linéaire des échanges et une élimination systématique de l'implicite. Elle précise : « En termes d'enseignement, on peut craindre que cette uniformité ne permette pas vraiment de préparer les apprenants à la diversité des discours oraux rencontrés en situation réelle. » (*Ibid.*)⁴⁶.

Néanmoins, l'exposition à des situations de communication authentiques n'est pas toujours facile. Les enseignants se plaignent souvent du manque de ressources, de ne pas avoir le bon document, au bon moment, avec les bons objectifs de communication, les bonnes thématiques, etc. En effet, malgré les innombrables ressources que propose Internet, il n'est pas rare que la ressource authentique adéquate n'existe pas. Parmi les ressources manquantes, les enseignants signalent souvent celles qui présenteraient des situations de communication ordinaires, des conversations entre amis ou encore des interactions dans des cafés. De plus, les ressources les plus appréciées étant des vidéos, des corpus tels que TCOF ou ESLO sont rarement exploités pour extraire une conversation ordinaire qui pourrait servir de support d'enseignement.

Afin de trouver une solution à ce problème, j'ai proposé à titre expérimental d'utiliser une série télévisée comme corpus multimodal en classe de FLE (André 2018b). Après de longues discussions, avec des enseignants et des étudiants du Master de Didactique des langues, concernant la série qui serait la plus appropriée, c'est-à-dire la plus apte à proposer des dialogues proches des interactions verbales authentiques, j'ai décidé de travailler avec la série *Plus belle la vie*, diffusée sur France Télévision. Ce travail a d'abord consisté à faire une analyse des dialogues de la série et à les comparer à des interactions verbales authentiques (notamment celles du corpus TCOF). Conformément aux études, notamment, sociologiques qui ont été réalisées à propos de la série et à ce que revendiquent les producteurs et les scénaristes, certains épisodes sont très réalistes et proches de la réalité des pratiques interactionnelles. Les épisodes qui ont constitué mon corpus de travail ont été choisis parce qu'ils évoquaient des thèmes d'actualité et parce qu'ils faisaient partie des préoccupations thématiques des apprenants (étudiants au

⁴⁶ Par ailleurs, dans ce même article Parpette insiste sur la nécessité de faire des « liens entre Sciences du langage et didactique » (1998 : 133), pour améliorer les documents supports conçus à des fins didactiques.

Département de FLE). Chaque épisode, à l'instar du travail de Quaglio (2009) avec la série *Friends*, a été analysé d'un point de vue interactionnel. J'ai écarté du corpus des séquences analysées comme étant trop éloignées des interactions authentiques (pour plus de précisions, voir André 2018b). L'utilisation de ce corpus a soulevé de vives réactions en salle de classe, notamment avec un groupe d'étudiantes kazakhs qui a été le premier groupe à expérimenter l'utilisation de la série à des fins d'apprentissage de la langue. En effet, sans détailler précisément les activités réalisées avec le corpus, je peux souligner certaines réactions des apprenantes démontrant l'importance de l'exposition à la langue. Ces apprenantes, de différents niveaux (de A1 à C1 selon le CECRL), ont été surprises par la langue qu'elles ont entendue et ont affirmé que cette langue ne ressemblait absolument pas à celle qu'elles avaient apprise dans leur pays. En outre, elles ont précisé que la langue présentée ressemblait davantage à celle qu'elles entendaient depuis leur arrivée en France et qu'elles peinaient à comprendre.

Parmi les éléments relevés par les apprenantes, le marqueur interactionnel *tu sais* produit par les acteurs de la série a été un élément révélateur. Les apprenantes ont affirmé ne pas comprendre les *tu sais* dans le discours de leurs interlocuteurs en France et ont avoué qu'à chaque fois qu'elles réussissent à le comprendre, elles cherchent à produire une réponse mais aucun locuteur ne leur laisse le temps de le faire. Ce marqueur a été la source d'un travail avec un concordancier (selon la méthodologie du *data-driven corpus* exposée dans le chapitre suivant). Ce travail a permis aux apprenantes de saisir les occurrences en interaction. La confrontation avec un corpus, ici non authentique mais s'approchant de la réalité des interactions, améliore la compréhension orale des apprenants. Plus surprenant, certaines apprenantes, après qu'elles ont compris le fonctionnement de certaines pratiques, se sont données comme tâche de les utiliser de façon appropriée lors de leurs prochaines interactions exolingues. Cette tâche peut hypothétiquement être liée au fait qu'une des actrices de la série avait relativement le même âge que les apprenantes. Il n'est pas rare que les apprenants identifient des phénomènes interactionnels correspondant à une certaine caractéristique ou facette identitaire des locuteurs (Beeching 2007) et choisissent de se les approprier.

Parmi les pratiques interactionnelles repérées dans le corpus *Plus belle la vie*, je peux également mentionner : *regarde* dans des énoncés du type *ouais mais regarde Abdel il est pas comme ça ; alors* en début de tour de parole ; *hein* ou *quoi* en fin de tour ; des répétitions pour insister dans des questions comme *tu es sérieux tu es sérieux vraiment ; etc.* Toutefois,

cette expérimentation demande un lourd travail de transcription et d'analyse afin de sélectionner et de constituer le corpus. Lebon (2018) a reproduit ce même travail avec une série espagnole *Los Serrano* pour un public d'adolescents en collège. Elle constate également l'importance de cette exposition à la langue cible au travers de thématiques appropriées aux préoccupations des apprenants ainsi que l'importance d'être confronté à des pratiques interactionnelles très utilisées par les locuteurs mais peu travaillées en salle de classe. Toutes les deux, nous concluons également que ce travail demande beaucoup de temps et d'énergie alors que les corpus ne sont pas des corpus de situations authentiques, même s'ils s'en rapprochent.

Afin de pallier une présentation partielle et erronée de la langue parlée, plusieurs autres initiatives, reposant sur des corpus authentiques, ont vu le jour. Ces dernières permettent d'exposer les apprenants à la langue parlée ou à des enseignants de trouver des ressources authentiques. Ces nouvelles propositions didactiques exploitent des corpus oraux et multimodaux à des fins d'enseignement et d'apprentissage du français. Je vais présenter rapidement les projets PFC-EF, FLORALE et CLAPI-FLE ([4.1.1.](#)), qui réfléchissent à des applications didactiques de leur corpus, initialement constitués par des linguistes, avant de m'attarder plus longuement sur le projet FLEURON ([4.1.2.](#)) qui a constitué un corpus multimodal à des fins didactiques.

4.1.1. PFC-EF, FLORALE et CLAPI-FLE

La rencontre entre les corpus et la didactique semble aujourd'hui petit à petit devenir une évidence justement parce que certains corpus rassemblent et organisent des données langagières et interactionnelles intéressantes pour des apprenants de langue. Cependant, la rencontre n'est encore qu'un effleurement. Seules quelques initiatives ont vu le jour. Elles viennent généralement de sociolinguistes de corpus qui envisagent des utilisations de leur corpus pour l'enseignement du français. Je ne mentionnerai que trois initiatives dans cette partie, celles qui me paraissent les plus avancées en ce qui concerne l'établissement de liens entre corpus, enseignement et apprentissage : PFC-EF, FLORALE et CLAPI-FLE.

◆ PFC-EF

Le projet PFC (Phonologie du français contemporain) propose un grand corpus oral de français (Durand, Laks, Lyche 2002). Les données de ce corpus sont recueillies en suivant un protocole identique dans tous les pays du monde, afin d'être facilement comparables. En plus des mots prononcés et des textes lus, le corpus est constitué d'entretiens guidés et de conversations informelles. Depuis 2006, le projet PFC-Enseignement du français⁴⁷ est destiné à l'exploitation pédagogique du corpus PFC (Detey *et al.* 2009, Detey *et al.* 2010). Il met à la disposition des enseignants et des apprenants des ressources pour l'apprentissage de l'oral et pour la familiarisation aux variétés de français dans le monde. Le site Internet du projet propose des ressources linguistiques, des fiches explicatives, des séquences didactisées, des fiches pédagogiques ainsi qu'un moteur de recherche (concordancier) pour accéder au corpus. Les ressources didactiques disponibles dans le cadre de ce projet soulignent également les oppositions entre français parlé et français écrit ou entre normes et variations.

Plusieurs travaux sont consacrés à l'exploitation du corpus du projet à des fins d'amélioration de la prononciation des apprenants de FLE ainsi que de leur compréhension de la variation phonético-phonologique (Detey *et al.* 2009). Plusieurs phénomènes de l'oral (comme les phénomènes de schwa ou de liaison mais aussi des phénomènes syntaxiques, pragmatiques ou culturels) sont expliqués et illustrés avec les données du corpus afin d'aider les enseignants et les apprenants à saisir les caractéristiques des français parlés. Dans une perspective de sensibilisation à la variation phonétique, Detey (2010 : 162) souligne que

les données des corpus, en particulier PFC, pourraient offrir aux apprenants un input phonétique dont les caractéristiques pourraient assurer un apprentissage phonologique plus résistant à la variation d'une part et à l'influence des représentations orthographiques d'autre part, sous réserve, naturellement, d'être employé de manière adéquate.

Ainsi, le projet PFC-EF propose une exposition aux variations qui s'appuie sur des recherches pointues émanant des données fournies par les corpus afin de soutenir l'enseignement et l'apprentissage de l'oral en FLE.

Parallèlement, le projet IPFC (Interphonologie du français contemporain) recueille un corpus d'apprenants, de locuteurs non natifs du français afin :

⁴⁷ Voir à l'adresse suivante : <https://www.projet-pfc.net/le-projet-pfc-ef/>.

1) de documenter et d'étudier l'apprentissage du système phonologique du français par des apprenants de diverses L1 ; 2) de documenter et d'étudier la variation (...) ; 3) de proposer des activités autour de la prononciation du français et de son apprentissage ; 4) de promouvoir l'image du français langue internationale. (Racine *et al.* 2012 : 4).

Ce corpus permet également des comparaisons avec le corpus PFC ainsi que des comparaisons intra-corpus en étudiant les productions langagières de natifs de diverses origines. Dans le domaine de la didactique de l'oral, l'étude de l'interphonologie des apprenants de FLE montre notamment l'importance de l'exposition des apprenants à la langue cible ainsi que celle de l'exploitation des corpus à des fins didactiques. La prononciation est un aspect important de la compétence socio-interactionnelle dans la mesure où elle garantit l'intelligibilité et la compréhensibilité des discours et parce qu'elle est également une marque de l'identité des locuteurs, qu'ils soient natifs ou non natifs (Detey *et al.* 2016). Sensibiliser les apprenants et les enseignants de FLE à la prononciation en leur proposant des données en corpus et des analyses de ces données, permet d'aborder la notion de variation.

◆ FLORALE

Afin de proposer aux apprenants une exposition au français parlé, la ressource FLORALE ⁴⁸ (français langue orale et FLE) a été élaborée. Comme le montrent ses concepteurs, Surcouf et Ausoni (2018) en examinant la partie orale des manuels de FLE, la langue à laquelle les apprenants sont exposés est souvent la langue écrite, héritée de la grammaire traditionnelle et de la norme. Ils soulignent : « Aussi les descriptions pédagogiques en FLE en viennent-elles paradoxalement à priver l'apprenant d'un nombre important d'informations fondamentales sur le fonctionnement de la langue cible » (*Ibid.* : 82). Les auteurs défendent l'utilisation d'une « ressource d'enseignement-apprentissage du français parlé sur corpus » (*Ibid.*). La ressource FLORALE est constituée d'émissions radiophoniques de France Culture. Un certain nombre de phénomènes caractéristiques de l'oral sont annotés (phonétiques/phonologiques, syntaxiques, interactionnels, lexicaux). L'objectif est bien la confrontation des apprenants à des exemples oraux authentiques, souvent éloignés des exemples construits des manuels voire éloignés des exemples dits

⁴⁸ Disponible à l'adresse suivante : <https://florale.unil.ch/>.

« authentiques » des manuels (Surcouf, Giroud 2016). La ressource FLORALE compte seulement cinq heures d'enregistrement mais l'interface du site Internet permet une navigation simple pour écouter des phénomènes rarement décrits par ailleurs. Bien que la ressource soit composée d'émissions de radio, elle présente tout de même des énoncés typiques de l'oral souvent négligés des enseignements de FLE. Ce constat est d'autant plus intéressant que les émissions de France Culture pourraient facilement être qualifiées d'élitistes et par conséquent faisant un « bon usage » du français, c'est-à-dire un français très normatif étant donné le haut niveau de scolarisation des participants (souvent des universitaires).

Quatre entrées sont proposées sur la page d'accueil du site : prononciation, structures de phrases, outils de la conversation, mots et expression. Chaque entrée propose plusieurs sous-catégories comme on le voit sur la capture d'écran ci-dessous pour l'entrée « outils de la conversation » :

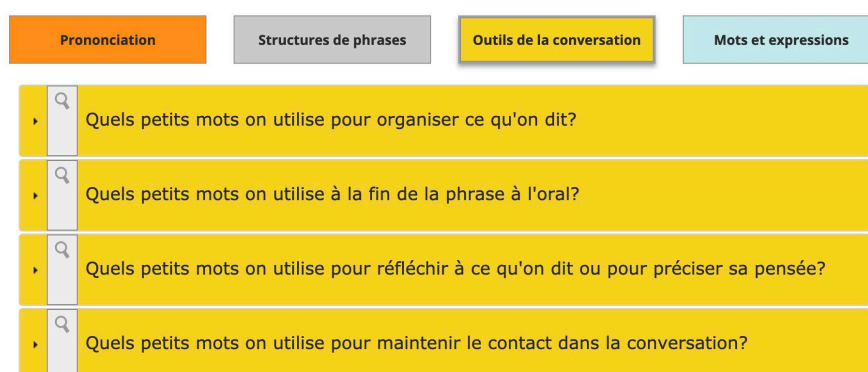


Figure 8 : Les sous-catégories de la catégorie « Outils de la conversation » dans FLORALE

Chaque sous-catégorie est illustrée avec des exemples extraits du corpus. Voici un extrait des énoncés proposés pour la sous-catégorie « Quels petits mots pour organiser ce qu'on dit ? » :

Quels petits mots on utilise pour organiser ce qu'on dit?

- alors**
alors euh quand je suis arrivée ici j'ai acheté quatre choses
- après**
après moi j'ai un rapport un peu bizarre avec l'argent
- bon**
du coup j'ai dit **bon** ben je vais c'est pas grave je vais bosser
- bref**
non mais tu te rends compte? / ça l'a marqué / mon premier truc, mais je m'en souviens, **bref** du coup première année je bosse bien
- donc**
et là on a eu les questions **donc** séparément toute une série de questions
- du coup**
enfin voilà donc **du coup**, il s'est un peu endetté euh chez notre voisin du dessous

Figure 9 : Extrait de la liste d'exemples qui illustrent la sous-catégorie « Quels petits mots pour organiser ce qu'on dit ? » dans FLORALE

L'apprenant ou l'enseignant peut écouter ces phénomènes dans un contexte assez court. Le site est également outillé d'un concordancier qui permet d'effectuer une recherche dans l'ensemble des émissions du corpus. Voici un extrait des résultats de la recherche de *genre* dans le concordancier :

30 Segments Premier Précédent 1 2 Suivant Dernier Aller Page 1 de 2

Exemple	Extrait	Tout
1 le genre de photo qui peut aussi devenir compromettant avec les années	▶ ▶▶	📄
2 basés sur le même genre d'amitié de solidarité de communauté d'intérêts ponctuels ou durables	▶ ▶▶	📄
3 le genre à être toujours chaud pour te suivre sur une connerie	▶ ▶▶	📄
4 et puis [3a] tiens [3o] tout d'un coup je [t] je tiens genre un sac	▶ ▶▶	📄
5 elles étaient du genre à se moquer du monde de tout le monde	▶ ▶▶	📄
6 ou bien non euh c'est genre c'est ma personnalité	▶ ▶▶	📄
7 le genre de fille capable de fondre en larmes pour un B	▶ ▶▶	📄
8 mais euh non c'est euh genre euh je suis infirmière en salle	▶ ▶▶	📄
9 euh terrible j'ai fait euh genre trois mois de dépression	▶ ▶▶	📄

Figure 10 : Extrait des résultats de la recherche de « genre » dans le concordancier de FLORALE

L'utilisateur peut écouter l'extrait, accéder à un contexte plus large ou écouter toute l'émission. Certains phénomènes de prononciation sont transcrits phonétiquement comme dans l'extrait 4 de la figure ci-dessus. Ainsi, la ressource FLORALE permet de sensibiliser les apprenants et les enseignants de FLE aux spécificités de la langue parlée en interaction en proposant différents aspects de la langue à observer.

◆ CLAPI-FLE

Une autre initiative qui mérite d'être mentionnée dans les efforts réalisés pour confronter les apprenants à la langue parlée en interaction est celle menée par l'équipe LIS du laboratoire ICAR à Lyon avec CLAPI-FLE⁴⁹ (Alberdi *et al.* 2018 ; Etienne, Jouin à paraître). La page d'accueil présente toutes les possibilités offertes par cette ressource :

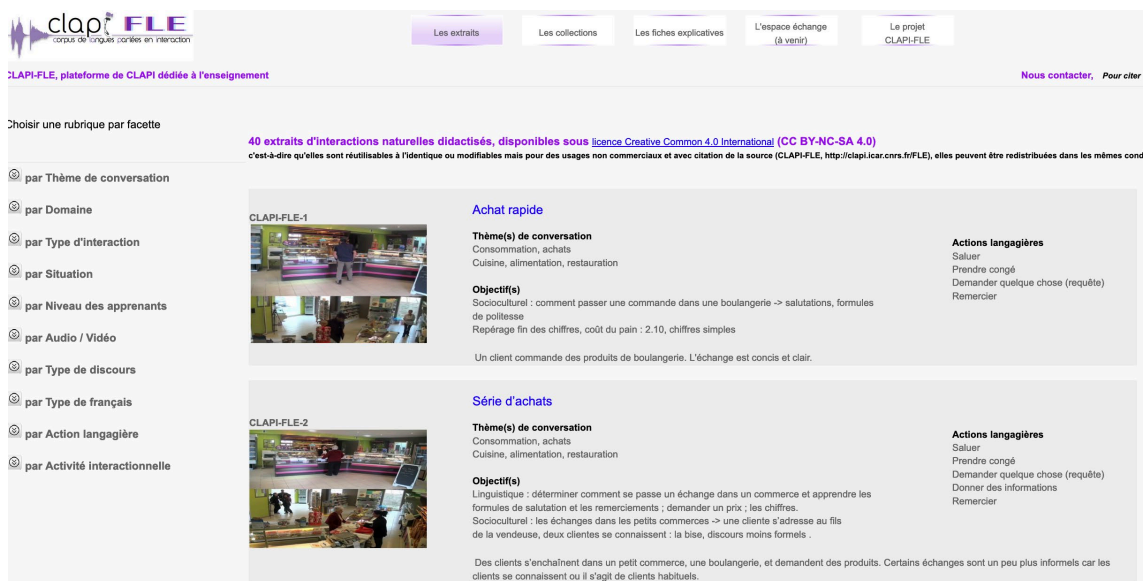


Figure 11 : Page d'accueil de CLAPI-FLE

La plateforme comporte plusieurs types de ressources :

- Quarante extraits du corpus d'interactions CLAPI ont été choisis, décrits et didactisés pour que des enseignants de français puissent les utiliser en salle de classe. Ces extraits sont « classés par type d'interaction (achats/ventes, réunions, conversations, repas, appels téléphoniques, jeux, visites...), thèmes de conversation, actions langagières, registre de langue plus ou moins formel, ainsi que par niveau de langue (A1-C2) » (Alberdi *et al.* 2018 : 58). Ces extraits sont accompagnés de fiches pédagogiques qui peuvent aider les enseignants à les exploiter à des fins didactiques.
- Des « collections » qui présentent des phénomènes spécifiques au français parlé en interaction. Par exemple, la collection « le truc c'est que » comporte

⁴⁹ Disponible à l'adresse suivante : <http://clapi.icar.cnrs.fr/FLE>.

une brève explication de cette pratique interactionnelle et des exemples extraits du corpus CLAPI viennent illustrer le fonctionnement de cette pratique.

- Des « fiches explicatives » appelées également « mini-cours » qui viennent documenter précisément un fait de langue. Par exemple, on trouve une fiche qui explique le fonctionnement du marqueur « trop », une autre celui des remerciements ou encore une autre celui des questions.

Ainsi, « la plateforme CLAPI-FLE tente aujourd’hui de répondre à certains des besoins exprimés par les didacticiens et les enseignants en familiarisant les apprenants au français oral tel qu’il est parlé dans les interactions sociales ordinaires, qu’elles soient professionnelles ou privées » (*Ibid.* : 67). Les fiches pédagogiques sont réalisées en collaboration avec des étudiants de deuxième année de Master de Didactique des langues de Paris Descartes, encadrés par Elodie Oursel, et bénéficient des recherches en interaction menées par l’équipe lyonnaise. Ces fiches tentent de transposer les analyses interactionnelles à la didactique, voici un exemple :

CLAPI-FLE-1

Achat rapide

Un client commande des produits de boulangerie. L'échange est concis et clair.



Fiche pré-pédagogique écrite par Martina Ronci - Master 2 DDL Paris Descartes

Objectif principal
Socioculturel : comment passer une commande dans une boulangerie -> salutations, formules de politesse

Objectif secondaire
Repérage fin des chiffres, coût du pain : 2.10, chiffres simples

Thème de conversation Principal Consommation, achats

Thème de conversation Secondaire Cuisine, alimentation, restauration

Niveau des apprenants A1 **Durée** 29 s **Type de discours** Formel **Type de français** Locuteurs natifs

Domaine Professionnel **Type d'interaction** Achat et vente de produits **Situation** Achat dans un commerce **Audio / Vidéo** Vidéo

Action langagière Acheter - Confirmer - Prendre congé - Remercier - Saluer

Activité interactionnelle
Achat d'un produit dans un commerce - Paiement dans les commerces - Salutations de début d'interaction - Salutations de fin d'interaction

Vocabulaire utilisé Produits de boulangerie ; salutations et remerciements.

Débit Normale **Variété** Standard **Clarté** Très clair

Locuteurs 2 locuteurs
VE1, la vendeuse : femme, 20-50 ans, française C32, le client : informations inconnues

Corpus BoulangerieRurale, G. Icor

Figure 12 : Fiche pédagogique dans CLAPI-FLE

Dans cet exemple, la vidéo exploitée est décrite selon les éléments principaux suivants : le domaine (professionnel) ; les thèmes abordés (la cuisine, l'alimentation et la restauration) ; le type de discours ou le ton de l'échange (formel) ; la situation ou le lieu (achat dans un commerce) ; les actions langagières réalisées (acheter, confirmer, prendre congé, remercier, saluer) ; le débit (normal) ; la variété (standard) ; les locuteurs. Cette fiche précise également des éléments didactiques tels que les objectifs qui peuvent être travaillés avec cette vidéo et le niveau des apprenants (selon le CECRL) avec qui cette vidéo peut être

travaillée⁵⁰. Chaque ressource est également accompagnée de sa transcription (avec la possibilité de la simplifier ou de la complexifier selon les conventions ICOR), de propositions d'exploitation et de données sur le vocabulaire (fréquence, répartition entre les locuteurs).

Cette plateforme est très riche en termes de transposition didactique. Elle est parfaitement complémentaire avec les deux précédentes ainsi qu'avec les travaux possibles à partir du dispositif FLEURON, au niveau des ressources proposées aux enseignants et aux apprenants ainsi que de la façon de les appréhender pour enseigner et apprendre à interagir en français.

4.1.2. FLEURON

Cette indispensable exposition à la langue cible pour des apprenants de langue a, en partie, donné naissance au projet FLEURON (André 2016). Ce projet a vu le jour dans les années 2000 au Centre de Recherches et d'Applications Pédagogiques en Langues (CRAPEL)⁵¹ à Nancy, au moment où ses membres cherchaient un projet fédérateur qui pouvait réunir toutes les thématiques de l'équipe. Parmi ces thématiques, les plus fortes et celles qui ont eu le plus d'influence dans le domaine de la didactique des langues sont notamment : les centres de ressources, l'autonomie, le conseil en langues, la séparation des aptitudes, les styles d'apprentissage, les nouvelles technologies au service de l'apprentissage et l'exploitation de documents authentiques. Le projet FLEURON, s'il est aujourd'hui concrétisé par un dispositif d'apprentissage en ligne⁵² visant à exploiter un corpus multimodal (André 2017), trouve son origine dans les travaux du CRAPEL concernant chacune de ces thématiques. En effet, en résumé, l'idée originale et originelle de FLEURON était de créer un dispositif numérique utilisable en autonomie pour apprendre à comprendre et à s'exprimer à l'oral grâce à des documents authentiques. Les étudiants étrangers étaient le public cible de ce dispositif pour au moins deux raisons : la première est l'accroissement du nombre d'étudiants étrangers qui viennent faire un séjour universitaire

⁵⁰ Je ne rentrerai pas ici dans le débat du niveau de langue attribué à des documents authentiques. En ce qui me concerne, comme je l'ai précisé en introduction, j'estime que les documents ne sont ni simples ni complexes. Ce sont les activités réalisées qui varient selon le niveau des apprenants.

⁵¹ Devenu aujourd'hui l'équipe Didactique des langues et sociolinguistique de l'ATILF.

⁵² Disponible à l'adresse suivante : <https://fleuron.atilf.fr/>.

en France et la seconde est la complexité des démarches administratives à l'université. Le projet a évolué lentement au sein de l'équipe, avant que je le reprenne en 2011⁵³.

Je vais m'intéresser tout d'abord plus particulièrement au fait que FLEURON a bénéficié des travaux pionniers du CRAPEL sur l'usage des documents authentiques dans l'enseignement et l'apprentissage des langues. En effet, avant le nouveau paradigme de la « révolution corpus » (McCarthy 2008 : 564) sur laquelle je reviendrai, l'authenticité a longuement été discutée en didactique et l'est encore aujourd'hui (Widdowson 1998, Adami 2009b, Chambers 2009). Elle a été prônée à Nancy dès les années 1970 (Duda *et al.* 1972 ; Abé *et al.* 1979 ; Holec 1990). L'exposition à la langue cible a été une des priorités, quasiment un combat, des chercheurs du CRAPEL. L'utilisation de documents authentiques constitue le fondement du projet FLEURON. Toutes les approches rigoureuses s'accordent aujourd'hui pour reconnaître que l'authenticité des données comme *input* est l'un des principes fondamentaux des démarches didactiques. En étant confrontés aux pratiques interactionnelles de différents locuteurs, les apprenants développent des compétences acquisitionnelles (Krashen, Terrell 1983 ; Krashen 1985 ; Duda, Tyne 2010).

Ainsi, le corpus FLEURON (décrit dans le chapitre 2, en [2.3.](#)) est né du souhait de pouvoir exposer les étudiants étrangers, apprenants de FLE, à la langue à laquelle ils allaient être confrontés dès leur arrivée en France pour pouvoir vivre comme un étudiant. Pour cela, aucun document ne peut être construit ou fabriqué à des fins didactiques. Le document authentique offre à voir aux apprenants une langue réellement utilisée par ses locuteurs. Ainsi, le corpus multimédia FLEURON propose un ensemble de documents authentiques de différentes natures. L'objectif est de montrer aux apprenants les usages réels de la langue dans des situations de communication réelles avec des locuteurs qui réalisent des activités interactionnelles spécifiques et qui visent des objectifs particuliers. Les apprenants, avec le dispositif FLEURON, sont exposés de différentes façons à la langue (j'y reviendrai) et peuvent observer son fonctionnement interactionnel, pragmatique et social.

Après plusieurs années de développement, le site FLEURON a été stabilisé en 2016 et mis à jour en 2019. Voici son interface actuelle :

⁵³ Pour avoir un historique du projet FLEURON, voir André (2016).

FLEURON Français Langue Etrangère Universitaire Ressources et Outils Numériques

Accueil Qu'est-ce que FLEURON ? Ressources multimédias Liens utiles S'inscrire A propos

*** L'équipe FLEURON se prononce fermement contre l'augmentation des frais d'inscription à l'université pour les étudiants étrangers hors UE. Cette augmentation est une mesure inique et profondément injuste parce qu'elle opère une sélection par l'argent.***

Bienvenue

Le site FLEURON vous propose des **ressources multimédias authentiques** qui illustrent un ensemble de situations de la vie d'un étudiant en France. Ces ressources permettent d'observer différentes situations, telles que :

- S'inscrire ou se ré-inscrire auprès d'un service administratif
- Obtenir sa carte d'étudiant et d'autres documents universitaires officiels
- S'informer sur sa réussite à un diplôme ou sur ses notes
- Obtenir des renseignements sur sa situation administrative et pédagogique
- Discuter avec des étudiants sur des sujets divers
- Discuter avec un enseignant sur son programme ou sur son travail

Ces documents vous permettent de découvrir les situations auxquelles vous serez confrontés à votre arrivée dans une université française. Vous pouvez ainsi **vous préparer à** :

- Mieux comprendre ce que dit votre interlocuteur
- Vous entraîner à mieux vous exprimer dans ces situations
- Interagir avec différents locuteurs dans différentes situations
- Comprendre le système universitaire français
- Comprendre des aspect de la vie d'un étudiant en France

Pour vous aider, chaque document est accompagné :

- D'un sous-titrage qui permet d'avoir accès au texte écrit de ce qui est dit
- De la transcription écrite complète de l'ensemble de la ressource
- De conseils pour vous donner des idées et vous aider à utiliser le document pour apprendre à comprendre ou apprendre à dire
- D'un glossaire vous expliquant les termes spécifiques utilisés

Pour développer votre compétence d'apprentissage et votre sens de l'observation de la langue, nous vous offrons la possibilité de travailler avec un **concordancier** (voir son utilisation dans la partie "Qu'est-ce FLEURON ?").

Français

Choisir cette langue

Utilisateur

Mot de passe

Connexion

Cliquez ici si vous avez oublié vos identifiants de connexion

Concordancier

Mot recherché

Options de recherche :

mot exact

Rechercher

Dernier media publié

Convention de stage (5)

Une étudiante se rend au bureau de la scolarité pour donner ses conventions de stage.

Figure 13 : Page d'accueil du site FLEURON

Le cœur de FLEURON est constitué des ressources multimodales authentiques et d'un concordancier (présenté précisément dans le chapitre suivant, point 5.3.1.). L'exploitation des ressources multimodales et du concordancier feront l'objet de descriptions précises dans certains des points suivants.

Le dispositif propose différents outils destinés à aider les apprenants à s'approprier le site et à l'utiliser à des fins d'apprentissage. Par exemple :

- chaque ressource est transcrite et sous-titrée
- l'utilisateur peut afficher ou masquer la transcription et/ou les sous-titres
- chaque ressource renvoie à d'autres ressources recommandées (par un algorithme)
- certaines ressources sont accompagnées de conseils pour l'apprentissage (comment travailler avec la ressource)
- en se créant un compte (gratuit), l'utilisateur peut :
 - o mettre des ressources de côté
 - o accéder à son historique

- prendre des notes (libres ou orientées par des questions)
- faire des commentaires sur les ressources et les partager avec tous les utilisateurs ou seulement avec les administrateurs
- accéder à un glossaire (principales abréviations et lexique universitaire)

Les outils proposés par le dispositif FLEURON permettent à la fois aux apprenants de travailler en autonomie et aux enseignants d'utiliser le dispositif en salle de classe, en agissant comme un guide ou un conseiller. Ces outils facilitent la prise en main du dispositif mais sont tous au service de l'exposition des apprenants à des données interactionnelles authentiques.

Les ressources présentées dans cette partie - PFC-EC, FLORALE, CLAPI-FLE et FLEURON – sont complémentaires. Elles présentent chacune des caractéristiques de l'oral et du français parlé en interaction utiles pour des enseignants et des apprenants de FLE. Elles sont didactisées de différentes façons (explications, fiches descriptives, outils d'interrogation des données, *etc.*) et peuvent être utilisées pour travailler les deux aptitudes orales, compréhension et expression (Carton 1995). Les auteurs de ces différentes ressources travaillent en collaboration pour offrir à leurs utilisateurs un panel varié de corpus oraux et multimodaux exploitables à des fins didactiques. L'objectif commun de l'élaboration de ces ressources est de proposer aux enseignants et aux apprenants des pratiques interactionnelles réelles.

4.2. Écart entre *langue prescrite* et *langue réelle*. Des pratiques interactionnelles non décrites pour le FLE

J'emprunte à l'analyse clinique de l'activité, comme je l'ai expliqué dans le premier chapitre, les notions de travail prescrit et travail réel et les réflexions menées autour ces notions. Ces dernières me paraissent tout à fait transposables à la problématique de l'exposition à la langue et à la didactique. D'un côté, il existe la *langue prescrite*, celle qui est présentée dans les grammaires, les manuels de langue et parfois par le discours des enseignants adressé aux apprenants. D'un autre côté, il existe la *langue réelle*, celle qui est effectivement utilisée par les locuteurs de la langue cible dans les situations de communication qu'ils rencontrent. La langue prescrite est souvent celle à laquelle les

apprenants de français sont exposés. La langue réelle entre encore trop rarement dans les salles de classe. Il y a parfois un écart important entre langue prescrite et langue réelle.

La langue prescrite relève de ce que l'on peut appeler le français standard, même s'il reste difficile de le circonscrire précisément. Les enseignants de FLE choisissent parfois volontairement d'enseigner la variété standard. Certains d'entre eux le verbalisent explicitement. Ils expliquent que c'est leur rôle d'enseignant de transmettre le français standard ainsi que la norme, et que les apprenants auront bien le temps d'acquérir d'autres variétés en situation d'immersion. Il est vrai que certaines études montrent que la variation peut être apprise dans un pays de la langue cible (voir par exemple Regan 1996, Mougeon et Rehner 2001 ou Thomas 2002). Cependant, comme je l'ai déjà mentionné précédemment, les apprenants habitués à la langue prescrite peuvent se trouver fort démunis face à la langue réelle. Cet écart entre langue prescrite et langue réelle peut être à l'origine de situations délicates, allant du simple malentendu à l'incompréhension totale d'une interaction ou d'un événement. Je mets ici l'accent sur la compréhension orale dans la mesure où c'est l'aptitude la moins travaillée par les enseignants (Carette 2001) et la plus concernée par la variation. En effet, même si l'expression orale d'un locuteur non natif est maladroite, très normée et proche de la grammaire traditionnelle, il peut être possible pour lui de se faire comprendre, notamment en s'appuyant sur son interlocuteur et sur la situation. Il en est autrement en compréhension. Comme le remarque justement Paternostro (2016 : 123), notamment à propos de la variation phonétique :

une distinction doit être opérée entre production et perception. Pour la très grande majorité des apprenants, en effet, le vernaculaire en production ne constitue pas un réel objectif d'apprentissage. En perception, en revanche, le développement d'une sensibilité à la dimension du diaphasique s'avère essentielle pour un décodage efficace au niveau (socio)-linguistique et pragmatique.

Parfois tout de même, les apprenants sont sensibles à la variation en production et souhaitent apprendre à s'exprimer, non pas comme un natif, mais comme un étudiant, comme un jeune, comme un ami dans une conversation, comme un professionnel, comme un marseillais. J'ai constaté ces souhaits lors de mes expérimentations consistant à utiliser des corpus oraux et multimodaux à des fins didactiques. Ces apprenants ont compris l'utilité des corpus et se sont rapidement fixés des objectifs particuliers liés à la langue réelle utilisée par les locuteurs.

De la même façon qu'il existe un continuum entre travail prescrit et travail réel (André 2009a), il existe des phénomènes interactionnels qui se trouvent sur une sorte de continuum entre langue prescrite et langue réelle. Même s'il est impossible de relever l'ensemble de ces phénomènes, je peux mentionner au moins ceux qui ont retenu mon attention depuis le début de mes travaux sur l'oral. Ces phénomènes ne font que rarement l'objet de cours de FLE et apparaissent rarement dans le discours des enseignants (même si souvent ils les produisent sans s'en rendre compte). Voici une liste non exhaustive de ces phénomènes :

- Les assimilations consonantiques (*je sais pas* prononcé [jepa], *je suis* [ʃi])
- La simplification de groupes consonantiques (*parce que* [paskə], *il dit* [idi])
- Les réductions vocalique (*tu as* [ta])
- Les hésitations (*euh*)
- Les amorces de mot (*tra- travail, dor- dormir*)
- Les répétitions et les reformulations
- L'absence du *ne* de négation
- Les petits mots, marqueurs ou ponctuels du discours (*alors, oui non, donc, du coup, quand même, voilà, hein, quoi, bon, ben, ah, etc.*)
- Les énoncés inachevés
- Les énonciations conjointes

Ces spécificités de l'oral sont considérées par certains analystes de la langue et encore plus par des enseignants de FLE comme des aspects déviants alors qu'ils « relèvent des aspects fondamentaux de l'organisation du discours oral et constituent des points privilégiés de la construction interactive du sens » (Debiasieux 2001 : 53). Ces phénomènes apparaissent dans des situations authentiques. L'utilisation des corpus à des fins didactiques tente en quelque sorte de simuler une situation d'immersion au sein de laquelle l'apprenant serait confronté à la langue réelle.

L'analyse sociolinguistique des interactions permet d'apporter à la didactique de l'oral en FLE une contribution indispensable dans la mesure où cette dernière n'est pas dotée de l'outillage descriptif approprié pour enseigner et apprendre une langue telle qu'elle est utilisée dans des situations de communication authentiques. Comme je l'ai précisé précédemment, il existe un écart entre la langue prescrite et la langue réelle. Il existe même des pratiques interactionnelles qui ne font pas l'objet d'une description en didactique

du FLE. Pour illustrer cette absence, je vais mentionner trois pratiques interactionnelles non décrites pour le FLE : l'expression *le truc c'est que*, déjà analysée et illustrée dans CLAPI-FLE (4.2.1.) ; le marqueur *genre*, analysé et illustré notamment grâce au corpus TCOF (4.2.2.) et le phénomène de dislocation *moi je* analysé et illustré avec les ressources CLAPI-FLE et FLEURON (4.2.3.). La présentation de ces pratiques interactionnelles représente une sorte de démonstration de transpositions didactiques possibles, de l'analyse sociolinguistique des interactions verbales vers l'exploitation pédagogique.

4.2.1. *le truc c'est que* dans CLAPI-FLE

Aucune grammaire et aucun manuel ne présentent les usages de l'expression figée *le truc c'est que*. Elle apparaît à l'oral dans des situations souvent informelles. Dans CLAPI-FLE, comme je l'ai expliqué précédemment, il existe un onglet « collection » qui présente une série de phénomènes caractéristiques du français parlé en interaction (Alberdi, Étienne, Juin-Chardon 2018). Cette série se divise en quatre catégories : syntaxe, grammaire, lexique et actions langagières. C'est dans la première⁵⁴ que *le truc c'est que* est traité, analysé et illustré. Une première partie est réservée à une courte explication du phénomène, comme on peut le voir ci-dessous :

L'expression figée *Le truc c'est que...* est typique de l'oral informel. Elle est utilisée pour introduire un élément problématique, contradictoire, gênant... Elle s'emploie dans deux cas : soit lorsque ce qui suit est objectivement contradictoire ou problématique, soit lorsque le locuteur veut montrer qu'il n'est pas à l'aise avec ce qu'il va dire (il n'est pas satisfait, il craint de blesser ou de gêner). Souvent, les deux dimensions, objective et subjective, se superposent dans l'emploi. Elle peut être accompagnée d'indices mimogestuels, comme le froncement de sourcils et les lèvres tirées, dents serrées, ou le haussement de sourcil, tête en arrière, bouche en « o ». Autres formulations, dans différents registres : *Le problème c'est que ; Il faut savoir que ; Ce qu'il y a, c'est que ; Alors, en fait ; Ça va pas le faire parce que...* Pour la culture : le mot *truc* (informel) est souvent utilisé pour remplacer le mot ou le nom d'une personne qui nous échappe. Il prend alors le sens de ce qu'il désigne, comme un joker (ex. : *J'ai trouvé un truc*). Le mot *truc* peut aussi désigner une situation, une histoire ou un problème (ex. : *C'est ça le truc*). C'est dans ces sens qu'il est employé dans l'expression *Le truc c'est que...*

La collection est organisée de la manière suivante :

1. *Le truc c'est que* dépréciatif
2. *Le truc c'est que* annonce une explication de type contrainte ou problème
3. *Le truc c'est que* avec une explication plus factuelle
4. *Le truc* seul
5. Variantes : *c'est ça le truc, ça c'est le truc, tu vois le truc*

Figure 14 : Explication de « *le truc c'est que* » dans CLAPI-FLE

Ensuite, chacune des cinq catégories mentionnées et listées dans cette partie explicative est illustrée avec des exemples extraits du corpus CLAPI. Ci-dessous le premier extrait qui illustre l'utilisation de l'expression en tant que dépréciatif :

⁵⁴ Disponible à l'adresse suivante : http://clapi.icar.cnrs.fr/FLE/collection_syntaxe_le_truc.php.

1. Le truc c'est que dépréciatif
Le truc c'est que... introduit une explication négative et a le même sens que *Le problème, c'est que...*

Exemple 1.1 - Corpus Repas - Kiwi

La jeune fille préférerait ne pas rester à l'université le vendredi où elle n'a pas cours d'analyse conversationnelle, mais elle y est plus ou moins contrainte (pour rendre service à son amie).

Vidéo (mp4)

Transcription (docx)



1 ELI	et le vendredi normalement j'ai une fois sur deux tu sais
2	int- euh hm comm[ent]?
3 BEA	[analyse conversationnelle (0.4)]
4 ELI	exact mais euh le truc c'est que du coup même si euh quand
5	je l'ai pas je reste en fait parce que Laurence vu qu'elle
6	est en formation?

Figure 15 : Premier exemple pour illustrer le fonctionnement de « le truc c'est que » dans CLAPI-FLE

L'enseignant ou l'apprenant qui souhaite exploiter cet extrait peut lire les éléments sur la situation de communication à gauche de la vidéo, écouter et regarder l'extrait, télécharger la vidéo et télécharger la transcription.

Cette collection comporte neuf extraits d'interactions dans lesquels apparaît l'expression *le truc c'est que* réparti dans les cinq catégories. Elle peut être utilisée pour travailler à la fois la compréhension et l'expression avec des apprenants. Cette façon de proposer une analyse, avec des extraits de corpus, me paraît tout à fait utile voire indispensable pour la didactique du FLE. D'une part, les enseignants, n'étant pas des analystes des interactions verbales, sont souvent en difficulté lorsque des apprenants leur demandent d'expliquer ce type de pratiques interactionnelles. D'autre part, même si le site CLAPI-FLE n'est pas prioritairement destiné aux apprenants, ces derniers peuvent tout de même bénéficier d'exemples authentiques qu'ils peuvent visualiser et tenter de comprendre. Les extraits présentés dans CLAPI-FLE sont liés aux demandes de didacticiens et d'enseignants de FLE, tout comme l'ensemble des analyses interactionnelles présentées dans cette synthèse. La présentation de l'utilisation de l'expression *le truc c'est que* peut, à mon avis, servir de modèle pour présenter une transposition didactique. L'idée des collections dans CLAPI-FLE peut être retenue et enrichie par l'explication et l'illustration d'autres spécificités de la langue parlée en interaction, telle que l'utilisation du marqueur *genre*.

4.2.2. *genre* et sa fiche descriptive

Parmi les demandes formulées en séances de formation linguistique par des enseignants et des apprenants, *genre* a retenu mon attention. Je me suis intéressée à *genre* parce qu'il a fait l'objet d'une interrogation sur son fonctionnement en salle de classe et cette interrogation a déstabilisé l'enseignant, à juste titre. Face à cette interrogation, il est complexe de répondre en termes de règles. Les interactions verbales comportent de nombreux petits mots spécifiques à l'oral, parfois empruntés à l'écrit. Ces petits mots accomplissent des activités interactionnelles, ont des effets pragmatiques, permettent de gérer le discours et l'interaction. Ils sont appelés de différentes façons en fonction de l'objet d'étude (marqueurs de discours, ponctuels, régulateurs, particules discursives, etc.)⁵⁵. Cependant, dans les grammaires et dans les manuels, ils sont souvent traités comme des « disfluences » voire des « dysfluences », avec un -y- relatif aux dysfonctionnements⁵⁶. En salle de classe, il n'est pas rare de constater que les enseignants de FLE les classent parmi les « tics de langage », une catégorie plutôt négative, parce qu'ils seraient non intéressants, spécifiques à un individu et à éviter.

L'analyse sociolinguistique de *genre* au sein d'interactions verbales montre que cette pratique accomplit des activités interactionnelles particulières et a des effets propres au genre de discours auquel les locuteurs participent. Tout d'abord, si des étudiants-apprenants de FLE souhaitent comprendre le fonctionnement de *genre*, c'est parce qu'ils l'entendent fréquemment dans des interactions quotidiennes. De plus, s'ils l'entendent aussi fréquemment, c'est parce que son utilisation – en tant que marqueur – n'a cessé de croître ces dernières années, quasiment exclusivement chez les jeunes de moins de vingt-cinq ans (Beeching 2016). Plus précisément, c'est l'utilisation de *genre* en tant que marqueur interactionnel qui est apparue récemment dans la langue réelle (André 2018/soumis). En effet, plusieurs études ont déjà été réalisées sur le marqueur *genre*, au-delà de ses utilisations dans des expressions telles que : *un genre de*, *du genre à* ou encore *ce genre de chose*. Yaguello (1998 : 18) remarquait, il y a un peu plus de vingt ans, que cette « nouvelle particule » se répandait « de plus en plus dans la génération de [ses] enfants ». Depuis, les valeurs pragmatiques de *genre* ont été identifiées. Il sert à exemplifier, illustrer, apporter une preuve, quantifier, introduire une explication, introduire un discours rapporté,

⁵⁵ Pour des précisions terminologiques, voir André (2006a : 345-249).

⁵⁶ Dans l'étude (André, Tyne 2012), mentionnée dans le chapitre 3, nous utilisons volontairement le terme de « dysfluence » à des fins polémiques.

catégoriser (Yaguello 1998 ; Fleischman 1998 ; Rosier 2002 ; Danon-Boileau, Morel 1997 ; Vladimirska 2016). Il permet également de jouer un rôle dans la construction syntaxique d'un énoncé (Rosier 2002 ; Chauveau-Thoumelin 2016), il peut remplacer *comme si* ou une conjonction de subordination. Il peut également introduire un discours métalinguistique (et serait l'équivalent d'une unité telle que *je veux dire par là*). Et enfin, il est également un marqueur interactionnel, c'est-à-dire qu'il joue un rôle au niveau de l'interaction et de l'implication de son interlocuteur dans son discours.

A l'instar des collections proposées par CLAPI-FLE, je propose une collection sur *genre*. Elle a pour objectif d'être une transposition didactique des analyses sociolinguistiques que j'ai effectuées afin d'aider les enseignants et les apprenants à comprendre cette spécificité du français parlé en interaction.

Fiche descriptive de *genre*

Le marqueur *genre* permet d'accomplir plusieurs activités à l'oral en interaction. Par exemple, parmi les activités interactionnelles les plus facilement identifiables, il permet d'introduire un exemple, de donner une explication, de catégoriser, de donner une approximation. Ce marqueur permet également d'introduire un discours indirect ou de relater une scène. Le marqueur *genre* permet également, un peu comme *tu vois*, de demander à son interlocuteur d'être attentif ou de l'impliquer dans son discours. D'un point de vue syntaxique, *genre* sert parfois de conjonction entre deux propositions. Le marqueur *genre* est essentiellement utilisé par des locuteurs jeunes (entre 15 et 25 ans) qui ont une relation amicale et qui participent à une conversation informelle.

La collection est organisée de la façon suivante :

1. *genre* : pour illustrer, exemplifier, apporter une preuve
2. *genre* : pour donner une approximation
3. *genre* : pour introduire une explication
4. *genre* : pour catégoriser
5. *genre* : pour introduire un discours indirect ou une scène recréée
6. *genre* : pour impliquer son interlocuteur dans son discours
7. *genre* : pour coordonner deux propositions
8. *genre* : pour introduire une question

Illustration avec des exemples extraits du corpus TCOF.

1. *genre* : pour illustrer, exemplifier, apporter une preuve

Exemple 1.1. Deux étudiantes préparent un exposé. Elles déterminent une stratégie pour essayer de respecter le temps qui leur est imparti. (Cours_MAU_15).

L2 donne un exemple de ce qu'elle peut faire pour signaler à son amie qu'il lui reste deux minutes de présentation.	L2 ben moi comme je serais en face ce que je fais bah je je te chronomètre et puis comme ça je te fais genre signe quand ça fait une minute deux minutes
	L1 ok on peut faire ça

Exemple 1.2. Deux étudiantes parlent de leur avenir professionnel. (Corpus_AMA_15)

L2 donne un exemple de ce qu'elle souhaite faire. L1 propose ensuite un autre exemple (en posant une question).	L1 tu veux faire quoi toi
	L2 moi je veux faire quelque chose en rapport avec la langue des signes + genre accompagner les enfants sourds
	(...)
	L1 et genre orthophoniste ça te plairait pas

Exemple 1.3. Une étudiante raconte à une amie son séjour Erasmus. (Erasmus_UZA_17)

L1 utilise <i>genre</i> comme interrogatif, pour demander à L1 de donner des exemples de matières qu'elle a suivies pendant son séjour universitaire. L2 commence son tour par <i>genre</i> pour donner un exemple.	L2 on a fait pas mal de matières qui n'ont pas forcément à voir avec la sociologie de façon directe
	L1 genre quoi +
	L2 genre euh + on avait euh démocratie et société civile

2. *genre* : pour donner une approximation

Exemple 2.1. Trois étudiantes racontent leur travail saisonnier. (Dimanche_NAG_16)

L3 prend *Pren1* (prénom anonymisé) comme exemple puis donne une approximation de ce qu'elle a volé.	L3 mais euh ouais genre *Pren1* nous elle euh elle a travaillé au McDo cet été et du coup elle a volé genre des caisses entières de crème Deluxe
	L1 là du coup elle en a plein chez elle

Exemple 2.2. Un jeune homme raconte son voyage en Angleterre. (Coree_GHU_14).

L1 donne une approximation de la distance entre deux villes.	L1 sur un site ouais on était parti clubbing avec les enfants à Worthing qui est à peu près à genre une heure de Brighton enfin de Ovingdean
--	---

3. *genre* : pour introduire une explication

Exemple 3.1. Un jeune homme raconte à ses amis son travail en Australie. (Garcons_TOU_17)

L2 fait une parenthèse pour expliquer pourquoi il faisait chaud.	L2 mais ouais ouais quand même là-dessus j'ai eu de la chance [rire] [toux] mais c'était chaud parce que genre euh dans la ferme de kiwis tu sais c'était pendant l'été L3 ouais L2 il faisait trente-cinq degrés et on devait ramasser des kiwis sous trente-cinq degrés euh toute la journée [toux]
--	--

4. *genre* : pour catégoriser

Exemple 4.1. Une jeune femme raconte son voyage en Inde. (Inde_BEL_ZIN_15).

L1 explique qu'il existe plein de moyens de transport différents qui peuvent s'apparenter à des taxis.	L1 c'est vraiment une des choses qui nous a le le plus euh le plus surpris et au niveau de la circulation et bien il y a des voitures des taxis mais il y a tout il y a des vélos il y a des tuks-tuks ces petites mobylettes avec des des places pour transporter genre taxis mais il y a aussi des animaux il y a des euh comment des chameaux des dromadaires qui tirent des charrettes
--	---

5. *genre* : pour introduire un discours indirect ou une scène recréée

Exemple 5.1. Une jeune femme raconte à une amie ce qu'il s'est passé lors de son trajet en voiture de chez elle à l'université. (Automobile_GUE_08)

L1 rapporte les paroles d'un automobiliste.	L1 et euh avant juste avant dans le village il y avait euh + il y avait une caisse devant moi et une autre derrière et celle de derrière lui collait au cul genre vas-y avance et tout L2 < ouais L1 mais > super nerveusement quoi
---	--

Exemple 5.2. Un jeune homme raconte à ses amis son travail en Australie. (Garcons_TOU_17)

L2 rapporte le mail qu'il a reçu.	L2 je devais quitter ma famille d'accueil j'avais nulle part où aller j'avais pas de taff et tout tu vois et genre euh ben à minuit moins dix j'ai reçu un mail de de la ferme qui me disait oh ben c'est bon si vous voulez vous pouvez venir euh demain on vous loge
-----------------------------------	---

Exemple 5.3. Une jeune femme raconte à une amie ce qu'il s'est passé lors de son trajet en voiture de chez elle à l'université. (Automobile_GUE_08)

L1 raconte la scène qu'elle a vécue. Elle accompagne son discours de gestes qui imitent ceux de l'automobiliste.	L1 ben quand on arrive au stop le matin tu sais il y a une queue L2 ouais L1 mais laisse tomber + elle s'arrête et euh là le mec je je commence il sort de la caisse et il va voir celui qui l'a doublé et tu sais genre trop vénère il commence à faire des gestes dans tous les sens tu vois qu'il crie tu vois qu'il est pas content tu ne sais pas pourquoi mais euh + voilà et la file de voitures qui avance qui avance et tout
--	---

6. *genre* : pour impliquer son interlocuteur dans son discours

Exemple 6.1. Une jeune femme raconte un film à un ami. (Film_TRE_15)

L1 ponctue son discours avec <i>genre</i> . Le marqueur accomplit plusieurs activités interactionnelles et, par la même occasion, en multipliant les apparitions de <i>genre</i> , il implique son interlocuteur dans son discours.	L1 ouais bah j- c'est ça qui rythme le truc que là <i>genre</i> il y en a un c'est un acteur et l'autre c'est un prof d'histoire euh à la fac je crois donc euh en gros il y en a qui a une vie de ouf l'autre qui a une vie de merde et euh et oui c'est le même principe en fait c'est un schizo tu vois mais il y a des trucs trop bizarres avec une histoire d'araignée pas compris pourquoi genre euh genre la fin et vraiment euh genre sa copine est enceinte enfin sa femme enfin bref elle est enceinte et genre euh tout à la fin il va la voir dans une pièce et en fait elle s'est transformée en araignée
---	---

7. *genre* : pour coordonner deux propositions, à la place d'une subordination

Exemple 7.1. Deux étudiantes parlent de la vie à Tahiti. (Tahiti_AIT_17)

L2 utilise <i>genre</i> à la place de <i>que</i> .	L2 ouais c'est ça la vie elle est chère là-bas L1 ou les gens ils sont pauvres c'est quoi ouais L2 hum + je crois genre on paye beaucoup là-bas L1 hum
--	---

8. *genre* : pour introduire une question

Exemple 8.1. Une jeune femme (L1) interroge son ami (L2) sur son ancien métier de militaire.
(Armee_MER_15)

L1 introduit sa question par <i>genre</i> .	L2	bah nous on a ses rations combat et on dort euh + sur des pauvres lits de merde et juste avec une petite moustiquaire au-dessus pour dire que on est protégé quoi
	L1	mais genre c'est parce qu'ils ont pas les budgets ou euh
	L2	oui l'armée française n'a pas du tout le budget

Exemple 8.2. Une jeune femme (L1) interroge son ami (L2) sur son ancien métier de militaire.
(Armee_MER_15)

L1 introduit sa question par <i>genre</i> . Les deux premiers sont une amorce, L2 interrompt L1 pour prendre la parole. Dès que L1 a terminé son tour et le message qu'il souhaitait communiquer, L1 reprend sa question en commençant par <i>mais genre</i> . On remarque ici que <i>genre</i> introduit les deux alternatives de la question.	L1	il y en a ou pas des mecs qui se tapent des putes
	L2	ah oui oui il y en a il y en a même eu un qui a attrapé le sida là-bas donc euh
	L1	sérieux par une pute
	L2	bah oui +
	L1	et genre euh < genre i-
	L2	une pute > c'était en fait euh + une euh une fille euh du village qui était venue pour ai- aider en cuisine +
	(...) ###	
	L1	mais genre il est tombé amoureux ou genre il voulait < juste la
	L2	non non > c'est voilà +
	L1	hum + chaud quand même +

Conclusion

Le marqueur *genre* peut accomplir différentes activités interactionnelles. Parfois, une seule occurrence vise plusieurs objectifs, tels qu'exemplifier et justifier ou expliquer et demander l'attention de l'interlocuteur.

Il apparaît à différents endroits dans le tour de parole, parfois au début pour prendre la parole, parfois entre deux unités et parfois en fin de tour.

Ainsi, la fiche descriptive de *genre* est réalisée à partir de l'analyse sociolinguistique d'interactions verbales (corpus TCOF). Elle est rédigée pour être utilisée soit par des enseignants, soit par des apprenants de FLE, essentiellement pour un travail en compréhension orale (même si, pour les apprenants qui le souhaitent, un travail en

expression est possible avec la fiche). Elle est présentée ici dans une version courte et simplifiée. Sa version finale comporte plus d'exemples pour chacune des catégories et donne la possibilité d'écouter les extraits sonores de chaque exemple. Elle présente également des liens vers des corpus oraux et multimodaux pour que les utilisateurs puissent effectuer leurs propres requêtes (*via* un concordancier), compléter leurs observations et analyser les occurrences qu'ils rencontrent (sur les principes de l'apprentissage sur corpus, développé dans le chapitre suivant) afin de les comprendre. Cette fiche a été testée et évaluée par des enseignants de FLE du DÉFLE à Nancy.

4.2.3. La duplication du sujet avec un pronom disloqué - *moi je*

Le sujet dupliqué avec pronom disloqué est un exemple typique de ce que les enseignants de FLE et les manuels n'enseignent pas à comprendre et à produire. Ce phénomène, pourtant fréquent à l'oral, est même parfois proscrit des salles de classe, jugé beaucoup trop éloigné d'un français standard et normatif. Cependant, si l'on observe la langue réelle, on constate que ce phénomène, analysé depuis longtemps dans les travaux sur le français parlé (même s'il n'y a pas consensus, voir par exemple Blasco-Dulbecco 1999 ou Coveney 2003), est fréquent, quel que soit l'identité du locuteur, ses relations avec ses interlocuteurs, le genre de discours ou encore le degré de formalité de l'interaction. Je vais m'intéresser ici à l'utilisation de *moi je* et à l'utilité de son observation pour des apprenants de FLE.

Tout d'abord, on peut regarder la fréquence de ce phénomène à l'oral. Par exemple, on se rend compte que même dans un corpus de dialogues d'une série télévisée, il est fréquemment utilisé. Les dialogues de la série *Plus belle la vie* que j'ai analysés afin de sélectionner ceux qui se rapprochaient le plus d'interactions authentiques (André 2018b) comportent ce type de dislocation avec une duplication du sujet. Le logiciel JConc permet de lister les occurrences⁵⁷ :

⁵⁷ Les épisodes qui constituent le corpus de travail ont été transcrits et alignés texte-son avec le logiciel *Transcriber*, ce qui permet d'interroger le corpus avec le concordancier JConc.

Num	Fichier	Contexte Gauche	Occurrence	Contexte Droit
1	3476	tu fais là/ bah écoute	moi je	je suis de retour à
2	3476	je suis un homme libre	moi je	suis un apache oui oui
3	3476	planche à roulettes non merci	moi je	préfère rester là à glander
4	3476	voir ça avec mon collègue	moi je	dois d'abord vous encaisser
5	3477	tu as plus envie de	moi je	préfère que tu me le
6	3477	ça change tout en fait	moi je	pensais qu'il en avait
7	3477	y en a mille mais	moi je	suis sûre qu'elle se
8	3477	vous présente mes excuses et	moi je	me sens mieux merci de
9	3477	humain que j'avais devant	moi je	me demande comment ma
10	3477	je vais faire de ça	moi je	sais pas tu connais pas
11	3478	je veux pas te bousculer	moi je	t'aime hein hum la
12	3478	couche t'a minée mais	moi je	veux juste m'assurer que
13	3478	un peu vaseux et toi	moi je	me suis pas pris une
14	3478	boule d'accord mais bon	moi je	relance mon activité j'ai
15	3478	à autre chose oui mais	moi je	me suis fait remonter les
16	3479	pas une promenade de santé	moi je	fais du sport hein oui
17	3479	là je suis plus rien	moi je	suis plus rien du tout
18	3481	me suis là parce que	moi je	suis pas sûr nan mais
19	3481	cette clause et euh dites	moi je	sais que: e Marc et
20	3481	il me fait entièrement	moi je	m'occupe de la partie
21	3481	Ariane est au courant lâche-	moi je	veux juste savoir si arianne
22	3482	vas-y toi profite pis	moi je	vais accompagner Noé attends

Figure 16 : Extrait des résultats de « moi je » dans le corpus Plus belle la vie (JConc)

D'autres corpus, authentiques cette fois, peuvent être utilisés pour se rendre compte de la fréquence d'utilisation de *moi je*. Par exemple, le corpus FLEURON compte 58 occurrences :

Concordancier moi je ☒ mot exact Rechercher		
Concordancier (58)	Ressources (0)	CNRTL
A gauche du mot	Mot recherché	A droite du mot
L2 : et euh au pôle européen	moi je	suis personnel de l'université Henri Poincaré
E :	moi je	suis je m'appelle Agatha *** je suis polonaise A : hum hum
A : comme ça c'est vrai que	moi j	'essaierai j'essaierai de la faire euh
A : et puis	moi je	m'occuperai de la partie administrative E : hum hum
A : je vais vous noter oh excusez-	moi je	vais vous noter sur le le dossier
A : donc vous pourrez vous pourrez les contacter et pis	moi je	vous attends avec le papier E : oui
E : oui d'accord A : et ensuite avec le cert-	moi j	'établis un certificat de préinscription
E : ce c'est ah! c'est c'est-à-dire	moi j	'habite où
A : 4 rue de (XX) E : oui oui c'est	moi j	'ai loué un appartement
A : et	moi je	pourrai faire la saisie
A :	moi j	'envoie les informations sur cette adresse d'accord
A :	moi je	t'env- je vous enverrai je vous envoie à tous les étudiants Erasmus
A : parce que pour	moi je	vais euh je vais en avoir besoin
A : le temps de faire les photocopies et puis	moi je	t'attends et
A :	moi je	t'attends et puis je te donne le papier
E : alors	moi j	'étais venue pour une demande toute simple
E : euh excusez-	moi j	'aurais voulu savoir où c'est que je peux me procurer euh
A2:	moi je	vais vous parler de la responsabilité civile
A : euh huit heures trente dix-sept heures	moi je	fais huit heures trente dix-sept heures
A : donc si vous venez me voir et que vous me dites voilà	moi j	'ai envie de
E1:	moi j	'étais allée donc oui c'était à la rentrée
A : et à partir de là	moi je	je fais les affiches
E : oui mais par exemple si elle finit un cours je peux pas y aller après son cours A : oui mais après	moi je	sais pas son emploi du temps hein
A : là	moi j	'ai rien d'indiqué là pour elle E : madame Voleri j'ai pas noté
A :	moi je	pense qu'il faudrait lui envoyer un mail hein E : lui envoyer un mail vous avez leur euh A : parce que pour les mails ils répondent toujours hein

Figure 17 : Extrait des résultats de la recherche de « moi je » dans FLEURON

Avec ce corpus, on constate que *moi je* apparaît dans différentes situations et différents genres de discours. Ce sujet dupliqué n'est pas réservé aux conversations informelles entre des personnes qui se connaissent bien et qui interagissent de façon « relâchée » (comme c'est souvent mentionné par les enseignants pour expliquer l'apparition de *moi je* quand il n'est pas banni des grammaires). Certains des locuteurs du corpus FLEURON sont engagés dans des relations de service et ne connaissent pas leurs interlocuteurs. Malgré ces éléments, la duplication du sujet est bien présente dans le corpus.

Une recherche des occurrences de *moi je* dans le corpus TCOF donne également des résultats intéressants. Dans les 65 heures d'enregistrement utilisées pour cette recherche⁵⁸, il y a 1675 occurrences de *moi je*. Cela correspond à une moyenne de 25 occurrences par heure ou encore une occurrence toutes les 2 minutes 30. Cette fréquence légitime grandement l'évocation de ce phénomène en formation linguistique.

Ensuite, il y a deux façons d'analyser simplement la duplication du sujet. La première est de nier les effets pragmatiques de la duplication. La seconde est de considérer la duplication comme une ressource qui permet d'accomplir des activités interactionnelles. Par exemple, l'énoncé « alors moi j'étais venue pour une demande toute simple » ouvre l'interaction suivante (ressource « Paiement en espèce » du corpus FLEURON⁵⁹) :

- | | |
|---|---|
| E | alors moi j'étais venue pour une demande toute simple hum en fait euh c'est pour mon inscription voilà je voulais payer en espèces |
| A | oui mais c'est pas le bureau là |

Cet énoncé peut être analysé de deux façons :

1. la dislocation à gauche de *moi* et la duplication du pronom est une caractéristique de l'oral et n'a aucune valeur pragmatique ;
2. la dislocation à gauche de *moi* et la duplication du pronom comportent des implicites interprétables avec la situation de communication. Ici, la locutrice arrive dans un bureau à la scolarité, elle a attendu son tour pour entrer dans le bureau et qu'on s'occupe d'elle, elle a une demande à formuler et sait que la réponse n'est pas simple. De plus, la dislocation permet de poser le topic ou le thème avant d'en dire quelque chose.

⁵⁸ Le corpus utilisé pour cette recherche comporte des corpus qui ne sont pas encore disponibles notamment parce que leur transcription n'est pas vérifiée et qu'ils ne sont pas anonymisés.

⁵⁹ Voir l'intégralité de la transcription de l'interaction en [4.3.1.](#)

Cette seconde analyse peut être complétée par la ressource CLAPI-FLE qui comporte une collection sur les dislocations⁶⁰. Comme toutes les collections, elle commence par une explication du phénomène :

La dislocation est le déplacement d'un (élément d'un) des syntagmes nominaux principaux de la phrase (sujet, compléments du nom ou compléments du verbe) et sa reprise (souvent pronominale) dans le cœur de la phrase. L'élément déplacé peut être placé au début ou à la fin de la phrase.

La dislocation est souvent utilisée à l'oral. Elle permet de marquer l'élément saillant de la phrase. Elle joue aussi probablement un rôle cognitif dans l'organisation et dans la structuration des idées et des informations.

D'autres phénomènes ont des fonctions similaires : les faux-départs et reformulations par exemple manifestent aussi le fait que le locuteur pense et produit en même temps ; les formulations clitiques (*c'est ... qui*) permettent également de rendre un élément saillant. Côté prosodie, la dislocation à droite porte des traits spécifiques : désaccentuation (volume sonore plus faible) et baisse du pitch (descente vers les graves, de 5 demi-tons en moyenne) par rapport au noyau de la phrase, et une pause possible, mais non obligatoire entre le noyau et la dislocation à droite. La dislocation à gauche porte également des traits spécifiques, mais leur identification est plus polémique. On remarque généralement une hausse de la fréquence fondamentale sur la syllabe accentuée de l'élément disloqué (montée dans les aigus), un allongement de cette même syllabe et un pic d'intensité (augmentation du volume) pour la voyelle de cette syllabe.

(Référence : De Cat 2007, *French Dislocation, Interpretation, Syntax and Acquisition*, Oxford : OUP.)

La collection est organisée de la manière suivante :

1. les dislocations à gauche et/ou à droite
2. les dislocations du sujet ou du complément
3. les dislocations du syntagme nominal entier ou partiel
4. l'élément disloqué conserve sa forme d'origine
5. l'élément disloqué ne conserve pas sa forme d'origine
6. l'accumulation de dislocations
7. L'apport pragmatique de la dislocation

Figure 18 : Explication des dislocations dans CLAPI-FLE

La septième catégorie présente l'apport pragmatique de la dislocation avec six exemples en contexte. Les exemples sont précédés d'une petite explication :

7. La dislocation : apport pragmatique

L'apport des dislocations au niveau pragmatique est souvent subtil. L'élément disloqué est en général le topique de la phrase, c'est-à-dire l'élément sur lequel elle porte. C'est une information déjà là, posée avant soit verbalement, soit par une orientation intersubjective commune (par ex. regarder un objet et orienter le regard de l'autre vers ce même objet pour qu'il devienne le centre de l'attention : même si on n'en a pas parlé avant, il ne constitue plus une information inconnue dans l'espace intersubjectif). Le locuteur veut la rendre saillante. Dans la plupart des exemples précédents, la dislocation reflète le processus de construction des idées et de la phrase, en posant l'élément saillant en premier, comme pour laisser le temps à l'interlocuteur (ou pour se laisser le temps) d'enregistrer le topique avant d'en dire quelque chose, ou en dernier, comme pour rendre l'élément saillant après-coup, ou pour le rappeler, le clarifier ou le désambigüiser. Parfois, cependant, elles marquent clairement une mise en contraste de deux éléments.

⁶⁰ http://clapi.icar.cnrs.fr/FLE/collection_syntaxe_dislocation.php



1 ELI [mais maintenant c'est quand même bizarre je trouve
 2 BEA [ouais
 3 MAR [ils apprennent [les lettres
 4 ELI [ouais mais même je trouve que c'est bizarre (0.3)
 5 alors moi je sais que qu- vous vous rappelez pas quand on était petites à moins
 6 que c'était pas comme ça vous; mais moi je savais pas écrire
 7 [hein m` enfin à la fin
 8 MAR [ben moi je me souviens plus
 9 BEA je sais plus;

Figure 19 : L'explication de l'apport pragmatique de la dislocation dans CLAPI-FLE et premier exemple

Ainsi, les apprenants et les enseignants de FLE peuvent alterner entre différentes ressources afin d'obtenir des informations sur la langue réelle, celle qui est effectivement utilisée par les locuteurs. Les ressources FLEURON et CLAPI-FLE sont complémentaires. Pour analyser en formation linguistique ces pratiques interactionnelles qui ne sont pas décrites ni dans les manuels ni dans les grammaires, les ressources qui exploitent les corpus oraux et multimodaux sont précieuses.

4.3. Apprendre à interagir avec les ressources de FLEURON

Les analyses présentées dans les points précédents proposent des transpositions didactiques qui expliquent l'apparition de certains phénomènes précis, tels que l'utilisation de l'expression *le truc c'est que*, du marqueur *genre* ou du sujet disloqué et dupliqué *moi je*. Néanmoins, ces phénomènes apparaissent dans la temporalité des interactions verbales de façon ponctuelle. Ils doivent donc pouvoir être observés, identifiés et compris en situation interactionnelle par des apprenants et des enseignants. Pour qu'un apprenant de langue puisse acquérir des compétences socio-interactionnelles, il paraît naturel d'utiliser comme documents supports des interactions verbales authentiques à partir desquelles l'enseignant va pouvoir travailler. D'une part, les apprenants observent des interactions afin d'apprendre à comprendre leur déroulement et leur fonctionnement. D'autre part, ces observations permettent aux apprenants d'apprendre à s'exprimer en interaction de façon appropriée. Je reviendrai plus précisément sur l'utilisation outillée des corpus dans le dernier chapitre mais je souhaiterais évoquer, en guise de transition, l'apport de l'analyse des interactions verbales à l'enseignement et à l'apprentissage du FLE.

Je présente ici l'exploitation du corpus multimédia FLEURON, décrit dans le chapitre 2, à des fins didactiques. Cette exploitation représente une première façon d'utiliser des corpus à des fins didactiques, en exploitant l'analyse sociolinguistique des interactions verbales. Elle consiste à faire observer les interactions verbales aux apprenants, et à analyser avec eux leur déroulement, tour par tour si besoin, au fur et à mesure de leur déroulement, dans la temporalité de leur élaboration. Il s'agit d'observer et de comprendre quelles sont les pratiques interactionnelles utilisées par les locuteurs pour accomplir leurs activités interactionnelles. Il s'agit parallèlement d'examiner les activités interactionnelles réalisées, leur potentiel enchevêtrement et leur diversité. Il s'agit également de découvrir les contraintes imposées par le genre de discours et les redéfinitions éventuelles de ce genre. Enfin, il s'agit de se rendre compte des variations de la langue, en fonction de la situation de communication. Chaque ressource de FLEURON peut être exploitée par une analyse sociolinguistique qui va faciliter la compréhension des interactions et l'appropriation de compétences socio-interactionnelles.

Je vais tout d'abord présenter l'utilisation des ressources multimodales de FLEURON par son public cible, c'est-à-dire des étudiants étrangers (4.3.1.) avant de présenter une utilisation « détournée » du corpus FLEURON avec des adultes migrants peu scolarisés (4.3.2.). Le corpus étant constitué de vidéos de situations de communication non universitaires, certaines d'entre elles sont utiles pour ce type de public.

4.3.1. Avec des étudiants étrangers

Les expérimentations relatées dans cette section représentent une première façon d'utiliser le corpus FLEURON à des fins didactiques. L'objectif est de transposer les analyses sociolinguistiques des interactions verbales à l'enseignement et l'apprentissage du FLE. Je vais examiner un premier exemple avec la vidéo intitulée « Paiement en espèces » (Figure 20 ci-dessous). Cette vidéo présente une interaction entre un agent administratif de la scolarité (A) et une étudiante (E) qui veut régler une partie de son inscription mais qui rencontre des problèmes pour effectuer le paiement car elle n'a pas d'espèces.

Retour
Païement en espèces
☆

Media
Description
Conseils
Glossaire
Notes

Cacher la transcription
Imprimer

E : alors moi j'étais venue pour une demande toute simple
E : hum en fait euh c'est pour mon inscription voilà
E : je voulais payer en espèces
A : oui mais c'est pas le bureau là
E : ben alors ce qu'il y a c'est que le bureau en face est fermé
A : oui il est fermé
A : il faut que
E : et vous pouvez pas du tout euh
A : non pas du tout parce que je suis pas régisseur de recettes donc je peux vraiment pas encaisser de d'argent
E : d'accord
A : alors ce que je vous propose c'est de
E : dites-moi
A : c'est de revenir euh demain
E : je peux pas je peux pas du tout
A : et
E : et vous avez vraiment pas du tout la possibilité de mettre ça dans un coin
E : parce que là ça m'en- ça m'ennuie en plus quoi

Masquer les sous-titres

E : alors moi j'étais venue pour une demande toute simple

Concordancier

Figure 20 : Début de l'interaction de la ressource « Paiement en espèces » dans FLEURON

Ce que je propose avec cette ressource et ce que j'ai expérimenté avec des apprenants de français à l'université, c'est une analyse sociolinguistique réalisée conjointement entre l'enseignant et l'apprenant. Concrètement, tour de parole par tour de parole, l'enseignant peut pointer les spécificités du français parlé en interaction (pas toutes pour ne pas surcharger cognitivement l'apprenant et pour s'adapter à l'apprenant). Pour illustrer cette démarche, je vais pointer plusieurs points de cette interaction, tel que j'ai pu le faire lors d'expérimentations pendant lesquelles j'ai pris le rôle d'une enseignante de FLE avec des étudiants étrangers (généralement autour du niveau B1 du CECRL).

Tout d'abord, l'observation des métadonnées indiquent que la situation est assez formelle, entre des locutrices qui ne se connaissent pas et qui se vouvoient lors d'une relation professionnelle. Après ces premières observations, l'analyse de la progression de l'interaction peut commencer. Pour faciliter la lecture, je reproduis la transcription de l'interaction analysée.

Extrait de la transcription de « Paiement en espèces » (FLEURON)

1. E : alors moi j'étais venue pour une demande toute simple hum en fait euh c'est pour mon inscription voilà je voulais payer en espèces
2. A : oui mais c'est pas le bureau là
3. E : ben alors ce qu'il y a c'est que le bureau en face est fermé
4. A : oui il est fermé il faut que
5. E : et vous pouvez pas du tout euh

6. A : non pas du tout parce que je suis pas régisseur de recettes donc je peux vraiment pas encaisser de d'argent
7. E : d'accord
8. A : alors ce que je vous propose c'est de
9. E : dites-moi
10. A : c'est de revenir euh demain
11. E : je peux pas je peux pas du tout
12. A : et
13. E : et vous avez vraiment pas du tout la possibilité de mettre ça dans un coin parce que là ça m'en- ça m'ennuie en plus quoi

Je vais analyser précisément mais de façon synthétique les deux premiers tours de cette interaction afin d'illustrer la démarche mise en œuvre avec des apprenants.

Tour 1 : alors moi j'étais venue pour une demande toute simple hum en fait euh c'est pour mon inscription voilà je voulais payer en espèces

- L'étudiante (E) commence son tour par *alors*, ce petit mot de l'oral permet de prendre la parole, il est également utilisé par l'agent (A) pour commencer son tour (ligne 8). Il est facile de faire remarquer à un apprenant (qui possède un métalangage) que cet adverbe, selon les grammaires, est détourné à l'oral, y compris dans une situation plutôt formelle et entre des locutrices qui ne se connaissent pas.
- Il est également possible de pointer la dislocation *moi je* qui apparaît dans cette situation plutôt formelle alors qu'elle est exclue des manuels et stigmatisée par les grammaires, comme je l'ai expliqué précédemment.
- Ensuite, il est intéressant de montrer à des apprenants que la locutrice utilise un temps du passé, le plus-que-parfait, pour expliquer sa présence actuelle et sa demande future avec *j'étais venue*. Lors de chaque utilisation de cette ressource, les apprenants sont surpris par cette pratique, puis comprennent son utilisation.
- La raison de la venue est une demande. La locutrice atténue cette demande en la qualifiant de *toute simple*.
- Ensuite, la locutrice expose le thème de la raison de sa venue *hum en fait euh c'est pour mon inscription voilà*. Cette unité est composée d'un noyau thématique « c'est pour mon inscription », de petits mots de l'oral « hum », « en fait », « voilà » et d'une hésitation. Cette façon d'exposer un élément, avec différents marqueurs, ne fait – (presque) évidemment – pas partie de la langue ou de l'interlangue des apprenants mais elle ne fait pas non plus partie des énoncés qu'ils comprennent facilement, même

au niveau B1. Dans ce tour, *hum en fait euh* sert de pivot entre la préparation de la demande et l'exposé du thème de la demande.

- La demande est exposée en fin de tour après la transition marquée par *voilà*.
- La locutrice produit de nouveau un verbe au passé – à l'imparfait – pour exprimer une action future : *je voulais payer en espèces*, ce qui représente la raison de sa venue. On constate ici l'utilisation de l'imparfait de politesse.
- En bref, ce tour est composé de plusieurs unités qui posent, au fur et à mesure de leur apparition, la raison de la venue de l'étudiante. Ce tour sert ici d'ouverture de l'interaction.

Tour 2 : oui mais c'est pas le bureau là

- L'agent répond en commençant son tour par « oui mais ». Dans d'autres études, j'ai analysé les mouvements effectués par « oui non mais » en début de tour (André 2005, 2010). Ces études peuvent être réinvesties pour analyser cet énoncé. L'agent montre à l'étudiante qu'elle a bien entendu sa requête avec « oui » puis elle introduit avec « mais » l'argument qui va la conduire à la refuser. Faire observer cet énoncé par des apprenants ne signifie pas que ceux-ci doivent être capables de le produire mais, dans un premier temps, ils peuvent apprendre à le comprendre.
- Il est possible de faire remarquer à un apprenant que la négation est réalisée sans *ne* dans *c'est pas*.
- L'utilisation du déictique *là* est également intéressante.

Je ne vais pas poursuivre cette analyse tour par tour mais je vais seulement relever des points intéressants pour des apprenants qui souhaitent acquérir une compétence socio-interactionnelle. Par exemple, à la ligne 5 on peut remarquer une question inachevée dans le tour de l'étudiante : *et vous pouvez pas du tout euh*. J'ai déjà analysé ce phénomène précédemment dans le point sur les questions (3.1.2.). L'énoncé est une question suspensive qui permet à l'étudiante de demander à l'agent de trouver des solutions sans le faire explicitement. Enfin, je vais mentionner le tour *je peux pas je peux pas du tout* (ligne 11), produit par l'étudiante qui présente une auto-reprise réitérative (André 2010b). Cette reprise permet d'insister sur l'impossibilité et de donner à celle-ci un caractère problématique. Cette pratique interactionnelle réalise ainsi plusieurs activités

interactionnelles (informer et insister) et a un véritable effet pragmatique puisque l'agent va tenter de trouver une solution (cette partie n'apparaît pas dans l'extrait de la figure 20). L'analyse des interactions a bien montré que la reprise participe à l'interaction, elle est pourtant souvent exclue de l'enseignement/apprentissage de la langue. Cette brève analyse, exposée ici rapidement et partiellement, donne à voir aux apprenants le fonctionnement du français parlé en interaction et la façon dont les locuteurs réalisent les activités visées par la communication.

En réalisant cette analyse, les apprenants peuvent observer les pratiques interactionnelles utilisées et leur cotexte d'apparition, les mécanismes de prises de paroles et d'alternance des tours, les activités interactionnelles et leurs effets pragmatiques, les gestes et les postures qui accompagnent le discours, les aspects extralinguistiques qui orientent l'interaction ou encore la construction collaborative du discours avec ses ajustements et ses coopérations. Il est possible également de procéder à ces analyses en se concentrant sur un aspect des interactions verbales.

Dans les exemples suivants, les apprenants ont eu pour tâche d'observer des activités interactionnelles précises. Dans le premier exemple, l'apprenant devait observer les formules de politesse et toutes les pratiques interactionnelles relevant de la politesse. Voici l'interaction de FLEURON qui a servi de ressource :

Retour
Commander un thé à la cafétéria

Media
Description



Masquer les sous-titres

E: bonjour je vais vous prendre un thé à la menthe s'il vous plait
A: d'accord 75 centimes
A: merci
(l'agent range la monnaie et prépare le thé)
A: et voilà
E: merci
A: bonne matinée
E: au revoir
A: au revoir

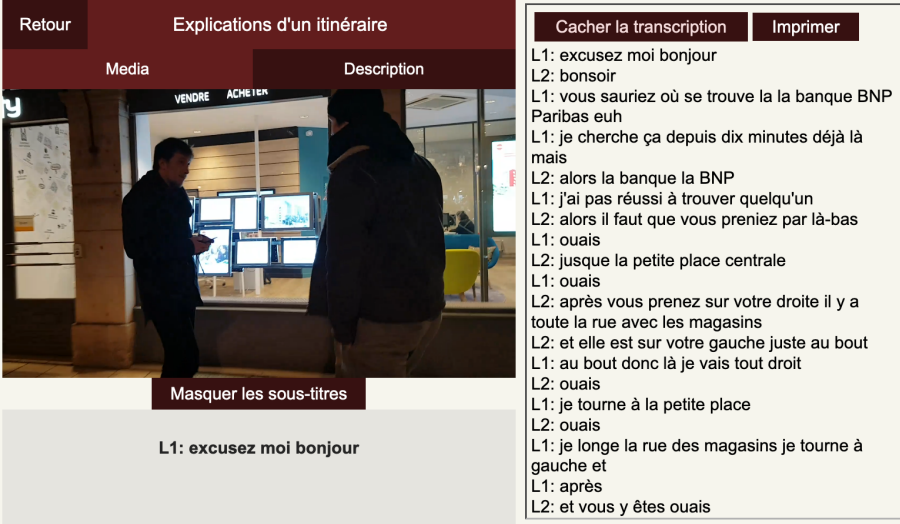
Cacher la transcription
Imprimer

Cacher la transcription
Imprimer

Figure 21 : Interaction dans la catégorie « Vie sur le campus » de FLEURON

Avec cette interaction, les apprenants peuvent observer la place occupée par la politesse dans les transactions commerciales (Kerbrat-Orecchioni 2001b, Traverso 2001), ici à la cafétéria du campus. Des observations concernant un objectif communicatif précis peuvent

également être effectuées selon le même procédé. Par exemple, les apprenants peuvent étudier comment demander son chemin et apprendre à comprendre les réponses avec la ressource suivante :



The screenshot displays a language learning interface. On the left, a video shows two men in a shop; one is looking at a screen while the other stands nearby. Above the video is a navigation bar with 'Retour', 'Explications d'un itinéraire', 'Media', and 'Description'. Below the video is a button labeled 'Masquer les sous-titres'. To the right of the video, a transcript of the conversation is shown, with buttons for 'Cacher la transcription' and 'Imprimer' at the top.

Retour **Explications d'un itinéraire** **Media** **Description**

Cacher la transcription **Imprimer**

L1: excusez moi bonjour
 L2: bonsoir
 L1: vous sauriez où se trouve la la banque BNP Paribas euh
 L1: je cherche ça depuis dix minutes déjà là mais
 L2: alors la banque la BNP
 L1: j'ai pas réussi à trouver quelqu'un
 L2: alors il faut que vous preniez par là-bas
 L1: ouais
 L2: jusque la petite place centrale
 L1: ouais
 L2: après vous prenez sur votre droite il y a toute la rue avec les magasins
 L2: et elle est sur votre gauche juste au bout
 L1: au bout donc là je vais tout droit
 L2: ouais
 L1: je tourne à la petite place
 L2: ouais
 L1: je longe la rue des magasins je tourne à gauche et
 L1: après
 L2: et vous y êtes ouais

Masquer les sous-titres

L1: excusez moi bonjour

Figure 22 : Interaction dans la catégorie « Vie en dehors du campus » de FLEURON

Dans une méthodologie plus traditionnelle, l'activité pédagogique qui consiste à demander son chemin est habituelle. Cependant, elle est rarement appuyée sur des interactions authentiques, notamment parce qu'il en existe peu de disponibles. Ces activités prennent souvent la forme de jeux de rôles. Les locuteurs-apprenants imaginent ce qu'ils pourraient dire dans une telle situation. L'analyse de la progression de cette interaction permet aux apprenants d'observer les usages réels de la langue dans cette situation précise.

Dans d'autres cas, les apprenants peuvent apprendre à réaliser une activité interactionnelle plus précise, telle qu'exposer un problème. Cette activité est fréquente en situation universitaire, beaucoup de ressources du corpus présentent des situations dans lesquelles un étudiant rencontre un problème, comme c'est le cas dans la ressource ci-dessous.

Retour
Problème d'emploi du temps (2)

Media
Description

Masquer les sous-titres

E: j'aimerais avoir un renseignement sur les emplois du temps de il y a deux semaines

Cacher la transcription
Imprimer

A: oui entrez
A: bonjour
E: bonjour
E: j'aimerais avoir un renseignement sur les emplois du temps de il y a deux semaines
E: parce que je vais avoir des conflits
A: bon
A: vous êtes en quelle licence
E: en L2 sociologie
A: en socio on va regarder si je peux pas vous renseigner je vous
A: je vous mettrais vers ma collègue mais déjà on va regarder là
A: est-ce que vous avez votre numéro étudiant s'il vous plait
E: numéro étudiant
E: XXX
A: XXX
A: XXX c'est ça
E: c'est ça hum
E: et c'est la semaine du cinq avril
A: alors je vais pas avoir l'em- votre emploi du temps exact mais je l'ai en général

Figure 23 : Interaction dans la catégorie « Démarches administratives » de FLEURON

L'analyse de cette interaction permet aux apprenants de repérer les pratiques interactionnelles qui permettent de demander un renseignement et d'exprimer un problème d'emploi du temps, au service scolarité d'une université.

Cette façon d'entrer dans le corpus FLEURON est utilisée fréquemment par les enseignants. Elle s'apparente à l'utilisation plus classique des documents authentiques mais ne propose pas d'activités ou d'exercices comme il est habituel de le faire, même dans des approches modernes, comme l'approche communicative ou actionnelle. Ce que je propose ici c'est une analyse de la progression et du déroulement de l'interaction, tour par tour, pratique interactionnelle par pratique interactionnelle, activité interactionnelle par activité interactionnelle, avec une attention au genre de discours, à la situation de communication et à la construction collaborative du discours. Cette analyse peut être complétée par des études détaillées de certains phénomènes spécifiques à l'oral et par l'utilisation d'autres ressources, telles que PFC-EF, FLORALE ou CLAPI-FLE. Ces ressources viennent préciser l'analyse du déroulement de l'interaction en renvoyant l'apprenant ou l'enseignant vers d'autres interactions. Il est également possible de compléter cette analyse par la recherche d'une pratique interactionnelle, d'un mot ou d'un groupe de mot, dans un corpus grâce à un concordancier, comme je le montrerai dans le chapitre 5.

4.3.2. Avec des adultes migrants peu scolarisés

La formation linguistique des adultes migrants fait partie de nombreuses questions de recherche et de travaux que je mène en collaboration avec Hervé Adami (Adami, André 2013 ; Adami, André 2012b ; Adami, André 2010). Ces questions ont également contribué à la naissance de notre groupe de recherche LTF. C'est donc naturellement que je me suis tournée vers ce terrain pour expérimenter l'exploitation de corpus avec un public FLI (français langue d'intégration). FLEURON me paraissait être un bon outil, bien qu'il n'ait pas été pensé pour ce public, pour deux principales raisons. Tout d'abord, FLEURON présente des situations de communication authentiques qui concernent des interactions auxquelles les migrants peuvent être amenés à participer. Par exemple, le corpus FLEURON comporte des interactions prenant place en dehors d'un campus universitaire, telles que des interactions à la gare, dans des bibliothèques, des musées ou plus largement dans des structures de la ville non réservées aux étudiants. Ces interactions authentiques sont rares et représentent une précieuse ressource pour les formateurs. Je ne reviendrai pas sur l'importance d'exposer les apprenants à la langue réelle et d'utiliser des documents authentiques mais je précise simplement qu'avec un public FLI cette importance devient capitale. Ces apprenants sont déjà en immersion dans la langue cible (Adami 2009) et ont conscience, en permanence, de la complexité des interactions verbales. Ils doivent être capable de la comprendre parce que les conséquences de leurs insécurités langagières sont considérables en termes d'intégration et de gestion du quotidien. Ensuite, FLEURON semble être un bon outil pour ce public dans la mesure où il présente une interface agréable avec des ressources multimodales. Les rares manuels existants pour ce public n'accordent pas une grande importance à ce type de ressources. Même les sites les plus avancés en termes d'application de la recherche rencontrent des difficultés pour en proposer. Je pense notamment au site « Français langue d'intégration et d'insertion » que le groupe LTF a développé⁶¹. Nous nous efforçons de proposer des activités pédagogiques à partir de documents authentiques mais ils restent difficiles à trouver. Ainsi, un corpus comme FLEURON peut venir en complément d'autres ressources pédagogiques.

L'Office Français de l'Immigration et de l'Intégration (OFII) est l'opérateur de l'État, sous la tutelle du ministère de l'intérieur, qui s'occupe de l'immigration. La prise en charge de l'accueil et de l'intégration des migrants fait partie de ses missions. Il fait signer un

⁶¹ <https://fli.atilf.fr/>

Contrat d'Intégration Républicaine (CIR) aux nouveaux arrivants souhaitant et étant autorisés à s'installer durablement en France. La signature du CIR implique pour les migrants, qui en ont besoin, de suivre une formation linguistique en français. Cette formation est dispensée par des organismes de formation qui doivent répondre à des appels d'offre. En Lorraine, l'appel d'offre de l'OFII a été obtenu en 2019 par l'ALAJI⁶², un organisme de formation, en partenariat avec le réseau des GRETA. C'est auprès des stagiaires⁶³ du GRETA Lorraine que j'ai pu mener les expérimentations que je vais présenter ici. Les groupes sont composés de quinze personnes, évaluées au niveau A2 par les tests d'entrée lors de la signature du CIR (ce qui nous donne en réalité peu d'informations sur leur véritable niveau en français). Les stagiaires sont originaires de différents pays : Afghanistan, Albanie, Érythrée, Italie, Niger, Nigeria, Surinam, Soudan, Tibet.

Les formations commanditées par l'OFII doivent strictement respecter les directives mentionnées dans l'appel d'offre. Parmi ces directives, sont spécifiés les thèmes qui doivent être abordés (ces thèmes sont, en partie, issus du Référentiel FLI⁶⁴). Ces thèmes sont notamment : le logement, la santé, les déplacements, la culture, la vie professionnelle et les symboles de la République. Les expérimentations menées au GRETA avec un public FLI devaient s'inscrire dans ces thématiques. C'est pourquoi, avec Charlène Weyh, en stage de master 2 au GRETA, nous avons choisi d'utiliser les ressources FLEURON correspondant aux thématiques du FLI en guise d'introduction de l'apprentissage sur corpus. Pour chacune des séances consacrées à l'utilisation de FLEURON, une ressource a été exploitée en utilisant l'analyse sociolinguistique des interactions. Nous avons fait observer le déroulement des interactions aux stagiaires, le genre de discours, les activités visées et les pratiques interactionnelles mises en œuvre.

Pour introduire la première séance, nous avons utilisé une vidéo de FLEURON qui présente une interaction à la gare, un homme s'adressant à une agente SNCF pour acheter un billet de train. Une activité traditionnelle de compréhension orale a été effectuée collectivement. Cette activité classique (compréhension globale puis détaillée, sans les sous-titres puis avec les sous-titres) a permis de faire connaissance avec le public. Ensuite, j'ai proposé de réaliser une activité collective d'analyse sociolinguistique des interactions avec cette même ressource. Ainsi, les apprenants, en fonction de leurs compétences socio-

⁶² <http://www.alaji.fr/>

⁶³ Les adultes qui suivent ce type de formation sont appelés « stagiaires ». Par commodité, j'utilise cette dénomination.

⁶⁴ Consultable ici : <https://fli.atilf.fr/wp-content/uploads/2015/04/R%c3%a9f%C3%A9rentiel-FLI.pdf>

interactionnelles et avec mon aide, ont commenté les tours de parole et ont pointé ceux qui les ont surpris.

Retour
Acheter un billet de train au guichet SNCF

Media
Description



Masquer les sous-titres

A: un aller-retour Epinal Thaon vous avez une carte de réduction

Cacher la transcription
Imprimer

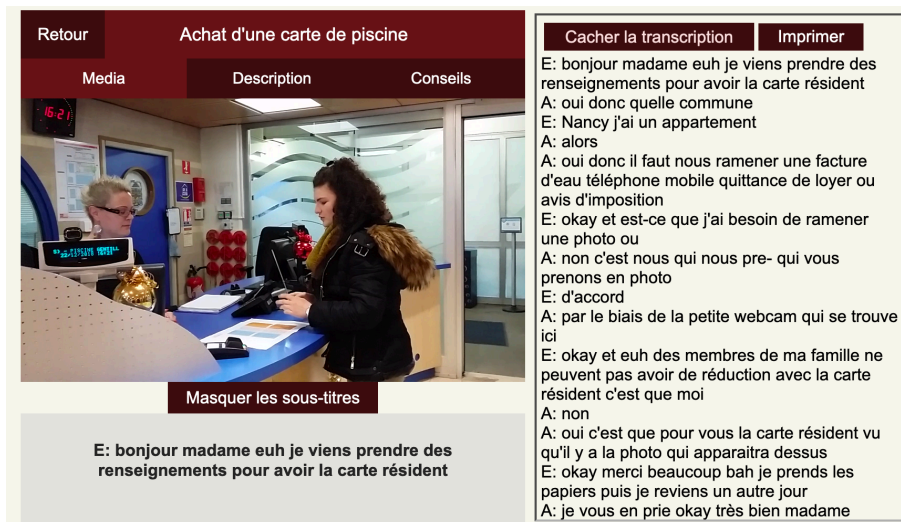
E: bonjour
A: bonsoir
E: ce serait pour un aller-retour Epinal Thaon s'il vous plait
A: un aller-retour Epinal Thaon vous avez une carte de réduction
E: pas du tout
(l'agent tape sur l'ordinateur)
A: donc ça fera 3 euros 80 s'il vous plait monsieur
A: merci
A: je vous dois 16 euros 20
A: pour vous
A: vous n'oubliez pas de composer avant le départ
A: d'Epinal et de Thaon
E: ok
A: voilà
E: au revoir
A: bonne soirée au revoir

Cacher la transcription
Imprimer

Figure 24 : Interaction à la SNCF dans FLEURON

Les stagiaires ont pu repérer les formules de politesse, l'ouverture et la clôture de l'interaction, la formulation de la demande, le prix du billet. En outre, ce fut l'occasion d'aborder des éléments culturels de la vie en France, tels que les transports en commun, le guichet de la SNCF et le compostage des billets de train. Les éléments repérés varient d'un stagiaire à l'autre, en fonction de leur niveau en langue et de leurs connaissances de la situation présentée. L'objectif est le même qu'avec des apprenants plus avancés, comme les étudiants des expérimentations présentées dans le point précédent, cependant les phénomènes pointés sont différents. Ces expérimentations avec un public migrant permettent ainsi de montrer qu'il n'existe pas de ressource compliquée ou réservée à des apprenants avancés. Ce sont les activités qui sont associées aux documents qui varient en fonction du niveau. L'analyse de la progression des interactions suit ce même procédé. En fonction du niveau des stagiaires, les pratiques interactionnelles observées et les activités réalisées ne seront pas les mêmes.

Toujours dans le cadre d'une activité collective d'analyse des interactions, lors d'une autre séance, la ressource intitulée « Achat d'une carte de piscine » a été utilisée (catégorie « Vie en dehors du campus »). Voici l'interaction :



Retour **Achat d'une carte de piscine**

Media **Description** **Conseils**

Masquer les sous-titres

Cacher la transcription **Imprimer**

E: bonjour madame euh je viens prendre des renseignements pour avoir la carte résident
A: oui donc quelle commune
E: Nancy j'ai un appartement
A: alors
A: oui donc il faut nous ramener une facture d'eau téléphone mobile quittance de loyer ou avis d'imposition
E: okay et est-ce que j'ai besoin de ramener une photo ou
A: non c'est nous qui nous pre- qui vous prenons en photo
E: d'accord
A: par le biais de la petite webcam qui se trouve ici
E: okay et euh des membres de ma famille ne peuvent pas avoir de réduction avec la carte résident c'est que moi
A: non
A: oui c'est que pour vous la carte résident vu qu'il y a la photo qui apparaîtra dessus
E: okay merci beaucoup bah je prends les papiers puis je reviens un autre jour
A: je vous en prie okay très bien madame

E: bonjour madame euh je viens prendre des renseignements pour avoir la carte résident

Figure 25 : Interaction dans la catégorie « Vie en dehors du campus » de FLEURON

Les apprenants ont analysé tour de parole par tour de parole cette relation de service. Ils ont observé les salutations, la façon d'exposer la demande, de poser des questions, de répondre aux questions, etc. On peut remarquer quelques éléments intéressants dans cette vidéo, des éléments qui peuvent être abordés par les enseignants mais leur contextualisation peut être utile :

- une demande de renseignement ;
- une assertion pour demander quelque chose : « je viens prendre des renseignements » ;
- le peu de politesse dans cet échange de service ne nuit pas à l'interaction (par exemple, l'agent (A) ne répond pas au bonjour de E) ;

et des éléments qui ne sont jamais abordés en formation linguistique :

- une question alternative inachevée en « ou » ;
- un retour sur l'axe syntagmatique « c'est nous qui nous pre- qui vous prenons en photo » ;
- des petits mots pour commencer un tour de parole « alors oui donc » ou « non oui ».

Ces éléments ont naturellement attiré l'attention des stagiaires. En outre, des éléments administratifs importants ont été abordés puisqu'il est question dans cette vidéo de documents qui permettent d'attester d'un domicile. Ces éléments font partie du quotidien des adultes migrants qui effectuent de nombreuses démarches administratives (recherche d'un logement, ouverture d'une ligne téléphonique, abonnement de transport, etc.).

Avec des stagiaires débutants, les ressources peuvent être utilisées pour travailler un point précis d'une interaction. Par exemple, la vidéo ci-dessous a été utilisée pour repérer les horaires d'ouverture de la médiathèque :

Retour Médiathèque de Nancy - Ouverture

Media Description

Masquer les sous-titres

donc la médiathèque vous pouvez venir ici aux horaires d'ouverture du mardi au samedi de 11h à 18h

Cacher la transcription Imprimer

donc la médiathèque vous pouvez venir ici aux horaires d'ouverture du mardi au samedi de 11h à 18h
le dimanche de 14h à 18h
sauf en été hein le dimanche nous fermons du 15 juin au 15 septembre
vous vous nous retrouvez aussi dans d'autres bibliothèques de quartier
au Haut-du-Lièvre et à celle de Nabécor
sachez aussi que vous pouvez donc venir visiter la médiathèque un samedi sur deux renseignez vous au niveau des de l'agenda culturel au CAN donc un samedi sur deux à 11h15 on peut vous faire une visite pendant une demi-heure
et on voit tous les tous les étages avec vous et vous pouv- vous aurez l'occasion de poser des questions
à bientôt

Cacher la transcription Imprimer

Figure 26 : Interaction dans la catégorie « Culture et lieux de culture » de FLEURON

La compréhension des horaires est un problème récurrent pour des migrants qui sont souvent peu habitués à des contraintes temporelles (souvent parce qu'elles n'existent pas dans leur pays d'origine). Par conséquent, ce travail de compréhension arrive très tôt dans leur formation linguistique afin d'éviter des rendez-vous manqués ou des files d'attente après les horaires d'ouverture des bâtiments.

Quelles que soient les vidéos utilisées avec un public en voie d'intégration langagière, si elles correspondent thématiquement à ses besoins, elles permettent de faire une analyse sociolinguistique de la progression des interactions verbales. Cette analyse est réalisée conjointement avec le ou la formateur·trice, comme avec un public plus avancé. En d'autres termes, quel soit le niveau des apprenants, les ressources de FLEURON peuvent être utilisées à des fins d'apprentissage de la langue. Elles sont néanmoins utilisées différemment selon le niveau des apprenants. Par exemples, les phénomènes interactionnels pointés sont différents. Le lexique est davantage travaillé. Les constructions syntaxiques le sont un peu moins. L'analyse des interactions permet également aux apprenants de repérer les éléments phatiques de la communication qui varient d'une culture à l'autre et, en France, d'une situation à l'autre.

Conclusion du chapitre

Exploiter les résultats de l'analyse sociolinguistique des interactions verbales à des fins d'enseignement et apprentissage de la langue permet à la didactique d'effectuer deux actions intimement liées. Tout d'abord, la première consiste à se doter d'outils utiles pour comprendre finement et scientifiquement le déroulement d'une interaction verbale et ses intrications à sa situation de communication. La didactique peut exploiter les analyses des corpus d'interactions effectuées dans différents travaux de recherche. Cette pratique existe dans certains paysages pédagogiques et semble bien répandue pour l'enseignement et l'apprentissage de l'anglais langue étrangère (Sinclair 2004, Breyer 2011), mais elle est quasiment inexistante en France. Comme le soulignent Kamber et Dubois (2016 : 5) : « du côté français, on constate qu'il existe un écart parfois abyssal entre les résultats de la recherche récente et les représentations de la langue dans les dictionnaires, les grammaires et les manuels ». Ensuite, la seconde consiste à emprunter à la sociolinguistique des interactions, sa méthodologie d'analyse. L'enseignant propose à l'apprenant, quel que soit son niveau, de procéder à une observation du déroulement de l'interaction en pointant certains phénomènes, certaines pratiques interactionnelles, certaines activités interactionnelles, au fur et à mesure de leur apparition dans la progression de l'échange. L'objectif n'est pas de transformer l'apprenant en analyste de la langue mais de l'aider à comprendre le fonctionnement de la langue parlée en interaction. Ainsi, la répétition de l'observation des situations de communication, c'est-à-dire l'observation des données en corpus, amène l'apprenant à prendre conscience des phénomènes interactionnels, de leur forme et de leur fonction, et à acquérir des compétences socio-interactionnelles. L'objectif n'est pas non plus de demander à des enseignants de FLE d'être des analystes de la langue. C'est pourquoi les transpositions des résultats des analyses des interactions à des fins didactiques, pour des non spécialistes de la langue, sont indispensables.

En outre, l'exposition à des corpus d'interactions permettrait de remplacer une immersion en milieu naturel qui, comme le souligne Howard (2010), constitue des conditions très propices à l'acquisition de la variation et des compétences socio-interactionnelles. Les enjeux résident également dans la formation des enseignants de langue. Elle est inévitablement indispensable. Néanmoins, les ressources qui voient le jour actuellement permettent d'aider les enseignants à exploiter les travaux des chercheurs.

Ces ressources sont complémentaires et peuvent être sollicitées lors de l'analyse d'une interaction. Le schéma ci-dessous tente d'illustrer cette démarche complémentaire.

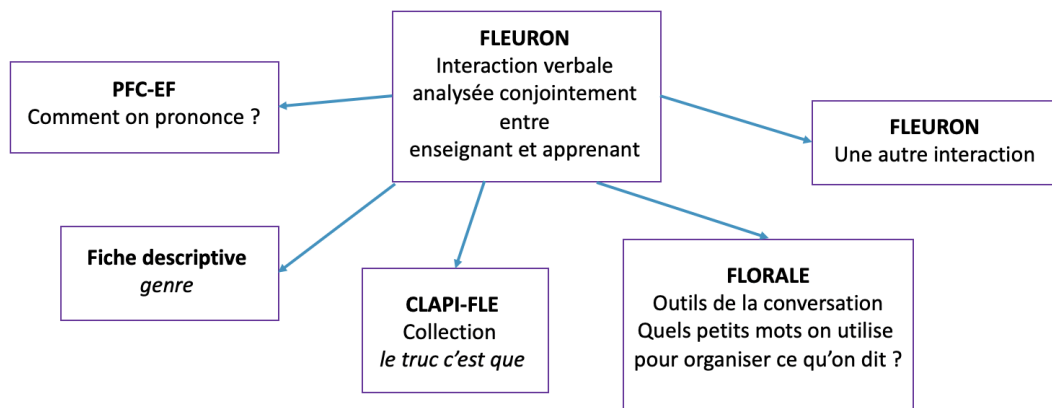


Figure 27 : Complémentarité des ressources pour une exploitation didactique

Ainsi, lors de l'analyse sociolinguistique d'une interaction verbale, réalisée conjointement entre un enseignant et un apprenant, il est possible de faire appel à d'autres ressources. Elles permettent d'observer plus amplement et de comprendre ces mêmes phénomènes, rencontrés dans d'autres corpus et dans d'autres situations. Ce schéma donne des exemples d'éléments à observer mais bien d'autres peuvent faire l'objet d'une complémentarité de données. Les pratiques interactionnelles rencontrées dans une situation de communication peuvent être expliquées de façon plus détaillée dans d'autres outils. Par ailleurs, il est également possible, pour compléter l'observation de l'interaction, de rechercher une pratique interactionnelle dans d'autres données en corpus par l'intermédiaire d'un concordancier. C'est cette démarche, intitulée *data-driven learning* ou apprentissage sur corpus, qui fera l'objet du chapitre suivant.

Chapitre 5. L'apprentissage sur corpus : un concordancier au service de l'enseignement et de l'apprentissage de la langue

Les chapitres précédents ont montré les liens entre l'analyse sociolinguistique des interactions verbales, les données à recueillir, certains phénomènes interactionnels analysés et la didactique de l'oral en FLE. Ces liens se concrétisent notamment en conduisant des analyses conjointes, des ressources multimodales du corpus FLEURON, entre des enseignants et des apprenants, afin que ces derniers acquièrent des compétences socio-interactionnelles. Ces liens se tissent donc pleinement dans l'exploitation de corpus à des fins didactiques et dans la transposition didactique des résultats de l'analyse sociolinguistique des interactions verbales. L'aboutissement de l'ensemble de mes travaux et de mes rencontres scientifiques trouve ses applications et ses implications dans ces propositions d'exploitation des corpus oraux et multimodaux en didactique. Dans ce dernier chapitre, je vais aborder une autre façon d'exploiter des corpus à des fins didactiques – en utilisant un concordancier, outil de la linguistique de corpus. Dans les travaux anglo-saxons, cette approche a été nommée *data-driven learning* par Johns (1991). Elle est née de la volonté de trouver des applications à l'exploitation linguistique des corpus dans le processus d'enseignement et d'apprentissage des langues. En d'autres termes, cette approche vise à mettre un concordancier au service de l'enseignement et de l'apprentissage de la langue. Elle a donné naissance à de nombreuses expériences et travaux scientifiques mais rarement concernant l'exploitation des corpus oraux, à part quelques exceptions (voir par exemple Zorzi 2001 ; Mauranen 2004 pour l'anglais ou Kerr 2013 pour le français), et encore plus rarement des corpus multimodaux (Braun 2007). A ma connaissance, cette approche n'a jamais été expérimentée en France avec des corpus d'interactions pour apprendre à interagir en français.

Dans ce chapitre, je vais présenter le *data-driven learning* (ou apprentissage sur corpus, en français) utilisé actuellement presque exclusivement à des fins d'enseignement et d'apprentissage de la langue anglaise écrite. L'objectif n'est pas de faire l'état de l'art ou l'histoire épistémologique de cette approche (à ce propos voir notamment Chambers 2010 ; Cobb, Boulton 2015 ; Boulton 2017 ; Boulton, Cobb 2017) mais d'en saisir ses principes pour les utiliser et les adapter à une exploitation didactique des corpus d'interactions. Je préciserai les principes essentiels du *data-driven learning*, notamment ceux qui m'ont été

utiles dans les expérimentations que j'ai menées auprès d'apprenants de FLE (5.1.). J'expliquerai ensuite comment l'analyse des résultats d'un concordancier, ou des lignes de concordances, peut être une analyse sociolinguistique, en redonnant toute son importance à la situation de communication (5.2.). Enfin, les derniers points seront consacrés à la présentation de diverses expérimentations qui mettent au jour le processus analytique déclenché par l'analyse de listes de concordances. L'objectif est d'observer comment les apprenants s'approprient les principes du *data-driven learning* et s'ils réussissent à en tirer profit. Je présenterai l'utilisation du corpus FLEURON interrogé par son concordancier avec des étudiants étrangers (5.3.) puis des usages détournés du corpus FLEURON (5.4.) avec des adultes migrants et des adultes francophones natifs en situation d'illettrisme. L'analyse de ces utilisations d'un concordancier à des fins d'enseignement et d'apprentissage du français parlé en interaction aboutira à différentes pistes pour une mise en œuvre de cette méthodologie didactique.

5.1. *Data-driven learning* (DDL) ou apprentissage sur corpus (ASC)

Le *data-driven learning* (DDL) est traduit en français par apprentissage sur corpus (ASC) par Boulton et Tyne (2014). Dans cette section, je vais présenter les principes fondamentaux de l'ASC qui peuvent être synthétisés et résumés de la façon suivante : accès direct aux données par l'intermédiaire d'un concordancier, approche inductive, redéfinition des rôles de l'apprenant et de l'enseignant.

La première mention de la possible exploitation outillée de corpus à des fins didactiques date du milieu des années 1980 avec la publication de Higgins et Johns (1984) et l'utilisation du terme *classroom concordancing*. Il s'agit d'utiliser un concordancier, un outil informatique créé pour faire de la linguistique de corpus qui permet de rechercher les occurrences d'un ou de plusieurs items dans une grande masse de données. Les résultats sont affichés verticalement, en alignant des lignes de concordances (voir un exemple figure 28, ci-dessous). Dans une perspective didactique, ces lignes sont analysées par des apprenants de langue en salle de classe. C'est le principe fondamental du *data-driven learning* (DDL), dont la première apparition du terme se trouve dans un article de Johns paru en 1990, article qu'il reprend un an plus tard dans l'ouvrage fondateur qu'il coédite avec

Philip King (Johns, King 1991). Dans leur introduction, ils posent une première définition du *data-driven learning* qui peut être décrit comme :

an application of computers to language-learning that has come to be known as 'classroom concordancing' or 'data-driven learning' (DDL) – the use in the classroom of computer-generated concordances to get students to explore the regularities of patterning in the target language, and the development of activities and exercises based on concordance output. (Johns, King 1991 : iii).

Même si les activités et les exercices issus de l'extraction des lignes de concordances sont nombreux et variés, le principe didactique fondateur de cette approche est de donner accès aux apprenants à des données de corpus, qu'ils doivent étudier afin de saisir les régularités et la structure de la langue. Les premières exploitations de corpus à des fins didactiques ne sont pas nées avec les travaux de Johns à partir de la fin des années 1980, mais de différentes initiatives méthodologiques en didactique (voir notamment Higgins, Johns 1984 ; Ahmad *et al.* 1985 ; Johns 1986, 1988). Ces premiers travaux sont tous anglo-saxons et concernent l'enseignement et l'apprentissage de l'anglais, et presque uniquement de l'anglais écrit.

L'association d'un corpus et d'un concordancier pour interroger les données est le principe technologique du DDL. L'objectif est d'exposer les apprenants à la langue réelle et de les amener à une sorte d'activité de découverte. Cette activité se fait grâce à des lignes de concordances, extraites par un concordancier et proposant les occurrences des items recherchés, avec leur contexte droit et gauche. Johns (1991 : 2) décrit ce procédé :

The most important computing tool for the data-driven approach is the concordancer, which is able to recover from text all the contexts for a particular item (morpheme, word or phrase) and to print them out in a way which facilitates rapid scanning and comparison. The most usual format is the keyword-in-context (KWIC) concordance in which the keywords are arranged one below the other down the centre of the page, with a fixed number of characters of context to the left and to the right.

Avec ces premiers principes du DDL, nous assistons à une méthodologie tout à fait nouvelle en didactique des langues.

Les utilisations qui en sont faites prennent de multiples formes et peuvent accueillir différentes approches didactiques. Le DDL est alors défini comme un couteau suisse, Johns et King le décrivent ainsi : « as useful a tool in their own way as the Swiss army knife » (1991 :

iii). Johns explique également à quel point le DDL est un outil puissant pour observer les usages de la langue dans la mesure où il donne aux apprenants de langue un accès direct aux données, ou plus précisément aux corpus de données langagières.

L'accès aux données est accompagné de certains principes qui régissent le DDL, même si celui-ci peut être abordé de différentes manières. Tout d'abord, l'approche offerte par l'observation des données en corpus est une approche inductive (Johns 1991 ; Tognini-Bonelli 2001). Cette approche est illustrée par une métaphore :

The central metaphors embodying the approach are those of the learner as 'linguistic researcher', testing and revising hypotheses, or as a 'language detective', learning to recognise and interpret clues from context ('Every student a Sherlock Holmes') (Johns 1997 : 101).

C'est cette conception méthodologique qui a attiré et retenu mon attention. Dans les phases préconisées par Johns (1991, 1997) pour pratiquer le DDL, je retiens et reformule les trois suivantes⁶⁵ :

- Observation des données d'un corpus *via* un concordancier ;
- Classification des données ;
- Généralisation des principes de fonctionnement à partir des observations et classifications effectuées.

Ainsi, les apprenants observent tout d'abord les items qu'ils ont recherchés ou qu'un enseignant leurs a suggéré ou demandé d'observer. Ils analysent leur distribution, c'est-à-dire qu'ils comparent leurs conditions d'apparition et leurs usages. Ensuite, ils identifient les items qui fonctionnent de la même façon, qui sont toujours précédés ou suivis des mêmes éléments, qui sont placés à tel ou tel endroit dans un énoncé, *etc.* et font des catégories. Enfin, ils induisent les règles de fonctionnement des items recherchés. Ainsi, les apprenants deviennent des détectives de la langue, des Sherlock Holmes voire des apprentis-chercheurs. Gavioli (2001) et Bernardini (2004) parlent de *learner as researcher* pour évoquer les activités dans lesquelles s'investissent les apprenants en analysant des lignes de concordances.

Les apprenants partent donc à la découverte des données puis analysent leur cotexte d'apparition dans des lignes de concordances. L'ASC amène à repenser

⁶⁵ Johns (1991 : 4) propose les phases suivantes : *Identify – Classify – Generalise*.

l'enseignement et l'apprentissage des langues, notamment la relation entre l'enseignant et l'apprenant (Leech 1997). Chambers (2010 : 355) l'explique clairement :

The role of the teacher also changes fundamentally, as s/he is no longer the sole source of knowledge about target language, but rather a facilitator of the learning process, helping the learners to interpret the data, and giving them advice on how best to search the corpus and analyse their results.

L'enseignant joue donc un rôle tout à fait particulier, différent de celui qu'il peut jouer dans des approches traditionnelles. Il accompagne les activités et guide l'apprenant. Johns (1991 : 1) explique également concrètement que :

the task of the learner is to "discover" the foreign language, and that the task of the language teacher is to provide a context in which the learner can develop strategies for discovery - strategies through which he or she can "learn how to learn".

L'enseignant devient un guide, un conseiller, « a coordinator of research, or facilitator » (Bernardini 2004 : 16). En apprenant comment apprendre, l'apprenant s'autonomise au fur et à mesure de ses expériences. A ce propos, Aston (2001 : 41) souligne : « Perhaps the greatest attraction of corpora for language pedagogy is their potential for autonomous learning ». Cet aspect de la méthodologie didactique est encore un point de rencontre entre mon environnement scientifique au sein de l'équipe « Didactique des langues et sociolinguistique » et mes propres intérêts. Cette équipe a mené des travaux pionniers dans ce que l'on peut appeler l'apprentissage des langues en autonomie ou en auto-direction⁶⁶. Même si je ne me suis jamais réellement penchée sur ce domaine, les liens me paraissent aujourd'hui évidents. L'ASC permet aux apprenants ainsi qu'aux enseignants de choisir les modalités d'utilisation des données, par exemple choisir un corpus approprié aux objectifs d'apprentissage fixés. L'apprenant est actif dans sa démarche d'apprentissage et apprend à apprendre.

Dans le cadre de cette méthodologie, la question des données et de leur authenticité se pose de la même façon que dans le chapitre précédent. Si l'importance de l'exposition à la langue cible ne fait plus aucun doute en didactique, il existe différentes façons d'être exposé ou de s'exposer. L'introduction des documents authentiques, notamment avec

⁶⁶ Voir les nombreux articles parus depuis les années 1970 dans la revue de l'équipe (ex-CRAPEL), disponible en ligne : <http://www.atilf.fr/spip.php?rubrique217>

l'approche communicative répondait déjà à cette exigence d'exposition à la langue réelle, telle qu'elle est pratiquée par les locuteurs dans des situations de communication différentes. Dans ce domaine, comme je l'ai déjà expliqué, mon équipe de recherche a largement contribué à promouvoir les documents authentiques (voir notamment Abé *et al.* 1979 ; Duda, Esch 1972 ; Holec 1990 ; le numéro spécial des *Mélanges* 2009). C'est dans ce bain que ma formation en didactique s'est faite, de façon naturelle étant donné mon approche sociolinguistique de la langue. Je ne discuterai pas ici du caractère authentique des données dès qu'elles sont recueillies en corpus, ou tout simplement prélevées à leur situation d'apparition. J'ai été invitée à Genève en 2016, à l'occasion des 125 ans du FLE, pour discuter cet aspect lors d'une table ronde entièrement consacré à l'authenticité (André à paraître 2019). Même si cette notion ainsi que le débat qui l'accompagne est ancien, il est toujours d'actualité ou le redevient régulièrement (Widdowson 1998, Adami 2009b, Chambers 2009) notamment dans l'enseignement et l'apprentissage d'autres langues que le FLE⁶⁷. Néanmoins, il est bien question de données authentiques (c'est-à-dire, non construites à des fins didactiques) lors que l'on choisit d'observer et de faire observer avec des corpus. Je reviendrai sur le choix des corpus opéré dans les différentes expérimentations que j'ai menées dans ce domaine. Ces données authentiques sont rassemblées en corpus, cela signifie que l'apprenant et l'enseignement accèdent à un grand nombre de données et les manipulent avec différents outils.

Je vais illustrer la démarche de l'ASC avec des exemples d'interrogation de corpus oraux avec des concordanciers appropriés. L'exploitation de corpus oraux nécessite des outils qui sont capables de traiter et de synchroniser le son et la transcription. Ces exemples de recherches d'occurrences ont été utilisés lors d'expérimentations avec des apprenants de FLE. Tout d'abord, à l'instar de la méthodologie décrite par Johns (1986) qui utilisait le logiciel *MicroCord* pour extraire les occurrences souhaitées, une liste de concordances est donnée aux apprenants, en voici un exemple que j'ai donné au cours d'une expérimentation :

⁶⁷ Par exemple, en 2019 avait lieu à l'université de Lorraine, un colloque intitulé : « Quelle authenticité pour l'enseignement de l'anglais oral ? », organisé par le laboratoire IDEA (Interdisciplinarité dans les études anglophones).

ref	Contexte gauche	Pivot	Contexte droit
Conv_cal_06...	ça la carte de bus tu peux la	prendre	< c'est pas le problème non non mais > bon je
Conversation...	alors elle m'envoyait un message je vais	prendre	me plasse clope ah moi aussi il y a pas de soucis
Conversation...	je l'avais appelé euh c'était pour	prendre	de ses nouvelles mais surtout pour parler pour euh pour m'alléger
Conversation...	termine ma valise j'ai pas envie de	prendre	le train de vingt-et-une heures ça fait trop tard dix heures et
Conversation...	c'est clair bon allez je vais aller	prendre	le train
Conversation...	> squatterais plutôt la terrasse si tu veux	prendre	un chocolat chaud ou un vin chaud mais euh faire du ski
Cro_cons_mu...	euh dans cet article trois l'hypothèse de	prendre	les deux opérateurs euh consultés déposent chacun un dossier les deux propositions
Cro_cons_mu...	il fallait à un moment donné euh ben	prendre	euh une décision qui permette d'avancer il me semble que celle
Deb_lm_mairie...	fait pas loucher quoi hein il il faut	prendre	le train en marche donc de toute façon on est on est
Deb_lm_mairie...	saison hein tout à fait ils acceptent de	prendre	en dehors donc c'est pour ça qu'on avait fait des
Deltaplane_tu...	mais enfin la plupart du temps tu peux	prendre	euh deux mille t- tu peux monter à deu- dans les Vosges
Deltaplane_tu...	rattraper tant pis bon tant pis je vais	prendre	mon petit vol j'arrive à Oulles je remonte à trois mille
Deltaplane_tu...	puisque je les vois plus ils ont dû	prendre	au moins trente kilomètres d'avance sur moi mais tant pis je
Educ_pot_08...	un film euh donc euh bosser dessus euh	prendre	vraiment un b- un bon paquet d'heures pour faire ça mais
Emploi_cha_0...	y rendre < vous avez il fallait >	prendre	le train hein ah ouais mais euh + euh vous non mais
Employeuducn...	quasiment presque tous les jours pour aller ou	prendre	une photo faire un article alors on passe son temps dans les
Employeuducn...	avoir moi je vais à la chasse pour	prendre	l'air d- avant tout parce que j'aime bien la nature
Employeuducn...	moi c'est de la détente c'est	prendre	l'air c'est je suis dans la nature je suis bien
Ergotherapie...	lui dis hein elle commence à savoir ouais	prendre	des choses < euh hum tout ça elle pousse < * *
Famille_fer_08...	la semaine ah bon hum ils ont dû	prendre	une semaine ah oui c'est bien il y a que trois
Famille_fer_08...	sale euh euh leur chaise ils essaient de	prendre	des des nâtes avec la * * ah mais c'est affreux

Figure 28 : Extrait des résultats de la recherche de « prendre » via TXM dans TCOF

Conformément à l'ASC, l'apprenant analyse chaque ligne de concordances pour comprendre le fonctionnement de l'occurrence recherchée. C'est ici que prend le sens de la métaphore du détective ou du Sherlock Holmes. L'objectif est d'observer le fonctionnement de « prendre » à l'oral. Selon les principes de l'ASC, l'apprenant observe chaque occurrence, les classe par catégorie, et induit le sens, le fonctionnement et les règles. Je ne préciserai pas ici le processus d'analyse que réalisent les apprenants, je le ferai précisément en [5.3.2.3](#).

Les outils numériques utilisés pour l'ASC permettent également aux apprenants d'observer un cotexte plus large. Par exemple, le logiciel TXM permet d'obtenir pour chaque occurrence son cotexte d'apparition, comme c'est le cas ci-dessous. La figure suivante montre l'extrait de transcription duquel est extrait la dix-huitième occurrence de la liste (pointée par une flèche) présentée précédemment (figure 28) :

L 2: 0:15:53 ouais

L 1: 0:15:53 tout le monde mange de l'agneau et et à Noël et Nouvel an t- 0:15:56 il y a les trois quarts des personnes achètent un gigot de chevreuil ou un gigot de de sanglier ou 0:16:02 pour passer les fêtes de fin d'année alors bon il faut 0:16:05 bon moi j'en prends j'en laisse euh ça ne me dérange pas outre mesure 0:16:09 mais euh bon voilà mais la chasse moi c'est plutôt de la détente euh 0:16:14 je préférerais emmener mon appareil photo bien sûr que 0:16:18 il m'arrive souvent de voir d- passer des animaux 0:16:20 je je je ne tire pas je les épaule même pas je les les je les regarde je les regarde passer 0:16:25 j'ai pas envie de tirer je tire pas et voilà c'est tout hein pour moi c'est de la détente c'est **prendre** l'air c'est 0:16:32 je suis dans la nature je suis bien assis sur mon tabouret tranquille 0:16:35 pendant que je suis là je ne fais rien dans mon jardin [rire du locuteur] ni ailleurs puis voilà quoi 0:16:39 et mais bon c'est tout hein je je < ne

Figure 29 : Extrait de la transcription qui comprend l'occurrence « prendre » en dix-huitième position dans la figure 28

Ainsi, si nécessaire, l'apprenant peut avoir accès à un extrait plus long que la simple ligne de concordances qui donne généralement sept ou huit mots à gauche et à droite de

l'occurrence. Les apprenants peuvent ainsi faire des allers-retours entre la liste des occurrences et leur apparition dans un cotexte plus large.

Cette façon d'aborder la langue permet à l'apprenant de s'approprier plus facilement les règles de fonctionnement puisque c'est lui qui les induit à partir de l'analyse des occurrences. Cette méthodologie permet d'associer le lexique et la grammaire, de comprendre la fonction de certains mots dans l'organisation syntaxique d'un énoncé. Comme l'explique Boulton (2018 : 72) : « Le travail rééquilibre des approches *top-down* (du sens à la forme) et *bottom-up* (de la forme au sens), en redonnant de l'importance à la langue elle-même, les formes et sens en contexte. ». Cette méthodologie facilite également la compréhension de blocs de mots ou de *chunks*. Pour illustrer ce travail, voici l'exemple d'une requête réalisée dans le corpus oral TCOF avec le concordancier JConc :

Num	Fichier	Contexte Gauche	Occurrence	Contexte Droit
227	Depart_ZOE_17	l'époque il fallait aussi	parce que	vous saviez pas qu'en
228	Depart_ZOE_17	les les les les ah	parce que	les filles elles pouvaient pas
229	Depart_ZOE_17	nan ils le faisaient pas	parce que	c'est mais je l'
230	Depart_ZOE_17	et et chez soi ah	parce que	chez toi tu continuais à
231	Depart_ZOE_17	soit en Italie ou ailleurs	parce que	je suis très curieuse d'
232	Depart_ZOE_17	dans lequel on se trouvait	parce que	ces écoles là étaient chauffées
233	Depart_ZOE_17	allaient pas à l'école	parce que	ils aidaient les familles aux
234	Depart_ZOE_17	était allée à l'école	parce que	c'était une famille où
235	Depart_ZOE_17	a eu beaucoup de chance	parce que	les conditions d'alors étaient
236	Depart_ZOE_17	quand même fait venir euh	parce que	ça les arrangeait nan parce
237	Depart_ZOE_17	quand même voilà et si	parce que	euh euh à l'époque
238	Erasmus_UZA_...	sais pas si c'est	parce que	il y avait que moi
239	Erasmus_UZA_...	m'a même plus motivée	parce que	ici j'ai l'impression
240	Erasmus_UZA_...	je pense que c'est	parce que	c'est Erasmus justement que
241	Erasmus_UZA_...	de turc c'était marrant	parce que	c'est pas euh comme
242	Erasmus_UZA_...	choisi cours euh d'anglais	parce que	là-bas on pratique vraiment
243	Erasmus_UZA_...	ca qui m'a étonné	parce que	du coup moi j'avais
244	Erasmus_UZA_...	vois ah ouais ben ouais	parce que	du coup on était obligé

Figure 30 : Extrait des résultats de la recherche de « parce que » via JConc dans TCOF

Les occurrences de *parce que* en interaction ne sont pas toujours faciles à comprendre dans la mesure où elles ne servent pas toujours à exprimer une cause (alors que c'est la fonction qui est généralement enseignée). Là encore, pour chaque occurrence, le concordancier permet d'obtenir un cotexte plus large, comme nous pouvons le voir dans la figure suivante. L'apprenant peut ainsi accéder à la progression de l'interaction.

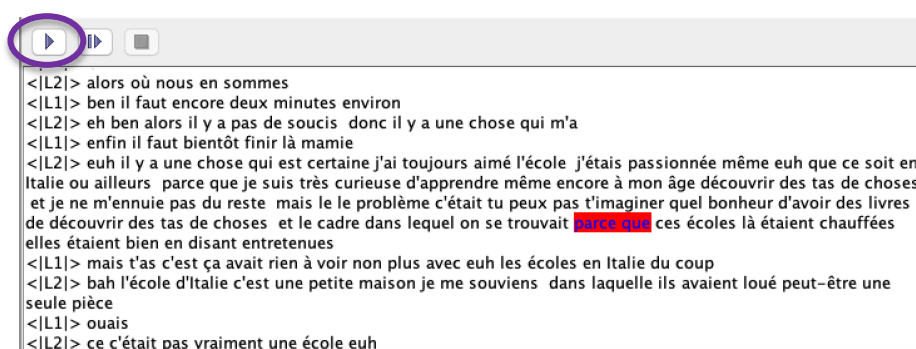


Figure 31 : Extrait de la transcription qui comprend l'occurrence « parce que » en sixième position dans la figure 30

Avec un extrait plus important, l'apprenant peut plus facilement saisir le sens de *parce que* et sa fonction. Avec cette occurrence, l'apprenant se rend compte que *parce que* sert à introduire un commentaire, une incise ou une précision. La conjonction ne suit pas une proposition principale contrairement aux règles de grammaire qui sont généralement fournies aux apprenants. Ce concordancier permet également d'écouter le segment correspondant, en cliquant sur la flèche de lecture (entourée en violet en haut à gauche de la figure). L'écoute de l'occurrence est importante, notamment parce que la prosodie influence le sens des énoncés en interaction.

Nous verrons, lors de mes différentes expérimentations, qu'en fonction du public d'apprenants avec lequel l'ASC est mis en place, l'accompagnement par un enseignant dans la démarche d'analyse est plus ou moins important. Néanmoins, l'objectif est d'atteindre les trois phases préconisées et présentées précédemment. L'apprenant, aidé ou non d'un enseignant ou d'un conseiller, observe les occurrences, analyse leurs apparitions en procédant à une classification puis tente de généraliser en dégagant des règles de fonctionnement. A propos de l'accompagnement des apprenants, je rejoins Johns (1986 : 159) : « It is important that teachers themselves should have experiences in using concordance output if they expect their students to make use of it. ». Si, comme plusieurs études le montrent, les apprenants réussissent bien à remplir leur rôle de détective de la langue, il est souvent nécessaire qu'un enseignant ou un accompagnateur montre le chemin de l'analyse, au moins lors de la première expérience. J'en profite pour insister ici sur la nécessaire formation des enseignants de langue. A Nancy, dans le Master FLE, j'ai formé la première promotion à l'exploitation de corpus à des fins didactiques lors de l'année universitaire 2018/2019. Cette formation a donné aux étudiants la possibilité d'utiliser

l'ASC, d'apprendre à analyser eux-mêmes les occurrences des concordanciers mais également à prendre en compte les éléments de la situation de communication pour aller vers une analyse sociolinguistique des données. Leńko-Szymańska (2017) mentionne également le besoin de former les enseignants pour qu'ils puissent exploiter des corpus à des fins didactiques. Son étude examine notamment « if and to what extent the trainees mastered the three sets of skills (i.e., technical, corpus linguistics, and pedagogical) necessary for independent use of corpora in language learning and instruction. » (*Ibid.* 219). L'objectif est de former les enseignants aux outils permettant d'exploiter et de sélectionner des corpus, aux analyses de données et aux activités didactiques à proposer aux apprenants, pour qu'ils acquièrent des compétences techniques, en sociolinguistique de corpus et pédagogiques.

5.2. De l'analyse de lignes de concordances à l'analyse sociolinguistique des données en interaction

La méthodologie impulsée par l'ASC consiste à faire analyser par des apprenants de langue le résultat de la recherche d'une chaîne de caractères (un ou plusieurs mots) *via* un concordancier⁶⁸, en présentant les résultats sous la forme de lignes de concordances ou dans un format appelé « KWIC » (*Keywords in context*). Ce procédé semble d'une grande efficacité pour l'apprentissage de l'écrit (pour rappel, presque la totalité des études portent sur la compréhension écrite de l'anglais et la production écrite en anglais, voir notamment Boulton 2012 ; Boulton, Cobb 2017). Le concordancier permet aux apprenants d'observer et d'analyser des dizaines d'occurrences afin de comprendre le fonctionnement de l'item ou des items recherchés. Ces recherches donnent accès à des contextes variés et authentiques. Certaines pratiques d'apprenants s'arrêtent à la consultation des listes de concordances et certaines pratiques d'enseignants s'arrêtent à la diffusion de ces listes (parfois en version imprimées sur papier). La démarche vise la description des faits de langue en contexte, mais quel est ce contexte ? Si les apprenants se restreignent aux lignes de concordances, ils analysent le cotexte du ou des mots recherchés. Cette approche de la langue représente déjà une grande avancée par rapport à une approche didactique explicite

⁶⁸ Parmi les concordanciers, *AntConc*, téléchargeable gratuitement, est certainement le plus utilisé. <http://www.laurenceanthony.net/software/antconc/>

qui consisterait à donner des règles de fonctionnement de la langue aux apprenants. Néanmoins, il est possible d'approfondir les analyses, de s'intéresser aux aspects extralinguistiques et de les corrélérer avec les phénomènes recherchés.

En effet, l'intérêt que présentent les corpus ne réside pas seulement dans le caractère authentique des données. Ces données ont été recueillies dans des situations de communication particulières, renseignées grâce aux métadonnées. L'ASC renouvelle la façon de traiter les faits langagiers dans le cadre de l'enseignement et de l'apprentissage d'une langue. Comme cela a déjà été mentionné dans des travaux et comme les apprenants le soulignent régulièrement, l'ASC est une nouvelle façon de faire de la grammaire. Presque tous les points grammaticaux peuvent être traités grâce aux corpus, et surtout ils peuvent être traités au regard des éléments de la situation de communication qui font varier leurs usages, leur fréquence ou leur récurrence. A propos de l'enseignement de l'anglais, Conrad (2000 : 558) explique :

Corpus linguistics offers the field of TESOL a new view of grammar. This view describes variation in the frequency and use of grammatical constructions and acknowledges the complexity involved in grammatical choices, which are not simply decisions about formal accuracy. Furthermore, corpus linguistics reveals that the variation in use of grammatical constructions is highly systematic: The ways that native speakers typically exploit their grammatical resources follow strong patterns, with systematic associations between grammatical choices and features of the social and linguistic context. Corpus linguistics leads to a description of grammar that is consistent with native speaker use, accounts for variation in a systematic way, and is useful for English language teachers and students. Incorporating this view of grammar into teaching will require a concerted effort on the part of researchers, materials writers, and teachers, but it is a logical, realistic, and beneficial outcome of corpus linguistics.

Dans le cadre de mon exploitation didactique des corpus oraux, je souhaite utiliser toutes les ressources qu'offrent les corpus, c'est-à-dire les données primaires et les données secondaires, qui décrivent les corpus. Ces dernières orientent les apprenants dans leur observation des premières. Les corpus présentent différents degrés de précision quant aux métadonnées. Comme je l'ai mentionné à plusieurs reprises, le genre de discours est un élément important qui influence largement les interactions. Ainsi, une des premières informations qui peut être requise dans l'exploitation de corpus à des fins d'enseignement et d'apprentissage des langues est le genre de discours auquel participent les locuteurs et au sein duquel apparaissent les chaînes de caractères recherchées par les apprenants *via* un concordancier.

L'accès aux métadonnées permet donc aux apprenants de faire de la sociolinguistique de corpus à des fins d'apprentissage de la langue (André 2018a). Au cours de l'année 2018, de récurrentes demandes de la part d'apprenants de FLE portaient sur la compréhension et la production de *du coup*. Ces demandes s'expliquent par l'utilisation récurrente de ce marqueur ces dernières années. Les enseignants de FLE peuvent se trouver mal aisés voire démunis pour répondre aux questions des apprenants concernant les règles de fonctionnement de ce marqueur, sa place dans l'interaction ou encore ses valeurs pragmatiques. Pour répondre à ces demandes, j'ai fait rechercher par les apprenants les occurrences du marqueur dans le corpus TCOF. Voici une capture d'écran de l'une des recherches :

Num	Fichier	Contexte Gauche	Occurrence	Contexte Droit
30	Coree_GHU_14	c'est pas cool et dans tout ça	du coup	on s'est fait taper sur les
31	Dos_COU_14	c'êt-ça rendait compliqué	du coup	le le l'opération pour le retirer
32	Dos_COU_14	dans l'oeil non donc voilà et	du coup	actuellement la ton ta boule
33	Dos_COU_14	santé un peu particuliers et	du coup	en quoi celui-ci est plus
34	Dos_COU_14	d'accord O. K. et sinon euh	du coup	euh depuis que euh qu'on fait
35	Dos_COU_14	euh qui sont pas faciles à	du coup	que ça que ça plus ça dure et
36	Dos_COU_14	qui a été euh quand même	du coup	parce que euh comme on
37	Forum_GOM_14	on leur vend une prestation et	du coup	elles tendent à répondre vu
38	Forum_GOM_14	une erreur qui n'était pas la	du coup	après aller négocier avec
39	Forum_GOM_14	toucher s-les comment les	du coup	voilà ça c'était vraiment très
40	Forum_GOM_14	tout le monde qui a	du coup	c'est vrai que euh de notre
41	Infirmier_AUD_...	un peu parti en hors sujet et	du coup	euh ça allait pas le faire mais
42	Infirmier_AUD_...	ça chamboulle tout dans la	du coup	je suis reparti puis je me suis
43	JEU_THI_14	sur mon sur mon ma partie	du coup	j'avais sauvegardé juste après
44	JEU_THI_14	me suis bien fait défoncé et	du coup	il va se passer exactement la
45	JEU_THI_14	exactement la même chose ou	du coup	c'est beaucoup d'aléatoire et
46	JEU_THI_14	être les mêmes en fait Oui oui	du coup	ce qui reste tout ce qui reste
47	JEU_THI_14	est une attaque euh d accord	du coup	euh la der-ben c'est marrant
48	JEU_THI_14	tout ça ils ont fait leur	du coup	je me suis dit autant aller les
49	JEU_THI_14	ça et ils se sont fait défoncés	du coup	je me suis dit arh j'avais pas
50	JEU_THI_14	qui sont sortis j'me suis fait	du coup	ben j'me suis dit que j'avais

Figure 32 : Extrait des résultats de la recherche de « du coup » via JConc dans TCOF

Comme je l'ai déjà expliqué, l'apprenant peut avoir un cotexte d'apparition plus large et accéder à la transcription du corpus pour saisir plus facilement le sens de ce qu'il observe. Ainsi, il est possible d'identifier le fonctionnement de *du coup* en tant que connecteur logique (qui exprime la conséquence) ou argumentatif, en tant qu'introducteur de question, d'explication ou de justification ou encore en tant qu'introducteur de tour de parole et permettant de gérer le déroulement des interactions. Ces activités interactionnelles et ces valeurs pragmatiques ont également été identifiées par le Groupe ICOR et Guinamard (2015) qui proposent d'exploiter les analyses de *du coup* en interaction à des fins didactiques. Les auteurs précisent à ce propos que : « Si l'on adopte une approche

communicative et actionnelle, « du coup » présente donc l'intérêt d'être un marqueur discursif de la langue française auxquels les apprenants seront confrontés dans leurs pratiques langagières sans pouvoir facilement l'interpréter » (*Ibid.* : 147). L'analyse du corpus TCOF montre que les occurrences de *du coup* sont nombreuses et fréquentes dans les interactions, même si comme dans le corpus CLAPI étudié par le groupe ICOR, les apparitions sont inégalement réparties entre les locuteurs et les genres de discours. Dans les 65 heures du corpus TCOF, il y a 1355 occurrences de *du coup*. Cela signifie qu'en moyenne, il y a une occurrence toutes les trois minutes⁶⁹. Évidemment, pour affiner ces résultats en fonction des éléments de la situation de communication, les métadonnées sont importantes.

Ainsi, en exploitant des corpus oraux, l'apprenant peut également avoir accès aux métadonnées et observer si l'item recherché est plutôt utilisé dans des conversations, par des locuteurs d'un certain âge, qui entretiennent une relation particulière, etc. Les trois captures d'écran suivantes montrent les métadonnées de TCOF qui sont susceptibles d'intéresser les apprenants. La première (figure 33) concerne des informations générales sur l'enregistrement et les deux suivantes (figures 34 et 35) des informations sur les locuteurs :

Nombre de locuteurs	2
Relation entre participants	Lien de parenté
Type de corpus	entre adultes
Modalités de recueil des données	Sollicitées par un chercheur
Canal de communication	Face à face
Cadre situationnel	Privé
Degré d'interactivité	Moyennement interactif
Genre de discours	Conversation
Typologie textuelle dominante	Narratif
Thème abordé / résumé	Un étudiant en école d'infirmier raconte le passage de ses concours à la fin de son année de terminale.

Figure 33 : Première partie des métadonnées du corpus *Infirmier_aud_14* dont sont extraites les occurrences 41 et 42 dans la liste de la figure précédente

⁶⁹ L'académie française explique que *du coup*, dans la plupart des emplois constaté en corpus, est un emploi fautif et précise sur son site : « On ne peut donc pas employer systématiquement *du coup*, ainsi qu'on l'entend souvent, en lieu et place de *donc*, *de ce fait*, ou *par conséquent*. » cf. <http://www.academie-francaise.fr/du-coup-au-sens-de-de-ce-fait>

Identification des locuteurs	
Étiquette	L1
Nombre de tours de parole	10
Age du locuteur	20
Sexe du locuteur	F
Niveau d'étude	Etudes Supérieures
Type de formation	L2 Sciences du langage
Profession(s) actuelle(s)	Animatrice
Rôle dans l'interaction	enquêteur
Degré de planification	Semi-planifié
Statut du français	Langue maternelle
Lieu de naissance	Nancy
Lieu de résidence (ville)	Bouxières aux dames

Figure 34 : Deuxième partie des métadonnées du corpus *Infirmier_aud_14*, celles qui concernent les locuteurs (L1)

Étiquette	L2
Nombre de tours de parole	10
Age du locuteur	20
Sexe du locuteur	M
Niveau d'étude	Etudes Supérieures
Type de formation	IFSI
Rôle dans l'interaction	interviewé
Degré de planification	Non planifié
Statut du français	Langue maternelle
Lieu de naissance	Reims
Lieu de résidence (ville)	Saulxures-les-Nancy

Figure 35 : Deuxième partie des métadonnées du corpus *Infirmier_aud_14*, celles qui concernent les locuteurs (L2)

En prenant connaissance des métadonnées du corpus, les apprenants peuvent compléter leurs analyses, faire des corrélations et finalement, faire de la sociolinguistique de corpus puisqu'ils analysent les occurrences à la lumière des éléments de la situation de communication.

Par exemple, il est possible de constater que l'usage de *du coup* est trans-générationnel, contrairement à l'usage de *genre*, présenté précédemment. Des locuteurs de tous âges utilisent *du coup*, mais tous ne l'utilisent pas et certains l'utilisent fréquemment. L'analyse sociolinguistique des occurrences montre également que dans certaines conversations, les occurrences de *du coup* sont très fréquentes, parfois réparties de façon égalitaire entre les locuteurs, parfois pas. La figure 32 montre le résultat dans trois

corpus différents (Dos_COU_14, Forum_GOM_14, Infirmier_AUD_14) et le début des résultats d'un quatrième corpus (JEU_THI_14). Il est facile de constater en observant ces lignes de concordances que tous les corpus ne sont pas égaux.

Ainsi, l'exemple de l'analyse de *du coup* montre que l'ASC peut se faire en ayant une approche sociolinguistique des interactions. Faire faire de la sociolinguistique de corpus à des apprenants leur permet d'observer la variation dans la mesure où « l'apprenant ne dispose pas de références sociales externes (ou plutôt il a à se les construire) » et où « il est très difficile de transmettre ce savoir subtil, privilège des natifs⁷⁰, la capacité à évaluer les situations » (Gadet 2001 : 66-67).

5.3. L'apprentissage sur corpus avec FLEURON

Ce point représente sans doute l'aboutissement et le cœur de mes travaux actuels et futurs. Par conséquent, il est long. Je m'excuse de ce déséquilibre dans cette synthèse mais son caractère original et novateur mérite, à mon avis, que je m'y attarde un peu plus longuement. Comme je l'ai précisé dans la présentation du corpus FLEURON, le dispositif d'apprentissage dans lequel il s'intègre comporte un concordancier texte-son-vidéo qui représente le cœur de ce dispositif. Je vais tout d'abord présenter ce concordancier avant de relater certaines des expérimentations d'ASC menées auprès d'apprenants de FLE.

5.3.1. Le concordancier de FLEURON

L'idée d'ajouter un concordancier au dispositif d'apprentissage FLEURON, est apparue au début des années 2010, dès les débuts de la concrétisation du projet (André 2016). Depuis 2010, le site Internet dédié à FLEURON a connu de nombreuses évolutions et plusieurs versions, avant sa version actuelle, mise à jour en 2019⁷¹. Les premiers sites Internet ont été entièrement développés par des étudiants en Licence d'informatique CISII (Concepteur Intégrateur de Systèmes Internet/Intranet pour l'Entreprise) de l'IUT

⁷⁰ J'apporterais une nuance à cette opposition étant donné que certains natifs en insécurité langagière rencontrent également des difficultés pour interagir de façon appropriée à la situation dans laquelle ils se trouvent.

⁷¹ Cette mise à jour comporte de nettes améliorations des fonctionnalités du site. Elles ont été réalisées par Ismaël Bada, contractuel à l'ATILF dans le cadre du CPER du 1^{er} septembre 2018 au 31 août 2019.

Charlemagne à Nancy encadrés par Yann Boniface⁷². Cette collaboration a pris la forme de projets tutorés et de stages pour ces étudiants en informatique au sein du projet FLEURON de 2010 à 2015. C'est au cours de l'un de ces stages, que deux étudiants ont développé le premier concordancier de FLEURON. Il a ensuite été repris en 2016 puis en 2019, il est unique puisqu'il propose d'observer les occurrences recherchées dans les ressources multimodales à partir de la liste des concordances (André, Ciekanski 2018).

Concrètement, le concordancier de FLEURON s'inspire des outils de la sociolinguistique de corpus mais il est adapté et simplifié pour une utilisation par des non spécialistes, c'est-à-dire des apprenants et des enseignants de FLE. Il permet de rechercher une chaîne de caractères dans un espace dédié accessible *via* l'interface du site (il est présent sur chacune des pages du site). Voici une image isolée :

Figure 36 : Concordancier de FLEURON – espace de recherche

Cette figure montre qu'il est possible de rechercher un mot exact, par exemple « même ». En tapant « même » dans le concordancier et en cochant « mot exact » tous les « même » et seulement les « même » seront affichés sous la forme de lignes de concordances, comme c'est le cas ci-dessous :

⁷² Maître de conférences en Informatique et membre du LORIA (UMR 7503).

Concordancier ☒ même ☐ mot exact Rechercher		
Concordancier (116)	Ressources (3)	CNRTL
A gauche du mot	Mot recherché	A droite du mot
L2 : qui font le	même	travail que moi en fait c'est c'est au niveau de c'est une politique des sites
A : en	même	temps E : hum hum
A : ah 16h bah très bien parfait il y a	même	5 minutes d'avance
A : j'accepte quand	même	
A : c'est quand	même	plus facile pour vous
A : le jour	même	
A : vous aurez la carte le jour	même	
A : pas pour l'instant bon bah si vous avez d'autres questions	même	plus tard
A : ah d'accord ou	même	par fax hein c'était E : ah oui A : vous pouvez envoyer le document par fax hein si c'est plus rapide euh en général
E : là on doit avoir le same la	même	chose et A : hum E : vous n'avez pas si je n'avais pas le X sécurité
A : c'est pour votre sécurité hein parce que s'il y a un souci	même	pour le médecin euh
A : ah non c'est c'est pas le	même	E : c'est pas le même A : non c'est hein pas pareil E : j'avais aussi euh
A : ah non c'est c'est pas le même E : c'est pas le	même	A : non c'est hein pas pareil E : j'avais aussi euh
E : ah ici pour moi c'est la	même	chose
A : donc il me faut quand	même	la la carte de séjour qui est qui est expirée
A : ah Epinal ça vous fait loin quand	même	pour venir E : oui
A : euh c'est une euh E : oui c'est la	même	chose
E : d'accord A : je fais la	même	chose pour euh toutes les 2
A : vous avez une étac ensemble dans la	même	résidence E : oui E : ouais

Figure 37 : Extrait de la liste des concordances de « même » dans FLEURON, en cochant « mot exact »

Si la case « mot exact » n'est pas cochée, il est possible de saisir « meme » – sans accent – et le concordancier affichera tous les « mêmes » ainsi que ses variantes, ici « mêmes » – au pluriel, comme c'est le cas ci-dessous :

A : donc nous avons quand	même	un sa- plus de mille personnes
A : mais bon je privilégie quand	même	les événements qui se passent ici à la au campus
E : avec les	mêmes	informations A : avec les mêmes informations donc on se partagera les informations
E : avec les mêmes informations A : avec les	mêmes	informations donc on se partagera les informations
E : donc qui voudraient faire du FLE ou	même	qui veulent apprendre une langue seules
E : hum j'ai choisi le cours par moi moi-	même	A : oui c'est bien

Figure 38 : Extrait du résultat de la recherche de « meme » dans FLEURON, sans cocher « mot exact »

Le concordancier trouve les variantes féminines et plurielles des mots recherchés ainsi que les variantes avec ou sans accent. Il est également possible de rechercher plusieurs mots successifs.

L'originalité et la force du concordancier de FLEURON réside dans le fait qu'il soit aligné texte-son-vidéo. Cela signifie qu'il est possible, en cliquant sur une occurrence, d'aller dans la ressource correspondante au moment même où cette occurrence est prononcée. Par exemple, si un apprenant souhaite observer dans un contexte plus large la troisième occurrence de la figure précédente, il lui suffit de cliquer dessus pour que la vidéo s'ouvre :

Retour Communication de la MDE – partie 1

Media Description Conseils

A: sinon je communique également sur les écrans télé
 A: on va peut-être avancer par là-bas ça sera plus simple
 A: donc nous avons des écrans de télé ici au rez-de-chaussée et à l'étage
 A: donc toujours pour communiquer sur les événements
 A: et donner également des informations aux étudiants
 E: d'accord
 A: alors ces écrans sont ici mais il y en a également à l'IUT Charlemagne
 A: il y en aura dans sur d'autres campus et on sera amené par la suite à communiquer sur tous les campus
 E: avec les **mêmes** informations
 A: avec les **mêmes** informations donc on se partagera les informations
 A: et on les diffusera sur tous les campus
 E: d'accord
 A: bon pour le moment c'est pas encore fait parce que l'université de Lorraine vient de se mettre en place
 A: et ça a besoin de beaucoup **
 E: ah c'est dans le cadre de l'université de

Masquer les sous-titres

E: avec les **mêmes** informations
 A: avec les mêmes informations donc on se partagera les informations

Concordancier
 mêmes

Figure 39 : Retour dans la ressource vidéo en cliquant sur la troisième occurrence dans la figure 38.

Ainsi, la ressource s'ouvre au moment où l'énoncé qui contient « même » est prononcé (ici à 2min24), avec le sous-titre correspondant et l'occurrence est en gras dans la transcription.

Cette fonctionnalité est particulièrement appréciable dans un travail avec un corpus de français parlé en interaction. Pour saisir le sens d'un mot, d'un bloc de mots ou d'un énoncé, il est parfois nécessaire de recourir à un contexte assez large et de visionner les quelques secondes ou les quelques minutes précédentes. Tout d'abord, le sens se construisant au fur et à mesure des interactions, la compréhension peut demander une contextualisation plus longue que quelques mots à droite et à gauche. Ensuite, les modes d'élaboration des discours oraux peuvent expliquer facilement cette nécessité de contexte. Les locuteurs peuvent évoquer des éléments puis les reconvoquer à un autre moment, en procédant à des rappels de thème qui peuvent prendre différentes formes (parfois un simple *au fait* ou *tu sais*, en fonction des relations entre les interactants). Les locuteurs peuvent également construire leur discours en piétinant, en répétant, en reformulant et en s'appuyant sur le discours de leurs interlocuteurs (et parfois de leurs gestes). Ces procédés sont parfois difficilement observables dans une ligne de concordances. Enfin, des aspects plus techniques rendent un accès plus large à l'interaction indispensable, notamment quand l'item recherché est en fin de tour et que le tour suivant n'apparaît pas dans la ligne de concordances, lorsque plusieurs locuteurs interagissent en même temps, lorsque les locuteurs produisent des énonciations conjointes, ou encore lorsque les apprenants

s'intéressent à certains aspects des interactions (la multimodalité, par exemple). Les apprenants utilisent souvent cette fonction, lorsqu'ils endossent leur rôle de Sherlock Holmes.

5.3.2. Les étudiants-apprenants-utilisateurs de FLEURON : des Sherlock Holmes

L'objectif de cette section est de rendre compte des expérimentations de l'exploitation didactique du corpus multimédia FLEURON avec des étudiants, apprenants de FLE à l'université. Dans la mesure où l'ASC peine à arriver en France (Boulton, Tyne 2014 ; Cobb 2014) et qu'il n'est que rarement mis en œuvre pour l'apprentissage de la langue parlée, je propose une analyse qualitative de son application avec FLEURON. Les expérimentations que j'ai menées concernent une centaine d'apprenants. Certaines ont été réalisées conjointement avec des stagiaires de deuxième année de Master, formés à l'ASC. D'autres ont fait partie des enseignements de Florence Poncet, enseignante de FLE et membre du projet FLEURON. Néanmoins, à ce stade de l'avancée des propositions didactiques en ASC, je ne produirai pas de données chiffrées. L'objectif est de montrer que les apprenants jouent le rôle d'apprentis-chercheurs ou de Sherlock Holmes, leur permettant d'entrer dans le triple processus d'observation et de catégorisation des données ainsi que de généralisation des règles induites.

Dès 2014, des expérimentations de FLEURON ont été mises en place, notamment au DÉFLE à Nancy mais également à Pékin, au département de français de l'université des langues et culture⁷³. Même si, jusqu'en 2016, le site Internet n'était pas stable, les premières expérimentations ont permis de mettre au jour un certain engouement pour le dispositif. Cet engouement avait deux origines : les ressources et le concordancier. D'une part, le corpus FLEURON présente des données authentiques très recherchées par les enseignants⁷⁴ et quasiment inexistantes, même sur Internet. D'autre part, l'outillage du site avec un concordancier intrigue les enseignants. Je dois bien avouer que ce n'est pas le concordancier qui les a le plus séduits dans le dispositif. Ce n'est donc pas par le concordancier qu'ils sont entrés dans le site. Il a fallu effectuer un travail de formation voire de publicité pour faire adhérer les enseignants au processus d'appropriation permis par le concordancier. En effet,

⁷³ Beijing Language and Culture University : <http://english.blcu.edu.cn/>.

⁷⁴ J'ai d'abord présenté le dispositif à des enseignants avant de l'expérimenter auprès d'apprenants.

l'utilisation d'un tel outil ne fait pas du tout partie des pratiques des enseignants. Les expérimentations menées par Kathia Varinot pendant son stage de Master et les formations que j'ai dispensées à Pékin, à la demande de l'Ambassade de France, ont été les premiers retours concrets sur le dispositif. Par exemple, les étudiants chinois ont souligné l'importance des titres et des résumés des ressources. Jusque-là, peu d'importance avait été accordée à ces éléments alors qu'ils permettent effectivement d'améliorer la compréhension orale, selon le modèle onomasiologique (Gremmo, Holec 1990). Ils ont souligné également que les vidéos authentiques du site étaient un véritable levier de motivation, phénomène relevé également par Peacock (1997). De plus, les premières observations de l'utilisation du concordancier ont pu être réalisées, notamment en observant les occurrences les plus recherchées par les apprenants.

A partir de 2016, une version plus stable du site va permettre des expérimentations nombreuses et à plus grande échelle. Ce sont certaines de ces expérimentations, menées en France, que je vais relater ici pour illustrer le processus analytique mis en place par les apprenants. Ce qui m'intéresse particulièrement, dans le cadre de ces études, c'est ce processus qui conduit les apprenants à faire des analyses sociolinguistiques pour induire les règles de fonctionnement des faits langagiers qu'ils recherchent dans le concordancier. L'objectif est d'observer si les apprenants se comportent effectivement comme des Sherlock Holmes et d'analyser comment ils y parviennent. En d'autres termes, l'objectif est d'examiner si l'ASC contribue au développement d'une conscience langagière et métalangagière qui mènera les apprenants vers une autonomie socio-interactionnelle.

5.3.2.1. Les consignes

Avant de présenter des expériences d'utilisation du concordancier de FLEURON, je vais proposer la transcription de certaines consignes données aux apprenants avant qu'ils se lancent seuls sur le site. Ces consignes sont d'autant plus importantes que les apprenants qui ont participé aux séances d'ASC ne l'ont jamais pratiqué auparavant, quel que soit leur pays d'origine et leur niveau en français. Chacune des consignes ci-dessous est produite après une présentation générale du dispositif FLEURON, et notamment de son corpus d'interactions.

Exemple 1 : Consigne donnée à un groupe d'étudiants, le site est projeté sur un écran via un vidéoprojecteur.

VA alors ce qu'on appelle un concordancier en fait c'est simple c'est un peu **comme un moteur de recherche** vous en utilisez tous les jours par exemple quand vous recherchez quelque chose dans **Google** ça va chercher dans plein de site et ben le concordancier que vous avez là si le site (montre le concordancier) c'est pareil mais ça va chercher dans toutes les vidéos du site par exemple **si vous taper merci** (tape merci dans le concordancier) et ben voilà **vous avez tous les merci qui sont prononcés par les locuteurs** par les personnes qui sont filmées dans dans toutes les vidéos vous avez la liste de tous les merci et ce qui est bien c'est que si vous voulez voir un peu la vidéo où il y a ce merci (montre une occurrence) vous pouvez cliquer sur merci mettons que vous vouliez voir celui-là hop **vous cliquez dessus et vous avez accès à la vidéo** et hop après vous pouvez revenir à la liste vous voyez + après qu'est-ce que vous faites de ça et ben pour comprendre comment on utilise merci où est-ce que ça apparaît dans un échange **vous allez observer tous les merci** et petit à petit au fur et à mesure de vos observations **vous allez comprendre comment merci fonctionne en français**

Le concordancier est comparé au moteur de recherche de Google puisque tous les apprenants l'utilisent ou utilisent un moteur de recherche équivalent (au moins les étudiants chinois avant leur arrivée en France). Rapidement au cours de l'explication, un exemple simple est donné afin de faciliter la compréhension et la prise en main du concordancier. Les consignes données à un groupe en salle de classe demandent parfois à être répétées ou reformulées pour garantir la compréhension de l'outil. J'ai pu remarquer que la manipulation technique des données opérée par le concordancier ne pose quasiment jamais de problème de compréhension. Les étudiants actuels maîtrisent parfaitement l'outil informatique et perçoivent facilement l'extraction réalisée par le concordancier. Il s'agit donc essentiellement, avec ce type d'apprenants, d'expliquer ce qu'ils peuvent ou doivent faire avec les résultats, comme le montre l'interaction suivante :

Exemple 2 : Interaction avec un étudiant lors d'une séance individuelle. L'apprenant a lancé la recherche de celle.

VA est-ce que vous avez envie de chercher quelque chose dans le concordancier
 E ben oui (VA regarde l'écran)
 VA donc celle en fait l'idée vous connaissez le principe hein l'idée là vous êtes allé chercher celle
 E oui
 VA et **vous avez tous les celle qui sont prononcé dans les vidéos de FLEURON**
 E ah oui
 VA il vous a expliqué Olivier
 E je sais comment ça fonctionne
 VA vous savez comment ça fonctionne donc maintenant **l'idée c'est d'essayer de comprendre pourquoi on a celle**
 E euh oui

- VA donc **il faut que vous regardiez toutes les lignes** (montre les lignes du doigt) et puis comprendre pourquoi vous avez celle
- E pour chercher le fonctionnement dans la phrase
- VA ouais et **si vous avez un doute** hein si vous comprenez pas là (montre une occurrence) c'est un juste voilà **vous pouvez cliquer dessus** (l'étudiant clique sur l'occurrence)

Dans ce cas, l'apprenant a bien compris le fonctionnement du concordancier. Il a eu besoin d'une petite explication supplémentaire pour se lancer dans l'analyse des occurrences. Parfois, les consignes sont condensées :

Exemple 3 : Interaction avec un étudiant lors d'une séance individuelle.

- VA **essayez de regarder ce qu'il y a avant ce qu'il y a après** vous pouvez aussi cliquer dessus (prend la main sur le site) et vous pouvez afficher la transcription et il va apparaître en gras quelque part là on l'a vu voilà essayez de regarder un petit peu avant donc à quoi ça sert à votre avis

Dans cet exemple, la consigne est une sorte de résumé des consignes précédentes mais produite en situation, en réalisant les actions mentionnées directement sur le site. Une fois que les consignes sont bien comprises, les apprenants peuvent effectuer leurs recherches.

L'analyse des différentes consignes produites avec différents apprenants montre une progression utile dans l'apport des informations. Un schéma récurrent se dégage et semble faciliter la prise en main du concordancier et son exploitation. Dans les consignes les plus efficaces, les explications sont fournies dans l'ordre ci-dessous et sont illustrées par les actions mentionnées réalisées sur le site :

1. Le concordancier fonctionne comme un moteur de recherche.
2. Il va extraire les items recherchés dans toutes les transcriptions des ressources.
3. Il affiche les résultats sous la forme d'une liste.
4. Au centre de la liste, se trouve l'item recherché.
5. De chaque côté, on voit son cotexte, c'est-à-dire les mots avant et les mots après, ou encore les mots à gauche et les mots à droite.
6. L'objectif est d'observer et d'analyser les occurrences, par exemple en regardant ce qu'il y a à droite et ce qu'il y a à gauche ou en regardant l'emplacement de l'occurrence dans le tour de parole.
7. Si besoin, on peut cliquer sur l'occurrence pour avoir un contexte plus important.

8. Dans ce cas, on peut écouter l'extrait de l'interaction dans lequel se trouve l'item recherché. On peut aussi accéder aux informations sur l'interaction (qui sont les locuteurs, qu'est-ce qu'ils font, où est-ce que ça se passe, etc.).
9. Puis on peut ensuite revenir à la liste des occurrences.
10. Et reprendre les analyses.

Toutefois, les consignes sont naturellement ajustées aux apprenants et aux conditions d'expérimentations. Elles sont également parfois négociées par les apprenants pour garantir leur intelligibilité et leur compréhension.

5.3.2.3. Le processus d'analyse : l'apprenant devient un détective

L'objectif est de présenter le cheminement et le processus d'analyse mis en œuvre par les apprenants afin de devenir des détectives au sens de Johns (1997), au cours de différentes expérimentations. Je cherche à définir et à asseoir une méthodologie afin de pouvoir en diffuser les clés, les atouts, les inconvénients, les échecs et les réussites. Beaucoup d'études sur l'ASC reposent sur des raisonnements théoriques et spéculatifs (Chambers 2019). Je cherche à proposer des pratiques didactiques, de tester leurs réalisations et leurs pertinences pour des apprenants de FLE. L'objectif est de montrer que l'ASC est « accessible » (Boulton 2018 : 79) et que des apprenants non spécialistes de sociolinguistique de corpus réussissent à en utiliser ses techniques et sa méthodologie.

L'objectif n'est pas d'évaluer l'ASC ou la progression didactique des apprenants et l'acquisition de nouvelles compétences socio-interactionnelles. Boulton et Cobb (2017) se sont lancé dans un énorme travail de méta-analyse de 64 études présentant des données quantitatives qui permettent de chiffrer les apports de l'ASC comme aide à l'apprentissage. Cette méta-analyse montre notamment que quasiment toutes les études portent sur l'apprentissage de l'anglais écrit. Je ne cherche pas ici à tester les effets à long terme sur l'interlangue des apprenants, même si parfois je peux en constater à court terme, je le mentionnerai. Cet objectif fera l'objet d'autres études, longitudinales avec des groupes témoins, afin de déterminer si l'ASC est efficace, plus efficace que d'autres approches, moins efficace ou inefficace. Les preuves de l'efficacité de l'ASC passent avant tout par une démonstration de la pertinence de sa méthodologie, au-delà d'une « idéologie corpus » ou

de la « corpus mission » (Römer 2006). Une fois que cette méthodologie sera précisément définie et aura été mise à l'épreuve, alors il sera possible d'éviter, comme le souligne Boulton (2008a : 44) à propos des méthodologies telles que la suggestopédie ou le *Silent Way*, que « le DDL risque de connaître le même sort que ces autres approches, de disparaître et ne représenter qu'une curiosité dans l'histoire de la didactique ».

L'ordre dans lequel sont présentées les analyses a été choisi parce qu'il correspond à une amélioration progressive de ma démarche et de ma recherche. Les différentes expérimentations ont permis de faire des ajustements en fonction de leur échec ou de leur succès. Les participants aux expérimentations ont également été associés à la démarche d'amélioration du protocole didactique, notamment en évaluant les activités que je leur proposais.

Je vais présenter le déroulement des recherches de *même*, *voulais*, *ainsi*, *mieux/meilleur*, *sans*, *tout/tous/toute*, *ce/cet*, *hein* ainsi que des recherches par thèmes et des analyses inattendues. L'objectif de ces présentations successives détaillées est d'apprécier la méthodologie, c'est-à-dire de montrer de quelles façons les apprenants réussissent à s'approprier les pratiques de l'ASC. Ces recherches ont parfois été réalisées par plusieurs apprenants. J'explique leurs démarches, parfois individuelles, parfois collectives, parfois de façon isolée (celle d'un seul apprenant) et parfois de façon synthétisée (celles de plusieurs apprenants). Cette présentation est sélective, cela ne signifie pas que les expérimentations se limitent à la recherche de ces occurrences mais le but est avant tout de présenter les pratiques effectives des apprenants et des enseignants.

◆ *même*

Lors de l'une des premières expérimentations de l'utilisation du concordancier de FLEURON au Département de français langue étrangère (DéFLE) de Nancy, un apprenant a fait part de son incompréhension du mot *même*. L'opportunité d'observer l'utilité du concordancier a été saisie. En accord avec les principes de l'ASC et afin de fournir un input compréhensible (Chambers 2010), j'ai proposé à cet étudiant de rechercher *même* dans le concordancier de FLEURON. Voici un extrait de la liste des 125 occurrences :

A gauche du mot	Mot recherché	A droite du mot
L2 : qui font le	même	travail que moi en fait c'est c'est au niveau de c'est une politique des sites
A : en	même	temps E : hum hum
A : ah 16h bah très bien parfait il y a	même	5 minutes d'avance
A : j'accepte quand	même	
A : c'est quand	même	plus facile pour vous
A : le jour	même	
A : vous aurez la carte le jour	même	
A : pas pour l'instant bon bah si vous avez d'autres questions	même	plus tard
A : ah d'accord ou	même	par fax hein c'était E : ah oui A : vous pouvez envoyer le document par fax hein si c'est plus rapide euh en général
E : là on doit avoir le same la	même	chose et A : hum E : vous n'avez pas si je n'avais pas le X sécurité
A : c'est pour votre sécurité hein parce que s'il y a un souci	même	pour le médecin euh
A : ah non c'est c'est pas le	même	E : c'est pas le même A : non c'est hein pas pareil E : j'avais aussi euh
A : ah non c'est c'est pas le même E : c'est pas le	même	A : non c'est hein pas pareil E : j'avais aussi euh
E : ah ici pour moi c'est la	même	chose
A : donc il me faut quand	même	la la carte de séjour qui est qui est expirée
A : ah Epinal ça vous fait loin quand	même	pour venir E : oui
A : euh c'est une euh E : oui c'est la	même	chose
E : d'accord A : je fais la	même	chose pour euh toutes les 2
A : vous avez... vous êtes ensemble dans la	même	résidence E : oui E : ouais
A : et puis tu auras quand	même	ta carte d'accord

Figure 40 : Extrait des résultats de la recherche de « même » dans le concordancier de FLEURON

Avec cette liste d'occurrences, l'étudiant est invité à procéder à une analyse distributionnelle en observant les cotextes gauche et droit d'apparition du mot, les différentes collocations possibles et, par induction, les différents sens de ce mot. Ce procédé correspond à ce que Johns (1997 : 102) formulait de la façon suivante : « one item, many contexts ». L'apprenant se livre à l'analyse de ces contextes. Il met en œuvre des activités métalinguistiques et métacognitives (Gombert 1996) en produisant différents raisonnements à partir de ses observations. Voici ce qu'il a écrit sur une feuille de papier au fur et à mesure de son analyse :

- même travail = comparaison
- même temps = simultanément
- même 5 minutes d'avance = déjà
- même par fax = aussi
- toi-même = tout seul
- quand même une liste de documents à fournir = je dois fournir une liste de documents
- même si = bien que
- quand même = très fréquent

Figure 41 : Notes d'un apprenant lors de l'analyse des occurrences de même dans le corpus FLEURON

L'apprenant a utilisé différentes stratégies pour déterminer des catégories d'emploi. Il a trouvé des synonymes (simultanément), identifié des activités interactionnelles (comparaison) et des reformulations ou des paraphrases. Il a identifié une collocation fréquente (quand même), sans en déterminer le sens (en tout cas, il ne l'a pas écrit sur sa feuille). L'apprenant a également verbalisé les règles de fonctionnement de *même*. Après avoir travaillé seul pour réaliser ses analyses, j'ai demandé à cet apprenant de m'expliquer de quelle façon il avait procédé. Il a également verbalisé une évaluation de l'approche proposée en expliquant :

je pensais que même était compliqué à comprendre mais quand on fait une analyse comme ça en fait c'est facile à comprendre en prenant un par un tous les mêmes je comprends.

En interrogeant l'apprenant sur les analyses distributionnelles qu'il a réalisées, j'ai pu comprendre quand et pourquoi il a écouté l'occurrence en contexte et quelles sont les occurrences qu'il a laissées de côté, qu'il n'a pas analysées (ce qui est permis par le nombre d'occurrences dans le corpus). Il n'est pas nécessaire d'analyser toutes les occurrences pour comprendre les sens et les fonctionnements les plus fréquents. Cette phase d'observation des pratiques interactionnelles ainsi que de leurs conditions d'apparition et d'utilisation semble favoriser l'appropriation de la langue dans la mesure où les apprenants dégagent eux-mêmes les règles de fonctionnement et les différents sens d'un mot, d'un bloc de mots ou d'une expression en interaction. Cobb (1999) a notamment montré que l'ASC était particulièrement efficace pour l'apprentissage du vocabulaire et Ellis (2001) a montré le rôle de l'appréhension des *chunks* (tel que *quand même*) dans le processus d'acquisition d'une langue. Cette expérience qui est apparue lors des premières expérimentations FLEURON m'a poussée à poursuivre dans cette voie et à déployer de très nombreuses expériences du même type.

◆ *voulais*

Parmi ces expériences, je vais présenter le travail effectué lors de la recherche de *voulais* dans le concordancier de FLEURON. Cette recherche est suggérée par Florence Poncet, l'enseignante avec qui je partage une séance de cours de FLE. Elle vient de terminer un cours sur le passé avec des observations en interaction de quelques exemples et, dans ce cadre, propose à la classe de chercher *voulais*. Avant la recherche, elle demande aux

apprenants de faire des hypothèses sur le fonctionnement et le sens de *voulais*. Trois réponses font consensus. Elles sont reproduites en substance si-dessous :

- *on peut le remplacer par 'souhaitais'*
- *sert à raconter une histoire, des actions terminées*
- *c'est de l'imparfait donc ça sert à exprimer du passé*

Parmi les multiples possibilités qu'offre l'ASC, ce processus consistant à faire faire des hypothèses et à proposer d'aller les vérifier en corpus est assez fréquent. Ensuite, les apprenants sont donc invités à se lancer dans l'analyse des occurrences de *voulais* :

A gauche du mot	Mot recherché	A droite du mot
A : euh je	voulais	juste vous demander vous suivez les cours de français intensif
E : je	voulais	payer en espèces
E: je voudrais découvrir un petit peu la France A: oui	voulais	savoir ce que vous pouvez me proposer pour euh
E: et je		
E : je	voulais	être sûre
E: c'est tout ce que je	voulais	savoir A: ouais E: je vous remercie
E1 : c'est pour ça je	voulais	savoir A : ils vont être bons hein on peut pas savoir enfin
A : c'est vrai que euh je ne	voulais	pas vous donner trop de documents E : oui A : au début
A : alors je	voulais	vous demander
A : alors ce que je	voulais	savoir est-ce que vous avez eu l'occasion de regarder un petit peu les
A1: voilà donc je	voulais	juste passer le relais à Sylvie
A: voilà ce que je	voulais	dire pour heu pour vous décrire un petit peu le le CROUS
A: non mais c'est pas E: d'ailleurs c'est ça que je	voulais	vous demander faut que je ramène quoi pour l'attestation de

Figure 42 : Liste des résultats de la recherche de « *voulais* » dans le concordancier de FLEURON

Le corpus FLEURON compte seulement douze occurrences de *voulais* (voir la figure ci-dessus). Les apprenants ont ainsi pu analyser toutes les occurrences. Voici pas à pas la démarche qu'un étudiant a effectué :

- Il a d'abord observé toutes les occurrences les unes après les autres, dans leur ordre d'apparition à l'écran.
- Il a commencé par faire une catégorie avec les occurrences qui étaient associées à une demande (en commençant par mettre dans cette catégorie les occurrences se trouvant dans des énoncés où le mot *demande* était présent). Les occurrences 1, 2, 3, 5, 6, 8, 9 et 12 ont été classées dans cette catégorie.
- Il est allé dans la vidéo pour comprendre l'occurrence 4. Le cotexte de la ligne de concordance est le suivant : *euh je voudrais savoir euh est-ce que euh j'ai vraiment validé mes deux années euh je voulais être sûre*. Avec la co-occurrence de *je voudrais*

savoir, l'étudiant a ajouté l'occurrence 4 dans la première catégorie (avec une demande).

- En allant visionner les 3 occurrences restantes, il s'est rendu compte que les 10 et 11 étaient prononcées lors de conférences d'accueil des étudiants étrangers. Pour ces deux occurrences, le locuteur utilise *voulais* alors qu'il est sur le point de faire quelque chose (passer le relai, dire quelque chose).
- L'occurrence 7 est la seule occurrence de *voulais* qui a le sens que l'apprenant connaît, un synonyme de *souhaitais*, et qui exprime un événement passé (la locutrice fait référence à un rendez-vous précédent).

Rapidement, les apprenants ont compris que le passé du verbe *vouloir* pouvait exprimer autre chose que des faits passés. Il permet d'exprimer une action future, comme je l'ai déjà expliqué lors de l'analyse de l'interaction de la ressource « Paiement en espèces » (en [4.3.1.](#)) dont l'occurrence correspond à la seconde occurrence de la liste ci-dessus. Les interventions collectives des apprenants sont reproduites en substance ci-dessous :

- | |
|---|
| <ul style="list-style-type: none"> - ça sert à faire une formule de politesse - ça sert dans des situations administratives - avec les locuteurs qui ne se connaissent pas - ça sert à préparer une question comme dans « je voulais vous demander » - c'est avant une question pour être plus poli - ça sert à dire qu'on est sur le point de faire quelque chose qu'on se prépare à faire quelque chose |
|---|

C'est en analysant chacune des occurrences de *voulais* que les apprenants ont réussi à comprendre ses usages et son fonctionnement. Dans le cadre de cette étude, comme je l'ai déjà précisé, je n'examine pas les effets à long terme de l'apprentissage sur corpus. Néanmoins, à l'issue de cet exercice, certains apprenants ont eux-mêmes créé des activités de systématisation (Carette 2001) et surtout de mises en pratique en utilisant, pendant la séance de formation, des énoncés appropriés comprenant *voulais*. De plus, certains des apprenants ont souligné qu'à partir de cette séance, non seulement ils comprenaient *voulais* mais qu'ils seraient désormais également capables de l'utiliser.

Lors d'une autre séance de travail concernant également *voulais*, j'ai invité les apprenants à se rendre sur le site CLAPI-FLE et à consulter la collection intitulée « Imparfait de politesse »⁷⁵. Les quatre premiers exemples illustrent le fonctionnement de *voulais*. Les deux premiers présentent la forme « je voulais + syntagme nominal » et les deux suivants « je voulais + proposition infinitive ». Ces formes sont listées dans la partie explicative de la collection. De plus le quatrième exemple comporte « je voulais + demander » :

Exemple 2.2 - Corpus Invitation téléphonique - Proposition de déjeuner

Dans le cadre d'une conversation amicale, l'amie semble gênée de demander ce qu'elle considère manifestement être une faveur

Audio (wav)

Transcription (docx)

1 Anaëlle allô Florence c'est Anaëlle (1.0)

2 Florence oui

3 Anaëlle ça va; (0.6)

4 Florence oui et toi;

5 Anaëlle ça va dis *je voulais te demander*; demain vu que euh on a qu'une heure pour

6 manger euh ça te dit que on aille manger dehors; (1.3)

7 Florence euh hm pf je sais pas en fait parce que le truc c'est que j'ai encore PAS

8 de monnaie (0.5)

0:00 / 0:22

Figure 43 : Exemple pour illustrer l'imparfait de politesse dans CLAPI-FLE

Cet exemple permet aux apprenants de constater que cette formule de politesse s'utilise également dans un autre contexte (ici une conversation téléphonique entre deux amies) que ceux qu'ils ont observés avec le corpus FLEURON. La complémentarité des ressources exploitant des corpus d'interactions est un outil précieux pour l'enseignement et l'apprentissage du FLE, j'y reviendrai.

◆ *ainsi*

Une apprenante lettone s'est lancée dans la recherche de *ainsi* dans le concordancier de FLEURON. Comme beaucoup d'utilisateurs, elle a fait des hypothèses quant à son sens et son utilisation avant de passer à la phase d'observation et d'analyse des données. Elle a utilisé l'anglais comme langue d'appui (Bailly *et al.* 2009) ainsi que la comparaison avec un potentiel synonyme et a verbalisé : *j'ai l'impression que ça fonctionne comme so donc peut-être comme donc en français*. La recherche dans FLEURON n'a extrait que six occurrences de *ainsi* :

⁷⁵ http://clapi.icar.cnrs.fr/FLE/collection_grammaire_imparfait_politesse.php

A gauche du mot	Mot recherché	A droite du mot
A : je vous donnerez le papier	ainsi	vous aurez votre carte E : hum hum
A :	ainsi	vous aurez l'accès à internet euh demain
A :	ainsi	vous pourrez aller au cours de français ça commence à 14h
A : euh et puis	ainsi	vous pourrez assister au cours euh tous les après-midi
A:	ainsi	que le complexe sportive heu sportif Lefebvre hein
A1: de vous renseigner auprès de la MGEL	ainsi	que de la LMDE

Figure 44 : Liste des occurrences de « ainsi » dans le corpus FLEURON

Malgré le peu d'occurrences, l'apprenante a pu constater que la comparaison avec *donc* pouvait fonctionner dans certains cas. Elle a également conclu que *ainsi* se rapprochait de *alors* qu'elle connaissait. Ce procédé de comparaison est mis en évidence par Johns (1991) qui propose de chercher des différences entre des mots (tels que *convince/persuade*) grâce aux lignes de concordances. Elle a remarqué également que deux occurrences de *ainsi* étaient suivies de *que*. En analysant toutes les occurrences, elle a saisi les sens présents dans le corpus.

Après avoir réalisé l'analyse de chacune des occurrences, je l'ai invitée à rechercher *ainsi* dans un corpus écrit. Elle a fait la recherche dans Frantext⁷⁶ (dans la partie sans abonnement) et remarque les très nombreuses occurrences. Je l'ai ensuite orientée vers le corpus du journal local *L'Est Républicain* qui compte 8894 numéros. Ce corpus est disponible et interrogeable sur le site ScienQuest⁷⁷. Voici le début de la liste des résultats :

Mot 1				
Forme ainsi				
Lemme				
Catégorie				
Chercher		17502 occurrences		
#	Contexte gauche	Occurrences	Contexte droit	Texte
1	Vingt-cinq exposants ont	ainsi	fait hier la démonstration devant des centaines de visiteurs qu'	# 1 - Bar le Duc - 1999 - Chapeau (accroche)
2	La participation de la commune est	ainsi	fixée : 5 F par habitant , 90 F par	# 1 - Bar le Duc - 1999 - Article
3	D'_autres échéances sont proches ,	ainsi	le triathlon USEP organisé cette semaine à Ligny , puis	# 1 - Bar le Duc - 1999 - Article
4	ce serait dommage de rester hors_du coup » , expliquait	ainsi	Dominique Cordel . L'artisan du boulevard de la Rochelle	# 1 - Bar le Duc - 1999 - Article
5	« Une très bonne idée » , commente	ainsi	Nicole . « Je ne suis pas une très grande	# 1 - Bar le Duc - 1999 - Article
6	Ce sont	ainsi	22 camions de ramassage qui n'ont pas cessé de	# 1 - Bar le Duc - 1999 - Article
7	Vous pourrez	ainsi	y trouver des meubles , bien_sûr , mais aussi des	# 2 - Belfort - 1999 - Article
8	amicale des commerçants du village était fortement représentée et venait	ainsi	encouragée ce nouveau commerçant à qui nous souhaitons également pleine	# 2 - Belfort - 1999 - Article
9	Laurence et Alain , les deux coachs vercellois voient	ainsi	leur travail , leurs conseils récompensés et les couleurs vercelloises	# 3 - Besançon - 1999 - Article
10	Il a pu	ainsi	s'ouvrir à l'ensemble du répertoire musical français et	# 3 - Besançon - 1999 - Article
11	Il faut se féliciter de la célérité	ainsi	manifestée . En souhaitant que tout ira aussi vite par_la_suite	# 3 - Besançon - 1999 - Article
12	Avec la Concorde de Saint-Ferjeux Le prêtre a	ainsi	profité de l'occasion pour rappeler « que la fête	# 3 - Besançon - 1999 - Article
13	Dans son stand , Jacques Duchesne a	ainsi	profité de la forte affluence en ce jour de clôture	# 3 - Besançon - 1999 - Article
14	Avant_de lancer cette affaire et devenir	ainsi	travailleur indépendant , Corinne a suivi un stage de gestion	# 3 - Besançon - 1999 - Article

Figure 45 : Extrait de la liste des résultats de la recherche de « ainsi » dans le corpus Est Républicain sur ScienQuest

⁷⁶ Disponible à l'adresse suivante : <https://www.frantext.fr/>.

⁷⁷ Disponible à l'adresse suivante : <https://corpora.aiakide.net/>.

Même si les corpus – FLEURON et Est Républicain – ne sont pas comparables en taille, il est possible d’observer une différence de fréquence d’utilisation et de place dans les énoncés. L’utilisation des corpus permet donc également aux apprenants d’observer et de comparer la fréquence d’apparition de certaines formes. Cette fréquence permet, par conséquent, aux apprenants de s’interroger sur la pertinence d’apprendre certaines formes ou sur les priorités d’apprentissage à l’oral.

◆ *mieux vs meilleur*

Lors d’une séance de travail en individuel, une étudiante chinoise s’est donnée pour tâche de comprendre la différence entre *mieux* et *meilleur*. Elle a observé certaines des 17 occurrences de *mieux* et les 3 occurrences de *meilleur*. En passant dans les rangs, l’enseignant (OB) demande à l’étudiante (E1) si elle a compris, voici un extrait de leur échange :

- OB tu as réussi à comprendre un petit peu tu as réussi à comprendre
 E1 oui pour les les les **les différentes situations pour utiliser mieux et pas meilleur**
 OB ouais
 (...) (l’étudiante refait la manipulation, elle cherche mieux puis meilleur, l’enseignant l’interroge sur un meilleur au féminin, l’étudiante visualise les résultats et certains contextes)
 E1 ok
 OB est-ce que tu comprends quelque chose avec tout ça
 E1 euh **je pense meilleur c’est pour décrire quelque chose peut-être mieux c’est plus que plus une comparaison entre deux choses**
 OB d’accord et comment tu vois ça
 E1 par exemple (montre une occurrence et va dans la ressource) ça c’est un rendez-vous ils organisent un rendez-vous et il dit **l’option pour le matin c’est mieux que autre option**
 OB oui
 E1 meilleur c’est juste pour parce que l’autre vidéo **il dit meilleure façon il décrit le type de façon peut-être**
 (...)
 OB tu as encore quelque chose à dire sur mieux et meilleur
 E1 **je pense que j’ai compris mieux avec les exemples**

Cet exemple montre bien le processus analytique dans lequel s’est engagée cette étudiante. Elle observe les résultats du concordancier, analyse et catégorise les occurrences qu’elle réussit à comprendre, va observer le contexte plus large lorsque c’est nécessaire puis revient à la liste des occurrences. Elle passe également d’une liste à l’autre, elle alterne ses analyses de *mieux* et de *meilleur*. Le dernier tour de l’extrait présenté ci-dessus est

intéressant, même s'il est ambigu. La première interprétation possible est le fait que le mot *mieux* soit compris par l'étudiante, grâce aux exemples qu'elle a observés. La seconde interprétation, ma préférée, est une utilisation du mot *mieux* pour montrer à l'enseignant qu'elle a effectivement compris son usage. Cette interprétation semble être la bonne, même si l'étudiante ne place pas *mieux* au bon endroit dans son énoncé. L'interaction avec l'enseignant s'achève sur ce tour. Étant donné le travail réalisé par l'étudiante, on peut faire l'hypothèse qu'elle a également compris le fonctionnement de *meilleur* et qu'elle aurait mentionné cette compréhension dans un énoncé qui réalise une sorte de bilan ou de conclusion de son travail. En outre, il n'est pas rare que des apprenants mettent en pratique, immédiatement à la fin de leur travail sur corpus, les items qu'ils ont recherchés pour montrer aux enseignants ou aux autres apprenants qu'ils ont effectivement compris leur fonctionnement. Cette pratique montre que les apprenants ont acquis une certaine autonomie langagière, notamment telle qu'elle est définie par Germain et Netten (2004, np), c'est-à-dire comme « la capacité de l'apprenant de prendre des initiatives langagières et d'utiliser avec spontanéité des énoncés nouveaux lors d'une situation authentique de communication en [langue-cible] ». L'autonomie langagière ou socio-interactionnelle se développe en soutenant une attention à la langue.

◆ *sans*

Lors d'une séance en binôme, deux étudiants ont souhaité analyser *sans* (il faisait partie de la liste de suggestions d'items à rechercher dans le concordancier). L'analyse de ce binôme est intéressante parce qu'elle met en évidence deux modes de rétention d'information différents. Après l'observation des données, les étudiants se mettent d'accord sur les deux premiers sens identifiés de *sans*. Néanmoins, la façon de mémoriser ces sens est différente selon les deux étudiants. Le premier choisit de retenir *sans* suivi d'un nom avec le syntagme « sans visa », qu'il note sur son cahier et le seconde préfère noter « quelque chose que l'on n'a pas ». Ensuite, ils ont analysé *sans* suivi d'un verbe. Le premier étudiant note « sans perturber » et le second note « quelque chose que l'on ne fait pas ». Interrogés sur ces notes, chacun confirme qu'il retiendra mieux ce qu'il a noté plutôt que ce que son voisin a noté. Ensuite, ils se sont intéressés au bloc *sans doute* et ils ont eu recours à tous les contextes pour parvenir à saisir son sens. Le travail en binôme permet de collaborer aux analyses et de partager des découvertes et des compréhensions (Aston 1997).

◆ *tous/tout/toute*

Une étudiante lettone a souhaité comprendre les différences entre *tous*, *tout* et *toute*. Elle s'est donc lancée dans la recherche de ces occurrences dans le concordancier de FLEURON. Cette expérience est intéressante parce que l'étudiante est très à l'aise avec la recherche bien qu'elle n'en ait jamais réalisée auparavant. Progressivement, elle a réussi à obtenir les informations qu'elle a jugées nécessaires pour progresser dans l'ASC. Sa compréhension des différences entre les trois items s'est faite avec de nombreux allers-retours entre les trois listes de concordances. Je vais relater brièvement son cheminement.

- Tout d'abord à la vue des 356 occurrences de *tout*, elle a déclaré que c'était trop compliqué.
- Ensuite, elle est passée à la recherche de *tous* (135 occurrences). En observant la liste moins longue des occurrences de *tous*, elle a constaté que *tous* était suivi de *les* ou de « quelque chose au pluriel ».
- Elle a observé de nouveau la liste de *tout* et a constaté que *tout* se trouvait devant un *le*, un verbe, une préposition (*au*, *à*), *de* et un nom masculin. Elle a noté cette liste.
- Puis elle a observé de nouveau la liste des *tous*. Elle a noté que *tous* n'était pas suivi de verbe mais que cette observation était à vérifier dans la mesure où elle effectuait des observations partielles des listes de concordances. Elle a également constaté que *tous* était suivi de prépositions, mais aussi de *en vélo* et de *dans le même cas*.
- Elle s'est ensuite intéressée à *toute*. Elle a d'abord constaté que *toute* était souvent suivie de *façon* avant de voir que *toute* était également souvent précédé de *de* pour former l'expression *de toute façon*. L'étudiante ne comprend pas cette expression. Elle essaie de la traduire en anglais avec *all maner* mais se rend compte que ça ne fonctionne pas. Pendant cette phase d'analyse, un étudiant syrien écoutait notre conversation. Il s'est approché et a soufflé *anyway* à l'étudiante lettone.

A partir de cette compréhension de *de toute façon* grâce à la traduction en anglais, l'étudiante a poursuivi ses investigations, en faisant de la sociolinguistique de corpus sans le savoir et de façon très concentrée. A la fin de la séance, elle a avoué être épuisée mais très satisfaite de la démarche, des possibles recours aux contextes et visualisations des ressources qui lui ont permis de « comprendre enfin les différences entre tout ça ».

◆ *ce/cet*

La recherche spontanée de plusieurs apprenants de *ce* et *cet* m'a un peu étonnée mais a renforcé l'utilité de l'ASC. L'origine de cette recherche est liée au fait que les apprenants qui l'ont effectuée ne parvenaient pas à comprendre la différence entre les deux formes et, par conséquent, qu'ils ne savaient pas les utiliser. En observant verticalement les occurrences et le cotexte droit de *ce* et *cet*, les apprenants ont saisi quasiment de façon instantanée la différence. L'un d'entre eux a verbalisé : « ben c'est facile ce après il y a une consonne et cet après il y a une voyelle ». Étonnée de cette recherche, j'essaie de comprendre en demandant à ce même apprenant de m'expliquer notamment s'il n'avait jamais entendu ou abordé cette règle dans le cadre d'une grammaire explicite. Il a répondu, comme les autres apprenants, qu'il avait déjà abordé la différence en cours et même de nombreuses fois, mais que c'est en visualisant les occurrences, nombreuses et verticalement, qu'il a compris.

◆ *hein*

Les spécificités de l'oral ne sont généralement pas connues des apprenants parce qu'elles n'ont pas été enseignées. Dans l'objectif d'apprendre à interagir en observant des corpus d'interactions, ces spécificités interrogent les apprenants. Dans une des expérimentations qui consistaient à analyser avec les apprenants le déroulement d'une interaction, le marqueur *hein* a les a intrigués. La recherche de *hein* dans le concordancier de FLEURON affiche 332 occurrences. Le travail a été réalisé de façon individuelle mais les quinze apprenants de la classe ont fait la même analyse. Collectivement, ils parviennent aux observations suivantes, classées en trois catégories. Le tableau ci-dessous présentent les analyses avec certaines paroles des apprenants reproduites en substance :

Pour solliciter	Pour exprimer une émotion	Pour donner du rythme
- Pour demander une question on n'utilise pas <i>est-ce que, quand ou comment</i> mais juste <i>hein</i> - Je peux utiliser ce mot pour poser une question - Il peut servir pour faire une proposition - Peut exprimer un doute	- Renforce le ton - Peut être utiliser pour une exagération - Peut remplacer <i>ah ?</i> - Pour montrer qu'on est surpris - Représente l'émotion des gens qui parlent	- Peut être utilisé quand on hésite - Pour faire une pause - Permet de s'arrêter, de réfléchir à la suite

- Signifie <i>c'est vrai</i> ou <i>comment</i> - En début de tour, il ne veut rien dire mais à la fin c'est pour questionner - Pour demander qu'on nous écoute		
--	--	--

Les apprenants concluent également que *hein* n'est utilisé qu'à l'oral, qu'ils l'entendent souvent, que parfois ils ne le reconnaissent pas dans les interactions et qu'ils ne l'avaient jamais abordé en cours. Les analyses des apprenants sont pertinentes. Le résultat de cette expérimentation est proche de celui auquel sont parvenus les étudiants étrangers dans les expérimentations de Braun (2019) avec des étudiants chinois. Les conclusions des apprenants s'approchent des études du marqueur *hein* à l'oral (Noda 2018) et réalisées dans le cadre d'analyse des interactions. Par exemple, dans André (2006a), j'ai relevé les activités interactionnelles et les valeurs suivantes : poser une question, demander une répétition, marquer sa surprise, demander une confirmation, demander l'attention de ses interlocuteurs, insister, accentuer, rappeler un savoir commun ou encore marquer un ton polémique. Une transposition didactique de ce type d'analyses pourrait venir compléter l'observation des apprenants effectuée grâce à la méthodologie de l'ASC et les aider à apprendre à interagir.

◆ *Des recherches de thème*

Lors d'une première séance d'ASC avec des étudiants chinois (Braun 2019), lorsqu'ils se sont lancés dans les recherches de leur choix, ils ont utilisé le concordancier pour rechercher des thèmes, comme ils l'auraient fait dans un moteur de recherche comme Google. Ces recherches sont sans doute orientées par la façon de présenter le concordancier. Comme je l'ai mentionné précédemment en présentant certaines consignes, pour que les apprenants comprennent le fonctionnement du concordancier, une comparaison est souvent faite avec un moteur de recherche. Néanmoins, ces recherches sont intéressantes. Parmi les items recherchés, les mots les plus fréquents sont : *université*, *SNCF*, *SIM*, *rendez-vous*, *campus*, (carte de) *séjour* et *médical*. Les étudiants ont expliqué les raisons de chacune de ces recherches. Je synthétise leurs explications ci-dessous :

- *université* : pour apprendre des choses sur l'administration, la vie étudiante, le déroulement des cours.

- *SNCF* : pour savoir comment prendre le train pour partir pendant les vacances, comment acheter des billets de train.
- *SIM* : pour savoir comment choisir un opérateur, connaître les démarches pour avoir un téléphone qui fonctionne en France.
- *rendez-vous* : parce qu'en France « il faut prendre des rendez-vous pour tout » de plusieurs façons (en face à face, au téléphone, par Internet).
- *campus* : pour avoir des informations sur les inscriptions, connaître les infrastructures du campus.
- *séjour* : pour connaître les démarches à faire pour renouveler une carte de séjour.
- *médical* : pour savoir quels sont les vaccins à effectuer, savoir à quoi sert le numéro de sécurité social.

Si ces recherches sont surprenantes en tant qu'item pour le concordancier, elles sont tout à fait cohérentes avec les préoccupations des étudiants et reflètent leurs besoins en termes d'apprentissage. Lors de ces recherches, il a pu être constaté que chaque recherche a été utile pour les apprenants. Chaque mot a plusieurs occurrences dans FLEURON : *université* 112, *SNCF* 9, *SIM* 5, *rendez-vous* 43, *campus* 65, *séjour* 59 et *médical* 11. Les apprenants se sont lancés dans une analyse de l'apparition de ces occurrences, mais pour obtenir des informations, pas pour apprendre comment elles fonctionnaient en interaction. Bien que ces recherches ne correspondent pas aux objectifs de l'ASC, les apprenants ont pu observer les interactions dans lesquelles ces unités lexicales sont utilisées. Ils ont ainsi été exposés à la langue en fonction de leurs objectifs d'apprentissage, plutôt socio-culturels dans ces cas.

◆ *Des analyses inattendues*

L'ASC est aussi parfois un apprentissage par « serendipité », c'est-à-dire par hasard (*serendipitous findings*) comme l'explique Bernardini (2000) ou un apprentissage « incidentiel » (*incidental learning*), c'est-à-dire non intentionnel comme l'expliquent Duda et Tyne (2010). En partant à la découverte de certaines données, les apprenants en explorent d'autres. Ils deviennent des explorateurs ou des « *travellers* » (Bernardini 2004 : 22). Par exemple, lors d'une séance individuelle, le fonctionnement du site est expliqué à un apprenant ukrainien. Une démonstration de l'utilisation du concordancier est réalisée avec le mot *depuis*. Dès que les résultats s'affichent, l'apprenant interroge

l'expérimentateur ⁷⁸ sur un autre mot. Voici la première ligne des résultats du concordancier :

A gauche du mot	Mot recherché	A droite du mot
E : euh	depuis	euh mardi

Figure 46 : Première ligne du résultat de la recherche de « depuis » dans FLEURON

Immédiatement, l'apprenant montre *euh*, comme on peut le voir sur la vidéo suivante :

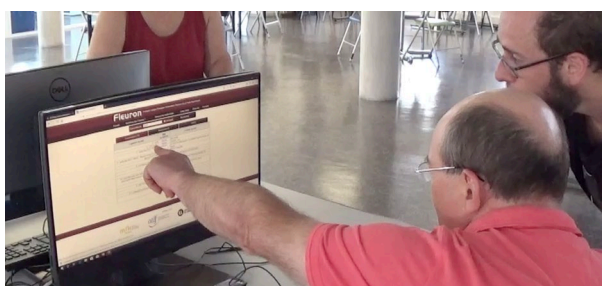


Figure 47 : Prise de vue du film de l'expérimentation

L'apprenant demande : *qu'est-ce que c'est euh*. L'enseignante (derrière l'écran) et l'expérimentateur s'engagent alors dans une explication en fournissant à l'apprenant un exemple contextualisé. Ensuite, *euh* est recherché dans le concordancier pour une rapide analyse et un surtout constat de sa fréquence élevée dans le corpus (1462 occurrences). De cette manière, dès la présentation de FLEURON, cet apprenant a été rassuré sur l'élaboration des discours à l'oral et sur la fréquence des hésitations chez les locuteurs natifs du corpus. Cette interaction montre également l'importance de la multimodalité lors de l'utilisation du concordancier accompagnée par un enseignant. Les apprenants n'ont pas toujours les compétences socio-interactionnelles pour verbaliser ce qu'ils observent, ce qui leur pose question, etc. Dans ces cas, le pointage est important. On verra qu'il l'est d'autant plus avec des apprenants de faible niveau (en [5.4.1.](#)).

Dans une autre séance, j'accompagne un apprenant qui se concentre sur l'analyse de la progression d'une interaction, intitulée « Réserver une salle à la maison de l'étudiant », entre un étudiant et le responsable de la maison de l'étudiant. L'apprenant s'arrête parce qu'il ne connaît pas le mot « créneaux » dans l'énoncé *je ne sais pas si vous avez encore des*

⁷⁸ Biagio Ursi, en contrat post-doctoral au sein du projet FLEURON.

créneaux pour lundi. Son premier réflexe est de rechercher *créneaux* dans un dictionnaire. Avant qu'il ne puisse y accéder, je l'arrête et lui demande de faire une recherche dans le concordancier de FLEURON. Il affiche donc les 11 occurrences de *créneaux*. Je l'aide à faire son analyse en observant les lignes avec lui, voici un extrait de notre interaction :

- VA **regardez ce qu'il y a avant ce qu'il y a après normalement vous pouvez déduire le sens des cré- si on prend le premier des créneaux 13 heures 20 heures 17 heures minuit deuxième vos créneaux de TD vous savez ce que c'est qu'un TD**
- E oui le travail dirigé-
- VA dirigé oui c'est un cours hein **ensuite vous avez des créneaux horaires +**
- E hum
- VA **ensuite pour choisir vos créneaux** donc pour l'instant on a quoi on a des heures + des choix + **ensuite vous n'avez pas beaucoup de contraintes il y a des créneaux qui sont assez libres** et ensuite ce sont **les créneaux de 18 heures** ensuite les **créneaux de vendredi après-midi** alors vous commencez à comprendre ce que c'est qu'un créneau
- E **ouais c'est comme réserver de de comme rendez-vous c'est un rendez-vous**
- VA oui alors un créneau c'est 14 heures 15 heures ça c'est un créneau par exemple ensuite vous avez le choix des créneaux horaires en fait c'est (fait des gestes pour montrer une durée dans le temps)
- E ah oui
- VA c'est une durée de 14 heures à 15 heures ou entre 16 heures et 17 heures ça ce sont des créneaux différents
- E ah oui oui oui j'ai compris

L'objectif de l'accompagnement aussi fort de l'analyse de *créneaux*, ligne par ligne, est de montrer à cet étudiant qui a le « réflexe dictionnaire », qu'il peut trouver le sens d'un mot grâce aux corpus. Je ne peux pas affirmer qu'après ce type de démonstration, cet étudiant aura, ce que j'appelle, le « réflexe corpus » mais il a compris qu'il était capable d'interpréter des lignes de concordances (Todd 2001) et de saisir le sens de l'item recherché. Par ailleurs, cet échange apparaît dans un cadre expérimental qui vise à former des apprenants à l'ASC. Cette approche n'est pas exclusive, si l'utilisation d'un dictionnaire est plus simple, plus rapide et plus efficace pour un apprenant, elle n'est évidemment pas à exclure.

5.3.4. Bilan des recherches d'occurrences

Les occurrences recherchées dans le concordanciers de FLEURON sont de différentes natures et ont différentes origines. Elles varient en fonction de plusieurs

éléments, notamment des objectifs d'apprentissage, de l'accompagnement de l'enseignant, du degré d'autonomie des apprenants ou encore des habitudes didactiques des apprenants. Je vais dresser un bilan de ces recherches en salle de classe, et de leurs exploitations par les enseignants et les apprenants. Ce bilan est réalisé à partir des pratiques observables, c'est-à-dire qu'elles qui ont fait l'objet d'un recueil de données (enregistrements, films, carnet de bord parfois) auprès de la centaine d'étudiants-apprenants-utilisateurs de FLEURON.

Tout d'abord, en ce qui concerne les pratiques professionnelles des enseignants, il y a deux façons de demander à des apprenants d'utiliser le concordancier : soit l'enseignant suggère ou prescrit les recherches, soit l'apprenant décide seul de ses recherches. Il n'est pas rare que les enseignants demandent aux apprenants de rechercher des éléments particuliers ou donnent des listes d'item à rechercher. Ce procédé n'est pas très conventionnel, si on se réfère à la définition prototypique du DDL proposée par Boulton (2011 : 572) :

The hands-on use of authentic corpus data (concordances) by advanced, sophisticated foreign or second language learners in higher education for inductive, self-directed language learning of advanced usage.

Cependant, plusieurs explications à ces guidages de la part des enseignants ont été constatées lors de mes expérimentations. La première est liée au fait que la méthodologie proposée par l'ASC est nouvelle et inconnue. Les apprenants n'ont jamais effectué un tel travail et les enseignants non plus (au moins lors des premières séances).

L'objectif est de ne pas mettre un apprenant en difficulté face au concordancier. Dans ce cas, un item ou une liste d'items peuvent être proposés à la recherche. C'est d'ailleurs ce que j'ai moi-même fait dans plusieurs de mes expérimentations, au moins au début pour faciliter la prise en main de l'ASC⁷⁹. La seconde explication est liée à la progression pédagogique prévue par les enseignants. Certains, pour ne pas rompre cette progression, proposent des items en lien avec le cours dans lequel s'inscrit la séance sur FLEURON. C'est le cas par exemple des recherches de « voulais » ou « j'étais » à la suite d'un cours sur l'imparfait, de « est-ce que » à la suite d'une présentation des questions, etc. Dans certaines situations, même lorsque les apprenants décident seuls de leurs recherches, elles sont influencées par le cours. Ainsi, les apprenants qui utilisent FLEURON à la suite d'un

⁷⁹ Voici un exemple de liste : *sans, comme, quoi, pour, chaque, ai, voulais, en, qui, bien, juste*.

cours traditionnel ont tendance à chercher dans le concordancier des éléments évoqués pendant ce cours. Le rôle de l'enseignant dans ce processus est doublement difficile à évaluer. D'une part, le degré d'autoritarisme de l'enseignant peut influencer la marge de liberté que les apprenants pensent pouvoir s'accorder. D'autre part, la figure de l'enseignant qui détient le pouvoir et telle qu'elle peut être perçue par les apprenants influence également cette marge de liberté.

Parfois les apprenants s'émancipent davantage du cours. Ils recherchent ce qui les intéressent. Il n'est pas rare également de constater que des apprenants, lorsqu'une liste d'items est proposée par l'enseignant, font d'autres recherches. Laisser la possibilité aux apprenants de faire leur propre choix a différentes conséquences, selon les apprenants et éventuellement leur style d'apprentissage (Duda, Riley 1990). Certains n'ont pas l'habitude d'avoir une certaine autonomie dans leur apprentissage et d'autres trouvent enthousiasmant le fait que l'enseignant leur donne la possibilité de travailler ce qu'ils souhaitent. Enfin, les enseignants peuvent proposer la recherche en corpus d'un élément pour lequel ils ne peuvent pas donner de règles de fonctionnement. Johns (1991) explique que l'ASC est utile pour ce qui n'est pas enseigné et ce qui n'est pas enseignable ou difficilement enseignable (*unteachable*). À différentes questions posées par des apprenants, les enseignants se trouvent fort démunis pour y répondre. Par exemple, les questions liées à des éléments de la gestion de l'interaction ou à des marqueurs interactionnels mettent les enseignants souvent dans l'embarras, notamment parce qu'ils ne sont pas formés à l'analyse des interactions. La réponse adéquate, dans ce genre de situations, est de demander aux apprenants de chercher ces éléments et de les analyser eux-mêmes. L'enseignant formé à l'ASC peut évidemment faire également la recherche ou accompagner l'apprenant dans la sienne. Par exemple, les apprenants interrogent souvent les enseignants sur le fonctionnement de *alors* ou de *donc* à l'oral. Ces petits mots de l'oral font partie des items les plus recherchés dans le FLEURON (voir tableau 3 ci-dessous). Pour rappel, des questionnements sur le fonctionnement de *genre* en interaction ont également conduit des enseignants à proposer des recherches d'occurrences dans FLEURON.

Quelle que soit la façon dont les enseignants proposent l'ASC à leurs apprenants, un des principaux objectifs est de réduire leur rôle et leur importance (Johns 1991) voire, petit à petit, de le supprimer (pour certaines activités liées à l'ASC) dans les cas où, au fur et à mesure, les apprenants deviennent autonomes (Aston 2001). En réalité, plusieurs postures peuvent être endossées par les enseignants selon les situations d'apprentissage

(équipement des salles, par exemple) et les apprenants (leur culture d'apprentissage, leur niveau en langue et d'autonomie). Une fois que les consignes sont exposées et comprises, les enseignants peuvent adopter différentes façons de faire. Ils peuvent accompagner l'ASC avec chaque apprenant en les aidant à interpréter les résultats du concordancier. Cette pratique peut être mise en œuvre pour tout type de recherche ou pour répondre à des questions plus compliquées (Kerr 2013). Dans ce cas, l'enseignant est une sorte de médiateur, il co-construit toutes les étapes de l'ASC avec l'apprenant : observation, classification, généralisation (Johns 1991, 1997). Dans mes expérimentations, j'ai moi-même endossé ce rôle de guide et d'aide à l'induction des règles de fonctionnement. Il existe des façons moins intrusives ou plus autonomisantes de faire ou de faire faire. L'enseignant peut impulser le début du travail, s'assurer que l'apprenant a bien compris ce qu'il devait faire puis le laisser terminer son activité. L'ASC demande souvent un peu de temps avant que les apprenants se familiarisent avec sa méthodologie et avant qu'ils aient le « réflexe corpus », c'est-à-dire qu'ils répondent à leur question sur la langue et son fonctionnement en allant observer sa réalisation en corpus. A l'opposé de ces pratiques, l'enseignant peut laisser les apprenants seuls face aux concordanciers. Ainsi, dans la mise en pratique de l'ASC, on peut constater un continuum entre un accompagnement guidé de l'apprenant et une absence d'accompagnement. Le succès ou l'échec de cette dernière option dépend largement des apprenants, de leurs objectifs, de leur niveau et de leurs pratiques.

En effet, en ce qui concerne plus particulièrement les apprenants, les pratiques de l'ASC vont également varier. Tout d'abord, selon leur niveau en langue, les recherches ne sont pas les mêmes. Les étudiants les moins avancés semblent rechercher plutôt des éléments lexicaux et les plus avancés des éléments grammaticaux puis des spécificités de l'oral. Parmi les avantages de l'ASC, on peut citer les possibilités de montrer le sens des mots, montrer que les mots ont des contextes d'utilisation différents (Kamber 2011) ou encore de lier lexicale et grammaire (Conrad 2000). Ces possibilités sont rapidement exploitées par les apprenants, parfois progressivement, en allant de la recherche lexicale à la recherche de mots grammaticaux et de pratiques interactionnelles. Certains apprenants ont plus ou moins besoin d'un accompagnement de la part d'un enseignant (Bernardini 2000). Par exemple, au risque d'alimenter de nombreux stéréotypes, les étudiants chinois apprécient particulièrement d'être accompagnés ou guidés par un enseignant (Guinamard, Zhang 2015), au moins au début de leur pratique de l'ASC. En tout cas, ces étudiants restent

souvent perplexes à l'idée qu'ils peuvent choisir ce qu'ils ont envie d'apprendre. Les expérimentations menées par Olivier Braun, stagiaire de Master 2, au sein d'une classe d'étudiants, exclusivement chinois, ont montré que pour rassurer ces étudiants, il fallait passer par une étape prescriptive (en leur proposant des items à rechercher). Cependant, dans d'autres expérimentations, également avec des étudiants chinois, si les débuts furent hésitants, ces étudiants se sont réjouis à l'idée de pouvoir effectuer leurs propres recherches et faire des choix d'apprentissage. Ainsi, l'ASC n'est pas manipulé avec le même niveau d'aisance selon les apprenants. Ce besoin d'accompagnement ou de liberté de l'apprenant vient se combiner avec ce même besoin de l'enseignant. Différentes combinaisons émergent et rendent l'ASC flexible et adaptable à différentes situations.

Les apprenants semblent tous développer différentes stratégies d'apprentissage en se prêtant à l'exercice de l'ASC (André, Ciekanski 2018). Ces stratégies relèvent de différents types, selon la typologie de Oxford (1990) :

- des stratégies *cognitives* : analyser les items des lignes de concordances, faire des comparaisons, trouver des synonymes, faire des hypothèses, faire des inférences, transférer d'une langue à l'autre, observer les items en interaction, induire les règles de fonctionnement ;
- des stratégies *métacognitives* : prendre une vue d'ensemble pour établir des liens, prêter attention, identifier le but d'une pratique interactionnelle (se préparer à interagir dans telle situation universitaire), planifier une activité interactionnelle (choisir des phénomènes et des items à explorer), évaluer une pratique interactionnelle (choisir de ne pas utiliser en production orale tel fait de langue recherché), s'autoévaluer en confrontant le fonctionnement de la langue déduit de l'observation à d'autres contextes dans le concordancier ; utiliser en situation un item recherché et analysé ;
- des stratégies *socio-affectives* : prendre des risques mesurés pour soutenir la motivation (exploration de faits de langues qui posent des problèmes et qui sont mal compris ; choix de ressources plus ou moins longues ; choix de ressources dans des contextes plus ou moins connus) ; coopérer avec autrui en comparant les règles avec d'autres apprenants ou en aidant les autres.

Par exemple, la stratégie métacognitive « prendre une vue d'ensemble pour établir des liens » met en exergue la capacité de l'apprenant à saisir les rapports entre un élément nouveau et ce qu'il connaît déjà et à réorganiser son apprentissage, à orienter ses

utilisations du corpus (focalisation sur le co-texte ou sur le contexte des vidéos), en s'appuyant pleinement sur la multimodalité de FLEURON. En cela, le concordancier multimodal apparaît être un atout pour le développement de la conscience métalangagière des apprenants.

Enfin, je terminerai ce bilan en évoquant les items recherchés par les apprenants. En d'autres termes, je vais tenter de comprendre quels sont les éléments que l'ASC, dans ce cadre, a permis d'observer et de comprendre. L'ASC étant mis en œuvre le plus fréquemment pour un travail sur les compétences écrites, les items recherchés pourraient être différents lorsqu'il s'agit d'apprendre à interagir. Le site FLEURON permet de savoir quels sont les items recherchés dans le concordancier.

La liste suivante présente les 20 items les plus recherchés par ordre décroissant, c'est-à-dire des plus recherchés au moins recherchés.

Rang	Item recherché	Rang	Item recherché
1	alors	11	euh
2	donc	12	genre
3	en fait	13	quoi
4	du coup	14	hein
5	bon	15	dessus
6	enfin	16	inscription
7	voilà	17	quand même
8	même	18	c'est
9	d'accord	19	mais
10	ben	20	je vous en prie

Tableau 3 : Liste des 20 items les plus recherchés dans FLEURON

Cette liste montre bien que les items recherchés dans un corpus oral sont différents de ceux qui pourraient être recherchés dans des corpus écrits. Le premier item (*alors*) est recherché parce qu'il est fréquemment entendu par les apprenants de FLE qui ne le comprennent pas. De plus, *alors* est également utilisé à l'écrit pour d'autres fonctions, ce qui rend sa compréhension plus compliquée à l'oral. En effet, les analyses d'interactions verbales montrent que *alors* est fréquemment employé en début de tour de paroles pour prendre son tour. Ces analyses sont observables en corpus. Par exemple, dans FLEURON, sur les 341 occurrences de *alors*, seulement 26 occurrences ne sont pas des débuts de tour de paroles (si on considère les combinaisons du type *bon alors*, *ben alors*, *ah alors*, *ah oui alors*, *donc alors*). La même analyse est réalisable avec *donc*. On peut remarquer que *donc* est le

deuxième item le plus recherché. Les interrogations des apprenants quant à son fonctionnement sont fréquentes et sont à mettre en perspective avec ses fréquentes utilisations. Le corpus FLEURON compte 1012 occurrences. Dans la très grande majorité des tours de parole qui comportent *donc*, il est placé en début de tour de parole.

Ainsi, il est facile de constater que les apprenants recherchent prioritairement des petits mots de l'oral (*alors, donc, en fait, du coup, bon, etc.*), c'est-à-dire les éléments qui permettent de structurer le discours ou gérer l'interaction. Cette recherche est fortement liée à l'absence de description de ces éléments dans les manuels et les grammaires (Zorzi 2001), voire dans le discours des enseignants (Bouchard 2002). Les apprenants entrent dans le concordancier en recherchant des mots qu'ils entendent fréquemment. Pour rappel, ces expérimentations ont été réalisées en France, par conséquent avec des étudiants étrangers en immersion. Ainsi, au quotidien, ces étudiants participent à des interactions verbales en français, mais ne comprennent pas le sens et la valeur pragmatique d'un certain nombre d'éléments. Les recherches dans FLEURON commencent donc souvent par des items qui leur posent question et qui nuisent à la compréhension des interactions auxquelles ils participent. Une étudiante m'a expliqué que dès son arrivée en France, elle a rencontré des problèmes avec « regarde », qui fonctionne comme un marqueur interactionnel. Elle m'expliquait qu'à chaque fois qu'elle l'entendait, elle scrutait autour d'elle pour trouver ce qu'il fallait qu'elle regarde. Pour comprendre le sens de ces marqueurs, ponctuels ou régulateurs, l'ASC est précieux.

Une hypothèse possible pour expliquer ces recherches est peut-être que les apprenants cherchent avant tout des éléments grammaticaux, qu'ils connaissent parce qu'ils sont utilisés à l'écrit. Ainsi les apprenants se servent-ils du concordancier pour saisir les liens entre lexique et grammaire, avant de se rendre compte que les éléments qu'ils recherchent relèvent davantage de pratiques interactionnelles permettant de structurer le discours à l'oral et de gérer l'interaction. Quelle soit la façon d'entrer dans le corpus *via* le concordancier, les apprenants évoluent dans leur utilisation de l'ASC. Ainsi, petit à petit, les aspects sociolinguistiques sont travaillés puisque les phénomènes interactionnels sont observés et analysés en contexte. De cette façon, les apprenants observent la variation de ces phénomènes en fonction de la situation de communication.

Dans les items les plus recherchés, on trouve également la marque d'hésitation *euh*. Cette recherche est effectuée par les apprenants, après qu'ils ont compris la transcription graphique de l'hésitation. Les nombreuses occurrences de *euh* (plus de 1500 dans FLEURON)

rassurent les apprenants, comme je l'ai expliqué précédemment. Ils se rendent compte que les locuteurs natifs produisent des hésitations et que ce ne sont pas des marques d'incompétences interactionnelles (André, Tyne 2012). Les recherches de *ben* et de *hein* permettent également aux apprenants de se rendre compte de la verbalisation des hésitations ou des pauses remplies (Blanche-Benveniste 2000, Henry *et al.* 2004).

Le mot *genre* apparaît en douzième position dans la liste. Comme je l'ai mentionné précédemment en 4.2.2., les occurrences fréquentes de *genre* sont récentes et concernent essentiellement des locuteurs jeunes (entre 15 et 25 ans) qui participent à des conversations informelles. Ces éléments situationnels expliquent le fort taux de recherche de *genre* dans le concordancier de FLEURON. Ce dispositif a été conçu pour des apprenants qui s'apprêtent à faire des études en France ou qui sont déjà en France pour un séjour universitaire. Le public cible, et ici expérimental, correspond parfaitement aux locuteurs utilisateurs de *genre* en tant que marqueur interactionnel (André 2018/soumis). Par conséquent, il n'est pas surprenant que les apprenants entendent *genre* dans leurs échanges quotidiens avec d'autres étudiants ou d'autres locuteurs du même âge et qu'ils cherchent à en comprendre ses sens et son fonctionnement en interaction.

Parmi les 20 items les plus recherchés, il y a également un mot lexical (*inscription*), un *chunk* ou un bloc de mots (*quand même*) et une unité entière, une locution ou une expression (*je vous en prie*). Ces recherches sont liées à certains des éléments déjà mentionnés précédemment. La recherche de mots lexicaux est généralement conduite par des apprenants les moins avancés en français. Toutefois, le contexte universitaire demande que le mot *inscription* soit bien maîtrisé par les étudiants. En outre, la procédure d'inscription est complexe. L'analyse des contextes et cotextes d'apparition est donc influencée par les objectifs d'apprentissage des apprenants. Les mots lexicaux les plus recherchés sont en lien avec le domaine universitaire ou la vie étudiante. Voici la liste des 10 mots lexicaux les plus recherchés après *inscription* : *travailler*, *semaine*, *rendez-vous*, *assurance*, *dossier*, *étudiant*, *train*, *étrangers*, *examen*, *langue*. Ensuite, les recherches de blocs de mots (*chunks*) sont assez fréquentes. Le fait que *quand même* apparaisse en dix-septième position est lié à la complexité de *même* et des problèmes de compréhension qu'il pose, comme je l'ai déjà évoqué précédemment. D'autres blocs de mots sont recherchés dans FLEURON, par exemple : *étant donné que*, *en fin de compte*, *des trucs qui sont*, *de toute façon*, *si vous voulez*, *en même temps*, ou encore des répétitions telles que *oui oui oui*. Enfin, des expressions liées à la politesse sont également recherchées dans le concordancier, telles

que *je vous en prie, s'il vous plait, bonjour, salut ou merci*. Ces recherches montrent que les apprenants se servent du concordancier pour apprendre à exprimer des formules de politesse, pour apprendre notamment où les placer dans les interactions, à qui les adresser et comment y répondre. Les aspects pragmatiques sont étudiés lors de ces recherches. Les apprenants de FLE en France se rendent rapidement compte que les échecs pragmatiques (Riley 2000) sont plus nuisibles que des erreurs de grammaire. Ces recherches semblent souvent être menées par les apprenants les moins avancés en français, comme je vais l'expliquer dans la section suivante.

5.4. Des usages détournés de FLEURON

Les résultats positifs des expérimentations de FLEURON avec son public cible – des étudiants étrangers – m'ont amenée à tenter de nouvelles expérimentations avec des publics qui ne sont absolument pas visés par le projet et généralement pas visés par l'ASC (Boulton 2008b). Mes travaux dans le domaine du français langue d'intégration et d'insertion (FL2I) m'ont naturellement orientée vers ces publics – des adultes migrants et des adultes francophones natifs en insécurité langagière. La « puissance de l'apprentissage sur corpus » (Johns 1991) mérite d'être testée avec des publics en difficulté. Comme lors des expérimentations précédentes, ce que je souhaite analyser ce sont les processus analytiques mis en œuvre par les utilisateurs du corpus avec le concordancier. J'examinerai tout d'abord les expérimentations avec des adultes migrants débutants en français et peu scolarisés ([5.4.1.](#)) et ensuite celles menées avec des adultes francophones natifs en insécurité langagière ([5.4.2.](#)).

5.4.1. Avec des adultes migrants peu scolarisés

Dans le chapitre précédent, j'ai présenté la possibilité d'exploiter les ressources du corpus FLEURON avec un public migrant, afin de réaliser conjointement (formateur/apprenants) une analyse sociolinguistique des interactions verbales (voir [4.3.2.](#)). A la suite de ces activités analytiques, j'ai mis en place des séances consacrées à l'introduction de l'ASC avec ce public. J'ai donc proposé à la formatrice, qui m'accueille dans

ses cours, et aux stagiaires de faire des recherches dans le concordancier de FLEURON. J'ai émis l'hypothèse que le concordancier, et les activités métalangagières associées, pouvaient être utilisés avec un public migrant débutant et peu scolarisé. C'est une sorte de double défi que j'ai souhaité relever : la mise en place d'activités relevant de l'apprentissage sur corpus oraux et multimodaux, d'une part, et avec un public de faible niveau en français et peu scolarisé, d'autre part.

Afin de faire comprendre les objectifs de l'ASC, une première séance collective a été réalisée. Elle a suivi l'analyse collective de la vidéo « Acheter un billet de train au guichet de la SNCF ». Au cours de cette interaction, la guichetière donne le montant du billet : *donc ça fera 3 euros 80 s'il vous plaît monsieur*. Avec Charlène Weyh, en stage de Master 2 au sein de groupe, nous avons utilisé l'apparition de *s'il vous plaît* pour leur expliquer le fonctionnement du concordancier. Cette recherche est liée aux questions posées par les stagiaires sur l'emplacement de *s'il vous plaît* dans un énoncé. Certains d'entre eux nous ont interrogés sur la possibilité de commencer un tour avec cette formule de politesse. Nous avons recherché toutes les occurrences de *s'il vous plaît*, affiché les résultats et analysé les 18 lignes de concordances.

Concordancier (18)	Ressources (0)	CNRTL
A gauche du mot	Mot recherché	A droite du mot
E : bonjour A : bonjour E : je voudrais m'inscrire	s'il vous plaît	hum
A : ou alors vous me met- vous me notiez votre nom là	s'il vous plaît	E : hum bien sûr
E : hum	s'il vous plaît	
A : donc une petite signature là	s'il vous plaît	
E: euh j'aurais voulu me renseigner euh pour euh savoir comment on s'inscrit au SIUAP	s'il vous plaît	
E: ce serait pour un aller-retour Epinal Thaon	s'il vous plaît	
A: donc ça fera 3 euros 80	s'il vous plaît	monsieur
E: oui A: voilà votre nom	s'il vous plaît	
E: bonjour je vais vous prendre un thé à la menthe	s'il vous plaît	
A: ok alors j'aurais besoin de votre carte d'étudiant	s'il vous plaît	E: d'accord
A: bon je peux vous redonner ça donc il y a un règlement de 170 euros donc vous payez par carte donc on va on va aller régler là-bas E: par carte	s'il vous plaît	si euh pas de soucis
A: d'accord vous me donnez votre numéro d'étudiant	s'il vous plaît	
A: s'il vous plaît E: euh est-ce que je pourrais avoir un américain sauce blanche	s'il vous plaît	E: euh est-ce que je pourrais avoir un américain sauce blanche s'il vous plaît
E: oui A: alors 3 euros et 20 centimes	s'il vous plaît	
E: je le montre A: oui	s'il vous plaît	
E: hum OK attendez un petit peu	s'il vous plaît	
A: est-ce que vous avez votre numéro étudiant	s'il vous plaît	E: numéro étudiant

Figure 48 : Liste des occurrences de « s'il vous plaît » dans FLEURON

Les stagiaires ont donc participé à une première expérience en groupe. Nous avons projeté sur un écran les résultats du concordancier et l'analyse s'est faite collectivement. Le problème du travail collectif dans un groupe hétérogène est bien connu, certains participent

pendant que d'autres attendent que ça se passe. J'en profite pour rappeler que les quinze stagiaires ont tous été évalués au niveau A1. Cependant, comme souvent à l'issue des tests de positionnement basés sur le CECRL, aucun n'a le même niveau, certains sont capables de s'exprimer un peu à l'oral et d'autres aucunement.

Après cette première expérience collective, j'ai proposé un travail individuel, ou en binôme selon les ordinateurs disponibles dans la salle. Nous avons fourni une liste de mots et de syntagmes de différentes natures, par exemple : *dessus, vu que, oui donc, je vous en prie, je voudrais, j'aimerais, ce serait, même, puis*. Cette liste a été établie en collaboration avec la formatrice et Charlène Weyh. Elles se sont appuyées sur les connaissances des stagiaires, c'est-à-dire sur des éléments interactionnels abordés dans les séances de formation précédentes, pour faire apparaître certains éléments dans la liste.

L'attention de plusieurs stagiaires a été portée sur « je vous en prie », notamment parce que l'expression est apparue à la fin de l'interaction qu'ils venaient d'analyser. Cette pratique interactionnelle n'était pas ou peu connue par les apprenants. Certains ont donc effectué une recherche *via* le concordancier de FLEURON, voici une capture d'écran des 24 résultats :

Concordancier je vous en prie ☒ mot exact Rechercher		
Concordancier (24)	Ressources (0)	CNRTL
A gauche du mot	Mot recherché	A droite du mot
E : merci A :	je vous en prie	
A : oui bah	je vous en prie	E : bonne après-midi A : hum
E : d'accord bah merci bien A :	je vous en prie	monsieur E : au revoir A : au revoir
E : bah merci A :	je vous en prie	
E : bon bah je vais aller voir les emplois du temps je vous remercie A : d'accord	je vous en prie	
E : ok merci A : voilà	je vous en prie	E : au revoir A : au revoir
A1: vous pouvez rentrer	je vous en prie	
E : je vous remercie A :	je vous en prie	
E2 : on vous remercie beaucoup E1 : merci A : bah	je vous en prie	au revoir
E : merci beaucoup A :	je vous en prie	
E : merci A :	je vous en prie	
E : merci beaucoup A :	je vous en prie	, au revoir E : au revoir
1: ben merci bien 2: ben	je vous en prie	bonne soirée 1: bonne soirée
A :	je vous en prie	bonne journée E: merci pareillement au revoir A: merci au revoir
E: je vous remercie A :	je vous en prie	bonne journée E: au revoir
E: merci beaucoup A :	je vous en prie	
E: merci beaucoup A :	je vous en prie	
E: okay merci beaucoup bah je prends les papiers puis je reviens un autre jour A :	je vous en prie	okay très bien madame
E: merci au revoir A :	je vous en prie	au revoir madame
A :	je vous en prie	au revoir E: au revoir
E: je vous remercie A :	je vous en prie	
E: euh j'ai A :	je vous en prie	asseyez-vous
A :	je vous en prie	E: bonne journée à vous
A :	je vous en prie	bonne journée monsieur E: OK au revoir et bonne journée

Figure 49 : Liste des occurrences de « je vous en prie » dans FLEURON

Avec la formatrice, nous avons parfois suggéré à des stagiaires de faire cette recherche (en plus de la liste de suggestions écrite au tableau). Cette occurrence n'est pas choisie au hasard. Avec un public pas ou peu scolarisé, comme avec un public d'étudiants débutants, les recherches dans le concordancier sont différentes. Les apprenants de FLE les plus avancés recherchent prioritairement des petits mots de l'oral ou des connecteurs, mais les moins avancés recherchent des pratiques interactionnelles ayant des valeurs pragmatiques qu'ils ne maîtrisent pas. Dans la séance précédente, l'analyse de *s'il vous plait* avait été particulièrement intéressante dans la mesure où les stagiaires ont pu saisir l'emplacement de l'expression, son rôle et ses valeurs pragmatiques. L'observation de l'apparition de *je vous en prie* a été réalisée avec succès avec un public FLI. L'analyse des occurrences dans la liste de concordances a été accompagnée mais les apprenants ont réussi à induire les règles de fonctionnement de cet énoncé (place dans l'interaction et dans le tour de parole, rôle, sens, valeur pragmatique).

Pour illustrer la démarche, je vais commenter la transcription d'une séquence d'ASC avec la recherche de *je vous en prie*. Ce travail est réalisé par un binôme de deux stagiaires, sur le même ordinateur. Les apprenants n'étaient pas obligés de rechercher cet énoncé, c'était une suggestion pendant la séance de travail à partir de l'interaction concernant l'achat d'une carte de piscine (figure 25, en [4.3.2.](#)). La formatrice avait laissé les stagiaires rechercher ce qu'ils souhaitaient mais leur avait tout de même fait des propositions (une liste de mots écrite au tableau). Ce travail, avec ces deux stagiaires, présenté ci-dessous dure six minutes, il intervient vingt-cinq minutes après le début de la séance qui consiste à utiliser le concordancier dans le cadre d'un travail individuel.

Les extraits 1 à 11 de l'enregistrement transcrits ci-dessous comptent trois participants, deux stagiaires (S2 et S3) et moi-même (VA). Chacun de ces extraits est commenté mais ils constituent une interaction continue lors d'une séance de formation. Le premier de ces extraits fait suite à un court échange sur le fait que la stagiaire 2 n'ait pas trouvé d'occurrence de « interagir » dans le corpus. J'explique les raisons de l'absence de ce mot aux stagiaires, le premier extrait commence avec la fin des explications :

1. VA donc en fait il faut que vous cherchiez des mots qui apparaissent qui qui qu'on prononce qu'on utilise souvent
- S2 qu'on utilise souvent
- VA est-ce qu'il y a un mot que vous entendez par exemple et que vous ne savez pas trop comment l'utiliser ou que vous ne comprenez pas bien est-ce qu'il y a quelque chose

que que vous aimeriez chercher /// alors est-ce qu'il y en a un dans la liste que Charlène vous a donné qui pourrait vous intéresser ///

Une liste de mots à rechercher dans le concordancier a été fournie afin de pouvoir amorcer la recherche au cas où les stagiaires ne sachent pas quoi rechercher. Il est fréquent que les apprenants, quel que soit leur niveau, ne sachent pas quoi rechercher lors de leur première utilisation du concordancier. C'est pourquoi, des séances collectives sont souvent proposées pour commencer et que des listes de mots leur sont également fournies.

2. S2 je vous **je vous en prie** *** qu'est-ce que c'est dans quelle question on peut poser ça je vous en prie je vous en prie oui
 VA oui on l'a vu hein tout à l'heure dans la je vous en prie alors c'est un s à la fin
 S2 s

La stagiaire choisit donc de chercher un énoncé entendu lors du travail effectué avec la ressource « Achat d'une carte de piscine ». La recherche de *je vous en prie* avait été donnée en exemple juste avant que les stagiaires se répartissent devant les ordinateurs de la salle.

3. VA voilà ah ben il y a quelqu'un d'autre vous entendez il y a quelqu'un d'autre qui travaille sur ça ah ben il y en a plein hein alors **l'idée ici c'est de regarder un petit peu quand est-ce qu'on l'utilise et qu'est-ce que ça veut dire** là vous avez en fait tous les je vous en prie qui sont utilisés dans toutes les vidéos du site donc elle vous a expliqué un petit peu Charlène hein c'est toutes des vidéos et on a écrit tout ce qui était prononcé
 S2 d'accord

Une fois que la liste des occurrences apparaît à l'écran, je réexplique le principe de l'activité et de l'utilisation du concordancier.

4. VA donc là en fait vous avez tous les je vous en prie qui sont prononcés dans les vidéos et pour vous pour comprendre comment ça fonctionne **ben regardez ce qu'il y a à gauche ce qu'il y a à gauche ce qu'il y a à droite et puis regardez un petit peu si vous réussissez à comprendre comment on l'utilise quand est-ce qu'on l'utilise**
 S2 d'accord oui
 VA **vous pouvez les prendre les uns après les autres**
 S2 oui

Comme souvent avec des publics faiblement scolarisés ou débutants, il faut répéter ou reformuler les consignes plusieurs fois. Ces consignes avaient déjà été données lors une séance collective précédente et répétées avant ce travail individuel.

5. VA par exemple le premier
 S2 merci je vous en prie

- VA ouais alors ça on va le trouver regardez on va en trouver plein des comme ça parce qu'en **fait à gauche là on va trouver plein [la stagiaire 2 pointe du doigt « merci » sur l'écran] de merci** regardez merci je vous en prie ah ben merci bien je vous en prie monsieur ah ben merci je vous en prie + je vais aller voir les emplois du temps je vous remercie d'accord je vous en prie alors à quoi ça sert à votre avis je vous en prie

La stagiaire 2 a repéré que le bloc de mots recherché était souvent précédé de « merci ». Elle pointe plusieurs « merci » sur l'écran, dans différentes lignes de concordances. Je saisis qu'elle a identifié un élément à gauche important et je verbalise son observation.

6. S2 **c'est comme la politesse** je vous en prie c'est voisin de dire merci je vous en prie d'accord
S3 **c'est comme merci à toi**

Les stagiaires commencent à analyser le sens de l'expression recherchée. La première identifie sa valeur de politesse, puis la compare à une autre pratique (« c'est voisin de dire merci ») et le second la compare également à « merci à toi ».

7. VA alors ça vient après merci oui
S3 oui **comme on dit merci**
VA ouais
S3 **on dit merci à toi**
VA oui
S3 **c'est un peu pareil**
VA alors c'est un peu pareil **ça veut dire qu'on l'utilise en réponse**
S3 ouais
VA à merci on en a plein
S2 oui alors je vous en prie oui **on répond à merci**

Dans ce passage, les stagiaires avancent encore un peu dans la compréhension de l'expression. Je leur apporte un élément de compréhension en précisant que ça peut être utilisé comme une réponse. Je fais cette proposition à la suite des trois premières interventions du stagiaire 3 (en gras dans l'extrait). En prononçant ces trois tours, il fait des gestes qui expriment un aller et un retour vers moi. Je comprends que le premier tour est un remerciement et le deuxième une réponse au remerciement. C'est pourquoi, je propose d'analyser *je vous remercie* comme une réponse. Ensuite, je reviens sur le sens, en interrogeant les stagiaires sur la signification.

8. VA alors **qu'est-ce que ça veut dire**
S2 je vous en prie
VA je vous remercie on en a plein hein je vous remercie je vous en prie
S2 je vous remercie je vous en prie bonne journée
VA ici aussi hein merci

- S2 je vous en prie **c'est comme attendu** quoi attendu merci je vous en prie
 VA alors
 S2 **j'ai entendu ce que vous avez dit**
 S3 c'est c'est **c'est des réponses pour merci**
 VA oui *** alors qu'est-ce que ça veut dire à peu près
 S3 ben

La compréhension se co-construit en observant les occurrences (en les lisant parfois) et en faisant des commentaires analytiques sur les occurrences. Ensuite, les stagiaires quittent l'écran des yeux et tentent de verbaliser leur compréhension lors d'un court échange.

9. VA je vous **je vous donne quelque chose**
 S3 **je te dis merci**
 VA oui et moi **je dis je vous en prie** [avec un ton particulier]
 S3 **c'est comme vous voulez pas**
 VA ben c'est comme si on dit
 S2 c'est comme un c'est comme un merci beaucoup je t'en prête et tu me redonnes **j'ai pris en charge qu'est-ce que vous m'avez**
 VA ouais + en fait c'est c'est une formule de politesse hein
 S2 oui
 VA en réponse à merci mais c'est pour dire euh
 S2 oui c'est comme on dit merci beaucoup
 VA voilà merci c'est qu- donc
 S2 je vous en prie
 VA je vous donne quelque chose tu me dis merci moi je te dis je t'en prie je vous en prie
 S2 **c'est comme de de rien**

A la suite de cet échange, accompagné d'un ton et de gestes appropriés pour l'illustrer, la stagiaire 2 comprend la signification et la compare à « de rien », une expression qu'elle connaît.

10. VA exactement ça c'est comme de rien
 S2 **je vous en prie de rien**
 VA exactement c'est exactement ça
 S2 ***
 VA et oui c'est exactement c'est comme de rien **donc vous voyez en fait chercher des mots ça sert à ça ça vous aide à comprendre**
 S3 *** vidéo sans vidéo ou sans
 VA ben vous pouvez après aller voir les vidéos
 S3 ouais

Je valide la compréhension de l'expression *je vous en prie* et j'en profite pour rappeler le fonctionnement du concordancier et son utilité. Ce rappel est destiné à les aider à poursuivre la séance expérimentale.

11. VA ou bien en fait
 S2 moi aussi j'entends tous les jours des mots euh oui donc comment on dit donc [tape donc dans le concordancier] c'est du qu'on entend souvent oui donc
 VA ah donc oui il y en a beaucoup dessus [le stagiaire 3 a écrit dessus dans le concordancier] + non il a voulu chercher dessus alors dessus

Les stagiaires ont bien compris le fonctionnement du concordancier et comment il peut les aider à appréhender des pratiques interactionnelles qu'ils ne maîtrisent pas. Immédiatement, la stagiaire 2 souhaite rechercher « donc » qui semble lui poser problème et le stagiaire 3 tape « dessus » dans le concordancier. La séquence se poursuit avec l'analyse des apparitions et des sens de « dessus » puis de « donc ». Enfin, les stagiaires vont chercher « selon » et consacrer 15 minutes à son analyse, jusqu'à la fin de la séance de formation du jour.

Pour illustrer d'autres démarches d'analyse distributionnelle réalisées par des adultes migrants peu scolarisés, je vais synthétiser des activités d'ASC en présentant simplement les étapes de leurs analyses. Par exemple, lors de plusieurs séances, les stagiaires ont cherché *donc*. Voici une capture d'écran du début de la liste des occurrences (1002 au total) :

Concordancier (1002)	Ressources (0)	CNRTL
A gauche du mot	Mot recherché	A droite du mot
L1 :	donc	oui est-ce que vous pourriez déjà m'expliquer un petit peu quelle est votre fonction ici au pôle européen
L2 : et	donc	pendant 5 ans j'ai été responsable de l'espace langue L1 : hum hum
L2 : et	donc	j'ai intégré le service accueil étudiants et euh
L2 :	donc	nous on les aide et puis il y a *** comme demande d'aides à la CAF
L2 : euh lorraine	donc	c'est c'est un peu ça
L2 : et	donc	c'est uniquement pour les étudiants et nous on est là
L2 :	donc	nous on explique les dossiers à les pièces à fournir
A : voilà	donc	je je ne précise pas professeur de polonais hein vous avez deviné E : hum hum
A :	donc	XX
A :	donc	c'est bon
A : professeur d'espagnol hein	donc	vous pouvez lui envoyer un petit mail aussi
A :	donc	je note l'adresse complète E : hum hum
A : je vais vous donner le formulaire	donc	pour la carte d'étudiante E : d'accord A : et puis on va prendre un rendez-vous ensemble
A :	donc	ça ira vite hein là non cette page non
A : il me faudrait les copies	donc	on peut mettre euh E : en fait est-ce que je peux le faire ici euh remplir ce document
A : c'est ce que vous avez déjà apporté	donc	on vous demande ce n'est pas très cher euh l'adhésion E : hum hum
A :	donc	vous pouvez euh vous pouvez vous inscrire E : d'accord
A :	donc	les cours en général commencent euh

Figure 50 : Extrait de la liste des occurrences de « donc » dans FLEURON

Et voici le résultat de l'analyse d'une stagiaire (ce qui est prononcé en substance par la stagiaire est en italique et ce qui est observé est entre parenthèses) :

1. *c'est un début / c'est pour prendre la parole*
2. *comme c'est comme ça alors je peux faire ça*
3. (« et donc ») *c'est un début et une suite*
4. (« voilà donc ») *c'est ce que je vous disais / je vous confirme*
5. (« oui donc ») *c'est une proposition / oui c'est qu'on avait dit avant*

Cette stagiaire a été scolarisée au Surinam jusque l'âge de 12 ans. Son analyse (partielle) des résultats du concordancier a duré douze minutes. Elle a sélectionné un certain nombre de *donc*. Je lui avais bien précisé qu'elle n'était pas obligée de tous les analyser et qu'elle pouvait passer ceux qu'elle ne comprenait pas du tout pour se concentrer sur ceux qu'elle parvenait à saisir avec moins de difficultés. Ainsi, au fur et à mesure de sa progression avec mon aide, elle réussit à observer les sens de *donc* en contexte, à analyser les différents aspects de ses apparitions, à les classer et à induire quelques règles. Avec cet exemple, on peut également constater que le processus d'analyse proposé par l'ASC peut être utilisé avec des apprenants peu scolarisés. Une des principales différences réside dans l'accompagnement nécessaire par un formateur. S'il faut du temps à des apprenants avancés et bien scolarisés pour s'approprier l'ASC, il semble qu'il en faille encore un peu plus pour des apprenants de faible niveau et peu habitués à réfléchir sur la langue.

Je vais présenter une autre recherche réalisée par une stagiaire ainsi que ses étapes d'analyse. Il s'agit de la recherche de *dessus* dont les résultats sont affichés ci-dessous :

A gauche du mot	Mot recherché	A droite du mot
A1: donc c'est bien aussi de communiquer là-	<i>dessus</i>	parce que c'est une occasion
E2: il y a un petit bouton un petit bouton vert j'appuie	<i>dessus</i>	
E1: ah oui c'est comme si tu étais toujours	<i>dessus</i>	E2: voilà E1: et que tu l'utilisais toujours
A : on vous a donné aucune aucune information là	<i>dessus</i>	E1 : hum pas du tout A : pas du tout
E1: donc je clique	<i>dessus</i>	E2: ouais
c'est un peu comme la carte étudiante que vous pouvez mettre de l'argent	<i>dessus</i>	
et ensuite c'est selon si tu fais beaucoup de photocopies ou pas tu peux mettre seulement deux euros ou cinq euros	<i>dessus</i>	
la photocopie ça doit être trente centimes donc ne mettez ne mettez pas dix euros	<i>dessus</i>	non
(l'agent passe la carte dans une machine) ah mais il y a rien du tout	<i>dessus</i>	
d'accord et c'est possible de mettre un peu de liquide	<i>dessus</i>	A: oui
je viens de visiter l'appart du	<i>dessus</i>	et du coup ben c'est bon il est pour moi donc je viens me présenter je suis la nouvelle voisine
A: le repas il est pas au-	<i>dessus</i>	de trois euros E: ok
E: bah je pense passer euh bah non mais je l'ai j'ai la j'ai toujours la carte de XX comme ça là-	<i>dessus</i>	A: ah d'accord okay
A: oui c'est que pour vous la carte résident vu qu'il y a la photo qui apparaîtra	<i>dessus</i>	

Figure 51 : Liste des occurrences de « dessus » dans FLEURON

Et voici la synthèse de l'analyse de la stagiaire (les occurrences observées sont entre parenthèses, les paroles de la stagiaire sont en italiques et les commentaires sur la situation entre crochets carrés) :

1. (« dessus ») *c'est quelque chose qui est un peu élevé*
2. (« pas au-dessus ») *je suis pas sûre / [elle fait des gestes : une main sert de repère et l'autre passe en-dessous et au-dessus] / je pense [fait le geste qui correspond à en-dessous]*
3. (« là-dessus ») *sur*
4. (« j'appuie dessus ») *si on touche le bouton le vélo c'est bon*
5. (« un peu de liquide dessus ») *pas compris [elle regarde plus de contexte dans la vidéo] veut acheter / liquide c'est quoi [la formatrice répond : c'est de l'argent] / ah d'accord [la formatrice explique le contexte et le fait de pouvoir mettre de l'argent sur une carte pour payer]*

Cette stagiaire qui a un faible niveau de français explique également qu'elle ne sait bien pas expliquer les significations mais qu'elle comprend le sens des occurrences. Elle explique à la formatrice : « je sais pas comment exprimer mais j'ai compris ».

Lors de l'analyse de *donc*, une stagiaire m'a demandé de l'aider à comprendre le syntagme *vous pouvez*. J'ai d'abord été surprise par cette demande et puis me doutant que la stagiaire l'avait entendu dans une situation particulière, je l'ai questionnée sur cette requête. Dans les formations FLI, il est très fréquent que les stagiaires sollicitent les formateurs en fonction de leurs besoins immédiats. Par exemple, il n'est pas rare qu'ils viennent avec un abonnement de bus ou un chèque à remplir, un planning à comprendre, etc. La recherche de *vous pouvez* était liée au code de la route. La stagiaire souhaitait que je l'aide à comprendre le sens de ce syntagme pour qu'elle puisse ensuite aider sa fille qui était en train de passer le code de la route. Le principal problème rencontré par cette stagiaire était de savoir dans quelles situations *vous pouvez* était une possibilité ou une obligation. Le corpus FLEURON compte 124 occurrences de *vous pouvez*. Après un travail d'observation, assez long (environ 20 minutes) nous avons réussi à identifier ses deux sens et les éléments cotextuels et situationnels qui permettaient de les différencier.

Avec un public FLI peu scolarisé, il est parfois nécessaire d'expliquer plusieurs fois ce que des apprenants-étudiants comprennent plus rapidement. Par exemple, il faut répéter ou reformuler des explications concernant le contenu du site, la définition d'une transcription, et surtout le fonctionnement du concordancier et l'activité d'analyse demandée. Pour certains stagiaires, il est difficile de comprendre ce qu'il faut observer. L'extrait ci-dessous illustre cette difficulté. Cet échange apparaît après cinq minutes d'explications déjà fournies par la formatrice et dure une minute et trente secondes :

Extrait – La formatrice reformule des explications.

- Stagiaire : **je ne comprends pas** [enlève et pose son casque]
 Formatrice : tu comprends pas ce que je te dis
 Stagiaire : non
 Formatrice : alors là tu as tapé je vous en prie dans le concordancier ça /t'a, tu as/ eu une liste d'accord et l'objectif moi l'exercice que je donne là **c'est de regarder à quoi il sert** je vous en prie **comment il est utilisé en français**
 Stagiaire : hum
 Formatrice : c'est ça en fait l'objectif c'est pour ça que tu dois regarder [pointe l'écran]
 Stagiaire : euh politesse
 Formatrice : ouais mh voilà c'est ça en fait que tu dois faire d'accord donc là effectivement ça sert pour la politesse **tu peux regarder quand est-ce qu'il est utilisé dans l'interaction au début à la fin** tu comprends ce que je dis
 Stagiaire : non
 Formatrice : non euh quand est-ce qu'il est utilisé dans l'interaction c'est-à-dire euh est-ce que si **par exemple on commence à parler toutes les deux si je dis je vous en prie est-ce que je vais plutôt le dire au début à la fin de la de notre conversation par exemple tu vois**
 Stagiaire : hum
 Formatrice : hum tu comprends et c'est ça en fait qu'il faut que tu regardes **à quoi il sert quand est-ce qu'il est utilisé etc** d'accord
 Stagiaire : oui
 Formatrice : ouais c'est un peu plus clair si tu veux on peut faire **on peut le faire ensemble** si tu veux
 Stagiaire : oui
 Formatrice : si c'est plus simple pour toi d'accord donc le tout premier

La formatrice reprend, répète et reformule, des explications qu'elle a déjà données à la fois en groupe et en individuel à cette stagiaire. Néanmoins, ce travail est nécessaire dans un premier temps pour que les apprenants saisissent bien ce qu'ils peuvent faire avec un concordancier. Pour chaque apprenant, quel que soit leur niveau de langue et leur niveau de scolarité, il faut une sorte de déclenchement de la compréhension de l'activité avec le concordancier pour qu'il puisse bénéficier des avantages de l'outil. A partir des explications reproduites ci-dessus, la stagiaire et la formatrice analysent ensemble les occurrences de

« je vous en prie » et parviennent (avec beaucoup de gestes et de bribes telles que : « après merci il faut poli aussi et renvoi ») à la compréhension de ses apparitions dans l'interaction.

L'objectif de ce travail et de ces expérimentations est de montrer que les activités métalangagières liées à l'utilisation du concordancier et l'induction des règles de fonctionnement par des apprenants sont possibles avec un public faiblement scolarisé et de faible niveau en français. Les activités d'ASC peuvent être associées, dans le cadre d'une progression pédagogique, à des activités de systématisation et de mise en pratique. Dans le cadre de mes expérimentations, je n'ai pas élaboré ce type d'activités. Je me suis concentrée sur le processus analytique mis en œuvre par les apprenants. Néanmoins, lors d'une séance, j'ai relevé un élément intéressant de mise en pratique spontanée de la part d'une stagiaire. Cette dernière était en train d'analyser le fonctionnement de « désolé ». Elle rencontrait des difficultés pour expliquer ce qu'elle comprenait ou pas des quatre occurrences du corpus. Lors du passage de la formatrice qui lui demande ce qu'elle a compris, elle lui répond : « par exemple désolée je comprends pas c'est ça désolé en français ». Elle attend la confirmation de la part de la formatrice puis elle reprend : « désolée j'arrive pas comprendre le français ». La séance se poursuit et lors d'une coupure de connexion Internet, la même stagiaire dit : « désolée on a pas de chance » face à son ordinateur qui affiche un message d'erreur. Cette utilisation appropriée de l'item recherché est un indice d'acquisition. L'attention portée aux interactions et aux usages de l'item semble fondamentale pour qu'il y ait apprentissage (Schmidt 1990). Ces expérimentations montrent néanmoins que l'accompagnement est plus intense pour les formateurs FLI que pour les enseignants de FLE qui interviennent auprès d'apprenants étudiants.

A l'issue d'une série d'expérimentations avec ces stagiaires de différentes nationalités, une rapide enquête sur l'utilisation de FLEURON a été réalisée. Les échanges qui ont lieu collectivement dans la salle de formation révèlent un enthousiasme pour le site et le concordancier. Quatre stagiaires relèvent les avantages de l'ASC pour comprendre la signification des mots (un stagiaire dit : *c'est un bon site pour sens des mots*). Un autre s'exprime plus précisément sur les sens des mots utilisés fréquemment (*c'est aussi il faut savoir les mots que l'on entend tous les jours mais qu'on sait pas ce que c'est alors qu'on entend tous les jours*). Un autre stagiaire explique que le travail avec le concordancier lui a permis d'apprendre à prendre la parole et à répondre à des questions (*c'est bien de savoir*

comment je dois répondre aux questions posées). Des aspects liés à l'importance de la politesse sont également évoqués (*si tu parles pas bien le français c'est bien de dire s'il vous plait ou je vous en prie*). Enfin, six stagiaires mentionnent les aspects socio-culturels liés à l'apprentissage de la langue en contexte (*c'est pas juste le mot c'est aussi savoir aller au cinéma à la médiathèque*). A la suite de cette appréciation, un stagiaire explique : *les vidéos c'est top*. En creusant un peu cette exclamation, il s'agit à la fois de la multimodalité et de l'authenticité des interactions proposées par le corpus FLEURON. Cette enquête fait également émerger un besoin, les stagiaires aimeraient *une application pour téléphone* pour un « ASC mobile » (Quan 2016). Même si FLEURON fonctionne sur tablette et smartphone, ces personnes migrantes n'ont souvent pas Internet à leur domicile ou dans les foyers dans lesquels elles sont hébergées, mais savent généralement utiliser les applications de leur téléphone.

5.4.2. Avec des adultes francophones natifs en insécurité langagière

L'exploitation de corpus à des fins didactiques, ou l'apprentissage sur corpus, suscitent de nombreuses questions quant à sa mise en œuvre, notamment avec un public de débutants (Boulton 2007), comme je l'ai déjà mentionné. Après avoir réalisé différentes expérimentations auprès d'un public débutant et intermédiaire en FLE, ainsi qu'avec un public FLI, j'ai souhaité expérimenter la méthodologie auprès d'un public d'adultes natifs en insécurité langagière. Cette idée détourne complètement l'utilité première et l'origine de l'apprentissage sur corpus. Néanmoins, les expérimentations précédentes m'ont encouragée à tenter une telle méthodologie avec ce public. Les participants aux expérimentations ont été choisis parce qu'ils suivaient une formation linguistique visant à améliorer leurs compétences scripturales. Ils se trouvent tous en insécurité langagière à l'écrit, c'est-à-dire en situation d'illettrisme, avec des degrés de difficultés différents. A titre d'exemple, voici le profil de cinq des dix stagiaires⁸⁰ qui ont participé à l'expérimentation :

⁸⁰ Les adultes francophones natifs qui suivent des formations linguistiques sont appelés « stagiaires », comme les migrants évoqués dans le point précédent.

Prénom	Sexe	Age	Niveau de scolarisation	Commentaires
Daniel	H	48	Jusque 3 ^{ème} Section d'Éducation Spécialisée	Gens du voyage sédentaire Recherche d'un emploi Ancien bucheron
Jordan	H	25	Jusqu'en 3 ^{ème} classique	Salarié d'une entreprise d'insertion (espaces verts). Souhaite passer le permis moto
Anthony	H	35	Section d'Éducation Spécialisée jusqu'à 18 ans	Recherche un emploi
François	H	53	CAP agriculture/élevage	Salarié d'une entreprise d'insertion (espaces verts). Travailleur handicapé (physique)
Tristan	H	26	CAP hôtellerie/restauration	Salarié d'une entreprise d'insertion (espaces verts). Pris en charge par l'Agence Régionale de Santé (ARS)

Tableau 4 : Présentation de cinq adultes francophones natifs en insécurité langagière

L'objectif premier de cette expérimentation était de savoir si l'introduction des principes de l'ASC allait faciliter l'entrée dans l'écrit et permettre aux stagiaires de s'approprier plus facilement des règles de fonctionnement de la langue dans la mesure où ils allaient les dégager eux-mêmes. Ce test est également lié à l'hypothèse selon laquelle observer de façon répétée la forme d'un mot (ce que permet l'observation verticale des lignes de concordances), c'est-à-dire la façon dont il est orthographié, va faciliter son acquisition. Ce fut un autre défi que j'ai souhaité relever tant les résultats avec des apprenants allophones étaient encourageants. Une grosse partie du défi réside dans le pari que ces personnes en insécurité langagière puissent réussir à endosser le rôle d'apprenti-linguistique ou de Sherlock Holmes.

5.4.2.1. Une transcription pour tester le degré d'illettrisme

Afin de me rendre compte du niveau des stagiaires, j'ai commencé les séances expérimentales par un exercice de transcription. L'activité pédagogique de transcription est considérée comme une activité de découverte de la langue. Elle est utilisée dans le cadre de l'enseignement et de l'apprentissage du FLE mais aussi d'autres langues (Lebre-Peytard *et al.* 1977, Conrad 2004, Lynch 2007). J'ai utilisé une vidéo de FLEURON dans laquelle une

personne présente quelques rayons de la médiathèque. Le passage à transcrire concerne les informations relatives au rayon presse (ressource « Rayon presse de la médiathèque ») et dure onze secondes. Voici sa transcription :

alors ici c'est l'espace presse vous pouvez trouver des revues des magazines au tout début ce sont les journaux donc l'actualité donc il y a l'Est Républicain notre journal régional

Les participants ont pu visionner l'extrait filmé et l'écouter plusieurs fois, y compris par petits extraits.

Voici cinq photos représentatives du niveau des dix participants à toutes les expérimentations avec des adultes francophones natifs :

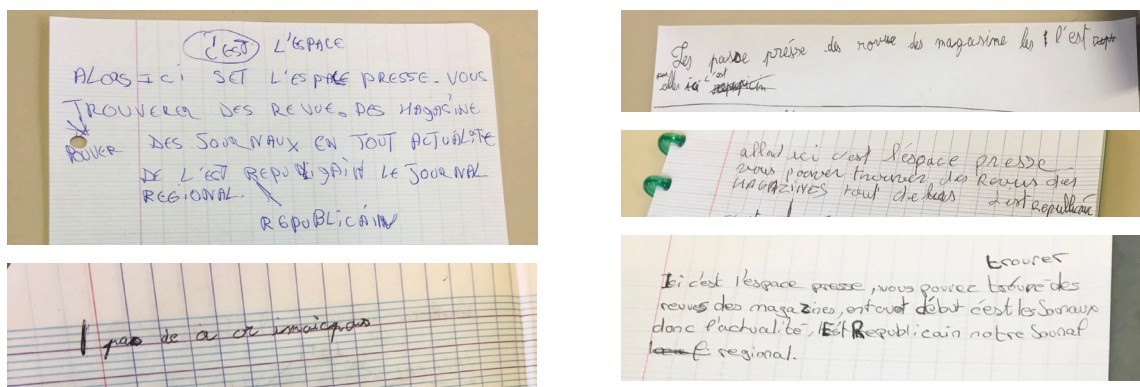


Figure 52 : Photos d'un exercice de transcription avec des personnes en insécurité langagière

Ce premier exercice s'est révélé ludique pour les participants. Ils ont visionné le film puis ils ont tenté de transcrire les onze secondes proposées. Si les participants ont produit des textes différents, il est aisé de constater qu'ils ont tous des difficultés pour écrire. Néanmoins, au-delà de leurs difficultés, avec la formatrice, nous avons pu constater que faire un exercice de transcription de français parlé avec ce type de public permet de faciliter l'entrée dans l'écrit. Le travail sur la transposition phonie-graphie est facilité et n'est pas assimilé à une dictée traditionnelle. Je ne vais pas détailler davantage cet exercice mais il aura été le point de départ de ces expérimentations consistant, finalement, à détourner FLEURON.

5.4.2.2. Une séquence pédagogique à partir de corpus

Dans un premier temps, j'ai souhaité expérimenter l'utilisation de plusieurs corpus à des fins d'apprentissage du français écrit (ou de remise à niveau) pour ces francophones natifs. Ainsi, avant d'opter pour un usage exclusif de FLEURON, j'ai utilisé d'autres corpus, écrits et oraux. Je vais donc d'abord décrire la mise en œuvre d'une séquence pédagogique à partir de plusieurs corpus. Cette séquence a remporté un succès mitigé auprès des stagiaires. Je pense néanmoins qu'il est intéressant d'en rendre compte dans la mesure où elle a permis des ajustements didactiques et méthodologiques dans un domaine jamais expérimenté.

Pour mener à bien ces expérimentations consistant à exploiter des corpus à des fins didactiques, j'ai souhaité utiliser des corpus dont l'interface était la moins impressionnante pour un public qui ne maîtrise pas les outils numériques (excepté l'usage de leur téléphone), même si je savais que j'allais moi-même manipuler les corpus⁸¹. Ainsi, avant d'opter pour l'utilisation exclusive du corpus FLEURON, j'ai utilisé deux autres corpus : l'Est Républicain, interrogeable *via* le site ScienQuest⁸² et OFROM qui possède son propre concordancier. La première expérimentation n'a pas été un immense succès. Je souhaite tout de même en rendre compte, comme je l'ai déjà précisé, dans la mesure où elle présente une première tentative d'exploitation de corpus à des fins didactiques. J'ai choisi d'adapter une séquence pédagogique que j'avais déjà réalisée avec un public d'étudiants en FLE et qui avait bien fonctionné, tant du point de vue de la mise en place de la méthodologie de l'ASC que de l'apprentissage de la langue. J'ai donc adapté et simplifié une séquence d'observation de la négation en corpus (André à paraître 2020). Une formatrice en charge de ce public en insécurité langagière m'avait également fait part du constat de la non maîtrise de la négation en *ne... pas* de la part de ses stagiaires. Forte de ce constat et de mon expérience passée, j'ai planifié une séquence pédagogique sur la négation. Ce choix peut paraître surprenant dans la mesure où cette difficulté peut paraître marginale pour ce public, mais il permettait une première séance d'expérimentation.

⁸¹ Au-delà de la non maîtrise de l'outil informatique de ce public, les organismes de formation qui assurent les séances ne possèdent que très rarement des salles contenant des ordinateurs pour les stagiaires. Le GRETA mentionné précédemment est un service de l'Éducation nationale, ce qui explique que les locaux soient bien équipés.

⁸² Disponible à l'adresse suivante : <https://corpora.aiakide.net/scientext20/>.

Tout d’abord, j’ai projeté le résultat de la recherche de « pas » dans le corpus de l’Est Républicain *via* le concordancier du site ScienQuest. Voici une capture d’écran :

1	Han par des candélabres en acier galvanisé noir n' a	pas	été prévu dans la délibération du 11 septembre 98 relative
2	Formule chère pour le cochon grillé Il n' en manquait	pas	« Après l' effort , le réconfort » , avait
3	5 décembre . Jacques Dodo , maire , n' a	pas	caché sa satisfaction de se trouver en présence d' une
4	une place de la mairie pourtant baignée de soleil ,	pas	la moindre ombre d' un bénévole , aussi bien du
5	un flash info déposé dans leur boîte aux lettres ,	pas	plus qu' un représentant du monde associatif prévenu lui par
6	bannière du « Printemps de l' environnement » n' a	pas	connu , à Fains-Véel , le succès que la municipalité
7	mettre un_peu de beurre dans ses épinards , n' hésite	pas	à sacrifier la sacro- sainte grasse matinée dominicale pour être
8	et l' Etat . « Bien_qu' il ne s' agisse	pas	d' un enterrement de première classe » , les agriculteurs
9	lors_de son séjour hivernal , le cirque Maximum n' a	pas	Depuis 4 ans , il sillonne les routes de France
10	bagages vers d'_autres cieux , la ville n' en fini	pas	Vendredi , le maximum sera une_fois encore réalisé pour séduire
11	une affinité plus marquée pour la pâtisserie , n' est	pas	loin_de regagner son fournil « parce_qu' il_y_a vraiment quelque_chose à
12	ne fait qu' une bouchée . « Ce n' est	pas	ça forcément qui nous ramène plus de clientèle , surtout
13	» , commente ainsi Nicole . « Je ne suis	pas	une très grande consommatrice de pain , mais je reste
14	l' accoutumée et sa réputation de sortie familiale n' a	pas	Le rallye a été préparé par l' équipe vainqueur l'
15	, boîtes , bouteilles et autres immondices très humains et	pas	naturels du tout , chaque équipe devait composer une statue
16	, Magali , l' épouse de Sébastien , ne cache	pas	sa surprise , car dans chacune des vitrines de la

Figure 53 : Extrait de la liste des occurrences de « pas » dans L’Est Républicain

Comme pour un public FLI, avec à un public natif en insécurité langagière, peu scolarisé et peu habitué à manipuler des textes écrits, il faut expliquer le principe du concordancier. Cette première liste devait permettre aux stagiaires de se rendre compte de la présence de *ne* pour chacune des occurrences de *pas* (en excluant, avec des explications, les occurrences 4, 5 et 15 de la liste). Le constat a été fait, avec mon aide, et après de multiples incitations à l’observation du contexte gauche. Voici quelques extraits de mon discours :

regardez bien regardez bien à gauche (en montrant le tableau) est-ce que vous repérez quelque chose un mot ou quelque chose qui est tout le temps là
 (...)
 regardez à gauche est-ce que vous voyez une forme qui revient tout le temps un peu avant pas une forme ou plusieurs formes qui se ressemblent ou même des lettres est-ce que vous réussissez à repérer des lettres à gauche qui sont présentes dans toutes les lignes
 (...)
 regardez avant pas même pas trop loin quand même il y a pas quelque chose qui revient + un mot + des lettres

A l’instar de l’explication du fonctionnement du concordancier, j’ai dû répéter de nombreuses fois les consignes d’observation en doutant encore aujourd’hui de leur totale compréhension. Cependant, au bout de quelques allers-retours entre mes explications et les lignes de concordances, la présence de la première partie de la négation, prenant la

forme de *ne* ou *n'*, a été remarquée. Un petit commentaire a été fait sur la différence entre les deux formes.

Ensuite, j'ai proposé l'observation d'un corpus oral pour que les stagiaires constatent des différences. J'ai donc projeté les résultats de la recherche de *pas* avec le concordancier du corpus OFROM, voici ce que les stagiaires pouvaient observer :

Locuteur	Texte
unifr11-esb	des choses où on n'arrive pas quoi
unifr11-esb	personne n'est pas bon en quelque chose il suffit de travailler mais
unifr11-esb	y en a qui sont pas f/ pas doués en math et d'autres en orthographe
unifr11-esb	oui je pense aussi mais après euh y a aussi euh des gens qui sont pas forts en orthographe et pis ça ben
unifr11-esb	ils comprennent pas beaucoup et pis euh
unifr11-esb	ouais le mot il leur vient pas directement juste alors c'est ça le gros problème en fait
unifr11-esb	euh oui assez quand même y a pas beaucoup de gens qui aiment l'ortographe euh
unifr11-esb	ça les gens ils vont pas non plus aller chercher dans un dictionnaire donc y a quand même des erreurs mais
unifr11-esa	ouais je pense ça va euh je sais pas hein mais je pense
unifr11-esa	des natels on on leur parle et puis ils nous donnent tout de suite ce qu'on veut je sais pas des trucs euh un peu
unifr11-esa	j'arrive pas trop à m'imaginer futur parce que
unifr11-esa	ouais je pense pis en plus ils avaient pas la DS la Gameboy euh
unifr11-esa	ouais euh on a ils ont pas des très bonnes notes non plus c'est
unifr11-esa	ceux qui ont MSN je trouve ouais ça influence euh pas mal
unifr11-esa	enfin le pas tout juste mais j'essaie de faire des efforts
unifr11-esa	non pas jus/ pas tout seuls comme ça
unifr11-esa	euh là je vais pas non plus aller intervenir chez tous les enfants qui vont sur internet mais

Figure 54 : Extrait de la liste des occurrences de « *pas* » dans OFROM

Avec cette liste, l'objectif était que les stagiaires constatent qu'à l'oral, la négation pouvait se réaliser sans *ne*, mais pas systématiquement. Cet exercice a été un échec. Les stagiaires n'ont pas compris où je voulais en venir et n'ont pas compris pas la différence, malgré les nombreux allers-retours entre les deux corpus et toutes les précisions sur le genre de discours (dans OFROM, il y a des narrations, des conférences et des discussions et ces métadonnées sont facilement accessibles), les locuteurs, les objectifs de la communication, etc. Ils ne perçoivent pas la différence entre *ne... pas* et *pas*. Je suis passée par une simulation pour tenter de faire comprendre au moins la différence entre l'oral et l'écrit, c'est-à-dire entre la possibilité de ne pas produire la négation complète et l'obligation de la

produire. Cette simulation a consisté, dans un premier temps, à écrire un texto à un employeur (tous les stagiaires ayant participé à l'expérimentation étaient en emploi ou en formation) pour prévenir d'une absence. Dans un second temps, ils devaient laisser un message vocal par téléphone à un ami proche qui venait de les inviter à manger alors qu'ils n'avaient pas faim. Cet exercice a été plutôt bien réussi. Le *ne* a été utilisé avec l'employeur (mais si des stratégies d'évitement ont parfois été utilisées) mais pas avec l'ami (quasiment tous ont produit : *j'ai pas faim*). La variation a été perçue et ils ont compris ce que je souhaitais leur montrer avec les listes de concordances. Ils ont eux-mêmes exprimé la variation écrit/oral en avouant qu'ils observaient davantage de la variation en fonction du genre de discours et de l'interlocuteur. Ce qui fut pour moi un excellent résultat. Cependant, lorsque je leur ai montré de nouveau les listes de concordances, ils ne saisissaient plus la différence. Je souhaitais aborder la variation à l'oral avec OFROM mais j'ai bien senti qu'ils allaient plutôt être perturbés par la démonstration. En outre, ce public a une capacité de concentration assez limitée. Cet exercice leur a demandé trop d'efforts et j'ai dû l'abandonner avant d'obtenir des résultats satisfaisants. Cette première expérimentation a tout de même été utile et a orienté les suivantes.

5.4.2.3. Le choix de FLEURON

Après la transcription pour saisir le niveau de maîtrise de l'écrit des individus, après l'expérimentation sur la négation, j'ai décidé de détourner le dispositif FLEURON avec ce public. L'expérimentation plutôt réussie de la transcription à partir des vidéos de FLEURON m'a poussée à tenter de nouvelles expérimentations mais exclusivement avec FLEURON. Pour rappel, l'objectif est d'observer si les stagiaires – des adultes francophones natifs en insécurité langagière – sont capables de se lancer dans un processus d'analyse distributionnelle malgré leurs problèmes de maîtrise de la langue.

◆ Les consignes

La présentation de FLEURON, de ses principes et de ces outils a été réalisée dès la première séance puis à plusieurs reprises ensuite. Voici des extraits de transcription d'explications et de consignes :

Exemple 1 : VA réexplique le fonctionnement de FLEURON

- VA alors vous vous souvenez un petit peu de mes jeux
 S1 euh
 VA mais c'est pas grave je vais vous réexpliquer tout ça (...) l'idée c'est toujours on va essayer de regarder donc ça c'est un site Internet euh qui est je sais pas si vous vous souvenez peut-être que vous vous aviez pas eu la présentation c'est un site Internet qui est fait pour des étudiants qui veulent voir un petit peu comment ça se passe quand on arrive dans une université en France mais en fait il y a plein de choses sur ce site qu'on peut utiliser différemment et en fait sur ce site il y a essentiellement des vidéos et il y a par exemple même des vidéos de personnes qui vont acheter qui vont chercher euh un sandwich sur le campus par exemple je vous montre une petite vidéo et ici vous voyez on a un jeune homme qui va acheter un sandwich (diffusion de la vidéo) alors **ce qui va nous intéresser nous ici l'avantage de ces petites vidéos c'est que vous voyez elles sont complètement sous-titrées**
 S2 hum
 VA donc en fait nous **ce qui va nous intéresser c'est de voir comment on écrit ce qui est dit**

Exemple 2 : Une formatrice (EJ) en stage qui participe aux expérimentations donne des explications.

- EJ on a vu la semaine dernière qu'on avait plusieurs mots comme ça qu'on prononce pareil mais à l'écrit ils sont différents souvent c'est ça qui pose problème c'est quand on va les écrire les reconnaître et **choisir le bon avec la bonne orthographe**

Exemple 3 : Des explications sur la transcription.

- VA avec FLEURON FLEURON c'est le nom du site Internet **tout ce qui est dit est aussi écrit donc on va essayer de voir comment ça s'écrit** justement on va travailler nous l'écrit ensemble à partir de ce genre de vidéo

Exemple 4 : Des explications du concordancier.

- VA alors pourquoi j'utilise ce truc là ce site Internet là parce qu'il y a un truc magique dans ce site Internet si je reviens au début on a ici un petit outil ici bon ça s'appelle un concordancier peu importe c'est comme un moteur de recherche quand vous tapez quelque chose dans google par exemple **et il va nous chercher tous les c'est comme ça écrit comme ça** et il va aller nous chercher comme tout est vous avez vu on écrit tout ce qui est prononcé donc ça va chercher dans tout ce qui est prononcé et nous alors on va regarder + **pour comprendre un petit peu comment ça fonctionne on va regarder à partir de celui du haut là donc on va regarder plusieurs c'est et on va essayer comment ça fonctionne c'est (...) si vous vous souvenez on peut aller voir dans la vidéo au moment où elle prononce ça**

Les activités métalangagières réalisées lors de l'utilisation du concordancier nécessitent des explications simples et répétées pour que les participants comprennent ce qu'il leur est demandé de faire. Les explications sont orientées par l'objectif de l'exercice. Il s'agit de faire comprendre aux stagiaires que les paroles sont transcrites et que l'observation des transcriptions va leur permettre de voir comment les mots sont orthographiés.

◆ Les analyses

Les activités proposées à partir du concordancier de FLEURON ont été choisies selon les difficultés habituellement rencontrées avec ce type de public et en collaboration avec la formatrice en charge d'un groupe-expérimentateur. La première séance a été consacrée à la différence entre *c'est* et *sait*. Ce couple fait partie des nombreux homophones qui posent des problèmes à l'écrit (pour les adultes en insécurité langagière comme pour les enfants en cours d'apprentissage de l'écriture, de nombreux travaux sont consacrés à ce sujet). J'ai choisi délibérément des mots éloignés linguistiquement pour faciliter la réussite de l'exercice, le but n'étant pas de piéger les stagiaires déjà en grandes difficultés.

La première étape de ce travail sur corpus a été de faire écouter dix extraits d'interaction aux stagiaires, comprenant l'une ou l'autre des occurrences (*c'est* ou *sait*), en leur expliquant pour chaque énoncé le contexte nécessaire à la compréhension. Les participants devaient cocher dans un tableau (dessiné préalablement au tableau) s'ils pensaient qu'il s'agissait de *c'est* ou de *sait*. La majorité de leurs réponses a été choisie au hasard, ils ont avoué qu'ils voyaient parfois qu'il y avait une différence mais ne savaient pas laquelle. Ce premier exercice a permis de valider le choix des occurrences à analyser avec eux. Ce type d'exercice a été réalisé lors de la première séance d'expérimentation de FLEURON avec ce public. Il a permis d'entrer en contact avec les participants en leur proposant un exercice semi-traditionnel ou qui pouvait ressembler à des exercices connus et déjà réalisés avant de les lancer dans l'analyse distributionnelle d'une liste d'occurrences affichée grâce à un concordancier. Cependant, cet exercice n'a pas été renouvelé. J'ai introduit ensuite l'ASC.

La seconde étape a donc été l'interrogation du corpus multimédia de FLEURON grâce à son concordancier. Voici des extraits des résultats du concordancier de FLEURON pour les recherches de *c'est* et *sait* :

A gauche du mot	Mot recherché	A droite du mot
L2 :	<i>c'est</i>	polyuniversitaire européen
L2 : qui font le même travail que moi en fait	<i>c'est</i>	c'est au niveau de c'est une politique des sites
L2 : qui font le même travail que moi en fait c'est	<i>c'est</i>	au niveau de c'est une politique des sites
L2 : qui font le même travail que moi en fait c'est c'est	<i>c'est</i>	une politique des sites
au niveau de		
L2 : euh lorraine donc	<i>c'est</i>	c'est un peu ça
L2 : euh lorraine donc c'est	<i>c'est</i>	un peu ça
L2 : et donc	<i>c'est</i>	uniquement pour les étudiants et nous on est là
L2 :	<i>c'est</i>	-à-dire tout ce qui est euh
A : mais	<i>c'est</i>	bien vous avez déjà pris de l'avance
A : j'ai je vois votre contrat d'études	<i>c'est</i>	bien
A :	<i>c'est</i>	madame XX Dalida
A : donc	<i>c'est</i>	bon

Figure 55 : Extrait de la liste des occurrences de « *c'est* » dans FLEURON

A gauche du mot	Mot recherché	A droite du mot
E2 : étudier on	sait	pas parce qu'elle peut rester que 3 mois E : étudier aussi
A : garde-le on	sait	jamais hein E : oui bah je le garde dans mes affaires
A : euh ouais non mais on est au courant du problème pour vous on	sait	que bon
A : vous inquiétez pas bon on	sait	que de toute façon vous devez avoir votre année d'acquise mais euh voilà
A1: en période d'examens et cetera on	sait	que les étudiants sont pas trop disponibles
A1: c'est bon à savoir oui et puis ça se	sait	pas forcément
A : j'essaierai de ne pas vous téléphoner euh mais on	sait	jamais
A: le C deux I c'est connu on le	sait	le contenu hein de de C deux I
A: donc quand vous dites que vous avez le C2i on	sait	

Figure 56 : Extrait de la liste des occurrences de « sait » dans FLEURON

Tout comme avec les adultes migrants faiblement scolarisés, le travail d'analyse doit être accompagné, médié par un formateur ou un enseignant. Cependant, malgré l'absence de métalangage, de conscience grammaticale, et malgré les problèmes de maîtrise de la langue de ce public, il parvient à analyser les occurrences et à induire des règles de fonctionnement. Voici neuf extraits de la transcription d'une séance collective qui illustrent le processus d'analyse mis en œuvre :

1. *c'est avec une suite ou plutôt ça sert à terminer + on le trouve à la fin comme avec c'est magnifique*
2. *sait + c'est le verbe savoir dans on sait jamais*
3. *c'est encore ouvert + ça veut dire il est encore ouvert le bureau*
4. *c'est on peut le remplacer par cela est*
5. *on le sait pas forcément + ça veut dire on n'est pas au courant*
6. *c'est on peut le remplacer par ceci est*
7. *c'est bon on dit ça quand on est invité à manger on peut le remplacer par c'était bon*
8. *pour faire la différence on peut changer de temps + mettre au passé*
9. *je sais si c'est bien euh c'est compliqué comment faire il y a un autre exemple je sais que c'est difficile euh je sais c'est je suis au courant que le travail est difficile avec être ça marche mais là il y a les deux alors faut plus de temps pour réfléchir*

Ces neuf étapes sont extraites d'un long processus d'analyse que j'ai accompagné en pointant certains éléments, en faisant des allers-retours dans les ressources afin de montrer un cotexte plus large et de donner des informations situationnelles, en posant des questions, etc. A la fin de la séance, les participant ont fait une synthèse :

en fait quand on peut dire c'était on met un c et quand on sait quelque chose quand ça veut dire qu'on est au courant on met sait avec s a i t (mot épelé)

Cette conclusion montre que les participants ont induit leurs propres règles, celles qui leur permettent d'écrire correctement les mots. A la séance suivante, à laquelle je n'ai pas assisté, la formatrice a proposé un exercice de systématisation⁸³. Les stagiaires devaient écrire *c'est* ou *sait* dans des énoncés (authentiques) à trous. Ils ont réinvesti les règles induites en remplaçant les mots recherchés par : *cela est*, *c'était* et *on est au courant*. Les résultats ont été positifs, à la surprise de tous, des stagiaires et de la formatrice.

Forte de ces premières expérimentations, d'autres ont suivi, sans activité traditionnelle préalable, en travaillant directement avec les lignes de concordances. Je vais rendre compte d'un autre exercice réalisé quelques semaines après l'analyse de *c'est* et *sait* (entre temps d'autres analyses du même type ont été effectuées par les stagiaires) mais qui s'inscrit dans la continuité. J'ai proposé d'analyser la différence entre *c'est* et *s'est*. Là encore, je projette les résultats des recherches. Voici celui de *s'est* :

A gauche du mot	Mot recherché	A droite du mot
E1: tu me disais... la dernière qu'on	s'est	vues
E : je ne sais pas ce qu'il	s'est	passé cette année
E1: mais maintenant je crois avec heu la création de l'université de Lorraine ça	s'est	étendu
E1: alors du coup Mélissa est-ce que tu peux me m'expliquer comment ça	s'est	passé ta recherche de directeur

Figure 57 : Liste des occurrences de « s'est » dans FLEURON

Tout d'abord, il est intéressant de mentionner que l'un des participants a rappelé la règle qui consiste à remplacer *c'est* par *c'était* pour savoir comment écrire le son entendu. Même si je ne m'intéresse pas ici à l'efficacité de l'ASC sur le long terme, je peux constater que les règles de fonctionnement induites sont retenues par certains participants. Ensuite, lors de cette analyse comparative, un nouveau fonctionnement a été identifié pour *c'est*, celui de présentatif qui n'avait pas été analysé aussi précisément lors de la séance précédente. Par ailleurs, l'analyse de *s'est* est facilitée par les seules quatre occurrences du corpus. Un vrai travail d'analyse peut être lancé sans que les participants se sentent noyés par les données.

⁸³ L'apprentissage sur corpus étant une méthodologie particulière, elle n'exclut pas la mise en œuvre d'activités de systématisation. Ces dernières viennent compléter les activités de découverte que seraient les activités d'analyse réalisées avec le concordancier. Cependant, j'ai fait le choix de laisser ces activités à la discrétion des formateurs qui m'ont accueillie pour les séances d'expérimentation sur corpus.

Dans ces situations de formation, la sélection de l'input présentée aux stagiaires peut être conseillée. La séance de travail dure une heure avec de nombreux allers-retours entre les deux listes d'occurrences (même si 45 minutes au total sont consacrées à l'analyse de *s'est*) et vers les ressources (et leur contextualisation). Au début de l'analyse de *s'est*, les participants avouent ne pas connaître cette forme et ne l'avoir jamais écrite. Néanmoins, après les différentes analyses réalisées des occurrences en contexte, les stagiaires parviennent à induire les règles suivantes à propos de *s'est* :

- c'est le verbe être
- après il y a aussi un verbe ou des actions
- le s apostrophe s'écrit s parce que c'est comme dans *se voir*
- au passé on peut dire *on s'est vus*
- devant il y a *on/il/ça* il y a un sujet ou quelque chose avant
- *se voir* c'est *voir* mais ensemble mutuellement
- c'est faire une action au passé

Ensuite, les stagiaires réalisent spontanément une activité de mise en pratique pour montrer qu'ils ont compris, notamment avec *se faire quelque chose*. Finalement, les stagiaires concluent en s'exprimant sur les avantages des listes de concordances. L'un d'entre eux explique : *c'est facile quand on regarde les exemples on peut trouver des astuces pour écrire*. Un autre stagiaire reprend en donnant les exemples suivants : *on se fait une soirée* et *on s'est fait une soirée* et en expliquant que le second exemple signifie que la soirée est déjà passée. Cette même construction est reprise par un autre stagiaire avec *se faire des crêpes*. Ainsi, les stagiaires trouvent leurs astuces et les résument en produisant la formule suivante :

quand on a *on/elle* ou une personne + la possibilité de mettre au passé = *s'est*

Les différents exemples que nous avons analysés avec les stagiaires ont ensuite conduit à l'analyse de la différence entre *on* et *ont*. Enfin, les analyses distributionnelles de *on* et *ont* nous ont amenés à revenir sur la distinction entre *peux/peut/peu* travaillée lors d'une séance précédente que je vais résumer ci-dessous.

Je terminerai donc par la présentation d'une expérimentation plus simple destinée à maîtriser les différences entre *peux*, *peut* et *peu*. Si l'on regarde les résultats du

concordancier pour les apparitions de *peux*, *peut* et *peu*, le cotexte gauche semble facilement identifiable :

E : sinon je	peux	faire un cours d'italien ou d'espagnol
A : c'est ce que vous avez déjà apporté donc on peut mettre euh E : en fait est-ce que je	peux	le faire ici euh remplir ce document
E1: et le et le en bas tu	peux	garder ton pantalon E2: oui E1: oui d'accord
A : je	peux	écrire en minuscule votre adresse électronique
A : hein si je	peux	je l'envoie en fin d'après-midi
E : vous voulez je	peux	l'écrire
E : et demain je	peux	je viens A : vous pouvez venir le chercher oui si vous voulez E : bien
A : euh parce que là c'est vrai qu'effectivement là c'est hum ça va poser problème je	peux	pas vous faire euh
E : rien A : je	peux	pas vous faire l'inscription pour la carte parce que j'ai vraiment besoin de ce document
A : euh bah je	peux	

Figure 58 : Extrait de la liste des occurrences de « *peux* » dans FLEURON

L2 : alors il est important de savoir qu'on	peut	avoir des pôles universitaires
A : c'est ce que vous avez déjà apporté donc on	peut	mettre euh E : en fait est-ce que je peux le faire ici euh remplir ce document
E : hum hum A : parfois on	peut	vous demander hein on peut faire un petit contrôle et demander
E : hum hum A : parfois on peut vous demander hein on	peut	faire un petit contrôle et demander
E2: ouais ça	peut	ça ça prend toute la journée
A : pour euh dire que vous êtes bien arrivé en France	peut	-être E : oui oui
A : euh entre temps vous aurez	peut	-être vu monsieur Smith
A : est-ce qu'on	peut	prendre rendez-vous jeudi 14 heures?
A : vos parents peuvent	peut	-être appeler
A : hein ça	peut	se faire rapidement je pense
E : oui A : ça	peut	se faire rapidement je pense hein
A : bah vous avez dû avoir	peut	-être un numéro d'étudiant non?
E : ouais donc c'est un visa d'études A : oui une fois que la personne a le visa elle	peut	venir s'inscrire
E : l'année prochaine A : ah oui alors à ce moment-là il faudrait	peut	-être revenir euh
E : elle	peut	rentr- elle elle va rentrer en France elle peut rentrer en France parce qu'elle a un passeport X A : hum
E : elle peut rentr- elle elle va rentrer en France elle	peut	rentrer en France parce qu'elle a un passeport X A : hum

Figure 59 : Extrait de la liste des occurrences de « *peut* » dans FLEURON

L1 : donc oui est-ce que vous pourriez déjà m'expliquer un petit	peu	quelle est votre fonction ici au pôle européen
L2 : euh lorraine donc c'est c'est un	peu	ça
A : toujours on vous laisse un petit	peu	de temps hein pour changer les cours hein E : hum hum super oui
A : euh vous pouvez oui si vous avez un petit	peu	de temps vous pouvez vous mettre euh
E1: d'accord oui pour faire les radios E2: oui c'était j'étais un	peu	mal à l'aise parce que
A : c'est bien donc attendez un petit	peu	et puis euh
A : mais si vous partez un petit	peu	avant je je vous ferai une attestation
A : donc Sarah donc je vais vous expliquer un petit	peu	la procédure hein
E : en français euh je connais un	peu	la langue parlée mais je sais pas très bien écrire en français en fait
A : parce que là c'est un	peu	tôt hein faut faire faut faire E : oui oui A : la demande quelques mois avant E : hum
A : vous pouvez attendre encore un petit	peu	euh E : hum hum A : il y a pas de problèmes euh
A : euh et on vous laisse un petit	peu	de temps pour faire le choix des cours
A : mais tu as encore un petit	peu	de temps hein pour choisir les cours euh
A : tu as encore un petit	peu	de temps on vous laisse le temps

Figure 60 : Extrait de la liste des occurrences de « *peu* » dans FLEURON

Les stagiaires identifient rapidement que :

- *peux* est immédiatement précédé de *je* ou *tu*
- *peut* est immédiatement précédé de *on*, *elle* ou *ça*
- *peu* est immédiatement précédé de *un* ou de *un petit*
- *peu* se trouve dans l'expression *un petit peu de temps*, *un peu tôt*, *un peu mal*
- *peu* s'oppose à *beaucoup*

Une fois encore, les stagiaires induisent les règles et les astuces qui leur permettent d'écrire correctement le mot. La récurrence et la régularité semblent favoriser l'induction des règles voire l'acquisition. L'objectif est le même que pour une langue étrangère, l'idée est d'exposer les individus à la langue réelle. Les données auxquelles les stagiaires sont exposés sont analysées, caractérisées et généralisées, selon les préconisations de Johns (1991, 1997). Cette approche paraît plus pertinente qu'une approche explicite au cours de laquelle un formateur donne les règles de fonctionnement. Le public cible de ces expérimentations a été confronté à des règles à de nombreuses reprises, à l'école, en formation pour adultes, en formation d'insertion. Malgré ces règles énoncées et répétées, ces personnes restent en situation d'illettrisme, avec des gros problèmes pour orthographier les mots.

Se pose de nouveau dans ces expérimentations avec un public peu scolarisé, la question de la grammaire. Adami (2018) tente de répondre à cette question, en minimisant l'utilité de la grammaire explicite, avec un public suivant des formations d'intégration et d'insertion sociale, professionnelle et citoyenne. Avec nos expérimentations et l'utilisation détournée de FLEURON, la grammaire est bien abordée mais de façon tout à fait particulière, et surtout de façon implicite. En tentant de rapprocher lexique et grammaire en utilisant la démarche de l'ASC dans le cadre de la maîtrise de la langue écrite, il est possible d'envisager des perspectives intéressantes. Le problème de l'homophonie que j'ai abordé avec des adultes en situation d'illettrisme a également été abordé par Leray et Tyne (2016) avec des enfants entre sept et douze ans du niveau CE2 à la sixième qui ont été soumis à l'ASC. Si le processus analytique et métalangagier mis en œuvre par les élèves n'est pas décrit, les résultats chiffrés de leur enquête sont encourageants. Ils soulignent par ailleurs que : « La question de l'exposition aux formes écrites comme moyen d'avancer dans la maîtrise de l'orthographe semble suffisamment pertinente pour que l'on s'y intéresse par le biais de l'ASC » (*Ibid.* : 134). Certains de leurs résultats les amènent également à conclure que « l'ASC constituerait ici une base plus solide pour que l'enfant puisse ensuite progresser

par lui-même » (*Ibid.* : 141). Je peux m'aventurer à faire le même constat avec un public adulte en situation d'illettrisme. Il s'agit de donner des habitudes analytiques métalangagières aux stagiaires et de leur donner quelques clés pour qu'ils identifient eux-mêmes les astuces qui leur seront utiles dans leurs futures activités de rédaction.

Conclusion du chapitre

Les différentes expérimentations visant à mettre un concordancier au service de l'enseignement et de l'apprentissage du FLE montrent que l'ASC peut prendre différentes formes. Ces dernières varient en fonction des apprenants, des enseignants, des objectifs et des conditions d'apprentissage, des habitudes didactiques. La variable apprenants peut être précisée par : le niveau en langue cible, le style d'apprentissage ou encore le degré d'autonomie. Tous ces éléments sont à prendre en compte pour introduire l'ASC. Ils influencent à la fois l'amorçage de l'ASC et son accompagnement. En ce qui concerne l'amorçage, le rôle de l'enseignant peut varier et, concrètement, il peut :

- proposer la recherche d'un item de son choix
- proposer la recherche d'un item dans le cadre d'une progression didactique
- proposer une liste d'items à rechercher
- proposer à l'apprenant de répondre à une question qu'il aurait posée de trouver lui-même la réponse en effectuant une recherche
- laisser libre l'apprenant de choisir les items à rechercher

Ces attitudes ne sont exclusives, elles s'adaptent aux éléments situationnels mentionnés plus haut. Elles peuvent également être adoptées de façon progressive, de la plus contraignante à la plus libre. En outre, à chacune de ces attitudes peut être associé un type d'accompagnement. Après que l'explicitation des consignes de recherche est formulée (en fonction des attitudes précédentes), c'est-à-dire dans quel type de recherche (libre ou contrainte) l'apprenant se lance, l'enseignant peut de nouveau adopter différentes attitudes (Pérez-Paredes *et al.* 2011). Il peut :

- accompagner l'apprenant pas à pas dans le processus de l'ASC, c'est-à-dire qu'il accompagne l'observation, la catégorisation et la généralisation (Johns 1991)
- accompagner le début du processus pour s'assurer qu'il est compris
- ne pas intervenir et laisser l'apprenant accomplir les différentes activités seul

Ces attitudes ne sont pas non plus exclusives et s'adaptent également aux éléments situationnels. Elles sont également possiblement mises en œuvre dans des configurations différentes dans lesquelles les apprenants sont :

- en groupe : le travail collectif permet de montrer à tous les apprenants d'une classe le fonctionnement du concordancier et les principes de l'ASC. Il met les apprenants en situation aidée (par l'enseignant et les autres apprenants) et permet également de faire émerger davantage de règles de fonctionnement.
- en binôme : le travail à deux permet une collaboration plus active qu'en groupe. Cela permet aux apprenants de partager leurs découvertes (Aston 1997) et de s'entraider (Landure 2011). Un binôme pousse les apprenants à verbaliser leurs analyses.
- seuls : les apprenants suivent leur propre rythme, font leurs propres choix et surmonte leurs propres difficultés.

Toutes ces possibilités sont combinables. Elles le sont également avec les attitudes possibles des enseignants, que ce soit pour amorcer l'ASC ou pour l'accompagner.

Ainsi, la mise en œuvre de l'ASC est modulable et adaptable. Certains apprenants préfèrent faire leurs propres recherches, même si cette étape est parfois compliquée (Bernardini 2000), et d'autres préfèrent être guidés voire contraints par des propositions de recherches et des consignes précises (Chang *et al.* 2003 ; Schmied 2006). Certains apprenants se réjouissent à l'idée de pouvoir prendre des initiatives et de travailler seuls. La recherche de leurs propres items se fait alors en fonction des objectifs qu'ils se sont fixés. En outre, plus les apprenants réalisent seuls leurs activités, plus il y a de chances pour qu'ils continuent l'ASC en dehors de la salle de classe. Quelles que soient les attitudes des enseignants et le profil des apprenants, il semble y avoir une corrélation entre l'enthousiasme de l'enseignant et la mise en œuvre du processus de l'ASC, voire de son efficacité (Lui, Jiang 2009), au moins au début. Des expérimentations longitudinales montreront peut-être une lassitude de la part à la fois des apprenants et des enseignants.

Les expérimentations d'ASC avec FLEURON permettent de constater que les apprenants font de la sociolinguistique de corpus sans le savoir. Ils peuvent observer tous les aspects de la langue (prononciation, prosodie, lexique, syntaxe, pragmatique, sémantique, etc.). Ils associent lexique et grammaire et se créent une interface entre syntaxe et sémantique. Un des apprenants a déclaré que les activités qu'il venait de réaliser étaient : « une façon sympa de faire de la grammaire ». En effet, si faire de la grammaire

pour faire de la grammaire n'a pas d'intérêt dans le cadre de l'apprentissage d'une langue, certains points grammaticaux méritent d'être abordés. La façon de les aborder reste encore souvent homogène. Les enseignements de FLE proposent encore des démarches explicite et déductive (voir par exemple Pescheux 2007), c'est-à-dire que les règles de grammaire sont énoncées (démarche explicite) et que des exercices permettent un entraînement à les appliquer (démarche déductive). Même lorsqu'une démarche inductive est proposée, elle reste basée sur un petit nombre d'exemples choisis par l'enseignant ou déterminés dans un manuel. Bento (2019) a analysé le traitement de la grammaire dans huit manuels de FLE publiés en France entre 2014 et 2017, par les quatre maisons d'éditions qui publient ce type de manuel (Didier, Clé International, Maison des langues et Hachette). Tous les manuels proposent une approche explicite et inductive de la grammaire mais aucun ne propose un recours aux corpus, tels qu'ils sont définis dans le cadre de ce travail. Les activités sont basées sur des exemples choisis par les enseignants, tout comme les questions auxquelles les apprenants doivent essayer de répondre. En outre, aucune initiative n'est permise pour les apprenants.

Avec un concordancier, il est possible d'observer de nombreux phénomènes interactionnels en situation. L'avantage du concordancier de FLEURON est sa multimodalité. Il rend possible d'observer les activités interactionnelles réalisées par les locuteurs, leurs actualisations en fonction du genre de discours et de la situation de communication. Pour le français, Kerr (2013 : 25) a proposé une liste d'observables classés par catégorie, reprise par Boulton et Tyne (2014 : 123). A partir de ces listes et en fonction des expérimentations avec le corpus FLEURON, je peux également proposer un inventaire des éléments les plus recherchés :

- des marqueurs interactionnels (*alors, donc, bon, hein, genre, tu sais, regarde*)
- du lexique, pour accéder au sens (*inscription, séjour*), pour comprendre le lien entre sens et syntaxe (*tout + verbe ou + préposition, même + si, se faire, se voir, penser, dont, ainsi*), pour comparer (*mieux/meilleur, ce/cet*)
- des blocs de mots ou des collocations (*quand même, de toute façon*)
- des constructions syntaxique (*moi je, imparfait de politesse, les clivées en c'est*)
- des connecteurs (*parce que, bien que*)
- des prépositions (*depuis, avant*)
- des expressions (*le truc c'est que*)
- des adverbes (*justement, évidemment*)

- des routines interactionnelles (salutations, remerciement)
- des mécanismes de prise de parole (*alors, euh, donc, voilà*)

Cette liste n'est pas exhaustive. Elle reflète les recherches les plus fréquemment effectuées lors des expérimentations présentées dans ce chapitre. On peut se rendre compte que les apprenants recherchent des éléments interactionnels pour lesquels ils n'ont guère d'explication. Mauranen (2004 : 97) expliquait déjà : « Spoken language is therefore a domain where learners need to work out many linguistic features on their own, because they cannot expect enough help from textbooks, teachers or reference materials ». L'ASC permet donc aux apprenants de découvrir et de trouver des réponses aux spécificités des discours oraux en interaction. L'ASC permet d'aller au-delà de l'exemplarité fournie par les documents de référence (dictionnaires, manuels) en interrogeant une masse de données et d'accéder aux régularités des usages, aux sens les plus fréquents, à la variation et à la contextualisation des données observées. En outre, il semble que le travail qu'implique l'ASC permet un traitement en profondeur de *l'input*.

Dans ce travail, je n'ai pas entrepris d'études sur l'efficacité à long terme de l'ASC. J'ai seulement constaté quelques réutilisations des phénomènes observés lors de la séance ou lors de séances ultérieures. A ce propos, l'étude de Landure (2011) est encourageante. Elle constate un meilleur taux de d'appropriation et de rétention des items consultés en corpus (68%) par rapport à la consultation de dictionnaires (48%). Elle précise que « Le travail sur corpus, par l'exploitation des items en contexte, favorise la construction des savoirs et engage une compréhension « profonde » (Briggs 2003) » (*Ibid.* : 174). Ici aussi des études longitudinales pourront confirmer ou infirmer l'efficacité à long terme de l'exploitation des corpus oraux et multimodaux à des fins d'apprentissage.

Le fait que les apprenants induisent eux-mêmes les règles de fonctionnement semble être une source de motivation et d'efficacité de l'ASC. A ce propos, Chambers (2005: 120) analyse les réactions de ses apprenants et conclut : « These positive reactions suggest that corpora and concordancing certainly have their place in a language-learning environment focusing on learner autonomy and discovery learning ». Les réactions positives des apprenants n'ont cependant pas toujours supprimé l'utilisation d'autres méthodologies ou d'autres stratégies (utilisation des dictionnaires, des grammaires par exemple), souvent moins laborieuses et moins chronophages. Pour tirer le meilleur profit possible de l'ASC, les apprenants doivent également être formés à sa méthodologie. L'implication de l'apprenant dans l'ASC développe ses capacités d'apprentissage, il apprend à apprendre.

Apprendre à partir du concordancier implique une démarche inductive et constructive de la part de l'apprenant qui devient un véritable acteur de son apprentissage. Ce travail créateur lui permet de se construire des repères socio-interactionnels mais aussi des repères d'apprentissage à partir des requêtes effectuées. Cette situation d'apprentissage semble développer l'autonomie des apprenants (Aston 2001). Par autonomie d'apprentissage, Germain et Netten (2004) font référence à « la capacité de mener, activement et de manière indépendante un apprentissage de langue ». Selon l'étude menée par Rivens Mompean et Eisenbeis (2009), le processus d'autonomisation des apprenants serait soutenu par le fait que l'apprentissage est motivé par la recherche conduite par les apprenants. C'est l'utilisation de stratégies d'observation et d'exploration qui contribuerait ainsi au développement de l'autonomie de l'apprenant. Il devient acteur responsable, conscient de ses besoins, préférences, objectifs et difficultés. En fonction des attitudes adoptées par les enseignants, les apprenants peuvent être aidés, soutenus et guidés, au début de leur découverte de l'ASC, mais l'objectif de celui-ci est de rendre les apprenants capables de l'exploiter seuls et que progressivement ils se sentent libres d'exploiter le corpus. Même si l'ASC semble engendrer un processus long (Park, Kinginger 2010) et que l'analyse des données peut prendre beaucoup de temps, plus les apprenants le pratiquent, plus le temps se réduit et l'ASC peut devenir une nouvelle pratique d'apprentissage, presque un réflexe, le « réflexe corpus ».

Un des objectifs de cette étude qualitative est de « permettre au DDL de sortir de ses confins universitaires » (Boulton 2008a : 43). Bien que ce vœu date de 2008, en dix ans peu de choses ont été faites pour que l'ASC puisse être une ressource méthodologique permettant un enseignement et un apprentissage du français parlé en interaction. Pour exaucer ce vœu, il reste un important travail de sensibilisation et de formation à réaliser. Cette formation des enseignants vise à réduire le fossé entre théorie et pratique (Chambers 2019) et de faire des ponts entre les théorisations des chercheurs et les pratiques des enseignants et des apprenants ordinaires (Boulton 2010). La formation des apprenants est également nécessaire, les activités suscitées par l'ASC ne vont pas de soi. Il s'agit d'apprendre petit à petit à apprendre avec des corpus.

Finalement, il le concordancier peut se mettre au service de l'enseignement et de l'apprentissage de la langue parlée en interaction. Les méthodologies de l'ASC peuvent s'ajuster aux enseignants, aux apprenants et aux éléments situationnels. Elles peuvent également, et surtout, venir compléter d'autres approches et d'autres méthodologies. Par

exemple, l'ASC peut intervenir efficacement au cours d'une analyse sociolinguistique des interactions verbales, telle que je l'ai présentée dans le chapitre précédent. En analysant le déroulement des interactions, si un élément soulève un problème de compréhension, un détour par l'ASC permet de le résoudre (ce procédé est illustré avec la recherche de *créneaux*, en [5.3.2.3.](#)). Il apparaît également que l'ASC puisse être utilisée avec des apprenants, quel que soit leur niveau, et par des enseignants, quel que soit leur expérience. La question de l'accessibilité des ressources demeure. Dans l'idéal, les enseignants et les apprenants disposeraient de ressources variées, appropriées, complémentaires, interrogeables simplement et de différentes façons pour répondre à tous leurs besoins socio-interactionnels.

Conclusion

Il est difficile de conclure ce travail parce qu'il représente une étape vers de nouveaux projets de recherche et de nouvelles expérimentations sur le terrain de la formation, de l'enseignement et de l'apprentissage du français (FLE et FL2I). Toutefois, avec le recul et la distance qu'a permis cette synthèse, j'ai pu clarifier, conforter et faire émerger des perspectives que je peux présenter en trois points. Il s'agit :

1. de poursuivre les analyses sociolinguistiques des interactions verbales,
2. de procéder à des études longitudinales et comparatives permettant d'évaluer les effets de l'exploitation didactique de ces analyses et des corpus à long terme,
3. de lancer un programme de recherche à partir d'une plateforme de ressources interactionnelles, documentées et accompagnées de propositions d'exploitations didactiques.

Tout d'abord, ce travail a permis de poser les fondements et d'affiner mon cadre d'analyse sociolinguistique des interactions verbales. La conception sociolinguistique et dynamique de la langue permet de mettre au jour les multiples ressources utilisées par les locuteurs et la façon dont elles se configurent dans des genres de discours et des situations de communication particulières pour réaliser des activités interactionnelles. Le français est assez peu étudié de ce point de vue complexe (contrairement à l'anglais notamment, voir par exemple Couper-Khulén et Selting 2018), il est donc important de **poursuivre les analyses**. Comme le souligne Traverso (2016 : 14) : « Le panorama du français parlé en interaction que l'on peut dresser à ce jour est loin d'être complet, les résultats sont encore fragmentaires et appellent à la poursuite de l'entreprise ». En choisissant une complémentarité des approches (interactionnelle, pronominal, syntaxique, pragmatique, discursive...), je vais poursuivre les travaux d'analyse des interactions verbales pour affiner leur description. En outre, ces travaux, qui sont à l'origine de mes recherches, ne cesseront jamais de me passionner. La complexité des interactions verbales demande à être analysée, expliquée, mise au jour.

De plus, comme je l'ai montré, ces études méritent de faire l'objet de transpositions didactiques afin de faciliter l'accès aux résultats des recherches scientifiques pour les praticiens de langue, c'est-à-dire ceux qui l'enseignent ou l'apprennent au quotidien. Ce travail doit se faire en collaboration avec des enseignants qui peuvent par exemple à la fois participer à la sélection des phénomènes interactionnels à décrire et à la façon de présenter

leurs analyses (par exemple sous forme de fiches descriptives, de collections). Cette collaboration permet également de tester la pertinence de ces transpositions auprès d'apprenants. Les liens étroits que j'entretiens avec Département de FLE de l'université de Lorraine, ainsi qu'avec sa directrice, Florence Poncet, membre de LTF et du projet FLEURON, vont me permettre de poursuivre mes travaux dans ce sens, tels que je les ai menés par exemple pour l'élaboration de la fiche descriptive du marqueur *genre*. L'équipe des vingt enseignants du département est toujours enthousiaste pour ce type de collaboration. Ce fut également le cas lors des formations que j'ai dispensées concernant l'utilisation de FLEURON et d'autres ressources (notamment CLAPI-FLE, Ofrom et ScienQuest pour l'écrit).

Au-delà des transpositions didactiques, l'enseignement et l'apprentissage d'une langue ne peut se passer d'une exposition à la langue (cible) réelle. Pourtant, l'introduction des corpus d'interactions en salle de classe ou dans un dispositif d'apprentissage (autonome par exemple) de la langue est encore rare. Les corpus permettent d'aborder tous les aspects des interactions verbales. Ils ne sont pas cantonnés à l'apprentissage du vocabulaire et de la grammaire. A propos de l'enseignement de l'anglais écrit, Timmis (2015 : 7) fait cette remarque : « we should note that there is no reason why the question should necessary be limited to revolutionising *grammar* teaching given that corpus research embraces other areas of linguistics study highly relevant to ELT [English Language Teaching] such as lexis, discourse and pragmatics ». Les corpus multimodaux permettent encore davantage : observer les façons de prendre la parole, de commencer un tour, de montrer son écoute, de poser des questions, de s'appuyer sur son interlocuteur, de marquer son accord ou son désaccord, *etc.* La liste est loin d'être exhaustive. Dans le cadre d'une exploitation du corpus CLAPI avec des apprenants suédois, Thomas *et al.* (2016) ont montré les avantages de proposer des interactions authentiques situées, variées et nombreuses. Contextualisées, les interactions authentiques sont plus faciles à comprendre que les dialogues de manuels qui ne prennent pas en compte les modes d'élaboration des discours oraux (Gillemore 2004 ; Kamber, Skupien 2009).

Par ailleurs, différents débats dans les travaux anglophones portent sur le sens du terme « révolution » en ce qui concerne l'introduction des corpus pour l'enseignement et l'apprentissage des langues. Cependant, dans le domaine de la didactique de l'oral en FLE, cette introduction dans les pratiques pédagogiques des enseignants ressemble bien à une révolution, puisque jusqu'à présent les corpus ne sont quasiment jamais utilisés, quel que soit le type de corpus (existant ou constitué spécifiquement). Toutefois, pour que les

enseignants introduisent des corpus dans les formations qu'ils dispensent, ils doivent eux-mêmes apprendre à les manipuler. Ils pourront ensuite transmettre à leurs apprenants la façon dont ils peuvent les utiliser. La technologie et l'accessibilité ne fait pas tout. Fligelston (1993 : 98) proposait déjà une pratique pédagogique basée sur trois types de catégories d'activités : *teachning about*, c'est-à-dire, en mes termes, se former et enseigner la sociolinguistique de corpus ; *teaching to exploit*, former les apprenants à exploiter les corpus ; et *exploiting to teach*, exploiter les corpus pour l'enseignement de la langue. Une triple formation – sociolinguistique, technique et didactique – est nécessaire et permet à l'enseignant d'avoir un rôle pilote (Breyer 2011) dans l'exploitation des corpus par les apprenants. De plus, beaucoup de recherches en didactique ne s'intéressent pas au corpus et à leurs exploitations pédagogiques. Pour cela, il est nécessaire de créer des liens entre les disciplines et des allers-retours entre théorie et pratique.

En outre, cette révolution peut être prise au sens marxiste, c'est-à-dire un bouleversement de l'ordre social et politique. Exploiter les corpus contribue au processus d'émancipation des apprenants. Leur usage régulier les rend plus autonomes, plus indépendants de la présence d'un enseignant. Quand les apprenants ont le « réflexe corpus », il travaille à leur auto-émancipation. Cet objectif n'est pas négligeable même s'il semble soumis à certaines conditions. Tout d'abord, dans le cadre de l'enseignement de l'anglais écrit, Landure (2013) montre il faut plus de temps pour rechercher et comprendre quelque chose dans un corpus que dans un dictionnaire, mais elle montre également que le travail sur corpus est plus efficace en termes de rétention. Ensuite, l'utilisation des corpus nécessite un temps d'apprentissage méthodologique. Ciekanski (2014 : 2) s'interroge sur le « manque d'intérêt apparent pour les corpus dans le champ de l'autonomie ». A cette question, elle répond « Le recours aux corpus en situation d'apprentissage en autonomie ne peut (...) faire l'impasse d'une médiation spécifique pour conduire à un apprentissage effectif. » (*Ibid.* : 9). En effet, lors de différentes expérimentations de FLEURON, malgré l'intégration de son corpus multimédia dans un dispositif d'apprentissage dédié et pensé pour être utilisé en autonomie, il a été constaté que, quel que soit le terrain et le public, une phase de médiation est fondamentale. Cette médiation permet de faire en sorte que les apprenants continuent à consulter les corpus en dehors de la salle de classe, pour une utilisation régulière. Avant d'être autonomes, les apprenants ont besoin d'apprendre à apprendre à utiliser les ressources qui sont à leur disposition. Dans le cadre d'un apprentissage formel, partir à la découverte des données, prendre conscience du

fonctionnement de la langue en interaction ou induire des règles de fonctionnement ne font pas partie des pratiques habituelles des apprenants.

Dans ce travail, j'ai voulu montrer que l'exploitation de l'analyse sociolinguistique des interactions verbales et des corpus était possible dans différentes situations et avec différents types d'apprenants. Bien que la majorité de mes expérimentations ait été menée auprès d'apprenants de FLE à l'université, celles qui ont été réalisées auprès d'un public FL2I se sont révélées pertinentes et encourageantes. J'ai proposé deux façons d'exploiter les données en corpus : en analysant le déroulement des interactions verbales et en utilisant un concordancier. Ces deux façons sont complémentaires. Un apprenant peut analyser une interaction puis apprendre un phénomène grâce au concordancier. Réciproquement, un apprenant peut d'abord tenter de comprendre un phénomène grâce au concordancier puis aller observer son utilisation en interaction. Ces premiers résultats ont permis d'évaluer la faisabilité et d'asseoir une méthodologie. Ils mettent au jour également son efficacité, au moins ponctuelle, c'est-à-dire que les apprenants comprennent les interactions et les pratiques interactionnelles qu'ils observent et leurs variations situationnelles. Parfois, ils s'approprient rapidement ces pratiques et les utilisent, de façon appropriée, comme ressources pour leurs propres interactions. Désormais, l'objectif est d'en étudier les effets à long terme.

Des études longitudinales, quantifiées (avec des évaluations socio-interactionnelles) et comparatives (avec des groupes témoins), doivent être menées. Je vais poursuivre mes collaborations, en France et à l'étranger, pour effectuer ces études, à la fois quantitatives et qualitatives pour confirmer ou infirmer les effets positifs de l'utilisation des corpus sur l'apprentissage, et notamment de l'ASC, c'est-à-dire de l'exploitation outillée d'un concordancier des corpus. Par ailleurs, mes liens avec différents organismes de formation pour adultes vont me permettre de poursuivre l'étude de l'introduction de l'ASC avec des publics FL2I. Si l'utilisation des corpus de données authentiques avec ce public est évidente, la méthodologie de l'ASC mérite d'être évaluée plus amplement. Il s'agit notamment d'observer si elle ne représente pas une surcharge de travail pour les formateurs et si elle peut s'inscrire dans la temporalité des formations courtes pour adultes.

L'objectif de ce travail a été de mettre en œuvre l'ASC pour apprendre à interagir à l'oral en FLE, de montrer des pratiques d'enseignants et d'apprenants, de proposer différentes façons de faire, de donner la possibilité de reproduire ces expérimentations, de faciliter l'introduction des corpus en didactique et, plus généralement d'évaluer la

méthodologie et la faisabilité. La poursuite de ce travail sera de diffuser ces éléments pour qu'ils soient facilement accessibles. Je rejoins Chambers (2019 : 5) qui regrette ce manque de diffusion :

Indeed, one of the reasons for the research–practice gap may well be that the means of dissemination of the research, namely the research articles or book chapters, are very often published in collections, such as special issues of journals or edited volumes, which are aimed primarily at other specialists and not at non-specialist practitioners.

Ainsi, j'ai tenté de réduire ce *gap* entre théorie et pratique afin de tester l'ASC. J'ai souhaité aller au-delà d'une approche théorique qui propose uniquement des explications et des principes à appliquer. J'ai tenté de montrer les « dessous » pratiques, c'est-à-dire d'examiner ce qu'il se passe d'un point de vue méthodologique. Mon objectif n'est pas de réinterroger l'ASC tel qu'il a été défini par Johns et par d'autres, de nombreuses études sont consacrées à la définition du terme *data-driven learning* (Chambers 2010) même si « its boundaries are fuzzy » (Boulton 2011 : 575). Mon objectif était d'analyser ces frontières floues et d'examiner si l'ASC peut effectivement être mis en œuvre par des apprenants pour qu'ils puissent observer des interactions et apprendre à interagir.

L'apprentissage sur corpus peut tout à fait être utilisé pour d'autres types de publics que celui pour lequel il a été prévu. J'ai montré que des apprenants de faible niveau en langue et faiblement scolarisé pouvaient devenir des détectives de la langue, comme les étudiants auprès desquels les premières expérimentations ont été réalisées. Il paraît important également de poursuivre l'utilisation des corpus avec des publics natifs en insécurité langagière. Il semble que la méthodologie de l'ASC facilite l'entrée dans l'écrit en permettant une visualisation orthographique massive et ordonnée (verticalement) de listes de mots. La poursuite de ces expérimentations permettrait de répondre à Adami (2018 : 91) qui propose la piste des corpus en s'interrogeant sur la façon d'« intégrer des formes de métalangues nécessaires sans passer par la grammaire de type scolaire » avec un public FL2I. L'ASC apporterait un complément aux formations actuelles de ce champ didactique. En outre, je ne suis pas spécialiste de l'apprentissage des langues par les enfants, mais je suis persuadée que l'ASC pourrait y contribuer (y compris à l'apprentissage de la langue maternelle). Une étude pilote a été menée en Italie en ce qui concerne l'apprentissage du FLE par des enfants de niveau collège (Di Vito 2018). Cette étude a certes été réalisée avec un corpus écrit de petite taille, interrogé avec *Antconc*, et avec un accompagnement

spécifique de la part des enseignantes qui l'ont mise en œuvre, mais les premières conclusions amènent Di Vito (2018 : 54) à affirmer : « nous sommes convaincue qu'une telle approche peut apporter des bénéfices également au niveau de l'enseignement au collège, dans la mesure où elle peut augmenter la motivation des apprenants et les encourager à être plus autonomes dans leur apprentissage. ». Les enjeux résident quand même dans la formation, des enseignants et des apprenants, dont j'ai mentionné la nécessité à plusieurs reprises. En ce qui concerne la formation des enseignants, au-delà de mes enseignements à l'université de Lorraine, je participe à des ateliers ponctuels dans d'autres universités. Une école d'été sur l'exploitation des corpus en didactique est également en préparation à Nancy pour 2020 (soutenue par Lorraine Université d'Excellence). Ces formations mettent les enseignants en situation d'utilisation des corpus, en pratiquant eux-mêmes la méthodologie de l'ASC notamment. Cette dernière n'est pas toujours une évidence, malgré les approches et les concepts développés actuellement en didactique. Un travail d'accompagnement est nécessaire. Il n'est pas question d'imposer des changements aux enseignements mais de leur présenter de nouvelles pratiques.

L'apprentissage sur corpus n'est pas une méthodologie exclusive. Elle n'a pas vocation à remplacer les autres méthodologies didactiques. Néanmoins, elle est, pour l'instant, absente des manuels de langue. Les éditeurs de manuels ne prennent pas de risque par rapport aux pratiques professionnelles des enseignants, ils ne s'aventurent pas dans des approches trop novatrices ou trop éloignées leurs habitudes. Il serait pourtant utile de rédiger un guide d'exploitation des corpus à des fins d'apprentissage du français parlé en interaction (un peu à l'instar de ce qui peut se faire pour l'anglais écrit, voir par exemple Liu et Lei 2017).

Enfin, de ces expérimentations, de ces premiers résultats, ainsi que des premières expériences menées autour de CLAPI-FLE, il apparaît qu'une mise à disposition, à la fois des résultats de la recherche et des corpus oraux et multimodaux, plus ambitieuse pourrait réellement faire évoluer la didactique de l'oral en FLE. Il s'agirait d'envisager une recherche à partir d'une **plateforme de ressources** dans laquelle les enseignants et les apprenants pourraient puiser pour construire leur enseignement en salle de classe ou leur apprentissage en autonomie. Cette ambition pourrait être concrétisée dans le cadre d'un grand programme de recherche qui analyserait les possibilités d'exploiter à la fois les résultats de travaux scientifiques et des ressources en corpus à des fins didactiques⁸⁴. Si les

⁸⁴ Cette recherche fera l'objet d'un dépôt de projet à l'ANR cette année.

travaux sont concluants, l'utilisation de cette plateforme pourrait être schématisée de la façon suivante :

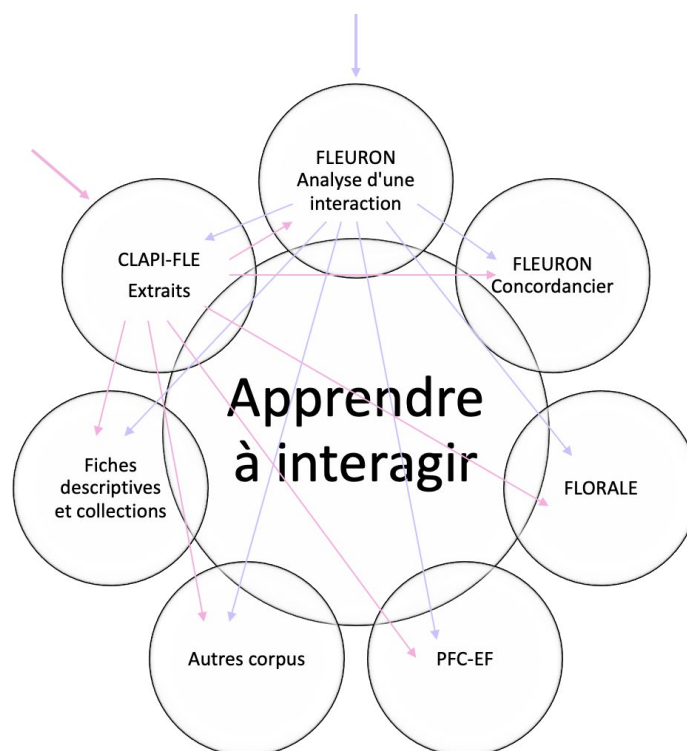


Figure 61 : Complémentarité des ressources pour apprendre à interagir

Pour apprendre à interagir, plusieurs ressources sont mises à disposition. Il est possible d'entrer par n'importe quelle ressource. Les flèches de couleur symbolisent ces entrées. Deux parcours seulement sont présentés pour des soucis de lisibilité du schéma mais chaque ressource peut être un point d'entrée dans la plateforme. Par exemple, un apprenant peut observer certains extraits de CLAPI-FLE puis naviguer vers les autres ressources (flèches roses). Une autre entrée possible est celle qui consiste à observer et à analyser une interaction du corpus FLEURON (flèches violettes). Cette première étape peut être complétée par chacune des ressources, présentées dans le schéma, ou par toutes les ressources. L'apprenant peut utiliser le concordancier pour rechercher le fonctionnement d'un élément qu'il a rencontré dans l'interaction, ensuite il peut aller consulter sa fiche descriptive ou écouter comment il se prononce ou encore l'observer dans d'autres corpus, etc.

Les ressources utilisées (y compris les corpus de linguistes) devront être proposées au sein d'une plateforme conviviale et dont l'utilisation sera modulable et expliquée. Si l'utilisation des corpus ne s'est pas répandue, c'est probablement parce que les applications

simples manquent. Dans un autre contexte, Tyne (2012 : 114) le remarquait déjà : « And whilst researchers and corpus linguists may be quite keen to underline the potential of their work for teaching and learning, there is a lack of user-friendly applications for general, everyday practice. ». Ainsi, ce projet vise à combler le manque fréquemment mentionné de corpus et à faciliter leur accessibilité par des enseignants et des apprenants de langue. D'autres technologies pourraient à terme venir compléter ces outils, je pense principalement à la lemmatisation des corpus (et celle de FLEURON en particulier) et à leur annotation, notamment en partie du discours ou pragmatique (comme c'est le cas pour les extraits de CLAPI-FLE). Le programme de recherche consiste à réfléchir aux données, à leur traitement et à leur accessibilité, en poursuivant notamment les travaux initiateurs du projet ORFEO⁸⁵, mais surtout à analyser les types de parcours, d'enseignement et d'apprentissage, possibles et leur efficacité. J'ai commencé à tester la complémentarité de certaines de ces ressources, notamment celle de FLEURON et CLAPI-FLE. Certaines études tendent à montrer les prémices des avantages à utiliser plusieurs ressources. Elles ont été présentées dans le cadre de différentes manifestations (par exemple aux colloques PRELA à Lyon et AFLS à Bristol) et font l'objet d'un article en préparation (avec Carole Etienne du projet CLAPI-FLE). Nous avons désormais besoin d'étendre ces études. Avec ce projet de grande envergure, il s'agit d'accroître les recherches sur la complémentarité des ressources et de la mettre au service de l'enseignement et de l'apprentissage du français parlé en interaction. L'exploitation combinée de ces ressources vise l'amélioration des compétences socio-interactionnelles des apprenants pour qu'ils puissent interagir de façon appropriée dans les situations de communication qu'ils rencontrent.

⁸⁵ Pour rappel : <https://www.projet-orfeo.fr/>

Bibliographie

- Abe D., Carton F., Cembalo S. M., et Régent O. (1979). Didactique et authentique : du document à la pédagogie. *Mélanges pédagogiques*, p.1-14.
- Abeillé A., Godard D. (à paraître). *La Grande Grammaire du français*. Arles : Actes Sud.
- Adami H. (2018), Existe-t-il une vie didactique sans la grammaire ? *Les langues Modernes*, 3, p.85-92.
- Adami H. (2015). De quoi les études plurilingues sont-elles le nom ? In Adami H., André V. (eds). *De l'idéologie monolingue à la doxa plurilingue : regards pluridisciplinaires*. Bern : Peter Lang, p.43-90.
- Adami H. (2009a). *La formation linguistique des migrants*. Paris : Clé international.
- Adami H. (2009b). Les documents authentiques dans la formation des adultes migrants : pratiques pédagogiques et contraintes institutionnelles, *Mélanges Crapel*, 31, p.159-172.
- Adami H. (2007). Dire le savoir-faire en contexte professionnel : problèmes, enjeux et perspectives. *Le français dans le monde. Recherches et applications. Langue et travail*, Juillet 2007, p.88-99.
- Adami H., André V. (2014). Les processus de sécurisation langagière des adultes : parcours sociaux et cursus d'apprentissage. *Revue française de linguistique appliquée*, Vol. XIX-2, p. 71-83.
- Adami H., André V. (2013). Corpus et Français Langue d'Intégration (FLI), *LINX*, n°68-69, p.135-158.
- Adami H., André V. (2012a). Formation aux écrits professionnels : vers une sociolinguistique impliquée. *Lidil. Revue de linguistique et de didactique des langues*, 45, p.51-71. En ligne : <https://journals.openedition.org/lidil/3178>
- Adami H., André V. (2012b). Vers le Français Langue d'Intégration et d'Insertion. In Adami H. et Leclercq V. (dir.), *Les migrants face aux langues des pays d'accueil. Acquisition en milieu naturel et formation*. Lille : Septentrion, p.277-289.
- Adami H., André V. (2010). Les migrants en insécurité linguistique au travail : analyses et perspectives de formation. *Points Communs*, n°40. [hal-00526362](https://hal.archives-ouvertes.fr/hal-00526362)
- Adami H., André V., Bailly S., Castillo D., Poncet F. et Riley P. (2003). L'étranger compétent. Un nouvel objectif pour la didactique des langues étrangères. In Groux, D., Lemeunier, M. (Eds). *Une identité plurielle : mélanges offerts à Louis Porcher*. Paris : Éditions L'Harmattan, p.535-550.
- Adami H., André V., Ciekanski M. (eds) (2015). La revue *Langage, Travail et Formation*. N° zéro. En ligne : <https://apps.atilf.fr/reseaultf/?cat=17>
- Adami H., André V., Langbach V. (dir.) (à paraître). *Les adultes en insécurité langagière. Enjeux sociaux et didactiques*. Villeneuve d'Ascq : Septentrion.
- Ahmad K., Corbett G., Rogers M. (1985). Using computers with advanced language learners : An example. *The Language Teacher*, 9(3), p.4-7.
- Alberdi C., Étienne C., Jouin-Chardon É. (2018). Les apports des corpus d'interactions naturelles en situation de classe : enjeux et pratiques. *Action Didactique*, 1, p.55-70. En ligne : <http://univ-bejaia.dz/action-didactique/ad1.html>

- André V. (à paraître 2020). Faire de la linguistique de corpus avec des apprenants de français langue étrangère. *Syntaxe et sémantique*.
- André V. (à paraître 2019). Des corpus oraux et multimodaux authentiques pour acquérir des compétences sociolangagières. In Gajo L., Luscher J.-M., Racine I., Zay F. (eds), *Variation, plurilinguisme et évaluation en FLE*. Bern : Peter Lang, p.209-223.
- André V. (2018/soumis). Étude micro-diachronique de genre dans le corpus TCOF entre 2006 et 2017. *50 ans de linguistique sur corpus oraux : apports à l'étude de la variation*, Orléans, 15-17 novembre. Article soumis à *Langages*.
- André V. (2018a). Nouvelles actions didactiques : faire de la sociolinguistique de corpus pour enseigner et apprendre à interagir en français langue étrangère. *Action didactique*, n°1, p.71-88.
- André V. (2018b). *Plus belle la vie* : analyse des interactions et sélection d'un corpus pour l'apprentissage du français langue étrangère. *Congrès Mondial de Linguistique Française – CMLF 2018*. SHS Web of Conference, 46, p.1-21. En ligne : <https://doi.org/10.1051/shsconf/20184607013>
- André V. (2017). Un corpus multimédia pour apprendre à interagir en situations universitaires en France. *Actes du troisième colloque international de l'ATPF « Enseigner le français : s'engager et innover »*, p.292-315.
- André V. (2016). FLEURON : Français Langue Étrangère Universitaire – Ressources et Outils Numériques. Origine, démarches et perspectives. *Mélanges Crapel*, 37, p.69-92. En ligne : <http://www.atilf.fr/IMG/pdf/art3.pdf>
- André V. (2015a). Sociolinguistique des interactions verbales : de l'analyse des situations de travail aux implications sociales, *Langage, Travail et Formation*, numéro zéro, p.1-10.
- André V. (2015b). L'hégémonie de l'anglais en situation de travail : une contrainte inéluctable ? In Adami H., André V. (Eds). *De l'idéologie monolingue à la doxa plurilingue : regards pluridisciplinaires*. Bern : Peter Lang, p.91-130.
- André V. (2014). L'énonciation conjointe : trace et ressource de la construction collaborative du discours. *Congrès Mondial de Linguistique Française – CMLF 2014*. SHS Web of Conference 8, p.1891-1904.
- André V. (2011). Analyse pragmatolinguistique de la co-construction du discours : le cas des énonciations conjointes et des reprises. *Actes du colloque XII International Symposium on Social Communication*, Organisé par le Centre de linguistique appliquée, 17-21 janvier, Santiago de Cuba, p.322-326.
- André V. (2010a). Emplois stratégiques des Formes Nominales d'Adresses au sein de réunions de travail en entreprise. In Kerbrat-Orecchioni C. (Ed.), *S'adresser à autrui. Les formes nominales d'adresse en français*. Chambéry : Université de Savoie, p.63-87.
- André V. (2010b). Éléments de construction collaborative du discours au sein de réunions de travail : la reprise et le couple oui non. *Pratiques*, n°147/148, p.199-222.
- André V. (2009a). Les compétences langagières des métiers de la propreté : de l'analyse des situations de communication à la formation professionnelle. *Bulletin Suisse de Linguistique Appliquée*, 90, p.149-165.
- André V. (2009b). La parole au travail : interdiction, valorisation, prescription. *Savoirs et Formations. Le français professionnel*, n°70, p36-38.

- André V. (ed.). (2008). Langues et relations de service. *Mélanges CRAPEL*, n°30.
- André V. (2006a). *Construction collaborative du discours au sein de réunions de travail en entreprise : de l'analyse micro-linguistique à l'analyse socio-interactionnelle*. Thèse de Doctorat. Sous la direction de Philip Riley. Université Nancy 2. Soutenue le 2 juin.
- André V. (2006b). Un chercheur en sociolinguistique dans une réunion de travail en entreprise. In Léglise I., Canut E., Desmet I., Garric N. (Eds). *Applications et Implications en Sciences du Langage*, Paris : L'Harmattan, p.41-52.
- André V. (2005). Oui non : une pratique discursive sous influence. *Marges Linguistiques*, n° 9, p.195-213. En ligne : http://www.revue-texto.net/Parutions/Marges/Marges_sommaire.html
- André V., Benzitoun C., Canut E., Debaisieux J.-M., Gaiffe B. (2010). Traitement informatique de données orales : quels outils pour quelles analyses ? *Recherches qualitatives – Hors Série « Logiciels pour l'analyse qualitative : innovations techniques et sociales »*, 9, p.131-150.
- André V., Boniface Y. (à paraître 2019). Des balbutiements à DECIDE : les tribulations de la lampe Psyphine. In Collectif Psyphine (ed.). *Que prêtons-nous aux machines ? Approches interdisciplinaires des interactions homme-robot*. Paris : Presses universitaires de Paris Nanterre.
- André V. et Boniface Y. (2018). Quelques considérations interactionnelles autour d'une expérience robotique. Workshop ACAI 2018. En ligne : http://www.lsis.org/wacai2018/proceedings/WACAI2018_paper_12.pdf
- André V., Canut E. (2010). Mise à disposition de corpus oraux interactifs : le projet TCOF (Traitement de Corpus Oraux en Français). *Pratiques*, n°147/148, p.35-51.
- André V. (ed.) (2008). Langues et relations de service. *Mélanges CRAPEL*, n°30.
- André V., Castillo D. (2005). The 'Competent Foreigner' : A new model for foreign language didactics ? In Preisler B., Fabricius A., Haberland H., Kjaerbeck S. and Risager K. (eds). *The Consequences of Mobility: Linguistic and Sociocultural Contact Zones*, Roskilde (Denmark): Roskilde University, Department of Language and Culture, p.154-162.
- André V., Ciekanski M. (2018). Apprendre à interagir à l'oral à partir d'un concordancier multimodal : effets sur le développement de la conscience langagière et sur l'autonomie de l'apprenant dans le dispositif FLEURON. In Dejean C., Mangenot F., Nissen E., Soubrié T. (coord.). *EPAL - Échanger Pour Apprendre en Ligne, Juin 2018, Grenoble, France*.
- André V., Collectif Psyphine (eds). (à paraître 2019). *Que prêtons-nous aux machines ? Approches interdisciplinaires des interactions homme-robot*. Nanterre : Presses Universitaires Paris Nanterre.
- André V., Pira C. (2006). Phénomènes de reprise : statut macrosyntaxique, effets pragmatiques et organisation interactionnelle. Communication présentée au Colloque international *Les linguistiques du détachement*, à Nancy, les 7-9 juin.
- André V., Tyne H. (2014). Sociolinguistics and the Study of French. In Tyne H., André V., Benzitoun C., Boulton A., Greub Y. (Eds.). *French through corpora: ecological and data-driven perspectives in French language studies*. Cambridge : Cambridge Scholars Publishing, p.132-138.

- André V., Tyne H. (2012). Compétence sociolinguistique et dysfluence en L2. In Kamber A., Skupens Dekens C. (éds). *Recherches récentes en FLE*. Berne : Peter Lang, p.21-46.
- André-Larochebouvy D. (1984). *La conversation quotidienne. Introduction à l'analyse sémiolinguistique de la conversation*. Paris : Didier.
- Arrivé M., Gadet F., Galimiche M. (1986). *La grammaire d'aujourd'hui. Guide alphabétique de linguistique française*. Paris : Flammarion.
- Aston G. (2001). Learning with corpora: an overview. In Aston G. (ed.). *Learning with corpora*. Houston: Athelstan, p.7-46.
- Aston G. (1997). Involving learners in developing learning methods: Exploiting text corpora in self-access. In Benson P., Voller P. (dir.). *Autonomy and independence in Language Learning*. Londres: Longman, p.204-263.
- Austin J. L. (1962/trad. française 1970). *Quand dire c'est faire*. Paris : Seuil.
- Avanzi M., Conti V., Corminboeuf G., Gachet F., Johnsen L. A., Montchaud P. (dir.) (2014). *Enseignement du français : les apports de la recherche en linguistique*. Bruxelles : Peter Lang.
- Bailly S., Boulton A., Château A., Duda R., Tyne H. (2009). L'anglais langue d'appui pour l'apprentissage du français langue étrangère. In Forlot G. (ed). *L'anglais et le Plurilinguisme : pour une didactique des contacts et des passerelles linguistiques*. Paris : L'Harmattan, p.35-57.
- Bakhtine M. (1984). *Esthétique de la création verbale*. Paris : Gallimard.
- Bakhtine M. (Volochnikov V.N.) (1977/1929). *Le marxisme et la philosophie du langage. Essai d'application de la méthode sociologique en linguistique*. Paris : Les Éditions de minuit.
- Baude O. (Coord.) (2006). *Corpus oraux, guide des bonnes pratiques 2006*. CNRS Éditions. Orléans : Presses Universitaires Orléans.
- Bautier E. (2016). Et si l'oral pouvait permettre de réduire les inégalités ? *Les Dossiers des Sciences de l'Éducation*, 36, p.95-108. En ligne : <https://journals.openedition.org/dse/1397>
- Becker J., André V., Dutech A. (2019). Qualcom : une expérience sur la qualification des comportements d'une lampe robotique. *Techniques & Culture, Varia*. En ligne : <https://journals.openedition.org/tc/10771>
- Beeching K. (2016). *Pragmatic markers in British English. Meaning in social interaction*. Cambridge: Cambridge University Press.
- Beeching K. (2007). La co-variation des marqueurs discursifs bon, c'est-à-dire, enfin, hein, quand même, quoi et si vous voulez : une question d'identité ? *Langue française*, 154 (2), p.78-93.
- Béguelin M.-J., Coveney A., Guryev A. (dir.) (2018). *L'interrogative en français*. Bern : Peter Lang.
- Bento M. (2019). Enseignement de la grammaire dans les manuels de français langue étrangère et approche actionnelle. *Lidil. Revue de linguistique et de didactique des langues*, 59. En ligne : <https://journals.openedition.org/lidil/6465>.
- Bentolila A. (1996). *De l'illettrisme en général et de l'école en particulier*. Paris : Plon.
- Benzitoun C., Fort K., Sagot B. (2012). TCOF-POS : un corpus libre de français parlé annoté en morphosyntaxe. *JEP-TALN 2012 - Journées d'Études sur la Parole et conférence*

- annuelle du Traitement Automatique des Langues Naturelles*, Juin 2012, Grenoble, France. p.99-112.
- Bernardini S. (2004). Corpora in the classroom : An overview and some reflexions on future developments. In Sinclair J. (ed.). *How to Use Corpora in Language Teaching*. Amsterdam: Johns Benjamins, p.15-36.
- Bernardini S. (2000). Systematising serendipity: proposals for concordancing large corpora with language learners. In Burnard L., McEnery T. (dir.). *Rethinking language pedagogy from a corpus perspective*. Frankfurt: Peter Lang. p.225-234.
- Bernstein B. (1975). *Langage et classe sociale*. Paris : Les Éditions de Minuit.
- Biber D. (1988). *Variation across Speech and Writing*. Cambridge : Cambridge University Press.
- Biber D., Conrad B. (2009). *Register, Gender and Style*. Cambridge : Cambridge University Press.
- Biber D., Douglas B., Conrad S. & Reppen R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge : Cambridge University Press.
- Biber D., Johansson S., Leech G., Conrad S., Finegan E. (1999). *Longman Grammar of Spoken and Written English*. Londres : Pearson.
- Bilger M. (coord.) (2008). *Données orales. Les enjeux de la transcription*. Perpignan : Presses Universitaires de Perpignan.
- Bilger M. (1999). Quelques problèmes autour de la « représentation » des données orales. In Barberis J.-M. (Ed.). *Le français parlé. Variété et discours*. Montpellier : Publication de l'Université Paul-Valéry, p.181-193.
- Blanche-Benveniste C. (2000). *Approches de la langue parlée en français*. Paris : Éditions Ophrys.
- Blanche-Benveniste C. (1990a). *Le français parlé. Études grammaticales*. Paris : CNRS Éditions.
- Blanche-Benveniste C. (1990b). Un modèle d'analyse syntaxique « en grille » pour les productions orales. *Anuario de Psicologia*, 47, p.11-28.
- Blanche-Benveniste C., Jeanjean C. (1987). *Le français parlé. Transcription et édition*. Paris : Didier Érudition.
- Blasco-Dulbecco M. (1999). *Les dislocations en français contemporain : étude syntaxique*. Paris : Honoré Champion.
- Bloomfield L. (1970). *Le langage*. Paris : Payot [Edition originale : 1933].
- Borzeix A. (1998). Genèse d'un réseau, effets d'une perspective. In Lazar A. (Ed.). *Langage(s) et travail : Enjeux de formation*. Actes du colloque INRP/CNAM/CNRS-LT, Octobre 1998. Paris : INRP, p.31-34.
- Borzeix A., Charles J., Zimmermann B. (2015). Réinventer le travail par la participation. Actualité nouvelle d'un vieux débat. *Sociologie du travail*, 57, p.1-19.
- Borzeix A., Fraenkel B. (dir.) (2001). *Langage et travail. Communication, cognition, action*. Paris : CNRS Editions.
- Bouchard, R. (2002). *Alors, donc, mais...*, « particules énonciatives » et / ou « connecteurs » : Quelques considérations sur leur emploi et leur acquisition. *Syntaxe et sémantique*, 3(1), p.63-73.

- Boulton A. (2018). Les corpus et les TIC comme aide à la découverte des langues. *Les Langues Modernes*, 3, p.71-84.
- Boulton A. (2017). Corpora in language teaching and learning. *Language Teaching*, 50(4), p.483–506.
- Boulton A. (2012). Corpus consultation for ESP: A review of empirical research. In Boulton A., Carter-Thomas S., Rowley-Jolivet E. (dir.). *Corpus-Informed Research and Learning in ESP: Issues and Applications*. Amsterdam: John Benjamins, p.261-291.
- Boulton A. (2011). Data-driven learning: The perpetual enigma. In S. Goźdz-Roszkowski (dir.). *Explorations across Languages and Corpora*. Francfort : Peter Lang, p.563-580.
- Boulton A. (2010). Data-Driven Learning: On Parper, in Practice. In Harris T., Moreno Jaén M. (eds). *Corpus Linguistics in Language Teaching*. Bern: Peter Lang, p.17-52.
- Boulton A. (2008a). Esprit de corpus : Promouvoir l'exploitation de corpus en apprentissage des langues. *Texte et Corpus*, 3, p.37-46.
- Boulton A. (2008b). Looking for empirical evidence of data-driven learning at lower levels. In B. Lewandowska-Tomaszczyk (ed.). *Corpus Linguistics, Computer Tools, and Applications: State of the Art*. Frankfurt: Peter Lang, p. 581-598.
- Boulton A., Cobb T. (2017). Corpus use in language learning: a meta-analysis. *Language Teaching*, 67, p.348-393.
- Boulton A., Tyne H. (2014). *Des Documents Authentiques aux Corpus. Démarches pour l'Apprentissage des Langues*. Paris : Didier.
- Bourdieu P. (1997). *Méditations pascaliennes*. Paris : Le Seuil.
- Bourdieu P. (1994). *Raisons pratiques*. Paris : Le Seuil.
- Bourdieu P. (1984). *Question de sociologie*. Paris : Les Éditions de minuit.
- Bourdieu P. (1983). Vous avez dit « populaire » ? *Actes de la recherche en sciences sociales*, 46, p.98-105.
- Bourdieu P. (1977). L'économie des échanges linguistiques. *Langue française*, 34, p.17-34.
- Bourdieu P., Passeron J.-C. (1970). *La reproduction. Éléments pour une théorie du système d'enseignement*. Paris : Les Éditions de minuit.
- Bourdieu P., Passeron J.-C. (1964). *Les héritiers. Les étudiants et la culture*. Paris : Les Éditions de minuit.
- Boutet J. (2001). La part langagière du travail : bilan et évolution, *Langage & société*, 2001/4, 98, p.17-42.
- Boutet J. (Ed.) (1995). *Paroles au travail*. Paris : L'Harmattan.
- Boutet J. (1994). *Construire le sens*. Bern : Peter Lang.
- Braun O. (2019). *L'apport de Fleuron dans l'apprentissage du FLE : mécanismes et accomplissements des apprenants*. Mémoire de Master Sciences du langage et didactique des langues. Université de Lorraine.
- Braun S. (2007). Integrating corpus work into secondary education: From data-driven learning to needs-driven corpora. *ReCALL*, 19(3), p.307–328.
- Breyer Y. A. (2011). *Corpora in Language Teaching and Learning. Potential, Evaluation, Challenges*. Frankfurt an Main : Peter Lang.

- Brié M. (2018). *Le traitement inférentiel d'une locutrice qualifiée en interaction avec des personnes en insécurité langagière*. Mémoire de Master Sciences du langage et didactique des langues. Université de Lorraine.
- Briggs J. (2003). *Teaching for quality learning at university: what the student does?* Buckingham: Open University Press/Society for Research into Higher Education.
- Bronckart J.-P. (2008). Discussion de quelques concepts pour une approche praxéologique du langage. *Actes du 1er Congrès mondial de linguistique française*, p.855-861.
- Burnard L. (Ed.). (2004). *BNC Baby* [CD-Rom]. Oxford: Research and Technology Service, Oxford University.
- Calame-Griaule G. (1965). *Ethnologie et langage : la parole chez les Dogon*. Paris : Gallimard.
- Cameron D. (1995). *Verbal Hygiene*. London : Routledge.
- Candéa M., Véron L. (2019). *Le français est à nous ! Petit manuel d'émancipation linguistique*. Paris : La Découverte.
- Cappeau P. (2016). Questions sur l'oral : médium, syntaxe, genre. *Le français d'aujourd'hui*, 195, p.23-36.
- Cappeau P., Gadet F. (2007). L'exploitation sociolinguistique des grands corpus. Maître-mot et pierre philosophale. *Revue française de linguistique appliquée*, Vol. XII-1 p.99-110
- Carette E. (2001). Mieux apprendre à comprendre l'oral en langue étrangère. *Le français dans le monde. Recherches et applications. Numéro spécial : Oral : variabilité et apprentissage*, p.126-142.
- Carton F. (1995). L'apprentissage différencié des quatre aptitudes. *Verbum*, p.63-74.
- Cefai D. (2003). Postface. L'enquête de terrain en sciences sociales. In Cefai D. (Ed.). *L'enquête de terrain*. Paris : Éditions La Découverte, p.465-615.
- Cerisara C., Mella O., Fohr D. (2009). JTrans: an open-source software for semi-automatic text-to-speech alignment. In *INTERSPEECH-2009*, p.1823-1826.
- Chambers A. (2019). Towards the corpus revolution? Bridging the research-practice gap. *Language Teaching*, p.1-16.
- Chambers A. (2010). What is data-driven learning? In O'Keeffe A., McCarthy M. (ed.). *The Routledge Handbook of Corpus Linguistics*. New York: Routledge, p.345-358.
- Chambers A. (2009). Les corpus oraux en français langue étrangère : authenticité et pédagogie, *Mélanges Crapel*, 31, p.15-33.
- Chambers A. (2005). Integrating corpus consultation in language studies. *Language Learning & Technology*, 9(2), p.111-125.
- Chambers A. (1997). What do we mean by fluency? *System*, 25/4, p.535-544.
- Chauveau-Thoumelin P. (2016). De l'exemplification à la catégorisation approximative : étude de la construction [[X] SN genre [Y] SN]. In *SHS Web of Conferences* (Vol. 27, p. 12005). EDP Sciences.
- Cheng W., Warren M., Xun-Feng X. (2003). The language learner as a researcher : Putting corpus linguistics on the time-table. *System*, 31(2), p.173-186.
- Clark H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.
- Clot Y. (1999). *La fonction psychologique du travail*. Paris : Presses Universitaires de France.
- Clot Y. (1995). *Le travail sans l'homme ? Pour une psychologie des milieux de travail et de vie*. Paris : La Découverte.

- Clot Y., Faïta D. (2000). Genres et styles en analyse du travail. Concepts et méthodes. *Travailler*, 4, p.7-42.
- Clot Y., Faïta D., Fernandez G., Scheller L. (2000). Entretiens en autoconfrontation croisée : une méthode en clinique de l'activité. *Perspectives interdisciplinaires sur le travail et la santé*, 2-1, p.1-9. DOI : 10.4000/pistes.3833.
- Coastes J. (1994). No gap, lots of overlap: turn-taking patterns in the talk of women friends. In Graddol D., Maybin J., Stierer B. (Eds.), *Researching language and Literacy in Social Context*. Multilingual Matters, Clevedon: Open University, p.177-192.
- Cobb T. (2014). A Resource Whih-list for Data-Driven Learning in French. In Tyne H., André V., Benzitoun C., Boulton A., Greub Y. (Eds.). *French through corpora: ecological and data-driven perspectives in French language studies*. Cambridge : Cambridge Scholars Publishing, p.257-292.
- Cobb T. (1999). Breadth and depth of lexical acquisition with hands-on concordancing. *Computer Assisted Language Learning*, 12(4), P.345-360.
- Cobb T., Boulton A. (2015). Classroom applications of corpus analysis. In Biber D., Reppen R. (Eds.). *The Cambridge handbook of English corpus linguistics*. Cambridge: Cambridge University Press, p.478–497.
- Cohen M. (1956). *Pour une sociologie du langage*. Paris : Albin Michel.
- Conrad S. (2004). Corpus linguistics, language variation and language teaching. In Sinclair J. (dir.). *How to Use Corpora in Language Teaching*. Amsterdam: Johns Benjamins, p.67-85.
- Conrad S. (2000). Will corpus linguistics revolutionize grammar teaching in the 21st century? *Tesol Quarterly*, 34(3), p.548-560.
- Couper-Kuhlen E., Selting M. (2018). *Interactional Linguistics. Stuying Language in Social Interaction*. Cambridge : Cambridge University Press.
- Coveney A. (2003). Le redoublement du sujet en français parlé : une approche variationniste. In Berit Hansen A., Mosegaard Hansen M.-B. (eds). *Structures linguistiques et interactionnelles du français parlé*. Copenhagen: Museum Tusculanum Press University of Copenhagen, p.111-143.
- Coveney A. (1996). *Variability in spoken French: A sociolinguistic study of interrogation and negation*. Exeter: Elm Bank Publications.
- Cuq J.-P., Gruca I. (2017). *Cours de didactique du français langue étrangère et seconde*. 4^{ème} édition. Grenoble : Presses universitaires de Grenoble.
- Danon-Boileau L., Morel M.-A. (1997). Question, point de vue, genre, style : les noms prepositionnels en français contemporain. *Faits de langues*, 9, p.192–200.
- De Fornel M. (1988). Constructions disloquées, mouvement thématique et organisation préférentielle dans la conversation. *Langue française*, 78, p.101-123.
- Debaisieux J.-M. (2009). Des documents authentiques oraux aux corpus : un défi pour la didactique du FLE. *Mélanges CRAPEL*, 31, p.35-56.
- Debaisieux J.-M. (2001). Vous avez dit « inachevé »... De quelques modes de construction du sens à l'oral. *Le français dans le monde. Recherches et applications. Numéro spécial : Oral : variabilité et apprentissage*, p.53-62.

- Debaisieux J.-M. (2000). Discours oral et didactique du FLE. *Le français dans le monde. Recherches et applications. Numéro spécial : Une didactique des langues pour demain*, p. 143-158.
- Debaisieux J.-M., Boulton A. (2007). Alors la question c'est...? Questions pragmatiques et annotation pédagogique des corpus. *Cahiers de l'AFLS*, 13(2), p.31-59.
- Dejours C. (2016). *Situations du travail*. Paris : Presses Universitaires de France.
- Dejours C. (2006). *L'évaluation du travail à l'épreuve du réel : critique des fondements de l'évaluation*. Paris : Éditions Quae.
- Dejours C. (1998). Souffrance en France. La banalisation de l'injustice sociale. Paris : Éditions du Seuil.
- Dejours C. (1995). *Le facteur humain*. Paris : Presses Universitaires de France.
- Demazière D., Zune M. (2019). Enquêter auprès de chômeurs exclus par décision administrative : repérer des invisibles, écouter des inaudibles, restituer des expériences. *Recherches Qualitatives*. Hors-série. Numéro 23, p.7-21.
- Detey S. (2010). Normes pédagogiques et corpus oraux en FLE : le curseur apprenabilité/acceptabilité et la variation phonético-phonologique dans l'espace francophone. In Bertrand O., Schaffner I. (dir.). *Quel français enseigner ? La question de la norme dans l'enseignement/apprentissage*. Palaiseau : Éditions de l'École Polytechnique, p.155-168.
- Detey S., Durand J., Laks B., Lyche C. (éds.) (2010). *Les variétés du français parlé dans l'espace francophone : ressources pour l'enseignement*. Paris : Ophrys.
- Detey S., Lyche C., Tchobanov A., Durand J., Laks B. (2009). Ressources phonologiques au service de la didactique de l'oral : le projet PFC-EF. *Mélanges CRAPEL*, 31, p.223-236.
- Detey S., Racine I., Kawaguchi J., Eychenne J. (dir.) (2016). *La prononciation du français dans le monde. Du natif à l'apprenant*. Paris : Clé international.
- Deulofeu J. (2000). Genre et typologie. Introduction. In Bilger M. (ed.). *Corpus. Méthodologie et applications linguistiques*. Paris : Honoré Champion, p.219-224.
- Di Vito S. (2018). Apprendre le FLE à partir d'un corpus. *Les langues modernes*, 3, p.51-60.
- Divoux A. (2018). La linguiste et les menuisiers, éthique et prise en compte du facteur humain dans une enquête sociolinguistique en entreprise. In Déléant L., Filet J., Jeanson L. (eds). *Questionner la recherche, contributions de jeunes chercheurs aux systèmes complexes*. Nancy : Éditions Universitaires de Lorraine, p.277-291.
- Divoux A. (2017). Corpus oraux et sociolinguistique des interactions au travail : de la constitution à l'analyse. *Actes des 9èmes Journées Internationales de la Linguistique de corpus*, p.10-13. En ligne : <https://jlc2017.univ-grenoble-alpes.fr/Contenu/LivretJLC2017.pdf#page=13>
- Drew P., Heritage J. (Eds) (1992). *Talk at Work. Interaction in Institutional Settings*. Cambridge: Cambridge University Press.
- Duda R., Esch E. et Laurens J. P. (1972). Documents non didactiques et formation en langues. *Mélanges pédagogiques*, p.1-48.
- Duda R., Riley P. (Eds.). (1990). *Learning Styles*. Nancy : Presses universitaires de Nancy.
- Duda R., Tyne H. (2010). Authenticity and autonomy in language learning. *Bulletin suisse de linguistique appliquée*, 92, p.86-106.

- Durán R., McCool G. (2003). If This Is French, Then What Did I Learn in School? *The French Review*, Vol. 77, No. 2, p.288-299. En ligne : <https://www.jstor.org/stable/3132776>
- Durand J., Laks B., Lyche C. (2002). La phonologie du français contemporain : usages, variétés et structure. In Pusch C., Raible W. (éds). *Romanistische Korpuslinguistik-Korpora und gesprochene Sprache/Romance Corpus Linguistics – Corpora and Spoken Language*. Tübingen : Gunter Narr Verlag, p.93-106.
- Duranti A. (1997). *Linguistic Anthropology*. Cambridge: Cambridge University Press.
- Durive L. (2016). *Compétence et activité de travail*. Paris : L'Harmattan.
- Ellis N. (2001). Constructions, chunking, and connectionism: the emergence of second language structure. In Doughty C., Long M. (dir.). *Handbook of second language acquisition*. Oxford: Blackwell. p. 63-103.
- Eme E., Nantes N., Delliaux C. (2011). Analyse cognitive et linguistique de l'illettrisme : bilan des études et implications pour la formation. *L'orientation scolaire et professionnelle*, 40/3, p.1-23.
- Eme E., Reilly J., Almecija Y. (2009). Compétences narratives et communicatives chez des personnes en situation d'illettrisme. *Revue européenne de psychologie appliquée*, 59, p.123-138.
- Encrevé P. (1977). Présentation. Linguistique et sociolinguistique. *Langue française*, 34, p.3-16.
- Etienne C., Jouin E. (à paraître, 2019). Constituer des ressources pédagogiques pour enseigner le français oral à partir des recherches menées en interaction. In Gajo L., Luscher J.-M., Racine I., Zay F. (eds). *Variation, plurilinguisme et évaluation en FLE*. Bern : Peter Lang.
- Falzon P. (1991). Les activités verbales dans le travail. In Amalberti R., Montmollin M. de, Theureau J. (Eds). *Modèles en analyse du travail*. Liège : Mardaga, p.229-250.
- Filliettaz L. (2002). *La parole en action. Éléments de pragmatique psycho-sociale*. Laval : Editions Nota Bene.
- Filliettaz L. (2004a). Négociation, textualisation et action. La notion de négociation dans le « modèle genevois » de l'organisation du discours ». In Grosjean M., Mondada L. (Eds). *La négociation au travail*. Lyon : Presses Universitaires de Lyon, p.69-96.
- Filliettaz L. (2004b). Présentation. *Cahiers de linguistique française. Les modèles du discours face au concept d'action*. 26, p.9-23.
- Fleischman S. (1998). Des jumeaux du discours. *La linguistique*, p.31-47.
- Fligelstone S. (1993). Some reflections on the question of teaching, from a corpus linguistics perspective. *ICAME journal*, 17, p.97-109.
- Gadet F. (2003a). *La variation sociale en français*. Paris : Ophrys.
- Gadet F. (2003b). Derrière les problèmes méthodologiques du recueil des données. *Texte !*
En ligne : <http://www.revue-texto.net/index.php?id=609>
- Gadet F. (2001). Enseigner le style. *Le français dans le monde. Recherches et applications. Numéro spécial : Oral : variabilité et apprentissage*, p.63-68.
- Gadet F. (2000). Des corpus pour (ne)... pas. In Bilger M. (ed.). *Corpus. Méthodologie et applications linguistiques*. Paris : Honoré Champion, p.156-167.

- Gadet F., Wachs S. (2015). Comparer des données de corpus : évidence, illusion ou construction ? *Langage et société*, 154, 4, p.33-49. doi:10.3917/lis.154.0033.
- Garfinkel H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs (N.J.) : Prentice-Hall.
- Gavioli L. (2001). The learner as researcher: Introducing corpus concordancing in the classroom. In Aston G. (ed.). *Learning with corpora*, Houston: Athelstan, p.108-137.
- Germain C., Netten, J. (2004). Facteurs de développement de l'autonomie langagière en FLE/FLS. *Alsic*, 7, p.55-69. En ligne : <https://journals.openedition.org/alsic/2280>
- Gilbert M. (2003). *Marcher ensemble. Essais sur les fondements des phénomènes collectifs*. Paris : Presses Universitaires de France.
- Gilmore A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, 40(2), p.97-118.
- Gilmore A. (2004). A comparaison of textbook and authentic interactions. *ELT Journal*, 58(4), p.363-374.
- Giroud A., Surcouf C. (2016). De « Pierre, combien de membres avez-vous ? » à « Nous nous appelons Marc et Christian » : réflexions autour de l'authenticité dans les documents oraux des manuels de FLE pour débutants. *SHS Web of Conferences* 27. CMLF 2016. En ligne: <https://doi.org/10.1051/shsconf/20162707017>
- Goffman E. (1987). *Façons de parler*. Paris : Les Éditions de minuit.
- Goffman E. (1974). *Les rites d'interaction*. Paris : Les Éditions de Minuit.
- Gohard-Radenkovic A., Jakavonyte-Staskuviene D., Skakova A. (2015). L'éducation plurilingue et les « approches plurielle » au service de quoi et au profit de qui ? Histoire d'un désenchantement. In Adami H., André V. (eds). *De l'idéologie monolingue à la doxa plurilingue : regards pluridisciplinaires*. Bern : Peter Lang, p.171-218.
- Golopentia S. (1988). Interaction et histoire conversationnelle. In Cosnier J., Gelas N., Kerbrat-Orecchioni C. (Dir.) *Échanges sur la conversation*. Paris : Éditions du CNRS, p.69-81.
- Gombert J-É. (1996). Activités métalinguistiques et acquisition d'une langue. *Acquisition et interaction en langue étrangère*, 8, p.41-55.
- Goodwin C. (1979). The interactive construction of a sentence in natural conversation. In Psathas G. (ed). *Everyday language: Studies in ethnomethodology*, New York : Irvington, p.97-121.
- Goody J. (2007). *Pouvoirs et savoirs de l'écrit*. Paris : La dispute.
- Goody J. (1979). *La raison graphique. La domestication de la pensée sauvage*. Paris : Les Éditions de Minuit.
- Gougenhein G., Michéa R., Rivenc P., Sauvageot A. (1964). *L'élaboration du français fondamental*. Paris : Didier.
- Gremmo M.-J., Holec H. (1990). La compréhension orale : un processus et un comportement. *Le français dans le monde. Recherches et applications. Numéro special : Acquisition et utilisation d'une langue étrangère*, p.30-40.
- Grignon C., Passeron J.-C. (1989). *Le savant et le populaire. Misérabilisme et populisme en sociologie et en littérature*. Paris : Galimard.

- Grünhage-Monetti M., Halewijn E. & Holland C. (2004). *ODYSSEUS : la deuxième langue sur le lieu de travail. Les besoins linguistiques des travailleurs migrants : l'organisation de l'apprentissage des langues à des fins professionnelles*. Strasbourg : Conseil de l'Europe.
- Grünhage-Monetti M., Holland C. & Szlabewski-Cavus P. (2005). *TRIM: Training for the integration of migrants and ethnic workers into the labor market and local community*. Baltmannsweiler : Schneider Verlag.
- Guespin L., Marcellesi J.-B. (1986). Pour la glottopolitique. *Langages*, 83, p.5-34.
- Guillot M.-N. (1999). *Fluency and its Teaching*. Clevedon: Multilingual Matters.
- Guinamard I. Jouin-Chardon E., Rispail M., Traverso V., Thai T. D. (dir.) (2015). *Langues parlées, interactions sociales. Une variété d'usages pour l'apprentissage du français*. Paris : L'Harmattan.
- Guinamard I, Zhanf C. (2015). L'utilisation d'une base de corpus de français parlé dans un contexte universitaire chinois : confrontation aux méthodes existantes. In Guinamard I. Jouin-Chardon E., Rispail M., Traverso V., Thai T. D. (dir.). *Langues parlées, interactions sociales. Une variété d'usages pour l'apprentissage du français*. Paris : L'Harmattan, p.179-201.
- Gülich E., Mondada L. (2001). Analyse conversationnelle. *Lexikon der Romanistischen Linguistik*, Vol I, 2, p.196-250.
- Gumperz J. (1982). *Discourse Strategies*. Cambridge: Cambridge University Press.
- Gumperz J., Hymes D. (Eds) (1972). *Directions in Sociolinguistics. The Ethnography of Communication*. Oxford : Basil Blackwell.
- Groupe ICOR, Guinamard I. (2015). « Du coup » dans l'interaction orale en français : description de ses usages situés à partir d'une base de données multimédia, et considérations didactiques. In Guinamard I., Jouin-Chardon E., Rispail M., Traverso V., Trinh D.T. (dir.). *Langues parlées, interactions sociales. Une variété d'usages pour l'apprentissage du français*. Paris : L'Harmattan, p.131-153..
- Haliday (1985). *Spoken and Written Language*. Oxford : Oxford University Press.
- Henry S., Campione E., & Véronis J. (2004). Répétitions et pauses (silencieuses et remplies) en français spontané. *Actes des XXVè Journées d'Etude sur la Parole (JEP'2004)*, p.1-4.
- Hermoso A. (2009). "Personnellement" et "franchement" : Deux attitudes énonciatives. *Langue française*, 161(1), p.23-38.
- Higgins J, Johns T. (1984). *Computers in Language Learning*. Londres : Collins.
- Holec H. (2000). L'apport historique du CRAPEL. *Le français dans le monde. Recherches et applications. Numéro spécial : Une didactique des langues pour demain*, p.175-179.
- Holec H. (1990). Des documents authentiques, pour quoi faire. *Mélanges pédagogiques*, p.65-74.
- Howard M. (2010). S'appropriier les normes sociolinguistiques en langue étrangère – quel rôle pour l'enseignant ? In Bertrand O., Schaffner I. (dir.). *Quel français enseigner ? La question de la norme dans l'enseignement/apprentissage*. Palaiseau : Éditions de l'École Polytechnique, p.283-296.
- Hubault F. (1995). A quoi sert l'analyse de l'activité en ergonomie ? *Performances Humaines et Techniques*, Hors-Série, p.79-85.
- Hunston S. (2002). *Corpora in applied linguistics*. Cambridge : Cambridge University Press.

- Hutchins E. (1990). The technology of team navigation. In Galegher J., Kraut B., Egidio C. (Eds). *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work*. Hillsdale : Lawrence Erlbaum, p.191-220.
- Hymes D. (ed.) (1964). *Language in Culture and Society : A Reader for Linguistics and Anthropology*. New York : Harper and Row.
- Hymes D. (1972a). Models of the interaction of language and social life. In Gumperz J., Hymes D. (Eds), *Directions in Sociolinguistics. The Ethnography of Communication*. Oxford : Basil Blackwell, p.35-71.
- Hymes D. (1972b). On Communicative Competence. In Pride J. B., Holmes J. (eds). *Sociolinguistics*. Harmondsworth : Penguin Books, p.269-293.
- Hymes D. (1974). *Foundations in Sociolinguistics. An Ethnographic Approach*. Philadelphia : University of Pennsylvania Press.
- Hymes D. (1984). *Vers la compétence de communication*. Paris : Hatier-Crédif.
- Janot S. (2017). *Constitution d'un corpus multimédia pour travailler la compréhension orale. Étude menée auprès d'un public migrant*. Mémoire de Master sous la direction de V. André, Université de Lorraine.
- Jeanneret T. (1999). *La coénonciation en français*. Bern : Peter Lang.
- Jefferson G. (1973). A Case of Precision Timing in Ordinary Conversation: Overlapped Tag-Positioned Address Terms in Closing Sequences. *Semiotica*, 9, p.47-96.
- Johns T. (1986). Micro-Concord: A language learner's research tool. *System*, 14(2), p.151-162.
- Johns T. (1988). Whence and whither classroom concordancing? In Bongaerts P., de Haan P., Lobbe S., Wekker H. (eds.) *Computer Applications in Language Learning*. Dordrecht : Foris, p.9-27.
- Johns T. (1991). Should you be persuaded: Two samples of data-driven learning materials. In Johns T., King P. (dir.). *Classroom Concordancing, English Language Research Journal*, 4, p.1-16.
- Johns T. (1997). Contexts: The background, development and trialling of a concordance-based CALL program. In Wichmann A., Fligelstone S., McEnery T., Knowles G. (eds). *Teaching an Language Corpora*. Harlow: Addison Wesley Longman, p.100-115.
- Kamber A. (2011). Contexte et sens : utilisation d'un corpus écrit dans l'enseignement/apprentissage du FLE. *Travaux neuchâtelois de linguistique*, 55, p.199-218.
- Kamber A. Dubois M. (2016). Corpus, grammaire et français langue étrangère : une concordance nécessaire. *Linguistik online*, 78(4), p.3-9.
- Kamber A., Skupien C. (2009). Les documents radiophoniques dans l'enseignement de la compréhension orale. *Mélanges CRAPEL*, 31, p.172-189.
- Kerbrat-Orecchioni C. (Ed.) (2014). *S'adresser à autrui : les formes nominales d'adresse dans une perspective interculturelle*. Chambéry : Université de Savoie.
- Kerbrat-Orecchioni C. (Ed.) (2010). *S'adresser à autrui. Les formes nominales d'adresse en français*. Chambéry : Université de Savoie.
- Kerbrat-Orecchioni C. (2005). *Le discours en interaction*. Paris : Armand Colin.
- Kerbrat-Orecchioni C. (2001a). *Les actes de langage dans le discours*. Paris : Nathan.

- Kerbrat-Orecchioni C. (2001b). « Je voudrais un p'tit bifteck » La politesse à la française en site commercial. *Les Carnets du Cediscor*, 7, p.105-118.
- Kerbrat-Orecchioni C. (1994). *Les interactions verbales*. Tome III. Paris : Armand Colin.
- Kerbrat-Orecchioni C. (1992). *Les interactions verbales*. Tome II. Paris : Armand Colin.
- Kerbrat-Orecchioni C. (1991). *La question*. Lyon : Presses universitaires de Lyon.
- Kerbrat-Orecchioni C. (1990). *Les interactions verbales*. Tome I. Paris : Armand Colin.
- Kerbrat-Orecchioni C. (1989). Le principe d'interprétation dialogique. *Cahiers de Praxématique*, 13, p.43-58.
- Kerbrat-Orecchioni C., Traverso V. (2004). Types d'interactions et genres de l'oral. *Langages*, 153, p.41-51.
- Kerr B. (2013). Grammatical description and classroom application. Theory and practice un data-driven learning. *Bulletin VALS-ASLA*, 97, p.17-39.
- Krashen S. (1985). *The Input Hypothesis: Issues and Implications*. Longman, New York.
- Krashen S. D., Terrell T. D. (1983). *The natural approach: Language acquisition in the classroom*. New York: Pergamon Press.
- Kübler N., Hamilton C. E. (2018). L'exploitation des corpus numériques dans les formations d'anglais scientifique en ligne : une étude de cas. *Lidil. Revue de linguistique et de didactique des langues*, 58, p.1-21. En ligne : <https://journals.openedition.org/lidil/5155>
- Labov W. (1969/1978). *Le parler ordinaire. La langue des ghettos noirs des États-Unis*. Trad. Franç. Paris : Les Éditions de Minuit.
- Labov W. (1976). *Sociolinguistique*. Paris : Les Éditions de Minuit.
- Lacoste M. (2001). Peut-on travailler sans communiquer ? In Borzeix A., Fraenkel B. (Eds). *Langage et travail. Communication, cognition, action*. Paris : CNRS Éditions, p.21-53.
- Lahire B. (1995). *Tableaux de famille. Heurs et malheurs scolaires en milieux populaires*. Paris : Le Seuil.
- Landure C. (2013). Corpus et dictionnaires : prise de conscience linguistique chez des apprenants LANSAD. *Bulletin suisse de linguistique appliquée*, 97, p.163-178.
- Landure C. (2011). Data-Driven Learning : Apprendre et enseigner à contre-courant. *Mélanges CRAPEL*, 32, p.163-178.
- Langbach V. (2014). *Analyse et mesure des insécurités langagières chez des adultes en situation d'insertion*. Thèse de doctorat sous la direction de Adami H. et André V. Université de Lorraine.
- Lebon L. (2018). *Los Serrano à la rescousse de l'apprentissage de l'accord en compréhension orale ? Étude menée auprès d'un public adolescent débutant d'espagnol langue étrangère en milieu institutionnel*. Mémoire de Master Sciences du langage et didactique des langues. Université de Lorraine.
- Lebre-Peytard M., Lieutaud S., Beacco J.-C., Malandain J.-L. (1977). *Le document oral brut dans la classe de français : la transcription de documents sonores authentiques*. Paris : Bureau d'études pour les langues et les cultures.
- Leech G. (1997). Teaching and language corpora: A convergence. In Wichmann A., Fligelstone S., McEnery T., Knowles G. (dir.). *Teaching and Language Corpora*. Harlow : Addison Wesley Longman, p. 1-23.

- Leńko-Szymańska A. (2017). Training teachers in data-driven learning : Tackling the challenge. *Language Learning & Technology*, 21(3), p.217-241.
- Leńko-Szymańska A., Boulton A. (eds) (2015). *Multiple Affordances of Language Corpora for Data-driven Learning*. Amsterdam : Johns Benjamins.
- Leplat J, Hoc J.-M. (1983). Tâche et activité dans l'analyse psychologique des situations. *Cahiers de psychologie cognitive*, 3, p.49-63.
- Leray M., Tyne H. (2016). Homophonie et maîtrise du français écrit : apport de l'apprentissage sur corpus. *Linguistik online*, 78(4), p.131-150.
- Lerner G. H. (1996). On the "semi-permeable" character of grammatical units in conversation: conditional entry into the turn space of another speaker. In Ochs E., Schegloff E. A., Thompson S. A. (Eds), *Interaction and Grammar*. Cambridge: Cambridge University Press, p.238-276.
- Lerner G. H. (1991). On the syntax of sentence-in-progress. *Language and Society*, 20, p.441- 458.
- Liu D., Jiang P. (2009). Using a corpus-based lexicogrammatical approach to grammar instruction in EFL and ESL contexts. *Modern Language Journal*, 93, p.61-78.
- Liu D., Lei L. (2017). *Using Corpora for Language Learning and Teaching*. Alexandria (VA): TESOL International Association.
- Lonchamp J. (2003). *Le travail coopératif et ses technologies*. Paris : Lavoisier.
- Longhi J. (coord) (2017). Linguistique de corpus appliquée. *Études de linguistique appliquée*, 188.
- Lotfi A., Baude O. (2005). Du Français Fondamental aux ESLO. Grand corpus de français parlé. Bilan historique et perspectives de recherche. *Cahiers de linguistique*, 33, p.131-146.
- Loufrani C. (1981). Le locuteur collectif ou locuteur tout court. *Recherches sur le français parlé*, 3, p.215-243.
- Mahmoudian M. (1998). Problèmes théoriques du travail de terrain. *Cahiers de l'ILSL*, 10, p.7-22.
- Malinowski B. (1967/1985). *Journal d'ethnographie*. trad. franç. Paris : Éditions du Seuil,
- Mangiante J.-M., Parpette C. (2011). *Le français sur objectif universitaire*. Grenoble : Presses Universitaire de Grenoble.
- Mangiante J.-M., Parpette C. (2004). *Le français sur objectif spécifique : de l'analyse des besoins à l'élaboration d'un cours*. Paris : Hachette FLE.
- Marcellesi C. (1976). Norme et enseignement du français. *Cahiers de Linguistique Sociale*, numéro 1. En ligne dans les Archives de Glottopol : <http://glottopol.univ-rouen.fr/archives.html>, p.7-10.
- Marcellesi J.-B. (1980). De la crise de la linguistique à la linguistique de la crise : la sociolinguistique. *La Pensée*, n° 209/janvier, p.4-21.
- Marui I., Nishijima Y., Noro K., Reinelt R., Yamashita H. (1996). Concepts of Communicative Virtues (CCV) in Japanese and German. In Hellinger M., Ammon U. (eds). *Contrastive Sociolinguistics*. New York: Mouton de Gruyter, p.385-409.
- Marx K. (1872). *Le Capital*. Traduction de M. J. Roy, entièrement révisée par l'auteur. Paris : M. Lachâtre. En ligne : <https://gallica.bnf.fr/ark:/12148/bpt6k1232830.r>

- Mauranen A. (2004). Spoken corpus for ordinary Learner. In Sinclair J. (ed.). *How to Use Corpora in Language Teaching*. Amsterdam: John Benjamins, p.89-105.
- Maurer B. (2011). *Enseignement des langues et construction européenne. Le plurilinguisme, nouvelle idéologie dominante*. Paris : Éditions des archives contemporaines.
- McCarthy M. (2008). Accessing and interpreting corpus information in the teacher education context. *Language Teaching*, 41(4), p.563-574.
- McCarthy M., O'Keeffe A. (2010). Historical perspective: What are corpora and how they evolved? In O'Keeffe A., McCarthy M. (eds). *The Routledge Handbook of Corpus Linguistics*. Abingdon: Routledge, p.3-13.
- McEnery T., Xiao R., & Tono Y. (2006). *Corpus-based language studies: An advanced resource book*. Londres : Routledge.
- Mélanges CRAPEL (2009). Des documents authentiques oraux aux corpus : questions d'apprentissage en didactique des langues, 31. En ligne : <http://www.atilf.fr/spip.php?rubrique584>
- Meunier F., Gouverneur C. (2009). New types of corpora for new educational challenges: Collecting, annotating and exploiting a corpus of textbook material. In Aijmer K. (Ed.). *Corpora and language teaching*. Amsterdam: John Benjamins, p. 179–201.
- Mondada L. (2008). La transcription dans la perspective de la linguistique interactionnelle. In Bilger M. (coord.). *Données orales. Les enjeux de la transcription*. Perpignan : Presses Universitaires de Perpignan. p.78-110.
- Mondada L. (2006). La pertinence du contexte : Contributions de l'ethnométhodologie et de l'analyse conversationnelle. *Verbum*, tome XXVIII, n° 2-3.
- Mondada L., Dubois D. (2004). Temporalité, séquentialité et multimodalité au fondement de l'organisation de l'interaction. Le pointage comme pratique de prise de tour. *Cahiers de linguistique française*, 26, p.269-292.
- Mondada L. (2002). Pour une linguistique interactionnelle. In Santacrose M. (Ed.). *Faits de langue - Faits de discours. Données, processus et modèles. Qu'est-ce qu'un fait linguistique ?* Vol 2. Paris : L'Harmattan, p.95-136.
- Mondada L. (2000). Les effets théoriques des pratiques de transcription. *LINX*, 42, p.131-149.
- Mondada L. (1998). Technologies et interactions dans la fabrication du terrain du linguiste. *Cahiers de l'ILSL*, 10, p39-68.
- Mougeon R., Rehner K. (2001). Acquisition of sociolinguistic variants by French immersion students: The case of restrictive expressions and more. *The Modern Language Journal*, 85/3, p.398-415.
- Mourlhon-Dallies F. (2008). *Enseigner une langue à des fins professionnelles*. Paris : Les Éditions Didier.
- Nassau G. (2016). *Les émotions en entretien de conseil dans un dispositif d'apprentissage de langue auto-dirigé : une analyse des interactions entre apprenant et conseillère*. Thèse de doctorat sous la direction de Bailly S. et André V. Université de Lorraine.
- Noda H. (2018). Enjeux intersubjectifs et interactifs. Autour des marqueurs : *hein, quoi, n'est-ce pas*. In Richard E. (éd.). *Des organisations dynamiques de l'oral*. Bern : Peter Lang, p.287-305.

- O'Keeffe A., McCarthy M. (ed.) (2010). *The Routledge Handbook of Corpus Linguistics*. New York: Routledge.
- Ochs E. (1979). Transcription as theory. In Ochs E., Schieffelin B. (Eds). *Developmental Pragmatics*. New York: Academy Press, p.43-72.
- Oloff F. (2018). L'hétéro-répétition comme validation des complétions collaboratives. Analyse séquentielle et multimodale de séquences de co-construction. In Richard E. (éd.). *Des organisations dynamiques de l'oral*. Bern : Peter Lang, p.267-285.
- Oxford R. (1990). *Language Learning Strategies: What Every Teacher Should Know*. New York: Newbury House Publisher.
- Park K., Kinginger C. (2010). Writing/thinking in real time: Digital video and corpus query analysis. *Language Learning & Technology*, 14(3), p.31-50.
- Parpette C. (1998). Le discours oral : des représentations à la réalité. In Springer C. (éd.). *Les linguistiques appliquées et les Sciences du langage. Actes du 2^e colloque de Linguistique Appliquée*. Strasbourg : COFDELA Éditions, p.126-134.
- Paternostro R. (2017). Peut-on enseigner la variation ? In Tyne H., Bilger M., Cappeau P., Guérin E. (dir.). *La variation en question(s)*. Bruxelles : Peter Lang, p.279-290.
- Paternostro R. (2016). *Diversité des accents en enseignement du français. Les parlers jeunes en région parisienne*. Paris : L'Harmattan.
- Payet J.-P. (2011). L'enquête sociologique et les acteurs faibles. *Sociologies, La recherche en actes, Champs de recherche et enjeux de terrain*. En ligne : <http://journals.openedition.org/sociologies/3629>
- Payet J.-P., Giuliani F., Laforgue D. (dir.) (2008). *La voix des acteurs faibles. De l'indignité à la reconnaissance*. Rennes : Presses Universitaires de Rennes.
- Peacock M. (1997). The effect of authenticity materials on the motivation of EFL Learners. *ELT Journal*, 51-2, p.144-156.
- Pekarek Doehler S. (2009). Démythifier les compétences : vers une pratique écologique d'évaluation. In Galatanu O., Pierrard M., Van Raemdonck D. (dir.). *Construction du sens et acquisition de la signification linguistique dans l'interaction*. Berne : Peter Lang, p.19-37.
- Pekarek Doehler S. (2006). Compétence et langage en action. *Bulletin Vals-Asla*, 84, p.9-45.
- Pérez-Paredes P., Sánchez Tornel M., Alcaraz Calero J. M., Aguada Jiménez P. (2011). Tracking learners' actual uses of corpora: Guided vs non-guided corpus consultation. *Computer Assisted Language Learning*, 24(3), p.233-253.
- Pescheux M. (2007). *Analyse de pratiques enseignantes : mémento pour une ergonomie didactique en FLE/S*. Paris : L'Harmattan.
- Quaglio P. (2009). *Television dialogue: The sitcom Friends vs. natural conversation*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Quan Z. (2016). Introducing 'mobile DDL (data-driven learning)' for vocabulary learning: An experiment in the context of EAP (English for Academic Purposes). *Journal of Computers in Education*, 3(3), p.273-287.
- Quignard M., Ursi B., Rossi-Gensane N., André V., Baldauf-Quilliatre H., Etienne C., Plantin C., Traverso V. (2016). Une méthode instrumentée pour l'analyse multidimensionnelle des tonalités émotionnelles dans l'interaction. *Congrès Mondial de Linguistique*

- Française – CMLF 2016*, SHS Web of Conference, 27, p.1-19. En ligne : <https://www.shs-conferences.org/>
- Racine I., Detey S., Zay F., Kawaguchi Y. (2012). Des atouts d'un corpus multitâches pour l'étude de la phonologie en L2 : l'exemple du projet « Interphonologie du français contemporain » (IPFC). In Kamber A., Skupens Dekens C. (éds). *Recherches récentes en FLE*. Berne : Peter Lang, p.1-19.
- Racine I., Zay F., Detey S., Kawaguchi Y. (2011). De la transcription de corpus à l'analyse interphonologique : enjeux méthodologiques en FLE. In Col G., Osu S. N. (eds). *Transcrire, écrire, formaliser*. Rennes : Presses Universitaires de Rennes, p. 13-30
- Ravazzolo E., Traverso V., Jouin E., Vigner G. (2015). *Interactions, dialogues, conversations : l'oral en français langue étrangère*. Paris : Hachette.
- Regan V. (1996). Variation in French interlanguage. A longitudinal study of sociolinguistic competence. In Bayley R., Preston D. R. (dir.). *Second Language Acquisition and Linguistic Variation*. Amsterdam: John Benjamins, p.177-201.
- Renaud P. (2004). Parole *dans* et parole *sur* les pratiques : pratiques langagières et sociolinguistique. In Légise I. (dir.). *Pratiques, langues et discours dans le travail social. Écrits formatés, oral débridé*. Paris : L'Harmattan, p.89-135.
- Richard E. (éd.), *Des organisations dynamiques de l'oral*. Bern : Peter Lang.
- Riegel M., Pellat J.-C., Rioul R. (1994). *Grammaire méthodique du français*. Paris : Presses Universitaires de France.
- Riley P. (2007). Language, Culture and Identity. An Ethnolinguistic Perspective. London : Continuum.
- Riley P. (2001). « Allô – je parle à qui ? » Salutation, communion phatique et négociations d'identités sociales. *Le français dans le monde. Recherches et applications. Numéro spécial : Oral : variabilité et apprentissage*, p.87-96.
- Riley P. (2000). « Je vous ai compris » : aspects ethnographiques de la compréhension. *Le français dans le monde. Numéro spécial : Une didactique des langues pour demain*, p.79-94.
- Rivens Mompean A., Eisenbeis M. (2009). Autoformation en langues : Quel guidage pour l'autonomisation ? *Les cahiers de l'Acedle*, 6(1), p. 221-244. <https://journals.openedition.org/rdlc/2204>
- Römer U. (2006). Pedagogical applications of corpora: Some reflections on the current scope and a wish list for future developments. *Zeitschrift für Anglistik und Amerikanistik*, 54(2) : p.121-134.
- Rosier L. (2002). Genre : le nuancier de sa grammaticalisation. *Travaux de linguistique*, (1), p.79-88.
- Rossari C., Beaulieu-Masson A., Cojocriu C., Razgouliaeva A. (dir.). *Les états de la question*. Montréal : Nota Bene.
- Roux S. (2010). La transparence du voile. Critique de l'anonymisation comme impératif déontologique. In Laurens S., Neyrat F. (Coord.). *Enquêter : de quel droit ? Menaces sur l'enquête en sciences sociales*. Brousseux : Editions du Croquant, p.139-153.
- Sacks H. (1992). *Lectures on Conversation. Edited by Gail Jefferson. (1964-1972)*. Oxford: Basil Blackwell, 2 volumes.

- Sacks H., Schegloff E. A., Jefferson G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language*, 50, p.696-735.
- Saville-Troike M. (2002). *The Ethnography of Communication : An Introduction* (Third edition). Oxford : Blackwell.
- Schegloff E. A. (2007). Sequence organization in interaction. Cambridge : Cambridge University Press.
- Schmidt R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), p.129-158.
- Schmied J. (2006). Corpus linguistics and grammar learning: Tutor versus learner perspectives. In Braun S., Kohn K., Mukherjee J. (dir.). *Corpus Technology and Language Pedagogy: New Resources, New Tools, New methods*. Frankfurt: Peter Lang, p.87-106.
- Schwartz O. (2011). Peut-on parler des classes populaires ? *La vie des idées*. En ligne : <https://laviedesidees.fr/Peut-on-parler-des-classes.html>
- Schwartz Y. (2000). *Le paradigme ergologique ou un métier de Philosophe*. Toulouse : Octares.
- Searle J. (1969/Trad. française 1972). *Les actes de langage*. Paris : Hermann.
- Sinclair J. (2005). Corpus and Text – Basic Principles. In Wynne M. (eds), *Developing Linguistic Corpora: a Guide to Good Practice*. Oxford: Oxbow Book. En ligne : <http://www.ahds.ac.uk/creating/guides/linguistic-corpora/index.htm>
- Sinclair J. (ed.) (2004). *How to Use Corpora in Language Teaching*. Amsterdam: John Benjamins.
- Surcouf C., Ausoni A. (2018). Création d'un corpus de français parlé à des fins pédagogiques en FLE : la genèse du projet FLORALE. *Études en didactique des langues*, 31, p.71-91. En ligne : <http://www.lairdil.fr/revue-edl-issn-2258-1510-626.html>
- Surcouf C., Giroud A. (2016). À quelle langue accède l'apprenant ? Examen critique du traitement de l'oral dans les premières leçons de manuels de français langue étrangère pour débutants. *Linguistik online*, 78(4), p.11-27.
- Terrail J.-P. (2009). *De l'oralité. Essai sur l'égalité des intelligences*. Paris : La Dispute.
- Teston-Bonnard S., Blanc N., Colon de Carvajal I., Markaki V. (coord.) (2017). *Recherches en linguistique et en didactique. Apports de l'analyse des corpus et des interactions*. Limoges : Lambert-Lucas.
- Thomas A. (2002). La variation phonétique en français langue seconde au niveau universitaire avancé. *Acquisition et interaction en langue étrangère*, 17, p.101-121.
- Thomas A., Granfeldt J., Jouin-Chardon E., Etienne C. (2016). Conversations authentiques et CECR : compréhension globale d'interactions naturelles par des apprenants de FLE. *Cahiers de l'AFLS*, 20(1), p.1-44.
- Timmis I. (2015). *Corpus Linguistics for ELT. Research and Practice*. London : Routledge.
- Todd R. (2001). Introduction from self-selected concordances and self-correction. *System*, 29(1), p.91-102.
- Tognini-Bonelli E. (2001). *Corpus linguistics at work*. Amsterdam: John Benjamins.
- Traverso V. (2016). *Décrire le français parlé en interaction*. Paris : Éditions Ophrys.

- Traverso V. (2004a). Grille d'analyse des discours interactifs oraux. In Beacco J.-C., Bouquet S., Porquier R. (eds). *Niveau B2 pour le français (utilisateur/apprenant indépendant). Textes et références*. Paris : Didier, p.119-148.
- Traverso V. (2004b). Un exemple d'analyse d'un genre interactif oral : la conversation à bâtons rompus. *Niveau B2 pour le français (utilisateur/apprenant indépendant). Textes et références*. Paris : Didier, p.149-164.
- Traverso V. (2001). Interactions ordinaires dans les petits commerces : éléments pour une comparaison interculturelle. *Langage et société*, 95(1), p.5-31.
- Traverso V. (1999). *L'analyse des conversations*. Paris : Nathan.
- Traverso V. (1996). *La conversation familière*. Lyon : Presses Universitaires de Lyon.
- Tyne H. André V., Benzitoun C., Boulton A., Greub Y. (éd.) (2014). *French through corpora: ecological and data-driven perspectives in French language studies*. Cambridge: Cambridge Scholars Publishing.
- Tyne H. (2012). Corpus work with ordinary teachers: Data-driven learning activities. In Boulton A., Thomas J. (eds.). *Input, Process and Product : Developments in Teaching and Language Corpora*. Brno: Masaryk University Press, p.114-129.
- Vialleton E., Lewis T. (2014). Reconsidering the authenticity of speech in French language teaching: theory, data, methodology, and practice. In Tyne H., André V., Benzitoun C., Boulton A., Greub Y. (Eds.). *French through corpora: ecological and data-driven perspectives in French language studies*. Cambridge : Cambridge Scholars Publishing, p.293-316.
- Vion R. (1992). *La communication verbale. Analyse des interactions*. Paris : Hachette Supérieur.
- Vladimirska E. (2016). *Entre le dire et le monde : le cas du marqueur discursif genre*. In Bat-Zeev Shyldkrot H., Adler S., Asnes M. (eds). *Nouveaux regards sur l'approximation et la précision*. Paris : Honoré Champion. p.195-208.
- Widdowson H. (1998). Context, community and authentic language. *TESOL Quarterly* 32/4, p.705-716.
- Williams G. (2005). Introduction. In Williams G. (ed.). *La linguistique de corpus*. Rennes : Presses Universitaires de Rennes, p.13-18.
- Winkin Y. (2001). Terrains urbains : l'héritage de l'École de Chicago. *Acta Romanica BAsiliensia*, 13, p75-78.
- Wynne M. (2005). Preface. In Wynne M. (eds). *Developing Linguistic Corpora: a Guide to Good Practice*. Oxford: Oxbow Book. En ligne : <http://www.ahds.ac.uk/creating/guides/linguistic-corpora/index.htm>
- Yaguello M. (1998). *Petits faits de langue*. Paris : Éditions du Seuil.
- Zarifian P. (1996). *Travail et communication. Essai sociologique sur le travail dans la grande entreprise industrielle*. Paris : Presses Universitaires de France.
- Zorzi D. (2001). The pedagogic use of spoken corpora: Learning discourse markers in Italian. In Aston G. (ed.) *Learning with corpora*. Houston: Athelstan, p.85-107.

Table des illustrations (figures et tableaux)

Figure 1 : Schéma - Intrications et influences dans les interactions verbales.....	18
Figure 2 : Illustration du schéma avec l'analyse d'une pratique interactionnelle	28
Figure 3 : Illustration du schéma avec une partie des analyses de la question	33
Figure 4 : La lampe robotisée Psyphine.....	35
Figure 5 : Illustration du schéma avec l'analyse d'une interaction homme/lampe robotisée	37
Figure 6 : Différents types de corpus dans l'approche sur corpus (André à paraître 2019) 68	
Figure 7 : Chaîne de traitement des corpus de données interactionnelles	70
Tableau 1 : Présentation synthétique d'exemples d'énonciations conjointes dans TCOF ..	80
Tableau 2 : Présentation synthétique d'exemples d'énoncés inachevés qui permettent de poser des questions.....	88
Figure 8 : Les sous-catégories de la catégorie « Outils de la conversation » dans FLORALE	120
Figure 9 : Extrait de la liste d'exemples qui illustrent la sous-catégorie « Quels petits mots pour organiser ce qu'on dit ? » dans FLORALE.....	121
Figure 10 : Extrait des résultats de la recherche de « genre » dans le concordancier de FLORALE	121
Figure 11 : Page d'accueil de CLAPI-FLE	122
Figure 12 : Fiche pédagogique dans CLAPI-FLE	123
Figure 13 : Page d'accueil du site FLEURON	126
Figure 14 : Explication de « le truc c'est que » dans CLAPI-FLE	130
Figure 15 : Premier exemple pour illustrer le fonctionnement de « le truc c'est que » dans CLAPI-FLE	131
Figure 16 : Extrait des résultats de « moi je » dans le corpus Plus belle la vie (JConc).....	139
Figure 17 : Extrait des résultats de la recherche de « moi je » dans FLEURON	139
Figure 18 : Explication des dislocations dans CLAPI-FLE	141
Figure 19 : L'explication de l'apport pragmatique de la dislocation dans CLAPI-FLE et premier exemple	142
Figure 20 : Début de l'interaction de la ressource « Paiement en espèces » dans FLEURON	144
Figure 21 : Interaction dans la catégorie « Vie sur le campus » de FLEURON	147
Figure 22 : Interaction dans la catégorie « Vie en dehors du campus » de FLEURON.....	148
Figure 23 : Interaction dans la catégorie « Démarches administratives » de FLEURON....	149
Figure 24 : Interaction à la SNCF dans FLEURON	152
Figure 25 : Interaction dans la catégorie « Vie en dehors du campus » de FLEURON.....	153
Figure 26 : Interaction dans la catégorie « Culture et lieux de culture » de FLEURON.....	154
Figure 27 : Complémentarité des ressources pour une exploitation didactique.....	156
Figure 28 : Extrait des résultats de la recherche de « prendre » via TXM dans TCOF	163
Figure 29 : Extrait de la transcription qui comprend l'occurrence « prendre » en dix-huitième position dans la figure 28	163
Figure 30 : Extrait des résultats de la recherche de « parce que » via JConc dans TCOF ..	164
Figure 31 : Extrait de la transcription qui comprend l'occurrence « parce que » en sixième position dans la figure 30	165
Figure 32 : Extrait des résultats de la recherche de « du coup » via JConc dans TCOF	168
Figure 33 : Première partie des métadonnées du corpus Infirmier_aud_14 dont sont extraites les occurrences 41 et 42 dans la liste de la figure précédente	169

Figure 34 : Deuxième partie des métadonnées du corpus Infirmier_aud_14, celles qui concernent les locuteurs (L1)	170
Figure 35 : Deuxième partie des métadonnées du corpus Infirmier_aud_14, celles qui concernent les locuteurs (L2)	170
Figure 36 : Concordancier de FLEURON – espace de recherche	172
Figure 37 : Extrait de la liste des concordances de « même » dans FLEURON, en cochant « mot exact »	173
Figure 38 : Extrait du résultat de la recherche de « meme » dans FLEURON, sans cocher « mot exact »	173
Figure 39 : Retour dans la ressource vidéo en cliquant sur la troisième occurrence dans la figure 38.....	174
Figure 40 : Extrait des résultats de la recherche de « même » dans le concordancier de FLEURON	181
Figure 41 : Notes d'un apprenant lors de l'analyse des occurrences de même dans le corpus FLEURON	181
Figure 42 : Liste des résultats de la recherche de « voulais » dans le concordancier de FLEURON	183
Figure 43 : Exemple pour illustrer l'imparfait de politesse dans CLAPI-FLE	185
Figure 44 : Liste des occurrences de « ainsi » dans le corpus FLEURON	186
Figure 45 : Extrait de la liste des résultats de la recherche de « ainsi » dans le corpus Est Républicain sur ScienQuest	186
Figure 46 : Première ligne du résultat de la recherche de « depuis » dans FLEURON.....	193
Figure 47 : Prise de vue du film de l'expérimentation	193
Tableau 3 : Liste des 20 items les plus recherchés dans FLEURON	199
Figure 49 : Liste des occurrences de « je vous en prie » dans FLEURON	204
Figure 50 : Extrait de la liste des occurrences de « donc » dans FLEURON	209
Figure 51 : Liste des occurrences de « dessus » dans FLEURON	210
Tableau 4 : Présentation de cinq adultes francophones natifs en insécurité langagière ..	215
Figure 52 : Photos d'un exercice de transcription avec des personnes en insécurité langagière	216
Figure 53 : Extrait de la liste des occurrences de « pas » dans L'Est Républicain	218
Figure 54 : Extrait de la liste des occurrences de « pas » dans OFROM	219
Figure 55 : Extrait de la liste des occurrences de « c'est » dans FLEURON	222
Figure 56 : Extrait de la liste des occurrences de « sait » dans FLEURON	223
Figure 57 : Liste des occurrences de « s'est » dans FLEURON	224
Figure 58 : Extrait de la liste des occurrences de « peux » dans FLEURON	226
Figure 59 : Extrait de la liste des occurrences de « peut » dans FLEURON	226
Figure 60 : Extrait de la liste des occurrences de « peu » dans FLEURON	226
Figure 61 : Complémentarité des ressources pour apprendre à interagir	240