

Un corpus multimédia pour apprendre à interagir en situations universitaires en France

Virginie André

▶ **To cite this version:**

Virginie André. Un corpus multimédia pour apprendre à interagir en situations universitaires en France. Troisième colloque international de l'ATPF " Enseigner le français : s'engager et innover ", Oct 2017, Bangkok, Thaïlande. hal-01862752

HAL Id: hal-01862752
https://hal.science/hal-01862752

Submitted on 29 Aug 2018

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.

Un corpus multimédia pour apprendre à interagir en situations universitaires en France

Virginie ANDRÉ
Université de Lorraine et CNRS

Résumé

La France accueille de plus en plus d'étudiants étrangers depuis une quinzaine d'années. Elle en a accueilli 310 000 en 2015, dont 17,5% étaient originaires des pays de la zone Asie-Océanie. Le nombre d'étudiants choisissant de venir faire un séjour universitaire en France ne cesse de croitre depuis plus de 15 ans. L'agence Campus France prend en charge ces étudiants pour les aider à trouver une université d'accueil, une bourse, un logement mais également à comprendre la culture française, à comprendre le système universitaire ou encore à suivre des cours magistraux. Cependant, aucune structure n'apprend aux futurs étudiants à interagir dans les situations de communication auxquelles ils vont participer lors de leur arrivée en France. C'est ce manque que nous cherchons à combler avec le projet FLEURON dont l'acronyme signifie « Français Langue Etrangère Universitaire : Ressources et Outils Numériques ». Le site Internet associé au projet (https://apps.atilf.fr/fleuron/) met à disposition des apprenants un corpus multimédia d'interactions authentiques se déroulant à la scolarité, à la bibliothèque universitaire, au restaurant universitaire mais aussi à la banque ou au guichet de vente de billets de transport en commun. Ces interactions sont essentielles pour les étudiants qui arrivent en France et qui doivent gérer leur installation, résidentielle et universitaire. Le site FLEURON est un dispositif d'apprentissage du français, en autoformation ou en formation guidée, qui permet d'éviter les problèmes d'intégration des étudiants. Le dispositif pédagogique s'inscrit dans le cadre du Français sur Objectif Universitaire et de l'approche originale de l'apprentissage sur corpus. Les apprenants visionnent et interrogent le corpus multimédia, notamment grâce à un concordancier qui leur permet d'observer les pratiques langagières des locuteurs. Cet outil, emprunté à la linguistique de corpus, est simplifié pour que les apprenants et les enseignants puissent l'utiliser à des fins d'apprentissage.

Mots-clés : Français sur objectifs universitaires,apprentissage sur corpus,interactions verbales authentiques

Introduction

Depuis une quinzaine d'années, le nombre d'étudiants étrangers venant faire un séjour universitaire en France ne cesse de croître. Au cours de l'année universitaire 2015-2016, la France a accueilli 310 000 étudiants étrangers[85]. Pour faciliter l'arrivée et l'accueil de ces étudiants, l'agence Campus France met en place un certain nombre d'actions. Néanmoins, l'intégration langagière des étudiants étrangers peut davantage être soutenue. En effet, lors de leur arrivée dans une université française, les étudiants sont souvent démunis d'un point de vue langagier dans la mesure où ils n'ont pas appris à participer aux situations de communication qu'ils rencontrent.

Notre objectif est de montrer de quelle façon un corpus multimédia de situations de communication authentiques peut aider les étudiants étrangers à acquérir des compétences sociolangagières facilitant leur intégration. Le projet Français Langue Etrangère Universitaire – Ressources et Outils Numériques (FLEURON) propose un dispositif d'apprentissage du français grâce à une interface numérique (André, 2016), disponible à l'adresse : https://apps.atilf.fr/fleuron/. Ce site Internet met à disposition un corpus multimédia que les apprenants peuvent exploiter, avec différents outils, à des fins d'apprentissage de la langue. Les recherches menées dans le cadre de ce projet montrent que la méthodologie innovante qu'il propose facilite l'appropriation sociolinguistique du français.

Dans un premier temps, nous expliquerons le contexte migratoire des étudiants étrangers en France, en insistant sur leurs besoins d'intégration. Dans un deuxième temps, nous préciserons le cadre scientifique de notre étude visant à permettre aux étudiants, apprenants de français, d'acquérir les compétences sociolangagières nécessaires pour interagir en situations universitaires en utilisant un corpus multimédia de situations de communication authentiques. Enfin, dans un troisième temps, nous montrerons comment ce corpus peut être exploité à des fins didactiques. Nous illustrerons cette démarche avec des exemples d'utilisation.

1. La mobilité étudiante : un contexte et des besoins

L'origine du projet FLEURON est liée à l'accroissement de la mobilité étudiante depuis les quinze dernières années. À Nancy (en France), lieu de naissance du projet, un constat régulier a été fait lors d'interactions entre le personnel, administratif et enseignant de l'université, et les étudiants étrangers : dans de nombreuses situations de la vie étudiante, des problèmes de communication sont fréquents.

[85] Les derniers chiffres disponibles sont produits par l'agence Campus France pour l'année universitaire 2015/2016. Ils sont disponibles à l'adresse suivante : http://ressources.campusfrance.org/publi_institu/etude_prospect/chiffres_cles/fr/chiffres_cles.pdf

1.1 Un contexte propice

Le contexte de la mobilité étudiante accrue est propice au développement de projets scientifiques visant à faciliter l'intégration des étudiants étrangers dans le pays qu'ils ont choisi pour faire un séjour universitaire. C'est dans ce cadre que le projet FLEURON voit le jour dès les années 2000. Les déplacements étudiants ne sont pas récents mais ils ne cessent de croître. En 2002, deux millions d'étudiants étaient en mobilité internationale, c'est-à-dire qu'ils effectuaient un séjour universitaire dans un pays étranger. En 2015, ils étaient plus de quatre millions.

La France est le premier pays, non anglophone, d'accueil d'étudiants étrangers. Ils étaient 310 000 en 2015. En Europe, les échanges internationaux ont été facilités au sein de l'espace européen avec le programme ERASMUS (*European Community Action Scheme for the Mobility of University Students*). Ce programme a été créé en 1987, il permettait aux étudiants de faire une partie de leurs études (entre 3 et 12 mois par cycle) dans un des 33 pays européens, ou entrés dans un processus d'adhésion à l'Union européenne. La France est le deuxième pays d'accueil des étudiants européens ERASMUS, derrière l'Espagne. En 2015, près de 30 000 étudiants européens étaient accueillis en France. Le programme ERASMUS est aujourd'hui un programme d'échange mondial. Rebaptisé ERASMUS+, il rassemble tous les programmes de mobilité internationale. Les étudiants européens ne représentent pas les plus importants flux d'étudiants en mobilité. Les quatre premiers pays d'origine des étudiants étrangers en France sont le Maroc avec 36 768 étudiants (ce qui représente 11,9 % des étudiants étrangers), la Chine avec 28 043 étudiants (9,1 %), l'Algérie avec 22 660 étudiants (7,3 %) et la Tunisie avec 12 077 étudiants (3,9 %). En dehors de l'Union Européenne, la zone Asie-Océanie représente la troisième zone mondiale d'origine des étudiants étrangers en France avec plus de 50 000 étudiants, après l'Afrique du Nord et l'Afrique.

Les étudiants étrangers viennent faire un séjour universitaire ou la totalité de leurs études en France. Certains viennent également pour obtenir un diplôme particulier qui enrichira ou complètera leur cursus initial dans leur pays d'origine. Les étudiants étrangers viennent très majoritairement rejoindre les universités françaises mais certains intègrent également les écoles d'ingénieurs et les grandes écoles (de management, de commerce, d'arts, par exemple). Un nombre important d'étudiants en mobilité s'inscrit en doctorat, 41 % des doctorants en France sont étrangers.

Ainsi, la mobilité étudiante est en constante progression. De plus en plus d'étudiants migrent pour effectuer des séjours universitaires en dehors de leur pays d'origine. À l'instar de la majorité des pays du monde, la France encourage ses étudiants à partir à l'étranger pour découvrir d'autres systèmes universitaires mais aussi d'autres langues et d'autres cultures. La France cherche également à attirer de plus en plus d'étudiants étrangers dans

ses universités (et ses écoles). En quinze ans, le nombre d'étudiants étrangers a doublé. Le ministère de l'Enseignement Supérieur et de la Recherche souhaite qu'il double de nouveau d'ici 2020. En 2013, la ministre de l'époque, Geneviève Fioraso, déclarait au Sénat que la France devait adopter une nouvelle politique d'attractivité universitaire et scientifique. Cette politique, destinée à attirer des étudiants mais également des enseignants-chercheurs et des chercheurs, comprend l'amélioration des conditions d'accueil et de séjour. Le projet FLEURON entre dans ce processus d'amélioration en proposant un dispositif innovant d'apprentissage du français qui permet de répondre aux besoins langagiers des étudiants étrangers arrivant en France.

Des besoins langagiers

Si 90 % des étudiants étrangers en France sont satisfaits de leur séjour universitaire, selon Campus France, il reste certains problèmes d'intégration liés à la maîtrise du français. En effet, ces problèmes sont fréquents bien que les étudiants étrangers se préparent à leur séjour universitaire en France. Les pays d'origine dispensent des cours de français, souvent spécialisés et sur objectifs universitaires. Il existe aussi des sites d'apprentissage du français en ligne. Les alliances françaises proposent également une offre de formations en français langue étrangère (FLE) dans tous les pays du monde. Certains étudiants suivent des cours intensifs de français dans des centres de langues en France, dès leur arrivée et avant de commencer leur année universitaire. Néanmoins, certaines compétences sociolangagières manquent aux étudiants et ces manques nuisent à une bonne intégration. Les échecs universitaires des étudiants migrants sont souvent dus à des problèmes d'intégration.

Les formations en français sur objectif universitaire (FOU) cherchent à faire acquérir aux étudiants étrangers des compétences langagières disciplinaires. Celles-ci concernent la compréhension d'un cours magistral ou plus généralement d'un discours scientifique, la participation à un exposé à l'oral, la construction d'une argumentation, la rédaction d'un écrit scientifique ou encore la prise de parole dans un cours. Le FOU concerne également l'acquisition de savoir et de savoir-faire méthodologiques tels que la compréhension d'une consigne lors d'un examen, la rédaction de réponses à des questions évaluatives, la restitution d'un travail ou encore la prise de notes de cours ou l'élaboration d'un résumé. Mangiante et Parpette (2011) ont très bien montré les enjeux du FOU ainsi que les domaines langagiers concernés.

Notre objectif n'est pas l'enseignement et l'apprentissage des compétences langagières précédemment décrites et habituellement traitées par le FOU. Nous avons recueilli des constats de différents collègues sur les problèmes de communication des étudiants étrangers en dehors des cours. Le personnel administratif et les enseignants-chercheurs remarquent régulièrement les problèmes d'interaction de la plupart des étudiants en mobilité. L'objectif du projet FLEURON est donc de proposer un dispositif d'apprentissage

du français utilisé dans des situations de communication rencontrées par des étudiants en France et en dehors des cours. Au-delà des compétences langagières nécessaires à la réalisation d'un cursus universitaire, les étudiants ont d'autres besoins langagiers qui leur permettent de s'intégrer en tant qu'étudiant au sein d'un système universitaire ainsi qu'en tant que citoyen migrant dans une société d'accueil.

Les besoins langagiers des étudiants que nous ciblons avec le dispositif FLEURON concernent donc ceux qui sont nécessaires pour participer aux situations de communication universitaires. Plus précisément, nous permettons aux étudiants de se préparer à participer aux situations administratives qui leur permettront d'organiser et de gérer leur séjour en France. Avec FLEURON, les étudiants peuvent par exemple apprendre à interagir avec le personnel de la scolarité, à comprendre les démarches administratives qu'ils doivent entreprendre, à poser les bonnes questions, à échanger avec le personnel de la bibliothèque, à emprunter un livre, à interagir avec les enseignants ou encore à prendre les transports en commun, et plus généralement à comprendre et à participer à la vie étudiante en France. Les compétences sociolangagières acquises grâce à ce dispositif permettent aux étudiants de participer aux situations de communication qu'ils rencontrent de façon efficace, d'un point de vue linguistique, pragmatique, interactionnel et culturel.

2. Cadre scientifique du projet

Le projet FLEURON est né de l'avancée des recherches en linguistique et en didactique des langues. Il s'inscrit à l'interface même de la linguistique et de la didactique. Nous verrons tout d'abord comment un corpus, qui est originellement un outil de la linguistique, peut être utilisé à des fins d'apprentissage. Ensuite, nous montrerons que les analyses (socio)linguistiques actuelles permettent de faciliter l'apprentissage de la langue.

2.1 L'apprentissage sur corpus

Les recherches récentes en didactiques des langues s'orientent vers l'utilisation des corpus à des fins didactiques (Johns, 1991), appelée également l'apprentissage sur corpus (Boulton & Tyne, 2014). Cette nouvelle méthodologie trouve son origine dans l'utilisation des documents authentiques, la linguistique de corpus et les nouvelles technologies.

Tout d'abord, l'importance de l'utilisation de documents authentiques, c'est-à-dire non fabriqués à des fins didactiques, n'est plus à démontrer. Depuis les années 1970, nous pouvons lire des plaidoyers en faveur des documents authentiques (Duda et al., 1972). Abé et al. (1979, p. 3) proposent une définition générale : « on peut définir le document authentique, de façon négative, comme tout document produit à des fins autres que l'apprentissage d'une langue seconde. En termes positifs, c'est un énoncé produit dans une situation réelle de communication ». Depuis, le caractère authentique des documents a été discuté (voir par exemple Widdowson, 1998 ; Adami, 2009 ; Chambers, 2009) sans remettre

en cause l'utilité des documents authentiques dans les processus d'acquisition et d'apprentissage des langues (Holec, 1990). L'acquisition de compétences langagières est favorisée par l'exposition à la langue cible, à condition que l'apprenant soit exposé à de la *vraie* langue, réellement produite par des locuteurs, dans des situations de communication réelles. Ce principe didactique remet en cause l'utilisation de manuels qui proposent des interactions jouées ou imaginées ou des écrits inventés ou fabriqués. Dans le cadre de notre étude, nous nous intéresserons uniquement aux données orales et multimodales proposées comme support d'apprentissage de la langue.

Ensuite, notre étude propose d'utiliser des corpus à des fins d'apprentissage. Nous pouvons définir un corpus comme « a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language » (Sinclair, 1996, p. 4). Il existe plusieurs corpus de français parlé qui sont des échantillons plus ou moins grands de la langue française, sélectionnés en fonction d'un certain nombre de critères extralinguistiques. Nous pensons par exemple aux corpus TCOF, CLAPI, ESLO 1 et 2, PFC[86]. Ces corpus rassemblent des enregistrements audio et/ou vidéo, accompagnés de leur transcription et de leurs métadonnées qui nous donnent des informations sur la situation de communication (locuteurs, genre de discours, lieu d'enregistrement, etc.). L'analyse du français parlé et du fonctionnement de la langue parlée en interaction est permise grâce aux différents corpus oraux (voir par exemple les travaux pionniers en France de Blanche-Benveniste & Jeanjean, 1987 ; Gadet, 1989 ; Kerbrat-Orecchioni, 1990-1994).

La linguistique de corpus[87] bénéficie de l'avancée du traitement informatique des données orales. Elle s'est également développée avec les nouvelles possibilités toujours plus performantes de recueil de données. Il est en effet possible aujourd'hui d'enregistrer et de filmer simplement, avec des outils qui ne sont pas trop intrusifs, et de produire des données d'excellente qualité. De plus, les corpus oraux sont désormais facilement manipulables et stockables. Les capacités actuelles des ordinateurs et des réseaux de communication permettent le stockage, le traitement et l'échange de telles données. Les corpus sont surtout interrogeables de différentes façons. Les outils de la linguistique facilitent l'observation et l'analyse de langue et de ses usages. En effet, nous y reviendrons avec un exemple d'exploitation de corpus à des fins d'apprentissage, des outils sont créés pour observer les données de différentes façons, notamment avec des concordanciers.

Les corpus oraux sont des ensembles de documents authentiques. Ils semblent donc offrir de nouvelles possibilités et de nouvelles perspectives à la didactique des langues. Les

[86] Disponibles respectivement aux adresses suivantes : http://www.cnrtl.fr/corpus/tcof/, http://clapi.ish-lyon.cnrs.fr, http://eslo.huma-num.fr, http://www.projet-pfc.net ou sur la plateforme ORTOLANG : https://www.ortolang.fr.

[87] Nous précisons de nouveau que nous nous intéressons uniquement aux données orales mais la linguistique de corpus concerne tous les domaines de la linguistique (Williams, 2005).

enseignants de langue qui mettent en œuvre des méthodologies qui s'appuient sur des documents authentiques sont souvent confrontés au manque de données[88]. Les corpus créés par des linguistes semblent alors pouvoir combler ce manque. Néanmoins, la manipulation de ces corpus n'est pas simple que ça soit pour les enseignants ou pour des apprenants de langue. Il existe plusieurs façons d'utiliser des corpus dans un cadre didactique (André à paraître). Nous verrons un exemple d'utilisation d'un corpus et des technologies associées à des fins d'apprentissage du français avec le corpus multimédia de FLEURON.

Ainsi, notre étude dans le cadre du projet FLEURON met en synergie différentes disciplines et différentes pratiques afin de proposer une nouvelle méthodologie pour l'enseignement et l'apprentissage du français. Nous faisons également appel aux nouvelles technologies de développement Web afin de rendre disponible et accessible cette méthodologie didactique et les outils qui lui sont associés. Au-delà de ce cadre général, notre étude s'inscrit également dans un cadre d'analyse plus fin, celui de l'analyse des interactions verbales.

2.2 Analyse sociolinguistique des interactions verbales

L'utilisation des documents authentiques en didactique des langues est apparue avec l'avancée des travaux en linguistique et notamment en linguistique interactionnelle (voir les travaux américains fondateurs de Sacks et al., 1974, Sacks, 1992 et Garfinkel, 1967 ; puis en France, par exemple, Kerbrat-Orecchioni 1990-1994 ; Traverso, 1999 ; Mondada, 2002). L'analyse des mécanismes interactionnels et du fonctionnement des interactions intéresse nécessairement la didactique des langues dans la mesure où mieux une langue est décrite mieux elle est enseignée. Les compétences orales sont souvent moins travaillées que les compétences écrites par les enseignants et les apprenants de langue justement parce que l'écrit est bien décrit et l'oral beaucoup moins. C'est notamment la notion de variation qui rend l'oral plus complexe et moins figé que l'écrit.

L'analyse sociolinguistique des interactions verbales (André, 2014) permet de saisir le fonctionnement des échanges et des multiples influences qui font varier la langue en fonction des situations de communication. Cette approche cherche « à appréhender les pratiques langagières, en analysant leurs spécificités linguistiques, interactionnelles, pragmatiques et contextuelles. En d'autres termes, les pratiques langagières (c'est-à-dire les productions linguistiques et interactionnelles) actualisées par les interlocuteurs sont appréhendées comme apparaissant dans une situation de communication toujours particulière. Ces pratiques langagières permettent d'accomplir des activités langagières (c'est-à-dire des actions communicatives telles que informer, questionner, remercier, etc.),

[88] La linguistique de corpus, et notamment les études sur l'oral, a été confrontée au même problème, c'est pourquoi nous voyons depuis quelques années maintenant naître des actions de diffusion de données (TCOF, CLAPI, etc.) et de mutualisation de données (c'est le cas par exemple avec la plateforme ORTOLANG que nous avons déjà mentionnée).

elles aussi particulières et dépendantes du genre de discours (conversation, réunion de travail, débat, etc.) dans lequel elles sont produites. En résumé, cette approche exige de saisir les multiples interinfluences entre les productions ou pratiques langagières et leur contexte d'apparition » (André 2015, p. 2). Nous reproduisons notre schéma pour illustrer ces influences :

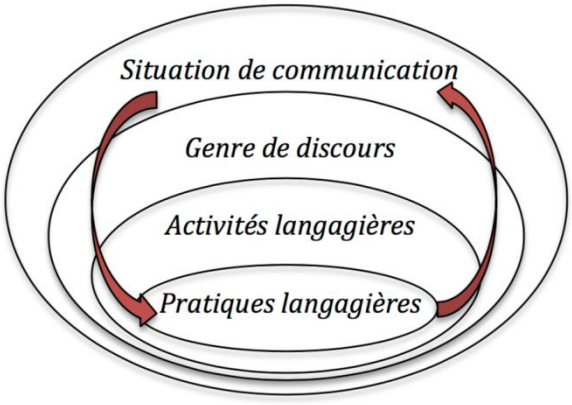


Figure 1 : Influences en interaction

Ce schéma permet d'identifier les influences présentes lors d'interactions verbales. Les pratiques langagières, les activités langagières, le genre de discours et la situation de communication sont imbriqués les uns dans les autres, chacun influençant la réalisation ou l'existence de l'autre. « Les flèches représentent la réciprocité des influences, de la situation vers les pratiques (la relation entre des interactants peut par exemple conditionner leur tutoiement) et des pratiques vers la situation (celle-ci peut être redéfinie au fur et à mesure de l'interaction, par exemple, à la suite de paroles injurieuses des locuteurs peuvent décider de rompre leur relation qui peut donc passer d'amicale à inamicale) » (*ibid*. p. 2-3). Cette façon d'appréhender la langue est en complète cohérence avec l'utilisation de corpus de documents authentiques en didactique des langues comme ressources pédagogiques.

Analyse des pratiques langagières

Dans le cadre de l'analyse sociolinguistique des interactions verbales, nous portons une attention particulière aux pratiques langagières (voir figure précédente). Ces dernières sont les réalisations effectives de la langue et donc ce que les apprenants de langue doivent pouvoir comprendre et produire. Cependant, comme nous l'avons précisé auparavant, elles varient en fonction de plusieurs éléments. Pour être compétent dans une langue, il faut savoir produire les bonnes pratiques langagières de façon appropriée. Plus précisément, cela signifie qu'un apprenant doit apprendre à prendre en compte, les activités langagières visées, le genre de discours auquel il participe et les éléments de la situation de communication afin de produire un discours approprié. La compétence en langue ne se

réduit pas à la connaissance de son vocabulaire et de ses règles de grammaire, les règles d'utilisation en contexte s'ajoutent et rendent les premières inutiles si celles-ci ne sont pas respectées. La « compétence de communication » (Hymes, 1984) est donc la capacité d'adapter son discours en fonction de la situation et de ce qu'elle comporte.

Cette conception sociolinguistique de la langue oblige les enseignants de FLE à utiliser différents documents, recueillis dans différentes situations, avec différents locuteurs, ayant des objectifs différents, etc. afin de montrer la variabilité des pratiques langagières. Nous verrons de quelle façon cet objectif didactique est rendu possible grâce à l'exploitation de corpus. Néanmoins, nous souhaitons tout de même illustrer notre approche avec des exemples de pratiques langagières qui varient d'une situation de communication à l'autre.

Par exemple, pour accomplir l'activité langagière « poser une question », différentes pratiques langagières sont possibles, le choix de leur production dépendant de la situation de communication. Dans les manuels de langue, sont mentionnés comme procédés langagiers permettant de poser une question : l'inversion du sujet et du verbe, l'utilisation d'un mot interrogatif et parfois l'ajout d'une intonation montante en fin d'énoncé. Les corpus montrent que certaines pratiques sont réservées à des genres de discours particuliers et montrent également que d'autres pratiques sont possibles. Debaisieux et Boulton (2007) ont déjà montré ce décalage entre la langue des manuels et la langue des corpus. Dans les corpus que nous avons analysés, TCOF et FLEURON, lors d'une conversation entre amis, nous ne trouvons aucune inversion sujet-verbe alors que nous pouvons en rencontrer quelques unes lors d'entretiens formels. Par contre, nous rencontrons d'autres pratiques langagières permettant de poser une question dans des extraits tels que :

Exemple 1 : deux étudiants parlent de linguistique (Corpus TCOF – How_PRO_14)[89]

|L2|> mais euh c'était plutôt l'écrit donc la linguistique de corpus euh destinée à l'apprentissage des langues < voilà
<|L1|> des > langues étrangères ou
<|L2|> bah surtout de l'anglais < hein

Dans cet extrait, nous rencontrons un phénomène assez fréquent en français. L1 pose une question avec un énoncé inachevé se finissant par la conjonction « ou ». Son interlocutrice saisit immédiatement qu'il s'agit d'une question et produit une réponse. Ce type de pratique langagière n'apparaît dans aucun manuel et sans analyse de corpus de données authentiques, il est peu probable que notre intuition de locuteur nous fasse penser à cette façon de poser une question. Nous rencontrons également un autre procédé qui pourrait être enseigné comme une stratégie de compensation à des apprenants de FLE. Voici un exemple :

[89] Les conventions de transcription sont celles du projet TCOF. Dans cet extrait, les chevrons délimitent les paroles superposées.

Exemple 2 : une locutrice raconte son mariage (Corpus TCOF – Mariage_ELO_14).
<|L1|> d'accord et euh si vous auriez quelque chose à à changer dans cet événement euh +
<|L3|> après si j'avais quelque < chose
<|L1|> de mieux >
<|L3|> à changer c'est que j'aimerais que mes parents justement euh soient (...)

Dans cet extrait, L1 produit une hésitation suivie d'une pause (+) pour signaler à son interlocutrice qu'elle a terminé la formulation de sa question et qu'elle lui laisse la parole. Ces énoncés inachevés accomplissent les effets pragmatiques escomptés par les locuteurs, ils servent à poser une question. Seule l'analyse de corpus de données orales authentiques permet d'observer ces phénomènes langagiers. En outre, l'analyse de grands corpus montre que ces phénomènes ne sont pas rares.

Nous prendrons un second exemple qui illustre bien la différence entre ce que les manuels de langue préconisent et les réels usages du français en interaction. Dans une étude précédente (André 2010), nous avons analysé les phénomènes de reprise. Nous avons montré que les reprises (allant de la répétition exacte de ses propres paroles ou de celles de son interlocuteur à la reformulation où seul le sens est repris, pas la forme) remplissent des objectifs pragmatiques particuliers et ne sont pas des lacunes ou un manque de compétence de la part des locuteurs. En effet, les reprises remplissent des objectifs pragmatiques et interactionnels tels que : poser une question, insister sur une information, marquer son accord avec un interlocuteur ou encore prendre son tour de paroles.

Depuis que l'analyse du français parlé est devenue un véritable objet d'étude, de nombreux travaux montrent les spécificités de l'oral. Ces dernières sont observables grâce aux corpus et à leurs modes d'interrogation. Les avancées de l'analyse des interactions sont utiles à la didactique des langues dans la mesure où elles orientent les apprentissages. Nous allons examiner plus précisément de quelle façon il est possible d'utiliser un corpus pour apprendre à interagir.

3 Utilisation du corpus multimédia FLEURON à des fins didactiques

Comme nous l'avons déjà expliqué, le projet FLEURON s'inscrit dans le cadre des recherches actuelles en linguistique et en didactique des langues. Nous allons maintenant préciser comment le corpus a été conçu et traité pour qu'il puisse servir à l'appropriation du français par des apprenants.

Présentation du corpus multimédia

Le corpus FLEURON comporte des vidéos de situations de communication authentiques que les étudiants étrangers sont susceptibles de rencontrer dès leur arrivée en France pour gérer

et organiser leur séjour[90]. Le corpus a été constitué (et ne cesse d'être enrichi) afin d'être un échantillon représentatif des interactions auxquelles les étudiants doivent ou peuvent participer lors de leur séjour universitaire. Le corpus FLEURON fait partie du cœur du dispositif d'apprentissage du français.

Ce corpus est composé d'interactions authentiques :

- avec le personnel administratif de l'université : scolarité, secrétariat, bibliothèque, santé, maison de l'étudiant, appariteur, service des relations internationales, centre de langues, service d'aide aux logements, service d'assistance sociale et médicale, service informatique, associations étudiantes, etc. ;
- avec des enseignants ;
- avec des services non universitaires mais indispensables à la vie étudiante : préfecture, transports en commun, lieux de culture, etc. ;
- avec des étudiants étrangers qui témoignent de leurs problèmes, surprises ou découvertes.

Les interactions sont de trois natures différentes :

- les situations de communication écologiques, c'est-à-dire qu'elles ont eu lieu en milieu naturel et qu'elles auraient eu lieu même si nous ne les avions pas filmées ;
- les situations de communication sollicitées pour obtenir des explications, elles auraient pu exister sans notre intervention ;
- les témoignages.

D'un point de vue méthodologique, les premières situations sont généralement filmées par des caméras installées de façon permanente dans le bureau de l'agent qui accepte d'être filmé. Tous les participants sont informés de l'enregistrement et peuvent refuser d'être filmés. Pour filmer les deuxièmes situations, nous nous déplaçons avec le matériel nécessaire pour réaliser les films. Enfin, pour les témoignages, nous filmons dans différents endroits, en fonction des disponibilités des étudiants.

Les vidéos sont entièrement sous-titrées et transcrites. Voici l'interface présentant un exemple de ressource sous-titrée et transcrite :

[90] Le corpus FLEURON contient également des ressources audio. Ce sont des documents que nous avons conservés pour leur pertinence. Cependant, désormais les ressources qui alimentent le site sont toutes des vidéos.

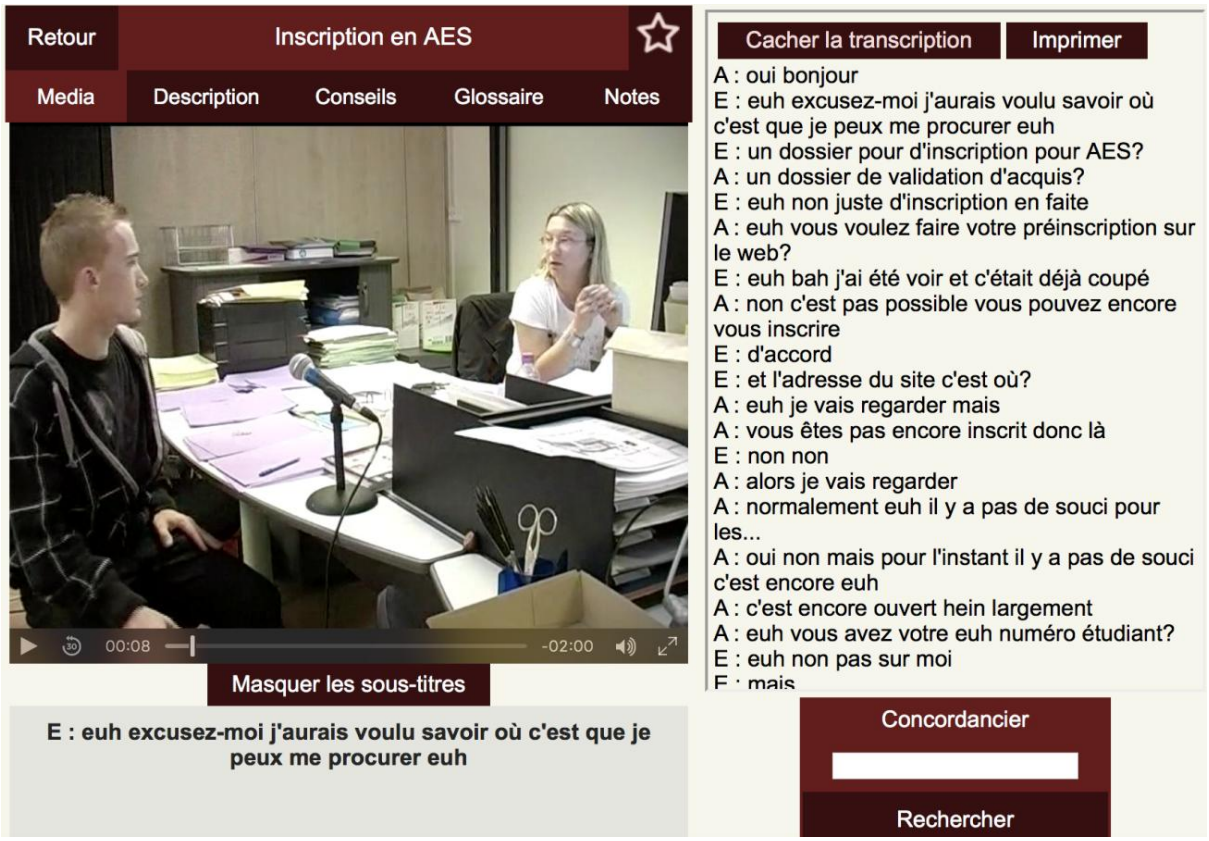


Figure 2 : Interface multimodale de la vidéo

L'apprenant peut choisir de faire apparaître ou de masquer les sous-titres ainsi que la transcription. Nous évoquerons l'utilité de ces outils dans le point suivant.

Le corpus est trié par catégorie, il en existe onze actuellement sur le site : procédure d'inscription, situations problèmes, demande d'informations, obtenir un document officiel, choisir ses cours, comprendre un itinéraire, explications du système universitaire, comprendre la vie étudiante, utiliser les transports, culture et lieux de culture et témoignages. Ces catégories sont amenées à évoluer en fonction des situations recueillies, des besoins des étudiants étrangers et des partenaires du projet.

Chaque ressource est décrite par un résumé qui présente les métadonnées utiles aux apprenants : lieu de l'interaction, locuteur, objectif de communication, etc. De plus, le titre donné à la ressource tente d'apporter également des informations pour orienter l'apprenant dans ses choix de visualisation et pour qu'il saisisse les éléments de la situation de communication. L'illustration suivante montre trois exemples de la catégorie « Culture et lieux de culture » :



Figure 3 : Titres et résumés de ressources du corpus

Lors d'une expérimentation du site FLEURON, des étudiants nous ont mentionné l'importance du choix des titres et des résumés. Nous y reviendrons dans le point suivant.

Nous souhaitons préciser également que le corpus FLEURON comporte des interactions exolingues, c'est-à-dire entre un locuteur natif du français et un locuteur non natif. C'est le cas de tous les témoignages d'étudiants mais également de certaines situations administratives, entre des agents de service français et des étudiants étrangers. Ces données sont recueillies et mises en ligne afin de montrer qu'un locuteur non natif peut participer à ces interactions. Les étudiants filmés dans ces situations sont des « étrangers compétents » (Adami et al., 2003), c'est-à-dire qu'ils réussissent à atteindre leurs objectifs communicationnels dans les interactions auxquelles ils participent. Nous pouvons préciser d'ailleurs que l'étranger compétent s'est souvent comporté en ethnographe avant de pouvoir interagir de façon appropriée. Il a observé les règles d'usages, les règles sociales et les comportements communicatifs adéquats (André & Castillo, 2005).

La totalité du corpus a été constitué à Nancy. De nombreuses situations de communication sont largement transposables dans une autre université française dans la mesure où les procédures d'inscription, de validation des acquis ou d'obtention des diplômes sont les mêmes. C'est également le cas pour les interactions à la bibliothèque ou dans les autres services administratifs de l'université ou d'ailleurs (SNCF pour les transports, préfecture, etc.), ces interactions pourraient avoir lieu dans toutes les universités et villes de France. Néanmoins, il existe des spécificités régionales. Nous pensons notamment à la variété langagière, au lexique spécifique de la région et à l'accent. En outre, si comprendre les

horaires d'ouverture d'un musée ne nécessite pas un document dans le musée que l'on souhaite visiter, il est toute fois plus facile de se projeter dans l'interaction si le musée de sa ville est représenté. Pour ces différentes raisons, nous souhaiterions enrichir le corpus avec des documents recueillis dans d'autres villes de France.

Accès au corpus et apprentissage du français

L'accès au corpus à des fins d'apprentissage de la langue n'est pas toujours aisé. Comme nous l'avons précisé, cet accès permet pourtant aux apprenants d'observer puis d'apprendre la langue en contexte, produite par des locuteurs réels. Cette confrontation à la langue permet de déconstruire le mythe du locuteur idéal, souvent visé en didactique des langues. L'apprenant constate qu'il est plus efficace, en situation de communication, de construire un énoncé approprié avec une syntaxe approximative ou non normative plutôt qu'un énoncé syntaxiquement bien construit mais inadapté à l'objectif pragmatique ou à son interlocuteur. En observant les corpus authentiques, l'apprenant apprend à construire un discours en interaction et à développer des compétences sociolangagières qui lui permettront d'interagir dans les situations qu'il va rencontrer.

Pour atteindre ces objectifs, l'utilisation de corpus est fortement appropriée. Néanmoins, cette nouvelle méthodologie s'accompagne pour que les voies d'accès au corpus ne soient pas un obstacle à l'apprentissage. Pour exploiter le corpus FLEURON, nous proposons deux voies d'accès : par ressource et par concordancier. L'accès est par conséquent non linéaire, il ne suit pas une hypothétique progression pédagogique.

3.1 Par ressource

Le corpus FLEURON est enrichi régulièrement mais reste encore de taille moyenne. Par conséquent, il est possible de naviguer dans le corpus en choisissant les documents à visualiser. Nous avons vu précédemment que les ressources du corpus étaient classées par catégories. Ces dernières représentent la première voie d'accès au corpus. L'interface présente les ressources de façon non linéaire et non ordonnée :

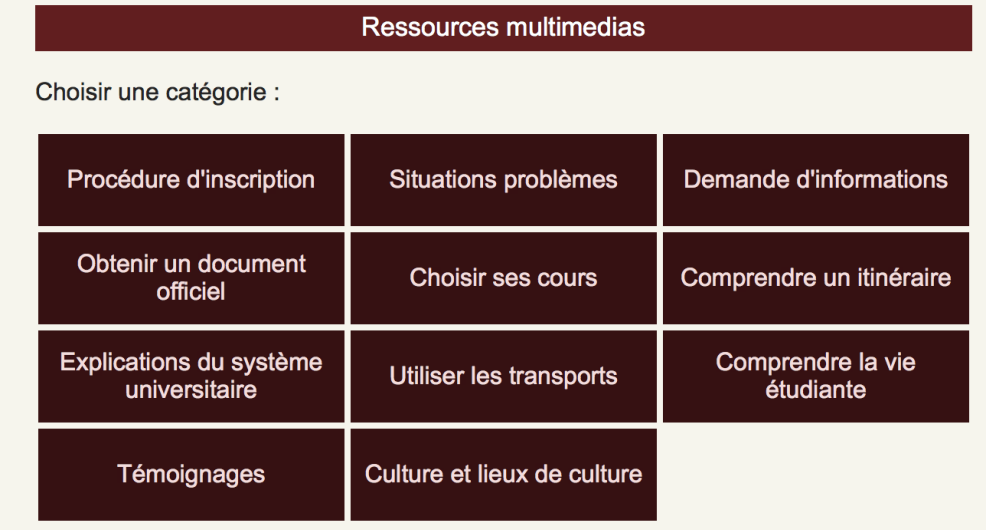


Figure 4 : Accès au corpus par catégorie

En procédant de cette façon, l'approche est davantage qualitative. L'apprenant choisit d'abord la catégorie qui l'intéresse puis le document qu'il souhaite visionner (la liste des documents avec titre et résumé apparaît une fois que la catégorie est choisie).

Une fois que le document est choisi, l'apprenant peut réaliser différentes activités. Certaines sont proposées par le dispositif sous forme de « conseils » associés à chaque ressource (voir les onglets de la figure 2). Ces conseils peuvent être par exemple :

- *Ecouter pour comprendre et dire : écoutez l'extrait et relevez la terminologie universitaire (les mots, les sigles). Si vous ne connaissez pas leurs significations, recherchez-les.*
- *Ecouter pour comprendre : essayez de relever les informations essentielles concernant la visite médicale (est-elle obligatoire ? qui vous demande de la faire ? où se déroule-t-elle ? etc.).*
- *Observer pour dire : observez la manière de se présenter et d'insister poliment mais fermement.*
- *Observer pour dire : l'étudiante exprime l'impossibilité de venir le lendemain pour régler son inscription : observez les formules qu'elle utilise (vraiment pas, pas du tout, non plus).*
- *Ecouter pour comprendre : essayez de relever les horaires d'ouverture du musée.*
- *Ecouter pour comprendre : quel est le moyen de transport utilisé ?*

L'enseignant ou l'apprenant peut choisir d'utiliser ces conseils pour orienter la visualisation du document.

D'autres activités sont possibles lors de la visualisation d'une ressource. Comme nous l'avons expliqué dans la présentation du corpus, ce dernier est intégralement sous-titré et transcrit. Les apprenants ont la possibilité de faire apparaître les sous-titres et/ou les transcriptions. Cette multimodalité aide les apprenants à comprendre les vidéos et à acquérir des compétences langagières. L'apprenant, guidé ou non par un enseignant, peut également observer, grâce à la transcription, le déroulement d'une interaction. Il peut observer l'ouverture et la clôture d'une interaction, en d'autres termes observer de quelles façons les locuteurs commencent et terminent un échange. Ensuite, l'apprenant peut observer le corps de l'interaction, de quelle façon poser des questions, demander des renseignements, se présenter, expliquer un problème, etc. dans un contexte particulier. L'apprenant devient un apprenti analyste des interactions verbales. Il explore les ressources pour observer le fonctionnement de la langue en situation.

3.2 Par concordancier

Le dispositif FLEURON permet un accès au corpus grâce à un concordancier. Cet outil emprunté à la linguistique de corpus permet de rechercher une occurrence (ou une chaîne de caractères) dans un texte. Comme tous les concordanciers qui interrogent des données orales, celui du site FLEURON va rechercher les occurrences souhaitées dans les transcriptions. L'avantage du concordancier FLEURON est qu'il permet un retour dans la ressource au moment même où l'occurrence a été prononcée. Par exemple, un apprenant recherche le mot « inscription » dans le concordancier[91] :

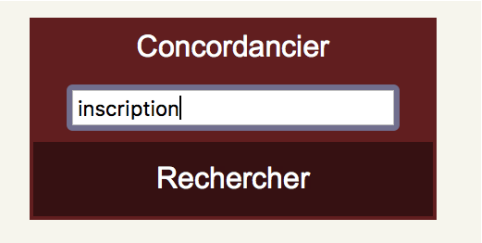


Figure 5 : Concordancier sur le site FLEURON[92]

Le résultat s'affiche sous forme de liste, l'occurrence au milieu et son contexte gauche et droit :

[91] Le mot « inscription » est actuellement le plus recherché sur le site.

[92] Nous pouvons remarquer que le concordancier est simple d'utilisation. Il est ici adapté aux besoins didactiques. Les enseignants et les apprenants qui utilisent le site ne sont pas supposés être formés à la linguistique de corpus.

A gauche du mot	Mot recherché	A droite du mot
je peux pas vous faire l'	inscription	pour la carte parce que j'ai vraiment besoin de ce document
et puis je vous ferez l'	inscription	entre deux rendez-vous
donc ce serait pour une	inscription	pour tout de suite ou pour le 2ème semestre?
et puis on fera l'	inscription	en fonction de de votre emploi du temps quand vous n'aurez pas cours
la fiche d'	inscription	aussi
quand je vais faire l'	inscription	je dois reprendre le numéro d'INE
ou c'est une	inscription	...
nous ce qu'on veut c'est	inscription	annuelle
elle peut faire une inscr-... non mais elle peut faire une	inscription	annuelle
donc il faut faire deux chèques parce qu'on fait l'	inscription	sur deux logiciels hein
donc la... la 1ère	inscription	dans l'enseignement supérieur c'était en 2006-2007 alors
donc là c'est une	inscription	pour un semestre
hein d'	inscription	il n'y a aucun euh vous n'avancez pas vous ne payez pas
alors je viens vous voir parce que j'arrive pas à faire mon	inscription	administrative
euh j'ai eu des problèmes avec mon	inscription	via le web
en... à la cellule d'	inscription	web
hum en fait euh c'est pour mon	inscription	voilà
un dossier pour d'	inscription	pour AES?
euh non juste d'	inscription	en faite
c'est vraiment un dossier d'	inscription	qu'il me faut

Figure 6 : Extrait des résultats de la requête « inscription » dans le concordancier

À partir de cette liste, l'apprenant peut faire différentes observations. Cette présentation permet une première phase d'observation, indispensable dans le processus d'apprentissage. L'apprenant peut travailler seul, en autonomie. Il peut ainsi observer « des réalisations linguistiques communicativement situées pour découvrir ce qu'il a à apprendre » (Gremmo, Riley, 1997). Il peut aussi être accompagné ou guidé par un enseignant. Nous reviendrons sur ces deux façons d'utiliser le dispositif.

Ensuite, si un apprenant souhaite observer une occurrence en contexte, il lui suffit de cliquer sur l'occurrence correspondante et la ressource apparaît. Voici un exemple lorsque l'on clique sur la huitième occurrence du mot « inscription » dans la liste précédente :

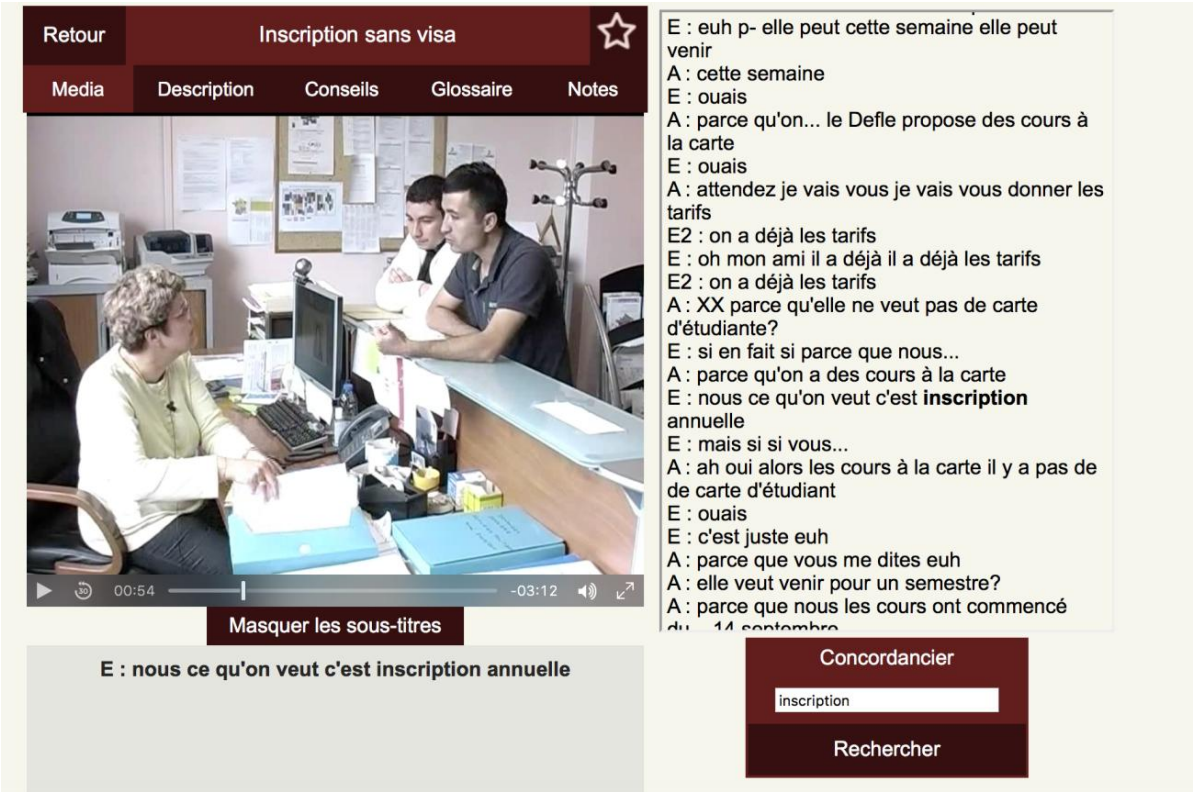


Figure 7 : Interface multimodale de la vidéo
renvoyant à l'une des occurrences du terme « inscription »

Les fonctionnalités du concordancier offrent donc la possibilité aux apprenants d'observer un mot dans son contexte textuel ainsi que dans son contexte situationnel. L'apprenant peut également remarquer des fréquences d'apparition de certains mots ou de collocations, « inscription annuelle » par exemple. Nous verrons dans le point suivant d'autres utilisations possibles du concordancier.

Expérimentations et résultats

Le site FLEURON a été expérimenté par différents apprenants dans différents pays. Les expérimentations ont porté sur l'ensemble des outils disponibles sur le site. Le site propose un glossaire, un compte personnel pour conserver des traces de son parcours de navigation, la possibilité de laisser des commentaires pour les utilisateurs et pour les administrateurs, un espace pour prendre des notes ou encore un moyen de mettre de coté des vidéos pour les visualiser plus tard. Dans cette étude, nous ne mentionnerons que les résultats des expérimentations qui concernent l'utilisation du corpus multimédia.

3.3 L'apprenant - analyste

L'objectif d'utiliser des corpus à des fins d'apprentissage semble fonctionner dans la mesure où la phase d'observation des données langagières est tout à fait remarquable (Johns 1991 ; Gavioli 2001 ; Cobb 2014). Elle va même bien plus loin que la phase d'observation telle que nous l'avions imaginée au départ. En effet, faire observer des

pratiques langagières, quelle que soit la voie d'accès au corpus, était notre première ambition avec un corpus comme celui de FLEURON. Nous avons pu constater, lors de différentes expérimentations, que cette étape est une réussite pour différentes raisons. Ainsi, l'apprenant, seul ou guidé par un enseignant, devient un observateur de la langue et des usages en situation. Il peut faire des observations de l'apparition de certaines pratiques langagières au sein d'un échange complet ou regarder les différentes actualisations et sens d'un terme dans plusieurs échanges. Il peut comparer différents emplois en fonction des locuteurs, du degré de formalité de l'échange, etc.

Le plus intéressant et peut-être le plus surprenant est que l'apprenant devient, sans s'en rendre compte, un analyste de la langue. Il est capable, souvent accompagné par un enseignant ou un formateur, de faire des analyses distributionnelles suffisamment fines pour qu'il comprenne le fonctionnement de certaines pratiques langagières, certains mots ou certaines constructions. Par exemple, avec Florence Poncet, Directrice du Département de Français Langue Etrangère à Nancy, nous avons expérimenté l'exploitation du corpus FLEURON pour faire observer à des étudiants le fonctionnement de la négation en « ne… pas » ou « pas ». Nous avons donc demandé aux étudiants d'analyser les conditions d'apparition du « ne » de négation en recherchant « pas » dans le concordancier[93]. Nos étudiants se sont littéralement pliés à l'exercice de l'analyse de la langue et on pu observer par eux mêmes les règles ou l'absence de règles qui régissent le fonctionnement de la négation.

Lors d'une autre expérimentation, suite à une question d'un apprenant sur le fonctionnement de « même », nous avons demandé aux apprenants de chercher ses occurrences dans le corpus FLEURON. Ils ont obtenu une liste d'occurrences :

[93] Dans une première étape de notre séquence pédagogique, nous avions utilisé plusieurs corpus, y compris des corpus écrits pour que les étudiants puissent faire des comparaisons.

A gauche du mot	Mot recherché	A droite du mot
qui font le	même	travail que moi en fait c'est c'est au niveau de c'est une politique des sites
en	même	temps
ah 16h bah très bien parfait il y a	même	5 minutes d'avance
c'est quand	même	plus facile pour vous
pas pour l'instant bon bah si vous avez d'autres questions	même	plus tard
ah d'accord ou	même	par fax hein c'était
là on doit avoir le same la	même	chose et
c'est pour votre sécurité hein parce que s'il y a un souci	même	pour le médecin euh
ah ici pour moi c'est la	même	chose
donc il me faut quand	même	la la carte de séjour qui est qui est expirée
ah Epinal ça vous fait loin quand	même	pour venir
oui c'est la	même	chose
je fais la	même	chose pour euh toutes les 2
vous avez... vous êtes ensemble dans la	même	résidence
si si si vous pourrez les faire quand	même	ouais
ils se retrouveront tous dans le	même	cas
d'accord bon oui il y avait juste ça à joindre il y avait peut-être quand	même	votre euh
qu'on vous donne pas non plus du	même	renseignement E : oui c'est pour ça A : donc ça bah je vous le redonne
nous avons eu quand	même	des vacances et donc euh
	même	si vous n'êtes pas inscrit vous avez la possibilité d'aller en cours
je l'ai pas c'est le	même	de mon mari
si si si vous pourrez les faire quand	même	ouais
ils se retrouveront tous dans le	même	cas
mais je viens quand	même	la première journée le vingt et un

Figure 8 : Extrait des résultats de la recherche de « même »

Avec cette liste d'occurrences, les étudiants ont procédé à une analyse distributionnelle en observant les contextes gauche et droit d'apparition du mot, les différentes collocations possibles (*même chose*, *quand même*, *même si*, *même cas*, *même + nom*, etc.) et par conséquent les différents sens de ce mot. Cette phase d'observation des pratiques langagières ainsi que de leur condition d'apparition et d'utilisation favorise l'appropriation de la langue. En outre, en donnant la possibilité à l'apprenant d'analyser lui-même les pratiques langagières, nous facilitons la compréhension du fonctionnement de la langue.

3.4 Autres résultats

Nous n'aborderons pas tous les résultats que nous avons pu conclure des différentes expérimentations de FLEURON mais nous en présenterons ici une sélection. Tout d'abord, le corpus FLEURON présente des situations de communication auxquelles les étudiants vont être confrontés, l'intérêt du corpus est alors immédiatement perçu. Dans le cadre d'une expérimentation en Chine (Varinot, 2017), des étudiants ont utilisé le site pour préparer leur séjour universitaire en France. Ils ont soulevé l'intérêt d'un tel corpus dans la mesure où aucune de ses ressources n'existe, même sur Internet[94]. Nous mentionnerons deux remarques supplémentaires faites par ces utilisateurs. Ils soulignent l'importance de se

[94] Si cette remarque est d'autant plus pertinente en Chine où la censure d'Internet accroit la difficulté de trouver des documents authentiques, il est également vrai que les ressources du corpus FLEURON sont introuvables.

rendre compte des situations réelles et insiste également sur le fait que les vidéos permettent de voir à « quoi ressemblent » des Français. Ainsi, les interactions verbales situées proposées par le corpus multimédia augmentent l'intérêt que les apprenants lui portent.

Ensuite, nous constatons que les apprenants choisissent les pratiques langagières qu'ils souhaitent apprendre et celles qu'ils ne jugent pas utiles. Ces choix sont parfois faits en fonction des locuteurs du corpus. En effet, certains apprenants ont fait le choix de n'apprendre que des pratiques langagières produites par des locuteurs de leur âge, d'autres ont souhaité apprendre des formes apparaissant dans des situations formelles en premier puis se sont intéressés aux situations moins formelles par la suite.

Enfin, l'apprentissage sur corpus est une méthodologie nouvelle, elle doit donc être présentée et expliquée aux enseignants et aux apprenants. En tant que linguiste, nous n'avons aucun problème pour utiliser le concordancier, pour imaginer quelles analyses peuvent faire les apprenants à partir des résultats en liste ou encore pour proposer une analyse sociolinguistique des interactions verbales. Cependant, nous constatons, même avec des enseignants de FLE très expérimentés, que l'application de ces méthodologies à des fins didactiques doit être accompagnée.

Conclusion

L'utilisation de corpus à des fins d'apprentissage, ou le *data-driven learning* (Johns, 1991 ; Aston, 2001), voient le jour dans des études anglo-saxonnes dès la fin des années 1980 mais elle n'en est qu'à ses débuts en France. L'utilisation de corpus multimédia pour apprendre à interagir est encore plus rare (voir l'expérience de Braun, 2007). Cette nouvelle méthodologie permet à l'apprenant d'observer les pratiques langagières en contexte puis de s'approprier ces pratiques et de se construire sa langue, en fonction de ses objectifs, de ses envies, de ses motivations et de son style. L'objectif du dispositif d'apprentissage FLEURON n'est pas de remplacer un enseignant, il apporte un complément tout en pouvant viser l'autonomisation de l'apprenant.

Notre projet propose donc une intrication entre la linguistique et la didactique. L'utilisation de corpus est née avec l'analyse du fonctionnement de la langue orale. Aujourd'hui, nous tentons d'exporter et d'adapter l'exploitation de corpus oraux et multimodaux à des fins d'apprentissage du français. Dans d'autres études (André à paraître), nous tentons de montrer comment il est possible d'utiliser des corpus constitués et traités pour des analyses linguistiques dans un cadre didactique. Dans cet article, nous nous sommes concentrée sur l'utilisation du corpus multimédia FLEURON par des apprenants de FLE ayant besoin d'interagir dans des situations universitaires en France.

Le dispositif FLEURON permet de faire face à l'hétérogénéité de ses utilisateurs. Ces derniers peuvent effectivement exploiter le corpus multimédia avec différents objectifs d'apprentissage. Les explorations qualitative, par ressource, et quantitative, par concordancier, permettent une exposition directe à la langue cible et à ses variations. Les futurs étudiants étrangers apprennent alors à interagir dans diverses situations universitaires en France grâce au corpus multimédia proposé. Ce travail permet alors de garantir à ces apprenants-migrants une intégration universitaire, sociale et citoyenne.

Références bibliographiques

ADAMI H. (2009), « Les documents authentiques dans la formation des adultes migrants : pratiques pédagogiques et contraintes institutionnelles », *Mélanges Crapel, 31*, p. 159-172.

ADAMI H., ANDRE V., BAILLY S., CASTILLO D., PONCET F. & RILEY P. (2003), « L'étranger compétent. Un nouvel objectif pour la didactique des langues étrangères » dans Groux D. & Lemeunier M. (Eds). *Une identité plurielle: mélanges offerts à Louis Porcher*, Paris, Editions L'Harmattan, p. 535-550.

ANDRE V. (à paraître), « Des corpus oraux et multimédia authentiques pour acquérir des compétences sociolangagières » dans Gajo L., Luscher J.-M., Racine I., *Variation, plurilinguisme et évaluation en FLE*, Genève, Peter Lang.

ANDRE V. (2016), « FLEURON : Français Langue Étrangère Universitaire - Ressources et Outils Numériques. Origine, démarches et perspectives ». *Mélanges Crapel*, 37, p. 69-92.

ANDRE V. (2015), « Sociolinguistique des interactions verbales : de l'analyse des situations de travail aux implications sociales », *Langage, Travail et Formation*, numéro zéro, p. 1-10, https://apps.atilf.fr/reseaultf/wp-content/uploads/2015/10/Virgine-Andre.pdf, (page consultée le 18/07/2017).

ANDRE V. (2014), « L'énonciation conjointe : trace et ressource de la construction collaborative du discours », *Congrès Mondial de Linguistique Française – CMLF 2014*, p. 1891-1904, http://www.linguistiquefrancaise.org/articles/shsconf/pdf/2014/05/shsconf_cmlf14_01323.pdf, (page consultée le 19/07/2017).

ANDRE V. (2010), « Eléments 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.

ANDRE V., CASTILLO D. (2005), « The 'Competent Foreigner' : A new model for foreign language didactics ? » dans Preisler B., Fabricius A., Haberland H., Kjaerbeck S. & Risager K., *The Consequences of Mobility : Linguistic and Sociocultural Contact Zones*, Roskilde (Denmark), Roskilde University, Department of Language and Culture, p. 154-162.

ASTON G. (ed) (2001), *Learning with corpora*, Houston, Athlestan.

BLANCHE-BENVENISTE C., & JEANJEAN C. (1987), *Le français parlé*, Paris, Didier Eruditions.

BOULTON A., TYNE H. (2014), *Des documents authentiques aux corpus. Démarches pour l'apprentissage des langues*, Paris, Didier.

BRAUN S. (2007), « Designing and exploiting small multimedia corpora for autonomous learning and teaching » dans Hidalgo E., Quereda L., Santana J. (ed), *Corpora in the Foreign Language Classroom*, Amsterdam, Rodopi, p. 31-46.

CHAMBERS A. (2009), « Les corpus oraux en français langue étrangère : authenticité et pédagogie », *Mélanges Crapel, 31*, p. 15-33.

COBB T. (2014), « A Ressource Wish-list for Data-driven Learning in French » dans 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.

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, http://afls.net/cahiers/13.2/Debaisieux.pdf, (page consultée le 18/07/2017).

DUDA R., ESCH E., LAURENS J.P., (1972), Documents non didactiques et formation en langues, *Mélanges pédagogiques*, p. 1-47.

GADET F. (1989), *Le français ordinaire*, Paris, Armand Colin.

GARFINKEL H. (1967), *Studies in Ethnomethodology*. Engelwood Cliffs, N.J., Prentice Hall.

GAVIOLI L. (2001), « The learner as researcher : Introducting corpus concordancing in the classroom » dans Aston G. (ed), *Learning with Corpora*, Houston, Athlestan, p. 108-137.

GREMMO M.-J., & RILEY, P. (1997), « Autonomie et apprentissage autodirigé : l'histoire d'une idée », *Mélanges Crapel*, *23*, p. 81-107.

HOLEC H. (1990), « Des documents authentiques, pour quoi faire ? », *Mélanges pédagogiques*, p. 65-74.

HYMES D. (1984), *Vers la compétence de communication*. Trad. franç. Paris, Hatier, Collection Langues et Apprentissage des langues.

JOHNS T. (1991), « Should you be persuaded: Two examples of data-driven learning » dans Johns T.& King P. (éds.), *Classroom Concordancing, English Language Research Journal, 4,* p. 1-16.

KERBRAT-ORECCHIONI C. (1990-1994), *Les interactions verbales*, 3 tomes, Paris, Armand Colin.

MANGIANTE J.-M., &PARPETTE C. (2011), *Le français sur objectif universitaire*, Grenoble, Presses universitaires de Grenoble.

MONDADA L. (2002), « Pour une linguistique interactionnelle » dans 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.

SACKS H. (1992), *Lectures on Conversation*. Edited by Gail Jefferson, (1964-1972). Oxford: Basil Blackwell, 2 volumes.

SACKS H., SCHEGLOFF E., JEFFERSON G. (1974), « A simplest systematics for the organization of turn-taking in conversation », *Language*, 50, p. 696-735.

SINCLAIR J. (1996), « Preliminary recommendations on corpus typology », *EAGLES Document TCWG-CTYP/P*, http://www.ilc.cnr.it/EAGLES/corpustyp/corpustyp.html, (page consultée le 18/07/2017).

TRAVERSO V. (1999), *L'analyse des conversations*, Paris, Nathan.

VARINOT K. (2017), *Développer des stratégies métadiscursives en compréhension orale grâce aux outils proposés sur le site FLEURON*, Mémoire de Master Sciences du langage et didactique des langues, sous la direction de Virginie André, Université de Lorraine.

WIDDOWSON H. (1998), "Context, community and authentic language", *TESOL Quarterly, 32/4*, p. 705-716.

WILLIAMS G. (dir) (2005), *La linguistique de corpus*, Rennes, Presses Universitaires de Rennes.