

OFROM, CORPUS ORAL DE FRANÇAIS DE SUISSE ROMANDE: UNE RESSOURCE POUR LA RECHERCHE... ET POUR L'ENSEIGNEMENT/APPRENTISSAGE DU FRANÇAIS

The aim of this paper is to offer French language teachers some instructional approaches on the use of OFROM, an online corpus of spoken French from the French-speaking region of Switzerland. The corpus data are recordings of speakers from that region and constitute a resource that is currently well-used by researchers interested in spoken French. The purpose of this contribution is to show that OFROM also offers a unique teaching resource for French teaching and learning which allows for contact with the "everyday" language spoken in this area of Switzerland.

● Laure Anne Johnsen | Uni Neuchâtel



Laure Anne Johnsen est maître d'enseignement à l'Université de Neuchâtel, où elle a soutenu son doctorat dans le domaine de la linguistique française et où elle enseigne actuellement le français langue étrangère (FLE). Ses intérêts de recherche portent sur la sémantique, la pragmatique et la syntaxe du français, notamment du français parlé, ainsi que sur l'enseignement/apprentissage du FLE.

Introduction

L'objectif de cet article est de proposer aux enseignants de français des pistes d'exploitation didactique d'OFROM, un corpus de français parlé en Suisse romande disponible en ligne¹. Les données mises à disposition proviennent d'enregistrements de locuteurs romands et constituent une ressource désormais bien exploitée par les chercheurs s'intéressant au français parlé. La présente contribution vise à montrer qu'OFROM représente également une ressource didactique unique en son genre pour aborder, dans l'enseignement/apprentissage du français, la langue «quotidienne» parlée en Suisse romande.

Dans cet article, nous présentons d'abord la base de données et la manière de l'utiliser, avant de proposer des pistes concrètes d'exploitation didactique, accompagnées de propositions d'activités sur la base d'extraits choisis.

1. OFROM: une base de français parlé par des Romands

La base de données OFROM (Oral de Français de Suisse ROMande), hébergée à l'Université de Neuchâtel (Avanzi, Béguelin & Diémoz 2012-2019), est la première ressource en ligne qui rassemble exclusivement des données de français parlé de Suisse romande, les locuteurs étant nés en Suisse et vivant dans une région de Suisse romande. Les données à disposition sont pour la plus grande partie des extraits de conversations² qu'il est possible d'écouter en ayant simultanément accès à la transcription. Les conversations peuvent être guidées par un enquêteur qui interroge un locuteur romand sur son parcours, ses intérêts, etc. ou alors plus libres, c'est-à-dire que les locuteurs discutent sans contraintes de sujets de leur choix. Quoi qu'il en soit, les personnes parlent de manière spontanée, autrement dit sans avoir préparé de réponses, et les thématiques abordées

1 Pour en savoir plus sur OFROM:



2 À ces extraits de conversation, il convient d'ajouter deux conférences universitaires.

sont variées et relativement communes.³ Mise en ligne en décembre 2012 et régulièrement alimentée par de nouveaux enregistrements, la base compte en l'état actuel 1'005'318 mots, répartis sur 64h de parole de 341 locuteurs.

1.1. Origine et buts scientifiques d'OFROM

À l'origine, la base de données a été conçue dans le but de rendre disponibles auprès de la communauté scientifique des enregistrements et transcriptions effectués en lien avec des cours de linguistique française à l'université, en vue d'étudier le français parlé en Suisse romande. Cette entreprise s'inscrit de manière plus générale dans la mouvance de la linguistique de corpus, vouée à répertorier des productions attestées des locuteurs, pour l'étude ou l'enseignement des langues. En ce qui concerne la constitution de corpus de français parlé, de plus en plus d'initiatives d'envergure se développent depuis un certain nombre d'années dans le monde francophone, traduisant l'intérêt grandissant des chercheurs pour cette problématique.⁴ D'autre part, le projet OFROM revêt également une dimension patrimoniale en tant qu'archive de la parole de Suisse romande. L'exploitation des données méthodiquement collectées permet ainsi de compléter et d'examiner des hypothèses antérieures et de documenter certains phénomènes relevés à propos du français de Suisse romande.

Désormais s'ajoute une orientation didactique en cours d'élaboration dont nous présentons ici un aperçu.

OFROM représente une ressource didactique unique en son genre pour aborder, dans l'enseignement/apprentissage du français, la langue «quotidienne» parlée en Suisse romande.

1.2. Comment utiliser OFROM?

Tout d'abord, il faut préciser que la ressource OFROM, son interface et son moteur de recherche sont simples à maîtriser et ne requièrent aucune compétence spécifique. Afin d'écouter des extraits⁵, il est nécessaire de lancer une recherche via le concordancier, qui permet de chercher des chaînes de caractères, des unités monolexicales (mots), des unités polylexicales (expressions), des lemmes (formes non fléchies, par défaut le masculin pour les adjectifs, le singulier pour les noms, l'infinitif pour les verbes) ou des étiquettes morphosyntaxiques (adjectif, pronom, verbe, etc.).

Un tutoriel succinct se trouve dans la rubrique et un mode d'emploi plus fourni est intégré au document de référence téléchargeable sur le site, qui contient également des informations sur les modalités de transcription et sur les données recueillies à propos des locuteurs (par ex. âge, sexe, etc.). Comme expliqué dans le tutoriel, il est possible de procéder à des recherches plus complexes par le biais de filtres⁶.

Figure 1: Fonction « Recherche » du concordancier.

- 3 Elles peuvent par exemple concerner les activités ou projets professionnels, les loisirs, les voyages, ou alors des expériences incongrues, la politique, la situation linguistique de la Suisse, les coutumes, etc.
- 4 Voir à cet égard le numéro 15 de la revue *Corpus* (2016), qui dresse un bilan sur l'état des corpus de français parlé dans la francophonie à travers la présentation de treize corpus.
- 5 Pour le bon fonctionnement de la ressource, un préalable est d'utiliser un navigateur compatible avec le HTML5 (comme Firefox ou Chrome).
- 6 Un filtre *contexte* permet de rechercher des cooccurrences d'éléments (par exemple la présence de la négation *ne* devant un verbe quelconque); un filtre *locuteur* permet d'orienter la requête en fonction de critères sociolinguistiques appliqués aux locuteurs (canton, ville, âge, sexe, etc.); un filtre *enregistrement* offre la possibilité de choisir la qualité sonore et le genre de parole souhaité pour la recherche.

Locuteur	Texte	Tout/Rien
unifr11-dla	moi Je suis j'ai huitante ans mais ça ça	Détails
unifr11-dla	d'ailleurs moi j'ai huitante ans je fais tout mes affaires je sais tout faire	Détails
unifr11-cmc	huitante centimètres c'est des gros arbres helin en foyard	Détails
unifr14-cba	par exemple en Suisse romande on va dire euh huitante on va dire euh adieu pour saluer	Détails
unifr16-012	cinquante soixante septante huitante vendangeurs	Détails
unine08-oca	bon on doit être présent à plus de huitante pour cent des cours pour que le cours nous soit validé	Détails
unine08-lha	en dix-neuf cent huitante-quatre	Détails
unine11-ffa	huitante-cinq pour cent est suisse allemand	Détails
unine11-smb	des groupes des années septante huitante donc euh groupes de rock groupes euh	Détails
unine11-nfa	et c'est une tour on est mangé euh on a été jusqu'au huitante-troisième étage	Détails

Figure 2: Dix premiers résultats de la recherche du mot huitante.



Figure 3: Fenêtre Détails montrant la frise temporelle à manipuler.

Critères

Type de données: Nombre d'occurrences

Par: Canton

Chaîne recherchée: huitante Mot entier

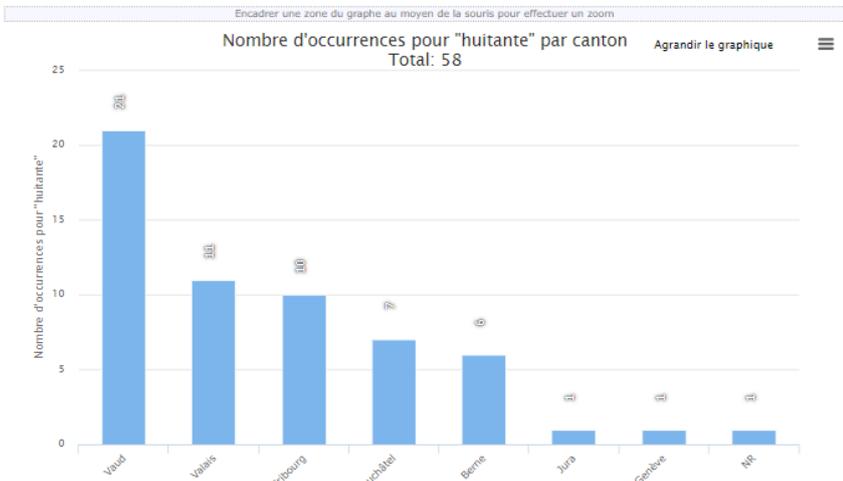


Figure 4: Rubrique Statistiques: graphique représentant le nombre d'occurrences du mot huitante par canton.

Une fois la requête lancée, les résultats apparaissent listés dans une nouvelle fenêtre. Ci-contre, on peut voir les dix premiers des cinquante-six résultats issus de la recherche du mot *huitante*.

Chaque résultat peut être écouté dans son contexte immédiat et est téléchargeable. En outre, le bouton *Détails* permet d'accéder à davantage d'informations à propos d'une occurrence en particulier. De là, une fonction particulièrement intéressante est celle qui permet d'élargir (ou de restreindre) le contexte selon la durée voulue (dans ce cas, ne pas oublier de «mettre à jour»). On peut donc étendre la durée d'un extrait à plusieurs minutes, à laquelle le téléchargement s'adapte automatiquement (à ma connaissance, c'est le seul corpus de français parlé qui propose cette option personnalisée de téléchargement).

Enfin, la rubrique *Statistiques* offre une visibilité en tout temps sur le contenu de la base. On peut par exemple obtenir un aperçu de la répartition des locuteurs ou visualiser la répartition par canton d'une occurrence donnée. Reprenons le cas de *huitante*:

Un rapide survol du graphique (figure 4) permet d'observer que les Neuchâtelois enregistrés, pourtant les locuteurs de loin les plus nombreux (90 devant les Fribourgeois au nombre de 46), sont peu enclins à employer cette variante, lui préférant la forme quatre-vingt (respectivement 7 contre 34 occurrences)⁷.

⁷ Toutefois, ceci n'est qu'une tendance sommairement relevée qui ne peut en l'état être présentée comme représentative de la population neuchâteloise, ce travail nécessitant des tests statistiques plus élaborés.

L'objectif est d'encourager les enseignants à exploiter des extraits authentiques de conversation comme matériel d'enseignement/apprentissage.

2. Exploitation didactique

Parmi les projets en lien avec la base OFROM, l'un d'entre eux est de mettre à disposition des ressources didactiques pour l'enseignement du français. L'objectif est d'encourager les enseignants à exploiter des extraits authentiques de conversation comme matériel d'enseignement/apprentissage.

Confronter les apprenants à une langue réelle par le biais d'un matériel authentique semble désormais faire consensus parmi les enseignants de langue. Par ailleurs, en cette ère du numérique, l'accès à toutes sortes de documents concrets se voit facilité par la manière dont se diffuse et se consomme l'information, la littérature, le divertissement, etc. Il existe désormais de nombreux moyens offerts aux enseignants pour accéder à des inputs authentiques: à partir de sites d'information, d'hébergement de vidéos, de banques d'images ou d'autres médias à fonction d'archivage.... En outre, des sites à vocation pédagogique proposent des séquences didactiques entières à partir de documents authentiques. Néanmoins, parmi ce matériel pédagogique, tout comme parmi les méthodes de français sur le marché, peu⁸ exploitent des supports représentant la langue telle qu'elle est parlée dans les échanges quotidiens entre usagers ordinaires. Il semble que les corpus de français oral, à visée prioritairement scientifique, demeurent peu connus des enseignants et des apprenants. Parmi les explications possibles,

outre la distance qui sépare le domaine de la recherche et celui de l'enseignement, on peut déplorer les difficultés d'emploi de certaines ressources (moteurs et terminologie complexes à maîtriser, etc.), des interfaces manquant d'intuition et de convivialité, etc. On peut également évoquer des jugements de valeur de la part des enseignants à l'égard de la parole quotidienne, susceptible de ne représenter que peu d'intérêt à leurs yeux, du point de vue de la forme (perçue comme «relâchée», «familiale» voire «incorrecte», etc.) comme du contenu (sujets de conversation peu pertinents). Or, la présente contribution veut les convaincre du contraire et leur donner un aperçu des apports de ce genre de ressources via la base OFROM.

J'ai choisi d'aborder deux axes d'exploitation d'OFROM pour l'enseignement/apprentissage du français: le travail au niveau des compétences ciblées de communication (§2.1) et l'approche de la variation linguistique (§2.2)

2.1. Travail sur des compétences de communication

Tout d'abord, je souhaite proposer une illustration de la manière dont l'acquisition de certaines compétences fonctionnelles «classiques» tant en français langue de l'école qu'en FLE peut être favorisée par des extraits de conversation quotidienne. Parmi les fonctions de communication à acquérir, on mentionne régulièrement dans les référentiels (voir par exemple les

⁸ On peut toutefois citer le projet PFC-EF <https://www.projet-pfc.net/le-projet-pfc-ef/> qui propose des fiches et séquences didactiques prêtes à l'emploi à partir de conversations ordinaires pour les enseignants et apprenants de FLM ou FLE; la base CLAPI-FLE <http://clapi.icar.cnrs.fr/FLE> met quant à elle à disposition des interactions naturelles audios et audiovisuelles en fournissant des pistes d'exploitation didactique en FLE. Le corpus FLORALE <https://florale.unil.ch/> quant à lui permet aux apprenants de FLE d'écouter des exemples de phénomènes caractéristiques du français parlé quotidien en vue de travailler la compréhension orale. Enfin, FLEURON <https://apps.atilf.fr/fleuron> propose des ressources audiovisuelles illustrant la vie d'un étudiant sur le campus en France.

descripteurs du CECRL, Conseil de l'Europe 2001) celles de description, d'argumentation, de narration, d'expression du goût, de la volonté ou de son avis, etc. Aborder ce genre de compétences discursives à travers l'observation et la pratique d'échanges naturels oraux entre interlocuteurs ordinaires, plutôt qu'à partir d'un modèle idéal ou de référence à reproduire, se profile comme une approche différente, originale et peut-être plus rassurante pour les apprenants...

Prenons le cas de la compétence rapporter les paroles d'autrui. Montrer que cette fonction est naturellement mise en pratique dans les échanges spontanés quotidiens peut en effet servir à apaiser certaines réticences d'apprenants envers sa maîtrise, liées à la mémorisation de règles de transposition grammaticale (pronoms, temps verbaux, expressions circonstanciées, etc.) et à leur application dans des phrases d'exercices sans contexte (du genre: *Il a dit: «j'ai faim» > Il a dit qu'il avait faim*, etc.). Ci-dessous, quelques activités sont proposées à partir d'un extrait d'OFROM. Bien entendu, la consigne est à adapter en fonction du public et des objectifs spécifiques. Les activités sont basées sur l'écoute des extraits correspondants, éventuellement avec la transcription à l'appui⁹. A noter que la lecture de la transcription seule est difficile si on n'y est pas habitué, et qu'elle ne permet pas de voir les tours de parole entre les différents locuteurs¹⁰:

Activité d'observation

1. Écoutez l'extrait ci-dessous (une fois sans pauses, une deuxième fois avec des pauses aux endroits opportuns) et repérez les moments où la locutrice ou son interlocuteur rapportent les paroles d'autrui ou leurs propres paroles: dites si le discours est introduit de manière directe ou indirecte en justifiant votre réponse. Réécoutez l'extrait en tout temps si nécessaire.
2. Pour aller plus loin: trouvez un texte écrit de type narratif contenant différentes formes de discours rapportés (discours direct et discours indirect) et faites le même travail de repérage (verbes introducteurs, propos rapportés, autres indices). Comparez la pratique du discours rapporté dans les deux situations.

Extrait 1:

ben en fait euh | _ | en fait si tu veux | _ | au travail | _ | y a un | _ | un | _ | mh je sais pas | % | c'est |
 | quoi un livreur | _ | _ | ouais un livreur | _ | qui est qui est venu nous déposer quelque chose | _
 | mh mh | _ | et en fait euh | _ | il est venu avec une palette | _ | % | et à | _ | la f/ quand il est parti
il a demandé si | _ | **il voulait** | _ | **si on voulait la garder la palette** **ou s'il devait la reprendre ou**
simplement | _ | exact/ | _ | _ | ouais exactement | hein | _ | okay | _ | _ | et pis du coup moi j'ai dit
ben si y a personne qui veut la garder moi je la prends | _ | *pis il m'a posé la question si j'en avais*
besoin de plus que d'une | _ | mh mh | _ | et pis euh je dis **ouais ben** | _ | **si vous en avez quinze**
euh je veux bien | _ | et pis en fait | _ | *il m'a dit ah mais oui mais je vais vous trouver ça* euh | _ | il
 en avait pas sur le moment mais *il m'a dit que il reviendrait* | _ | et en fait euh ben deux semaines
 après | _ | j'ai ma collègue qui m'appelle | _ | et *qui me dit* **ben tes palettes sont arrivées elles sont**
sur ta place de parc euh **tu** | _ | d'accord | _ | **tu viens les chercher quand tu veux quoi** | _ | quinze
 palettes | _ | quinze palettes ouais | _ | c'est beaucoup hein (unine15-011)

9 Quelques informations sur les transcriptions: | _ | marque une pause silencieuse, % représente un segment non transcrit car incompréhensible. # remplace un élément anonymisé. @ signalent un segment prononcé par un enquêteur. Enfin, / indique un segment inachevé.

10 Il est tout à fait possible d'adapter la transcription pour plus de lisibilité à des fins didactiques, en créant par exemple une ligne différente par locuteur. Nous ne conseillons toutefois pas de supprimer les traces de dysfluences (*euh*, répétition, interruptions, etc.) ni de procéder à la correction de phénomènes grammaticaux (par ex. ajouter des *ne*), ni encore à remplacer/ajouter une forme de ponctuation (propre à l'écrit), la transcription se voulant d'être au plus près de ce qui a été prononcé.

La base OFROM représente
ainsi une ressource
idéale pour approcher
globalement la variation.

Activité de production

1. Après plusieurs écoutes de cet extrait, racontez cette scène à un camarade (Par ex.: une femme raconte à un ami qu'un jour au travail [...]).
2. Imaginez et jouez la scène d'origine (au discours direct) entre la locutrice et le livreur, en ajoutant des propos adaptés au contexte.
3. Mettez-vous dans la peau du livreur qui raconte cette même anecdote à un collègue. Trouvez une suite et une fin à la scène.

Aborder le discours rapporté à travers cet extrait de conversation ordinaire devrait permettre un certain nombre d'observations, parmi lesquelles, de manière non exhaustive: la façon dont le discours direct est réalisé (traces de spontanéité ou d'expressivité comme *ben, ouais ben, ah*, etc., contrefaçon de la voix), la distinction et l'alternance entre discours direct et indirect (conjonctions ou autre moyens de balisage à l'oral des propos rapportés), le fait que les paroles rapportées (même au discours direct) ne sont pas forcément une reproduction à l'identique des propos d'origine (il réside une forme d'interprétation dans le discours rapporté). Les activités de production permettent quant à elles de développer la compétence de prise en charge des paroles d'autrui (ou de soi-même) à travers l'adoption de différents points de vue sur une même scène, facilitant par-là l'appropriation des fameuses «contraintes» de transposition (pronoms, temps verbaux, déictiques).

Le corpus OFROM ne permet pas pour l'instant de faire des recherches d'extraits par «compétence fonctionnelle» ou «acte de langage», la perspective de créer ce genre de collections d'extraits étant actuellement au stade de la réflexion¹¹. Pour l'heure, l'enseignant intéressé doit entreprendre par lui-même le travail de

repérage d'extraits adaptés à la compétence sur laquelle il souhaite travailler, soit en écoutant des extraits «au hasard», soit en faisant des recherches plus ciblées (par exemple pour la fonction «discours rapporté», rechercher le lemme «dire» ou «demander» permet de trouver assez facilement des extraits adaptés).

2.2. Approche de la variation

L'un des autres intérêts didactiques d'OFROM est de faciliter l'approche de la variation linguistique sous plusieurs de ses aspects. Dans l'enseignement du français, langue maternelle comme étrangère, on sépare traditionnellement les compétences sur la base du médium (oral vs écrit), le médium écrit ou sa version «standardisée» étant considérée comme la norme à enseigner, le modèle à suivre et à reproduire (souvent même dans des scènes d'oral de tous les jours, cf. la réalisation du *ne* de négation). Or cette opposition dominante tend à délaissier les autres aspects de la variation linguistique, notamment les axes diaphasique (en termes de «style», de «registre», etc.), diatopique (régionale) et diastratique (niveau socio-éducatif des locuteurs), qui sont des aspects particulièrement difficiles à maîtriser pour un apprenant de FLE. La base OFROM représente ainsi une ressource idéale pour approcher globalement la variation en examinant l'interaction des différentes variables de production de la parole (lieu, moment, bagage socio-culturel, intentions de communication, type de situation, etc.).

Le français parlé en «situations»

Dans le futur, OFROM devrait fournir une palette plus diversifiée de situations de parole. Cette diversité permettra de comparer plus finement différents genres de l'oral ainsi que les conditions de production propres à chaque situation: Que

¹¹ Je me tiens à disposition pour fournir des extraits d'OFROM pertinents pour aborder les compétences d'argumentation et de présentation d'un exposé.

dit-on, que ne dit-on pas dans quels types de situations? Quelles en sont les caractéristiques récurrentes? Comment les locuteurs réalisent-ils certains actes de langage dans ces diverses situations?, etc. Autant de questions auxquelles réfléchir et tenter de répondre dans un objectif de sensibilisation des apprenants aux différentes situations de l'oral, tout en nuanciant l'opposition trop stricte que l'on établit parfois entre oral-écrit.

Pour l'heure, OFROM permet de situer, sur une échelle de degré de préparation de la parole, l'interaction libre d'un côté, (discussion «à bâtons rompus» entre des proches) et la conférence universitaire de l'autre (pour l'instant au nombre de deux). En sélectionnant ces deux genres sur OFROM (via le filtre *enregistrement*) et en les confrontant, bien des caractéristiques propres à l'un ou l'autre se dégagent, du point de vue des dimensions interactionnelle (marques de relations sociales, de politesse), lexicale, morphosyntaxique (négation, forme des questions, dislocations), phonétique ou para-verbale. Un travail de classement des faits observés selon ces différentes dimensions permet ainsi de nuancer et d'illustrer ce que l'on entend par variation diaphasique, généralement abordée dans l'enseignement en termes de registres de langue, mais souvent peu explicitée et illustrée par quelques exemples de variantes types (du genre *ça va?* vs *comment allez-vous?*).

Nous avons proposé ce genre d'activité dans le cadre d'un cours d'expression et compréhension orales destiné à des apprenants universitaires de FLE (niveau intermédiaire). Dans une phase d'observation à propos de deux courts extraits (l'un issu de unine15-045 pour l'interaction libre, l'autre de unifr12-swa pour la conférence), les consignes soumises aux étudiants étaient les suivantes:

1. Écoutez les extraits: quels sont leurs points communs, leurs différences?
2. Dans chacun des extraits, pou-

vez-vous déterminer qui parle, à qui, à propos de quoi, dans quelle situation de communication?

3. Repérez les traces d'oral spontané ou à l'inverse les indices d'un discours préparé ou tout autre phénomène remarquable de la parole. Classez ces marques selon les dimensions présentées (interactionnelle, lexicale, morphosyntaxique, phonétique, para-verbale).

4. (Pour aller plus loin) Par groupe de trois, échangez sur les marques de spontanéité dans vos langues d'origine respectives, et discutez des jugements dont elles font l'objet.

Ce travail d'observation et de réflexion permet de faire ressortir d'une part, que les aspects généralement catégorisés comme des «dysfluences» et stigmatisés (chevauchements, hésitations, répétitions, interruptions, etc.), sont la trace de la fabrication du discours sur le vif (plus visible dans des échanges libres entre proches, par opposition au genre de la conférence¹²) et d'autre part, que les différences relevées entre les deux extraits sont liées aux conditions de production propres à chaque genre¹³. L'objectif consiste ainsi à aborder la variation non pas en termes de jugement de valeur (*bien/pas bien, juste/faux, correct/incorrect*) ou de différence entre oral/écrit, mais plutôt en termes d'adéquation à la situation de parole: des locuteurs francophones parlent ainsi dans telle ou telle situation.

Le français parlé en Suisse romande

Concernant l'aspect régional de la variation, une même approche critique peut être adoptée, afin de réfléchir au rapport entre cette variété et d'autres variétés du français, au rapport des apprenants à cette variété, à ce qui constitue une langue, etc. Selon de Pietro (2008), les activités autour du français régional, dans le cadre de l'école obligatoire en Suisse romande, devraient également servir à construire des références culturelles et identitaires

12 Ceci dit, au sein du genre conférence (de même qu'au sein de chaque «genre» ou «sous-genre» de parole) on observe à nouveau de la variation, selon le type de conférence, les disciplines, les orateurs, le public, etc.

13 Ces différences liées aux conditions de production, notamment en termes de degré de préparation, s'observent également à l'écrit. De manière similaire, on peut proposer aux apprenants deux extraits provenant du médium écrit, l'un d'un genre de «référence» (scientifique ou prose littéraire par exemple) et l'autre d'un genre écrit spontané (chat, forum, SMS ou autre genre conversationnel en ligne).

OFROM pourrait servir, en complément d'autres types de documents authentiques (émissions radio, tv, etc.), à illustrer ces spécificités du français régional dans un contexte réel, et à en montrer la vitalité et l'utilité.

et aussi, à lutter contre le sentiment d'insécurité linguistique bien connu des Romands (entre autres Jolivet 1984, Singy 1996), comme tout locuteur périphérique, par rapport à un français qui servirait de «référence», généralement assimilé au français hexagonal, en particulier celui de Paris.

La base OFROM contient bon nombre d'observations ou de jugements de la part de locuteurs romands à propos de la variation régionale, les enquêteurs les ayant parfois délibérément interrogés sur ce thème. Ci-dessous l'extrait audio dont est proposée la transcription peut par exemple servir de «déclencheur» de réflexions ou débats sur ces questions auprès des apprenants:

Extrait 2:

j'aime pas non plus euh | _ | ça a rien à voir mais c'est dans les chiffres | _ | j'aime pas le quatre-vingt-seize le | _ | les les les français ils sont ils font un blocage | _ | je je pense qu'on a raison | _ | avec les belges | _ | hein on doit compter en base dix pas en | _ | en base dix mais pas en quatre-vingts en base vingt ou | _ | hein | _ | @ | _ | @ | _ | ah ben ouais ben ça ma foi je s/ là j'ai je suis faux là je suis faux je suis suisse | _ | donc je prends pas le petit-déjeuner euh | _ | le déjeuner et le diner moi je suis j'ai déjeuner diner souper mais ils me comprennent les français quand je leur parle de | _ | ils comprennent | _ | alors je suis je pense que je suis faux | _ | effectivement là par rapport aux français | _ | mais ça me pose pas de problème | _ | ça me pose pas de problème je reste faux | _ | quand même | _ | ouais | _ | ah et peut-être que | quoi eh je demanderais un petit-déjeuner non si je suis à l'hôtel en France j'irais prendre mon petit-déjeuner tu vois | _ | mais ça reviendrait vite hein | _ | par contre à midi j'arrive pas à dire déjeuner je crois pas | _ | à midi j'arrive pas (unine15-008)

On peut également confronter plusieurs extraits révélateurs d'opinions (tirés par exemple de unine15-016 , unine15-021, unine16-020, unine15-020) offrant l'occasion aux apprenants de questionner leur rapport au français romand, les rapports entre le français dit «standard» et d'autres variétés régionales, la notion de norme ou de légitimité des locuteurs francophones, etc.

Quant à l'apprentissage proprement dit des principales variantes lexicales régionales, il peut s'avérer pertinent pour les apprenants de FLE en immersion dans une région hôte ou provenant de régions limitrophes (pour les apprenants suisses-alsaciens ou italo-romands par exemple). En effet, il prend part à un projet d'intégration à la société ou région d'accueil (ou voisine). On pense spontanément aux régionalismes en lien avec des aspects très concrets de la vie quotidienne (les expressions numériques, les noms des repas de la journée, les formules de politesse, les expressions liées à l'administration et à la politique, etc.). OFROM pourrait servir, en complément d'autres types de documents authentiques (émissions radio, tv, etc.), à illustrer ces spécificités du français régional dans un contexte réel, et à en montrer la vitalité et l'utilité, sans risquer de tomber dans la caricature, le cliché ou le folklore comme c'est parfois le cas (de Pietro 2008).

Voici une proposition d'activité, pensée pour des apprenants de FLE adultes ou du secondaire, autour d'un extrait d'OFROM qui contient un certain nombre de «statalismes» (expressions liées à l'administration ou la politique d'un pays, d'une région délimitée):

1. Écoutez l'extrait suivant et notez les termes liés à l'administration, à la politique ou à la société suisse.
2. Pouvez-vous donner des exemples concrets (personnes, institutions, lieux, etc.) des catégories politiques ou administratives relevées (cantons, communes, districts, partis politique, conseillers fédéraux, etc.)?

OFROM représente une
ressource idéale pour
stimuler des réflexions sur
la langue et la variation.

Extrait 3:

et pis ben ça c'est c'est je pense moi qui veux travailler euh | _ | plus tard dans l'administration publique | que je m'en fous euh d'une **ville cantonale** ou euh même si un jour où | _ | où on cherche une **conseillère fédérale** euh | _ | je deviendrai **UDC** si il faut | _ | je m'a/ je m'appellerai Trudi | _ | je me ferai des tresses | _ | [...] je fais partie du **conseil communal** euh [...] je fais partie du **conseil général** de mon village parce qu'y a une différence entre | **conseil général et conseil communal** alors **conseil communal** on est élu | _ | et une **commune** de moins de je sais pas combien d'habitants exactement ben | on a un conseil général | _ | je crois que c'est moins de mille | _ | [...] les gens sont de plus en plus vieux peuvent rester de plus en plus à la longtemp à la maison | _ | et pis y a des gens qu'on met en **EMS** mais qu'on pourrait garder à la maison ou dans un | un un lieu entre deux euh | _ | qui leur conviendrait mieux | _ | qu'un **EMS** qui est un lieu pour des gens qui ont quand même des problèmes de santé euh [...] et y a eu l'idée de faire des **appartements protégés** | _ | [...] ce qui est assez cool c'est que ça a lancé un débat et une commission | _ | je sais pas si on dit vraiment une commission mais disons que **les préfets** | _ | **le préfet** de la région et pis **les les syndics** se sont alliés et planchent sur le sujet | _ | et suite à notre entre guillemets impulsion | _ | y a des des une étude plus approfondie qui est en cours dans le **district** de | # | (unine08-jla)

Cet extrait offre ainsi l'occasion d'enchaîner sur des activités en lien avec des aspects culturels concrets (société, système politique, histoire, etc.).

Pour le moment, la recherche de régionalismes lexicaux dans OFROM se fait manuellement, en recherchant le mot/lemme souhaité via le concordancier. L'une des pistes de recherche future est de parvenir à un étiquetage automatique des mots selon leur caractère régional, ce qui permettrait de faire des recherches plus systématiques sur ce point, pour identifier par exemple des extraits contenant de nombreux régionalismes, pour examiner quelles spécificités lexicales apparaissent dans quelles régions, etc.

Une autre piste pour valoriser la dimension patrimoniale de ce corpus consisterait à proposer au grand public des collections sonores illustrant en contexte quelques phénomènes régionaux de Suisse romande (mais souvent attestés au-delà!) et à en fournir une brève explication à caractère vulgarisé. On peut

penser par exemple à l'illustration et l'explication de l'emploi de *vouloir* comme auxiliaire du futur (du type *il veut pleuvoir* pour *il va pleuvoir*), à celles de l'emploi des formes surcomposées (*j'ai eu été jouer au volley avec des copains* > unine11-fdb) (Avanzi, Béguelin & Diémoz 2016).

Il ne faut toutefois pas oublier que la variété romande (d'ailleurs elle-même non homogène, des différences existant entre les régions) appartient à la même langue que le français «de référence» et qu'en Suisse romande, les variantes régionales coexistent avec les variantes standards (Avanzi, Béguelin & Diémoz 2016). On ne s'étonnera donc pas que la base OFROM ne contienne pas tous les régionalismes lexicaux répertoriés dans les ouvrages de référence (e.g. Thibault 1997), le corpus étant en l'état limité à 64h d'enregistrement (mais voué à s'agrandir) et les sujets de conversations n'épuisant évidemment pas tous les domaines sémantiques concernés par les régionalismes.

La démarche consiste à
confronter les apprenants
à la langue telle qu'elle
est réellement pratiquée
autour ou près d'eux.

Conclusion

Le recours aux corpus en général étant encore peu courant dans l'enseignement du français, surtout au niveau de l'école obligatoire ou secondaire, le but de cet article était de présenter OFROM comme une ressource facile d'emploi et unique en son genre et surtout de fournir des pistes didactiques concrètes pour l'exploiter. J'espère avoir montré la pertinence de travailler en classe de langue à partir d'extraits de situations diverses et authentiques de parole entre locuteurs romands: la démarche consiste à confronter les apprenants à la langue telle qu'elle est réellement pratiquée autour ou près d'eux. L'observation de pratiques langagières orales quotidiennes en contexte offre un angle d'attaque original et complémentaire à une approche plus traditionnelle qui demeure souvent basée sur les genres considérés comme plus «nobles» ou plus «dignes d'intérêt». Montrer aux apprenants que les compétences de communication à acquérir se

retrouvent, certes avec des conditions d'emploi différentes, dans toutes sortes de situations, notamment des situations qui leur sont familières et dans lesquelles ils peuvent plus aisément se projeter, pourrait me semble-t-il favoriser l'acquisition de certaines pratiques langagières (abordées généralement en termes de «règles», «contraintes d'emploi», etc.). En outre, OFROM représente une ressource idéale pour stimuler des réflexions sur la langue et la variation qui en ressort: le recours à la base permet d'observer l'utilisation des variantes à différents niveaux et de réfléchir à leurs circonstances d'emploi appropriées, de se situer en tant que locuteur par rapport à diverses variétés, y compris au français dit «de référence», de mesurer les enjeux culturels et identitaires en lien avec la variation, etc. Bref, le corpus OFROM a beaucoup à offrir sur les plans scientifiques et didactiques pour quiconque souhaite l'exploiter et il ne manque pas de perspectives pour son développement futur.

Bibliographie

- Avanzi, M., Béguelin, M.-J. & Diémoz, F. (2012-2019). Présentation du corpus OFROM – corpus oral de français de Suisse romande. Université de Neuchâtel. <http://www.unine.ch/ofrom>
- Avanzi, M., Béguelin, M.-J. & Diémoz, F. (2016). De l'archive de parole au corpus de référence: la base de données orales du français de Suisse romande (OFROM). *Corpus*, 15. <https://journals.openedition.org/corpus/3060>
- Avanzi, M., Béguelin, M.-J. & Diémoz, F. (2016). Introduction. *Corpus*, 15. <http://journals.openedition.org/corpus/2913>
- Cappeau, P. & Gadet, F. (2007). Où en sont les corpus sur les français parlés? *Revue française de linguistique appliquée*, 12, 129-133.
- Conseil de l'Europe (2001). *Un cadre européen commun de référence pour les langues: apprendre, enseigner, évaluer*. Paris: Didier.
- Jolivet, R. (1984). L'acceptabilité des formes verbales surcomposées. *Le Français moderne*, 52, 159-176.
- de Pietro, J.-F. (2008). L'école face aux variétés du français: réflexions à partir de la situation en Suisse francophone. In: *La langue française dans sa diversité. Actes du Colloque tenu à Québec les 21, 22 et 23 septembre 2008 en marge de la Rencontre des organismes de gestion linguistique*. Québec: Secrétariat à la politique linguistique, pp. 181-208.
- Singy, P. (1996). *L'image du français en Suisse romande. Une enquête sociolinguistique en Pays de Vaud*. Paris: L'Harmattan.
- Thibault, A. (1997). *Dictionnaire suisse romand*. Carouge: Éditions Zoé.