



HAL
open science

Création d'un corpus de français parlé à des fins pédagogiques en FLE : la genèse du projet FLORALE

Christian Surcouf, Alain Ausoni

► To cite this version:

Christian Surcouf, Alain Ausoni. Création d'un corpus de français parlé à des fins pédagogiques en FLE : la genèse du projet FLORALE. *Etudes en didactique des langues*, 2018, *Les corpus / Corpora*, 31, pp.71-91. hal-04034793

HAL Id: hal-04034793

<https://ut3-toulouseinp.hal.science/hal-04034793>

Submitted on 17 Mar 2023

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.

Création d'un corpus de français parlé à des fins pédagogiques en FLE: la genèse du projet FLORALE

Christian SURCOUF & Alain AUSONI

Maîtres d'enseignement et de recherche

École de français langue étrangère (EFLE)

Faculté des Lettres, Université de Lausanne (Suisse)



Introduction

Commençons par le récit d'une anecdote. Dans un article explicitement intitulé "If this is French, then what did I learn in school?", l'un des auteurs relate une de ses expériences langagières personnelles alors qu'il séjournait à Paris après trois ans d'apprentissage universitaire du français aux États-Unis. Ce jour-là, l'auteur se rend dans un grand magasin pour y faire des achats.

I found a salesperson and asked in my prerehearsed French, "Pardon, Madame, où sont les ... (whatever it was that I was looking for)?" After quick reflection, she shook her head and replied "Chépas, m'sieur." Her response bewildered me until I recalled a conversation with a friend in college who had spent a summer studying in Paris. She happened to mention that Parisians frequently uttered "Chépas" instead of enunciating "Je ne sais pas." At the time, I failed to see any connection between the two dissimilar sounding phrases. Yet, if not for that chance conversation, I would have been entirely at a loss as to what that salesperson had said¹ (Durán & McCool, 2003: 288).

D'un point de vue didactique, un tel témoignage ne peut qu'interpeller. Comment se peut-il qu'après trois années d'études du français cet apprenant n'ait pas eu connaissance de l'existence d'une expression aussi fréquente que /ʃsepɑ/? Comment expliquer son incapacité à percevoir la relation entre /ʒənəsɛpa/ et /ʃsepɑ/? Pourquoi, comme il le relève lui-même: "No instructor in any of my French courses had ever alluded to the transformation of the formal *je ne sais pas* to the colloquial *chépas (j'sais pas)*" (*ibid.*: 288)? Pédagogiquement tout aussi préoccupante s'avère par ailleurs l'observation selon laquelle l'auteur n'éprouvait

¹ Signalons qu'ici le problème ne résulte nullement d'une trop grande complexité lexicale, morphosyntaxique, sémantique ou pragmatique, mais simplement d'un manque d'habitude à une forme pourtant fréquente en français.

aucune difficulté à comprendre ses enseignants à l'Alliance française alors que sa compréhension des émissions de radio ou de télévision, des films, des chansons et des gens dans la rue se révélait très insuffisante (*ibid.*). Pourquoi une telle disparité? Ses enseignants de français langue étrangère (FLE) pratiquaient-ils une variante particulière de l'oral se démarquant considérablement du "français ordinaire"²? De telles difficultés seraient-elles inhérentes à l'apprentissage même d'une langue? Et comment expliquer qu'après les débats des années 70 autour de l'intégration des documents authentiques³ (à ce propos, voir par exemple Cuq & Gruca, 2002: 389; Giroud & Surcouf, 2016; Vialleton & Lewis, 2014: 299-303), notamment oraux, un apprenant de la fin du XX^e siècle se retrouve dans cette situation inconfortable, conséquence, *a priori*, d'un enseignement de l'oral inadéquat, ou pour le moins insatisfaisant? Si l'on peut éventuellement envisager que l'apprenant n'ait pas pour impératif immédiat de maîtriser la production de telles particularités langagières, n'est-il pas en revanche du ressort de l'enseignement de l'y sensibiliser, pour la simple raison que sa compréhension de la langue cible en dépend singulièrement? En définitive, aussi anecdotique puisse-t-il paraître, ce témoignage ne peut qu'interpeller didacticiens et pédagogues sur la problématique de l'enseignement/apprentissage du FLE dans sa dimension orale. Signalons qu'un tel problème ne se restreint pas aux apprenants en milieu hétéroglotte. En effet, comme nous le montre notre pratique quotidienne de l'enseignement du FLE en Suisse romande, la compréhension de l'oral spontané⁴ reste souvent difficile pour nombre d'étudiants résidant pourtant en milieu francophone.

De tels constats nous conduiront à réfléchir, dans un premier temps, à l'origine possible de ces difficultés de compréhension et, *a fortiori*, d'appropriation. À cet effet, après une réflexion épistémologique sur la description linguistique de l'oral, nous détaillerons, à l'aide d'exemples concrets trouvés dans des ouvrages de référence de français langue étrangère ou maternelle (FLM), certaines des faiblesses de la présentation didactique des fonctionnements de l'oral. La seconde partie exposera, quant à elle, la solution que nous avons adoptée en créant FLORALE (français langue orale et FLE), une interface informatique permettant à des apprenants (dès le niveau A2) d'explorer un corpus de français parlé authentique, annoté à des fins pédagogiques. Nous évoquerons succinctement les défis, à la fois informatiques et pédagogiques, qu'il nous a fallu relever pour mettre en œuvre cette ressource encore récente.

² "Ce n'est bien sûr pas le français soutenu, ni recherché, ni littéraire, ni normé. [...]. C'est donc surtout le français familier, celui dont chacun est porteur dans son fonctionnement quotidien, dans le minimum de surveillance sociale: *la langue de tous les jours*" (nous soulignons) (Gadet, 1989: v).

³ Un document authentique, se définit "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.. En sont exclus les textes écrits et oraux fabriqués ou modifiés en vue d'enseigner une langue, le discours d'un enseignant en langue ou d'un *native speaker*, si ce discours a pour fonction d'enseigner une langue, et les productions en langue étrangère des apprenants (échanges entre apprenants et enseignants ou entre apprenants) si ces productions ont pour but d'apprendre la langue étrangère" (Abe *et al.*, 1979: 2-3).

⁴ Sur la question du terme "oral spontané", voir Blanche-Benveniste & Bilger (1999).

Si, dans la perspective de l'enseignement/apprentissage de l'oral en FLE, le didacticien et le pédagogue puisent dans les descriptions fournies par les linguistes et les grammairiens, voyons quels sont les enjeux épistémologiques inhérents à toute tentative de description de l'oral.

BRÈVES CONSIDÉRATIONS ÉPISTÉMOLOGIQUES SUR LA PRÉSENTATION OU LA DESCRIPTION DE L'ORAL

Rappelons tout d'abord que, quels que soient le prestige et le statut accordés à l'écrit dans nos sociétés, l'oral reste premier phylogénétiquement et ontogénétiquement (Martinet, 1960/2008: 32-33) ainsi que, pour la plupart des individus, quantitativement (Blanche-Benveniste, 2003: 317). En dépit de son évidence, ce constat, *a priori* consensuel en linguistique, n'en demeure pas moins soumis au biais induit par le truchement de l'écrit dans les analyses⁵.

The [Written Language Bias] heritage is still pervasive, despite the facts that linguists, with few exceptions, claim the primacy of speech and spoken language over writing and written language, and that, in more recent times, they have indeed been explicitly concerned with spoken language too. We can talk about a paradox in modern linguistics: one claims the absolute primacy of spoken language, yet one goes on building theories and methods on ideas and experiences of a regimented, partly made-up language designed for literate purposes and overlaid with norms proposed by language cultivators, standardisers and pedagogues. All this amounts to a deeply ingrained contradiction based on a veritable reversal of priorities (Linell, 2005: 30).

Retraçons brièvement certains des écueils auxquels risque de se heurter le linguiste dans son entreprise de description de la langue orale.

Les entités en jeu dans la présentation ou la description de la langue

Lorsqu'il entreprend de décrire son objet – la langue –, le linguiste recourt à un outil de description/présentation, qui se trouve être également la langue, alors investie dans cette fonction du statut de métalangue⁶. Or, dans nos sociétés alphabétisées, l'objet-langue tout comme l'outil de description ou de présentation (*médium* dans le Tableau 1 ci-dessous) existent sous deux formes: orale et écrite. On obtient alors quatre configurations susceptibles d'influer sur les présentations linguistiques, grammaticales et pédagogiques. Il y a congruence entre l'objet et le médium lorsque les deux sont de même nature.

⁵ Sur le très long terme, Auroux *et al.* (2004: 60-61) rappellent le rôle fondamental de l'écriture dans l'émergence d'une réflexion linguistique: "l'histoire nous montre incontestablement que dans les civilisations où il y eut apparition d'un savoir linguistique, celui-ci est indubitablement né après que l'écriture eut été une technique bien dominée".

⁶ Entendue comme "la langue (naturelle ou formelle) servant à décrire la langue et le langage" (Neveu, 2004: 188).

	Objet	Médium	
(a)	oral	oral	congruent
(b)	écrit	écrit	congruent
(c)	oral	écrit	non-congruent
(d)	écrit	oral	non-congruent

Tableau 1 – Les effets de congruence entre l’objet et le médium oraux et écrits

Seules les configurations (a) et (b) s’avèrent congruentes. Ainsi une description orale est-elle en mesure de reproduire fidèlement la caractéristique sonore définitoire de la langue orale qu’elle décrit. Il en serait de même pour l’écrit en (b). Quant aux configurations (c) et (d), elles posent immédiatement le problème du transcodage de l’oral à l’écrit ou inversement (voir *infra*). La grammaire et, dans son prolongement, la linguistique, ont largement tiré profit de la congruence offerte par la configuration (b) en commentant à l’écrit des exemples écrits, traditionnellement littéraires, et ce au détriment de l’oral. En définitive,

[l]orsqu’il est question de la langue française, de sa grammaire et de son lexique, c’est en général de la langue écrite qu’il s’agit. La grammaire et le lexique de langue parlée n’apparaissent dans les ouvrages de référence, la plupart du temps, que comme des curiosités marginales dignes d’un petit musée des horreurs de la langue (Blanche-Benveniste, 2003: 317).

L’écrit, par sa capacité à fixer définitivement une production langagière, la rend observable et analysable à souhait. Il s’offre alors naturellement comme objet privilégié de description pour le linguiste, le grammairien et le pédagogue⁷. Par ailleurs, depuis l’avènement de l’imprimerie, l’écrit constitue l’instrument didactique privilégié de transmission de tels savoirs grammaticaux et linguistiques. Cette emprise de l’écrit s’est vue d’autant plus aisément confortée que l’enregistrement et le stockage de données orales ne sont devenus possibles que ces dernières décennies grâce aux progrès de l’informatique. En somme, tous ces facteurs tendent à conférer à l’écrit le triple statut dont il jouit encore amplement dans la linguistique contemporaine, à savoir celui d’objet à décrire, d’outil de description et d’outil de transmission. À cette suprématie de l’écrit sur l’oral, s’ajoute une autre limitation.

Pour travailler sur la langue parlée, il faut pouvoir se la représenter et la conserver en mémoire suffisamment longtemps. La mémoire à court terme ne permet pas d’en observer de longs fragments de façon fiable. Lorsque l’énoncé parlé est transféré dans la mémoire épisodique, nous en conservons le contenu plus que la forme. Pour étudier le parlé, nous devons donc nous en donner des représentations graphiques: par des transcriptions écrites et par des transpositions de la matière sonore (Blanche-Benveniste, 2003: 321).

⁷ Auroux *et al.* (2004: 62-63) expliquent que: “pour qu’il y ait sciences du langage, il faut que le langage soit placé en position d’objet. Il faut qu’il soit là, devant nous, manifestation de lui-même et non pas d’autre chose, contrairement à ce qui se passe dans son usage quotidien. Le processus d’apparition de l’écriture [...] est un processus d’objectivation considérable et sans équivalent antérieur. [...] ce qui fait véritablement démarrer la réflexion linguistique, c’est l’altérité, envisagée essentiellement du point de vue de l’écrit”.

Le problème du transcodage

Un des caractères définitoires fondamentaux de l'oral repose sur son caractère vocal, l'écrit étant quant à lui toujours silencieux. En ce sens, “la notation de l'oral par écrit constitue une contradiction irréductible: l'écrit ne présentera jamais qu'une image approximative de la réalité linguistique orale” (Gadet, 1989: 43). Intransposable s'avère – entre autres – la dimension sonore de l'oral, menant inexorablement à la perte du contour mélodique, du rythme, des accentuations, de la qualité de la voix, etc. (voir Linell, 2005: 34). Cependant, une fois admise l'incontournabilité d'une forme de transcription, trois grandes options s'offrent au linguiste (pour un aperçu des possibilités, voir Gadet, 1989: 43-49): l'orthographe conventionnelle, une orthographe aménagée et l'alphabet phonétique international. La deuxième option, choisie par Durán & McCool (2003: 288) avec la graphie “chépas” pour transcrire l'une des quatre formes orales [ʃepa], [ʃʃepa], [ʃʃɛpa] ou [ʃɛpa], s'avère immédiatement problématique parce qu'elle n'est pas normée et requerra fréquemment la création de graphies nouvelles, non-partageables, entre chercheurs et, *a fortiori*, difficiles à adopter dans l'enseignement/apprentissage du FLE. En somme, qu'il s'agisse de descriptions linguistiques ou de leur adaptation et utilisation pédagogique en aval, seules les deux autres options semblent recevables. Lors de la constitution de corpus oraux informatiques de grande étendue, l'orthographe conventionnelle présente l'avantage de garantir la facilité de transcription et de lecture, la formulation de requêtes informatiques, l'interopérabilité et l'échange de données entre corpus.

Cependant, aussi indispensable soit-elle pour l'analyse linguistique, comme le remarque Hagège (1985: 102), en soi, “l'écriture est une analyse linguistique à des degrés divers de conscience”, et Gadet (2007: 44) de “se demander si les linguistes se donnent vraiment les moyens d'appréhender l'oral autrement que domestiqué par les normes de l'écrit”. En effet, si l'objectif poursuivi par la description est avant tout phonétique ou morphologique, la transcription phonétique s'impose, et évitera certaines incohérences descriptives. Par exemple, Maingueneau (1994/1999: 101) stipule, à propos de la forme *boxera*, que “le futur simple peut s'analyser comme la combinaison de l'infinitif *boxer* et du présent de l'auxiliaire *avoir*: *ai, as, a, av-ons, av-ez, ont*”. Opérationnelle orthographiquement (*boxer+a > boxera*), la transcription phonétique révèle que, du point de vue morphologique, l'ajout de /e/, /a/, etc. à l'infinitif /bɔkse/ donnerait */bɔksee/, */bɔksea/, etc. et non pas les formes /bɔks(ə)ʁe/ /bɔks(ə)ʁa/, etc. attendues⁸ (voir Giroud *et al.*, 2018). Comme le montre cet exemple, la complexité du rapport graphème-phonème de l'orthographe française (voir Catach *et al.*, 1995/2010; Fayol & Jaffré, 2014; Honvault-Ducrocq, 2006) la rend peu propice aux explications touchant la prononciation et la morphologie.

⁸ Cette conception orthographique de la morphologie verbale se retrouve dans de nombreux ouvrages de conjugaison, où elle constitue *la* règle de formation du futur (voir, par ex., Bled *et al.*, 2010: 119).

Dans une problématique pédagogique proche de la nôtre, dans un article intitulé “Enseigner une grammaire de l’oral aux débutants”, Capelle remarque:

La difficulté commence dès le premier contact avec le français. Les étudiants voient la transcription écrite des énoncés qu’ils apprennent oralement. Les explications qu’ils reçoivent ne les aident nullement à comprendre car on leur donne des règles valables pour le seul écrit. Ce qu’ils entendent n’a que peu à voir avec ces règles et ils se trouvent confrontés à de multiples contradictions (Capelle, 2016: 26).

L’auteur utilise alors l’énoncé “le nouvel appartement de mes amis est au troisième étage” pour illustrer son propos:

Oralement, en français, la structuration des énoncés est assurée par les fonctions de l’intonation et de l’accent tonique. [...] Il ressort de ce qui précède que l’unité d’enseignement oral n’est pas le mot, mais le groupe phonique. La cohésion à l’intérieur du groupe est assurée ou renforcée par la liaison: “mes amis”, et l’enchaînement “nouvel appartement”. Sous l’effet de la liaison et de l’enchaînement, la proportion des syllabes ouvertes (c’est-à-dire se terminant par une voyelle) dépasse les 70 % estimés et tend vers 100 %.

Le/ nou/ ve/ la/ ppar/ te/ ment/ de/ me/ sa/ mis/// /e/ tau/ troi / iè/ mé/ ta/ ge [*sic*] (*Ibid.*: 26 & 27).

Outre le fait que la décomposition syllabique semble en partie discutable (par exemple /taʒ/ ne forme, *a priori*, qu’une seule syllabe en français standard), cet exemple illustre la difficulté de recourir à l’orthographe pour expliquer des phénomènes ayant trait à la prononciation (“ve” se lirait naturellement /və/ au lieu du /vɛ/ ciblé, “sa” /sa/ au lieu de /za/, etc.).

Ce bref examen montre certains des problèmes auxquels se retrouve confronté le linguiste lorsqu’il décrit son objet, en raison même de la nature de l’outil de description que constitue l’écrit. Et si le recours à l’alphabet phonétique se révèle d’un précieux secours dans l’analyse des phénomènes morphologiques ou phonétiques, il faut cependant garder à l’esprit que jamais la transcription, quelle qu’en soit la qualité, ne peut se substituer à l’écoute de l’original oral. Mais quel est précisément le type d’oral auquel sont soumis les apprenants durant leur apprentissage du FLE?

LES MODÈLES DE L’ORAL AUXQUELS L’APPRENANT EST CONFRONTÉ

La présentation congruente de l’oral

Examinons, en premier lieu, la configuration, *a priori* optimale, consistant à présenter l’oral à l’aide de l’oral. Dans le cas de l’oral pratiqué par l’enseignant face à ses apprenants, il s’agirait en somme d’une forme d’incarnation de l’oral, alors que les enregistrements audiovisuels des manuels de FLE constitueraient davantage une présentation de l’oral.

L’oral de l’enseignant de langues

Souvenons-nous de la remarque pédagogiquement préoccupante de Durán concernant ses enseignants. Non seulement aucun d’entre eux n’avait mentionné

l'existence de la forme /ʃ(ʃ)epa/ mais, plus intrigant encore, l'auteur rapporte n'avoir éprouvé aucune difficulté à comprendre ses enseignants de l'Alliance Française (Durán & McCool, 2003: 288). En d'autres termes, durant ses trois années d'études universitaires du français, aucun enseignant n'aurait jugé nécessaire de sensibiliser explicitement ses étudiants à l'absence de *ne* dans la négation, et à l'assimilation en [ʃ] qu'elle entraîne fréquemment sur la prononciation du pronom [ʒə]⁹. Implicitement, par ailleurs, Durán n'avait, *a priori*, jamais eu affaire à des enseignants recourant suffisamment souvent à la formulation [ʃ(ʃ)εpa] ou toute autre structure analogue ([ʃ(ʃ)ɥipa], [ʃkʁwapa], etc.) pour les percevoir comme usuelles et lui en faciliter la compréhension. Il semble alors légitime de s'interroger sur le type d'oral qu'utilisaient ces enseignants. Comment expliquer que cet oral ne lui ait pas permis de comprendre aisément le français pratiqué par des millions de francophones dans leur quotidien?

Ellis (1997: 45-46) rapporte que, d'une manière générale, lorsqu'ils s'adressent à des étrangers, les locuteurs natifs tendent à parler plus lentement en simplifiant les structures utilisées, en recourant davantage à des phrases courtes et en évitant notamment les subordonnées. Dans une étude portant sur le discours adressé à des apprenants par des enseignants de tchèque, allemand et anglais langues étrangères, et comparé à des conversations entre natifs de ces trois langues, Henzl (1979)¹⁰ relève plusieurs tendances. Au niveau du lexique, le vocabulaire est plus simple, et les mots spécifiques sont remplacés par des mots plus généraux; il comprend peu de mots composés et d'expressions idiomatiques, et il est stylistiquement plus neutre, évitant notamment les expressions marquées socialement, régionalement et émotionnellement. Pour ce qui est de la morphosyntaxe, les phrases sont plus courtes et plus simples, leur longueur étant corrélée au niveau des apprenants. Enfin, pour ce qui concerne les caractéristiques sonores de ce type d'oral, le volume de la voix est plus fort, le débit plus lent et corrélé au niveau des apprenants. Les pauses, plus fréquentes, coïncident avec les frontières entre constituants. La prononciation reste très proche du standard, avec peu de réductions de voyelles ou de groupes consonantiques. Par ailleurs, l'auteure constate une présence plus élevée de coups de glotte devant les initiales vocaliques (résumé d'après Henzl, 1979: 161-163). Bien que cette étude ne porte pas sur le français, elle n'en montre pas moins des tendances claires, probablement transférables au contexte de l'enseignement du FLE, qui expliqueraient notamment la raison pour laquelle de nombreux

⁹ Cette remarque s'avère d'autant plus surprenante que 1) le *ne* de négation n'apparaît guère en conversation (utilisé dans 6,3% des 5 092 occurrences devant clitique selon Culbertson, 2010: 95), 2) l'assimilation qui s'ensuit est également courante notamment en raison du fait que "the subject clitics have fully morphologized into inflectional prefixes in the 1st and 2nd person; [...] prefixes undergo systematic allomorphy [...]: e.g. devoicing before voiceless consonants" (Fonseca-Greber & Waugh, 2003: 102, 104).

¹⁰ L'auteure utilise le terme "foreigner talk" décrit par Ferguson (1975) pour l'anglais. Toutefois, en situation pédagogique et, *a fortiori*, en contexte universitaire, l'enseignant ne recourra jamais à ce type de "xénolecte" (Klein, 1987/1989: 65-66) consistant à omettre des déterminants, des copules, éviter les conjugaisons, etc.

apprenants, à l’instar de Durán, comprennent assez aisément leurs enseignants pendant les cours mais éprouvent des difficultés face au français oral pratiqué par les locuteurs natifs en situation écologique. Par exemple, le débit plus lent évoqué dans cette recherche (voir également Hazan *et al.*, 2015), parallèlement à une diction plus soignée, conduirait naturellement en français à la préservation des schwas, privilégiant alors des prononciations telles que /ʒəse/, /ʒəsɥi/, etc. plutôt que /ʃʃɛ/, /ʃɥi/, etc.

Seule une véritable prise de conscience de ce phénomène pourra permettre une meilleure préparation des apprenants dans leur confrontation au français tel qu’il se parle effectivement en dehors de la classe. Néanmoins, il est probable que subsistera une certaine inertie engendrée par le français oral “incarné” par le style oratoire de l’enseignant lui-même, ses représentations du “français standard” et la prégnance particulièrement forte dans le monde francophone du modèle écrit.

En pays de littératie et d’idéologie du standard, le pouvoir prêté à l’écrit minorise le statut de l’oral. Tout spécialement en français, l’écrit, lieu essentiel où a porté la standardisation, apparaît plus homogène que l’oral, où le foisonnement variationnel peut difficilement être jugulé, ce qui réaffirme l’opposition convenue entre écrit normé et oral instable (Gadet, 2007: 45).

Cette importance de l’écrit se voit, de surcroît, accrue dans de nombreux systèmes scolaires et universitaires, où:

[l]’apprenant écrit plus qu’il ne parle, doit se taire quand les autres s’expriment et l’enseignant parle plus que ne s’expriment les apprenants! La puissance du modèle de l’écrit fait donc largement partie de l’imaginaire des apprenants et des enseignants (Weber, 2006: 31).

Qu’en est-il maintenant d’un autre versant de la congruence oral-oral, fournie cette fois-ci par les enregistrements audiovisuels présentés dans les manuels de FLE?

L’oral des enregistrements des manuels

Depuis l’avènement du phonographe (à cet égard, voir déjà les remarques critiques de Clarke, 1918) et ses nombreux raffinements audiovisuels ultérieurs, l’enseignement/apprentissage des langues peut s’étayer sur les enregistrements sonores proposés par les manuels. Mais le français des extraits sonores proposés dans les manuels est-il authentique? S’il est fabriqué, s’approche-t-il néanmoins de la réalité du français conversationnel que les dialogues présentés sont censés incarner?

Prenons un cas très précis. Une analyse des trois premières leçons de dix manuels de FLE pour débutants révèle, par exemple, que dans les dialogues mettant en scène des personnages se tutoyant (excluant donc par nature le registre formel), la réduction de *tu* à /t/ devant voyelle (/tabitu/) n’est présente que dans 7% des cas dans les enregistrements des manuels (N=127) alors qu’un corpus de conversations issu de CLAPI¹¹ révèle au contraire 88% d’élisions (N=209) (Surcouf

¹¹ Corpus de langue parlée en interaction, Laboratoire ICAR, Université de Lyon. URL: <http://clapi.ish-lyon.cnrs.fr/>.

& Giroud, 2016: 18). En d'autres termes, paradoxalement, la variante de loin la plus fréquente à l'oral est ignorée dans les manuels au profit de la plus rare. On peut supposer que le débit de parole est en partie responsable de cet état de fait, comme en témoigne également la proportion beaucoup plus importante de schwas dans les enregistrements (voir Bento, 2007: 197; Surcouf & Giroud, 2016: 13-16; Vialleton & Lewis, 2014: 309).

Une partie du problème réside sans doute dans le processus de conception des dialogues pédagogiques présentés par ces manuels. En effet, ils sont le plus souvent basés sur de l'écrit, c'est-à-dire lus puis interprétés par des comédiens: "avant d'être oralisé, le dialogue pédagogique est écrit [et] s'inscrit dans une logique de présentation propre à l'écrit [...]" de sorte que, en plus des phénomènes phonologiques, toute une gamme de traits propres à l'oral spontané tels que "les dislocations, les reprises, les extractions, les hésitations, les interruptions multiples mobilisent un matériel linguistique peu familier aux apprenants et aux professeurs" (Ravazzolo *et al.*, 2015: 111). Le manque de naturel des enregistrements est effectivement relevé par Vialleton & Lewis (2014) dans leur examen de 17 manuels de FLE publiés en France et au Royaume-Uni après 2000.

The slow elocution evidenced by the much lower speech rate¹² in the textbook samples, the artificial segmentation resulting from the higher proportion and greater length of silent pauses within speech turns, usually between constituents, and the clearer articulation shown by the less frequent realisation of elisions and assimilations in possible environments, all give a false acoustic image of what spoken French sounds like in real interactions (Vialleton & Lewis, 2014: 310).

En définitive, en dépit de la congruence oral-oral, *a priori* favorable à la présentation d'un oral proche de la réalité des usages, là encore, comme dans le cas du style oratoire particulier des enseignants, les enregistrements des manuels de FLE peinent à donner une image fidèle du français tel qu'il est normalement pratiqué. Le bon sens n'inviterait-il pourtant pas à penser, dans la lignée d'Holec (1990: 68), que "l'écoute, même intensive, de textes pseudo-oraux [...] ne prépare guère à comprendre la radio ou un interlocuteur en chair et en os. Comme cela a déjà été dit à propos des savoirs, on ne peut acquérir que ce que les supports permettent d'acquérir".

Envisageons maintenant le versant non-congruent, où l'écrit sert à la description de l'oral.

La présentation non-congruente de l'oral dans les ouvrages de référence de FLE et de FLM

Comme le rappellent Koch & Oesterreicher (2001: 590), "dès qu'une communauté linguistique dispose d'une langue écrite en tant que norme prescriptive, celle-ci jouit d'un prestige absolu qui lui confère la primauté par

¹² Selon Vialleton & Lewis (2014: 308) le débit dans les enregistrements des manuels est de 3,25 syllabes par seconde et s'élève à 4,78 syllabes par seconde pour des conversations écologiques entre natifs. En situation naturelle, l'apprenant devra donc s'adapter à une augmentation de presque 50% du débit.

rapport aux variétés orales préexistantes”. Aussi “[t]ous les exemples de grammaire étaient pris, traditionnellement, dans les œuvres des grands écrivains” (Blanche-Benveniste, 2003: 320). Si les grammaires de FLE ne citent guère d’exemples littéraires, il n’en demeure pas moins qu’elles s’inscrivent dans la continuité de celles du FLM en véhiculant, elles aussi, les représentations traditionnelles sur l’écrit. L’information se voit alors hiérarchisée: les descriptions portant sur la forme écrite apparaissent en premier et sont présentées comme la norme à suivre. La forme orale, lorsqu’elle est mentionnée, est souvent disqualifiée, voire bannie. Voyons donc quelques exemples illustrant ces divers cas de figure.

Comment figurer l’acoustique dans le silence de l’écrit?

Le premier écueil auquel se retrouvera confronté l’apprenant résulte de l’impossibilité de faire figurer une quelconque information sonore dans le silence du livre. Disparaissent ainsi contours prosodiques, rythmes, accents, etc. Certains artifices, comme des flèches ou des prosogrammes, sont cependant utilisés pour représenter l’intonation. Mais, quel que soit le raffinement de la présentation graphique, ces informations demeurent silencieuses et difficiles à transposer oralement en un contour mélodique.

La prescription explicite ou implicite

Nous avons vu plus haut que le clitique /ty/ avait une très forte tendance à se réduire à /t/ devant une initiale vocalique (/taʁiv/). Si Delatour *et al.* (1991: 163) mentionnent effectivement cette option, c’est pour la condamner explicitement: “ne dites pas ‘T’aimes’, dites ‘Tu aimes’”, en contradiction avec la réalité de l’usage le plus fréquent de la langue parlée. La prescription n’est cependant pas toujours ouvertement formulée. Ainsi, dans leur *Grammaire expliquée du français*, Poisson-Quinton *et al.* (2003: 34) écartent implicitement les prononciations /i/ et /ɛ/, dont la première est pourtant très courante à l’oral: “*Ils, elles*, se prononcent [il] et [ɛl] devant une consonne: *Ils parlent* [ilparl]; *elles parlent* [ɛlparl]”, et ce contre les résultats des études sur corpus oraux (voir Ashby, 1984; Brognaux & Avanzi, 2015).

La stigmatisation indirecte

Outre ces prescriptions, l’oral souffre de ce qu’on pourrait appeler une stigmatisation indirecte. À titre d’illustration, prenons l’exemple de l’usage de *on*, qualifié de “familier” dans *La Grammaire méthodique du français*: “[l]a première personne du pluriel (*nous partons*) est fréquemment remplacée à l’oral, familier surtout, par *on* (*on part*)”¹³ (Riegel *et al.*, 1994: 34). La priorité accordée à l’écrit transparaît dans l’ordre même de la formulation: *nous*, forme attendue (celle de l’écrit), serait “remplacée à l’oral” par *on*. Pourtant, si avec Blanche-Benveniste (2003: 317), on reconnaît que “c’est [...] sous sa forme parlée que la langue est le plus largement partagée” et, par ailleurs, “qu’on apprend à parler avant d’apprendre à

¹³ L’édition de 2009 modifie seulement l’exemple: “on va au cinéma?” (Riegel *et al.*, 2009: 62).

écrire” (Saussure (de), 1916/1994: 47), alors la prise en compte des pratiques de l’oral s’avère fondamentale, et ce quel que soit le prestige accordé à l’écrit dans nos sociétés. En ce qui concerne l’usage du *on*, sur la base de leur *Corpus of Everyday Conversational European French (ECEP)*, 194 000 mots), Waugh & Fonseca-Greber (2002: 117) établissent que la première personne du pluriel est exprimée par *on* dans 99% des cas, et pas *nous* (N=1 348), conduisant les auteurs à conclure: “*nous*, the 1st Pl. [plural] of written French, is no longer applicable to a discussion of spoken French. In its stead, the form *on* has undergone a change in its basic meaning, which is now the personal ‘we’” (*ibid.*: 125). Dans les descriptions du français des ouvrages de référence, ce résultat statistique devrait logiquement conduire à inverser la formulation de Riegel *et al.* (1994: 34) pour aboutir à l’alternative: “à l’écrit, le *on* est souvent remplacé par *nous*”.

La mention différée du fonctionnement oral

La *Grammaire expliquée du français* (niveau intermédiaire) présente la négation sous sa forme “ne ... pas” dans un encadré sous lequel Poisson-Quinton *et al.* (2002: 225) expliquent que “*Ne* est historiquement la marque essentielle de la négation; *pas* est le terme le plus fréquemment associé à *ne*”. L’un des trois exemples semble manifestement oral: “– Paul est là? – Non, il n’est pas là” sans pour autant en refléter l’usage puisque que le *ne* est absent “dans 80 % des cas dans les conversations” (Blanche-Benveniste, 2003: 322). Pour accéder au fonctionnement de l’oral, l’apprenant devra parcourir sept pages de descriptions relatives aux divers usages de la négation avant de lire dans la marge: “dans la langue orale, le *ne* disparaît” (Poisson-Quinton *et al.*, 2002: 231). Là encore, comme dans la description des usages de *on* et de *nous*, le fonctionnement écrit est présenté en premier. C’est d’ailleurs ce que sous-entend conceptuellement la formulation “le *ne* disparaît”, formulation qui érige *ne ... pas* comme forme première, menant à une “disparition” de *ne* à l’oral.

Absence de mention des fonctionnements de l’oral

Venons-en à notre dernier cas de figure, extrême en l’occurrence. Lorsqu’ils abordent la négation dans leur *Grammaire essentielle du français B1*, Glaud *et al.* (2015) proposent l’enregistrement d’un dialogue entre une mère et son fils, comprenant douze négations. En voici les premiers échanges.

- Oh ce **n**’est **pas** possible! Vous **n**’avez toujours **pas** rangé votre chambre! Vous **ne** jouerez **plus** aux jeux vidéo tant que cette chambre **ne** sera **pas** rangée!
- Maman, je dois aller chez Léo!
- Tu **ne** vas aller **nette part** et tu **ne** vas voir **personne**! Tu **n**’iras **ni** chez Léo **ni** à l’entraînement de football! (*ibid.*, 2015: 156.)

À la page suivante, l’apprenant découvre la rubrique “Mémorisez”, où apparaissent laconiquement les structures de fonctionnement du point grammatical traité: “la négation simple: *ne...pas*” illustrée par l’exemple, reprise du dialogue de la page précédente “Ce n’est pas possible!” (*ibid.*: 157). L’apprenant est plus tard invité à mettre en pratique ses nouvelles connaissances sur la question à l’aide d’un

exercice oral avec la consigne: “À l’oral, répondez aux questions à la forme négative”. Le modèle fourni est le suivant:

- Vous parlez toujours à vos voisins?
- ⇒ – Non, nous ne parlons plus à nos voisins! (*ibid.* : 158)

Étonnamment, alors que cet ouvrage de référence s’adresse à un public B1, aucune mention n’est faite du fonctionnement oral de la négation durant les six pages d’exemples, d’explications et d’exercices.

Bilan de cet examen

Comme le montre cet examen des présentations de l’oral, qu’il s’agisse d’enregistrements ou de descriptions dans les manuels de FLE (ou de FLM), le biais scriptural dénoncé par Linell (2005) s’avère bien présent. Pour un francophone natif compulsant une grammaire de FLM, la présence d’un tel biais ne porte pas à conséquence dans la mesure où il peut, sans crainte, se fier à sa propre intuition sur la question. En revanche, tel n’est pas le cas de l’apprenant de FLE.

En s’inspirant des grammaires pédagogiques ou linguistiques de FLM, les manuels et les grammaires de FLE ne font que reproduire le biais scriptural hérité de la tradition grammaticale, selon laquelle l’écrit constitue *la* norme. Aussi les descriptions pédagogiques en FLE en viennent-elles paradoxalement à priver l’apprenant d’un nombre important d’informations fondamentales sur le fonctionnement de la langue cible, pouvant alors mener à la situation évoquée par Durán & McCool (2003: 288) dans l’introduction. En ce sens, il semble fondamental d’offrir aux apprenants des outils leur permettant de découvrir les fonctionnements du français oral quotidien tout en ayant la possibilité de l’écouter. C’est à cette finalité que nous avons travaillé en élaborant une ressource d’enseignement-apprentissage du français parlé sur corpus, qui s’inscrit dans la lignée de la première norme pédagogique proposée par Valdman (1989: 21): “[the pedagogical norms] should reflect the actual speech of [target language] speakers in authentic communicative situations”.

PRÉSENTATION DE FLORALE, UNE RESSOURCE POUR L’ENSEIGNEMENT/APPRENTISSAGE DU FRANÇAIS ORAL SUR CORPUS

En réaction aux problèmes évoqués jusqu’ici, nous avons entrepris de créer FLORALE (français langue orale et FLE), une ressource pour l’enseignement/apprentissage sur corpus oral visant à aider les apprenants de FLE (dès le niveau A2) à se construire une représentation pertinente du français parlé tel qu’il se pratique au quotidien en France et dans ses régions limitrophes. Dans ce but, et c’est ce qui fait la spécificité de cette ressource, un nombre important de phénomènes caractéristiques de la langue parlée ont été annotés manuellement pour permettre leur recherche via une interface informatique. FLORALE se distingue ainsi d’autres ressources pédagogiques exploitant des corpus oraux ou audiovisuels comme Fleuron, qui se donne pour but d’“illustrer[r] un ensemble de

situations de la vie d'un étudiant"¹⁴, c'est-à-dire de faire le point sur des situations de communication particulières, ou SACODEYL, un corpus d'entretiens en différentes langues avec des jeunes de 13 à 18 ans, dont les annotations pédagogiques ne sont pas centrées sur les caractéristiques de l'oral¹⁵.

Comme ressources orales, nous avons retenu des émissions radiophoniques provenant de France Culture et constituées, pour la plupart, de documentaires enregistrés sur le vif, comme *Les pieds sur terre*. Dans l'optique pédagogique qui est la nôtre, ces émissions présentent l'avantage d'être en libre-accès sur internet, d'offrir une qualité professionnelle d'enregistrement et, de surcroît, la possibilité de faire entendre des locuteurs d'âges et de niveaux socio-culturels variés, s'exprimant en situation "naturelle"¹⁶.

À l'heure actuelle, nous disposons de cinq heures d'émissions transcrites, segmentées par groupes de cohérence morphosyntaxique et prosodique d'environ deux secondes, alignés sur le signal sonore (au segment) et annotés à l'aide du logiciel ELAN¹⁷. Si "un corpus oral de plusieurs minutes peut suffire pour l'étude de certains phénomènes très récurrents sur le plan phonique" (Boulton & Tyne, 2014: 131), l'objectif à court terme est de doubler la taille du corpus pour massifier les occurrences des phénomènes de moindre fréquence. Constatant qu'aujourd'hui encore "la composante orale est généralement réduite", même dans les grands corpus disponibles, Boulton & Tyne (*ibid.*: 133) remarquent que les transcriptions y sont, de surcroît, "peu soignées" et n'ont bien souvent "pas bénéficié de toutes les considérations théoriques et méthodologiques nécessaires". Pour éviter cet écueil, et d'autant que notre ressource a été créée à des fins pédagogiques et s'adresse à des locuteurs non-natifs, un soin particulier est apporté aux transcriptions de FLORALE: elles se basent sur des normes détaillées et font l'objet de deux supervisions. L'annotation des phénomènes de l'oral répertoriés est, quant à elle, supervisée une fois. On estime ainsi que la transcription, l'annotation pédagogique, et l'ensemble des supervisions d'une minute de données sonores requièrent près de cent minutes de travail et que, sur cette minute d'émission, une trentaine de phénomènes de l'oral sont annotés. Voyons maintenant quels sont ces phénomènes et quelles sont les raisons de leur présentation sur FLORALE.

¹⁴ Fleuron (Français Langue Étrangère Universitaire Ressources et Outils Numériques). URL: <https://apps.atilf.fr/fleuron/>.

¹⁵ SACODEYL (System-Aided Compilation and Open Distribution of European Youth Language). URL: <http://www.um.es/sacodeyl/>.

¹⁶ On ne peut cependant pas exclure le biais possible induit par l'acte même d'enregistrement (voir la problématisation du "paradoxe de l'observateur" chez Gadet, 2007: 38). Signalons que le mixage des émissions et l'impossibilité d'accéder aux métadonnées complètes sur les locuteurs ne constituent guère de problème d'un point de vue pédagogique puisque notre finalité n'est ni l'analyse des interactions dans leur intégralité ni l'étude de la variation dans son sens socio-linguistique.

¹⁷ ELAN: Max Planck Institute for psycholinguistics, Nijmegen, URL: <http://tla.mpi.nl/tools/tla-tools/elan/>.

Les phénomènes de l'oral présentés sur FLORALE

Les phénomènes annotés, et donc recherchables, couvrent quatre dimensions du français oral. La première, phonétique, comprend une trentaine d'items, parmi lesquels on trouve, entre autres, les assimilations de /ʒə/ en /ʃ/ (ʃkɛwa) et /də/ en /t/ devant consonne sourde (patpɛɔblɛm/), la réduction de /ty/ à /t/ devant voyelle (/takɛlaʒ/), le rôle de l'intonation dans la formulation des questions, les liaisons obligatoires, etc. La dimension syntaxique compte, quant à elle, une soixantaine de phénomènes, dont la formation de la phrase négative à l'oral, les dislocations, les (pseudo-)clivées, les usages de *falloir* et *valoir* sans *il* (/fokʃpaʁt/), les verbes introducteurs du discours direct (“il *fait*...”, “il *est là*...”), la parataxe (“je suis sorti il pleuvait”). Pour la dimension conversationnelle, qui comprend une quarantaine de phénomènes, nous avons retenu certains marqueurs discursifs comme, par exemple, les particules *bon*, *du coup*, *quoi*, des reformulateurs (*enfin*, *je dirais*, etc.), des particules d'extension (*un truc comme ça*, *et tout*, *machin*, etc.). Enfin, la dimension lexicale concerne une trentaine de phénomènes regroupant notamment les mots familiers ou argotiques, le verlan, les expressions idiomatiques, les troncations, l'usage particulier de mots comme *genre*, *style*, *limite* (“il était *limite* vulgaire”).

Si certains de ces traits langagiers – la liaison, les questions et leurs spécificités prosodiques – font déjà l'objet de descriptions dans les ouvrages de référence des apprenants de FLE, FLORALE présente l'avantage considérable d'en proposer des illustrations sonores authentiques en grand nombre. D'autres caractéristiques du français parlé, comme l'assimilation de /ʒə/ en /ʃ/, la négation en /pa/ ou l'usage de *on*, ont été retenues car, on l'a vu, elles sont rarement présentées comme courantes dans les grammaires et les manuels de FLE ou de FLM.

D'autres phénomènes encore s'avèrent intéressants en ce qu'ils divergent assez radicalement de ce qu'on trouve dans les écrits normés. Ainsi en est-il, par exemple, de l'usage de certains verbes introducteurs de discours direct comme *faire*: “le prof de physique il *fait* Théo un génie le mec il dit un génie le mec s'il te plaît”¹⁸. Enfin, nous avons choisi de présenter certains traits langagiers souvent difficiles à expliquer pour l'enseignant, le didacticien, voire le linguiste, comme les différents usages des marqueurs discursifs *quoi* ou *en fait*. Dans ce cas de figure, le pari de FLORALE est celui d'aider l'apprenant à se forger progressivement une intuition des phénomènes concernés grâce à la massification des occurrences orales ciblées auxquelles le corpus donne accès et grâce aux activités épilinguistiques¹⁹ que leur exploration peut susciter. Si, à l'instar de Cuq & Gruca (2002: 358), on considère qu'une grammaire de FLE doit “avoir pour priorité d'aider par le raisonnement à accélérer l'appropriation linguistique, et à créer des réflexes intuitifs”, il semble raisonnable de faire l'hypothèse que la confrontation à un grand nombre d'exemples oraux authentiques sera susceptible d'établir progressivement ces

¹⁸ Extrait de France Culture, *Les Pieds sur terre* (22/12/2016).

¹⁹ L'épilinguistique représente l'“activité métalinguistique non-consciente” (Culioli, 1999: 19).

“réflexes intuitifs” permettant à l’apprenant de comprendre et, au final, de s’approprier l’usage des phénomènes de l’oral en question²⁰.

Une fois résolue l’épineuse question du choix des phénomènes à annoter pour construire notre base de données, l’autre grand défi portait sur la manière de les présenter à l’apprenant de FLE. Comment lui permettre d’accéder à l’ensemble des phénomènes annotés? Quelle interface informatique proposer pour une navigation (si possible) aisée au sein du corpus²¹?

L’allure et les fonctions de l’interface informatique

Dès l’origine du projet FLORALE, nous avons tenu à réfléchir au défi consistant à proposer une interface informatique susceptible d’être utilisée par des apprenants débutants (idéalement dès le niveau A2). Durant tout le processus de conception s’est alors posée la question de savoir à quel type de métalangage recourir pour présenter les phénomènes de l’oral.

Sur la page d’accueil, le mode de recherche paramétré par défaut implique une navigation au clic dans les quatre principaux types de phénomènes. Dans un usage minimaliste du métalangage, ceux-ci sont nommés *Prononciation*, *Structures de phrases*, *Outils de la conversation*, et *Mots et expressions*. Ils sont identifiés par un code couleur sur les différentes pages du site.

Les phénomènes qui dépendent de chacune des quatre grandes catégories sont présentés sous la forme d’une question simple, formulée dans un vocabulaire voulu accessible tout en restant suffisamment clair (de même pour les réponses à ces questions). L’apprenant trouve ainsi l’assimilation de /ʒə/ en /ʃ/ via la question: “Comment sont souvent prononcés les pronoms *je*, *tu*, *il(s)*, et des fois *elle(s)*?”, dont une des réponses est “*je* est prononcé /ʃ/ devant les consonnes /p/, /t/, /k/, /f/, /s/, /ʃ/”.

Comme pour tous les autres phénomènes répertoriés, l’apprenant peut écouter un exemple sonore extrait de la base de données et lire parallèlement sa transcription orthographique, enrichie au besoin d’indications en alphabet phonétique international. Dans le futur, pour chaque phénomène, nous prévoyons d’inclure des fichiers d’explications pédagogiques téléchargeables sous format PDF, avec exemples sonores et indication de références théoriques.

²⁰ Ce processus ferait écho au processus d’acquisition du bébé. Sur le lien entre la fréquence d’exposition à une construction et la représentation cognitive de la langue, Bybee (2008: 217-218) écrit: “When constructions are combined with a Usage-Based model the result is a theory that proposes that grammatical structures are built up through experience with specific examples of constructions which are categorized in memory by a mapping process that matches strings for similarity and difference. The resulting cognitive representations are abstractions over one’s cumulative experience with language. However [...] there is evidence that knowledge of specific instances of language use is not entirely lost in this abstraction process and especially with reinforcement through repetition, specific instances of constructions can have memory representations. In this view of grammar, then, frequency of use plays an important role in determining cognitive structures”.

²¹ Pour des raisons éditoriales, il n’a pas été possible d’introduire des captures d’écran pour présenter l’interface décrite dans ce qui suit. Nous renvoyons les lecteurs à: <http://florale.unil.ch>.

Pour accéder à toutes les occurrences similaires contenues dans la base de données FLORALE, il suffit à l'apprenant de cliquer sur l'exemple pour lancer la recherche. L'ensemble des résultats s'affiche alors sous la forme d'une liste numérotée de segments transcrits, où le phénomène recherché est mis en évidence en couleur et identifié dans une infobulle.

L'apprenant peut alors, non seulement écouter les extraits sonores correspondants, mais aussi les découvrir dans leur contexte élargi, ou encore accéder à une page proposant l'écoute de l'intégralité de l'émission radiophonique dont ils proviennent, avec défilement de sa transcription orthographique et possibilité de mettre en évidence les différentes occurrences d'un phénomène choisi

En recourant au mode de présentation "Extrait", l'apprenant peut se familiariser avec le sens et le contexte d'apparition des phénomènes qui l'intéressent. Possibilité lui est donnée d'étendre vers la gauche et/ou vers la droite la durée de l'extrait sonore. Comme sur toutes les pages de l'interface, il lui suffira de double-cliquer sur un mot pour accéder aux définitions (synonymes, conjugaisons, etc.) du dictionnaire Alexandria²². Des réglages plus fins sont en outre disponibles pour la lecture du son, susceptible d'être ralenti, ou lu en boucle, l'objectif de ces paramètres étant de favoriser la compréhension et, idéalement, sur le plus long terme, selon la difficulté, l'appropriation du phénomène concerné.

Ce premier mode de recherche, voulu très intuitif, doit garantir une utilisation de FLORALE avec un guidage minimal de l'enseignant. Il en existe d'autres, disponibles sur la page "Recherche", qui devraient permettre à l'enseignant de créer différents types d'exercices et de tâches d'apprentissage.

Le mode "Chercher un phénomène de l'oral" donne une vue d'ensemble de tous les phénomènes annotés autorisant dès lors l'accès rapide à l'un d'entre eux. Le bouton "Chercher une émission" permet, quant à lui, de parcourir les diverses thématiques présentes dans le corpus; il propose, par ailleurs, des outils de recherche sur la base des métadonnées des locuteurs (sexe, âge, origine géographique, catégorie socioprofessionnelle). Enfin, "Chercher un mot" offre la possibilité de chercher un (groupe de) mot(s) ou un morphème dans l'ensemble du corpus. À titre d'illustration, voyons quelques utilisations possibles. En combinant ces différents modes de recherche, il est envisageable, par exemple, d'étudier la prononciation des adverbes en *-emment* (/amã/), de se faire une idée de la fréquence des variantes pour les différents types de négations (*ne...pas* ou *pas*), d'écouter les diverses manières dont *il y a* est articulé (/ilija/, /ilja/ ou /ja/), ou encore de voir si *il faut* s'utilise plus souvent que *faut* seul.

Le projet FLORALE n'étant qu'à ses débuts, nombre de questions restent en suspens. Quelle en sera la perception par les apprenants? Quel usage en feront-ils? Pourront-ils parvenir à mener des requêtes en fonction de leur souhait originel? Miser sur leur intuition en cas de difficulté à fournir des explications claires sur un phénomène donné est-il raisonnable? Si l'écoute de nombreux exemples similaires

²² URL: <http://www.memodata.com/2004/fr/alexandria/install.shtml>.

permet d'atteindre en quelques minutes une fréquence élevée et devrait, *a priori*, renforcer la compréhension du phénomène en question, leur permettra-t-elle, sur le long terme, de mieux l'entendre dans des conditions d'écoute normale sans la transcription? En viendront-ils eux-mêmes à utiliser certaines de ces formulations de manière socio-linguistiquement appropriée? Dans le but de répondre en partie à ces interrogations, un suivi des expériences d'utilisation de FLORALE par des apprenants s'avère nécessaire. Ce travail ne fait que commencer. Il s'agira, dans un premier temps, de recueillir les observations des enseignants guidant l'utilisation de FLORALE ou supervisant son usage en autonomie et d'enquêter auprès de leurs apprenants.

Conclusion

Comme l'illustre notre anecdote de départ, pour plusieurs raisons évoquées dans cet article, les modèles auxquels l'apprenant de FLE est traditionnellement exposé ne lui permettent pas de se construire une représentation vraiment pertinente des fonctionnements du français oral quotidien, ce qui n'est pas sans conséquence sur sa capacité à comprendre les locuteurs natifs en situation naturelle et, *a fortiori*, à s'approprier l'usage de certains phénomènes caractéristiques de la langue parlée. Nous espérons avoir montré que l'utilisation d'une ressource d'enseignement-apprentissage du français parlé sur corpus telle que FLORALE – novatrice quant aux outils de recherche proposés et présentant un nombre inédit de phénomènes caractéristiques du français parlé – est susceptible de pallier certains de ces problèmes.

Les apprenants et les enseignants de FLE se saisiront-ils efficacement de cette ressource? Il y a vingt ans, Sinclair s'enthousiasmait devant les perspectives ouvertes par les corpus, amenés selon lui à occuper rapidement une place centrale dans les pratiques d'enseignement-apprentissage. Il prévoyait que les jeunes étudiants y ayant accès s'empareraient des corpus pour se former en autonomie, et anticipait surtout avec plaisir l'émergence de "problèmes" pour la didactique des langues: "problems will arise when the textual evidence does not fit the precepts of classroom and textbook, and the mythology will prove no match for the facts" (Sinclair, 1997: 30). Force est de constater que sa prévision était fautive. Qu'en sera-t-il dans vingt ans? Aura-t-on assisté à une démocratisation de l'enseignement/apprentissage du français oral sur corpus? Nul doute qu'un élément de réponse réside dans l'évolution des représentations de la langue que se font les didacticiens, les enseignants et les apprenants. À cet égard, en guise de conclusion, citons cette remarque orale d'une participante à un colloque alors que nous défendions le point de vue selon lequel il est fondamental d'enseigner la négation orale en /pa/ dès le début de l'apprentissage puisqu'elle constitue la forme la plus fréquente. Dans la situation relativement formelle d'un échange entre inconnus au cours d'une manifestation scientifique, cette participante s'est alors exclamée: "moi, j'oserais pas!"

Références bibliographiques

- ABE, DANIELLE, FRANCIS CARTON, SAM MICHEL CEMBALO & ODILE RÉGENT. 1979. Didactique et authentique: du document à la pédagogie. *Mélanges Crapel* 9, 1-14.
- ASHBY, WILLIAM J. 1984. The elision of /l/ in French clitic pronouns and articles. *Romanitas. Studies in Romance Linguistics* 4, 1-16.
- AUROUX, SYLVAIN, JACQUES DESCHAMPS & DJAMEL KOULOUGHLI. 2004. *La philosophie du langage*. Paris: Presses universitaires de France.
- BENTO, MARGARET. 2007. Le français parlé: une analyse de méthodes de français langue étrangère. MICHAËL ABECASSIS, ÉLODIE VIALLETON & LAURE AYOSSO (dir.). *Le français parlé au 21^{ème} siècle: normes et variations dans les discours et en interaction. Annales du Colloque d'Oxford (juin 2005). Volume 2*. Paris: L'Harmattan, 191-212.
- BLANCHE-BENVENISTE, CLAIRE & MIREILLE BILGER. 1999. Français parlé-oral spontané. Quelques réflexions. *Revue française de linguistique appliquée* 4: 2, 21-30.
- BLANCHE-BENVENISTE, CLAIRE. 2003. La langue parlée. MARINA YAGUELLO (dir.) *Le grand livre de la langue française*. Paris: Seuil, 317-344.
- BLED, ÉDOUARD, ODETTE BLED & DANIEL BERLION. 2010. *Le Bled*. Paris: Hachette.
- BOULTON, ALEX & HENRY TYNE. 2014. *Des documents authentiques aux corpus. Démarches pour l'apprentissage des langues*. Paris: Didier.
- BROGNAUX, SANDRINE & MATHIEU AVANZI. 2015. Sociophonetics of phonotactic phenomena in French. *Proceedings of the international conference on phonetic sciences (ICPhS)*, 1-5.
- BYBEE, JOAN. 2008. Usage-based grammar and second language acquisition. PETER ROBINSON & NICK C. ELLIS (dir.). *Handbook of cognitive linguistics and second language acquisition*. New York: Routledge, 216-236.
- CAPELLE, GUY. 2016. Enseigner une grammaire de l'oral aux débutants. *Le français dans le monde* 406, 26-27.
- CATACH, NINA, CLAUDE GRUAZ & DANIEL DUPREZ. 1995/2010. *L'orthographe française. Traité théorique et pratique avec des travaux d'application et leurs corrigés*. Paris: Nathan.
- CLARKE, CHARLES C. 1918. The phonograph in modern language teaching. *The Modern Language Journal* 3: 3, 116-122.
- CULBERTSON, JENNIFER. 2010. Convergent evidence for categorial change in French: from subject clitic to agreement marker. *Language* 86: 1, 85-132.
- CULIOLI, ANTOINE. 1999. *Pour une linguistique de l'énonciation. Formalisation et opérations de repérage. Tome 2*. Paris/Gap: Ophrys.

- CUQ, JEAN-PIERRE & ISABELLE GRUCA. 2002. *Cours de didactique du français langue étrangère et seconde*. Grenoble: Presses universitaires de Grenoble.
- DELATOUR, YVONNE, DOMINIQUE JENNEPIN, MAYLIS LÉON-DUFOUR & ANNE MATTLE-YEGANEH. 1991. *Grammaire du français*. Paris: Hachette.
- DURÁN, RICHARD & GEORGE MCCOOL. 2003. If this is French, then what did I learn in school?. *The French Review* 77: 2, 288-299.
- ELLIS, ROD. 1997. *Second language acquisition*. Oxford: Oxford University Press.
- FAYOL, MICHEL & JEAN-PIERRE JAFFRÉ. 2014. *L'orthographe*. Paris: Presses Universitaires de France. Que sais-je?
- FERGUSON, CHARLES A. 1975. Toward a characterization of English foreigner talk. *Anthropological Linguistics* 17: 1, 1-14.
- FONSECA-GREBER, BONNIE & LINDA R. WAUGH. 2003. The subject clitics of conversational European French. RAFAEL NÚÑEZ-CEDENO, LUIS LÓPEZ, & RICHARD CAMERON (dir.). *A romance perspective on language knowledge and use. Selected papers from the 31st linguistic symposium on romance languages (LRS), Chicago, 19-22 April 2001*. Amsterdam/Philadelphia: John Benjamins, 99-117.
- GADET, FRANÇOISE. 1989. *Le français ordinaire*. Paris: Armand Colin.
- GADET, FRANÇOISE. 2007. *La variation sociale en français*. Paris: Ophrys.
- GIROUD, ANICK & CHRISTIAN SURCOUF. 2016. De “Pierre, combien de membres avez-vous?” à “Nous nous appelons Marc et Christian”: réflexions autour de l'authenticité dans les documents oraux des manuels de FLE pour débutants. *Congrès Mondial de Linguistique Française 2016*, 1-18.
- GIROUD, ANICK, CHRISTIAN SURCOUF & MALIKA BEN HARRAT. 2018. Conceptualiser à partir d'un corpus oral: l'influence de la “doxa” grammaticale. *Recherches en didactiques des langues et des cultures. Cahiers de l'Acedle* 15: 1. URL: <http://journals.openedition.org/rdlc/2703>.
- GLAUD, LUDIVINE, YVES LOISEAU, ÉLISE MERLET, MARION PERRARD & ODILE RIMBERT. 2015. *Grammaire essentielle du français B1*. Paris: Didier.
- HAGÈGE, CLAUDE. 1985. *L'homme de paroles*. Paris: Fayard.
- HAZAN, VALÉRIE, MARIA UTHER & SONIA GRANLUND. 2015. How does foreigner-directed speech differ from other forms of listener-directed clear speaking styles? *Proceedings of the 18th International Congress of Phonetic Sciences*. URL: <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0346.pdf>.
- HENZL, VĚRA M. 1979. Foreign talk in the classroom. *International Review of Applied Linguistics* 17: 2, 159-167.
- HOLEC, HENRI. 1990. Des documents authentiques, pour quoi faire? *Mélanges pédagogiques*, 65-74.

- HONVAULT-DUCROCQ, RENÉE (dir.). 2006. *L'orthographe en questions*. Mont-Saint-Aignan: Publications des universités de Rouen et du Havre.
- KLEIN, WOLFGANG. 1987/1989. *L'acquisition de langue étrangère*. Paris: Armand Colin (traduction de *Zweitspracherwerb: Eine Einführung* par COLETTE NOYAU).
- KOCH, PETER & WULF OESTERREICHER. 2001. Langage parlé et langage écrit. GÜNTER HOLTUS, MICHAEL METZELTIN & CHRISTIAN SCHMITT (dir.). *Lexikon der Romanistischen Linguistik. Band I/2 - Methodologie (Sprache in der Gesellschaft / Sprache und Klassifikation / Datensammlung und verarbeitung)*. Tübingen: Max Niemeyer, 584-627.
- LINELL, PER. 2005. *The written language bias in linguistics: its nature, origins, and transformations*. New York: Routledge.
- MAINGUENEAU, DOMINIQUE. 1994/1999. *L'énonciation en linguistique française*. Paris: Hachette.
- MARTINET, ANDRÉ. 1960/2008. *Éléments de linguistique générale*. Paris: Armand Colin.
- NEVEU, FRANCK. 2004. *Dictionnaire des sciences du langage*. Paris: Armand Colin.
- POISSON-QUINTON, SYLVIE, REINE MIMRAM & MICHÈLE MAHÉO-LE COADIC. 2002. *Grammaire expliquée du français. Niveau intermédiaire*. Paris: CLE International.
- POISSON-QUINTON, SYLVIE, CÉLYNE HUET-OGLE, ROXANE BOULET & ANNE VERGNE-SIRIEYS. 2003. *Grammaire expliquée du français. Niveau débutant*. Paris: CLE International.
- RAVAZZOLO, ELISA, VÉRONIQUE TRAVERSO, ÉMILIE JOUIN & GÉRARD VIGNER. 2015. *Interactions, dialogues, conversations: l'oral en français langue étrangère*. Paris: Hachette.
- RIEGEL, MARTIN, JEAN-CHRISTOPHE PELLAT & RENÉ RIOUL. 1994. *Grammaire méthodique du français*. Paris: Presses universitaires de France.
- RIEGEL, MARTIN, JEAN-CHRISTOPHE PELLAT & RENÉ RIOUL. 2009. *Grammaire méthodique du français*. Paris: Presses universitaires de France.
- SAUSSURE (DE), FERDINAND. 1916/1994. *Cours de linguistique générale*. Paris: Payot.
- SINCLAIR, JOHN. 1997. Corpus evidence in language description. ANNE WICHMANN STEVEN FLIGELSTONE, TONY MCENERY & GERRY KNOWLES (dir.). *Teaching and language corpora*. London: Routledge, 27-39.
- SURCOUF, CHRISTIAN & ANICK GIROUD. 2016. À quelle langue accède l'apprenant? Examen critique du traitement de l'oral dans les premières leçons de manuels de français langue étrangère pour débutants. *Linguistik Online* 78: 4, 11-27.
- VALDMAN, ALBERT. 1989. The elaboration of pedagogical norms for second language learners in a conflictual diglossia situation. SUSAN GASS, CAROLYN MADDEN, DENNIS PRESTON & LARRY SELINKER (dir.). *Variation in second*

language acquisition. Volume 1: Discourse and pragmatics. Clevedon: Multilingual Matters, 15-34.

VIALLETON, ÉLODIE & TIM LEWIS. 2014. Reconsidering the authenticity of speech in French language teaching: theory, data, methodology, and practice. HENRY TYNE, VIRGINIE ANDRE, CHRISTOPHE BENZITOUN, ALEX BOULTON & YAN GREUB (dir.). *French through corpora: ecological and data-driven perspectives in French language studies.* Newcastle-upon-Tyne: Cambridge Scholars Publishing, 293-316.

WAUGH, LINDA R. & BONNIE FONSECA-GREBER. 2002. Authentic materials for everyday spoken French: corpus linguistics vs. French textbooks. *Arizona Working Papers in SLAT 9*, 114-127.

WEBER, CORINNE. 2006. Pourquoi les Français ne parlent-ils pas comme je l'ai appris? *Le français dans le monde* 345, 31-33.