



<https://bulletin.vals-asla.ch> · e-ISSN: 2813-7973 · p-ISSN: 1023-2044

Thomas, A. (2021). Corpus et exercices en ligne, complémentarité ou redondance? L'exemple des pronoms relatifs qui et que. *Bulletin suisse de linguistique appliquée* (Numéro spécial 2), 85-105. <https://doi.org/10.26034/ne.vals-asla.2021.1267>

Cet article est publié sous la licence *Creative Commons Attribution 4.0 International* (CC BY):
<https://creativecommons.org/licenses/by/4.0>



© Anita Thomas, 2021

Corpus et exercices en ligne, complémentarité ou redondance? L'exemple des pronoms relatifs *qui* et *que*

Anita THOMAS

Université de Fribourg

Département de Plurilinguisme et Didactiques des Langues Étrangères

Avenue de l'Europe 20, 1700 Fribourg, Suisse

anita.thomas@unifr.ch

The aim of this article is to discuss the use of online corpora as a resource for foreign language teaching. Corpora allow access to the language in its use, which can be considered more authentic than that found in exercises. This study is based on an analysis of the French relative pronouns *qui* ('who') and *que* ('which/that') in their most basic use. An analysis of data from learners of L2 French indicates a preference for the use of *qui* with animate antecedents and *que* for inanimates, but there are differences according to the tasks and the learners' L1. Correct production of relative pronouns seems to be more difficult in written than in oral production and less correct with *que* than with *qui*. A comparison between the teaching material available on the *Point du FLE* website and a sample of the ScienQuest written corpus reveals a similar distribution of the two relative pronouns. The discussion is focussing on the contribution of the comparison of different data sources for the development of corpus-informed teaching materials.

Keywords:

corpus informed teaching, learner corpus research, French L2, relative clause, foreign language teaching.

Mots-clés:

corpus et didactique, corpus d'apprenants, FLE, pronoms relatifs, didactique des langues étrangères.

1. Introduction: les corpus comme input de qualité

Pendant longtemps l'enseignement des langues étrangères a été basé sur le matériel fourni dans le manuel de cours ainsi que sur celui transmis par l'enseignant·e. Les personnes qui souhaitent avoir plus d'input, sans se rendre dans une région où la langue est parlée, pouvaient essayer de lire des romans, des journaux ou écouter la radio. À l'ère digitale, toutes ces ressources supplémentaires sont disponibles en un clic. Cette situation est une aubaine pour les apprenant·e·s motivé·e·s qui souhaitent s'exposer le plus possible à la langue cible. Pour les enseignant·e·s, elle demande aussi des compétences pour évaluer la qualité des ressources à disposition et parfois aussi pour rendre ces ressources accessibles en classe de langue. Il est donc important de mener des études évaluant la qualité de celles-ci.

L'objectif de l'étude présentée dans cet article est d'analyser et de discuter la qualité de ressources disponibles en ligne, et plus particulièrement de l'apport des corpus de langue cible comparé à des exercices en libre accès. Pour cela, je me concentrerai sur l'enseignement-apprentissage d'un phénomène précis

en français langue étrangère, à savoir les pronoms relatifs *qui* et *que* dans leur utilisation la plus simple. L'objectif de la comparaison entre les deux types de ressource est de contribuer à la discussion sur l'utilisation des corpus pour l'enseignement, en termes de qualité d'input.

2. Considérations théoriques

2.1 Corpus et enseignement des langues étrangères

Le nombre de publications traitant de l'utilisation des corpus pour l'enseignement n'a cessé d'augmenter durant les dix dernières années, à tel point qu'on a vu en 2017 la parution d'une *timeline* et d'une méta-analyse (Boulton & Cobb 2017). Par corpus, on entend traditionnellement "a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language" (Sinclair 1996: 4). Alors que les corpus sont des outils utilisés depuis longtemps pour l'analyse des conversations (voir par exemple le corpus CLAPI¹) ou du développement langagier (voir par exemple la base de données CHILDES²), ce n'est qu'avec l'avènement d'internet que les corpus ont été rendus accessibles et donc exploitables par le public. L'accès à de larges collections de langue dans son usage a donné lieu à une forme de regain d'intérêt pour la description linguistique, cette fois à partir de données empiriques (Tyne et al. 2014; Williams 2005), renouvelant ainsi la traditionnelle opposition entre la langue comme système et la langue comme corpus (Mayaffre 2005). Comme l'expliquent Boulton & Tyne (2014: 42-45), la linguistique traditionnelle est plutôt basée sur une réflexion menant à une théorie qui reflète la langue comme système. Dans cette approche, l'introspection et l'intuition jouent un rôle important. La linguistique de corpus se base en revanche sur l'observation de données qui reflètent la langue dans son usage, ce qui implique un certain niveau de variation. Ces descriptions-ci sont par contre basées sur des données empiriques. D'une certaine manière, ces deux approches correspondent aussi aux deux sortes de ressources traitées dans la présente étude: les exercices sont souvent construits à partir d'une réflexion sur la grammaire de la langue cible, alors que les données de corpus correspondent à des données d'usage. Ainsi, dans une visée de didactique des langues étrangères, plutôt que de considérer ces deux approches de manière opposée, on pourrait les voir comme complémentaires comme le suggèrent Boulton & Tyne (2014: 42).

Dans sa *research timeline* sur l'utilisation des corpus pour l'enseignement, Boulton (2017) présente 46 articles publiés touchant à différents aspects, tels que les attitudes des apprenant·e·s par rapport aux corpus, la manière d'utiliser les corpus ainsi que les effets de ceux-ci sur l'apprentissage ou comme

¹ <http://clapi.ish-lyon.cnrs.fr> (page consultée le 10 mars 2021).

² <https://childes.talkbank.org/access/> (page consultée le 10 mars 2021).

ressource de référence. Les avantages présupposés de l'utilisation des corpus sont liés au fait que "such contact [with language through corpora] provides the massive contextualised exposure needed for language learning, but in a more controlled way than purely haphazard exposure via regular reading or listening, thus promoting or enhancing noticing, language awareness, autonomy, and ultimately producing 'better learners'." (Boulton 2017: 483).

Globalement, les résultats s'avèrent plutôt positifs. La *timeline* révèle que les corpus sont surtout utilisés pour le développement de l'écrit (académique), dans le cadre de séquences dans lesquelles les apprenant·e·s sont invité·e·s à comparer leur production avec des textes de corpus. Elle révèle aussi que l'utilisation des corpus donne de bons résultats avec des apprenant·e·s débutant·e·s. La méta-analyse de Boulton & Cobb (2017) confirme les aspects positifs de l'utilisation des corpus pour l'enseignement, en particulier pour le développement du vocabulaire et de *patterns* lexico-grammaticaux. Par contre, les études ayant examiné l'effet de l'utilisation des corpus sur le développement de la grammaire sont rares (Boulton & Cobb 2017: 383). Néanmoins, cette méta-analyse permet aux auteurs de constater que: "... we reach the somewhat surprising and possibly encouraging conclusion that DDL [Data Driven Learning] works pretty well in almost any context where it has been extensively tried." (Boulton & Cobb 2017: 386).

D'autres études actuelles soulignent les avantages de l'utilisation des corpus pour l'enseignement des langues étrangères: ils permettent notamment d'avoir un contact avec la variation et la culture de la langue cible (André 2018; Gajo et al. 2019: 189ss) et on voit de plus en plus d'études traitant du développement de l'enseignement des langues étrangères sur la base de données d'usage de la langue (Larrivée & Lefevre 2020; Tyler et al. 2018).

Un autre avantage de l'utilisation des corpus est qu'ils permettent "l'observation et la description des usages authentiques de la langue." (Jacques 2005: 24 en gras dans le texte). L'aspect authentique des corpus est également un argument pour leur utilisation en didactique des langues étrangères (Boulton & Tyne 2014: 22; Gajo et al. 2019: 189ss; Gilmore 2004; Vialleton & Lewis 2014). De ce point de vue, les corpus permettent un contact direct avec une utilisation dite réelle de la langue, ce qui est considéré par beaucoup comme motivant pour les apprenant·e·s. Le contact avec la langue dans son utilisation permet aussi aux enseignant·e·s de vérifier leurs intuitions au niveau de la langue dans son usage. Finalement, l'utilisation des corpus pourrait être particulièrement pertinente pour l'enseignement de l'oral et de l'interaction; paradoxalement, c'est un aspect de la langue qui est souvent sous- ou mal représenté dans les manuels de langue (Ravazzolo & Etienne 2019).

Mais l'utilisation des corpus comme matériel authentique ne présente pas que des avantages. Une des caractéristiques des corpus est qu'ils contiennent beaucoup de "bruit", des données qui ressemblent à ce que l'on cherche mais

qui ne le sont pas. Ainsi le travail avec les corpus demande un effort de sélection des données ou ce que Cuq & Gruca appellent une analyse "prépédagogique" (Cuq & Gruca 2005: 433). Les utilisateurs sont aussi confrontés à des difficultés techniques et il est parfois difficile de trouver ce que l'on cherche. Il s'agit donc d'une approche qui peut s'avérer chronophage. Ainsi, malgré l'enthousiasme pour l'utilisation des corpus comme ressource didactique, il est nécessaire de voir dans quelle mesure ils sont complémentaires à des exercices en libre accès, voire meilleurs que ceux-ci. Cette étude a ainsi pour but de comparer ces deux sources d'input pour les apprenant·e·s. Pour cela, je prendrai l'exemple d'un phénomène grammatical précis, à savoir les pronoms *qui* et *que* dans leur utilisation la plus simple (voir plus bas).

2.2 Apprentissage d'une langue étrangère et simplification

Les recherches sur l'apprentissage des langues étrangères ont montré que les apprenant·e·s passent par différents stades de production de la L2 et que ces stades reflètent différentes manières de traiter la langue cible. Parmi celles-ci, on trouve des processus de simplification de la langue cible (Ellis et al. 2016; Ortega 2009: 116ss), probablement pour pallier la surcharge cognitive liée au traitement de la nouvelle langue à apprendre. Cette simplification donne lieu à l'assignation d'une forme par fonction, un processus connu en tant que *one-to-one principle* (Andersen 1984). Ce processus est basé sur la sensibilité à la fréquence des formes et aux biais disponibles dans l'input. Selon ce principe, les apprenant·e·s en début d'apprentissage simplifient la langue cible en reproduisant ces biais, ce qui mène à une sur-généralisation des formes. Ce n'est qu'avec une certaine expérience de la langue (Bybee 2008) que les apprenant·e·s réussiront à percevoir des distinctions plus fines qui mèneront à leur tour à des restructurations et ainsi à une progression vers une langue reflétant de plus en plus les usages de la langue cible.

Certaines études suggèrent que les apprenant·e·s ont parfois de la peine à quitter des réflexes de simplification et de sur-généralisation (cf. fossilisation Klein & Perdue 1997). En ce sens, la classe de langue est une chance, car le matériel didactique proposé peut être construit de manière à soutenir la progression en veillant à ce que les apprenant·e·s aient la possibilité de dépasser les simplifications et les sur-généralisations en leur présentant plus spécifiquement ce qu'ils ne semblent pas noter (Collins et al. 2009; Lyster & Rebuffot 2002).

2.3 Les pronoms relatifs simples *qui* et *que*

Cette étude se concentre sur l'utilisation des deux pronoms relatifs simples *qui* et *que* (Riegel et al. 1994: 794). Néanmoins, la présentation comprendra aussi quelques autres formes, en particulier dans le but de montrer les difficultés auxquelles les apprenant·e·s du français L2 doivent faire face.

La fonction des pronoms relatifs est d'introduire une proposition relative et, dans les cas les plus simples, le pronom relatif reprend un antécédent qui se trouve dans la phrase principale³. En français, le choix du pronom relatif est lié à sa fonction, de sujet, par exemple "les livres *qui* sont sur la table me dérangent" ou d'objet, par exemple "les étudiants *que* je connais sont très engagés", dans la phrase subordonnée. Le pronom relatif simple ne s'accorde pas en genre, en nombre ou en personne mais reprend néanmoins les accords qu'on aurait eu à faire avec l'antécédent (Riegel et al. 1994: 795), par exemple "les photos que j'ai prises...".

Utilisé seul, le pronom relatif *qui* est utilisé que l'antécédent soit animé/humain ou non (Riegel et al. 1994: 794). Par contre, dans sa fonction de pronom interrogatif (direct ou indirect), par exemple "*Qui* reprend du dessert ?", "je me demande *qui* je vais voir ce soir", ou lorsqu'il se trouve dans une construction prépositionnelle, par exemple "voici la collègue avec *qui* je ne parle jamais", *qui* représente toujours un humain (Riegel et al. 1994: 794) (ou un animal) quelle que soit la fonction du pronom (sujet ou objet).

Le mot *que* n'est pas seulement un pronom relatif, il est également utilisé comme conjonction de subordination (Riegel et al. 1994: 788, 824), par exemple "je pense *que* la vue est belle". Selon la liste des mots les plus fréquents publiée sur le site *Eduscol*⁴ du ministère de l'éducation nationale, *que* comme conjonction serait presque deux fois plus fréquent que *que* comme pronom (176'161 *que* conjonction versus 99'412 *que* pronom; *qui* pronom 148'392).

Les recherches sur l'apprentissage des relatives suggèrent que la production des pronoms *qui* et *que* comme relatifs apparaît assez tôt, au stade 2 selon l'itinéraire d'apprentissage du français L2 proposé par Bartning & Schlyter (2004). Selon Bartning & Kirchmeyer (2003: 15), l'itinéraire de l'apprentissage des relatives se fait en trois étapes principales. Dans la première, les apprenant·e·s utilisent une forme, de base ou passe-partout, qui sert pour toutes les fonctions. Pour le français L2, il s'agit de *qui* (alors qu'en italien c'est le *che*). Selon les auteures, "[c]ette surgénéralisation de *qui* peut être liée à la saillance de la forme, la plus fréquente dans l'input." Dans un deuxième temps, les apprenant·e·s utilisent une forme pour une fonction, donc le *one-to-one principle* décrit plus haut⁵ et, finalement, ils/elles réussissent à utiliser les formes dans leurs diverses fonctions. Cette progression est basée sur des

³ Voir Cappeau (2015) pour une discussion concernant les différentes manières de considérer les relatives, ainsi que sa proposition de les traiter comme des expansions du groupe nominal (tel un adjectif).

⁴ https://cache.media.eduscol.education.fr/file/ecole/20/6/liste-mots-par-frequence_115206.pdf (page consultée le 10 mars 2021).

⁵ A noter que Bartning & Kirchmeyer (2003) s'intéressent à des fonctions discursives dans leur étude.

apprenant·e·s suédophones du français, mais elle a été confirmée par des études sur des apprenant·e·s japonais·e·s (voir Trévisiol-Okamura 2015: 109).

Les études sur le français L1 ont montré que les relatives en *qui* ainsi que les clivées forment la majorité des relatives produites (Welcomme 2013). Roubaud (2015: 133) montre que les enfants des classes du primaire utilisent les relatives avec des antécédents humains dans près de 80% des cas. Elle signale que ces données sont néanmoins toutes basées sur des tâches dans lesquelles on demande de décrire des situations liées à des personnes. Selon elle, des tâches de type explicatif ou argumentatif pourraient mener à la production d'autres pronoms.

Welcomme (2013) a mené une étude sur les relatives à l'oral avec trois groupes d'apprenant·e·s néerlandophones, âgés entre 12 et 18 ans, ayant eu entre 2 et 7 ans de français à l'école. Les résultats de son étude, basés sur une tâche de narration orale (*Frog Story*), montrent que la production des relatives augmente avec le niveau scolaire. Les relatives introduites par *qui* sont de loin les plus fréquentes alors que celles introduites par *que* sont peu fréquentes. Le nombre de relatives en *qui* augmente nettement selon le niveau des groupes, alors que le nombre de relatives en *que* stagne, un résultat qui reflète la fréquence de ces formes en français.

En résumé, le pronom relatif *qui* apparaît comme le pronom le plus fréquent et peut être considéré comme le pronom relatif par excellence. On s'attend donc à une production plus fréquente et correcte de celui-ci.

Étonnamment, il n'existe pas, à ma connaissance, d'étude qui s'est penchée sur l'influence de l'antécédent pour le choix du pronom relatif par les apprenant·e·s. Cet élément est pourtant fortement mis en avant dans les grammaires pour apprenant·e·s du FLE (voir par exemple Delatour et al. 2004: 204), même si, ou justement parce que l'antécédent ne joue pas de rôle déterminant pour le choix des pronoms *qui* et *que* dans leur utilisation la plus simple. Pourtant, pour des raisons de construction syntaxique – l'antécédent précède le pronom – l'antécédent pourrait être saillant, surtout pour les apprenant·e·s ayant une L1 ou une L2 (par exemple l'anglais) dans laquelle le pronom prend différentes formes selon la nature de l'antécédent.

Ainsi, à partir des études antérieures sur l'apprentissage des pronoms relatifs, on pourrait s'attendre à ce que les apprenant·e·s commencent par utiliser le pronom *qui* comme forme par défaut. Ensuite, par extension de l'utilisation de *qui* comme interrogatif pour les humains et de l'utilisation de *que* pour "d'autres choses" (il est peu probable que les apprenant·e·s débutant·e·s réussissent à faire la différence entre le *que* relatif et le *que* complétif), notamment les non animés. Si tel est le cas, les inputs et exercices devraient fournir aux apprenant·e·s de nombreuses occasions de réaliser que cette distribution est

simplificatrice et attirer l'attention sur la fonction des pronoms dans la subordonnée.

Ainsi, dans le but de discuter du possible apport des corpus comme ressource pour l'enseignement des langues, cette étude traitera les trois questions suivantes:

1. Observe-t-on une distribution dans l'utilisation des pronoms *qui* et *que* selon leur antécédent dans les données d'apprenant·e·s ?
2. La distribution des formes présente dans du matériel didactique en libre accès permet-elle de dépasser les préférences/biais observés dans les données des apprenant·e·s ou au contraire les renforce-t-elle ?
3. Les biais observés dans le matériel didactique en libre accès correspondent-ils à la distribution que l'on trouve en français tel que représenté par le corpus écrit ScienQuest ou l'usage représenté dans ce corpus est-il différent ?

3. Données des corpus d'apprenants

3.1 Choix des corpus

Les données des apprenant·e·s du français L2 sont des données de production qui proviennent d'un corpus d'apprenant·e·s ayant le suédois comme L1, le corpus CEFLE, et d'un corpus d'apprenant·e·s ayant l'allemand comme langue de scolarisation et le suisse allemand comme langue du quotidien, le corpus HarMoS (voir prochaines sections).

Les langues en présence sont donc toutes des langues germaniques, mais la manière de produire le pronom relatif le plus simple dans celles-ci diffère. En suisse allemand, on peut utiliser le pronom relatif *wo* pour les fonctions correspondant au *qui* et *que* du français ("d'Frau *wo* wartet" 'la femme qui attend', "d'Frau *wo* ich gseh" 'la femme que je vois'). En allemand, l'antécédent détermine le genre du pronom relatif, mais la construction verbale de la subordonnée détermine le cas de celui-ci ("der Mann, *der* wartet" 'l'homme qui [der = nominatif] attend', "der Man, *den* ich sehe" 'l'homme que je vois' [den = accusatif/objet direct]). L'attention à la fonction du relatif dans la subordonnée est donc un phénomène syntaxique auquel les apprenant·e·s ayant une bonne maîtrise de l'allemand sont déjà sensibilisés. En suédois, le pronom relatif *som* est le plus fréquent et commun. Il est utilisé quel que soit l'antécédent et quelle que soit la fonction du relatif dans la subordonnée. Toutefois, lorsqu'il prend la fonction de sujet, il est obligatoirement exprimé, alors qu'il peut être omis dans sa fonction d'objet ("kvinnan *som* väntar" 'la femme qui attend', "kvinnan (*som*) jag ser" 'la femme que [optionnel] je vois'). Le suédois connaît un autre pronom relatif correspondant aux utilisations les plus simples de *qui* et de *que*, mais qui est stylistiquement plus formel (*vilken*, *vilket*, *vilka*). Ce pronom s'accorde en genre et en nombre avec l'antécédent; comme *qui* et *que*, ce pronom peut aussi

être utilisé pour introduire une phrase interrogative. Finalement, en anglais, la première langue étrangère des suédophones, le pronom relatif est choisi selon l'antécédent, l'humain étant repris avec un *who*, alors que *which/that* reprend les non animés et les phrases.

Ces corpus ont été choisis afin de permettre une certaine variation dans les données en termes de niveau de français, de modalité de production et de tâche. Ils seront présentés plus en détail dans les deux sections suivantes.

3.2. Le corpus CEFLE

Le corpus CEFLE est un corpus en libre accès en ligne⁶ (Ågren 2008) comprenant des textes écrits par des apprenant·e·s suédophones de français langue étrangère qui suivent des cours de français au secondaire II (16 à 19 ans) ainsi que par un groupe contrôle de natifs de Paris du même âge. Les données utilisées pour la présente étude sont celles du sous-corpus transversal pour lequel chaque apprenant·e a produit un texte à partir de la série de 16 images *Le voyage en Italie*. Il s'agit d'une tâche de narration parlant de deux femmes qui font un voyage en Italie et qui y rencontrent deux hommes italiens. Les apprenant·e·s n'avaient aucune aide à leur disposition et ont pris environ 30 à 40 minutes pour produire leur texte. Les textes ont été répartis selon les stades de développement du français proposés par Bartning & Schlyter (2004): 16 textes au stade 1 (débutant) et 30 textes pour chacun des stades 2 à 4, donc 106 textes au total. Le groupe contrôle comprend également 30 textes.

3.3 Le corpus HarmoS

Le corpus HarmoS a été constitué en Suisse en 2007, dans le cadre d'un projet de recherche dont l'objectif était d'évaluer le niveau de compétence en langue étrangère au terme du primaire (12 ans) et du secondaire I (15 ans), à l'écrit comme à l'oral (Rey 2020; Schneider et al. 2009). Les apprenant·e·s ont donc un niveau débutant à intermédiaire. Le corpus comprend des productions orales en interaction et en monologue de 30 élèves, basées sur différentes tâches ciblant des compétences actionnelles décrites dans le CECR. Il comprend en outre les textes de 90 élèves de 6^e et de 60 élèves de 9^e qui ont écrit un texte à partir d'une série de quatre images, *Lisa et son chien Pacco*, traitant d'une jeune fille, de son père et d'un chien, ainsi que les textes de 175 élèves de 6^e et de 112 élèves de 9^e, dans laquelle les élèves devaient donner leur avis sur les devoirs (avantages et inconvénients, devoirs idéaux).

3.4 Compilation des données des apprenant·e·s

Les données des apprenant·e·s (ainsi que du groupe contrôle) ont été compilées manuellement en cherchant les occurrences de *qui*, *que* et *qu'* dans leur fonction de pronom relatif dans les textes et les transcriptions. Elles ont ensuite été reportées dans un tableau et codées au niveau de l'antécédent –

⁶ <https://projekt.ht.lu.se/cefle/> (page consultée le 10 mars 2021).

humain ou non animé – de la fonction du pronom relatif – sujet ou objet – et finalement de la forme – correcte ou incorrecte. Les pronoms relatifs produits à la place d'un pronom ou d'un subordonnant différent (*où*, *ce que*, *ce qui*, *quand*, *auxquelles* ainsi que *dont*) n'entrent pas dans les calculs, ni les utilisations ambiguës ("elles boivent du café et parlent beaucoup de ce voyage *qui* se rendre" texte original non retouché). La production de l'élision avec *qui* ("qu'est jaune") a été comptée comme incorrecte même si cette forme est attestée en français parlé. Dans le cas de répétitions du pronom seul (hésitation à l'oral, données HarMoS), seule la dernière production a été codée. Par contre, si le pronom a été produit avec un verbe, il a été compté ("qu'on peut, qu'on peut acheter", 2 occurrences). Les utilisations de *que* dans un contexte pour *ce que*, mais pour lesquelles le *ce* a été omis, n'ont pas été prises en compte. Les occurrences sans verbe qui suit ont également été exclues, puisque c'est le verbe qui permet de vérifier la fonction prévue.

La production a été évaluée comme correcte lorsque le pronom attendu a été produit selon la fonction qu'il prend dans la phrase, donc *qui* sujet et *que* objet. Les occurrences incorrectes correspondent donc à une erreur au niveau du choix du pronom selon la fonction que celui devrait avoir. Par exemple, dans l'énoncé "Le garde de nationale, qui picko conetre" (CEFLE, Billy), le *qui* est utilisé pour un antécédent humain (le garde national), mais la fonction est objet (que Picko connaît).

3.5 Résultats des apprenant·e·s

3.5.1 Corpus écrit CEFLE – apprenant·e·s

Au total, les apprenant·e·s du corpus CEFLE ont produit 151 pronoms, dont 130 *qui*, produits correctement dans 95% des cas, et 21 *que*, produits correctement à 62%. On voit donc une dominance de l'utilisation du pronom *qui* et on voit aussi que le pronom *que* est non seulement rare mais aussi plus difficile à produire correctement.

Le Tab. 1 présente la distribution des pronoms selon leur antécédent. Même si la différence au niveau des occurrences est importante (voir Fig. 1), on observe la distribution suivante des pronoms: *qui* est préféré avec les antécédents humains et *que* avec les non animés. On observe également que l'utilisation de *qui* avec les antécédents humains est très correcte dès le début, alors que l'utilisation de *qui* avec les non animés est un peu moins correcte. Ces résultats suggèrent que *qui* pourrait servir de pronom relatif par défaut et confirment ainsi les études antérieures.

On ne peut pas vraiment se prononcer sur l'utilisation du pronom *que* avec les humains vu leur rareté, mais on peut voir que les apprenant·e·s utilisent *que* avec les non animés même lorsqu'ils ont la fonction de sujet et que cette difficulté perdure jusqu'à un niveau avancé de français (stade 4). L'utilisation du pronom avec les antécédents non animés semble généralement plus difficile.

Pronom produit Total	stade 1 (16 textes)	stade 2 (30 textes)	stade 3 (30 textes)	stade 4 (30 textes)
<i>qui</i> avec humain 98% (101/103)		94% (15/16)	97% (37/38)	100% (49/49)
<i>qui</i> avec non animés 79% (23/29)		70% (7/10)	83% (5/6)	85% (11/13)
<i>que</i> avec humain 50% (2/4)	0% (0/2)	100% (1/1)		100% (1/1)
<i>que</i> avec non animés 61% (11/18)		0% (0/1)	33% (1/3)	71% (10/14)

Tab. 1: Proportion de production correcte (et occurrences) des apprenant·e·s du corpus écrit CEFLE

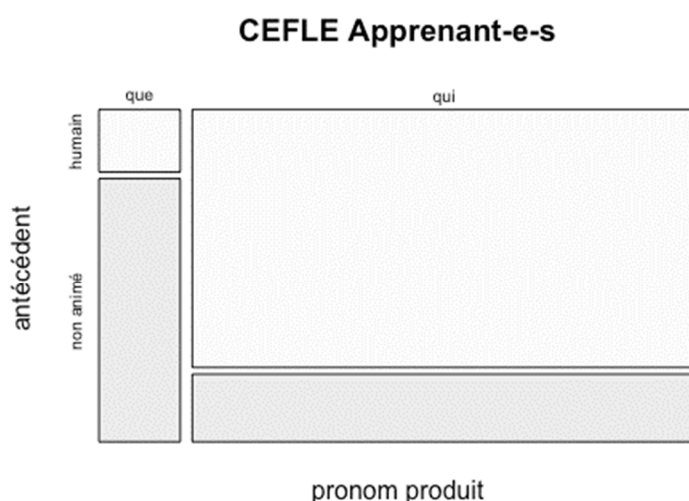


Fig. 1: Distribution des pronoms selon l'antécédent et le nombre d'occurrences chez les apprenant·e·s du Corpus écrit CEFLE

3.5.2 Corpus CEFLE – Groupe contrôle

La Fig. 2 présente la distribution des pronoms produits par le groupe contrôle (personnes natives du français). On voit que comme chez les apprenant·e·s le nombre d'utilisations du pronom *qui* (43 occurrences) est ici aussi nettement plus important que celles de *que* (6 occurrences), alors que le pronom *qui* est utilisé autant avec des antécédents humains (21 occurrences) que non animés (22 occurrences). La distribution entre les antécédents humains et non animés est donc plus équilibrée dans les données du groupe de contrôle que des

apprenant·e·s. On voit par contre que la proportion de *qui* est similaire à ce que l'on a vu chez les apprenant·e·s, ce qui suggère une influence de la tâche.

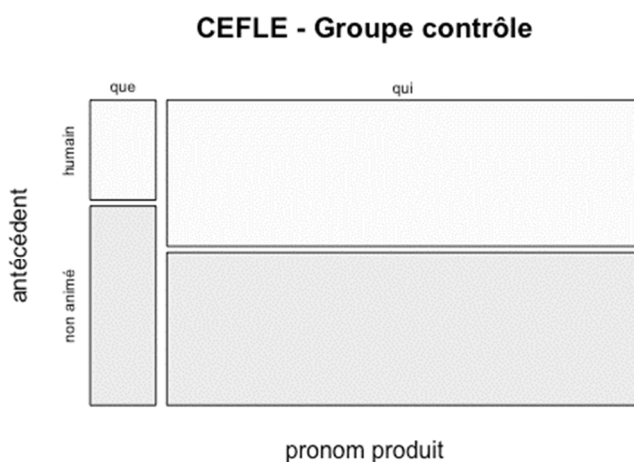


Fig. 2: Distribution des pronoms selon l'antécédent et le nombre d'occurrences chez les natifs du Corpus écrit CEFLE

3.5.3 Résultats des apprenant·e·s du corpus HarmoS

Les données du corpus HarmoS proviennent de trois tâches différentes (cf. Tab. 2). Au total, 177 pronoms relatifs ont été produits, dont 123 en production orale. Peu de pronoms relatifs ont été produits à l'écrit. Comme dans le corpus CEFLE, le pronom *qui* (130 occurrences) est plus fréquent que *que* (47 occurrences) et le pronom *qui* est aussi plus correctement utilisé (94%) que *que* (85%), même si la différence est plus faible que dans CEFLE. Les apprenant·e·s du corpus HarmoS ont moins de difficultés avec le pronom *que* que ceux du corpus CEFLE, ce qui pourrait s'expliquer par la L1 ou la langue de scolarité.

La Fig. 3, qui comprend les données des trois tâches, révèle une distribution des pronoms selon les antécédents, même si la différence n'est pas aussi marquée avec le pronom *qui* qu'avec le pronom *que*.

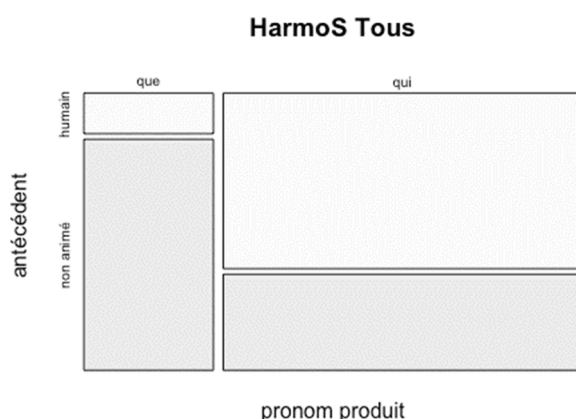


Fig. 3: Distribution des pronoms selon l'antécédent et le nombre d'occurrences dans les trois tâches du corpus HarmoS

Le Tab. 2 révèle qu'il existe des différences entre les tâches. À l'oral, les pronoms sont plutôt bien maîtrisés, sauf le *que* avec les humains, et il y a une distribution assez claire des pronoms selon les antécédents. La production des pronoms est un peu plus difficile à l'écrit. La Fig. 4 illustre les différentes proportions d'utilisation des pronoms: la tâche narrative, comprenant principalement des êtres animés, semble favoriser l'utilisation du pronom *qui*, alors que la tâche plus argumentative donne lieu à une production relativement équitable des deux pronoms. La tâche plus argumentative semble donc favoriser la production du pronom relatif *que* mais aussi l'utilisation des deux pronoms avec un antécédent non animé. On voit aussi que la production de *qui* est un peu plus difficile avec les antécédents humains dans cette tâche (mais il n'y a presque pas d'occurrences).

Même si les conclusions doivent rester modestes en raison du petit nombre de données, on peut malgré tout constater une distribution de l'utilisation des deux pronoms relatifs selon l'antécédent: l'utilisation de *que* avec un antécédent humain est rare et peu réussie et l'utilisation de *qui* avec un antécédent non animé est également plus rare et parfois moins bien réussie qu'avec un antécédent humain. On observe aussi des différences selon la tâche, ce qui pourrait confirmer la suggestion de Roubaud (2015).

Pronom produit Total	Oral (30 transcriptions)	Écrit Narratif (150 textes)	Écrit Opinion (287 textes)
<i>qui</i> avec humain 98% (82/84)	100% (72/72)	88% (7/8)	75% (3/4)
<i>qui</i> avec non animés 87% (40/46)	96% (26/27)	100% (3/3)	69% (11/16)
<i>que</i> avec humain 14% (1/7)	14% (1/7)	0	0
<i>que</i> avec non animés 98% (39/40)	100% (17/17)	75% (3/4)	100% (19/19)

Tab. 2: Proportion de production correcte (et occurrences) des apprenant·e·s du corpus HarmoS

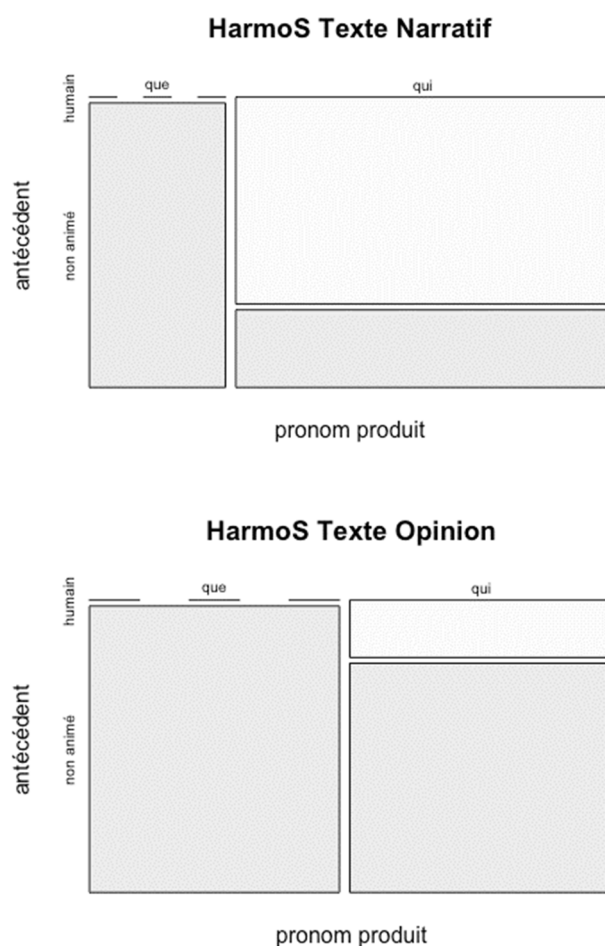


Fig. 4: Distribution des pronoms selon l'antécédent et le nombre d'occurrences dans les données HarmoS dans les tâches écrites

3.5.4 Synthèse des données d'apprenant·e·s

L'analyse des données des apprenant·e·s a globalement montré que le pronom *qui* pose peu de problèmes, en particulier lorsqu'il est utilisé avec un antécédent humain (ou un animal). Le pronom *que* est utilisé beaucoup plus rarement et de manière moins correcte que *qui*. Les données ont globalement confirmé l'hypothèse d'une préférence de *qui* avec les antécédents humains et de *que* avec les antécédents non animés. Mais on a également pu observer une différence selon la L1 des apprenant·e·s, la production étant généralement plus correcte chez les apprenant·e·s des écoles suisses que chez les suédophones, et surtout selon la tâche, une tâche argumentative favorisant l'utilisation du pronom relatif *que*. Alors que le pronom *qui* semble servir de forme par défaut chez les apprenant·e·s des stades 2 et 3 du corpus CEFLE, cela ne semble pas être le cas chez les apprenant·e·s du corpus HarMoS.

4. Données d'input: matériel didactique et corpus

Voyons maintenant quelle est la distribution de ces formes dans des données de français accessibles gratuitement en ligne. Pour cela, j'ai choisi les explications théoriques et les exercices disponibles de la base de données du site du *Point du FLE*⁷ et les données du corpus de français écrit ScienQuest⁸ (Falaise et al. 2011). Ces données peuvent être considérées comme des données d'input, dans la mesure où elles représentent différentes formes de français auxquelles les apprenant·e·s sont exposé·e·s.

4.1 Données du Point du FLE

4.1.1 Compilation des données

Les données du site du *Point du FLE* ont été compilées manuellement en janvier 2019. Tous les exercices et présentations théoriques ciblant notamment les pronoms relatifs *qui* et *que* aux niveaux débutants ont été récupérés, soit sous forme de pdf soit sous forme de capture d'écran. Au total, 26 documents de différentes longueurs ont ainsi été retenus. Les occurrences de *qui* et de *que* ont été compilées en calculant chaque occurrence dans les explications et introductions théoriques et chaque occurrence présentée ou attendue dans les exercices.

4.1.2 Résultats

Tout d'abord il est intéressant de constater que la question de l'antécédent est souvent thématifiée dans les explications et introductions théoriques. Dans l'exemple reproduit dans la Fig. 5, on voit qu'on s'est appliqué à montrer un exemple avec un antécédent humain et un avec un non animé pour chaque pronom. Par contre, le texte précédant les exemples est ambigu "QUI: reprend

⁷ <https://www.lepointdufle.net> (page consultée le 10 mars 2021).

⁸ <https://corpora.aiakide.net/scientext20/> (page consultée le 10 mars 2021).

le sujet du verbe qui suit", en effet *qui* "reprend" l'antécédent au niveau référentiel, mais prend la fonction de sujet à partir du verbe qui suit.

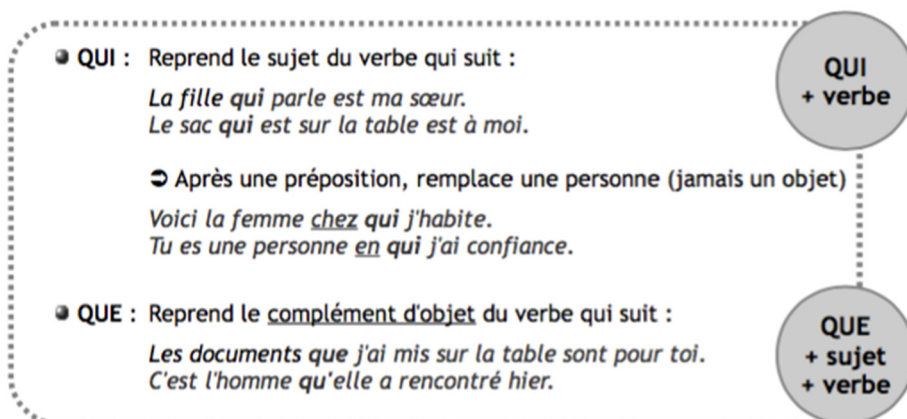


Fig. 5: Exemple de présentation théorique⁹

Dans l'exemple reproduit dans la Fig. 6, le lien vers l'antécédent est présenté de manière explicite, tout comme le fait que *qui* et *que* sont tous deux utilisés avec des animés et des non animés. Le seul problème est que ce n'est pas la fonction de l'antécédent qui compte, mais celle du pronom, et que c'est l'antécédent qui est animé ou non. En examinant les différentes propositions théoriques, on constate qu'il n'y en a en fait que très peu qui présentent le phénomène de manière non ambiguë.

antécédent	forme employée	exemple
sujet	animé et non animé : QUI	Le garçon <u>qui</u> joue là-bas est mon neveu. La chaise <u>qui</u> est là est cassée.
C.O.D.	animé et non animé : QUE	L'homme <u>que</u> j'aime s'appelle Jo. La pomme <u>que</u> je mange est délicieuse.

Fig. 6: Exemple de présentation théorique¹⁰

Au total, 392 pronoms ont été relevés dans les données du *Point du FLE*. Comme on le voit dans le Tab. 3, le nombre d'occurrences des deux pronoms est assez égal, mais le pronom *que* est légèrement plus fréquent (206 occurrences contre 186 pour *qui* au total). On observe aussi que les antécédents sont le plus souvent non animés (71% des données). Même si elle n'est pas très marquée (Fig. 7), on constate tout de même une plus haute proportion d'antécédents humains avec *qui* et surtout une très faible utilisation de *que* avec les antécédents humains, 21% (11/53 occurrences) dans les

⁹ https://flenantes.files.wordpress.com/2013/10/pronoms_relatifs_simples1.pdf (page consultée le 10 mars 2021).

¹⁰ https://www.lepointdufle.net/ressources_fle/a/pdf/pronoms_relatifs.pdf (page consultée le 10 mars 2021).

présentations théoriques et 16% (25/153 occurrences) dans les exercices. On ne retrouve donc pas la dominance de *qui* en termes d'occurrence et la distribution des pronoms selon les antécédents est moins marquée que chez les apprenant·e·s. Néanmoins, la très faible utilisation de *que* avec les humains, en particulier dans les exercices, ne vient pas contredire le système de simplification des apprenant·e·s.

	Théorie		Exercices		<i>Totaux</i>
	qui	que	qui	que	
humains	17	11	60	25	113
non animés	31	42	78	128	279
<i>Totaux</i>	48	53	138	153	392

Tab. 3: Nombre d'occurrences dans les données du *Point du FLE*

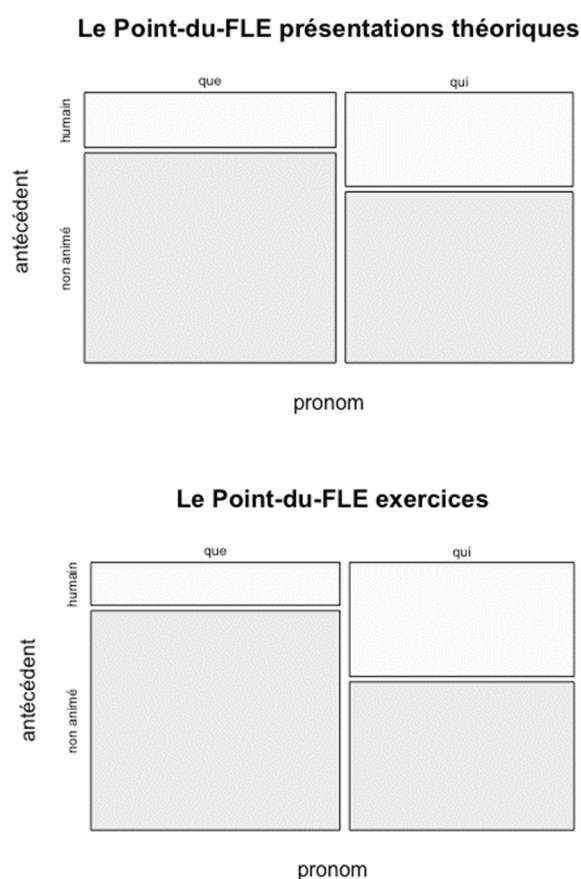


Fig. 7: Distribution des pronoms selon l'antécédent et le nombre d'occurrences dans les données du *Point du FLE* selon le type d'input

4.2 Données du corpus écrit de ScienQuest

J'ai choisi le corpus écrit *ScienQuest* (Falaise et al. 2011) comme élément de comparaison avec les exercices, car ce corpus est un corpus écrit facile d'accès et d'utilisation. Il pourrait donc être utilisé sans trop de difficultés pour l'enseignement du FLE (voir par exemple Yan et al. 2018). En outre, un travail didactique sur les pronoms relatifs se fait plus aisément sur des données écrites qu'orales.

4.2.1 Compilation des données

Le matériel pour cette analyse provient du sous-corpus *Démo*, qui comprend les données de cinq genres textuels (encyclopédie, littérature, presse, scientifiques, tourisme; voir la description du corpus sur leur site) pour un total de 903 textes et 6 millions de mots. La compilation des données a été faite au moyen de l'outil de recherche disponible sur le site. Ainsi 28'303 occurrences du pronom *qui* et 11'178 occurrences de *que* ont été trouvées. Les données pour la présente étude sont basées sur les 20 premières pages de résultats, correspondant à 350 lignes de résultats. Le codage a été effectué manuellement sur un *pdf* comprenant 57 pages pour chacun des deux pronoms. La Fig. 8 montre un extrait du codage, qui révèle la variété des utilisations du pronom *que*. Les antécédents des lignes 1 et 7 ont été codés comme *non animés*, ceux des lignes 3, 5 et 8 comme *humains*. L'antécédent "deux ans" a été codé comme *autre utilisation* (exclu des calculs).

Et avec le sourire , ainsi que le petit cadeau	qu'	il remet à chaque lecteur , en l' occurrence hier
Au lard et fromage " C' est la première fois	que	j' entends parler d' une opération aussi importan
il l' a été deux fois , ce saint éponyme	que	se sont choisi les boulangers pour patron . Car ho
faut mettre en valeur ce qui est bon , ce	qu'	on fait bien . Et c' est le cas du
us les ans d' environ 3.400 nouveaux marins ,	qu'	elle forme à une trentaine de métiers . Documen
Depuis deux ans	qu'	ils jouent ensemble , ils ont sillonné une bonne p
C' est ce travail	qu'	elle exercera à Cravanche tout en élevant ses tro
épousé , en deuxièmes noces , Julien Mettétal	qu'	elle avait eu le grand chagrin de perdre en 1974

Fig. 8: Extrait des données de ScienQuest

4.2.2 Résultats

Les résultats sont ainsi basés sur 447 utilisations simples des pronoms dans le corpus *ScienQuest* (Tab. 4). Le pronom *qui* est plus fréquent (66%) que *que* et les antécédents non animés sont plus fréquents (63%) que les humains. La distribution des antécédents est par contre plus claire que dans les données du *Point du FLE*, la proportion de *que* avec des antécédents humains ne s'élevant qu'à 17% (26/153 occurrences) et la proportion de *qui* avec les humains à 47%

(138/294 occurrences) (voir Fig. 9). Ces données confirment la fréquence du pronom *qui* ainsi que la distribution des pronoms selon les antécédents.

	qui	que	Totaux
humains	138	26	164
non animés	156	127	283
Totaux	294	153	447

Tab. 4: Occurrences compilées dans les données de ScienQuest

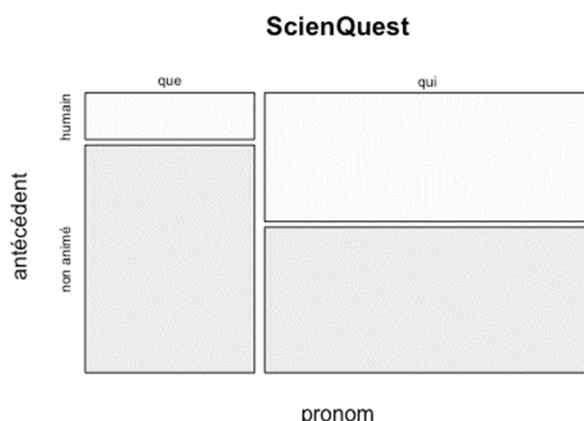


Fig. 9: Distribution des pronoms selon l'antécédent et le nombre d'occurrences dans les données du corpus écrit ScienQuest

4.3 Résumé des données d'input

L'analyse des données d'input a montré une distribution des pronoms selon les antécédents, même si l'association entre *qui* et *humain* était plus faible que chez les apprenant·e·s. En outre, la plus grande fréquence du pronom *qui* est apparue dans les données du corpus, mais pas dans les exercices.

5. Discussion

L'analyse de différentes données d'apprenant·e·s et d'input au niveau de la présence des pronoms *qui* et *que* et de leur distribution selon les antécédents *humain* et *non animé* a montré qu'une telle distribution était visible dans toutes les données, mais à différents degrés. Globalement, ces résultats confirment les approches basées sur l'usage, qui font l'hypothèse d'une relation assez étroite entre ce que l'on trouve dans l'input et la production des apprenant·e·s (Bybee 2008; Ellis et al. 2016).

Pour les apprenant·e·s, les résultats semblent varier selon la L1, le mode de production et surtout la tâche. Les tâches narratives, dans lesquelles des humains/animaux jouent un rôle central, donnent en effet lieu à une forte

production du pronom *qui*, qui semble parfois servir de pronom relatif par défaut. Les apprenant·e·s germanophones semblent aussi avoir plus de facilité à produire les pronoms relatifs correctement. Ces résultats montrent qu'il est nécessaire de conduire d'autres études, avec une variété de L1 et de tâches.

La fréquence de *qui* relevée dans les études antérieures (Bartning & Kirchmeyer 2003; Trévisiol-Okamura 2015) a été constatée dans l'échantillon des données corpus de français écrit ScienQuest, mais pas dans les données du *Point du FLE*. Cette différence pourrait refléter la volonté des personnes créant des exercices d'exposer les apprenant·e·s aux deux formes de pronoms relatifs.

Outre son apport empirique, cette étude permet de rappeler la pertinence de la mise à disposition de corpus de différents types, d'apprenant·e·s comme d'input, dans le but de les mettre en lien en vue d'une amélioration du matériel de didactique de langue étrangère. En même temps, elle révèle aussi la difficulté de manipuler ces données. Les outils de recherche automatique pour détecter certaines erreurs typiques des apprenant·e·s doivent encore être développés. Comme le suggère Granger (2009: 15), développer un enseignement informé par les corpus (*corpus informed teaching*) demande des connaissances dans plusieurs champs adjacents: linguistique de corpus, théorie linguistique, acquisition des langues étrangères/secondes, didactique des langues étrangères, recherche sur corpus d'apprenant·e·s. Une telle palette de connaissances risque de décourager plus d'un·e enseignant·e. Une des tâches de la linguistique appliquée à l'heure digitale, dans un effort conjoint de linguistes et de didacticien·ne·s, serait donc de préparer du matériel didactique sur base empirique, à l'instar de ce qui se fait par exemple sur le site de CLAPI-FLE¹¹.

Pour conclure, concernant l'étude présentée ici, on peut considérer que les données des exercices correspondent finalement bien à l'usage tel qu'il apparaît dans l'échantillon de ScienQuest. En ce sens, l'utilisation des corpus pour l'enseignement des pronoms relatifs *qui/que* semble avant tout redondante au niveau de ce qui est ciblé, sans compter que les données de corpus contiennent bien d'autres utilisations de ces relatifs. Alors que la richesse de telles données peut être fructueuse pour faire découvrir les usages multiples d'une même forme, elle exige un bon niveau de formation linguistique pour en faire le tri pour une application ciblée sur un point de grammaire précis (Thomas 2020).

Une autre piste pourrait être d'utiliser les corpus pour vérifier la présence de biais simplificateurs dans la langue cible et sur cette base construire des exercices qui permettent de les dépasser. Ainsi, plutôt que d'en faire des sources de données du français tel qu'il est vraiment utilisé, les corpus pourraient soutenir les didacticien·ne·s dans leur effort de créer des exercices

¹¹ <http://clapi.icar.cnrs.fr/FLE/> (page consultée le 10 mars 2021).

dans lesquels l'input ne vient pas soutenir les processus de simplification de la langue cible par les apprenant·e·s, comme le suggère VanPatten (2002 par exemple) dans ses travaux sur l'input.

BIBLIOGRAPHIE

- Ågren, M. (2008). *A la recherche de la morphologie silencieuse. Sur le développement du pluriel en français L2 écrit*. Thèse de doctorat, Université de Lund, Suède.
- Andersen, R. W. (1984). The one to one principle of interlanguage construction. *Language Learning*, 34(4), 77-95.
- André, V. (2018). Nouvelles actions didactiques: faire de la sociolinguistique de corpus pour enseigner et apprendre à interagir en français langue étrangère. *Action Didactique*, 1, 71-88.
- Bartning, I. & Kirchmeyer, N. (2003). Le développement de la compétence textuelle à travers les stades acquisitionnels en français L2. *Acquisition et interaction en langue étrangère*, 19, 9-39.
- Bartning, I. & Schlyter, S. (2004). Itinéraires acquisitionnels et stades de développement en français L2. *Journal of French Language Studies*, 14(3), 281-299.
- Boulton, A. (2017). Corpora in language teaching and learning. *Language Teaching*, 50(4), 483-506.
- Boulton, A. & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348-393.
- Boulton, A. & Tyne, H. (2014). *Des documents authentiques aux corpus: démarches pour l'apprentissage des langues*. Paris: Didier.
- Bybee, J. (2008). Usage-based grammar and second language acquisition. In P. Robinson & N. C. Ellis (éds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 216-236). New-York: Routledge.
- Cappeau, P. (2015). La relative: observation critique des deux présentations proposées par les manuels de FLE et les grammaires d'apprenants. In P. Trévisiol-Okamura & M. Kaheraoui (éds.), *Les subordonnées, corpus, acquisition et didactique* (pp. 21-34). Rennes: Presses universitaires de Rennes.
- Collins, L., Trofimovich, P., White, J., Cardoso, W. & Horst, M. (2009). Some input on the easy/difficult grammar question: An empirical study. *Modern Language Journal*, 93(3), 336-353.
- Cuq, J.-P. & Gruca, I. (2005). *Cours de didactique du français langue étrangère et seconde*. Grenoble: Presses universitaires de Grenoble.
- Delatour, Y., Jennepin, D., Léon-Dufour, M. & Teyssier, B. (2004). *Nouvelle grammaire du Français. Cours de civilisation française de la Sorbonne*. Paris: Hachette.
- Ellis, N. C., Römer, U. & O'Donnell, M. B. (2016). *Usage-based approaches to language acquisition and processing: Cognitive and corpus investigations of construction grammar*. Chichester: Wiley.
- Falaise, A., Tutin, A. & Kraif, O. (2011). Définition et conception d'une interface pour l'exploitation de corpus arborés pour non-informaticiens: la plateforme ScienQuest du projet Scientext. *Traitement Automatique des Langues*, 53(3), 103-128.
- Gajo, L., Luscher, J.-M., Racine, I. & Zay, F. (2019). *Variation, plurilinguisme et évaluation en français langue étrangère*. Bern: Peter Lang.
- Gilmore, A. (2004). A comparison of textbook and authentic interactions. *ELT Journal*, 58(4), 363-374.
- Granger, S. (2009). The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation. In K. Aijmer (éd.), *Corpora and language teaching* (Vol. 33, pp. 13-32). Amsterdam/Philadelphia: John Benjamins.

- Jacques, M.-P. (2005). Pourquoi une linguistique de corpus? In G. Williams (éd.), *La linguistique de corpus* (pp. 21-30). Rennes: Presses Universitaires de Rennes.
- Klein, W. & Perdue, C. (1997). The Basic Variety (or: Couldn't natural languages be much simpler?). *Second Language Research*, 13(4), 301-347.
- Larrivée, P. & Lefeuve, F. (éds.). (2020). *La didactisation du français vernaculaire*. Caen: Presses Universitaires de Caen.
- Lyster, R. & Rebuffot, J. (2002). Acquisition des pronoms d'allocution en classe de français immersif. *Acquisition et interaction en langue étrangère*, 17, 51-72.
- Mayaffre, D. (2005). Rôle et place du corpus en linguistique. Réflexions introductives. *Textol!*, X(4), 5-17.
- Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.
- Ravazzolo, E. & Etienne, C. (2019). Nouvelles ressources pour le FLE à partir des études en interaction. *Linx. Revue des linguistes de l'université Paris X Nanterre*, 79. <https://doi.org/10.4000/linx.3454>
- Rey, K. (2020). *Acquisition des propositions subordonnées en français L2 par des apprenants germanophones. Résultats d'une étude transversale dans des productions écrites narratives et argumentatives* [Thèse de Master]. Université de Fribourg, Suisse.
- Riegel, M., Pellat, J. C. & Rioul, R. (1994). *Grammaire méthodique du français* (4^e éd.). Paris: Presses Universitaires de France.
- Roubaud, M.-N. (2015). Accéder aux compétences syntaxiques des élèves (8-11 ans): le cas des relatives. In P. Trévisiol-Okamura & M. Kaheraoui (éds.), *Les subordonnées: corpus, acquisition et didactique* (pp. 123-138). Rennes: Presses universitaires de Rennes.
- Schneider, G., Lenz, P. & Studer, T. (2009). *Langues étrangères. Rapport scientifique de synthèse et modèle de compétences. Version provisoire (avant adoption des standards de base)*. Bern: CDIP.
- Sinclair, J. (1996). Preliminary recommendations on corpus typology. *EAGLES Document*. <http://www.ilc.cnr.it/EAGLES/pub/eagles/corpora/corpusyp.ps.gz>
- Thomas, A. (2020). Didactisation du français ordinaire: les défis de la variation dans l'enseignement des temps du passé en français langue étrangère. In P. Larrivée & F. Lefeuve (éds.), *La didactisation du français vernaculaire* (pp. 153-186). Caen: Presses universitaires de Caen.
- Trévisiol-Okamura, P. (2015). L'acquisition et l'enseignement des relatives en FLE: regards croisés. In P. Trévisiol-Okamura & M. Kaheraoui (éds.), *Les subordonnées: corpus, acquisition et didactique* (pp. 103-120). Rennes: Presses universitaires de Rennes.
- Tyler, A. E., Ortega, L., Uno, M. & Park, H. I. (2018). *Usage-inspired L2 instruction: Researched pedagogy*. Amsterdam/Philadelphia: John Benjamins.
- Tyne, H., André, V., Benzitoun, C. & Greub, Y. (éds.). (2014). *French through corpora: Ecological and data-driven perspectives in French language studies*. Newcastle: Cambridge Scholars.
- VanPatten, B. (2002). Processing instruction: An update. *Language Learning*, 52(4), 755-803.
- Vialleton, É. & Lewis, T. (2014). Reconsidering the authenticity of speech in French language teaching: Theory, data, methodology, and practice. In H. Tyne, V. André, C. Benzitoun, A. Boulton, & Y. Greub (éds.), *French through corpora: Ecological and data-driven perspectives in French language studies* (pp. 293-316). Newcastle: Cambridge Scholars.
- Welcomme, A. (2013). Les subordonnées relatives à l'oral en français langue étrangère et en français langue maternelle. *Mémoires de la Société Néophilologique*, LXXXVI, 97-136.
- Williams, G. (2005). *La linguistique de corpus*. Rennes: Presses Universitaires de Rennes.
- Yan, R., Tutin, A. & Tran, T. T. H. (2018). Routines verbales pour les français langue étrangère: des corpus d'experts aux corpus d'apprenants. *Lidil. Revue de linguistique et de didactique des langues*, 58. <https://doi.org/10.4000/lidil.5411>