



ISSN 1901-3809

ISSN en ligne 2261- 2807

Complexité et relatives en français vernaculaire langue maternelle et langue seconde

Pierre Larrivée

Normandie Université, Université de Caen, CRISCO EA4255, France
 pierre.larrivee@unicaen.fr

Résumé

Ce travail a pour but de cerner le rôle de la complexité pour les subordonnées relatives en français langue maternelle et langue seconde. La complexité déterminerait en effet les manifestations des relatives suivant la hiérarchie d'accessibilité (Keenan, Comrie 1977). Si c'est le cas, ces relatives devraient avoir des comportements comparables en L1 et L2, et quelle que soit la langue maternelle dans ce dernier cas. Ces comportements sont ici examinés sous le rapport du nombre respectif de pronoms relatifs produits, de la proportion de leurs collocations et par les taux de disfluences. On s'attend en effet qu'un relatif plus accessible soit plus fréquent, entrent moins dans des collocations et soit moins l'objet de disfluences qu'un relatif moins accessible. Ces attentes sont confirmées par l'analyse des productions de natifs et d'apprenants norvégiens du français dans un corpus de conversations libres, sauf pour les collocations, dont le nombre est inversement proportionnel à l'accessibilité des marqueurs. Cela amène à discuter la cause de cette hiérarchie d'accessibilité dans son rapport avec la notion de fréquence.

Mots-clés : complexité, fréquence, relatif, hiérarchie d'accessibilité, collocation, disfluence

Complexity and relative clauses in vernacular French as a first and second language

Abstract

The objective of this article is to determine the role of complexity for relative clauses in vernacular French as a first and second language. Complexity is expected to govern the behavior of relative clauses through the accessibility hierarchy (Keenan and Comrie 1977). The hierarchy predict comparable behaviors of relatives in L1 and L2, whatever the pairs of second languages concerned. These behaviors are studied through the respective number of relative pronouns, their collocations and disfluency rates. A more accessible relative pronoun should be more frequent, enter in fewer collocation and give rise to fewer disfluencies than a less accessible relative pronoun. These expectations are confirmed by the analysis of both L1 and L2 speaker by the analysis of spontaneous conversations between French and Norwegian speakers of French, except for collocations, the number of which is inversely proportional to the pronoun's accessibility. This leads to reappraise the cause of the hierarchy in relation to frequency.

Keywords: complexity, frequency, relative clauses, accessibility hierarchy, collocations, disfluencies

1. Introduction¹

L'acquisition de la grammaire d'une langue seconde est souvent marquée par le transfert, soit l'influence d'une langue sur les productions dans une autre (voir Yu et Odlin, 2017 pour des discussions récentes). Cette influence devrait être moins importante quand l'acquisition dépend de la complexité. Ainsi des subordinées relatives dont l'acquisition dépendrait d'une hiérarchie d'accessibilité établie selon la complexité de la fonction syntaxique occupée par le pronom tête (Keenan et Comrie 1977). Cette hiérarchie propose que le sujet est plus accessible que l'objet direct, l'objet direct plus que l'objet indirect, l'objet indirect plus que l'oblique, et ainsi de suite, comme le montre ce schéma.

SU > DO > IO > OBL > GEN > OCOMP
 (qui > que > où / dont / (prép.) + lequel > où / dont / (prép.) + lequel > dont > 0)

La compétence des apprenants concernant la subordinée relative ne devrait donc guère être influencée par la langue d'origine, et leur performance devrait être comparable à celle des natifs. Ce sont ces hypothèses que teste le présent travail.

En s'appuyant sur les travaux actuels, ce travail vise à comparer l'acquisition des relatives françaises par les locuteurs langue seconde avec celle des locuteurs natifs. Il part de la hiérarchie d'accessibilité de Keenan et Comrie pour vérifier trois de ses prédictions. La première concerne la fréquence respective des pronoms relatifs. La deuxième, que les relatifs moins accessibles seront plus liés à des collocations, les collocations réduisant la charge de traitement plus importante qu'imposent les relatifs moins accessibles. La dernière est la corrélation entre l'accessibilité de chaque marqueur et la proportion de ses disfluences (les « erreurs »). La vérification des prédictions est faite à partir du corpus *Apprenants norvégiens du français*, qui réunit des locuteurs natifs et de français langue seconde dans des conversations libres. Il permet de comparer les productions des uns et des autres dans le même environnement. Après une présentation d'études actuelles permettant de repérer certains paramètres pertinents pour l'acquisition des relatives, le travail présente les résultats de l'analyse du corpus. Les trois dimensions de la fréquence respective des relatifs, des proportions d'emplois en collocation et des taux d'erreurs sont successivement comparées pour chaque marqueur et pour les populations des apprenants et des natifs. Le travail contribue donc à articuler le rôle respectif de la fréquence et de la complexité dans l'acquisition, en se fondant sur des données du français L2.

2. Arrière-plan théorique

Le but de ce travail est d'interroger le rôle d'une hiérarchie d'accessibilité dans l'acquisition des subordonnées relatives. Une telle hiérarchie laisse attendre une différenciation de comportement selon les marqueurs. Par contre, aucune différenciation majeure ne devrait se manifester selon que le locuteur est de langue seconde ou de langue maternelle. Autrement dit, les questions que je me pose ici sont de deux ordres :

- La hiérarchie d'accessibilité prédit-elle le nombre respectif d'occurrences des relatifs, leurs taux de collocations et de disfluences ?
- Les facteurs découlant de la complexité se retrouvent-il de façon comparable chez les natifs et chez les langue seconde ?

Ces questions trouvent leur bien-fondé dans les éléments fournis par certaines des études récentes. Quelques-unes d'entre elles concernant la langue maternelle adulte, la langue maternelle des enfants et la langue seconde d'apprenants avancés sont rappelés ci-dessous.

Les études de la pratique vernaculaire, spontanée et non surveillée, du français sont largement fondées sur des démarches qualitatives. Ces approches ne permettent cependant pas de définir le poids respectif des marqueurs, et semblent grossir l'importance réelle de manifestations spectaculaires (comme le très stigmatisé décumul du relatif, de type *La femme que j't'en ai parlé, c'est ma sœur*, qui se révèle tellement rare dans les corpus qu'il n'est pas attesté dans ESL02). Une approche quantifiée des relatives en français vernaculaire langue maternelle est fournie par Larrivée et Skrovec (2016). A partir de la hiérarchie d'accessibilité, ils considèrent le poids respectif et le taux de disfluence des relatifs dans une partie du corpus ESL02. Le nombre d'occurrences brut des marqueurs est fourni par le tableau suivant.

qui	7623
que / qu'	2373
où	841
dont	124
lesquelles, lesquelles, auxquels, auxquelles	28

Tableau 1

Il montre la dominance marquée de *qui*, suivi de *que*, *où*, *dont* et des dérivés de *lequel*. Le poids respectif des relatifs suit clairement la hiérarchie. Si Larrivée et Skrovec ne donnent pas une analyse systématique des collocations, ils se penchent sur les taux de disfluences pour chaque marqueur. Les emplois syntaxiquement atypiques, les hésitations, et les tronctions de propositions relatives qui ne sont pas terminées se retrouvent aux taux suivants.

Relatif	Emploi syntaxiquement atypique	Hésitation	Troncation	Total « erreurs »	Total occurrences étudiées
qui	1	0	17	18 (3,2%)	563
que	3	3	23	29 (11,9%)	244
où	0	3	17	20 (11,4%)	175
dont	1	2	0	3 (17,64%)	19
lequel et associés	0	1	0	1 (14,3%)	7

Tableau 2

Encore une fois, la proportion d’emplois disfluents est plus importante pour les relatifs moins accessibles. Ainsi, sous les deux rapports documentés de façon suivie, il y a une relation avec la hiérarchie d’accessibilité : plus un terme est accessible, plus il est fréquent, et moins il est sujet à l’erreur.

Les tendances observées se retrouvent dans l’apprentissage du français langue maternelle de la part d’enfants du primaire. C’est ce que documente Roubaud (2015) à partir du corpus Charolles de productions écrites de textes narratifs par des enfants de trois niveaux du cycle primaire appartenant à des écoles de deux milieux sociaux. Si elle ne décompte ni les erreurs ni les collocations, elle offre un chiffrage pour le poids des marqueurs (*qui*, *où*, *que*, *dont*), dans les productions de chacun des trois groupes de scolarité (CE2, CM1 et CM2). Les résultats sont résumés par son tableau 5 (2015 : 129). Ces chiffrages établissent la prépondérance de *qui* dans les textes, à tous les niveaux. On note en outre une dominance inattendue de *où* par rapport à *que*, qui se manifeste dès le premier niveau où *que* n’est pas attesté. Enfin, *dont* ne se manifeste qu’au niveau intermédiaire et reste rare. Dans l’ensemble, la hiérarchie est respectée, sauf pour cet échange de place entre *que* et *où*.

Ces tendances se retrouvent-elles dans l’acquisition de la langue seconde ? Les travaux de Nathalie Kirchmeyer sur la relative sont fondés sur du corpus InterFra qui documente des entretiens et productions narratives d’apprenants suédois

universitaires du français ainsi que de locuteurs natifs. Il est donc possible de comparer les productions d'apprenants et de natifs. Méthodologiquement, le travail a l'avantage de compter séparément les emplois particuliers des relatives qui pourraient brouiller les cartes. Les clivés en *il y a* et *c'est que*, qui sont particulièrement fréquentes en français vernaculaire ainsi que les relatives périphrastiques de type *ce qu-* sont comptabilisées indépendamment. Les résultats (Hancock et Kirchmeyer 2005 : 19) montrent que la hiérarchie d'accessibilité se manifeste en langue maternelle comme en langue seconde. Si on retire les clivées et les relatives périphrastiques, on obtient des pourcentages suivants : pour les apprenants 68 % de *qui*, 18 % de *que* et 13,6 % de *où* ; pour les natifs, 63 % de *qui*, 18,7 % de *que* et 18,2 % de *où*. Le rapport entre *que* et *où* est moins différencié qu'on pourrait l'attendre. Les constructions comptabilisées séparément ont un comportement divergent, avec un pourcentage plus important de clivées chez les apprenants (9 % de plus) et moins de relatives périphrastiques (7 % de moins). On notera cependant que la proportion des erreurs n'est pas documentée. Des travaux sur l'acquisition des relatives par des locuteurs de langues germaniques comme Welcomme (2003) ou Olsson (2013) accréditent l'idée d'une convergence dans le parcours entre langue maternelle et langue seconde, qui restent à valider. Cette convergence supposée entre natifs et apprenants avancés est ce à qu'évaluent les sections qui suivent.

3. But et méthode

Le but de ce travail est d'établir le rôle de la complexité pour l'acquisition des subordonnées relatives. Si la complexité telle que décrite par la hiérarchie d'accessibilité est le facteur central pour l'acquisition des relatives, on s'attendrait à ce que l'acquisition soit comparable à travers les langues secondes, quelle que soit la langue maternelle de l'apprenant. En outre, cette acquisition devrait également être comparable à celle qu'on retrouve en langue maternelle. C'est dans ce sens que vont les données des études existantes dans la section précédente. Pour les confirmer, on voudrait pouvoir comparer le comportement des locuteurs natifs et des langues secondes dans un même environnement discursif. L'environnement discursif idéal est la production orale spontanée non-surveillée qui permet, comme le souligne Myles (2015), de mettre au jour la compétence syntaxique immédiate des locuteurs.

Les corpus de pratique vernaculaire mettant en rapport locuteurs natifs et de langue seconde sont rares, en tout cas pour le français. Il est cependant documenté par le corpus *Apprenants norvégiens du français*. Il s'agit d'une ressource constituée pendant l'année à l'été 2015 sous la coordination de Catrine Bang Nilsen. Elle se compose de cinq enregistrements d'environ une heure chacun de conversations

libres entre des apprenants norvégiens du français et un francophone. Les apprenants norvégiens étaient avancés au sens où c'étaient des étudiants universitaires engagés dans des études de français à l'Office Franco-Norvégien d'Étude et de Coopération de l'Université de Caen Normandie (voir Bartning 1997 pour une discussion de la notion d'apprenants avancés). Chaque enregistrement a été transcrit selon les normes du CFPQ, qui reste le seul corpus librement accessible de conversations spontanées pour le français, et les transcriptions sont librement accessibles sur le site du laboratoire CRISCO.

La méthode consistait à identifier tous les pronoms relatifs de chaque transcription. Un pronom relatif est un élément à la périphérie gauche d'une subordonnée, qui a généralement un antécédent, et dont la forme est déterminée par la fonction que l'antécédent aurait dans la subordonnée. Par exemple, dans l'énoncé d'une apprenante *D'autant que c'est le <f<SEULE>> chose que : dont on PARle en norvège*, le relatif est *dont* car l'antécédent *chose* serait le complément indirect en *de* du verbe *parler* de la subordonnée. Chaque relatif a été ventilé selon qu'il était produit par un locuteur L1 ou L2. Cela permettait de répondre à la question sur la convergence du comportement des formes en langue maternelle et seconde. Les autres paramètres pertinents pour évaluer la convergence étaient le nombre respectif d'occurrences des relatifs, les collocations dans lesquelles entraient les relatifs, et le taux de disflue auquel ils donnaient lieu. À partir de ces données, il devenait possible de répondre aux questions explicitées au début de la section 2.

Les réponses sont apportées en trois parties. La première section considère la question de la prépondérance respective des marqueurs. La seconde celle des collocations. La troisième enfin envisage la dimension de la disflue. Chaque partie compare les productions des natifs et des langue seconde. Ainsi, l'idée que la complexité des marqueurs prédit leur comportement peut être pleinement évaluée.

4. Analyse des données

4.1. Prépondérance des marqueurs

Cette section présente les résultats de l'analyse du nombre respectif des pronoms relatifs dans la pratique vernaculaire des apprenants et des natifs du corpus *Apprentissage du français*. La prépondérance des marqueurs est réputée suivre directement la hiérarchie d'accessibilité proposée par Keenan et Comrie (1977). Plus la fonction typiquement représentée par un relatif est accessible sur cette hiérarchie, plus il devrait être fréquent. La chose est vérifiée dans la langue maternelle de locuteurs enfants, adultes et chez les locuteurs langue seconde, comme on l'a montré ci-dessus. Qu'en sera-t-il dans le corpus étudié ? La réponse à cette question est apportée dans le tableau suivant.

Relatif	Apprenants		Natifs	
	%	n	%	n
<i>qui</i>	63,2 %	136	57,9 %	173
<i>que</i>	31,2 %	67	33,1 %	99
<i>quoi</i>	-	0	-	0
<i>où</i>	4,2 %	9	8 %	24
<i>dont</i>	0,5 %	1	1 %	3
<i>lequel et dérivés</i>	0,9 %	2	-	0
Totaux	100 %	215	100 %	299

Tableau 3. Proportion des pronoms relatifs dans le corpus *Apprentissage du français* pour les apprenants norvégiens et les locuteurs natifs du français

Les chiffres et pourcentages rendent possible la comparaison des marqueurs entre eux, et pour chacune des deux populations. De façon globale, la hiérarchie est clairement suivie. Chez les apprenants, *qui* emporte près des deux-tiers des emplois à 63,2 % des 215 relatifs ; il est suivi de *que* qui représente pratiquement un autre tiers à 31,2 % des emplois ; avec pour les emplois restants 4,2 % de *où*, 0,5% de *dont* et 0,9 % de *lequel*. Les pourcentages sont proches de ceux des natifs. Chez ces derniers en effet, on retrouve pratiquement 58 % de *qui*, 33 % de *que*, 8 % de *où* et 1 % de *dont*. On notera que le pourcentage de *qui* est plus faible que dans d'autres corpus où il se monte à 80 % et plus des occurrences. La raison pourrait en être un fait observé par Larrivée et Skrovec (2017) : plus l'échange est marqué par une proximité sociolinguistique, plus le taux relatif de *qui* est bas. Les chiffres cités dans le sous-corpus d'ESLO2 sont de 495 *qui*, 65 *que*, 36 *où*, 21 *dont* et 10 *lequel* dans les Conférences, et de 123 *qui*, 53 *que*, 11 *où*, 1 *dont* et 1 *lequel* dans la vie quotidienne de l'étudiant. Autrement dit, le premier genre, caractérisé par la distance, compte parmi les relatifs 78,9 % de *qui*, alors que le second, marqué par la proximité, en compte 65 %.

4.2. Collocations

La section précédente a montré comment le nombre respectif d'occurrences des relatifs suit la hiérarchie d'accessibilité, comme c'est le cas dans les analyses de corpus disponibles. Concernant les relatifs moins accessibles, on pourrait croire que la difficulté de leur traitement les amène à s'associer à des collocations. L'usage dans ces contextes permettrait présumément de diminuer la difficulté de traitement, puisque les collocations, étant produites globalement, sont d'un traitement plus

facile que les séquences compositionnelles, assemblées item par item. Or, les collocations associées aux relatifs ne sont pas particulièrement parlante dans le corpus de vernaculaire langue maternelle de Larrivée et Skrovec (2016). Certes, *où* s’associe à l’antécédent *le moment* à près de 15 %, et 35 % des emplois de *dont* introduisent le verbe *parler*, mais on ne note pas de collocations particulières pour les relatifs les plus rares que sont *lequel* et *quoi*. Un type de collocation plus fréquent est celui des clivées. En effet, en comptant à la fois les clivées en *c’est* et les présentatifs en *il y a*, on remarque que 41,5 % des emplois de *qui*, 20 % des *que*, 22 % des *où*, 11,8 % de *dont* et 43 % des *lequel*. La tendance est donc contraire aux attentes d’une relation entre emploi en clivées et faible accessibilité du relatif, puisque c’est la relation inverse qui dans l’ensemble se manifeste, si on met de côté le cas de *lequel* attesté par peu de formes. Plus une forme est accessible, plus elle s’emploie en clivée. Qu’en est-il en langue seconde ? Si Kirchmeyer et Welcomme recense les clivées en général, il n’y a pas à ma connaissance d’étude qui donne la proportion de clivées pour chaque relatif. Les clivées représentent globalement une proportion importante des emplois des relatifs. Hancock et Kirchmeyer (2005) et Welcomme (2003) notent que les clivées sont sur-représentées chez les apprenants par rapport aux natifs. A partir de ces observations, il y a donc lieu de se demander si l’emploi des relatifs en clivée varie en fonction de l’accessibilité du relatif, et s’il y a un décalage entre apprenants et natifs. Ce sont deux questions auxquelles le corpus permet de répondre directement. Pour chacune des populations, les occurrences de chaque relatif ont été examinées et les emplois dans les clivées en *c’est* ou en *il y a* ont été comptabilisés. Le pourcentage du nombre d’emplois en clivées par rapport au nombre d’occurrence de chaque relatif est donné dans le tableau suivant.

Relatif	Totaux apprenants			Totaux natifs		
	<i>Il y a</i>	<i>C’est</i>	Totaux des deux contextes	<i>Il y a</i>	<i>C’est</i>	Totaux des deux contextes
qui	12,5 %	18,4 %	30,9 %	21,4 %	21,4 %	42,7 %
que	3 %	16,4 %	19,4 %	7 %	16,2 %	23,2 %
quoi	-	-	-	-	0	-
où	11 %	0	11 %	8,3 %	4,2 %	12,5 %
dont	0	100 %	100 %	0	0	0
lequel et dérivés	0	0	0	-	-	-
Totaux	9,3 %	17,2 %	26,5 %	15,4 %	18 %	33,4 %

Tableau 4. Proportion de chaque pronom relatif utilisé dans les clivées dans le corpus *Apprentissage du français* pour les apprenants norvégiens et les locuteurs natifs du français

Ces résultats nous permettent de considérer la dimension des relatifs et celle des populations apprenants et natifs. Pour ce qui est des relatifs, les tendances globales sont contraires aux attentes selon lesquelles moins le terme est accessible, plus il entrerait dans une collocation. Ici, au contraire, et comme dans les données de Larrivée et Skrovec (2016), ce sont les relatifs les plus accessibles qui entrent le plus en collocation avec une clivée. Il y a des accrocs à cette tendance, avec 100 % de *dont* dans la clivée en *c'est*, pour une seule occurrence. Les emplois avec *il y a* chez les apprenants sont erratiques passant de 12,5 % de *qui*, à 3 % de *que* à 11 % de *où*. Pour la clivée en *c'est* chez les apprenants, la chute de 18,4 % de *qui* à 16,4 % de *que* n'est probablement pas significative. Néanmoins, les totaux de 30,9 % de *qui* à 19,4 % de *que* à 11 % de *où* chez les apprenants sont parlants. Ils sont remarquables chez les natifs, où les totaux vont de 42,7 % de *qui*, à 23,2 % de *que*, à 12,5 % de *où*. La pente est encore plus affirmée que dans l'étude du vernaculaire dans Larrivée et Skrovec (2016). On note comme chez les apprenants un *que* moins fréquent que *où* avec *il y a*, qui ne dépare cependant pas la tendance générale.

4.3 Emplois disfluents

La section précédente a montré que la distribution des relatifs dans les contextes clivés était en rapport inverse avec leur accessibilité. Cela contredit donc l'attente que moins un marqueur est accessible, plus il sera utilisé dans des contextes collocatifs comme les clivées. Cette attente reposait sur l'idée que la lourdeur de traitement d'un marqueur pouvait être compensée par le caractère formulaire de l'environnement d'emploi. C'est l'attente d'un rapport entre accessibilité et taux de disfluence que vérifie cette section. L'idée est que la lourdeur de traitement devrait entraîner une plus grande proportion de disfluences, à la fois chez les natifs et chez les apprenants. On s'attendrait donc que la proportion d'erreur caractérisant *dont* soit plus grande que celle qu'on retrouve avec *qui* par exemple. Afin d'établir si cela était le cas, il a donc fallu identifier les disfluences manifestées dans l'emploi des relatifs dépouillés dans le corpus. Les catégories de disfluences utilisées reprennent celles de Larrivée et Skrovec (2017, 2016), à savoir les hésitations, les productions syntaxiques atypiques et les tronctions. Les hésitations concernent le choix du pronom relatif, répété à l'identique ou en alternance avec d'autres relatifs. L'erreur correspond au choix d'un pronom relatif inattendu étant donné la fonction remplie. La tronction enfin est le fait pour la proposition subordonnée de ne pas se voir complétée. Illustrées par les productions d'apprenants (en a.) et de natifs (en b.) qui suivent, ces catégories épuisaient les occurrences trouvées dans le corpus.

- (1) a. c'est le <f-SEULE> chose **que**: dont on PARle en norvège
 b. y'a bien une question **que** je me- **que** je me posais
- (2) a. j'ai pensé à qu'est-ce **que** j'ai besoin
 b. quand les employER c'**qu'**on a pas vu en cours
- (3) a. Mon pERE il fait parti dans une gROUPE à faceBOOK **qui** s'appelle euh (.)°
 c. c'est quelqu'un **qui** euh : (.) c'est un archéologue des traces écrites (.)
 c'est - c'est quelqu'un **qui** va LIRE des TRES TRES vieilles inscriptions

Les types de disfluences identifiées se distribuent comme suit selon le relatif et le groupe de locuteurs.

Relatif	Totaux apprenants			Totaux natifs		
	Hésitation	Erreur	Troncation	Hésitation	Erreur	Troncation
qui	15 %	1,5 %	14 %	17,9 %	2,3 %	5,7 %
que	10 %	8,9 %	7,5 %	18,2 %	4 %	7 %
quoi	-	-	-	-	-	-
où	22,2 %	-	11,1 %	16,7 %	4,2 %	8,3 %
dont	100 %	-	-	0	0	0
lequel et dérivés	0	100 %	0	-	-	-
Totaux	14,5 %	4,7 %	11,6 %	17,7 %	3 %	6,4 %

Tableau 5. Proportion des emplois disfluents de chaque pronom relatif dans le corpus *Apprentissage du français* pour les apprenants norvégiens et les locuteurs natifs

Considérons les types de disfluence, au regard des items d'une part et des populations d'autre part. Concernant les hésitations, elles présentent chez les apprenants un schéma de progression inégale suivant l'accessibilité du relatif. Si on fait exception du taux d'hésitation pour *que* à 10 %, on passe de 15 % pour *qui* à 22 % pour *où* à 100 % pour *dont*. Chez les natifs, les hésitations sont plus nombreuses que chez les apprenants, à presque 18 % contre 14,5 %. Aucun schéma de progression n'est pourtant manifeste, les pourcentages tournant autour de la moyenne de 17,7 %. Les hésitations ne présentent pas de schémas de variation significatifs. La même conclusion semble s'appliquer aux troncations. Chez les apprenants, 14 % des relatives avec *qui* ne sont pas complétées, 7,5 % de celles avec *que* et 11,1 % avec *où*. La moyenne de 11,6 % est pratiquement le double du taux des natifs. Chez ces derniers, il y a une faible progression des troncations selon le relatif, de 5,7 % avec *qui*, à 7 % avec *que*, à 8,3 % avec *où*.

En ce qui concerne les erreurs, chez les apprenants, elles marquent une progression selon l'accessibilité du relatif. Alors que *qui* est l'objet d'un taux de 1,5 % de productions syntaxiquement inattendues, on en est à 8,9 % avec *que* et à 100 % avec *lequel*. La moyenne de 4,7 % du nombre total d'occurrences produites par les apprenants est plus élevée que celle des natifs à 3 %. Chez ces derniers, il y a une modeste progression des erreurs : 2,3 % des occurrences de *qui* sont des productions syntaxiques atypiques, face à 4 % et 4,2 % respectivement pour *que* et *où*.

Ainsi, les hésitations et les troncations ne présentent pas de tendance affirmée. Du point de vue des populations, les apprenants produisent un peu plus d'occurrences disfluentes à 30,8 % que les natifs à 27,1 %. S'ils produisent globalement moins d'hésitation, les troncations et les erreurs sont plus nombreuses que chez les natifs, soit moitié plus pour les erreurs et pratiquement deux fois plus pour les troncations. De là on peut proposer que leur manifestation n'est pas spécialement déterminée par la hiérarchie d'accessibilité de Keenan et Comrie (1977). Cette hiérarchie rend compte des tendances caractérisant les erreurs. Les productions syntaxiques inattendues représentent un taux plus élevé à mesure que les relatifs sont moins accessibles. C'est également ce qu'avaient établi Larrivée et Skrovec (2016) pour les locuteurs de langue maternelle. Le taux d'erreur va donc dans le sens d'un rapport entre relatifs et complexité.

5. Discussion conclusive

Ce travail s'intéresse à l'acquisition des relatives. La raison de cet intérêt est que son rapport avec la notion de complexité prédit un profil comparable entre acquisition langue maternelle et langue seconde. Elle est en outre suffisamment bien documentée pour que le rapport avec la hiérarchie d'accessibilité soit crédible, et que ses manifestations soient repérables. Même si certains auteurs identifient des faits de transferts (Hancock et Kirchmeyer 2005), ils restent mineurs, et notre corpus ne contient face à 215 relatifs qu'une seule occurrence d'un *comme* relatif vraisemblablement influencée par le relatif universel *som* du norvégien. Les questions que s'est posé ce travail dans ce cadre étaient les suivantes.

- La hiérarchie d'accessibilité prédit-elle le nombre respectif d'occurrences des relatifs, les taux de collocations et de disfluences ?
- Les facteurs découlant de la complexité se retrouvent-ils de façon comparable chez les natifs et chez les langue seconde ?

Sur le premier point, force est de constater que la proportion de relatifs et le taux de disfluence suivent effectivement la hiérarchie d'accessibilité, à la fois pour

les apprenants norvégiens du français et pour les francophones du corpus étudiés. Par contre, le taux de collocation va exactement dans le sens contraire de ce qui est attendu. Le taux global d'emploi dans une clivée est directement proportionnel à l'accessibilité du marqueur, alors qu'on aurait pu croire à une relation inverse où la faible accessibilité amenait plus d'emploi en collocation pour alléger la lourdeur du traitement. Ce résultat s'observe à la fois chez les apprenants et les natifs. On ne peut que conclure que les apprenants comme les natifs sont face au même déterminisme qui informe leur production.

Cependant, la hiérarchie d'accessibilité n'explique pas tout. On a vu qu'elle ne rendait pas compte des taux de troncation et d'hésitation. On a également constaté que contrairement aux prédictions de la hiérarchie, le relatif *où* manifeste une prépondérance parfois plus importante que *que*, alors que c'est bien ce dernier qui devrait être plus accessible que le premier. Enfin, notre étude transversale amène à se demander si la convergence entre natifs et langue seconde se maintient à tous les niveaux d'acquisition.

De façon plus générale, il y aurait lieu de se demander ce qui explique cette hiérarchie d'accessibilité. On admet aisément que le sujet est moins complexe que l'objet, l'objet direct que l'objet indirect, et l'objet indirect que l'oblique. Cela est-il dû à des facteurs syntaxiques donc, qui ferait dépendre la complexité de la distance entre le relatif et sa position de base ? On peut alors concevoir que le sujet est plus proche de sa position de base, mais il n'est pas clair que cela expliquerait une distinction entre l'objet direct et l'objet indirect. Ou peut-on envisager un facteur de fréquence ? Un exemple en est donné par le relatif *quoi*. Fort rare dans l'usage contemporain, il devrait en tant qu'objet indirect avoir un profil similaire à celui de *dont* dans ces fonctions. Il faudrait un corpus considérable pour pouvoir tester son taux de disfluente par exemple, qui pourrait être évalué par une procédure expérimentale face à *dont* objet indirect. Si les comportements divergent, il faudrait conclure que la fréquence joue un rôle important. Pour prendre un autre exemple, le relatif *où* est assez fréquent, et conjoint des fonctions d'objet indirect et d'oblique. En étudiant ses seules occurrences d'objet indirect, on pourrait, en les comparant avec celles plus rares de *dont* sous cette fonction, voir si les deux formes convergent ou, comme on peut le croire, divergent. Cela montrerait que la fréquence joue un rôle considérable. Ce rôle suffirait-il en retour à expliquer la hiérarchie d'accessibilité ? Est-ce que la fréquence du sujet dans les langues par rapport à l'objet direct, pour autant qu'elle soit avérée, peut servir d'explication pour ce qui a été interprété comme une hiérarchie structurale ? Et si non, comment articuler facteurs structuraux et facteurs de fréquence ? C'est la question qui semble se poser à une approche explicative de l'acquisition des langues.

Références

Corpus

<http://www.crisco.unicaen.fr/Corpus-Apprenants-norvegiens-du.html>
[consulté le 15 juin 2018]

Bibliographie

Bartning, I. 1997. « L'apprenant dit avancé et son acquisition d'une langue étrangère. Tour d'horizon et esquisse d'une caractérisation de la variété avancée ». *AILE* n° 9, p. 9-50.

Hancock, V., N. Kirchmeyer. 2005. Discourse structuring in Advanced L2 French : The relative clause. In: *JFocus on French as a Foreign Language*. Clevedon: Multilingual Matters.

Keenan, E. L., B. Comrie. 1977. « Noun Phrase Accessibility and Universal Grammar ». *Linguistic Inquiry* 8,1, 63-99.

Larrivée, P., M. Skrovec. 2017. Variation en français vernaculaire : le cas des relatives. Manuscrit, 18 pages.

Larrivée, P., M. Skrovec 2016. Les relatives en français vernaculaire. In: *Congrès Mondial de Linguistique Française 2016*. SHS Web of Conferences.

Myles, F. 2015. « Second language acquisition theory and learner corpus research ». In: *The Cambridge Handbook of Corpus Research*. Cambridge: Cambridge University Press.

Olsson, I. 2013. *La relative relativement acquise*. Mémoire de maîtrise, Université de Lund.

Yu, L., T. Odlin (dirs). 2016. *New Perspectives on Transfer in Second Language Learning*. Bristol: Multilingual Matters.

Roubaud, M.-N. 2015. « Accéder aux compétences syntaxiques des élèves (8-11 ans) : le cas des relatives ». In : *Les subordonnées. Corpus, acquisition et didactique*. Rennes : Presses Universitaires de Rennes.

Trévisiol-Okamura, P. 2015. « L'acquisition et l'enseignement des relatives en FLE : regards croisés ». In: *Les subordonnées. Corpus, acquisition et didactique*. Rennes : Presses Universitaires de Rennes.

Welcomme, A. 2013. « Les subordonnées relatives à l'oral en français langue étrangère et en français langue maternelle ». *Mémoires de la Société Néophilologique* n° LXXXVI, p. 97-136.

Note

1. Ce travail a été présenté à l'atelier *L'apprentissage du français et du norvégien à Caen* en janvier 2017, au séminaire d'Anita Thomas à l'Université de Fribourg en mars 2017 et au Congrès de l'Association Pour les Études Nordiques en juin 2017. Je remercie les participants de leurs observations, Catrine Bang Nilsen et Hans Petter Helland, en notant que les erreurs restantes sont toutes de mon fait.