

La fortune des mots : grandeur et décadence de *car*

Benjamin Fagard

Langues, textes, traitement informatique, cognition (LaTTice)
benjamin.fagard@free.fr

Liesbeth Degand

Université Catholique de Louvain
liesbeth.degand@uclouvain.be

1 Introduction

« Qui pourrait rendre raison de la fortune de certains mots et de la proscription de quelques autres ? » se demandait La Bruyère (*Caractères*, 14). Nous proposons une réflexion sur l'histoire mouvementée d'un mot, *car*, que La Bruyère prend précisément comme exemple de ce phénomène, puisqu'il poursuit : « Quelle persécution le *car* n'a-t-il pas essuyée ! et s'il n'eût trouvé de la protection parmi les gens polis, n'était-il pas banni honteusement d'une langue à qui il a rendu de si longs services, sans qu'on sût quel mot lui substituer ? »

Nous analyserons dans un premier temps les phénomènes en jeu dans l'évolution du latin *quā rē* (puis *quārē*) au gallo-roman *car*, à la lumière de la théorie de la grammaticalisation. Nous comparerons en outre cette évolution avec celles déjà observées pour l'expression de la cause dans diverses langues (Heine & Kuteva 2002, Vogl 2007, Evers-Vermeul 2005, Verhagen 2000) ; nous postulons qu'il s'agit d'un phénomène complexe de lexicalisation en latin, avec limitation des contextes d'emploi, puis d'une grammaticalisation avec généralisation progressive d'un sens à des contextes de plus en plus divers, selon une chaîne sémantique [MOYEN > CAUSE]¹. Nous détaillerons ensuite l'évolution sémantique et fonctionnelle de *car*, de l'ancien français au français moderne, en notant les différences par rapport au catalan et à l'occitan *car*.

Dans un deuxième temps, nous étudierons une explication possible de la disparition de *car* en français moderne, à savoir son remplacement par *parce que*². Nous avons conduit à cet effet une étude sur corpus³ des deux termes en français médiéval, classique et moderne, en utilisant diverses bases de données. Notre étude montre une régression progressive des emplois de *car*, et une augmentation des emplois de *parce que* ; cependant, ce n'est qu'en français moderne parlé que le remplacement est acquis, *parce que* ayant une fréquence et des emplois semblables à ceux de *car* en ancien français. Ce remplacement semble donc plutôt une *conséquence* qu'une *cause* de la disparition de *car*.

2 Grammaticalisation de *car*

2.1 De *quā rē* à *car*

L'origine de *car* est à chercher dans une construction latine en syntaxe libre, *quā rē* « pour quelle chose/raison », « pour cette chose/raison », avec un ablatif de cause (voir Ernout & Thomas, 1964 : 94). L'expression de la cause par un élément grammaticalisé constitue un exemple classique de grammaticalisation (voir par exemple, Hopper & Traugott 2003 : 33, 80), dont les textes latins ont conservé les étapes principales. Comme le rappellent Leumann, Hofmann & Szantyr (1933 : II.2.2, 540 sq.), *quārē* (avec cette graphie) apparaît comme interrogatif au début du 2^{ème} siècle avant J.-C. ; l'ex. 1 ci-dessous illustre ce type d'emploi. Il peut également être employé comme relatif complexe (ex. 2).

1. *Quare* negasti ?
« *pourquoi* as-tu nié ? »

2. Multas res novas in edictum addidit, **quare** luxuria reprimeretur
 < ... beaucoup de nouveautés **par quoi** le luxe ... >
 « il ajouta beaucoup de dispositions à l'édit pour réprimer le luxe »

Il évolue ensuite par étapes, à partir de son emploi comme interrogatif : il introduit des interrogations indirectes au 1^{er} siècle av. J.-C., et ce n'est qu'au siècle suivant qu'il est employé comme conjonction de coordination. Cet emploi est le résultat, d'après ces mêmes auteurs, de la disparition de la pause et de l'intonation interrogative ; Väänänen (1981 [1963], § 370) en fournit quelques exemples (notamment l'ex. 3 ci-dessous) rendant possible le glissement vers le sens causal (ex. 4). On trouve également en latin des emplois où *quare* introduit un impératif (ex. 5), qui annoncent les emplois de *car* en ancien français comme renforcement de l'impératif.

3. non es eques ; **quare** ? non sunt tibi milia centum (Suétone, *Tibère*, 59)
 « tu n'es pas chevalier, pourquoi ? tu n'as pas cent mille sesterces »
 4. ita tu qui legis, bona vita vive, sodalis, **quare** post obitum nec risus nec lusus nec ulla voluptus erit (CIL IX 3473)
 « et toi qui lis, mène une belle vie, mon ami, parce qu'après la mort point de rire, de jeu ni de volupté »
 5. **Quare** agite
 « Donc agissez »

Cette grammaticalisation est donc une évolution graduelle, avec réanalyse (cf. ex. 3). Elle s'accompagne d'une augmentation de la fréquence, typique de la grammaticalisation (voir Hopper & Traugott, 1993 : 103). On peut mesurer cette augmentation en comparant un corpus de textes latins⁴ et un autre d'ancien français : en latin, la fréquence relative de *quārē* est de 1,74 occurrence pour 10 000 mots. Or si l'on compare cela aux données sur corpus dont nous disposons pour l'ancien français, on constate qu'il y a une nette augmentation de fréquence du latin au français médiéval (voir tableau 1 ci-dessous).

période	latin classique	ancien français	moyen français
morphème	<i>quārē</i>	<i>quar, quer, car</i>	<i>car</i>
fréquence	1,74	40	41,9

Tableau 1 : fréquence de *quare* et *car* en latin classique, en ancien et en moyen français (en nombre d'occurrences pour 10 000 mots ; Perseus et BFM).

Un autre argument en faveur de la grammaticalisation de *quārē* est l'extension des contextes d'emploi. Ainsi, Ernout & Thomas (ibid. : 315) notent pour l'emploi interrogatif de *quare* qu'il « appartient plutôt à la langue parlée » : le passage à la langue écrite se fera entre le latin classique et l'ancien français, cette extension des contextes d'emploi constituant une caractéristique typique de la grammaticalisation.

Enfin, cette grammaticalisation s'accompagne d'une réduction de la distribution sémantique et fonctionnelle de *car*, ce qui va également dans le sens d'une grammaticalisation croissante. Cette réduction est très nette. L'emploi en syntaxe libre *quā rē* « raison pour laquelle », possible en latin, disparaît entièrement en ancien français. Par ailleurs, même les emplois grammaticalisés subissent une réduction. Ainsi, l'ancien français *car* (ou *quar, quer...*) conserve les emplois cités ci-dessus, comme adverbe interrogatif (ex. 1), adverbe relatif (ex. 2) et conjonction de coordination (ex. 3), mais les seuls emplois réellement fréquents sont ses emplois comme conjonction causale (exemples 6 et 7) d'une part, et comme particule de renforcement de l'impératif d'autre part (ex. 8). De plus, d'un point de vue sémantique, il y a une évolution nette : il n'est plus possible de gloser *quar* par « raison pour laquelle » ou « pour quelle raison » (voir les ex. 6 à 8), comme en latin ; le substantif *res* « chose, objet, affaire » a perdu, en même temps que sa substance phonétique, sa capacité référentielle.

6. Forment m'en poise, **quar** mout l'avoie amé
 « J'en suis fort peiné, **car** je l'avais aimé tendrement » (*Aliscans*, 12^{ème} siècle)
 7. Ferez i, Francs, **kar** très ben les veintrum !

« Frappez, Francs, **car** nous les vaincrons sans peine ! » (*Chanson de Roland*, 11^{ème} siècle)

8. Filz, **quar** t'en vas colcer avoc ta spuse

« Fils, va **donc** te coucher auprès de ton épouse » (*saint Alexis*, 11^{ème} siècle)

L'emploi comme conjonction causale se retrouve dans d'autres langues du domaine gallo-roman, ce qui indique qu'il s'est probablement imposé assez tôt ; on le trouve ainsi en catalan (dès le 12^{ème} siècle, soit les plus anciens textes, mais encore en catalan moderne, cf. ex. 9), en occitan (*car*, *cor* « car, parce que, puisque, afin que, pourquoi » et également dans la construction *per so car* « parce que »). En revanche, on n'en trouve apparemment pas trace dans les autres domaines de la Romania (ibéro-roman, italo-roman).

9. No refusen el seu ajut, car, altrament, s'ofendria

« Ne refusons pas son aide, **car**, autrement, il en prendrait ombrage » (*Enciclopèdia Catalana*⁵)

En ancien français, *car* est donc nettement plus grammaticalisé : d'un point de vue morphologique, on voit qu'il y a fusion des deux mots ; d'un point de vue phonétique, il y a la réduction typique de la grammaticalisation (*qua re* > *quare* > *quar* > *car*) ; d'un point de vue sémantique, il y a généralisation du sens (ou *semantic bleaching*, voir Hopper & Traugott 1993 : 32). Les emplois comme adverbe interrogatif et relatif disparaissent tout à fait, tandis que les emplois comme conjonction et comme particule de renforcement de l'impératif se grammaticalisent. Il y a ainsi à la fois élimination (facultative il est vrai dans le processus de grammaticalisation) de certains emplois, et une grammaticalisation accrue avec généralisation du sens. Ceci permet l'emploi de *quārē* dans des contextes de plus en plus variés, ce qui contribue à expliquer l'augmentation de fréquence.

Nous noterons pour finir cette section que le glissement de l'interrogatif au causal illustré par *quare* se retrouve par ailleurs en italien (*perché*), en allemand (moyen haut all. *wande* « car » < ancien haut all. *hvanda* « pourquoi ? », qui donne d'ailleurs le néerlandais *want* « car » ; dans la langue moderne, *warum*, fam. « parce que »), et en grec moderne (γάρτι) (voir Leumann, Hofmann & Szantyr, ibid.). Ce n'est cependant qu'une des origines possibles pour la grammaticalisation de la cause, comme nous le verrons dans la section suivante.

2.2 Approche typologique : grammaticalisation de la cause

La grammaticalisation des conjonctions de cause peut en effet se faire suivant un certain nombre de schémas connus, qui se retrouvent d'une langue à l'autre et, plus important, d'un groupe linguistique à l'autre. Heine & Kuteva (2002 : 328⁶) notent les chaînes de grammaticalisation partant de notions spatiales, temporelles et *autres* ; Vogl (2007) soutient pour sa part que la chaîne de grammaticalisation [temps > cause] est en fait plus complexe. Il y a selon elle une étape intermédiaire : il faut poser une chaîne du type [temps > condition > cause] – constituant en fait un sous-ensemble des extensions du type [condition > cause].

Cependant, ces chaînes de grammaticalisation ne permettent pas de rendre compte de l'évolution illustrée par *car*, c'est-à-dire le passage de l'interrogatif à la cause. En ajoutant cette possibilité à celles que proposent les auteurs cités plus haut, on voit (tableau 2) la multiplicité des évolutions sémantiques aboutissant à l'expression de la cause.

Domaine sémantique de départ	Chaîne sémantique	Langues concernées
SPATIAL	« ici » > temporel « pendant, quand » > cause	lingala ⁷
	« lieu » > cause	kono, bambara (une seule famille linguistique)
TEMPOREL	« depuis » > cause	français, anglais, basque, aranda
	temporel > cause	vieux haut allemand,

NOTIONNEL		français, latin, finnois, roumain, etc.
	temporel > condition > cause	anglais, allemand, néerlandais
	« matière » > cause	baka, vai, lingala, bulu
	but > cause	to'aba'ita, twi
	« dire » > cause	baka, lezgien
	« pourquoi » (interrogatif) > cause	moyen haut allemand, néerlandais, latin, grec moderne, italien

Tableau 2 : quelques chaînes sémantiques (éventuellement ‘chaînes de grammaticalisation’⁸) aboutissant à la création de conjonctions causales

La grammaticalisation de la cause à partir d’un interrogatif n’est donc clairement qu’une possibilité parmi d’autres. Il reste à voir comment cette grammaticalisation s’est poursuivie, à partir de l’ancien français.

2.3 Evolution sémantique et fonctionnelle de *car*, de l’ancien français au français moderne

En ancien français *car* a, comme nous l’avons vu, des emplois comme conjonction causale et comme ‘particule’ de renforcement accompagnant l’impératif. Cette dernière fonction disparaît assez vite (dès le moyen français), mais l’emploi comme conjonction causale perdure apparemment jusqu’en français moderne – « apparemment » parce que, s’il est certain que *car* est encore employé dans la norme écrite, il nous semble possible d’affirmer qu’il est en passe de disparaître de la langue orale, et ce pour plusieurs raisons. D’abord parce que, d’après une étude de Simon & Degand (2007 : 327-328), il a une fréquence à l’oral qui est extrêmement faible, bien moindre notamment qu’à l’écrit (0,2 occurrences pour 10 000 mots à l’oral contre 4 occurrences pour 10 000 mots à l’écrit – avec pour corpus respectivement la base Valibel⁹ et *Le Soir*). D’autre part parce que, d’après la même étude, le sens de *car* à l’écrit et à l’oral est le même (ce qui n’est pas nécessairement le cas pour tous les connecteurs: voir *parce que* dans la même étude), ce qui indique selon nous que les seuls emplois de *car* à l’oral sont des calques de la langue écrite. Enfin, parce qu’il semble n’y avoir aucune nouvelle extension sémantique depuis l’ancien français, ce qui semble pour certaines classes de mots être lié à la disparition des lexèmes (par exemple pour la disparition des prépositions *coste*, *jouste*, *lez* et leurs dérivés, voir Fagard 2006).

3 Disparition et remplacement de *car*

3.1 Evolution quantitative de quelques conjonctions causales (étude sur corpus)

Le tableau suivant ne suggère pas spécialement un remplacement de *car* par *parce que* ou *puisque*. A l’écrit, la fréquence de *car* semble avoir diminué graduellement de l’ancien français au français moderne (voir les notes 3 et 9 pour la constitution du corpus), perdant plus de 15% de ses emplois d’un siècle à l’autre à partir du moyen français – à l’exception cependant du 19^{ème} siècle, où il y a une légère hausse. Au total, *car* passe d’une fréquence de 40 occurrences pour 10 000 mots en ancien et moyen français (jusqu’au 14^{ème} siècle inclus) à une fréquence de moins de 5 pour 10 000 au 20^{ème} siècle, dans notre corpus. Les conjonctions *puisque* et *parce que*, pour leur part, présentent une évolution bien différente. La première, *puisque*, semble d’abord augmenter du moyen français au français classique, passant de 0,1 à 3 occurrences pour 10 000, mais retombe ensuite légèrement, et se stabilise autour de 2 occurrences pour 10 000 au 20^{ème} siècle (2,66 occurrences pour 10 000 mots, dans la base Valibel). La seconde, *parce que*, présente une augmentation très importante de l’ancien français (0,2 occurrences pour 10 000) au français moderne (près de 5 occurrences pour 10 000), avec un creux au 19^{ème} siècle.

Fréquences (pour 10 000 mots)	1001-1329	1330- 1400	1401- 1502	1501- 1600	1601- 1700	1701- 1800	1801- 1900	1901- 2000
<i>car</i>	40,0 ¹⁰	41,9	30,0	25,7	13,7	6,0	7,4	5,4 ¹¹
<i>parce que</i> (total)	7,2	8,6	6,9	6,6	6,0	5,2	3,5	4,7
<i>par ce que, por ce que</i> (graphie détachée)	7,1	8,2	6,2	1,7	0,4	0,2	0,1	0,1
<i>parce que</i> (graphie liée)	0,2	0,4	0,7	4,9	5,6	5,0	3,4	4,6

Tableau 3 : fréquence de *car* et *parce que*, de l'ancien français au français moderne.

Cependant, malgré son augmentation très nette, *parce que* ne semble pas en mesure de remplacer entièrement *car*. Sa fréquence au 20^{ème} siècle reste en effet près de dix fois inférieure à celle de *car* en ancien français. Cependant, le corpus Valibel (langue parlée) nous révèle que la fréquence de *parce que* est très proche de celle de *car* en ancien français, avec 37 occurrences pour 10 000 (tableau 4 ci-dessous), tandis que *car* semble avoir quasiment disparu du langage oral. Bien sûr, malgré la proximité frappante des données (37 et 40 occurrences pour 10 000 mots), il ne faut pas oublier qu'il s'agit de deux corpus tout à fait différents : écrit des 11^{ème}-14^{ème} siècles d'une part, oral du 20^{ème} siècle d'autre part ; textes littéraires et techniques d'une part, interviews et conversations spontanées d'autre part... Ces résultats sont donc à manipuler avec précaution.

Fréquence (en occurrences pour 10 000 mots)	Car	Parce que
corpus journalistique écrit (<i>Le Soir</i>)	4	3,2
corpus oral (Valibel)	0,2	37

Tableau 4 : fréquence de *car* et *parce que*, à l'oral et à l'écrit, en français moderne.

Il reste à voir, bien sûr, si l'analyse sémantique et fonctionnelle de ces conjonctions permet de renforcer cette hypothèse du remplacement de *car* par *parce que* dans la langue parlée, avec un « retard » très net de la langue écrite. C'est ce que nous allons tenter de faire dans la prochaine section.

3.2 Evolution sémantique

3.2.1 La subjectification

Notre argumentation repose principalement sur une analyse de l'expression de la cause en termes de subjectivité. Si l'on suit Traugott 2005 et Sweetser 1990, la subjectification est un phénomène qui accompagne dans certains cas la grammaticalisation (mais peut également être observé indépendamment) et consiste en l'ancrage progressif d'un terme ou d'une construction dans l'univers de croyances du locuteur : on parle aussi de *degré d'implication du locuteur* (« speaker involvement », voir Pander Maat & Degand 2001)¹². Pour ne donner qu'un exemple, lorsqu'on dit (10), la relation de *cause* employée n'a

rien à voir avec une cause réelle, objective : s'il y a relation de cause à effet, c'est uniquement dans l'esprit du locuteur. Sweetser (1990 : 77-78) parle dans ce cas de conjonction épistémique.

10. Il a plu, parce que la route est mouillée.

La relation de cause « objective » repose sur la relation de cause à effet réelle : la route est mouillée parce qu'il a plu. On a affaire ici à une conjonction de contenu (Sweetser, *ibid.*). La subjectification de (10) vient de l'intégration au terme *parce que* d'un « je conclus que » implicite (11).

11. Je conclus qu'il a plu, parce que la route est mouillée.

On arrive ainsi à des emplois où *car* ou *parce que* n'expriment plus la cause, mais servent à justifier l'énonciation (y compris la « justification d'un dire partiel » pour Bentolila 1986 : 101 ou la « speech-act conjunction » de Sweetser 1990 : 78), ce qui est illustré en (12). On parle alors d'intersubjectification (Traugott & Dasher 2002 : 23-24).

12. Dépêchons-nous, parce qu'il va pleuvoir.

Sweetser défend l'idée que la plupart des conjonctions causales (et adversatives) peuvent exprimer des relations à la fois dans le domaine du contenu (« content domain »), épistémique (« epistemic domain ») et interactionnel (« speech-act domain »). C'est le contexte qui détermine l'interprétation à donner plutôt que la forme linguistique en présence. Ce contexte peut néanmoins porter des indices linguistiques facilitant l'interprétation dans un domaine plutôt que l'autre (voir la section suivante). Pour le français (et le néerlandais), nous avons montré ailleurs (Simon & Degand 2007, Degand & Pander Maat 2003, Pander Maat & Degand 2001) que les différents connecteurs causaux encodent un certain degré d'Implication du Locuteur qui doit être compatible avec son contexte. Si le niveau d'IdL est trop élevé ou trop bas, l'énoncé résultant sera inadéquat. Ainsi, les connecteurs peuvent être rangés sur une échelle scalaire de subjectivité croissante comportant différents niveaux de subjectivité (ou Implication du Locuteur), à savoir les relations causales non-volitives (factuelles), volitives, épistémiques et interactionnelles (actes de langage). De la sorte, ces études ont montré qu'à l'écrit *puisque* exprime des relations causales plus subjectives (surtout épistémiques) que *car*, qui est à son tour plus subjectif (relations surtout volitives et épistémiques) que *parce que* (relations surtout non-volitives et volitives) (Degand & Pander Maat 2003). Ces connecteurs ne peuvent dès lors pas se substituer les uns aux autres dans tous les contextes sans réinterprétation, même si des zones de « chevauchement » existent où la substitution est possible¹³. À l'oral, par contre, la différence sémantique entre *parce que* et *car* disparaît, *parce que* pouvant exprimer toutes les relations réservées à *car* à l'écrit (Simon & Degand 2007)¹⁴. C'est dans ce cadre que nous proposons de parler de subjectification de *parce que*.

3.2.2 Indices de subjectification

On peut mettre en évidence le degré de subjectification d'un terme à l'aide de différents marqueurs ou indices, principalement le type d'énonciation (récit, discours indirect, discours direct), la nature des segments reliés (fait, opinion, action), l'emploi du temps présent et de la première ou deuxième personne (voir Pander Maat & Degand 2001, Pander Maat & Sanders 2001, Degand & Bestgen 2004, Pit 2004)¹⁵. On peut ainsi établir une échelle de subjectivité pour différents emplois d'un même terme, ou plus généralement pour différents types de causalité. Nous avons ainsi défini quatre degrés principaux de subjectivité dans la causalité, de la causalité objective aux actes de langage. Le schéma 1 ci-dessous et les exemples (13-18) correspondants illustrent et précisent cette approche.

cause objective/non-volitive : 1. met en relation de manière factuelle deux « états des choses » (ex. 13)

> cause volitive : 2. explique/justifie la décision / l'action du locuteur (ex. 14-15)

> cause épistémique : 3. explique/justifie l'état d'esprit du locuteur (ex. 16-17)

> cause intersubjective : 4. explique/justifie un acte de discours (ex. 18)

Schéma 1 : échelle de subjectivité pour différents types de causalité (sur la base de Pander Maat & Degand 2001)

13. En la teste ad e dulong grant mal : Rumpu est li temple, **por ço que** il cornat.
« Sa tête lui fait horriblement mal : il a la tempe rompue, **parce qu'**il a sonné le cor ». (*Chanson de Roland*, v. 2102, 11^{ème} siècle)
14. Li duze per, Li païen controverent Les nuns que as jurz dunerent Li premier, que apelum Diemeine par num, Al soleil le dunerent, E sun num li poserent **Pur ço que** enluminout Le mund e nuit chazot.
« Ils lui donnèrent son nom parce qu'il enlumina le monde » (*Comput*, v.429, 12^{ème} siècle)
15. Reposent sei **quar** lassét sunt.
« Ils se reposent, car ils sont las » (*Saint Brendan*, v. 638, 12^{ème} siècle)
16. Sire, dist Guenes, ço ad tut fait Rollant ! Ne l'amerai a trestut mun vivant, Ne Oliver, **por ço qu'il** est si cumpainz.
« Je le haïrai jusqu'à ma mort, et Olivier aussi, parce qu'il lui est proche » (*Chanson de Roland*, v. 324, 11^{ème} siècle)
17. Forment m'en poise, **quar** mout l'avoie amé ; Mes par mon chief ja sera comparé.
« Cela me pèse fort, **car** je l'aimais tendrement ». (*Aliscans*, laisse XLVI, 12^{ème} siècle)
18. Ensembl'od lui i ferrunt veirement. De ço qui calt ? **car** ne lur valt nient. Demurent trop, n'i poedent estre a tens.
« Mais qu'importe ? **Car** cela ne leur sert à rien » (*Chanson de Roland*, v. 1840, 11^{ème} siècle)

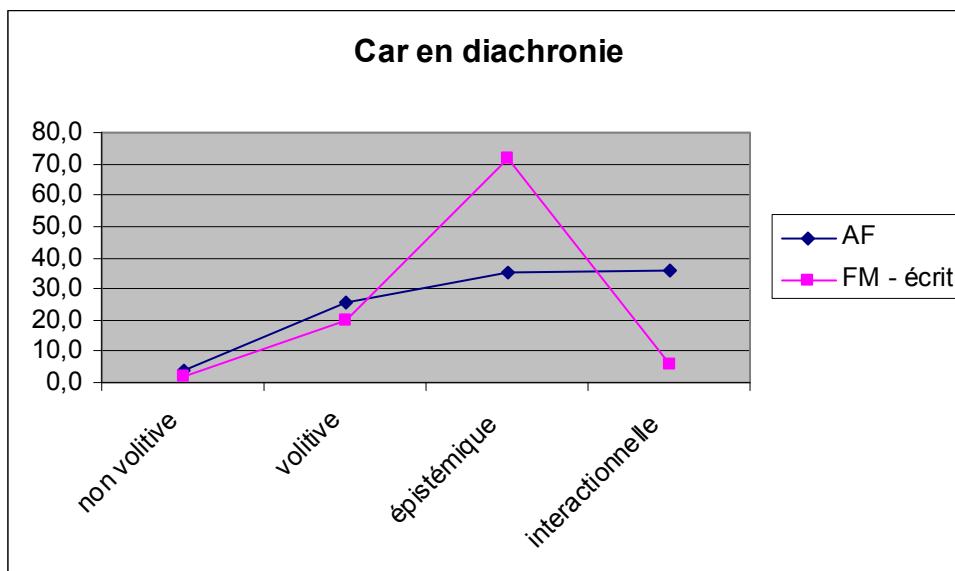
Comme on le voit, la grammaticalisation de *car* est achevée dès l'ancien français : d'adverbe et relatif complexe, *car* est devenu conjonction de subordination puis de coordination, avec des emplois aussi bien subjectifs (ex. 17) qu'intersubjectifs¹⁶ (ex. 18). Son évolution jusqu'en français moderne ne s'accompagne pas de changement morpho-syntaxique majeur, et se résume à la perte d'une partie de ses emplois, comme nous allons le voir.

3.2.3 Sémantique de *car* et *parce que* en ancien et en moyen français

Si l'on observe de près le sémantisme de *car* et *parce que* en ancien français, on constate qu'il y a un net déséquilibre entre les emplois des deux conjonctions : *car* est nettement plus (inter-)subjectif que *parce que*¹⁷. On remarque ainsi que la cause intersubjective n'est jamais représentée dans notre corpus pour *parce que* – on ne trouve que des emplois de cause objective (ex. 13) et, rarement, subjective (ex. 14 – cf. aussi l'emploi de *car* dans l'ex. 15). Les emplois de *car*, en revanche, sont nettement subjectifs et intersubjectifs (ex. 4, 5, 15, 17, 18), et on trouve très peu d'emplois objectifs. On a affaire ici à deux marqueurs bien distincts, qui contrairement à ce qui se passe en français moderne, ne semblent pas pouvoir se substituer facilement l'un à l'autre, si ce n'est dans quelques rares cas de relations causales non-volitives et volitives¹⁸. Nous verrons dans la section suivante comment cette situation de départ a progressivement évolué jusqu'au français moderne parlé.

4 Etude sur corpus : remplacement de *car* par *parce que* ? De l'ancien français au français moderne : écrit et oral

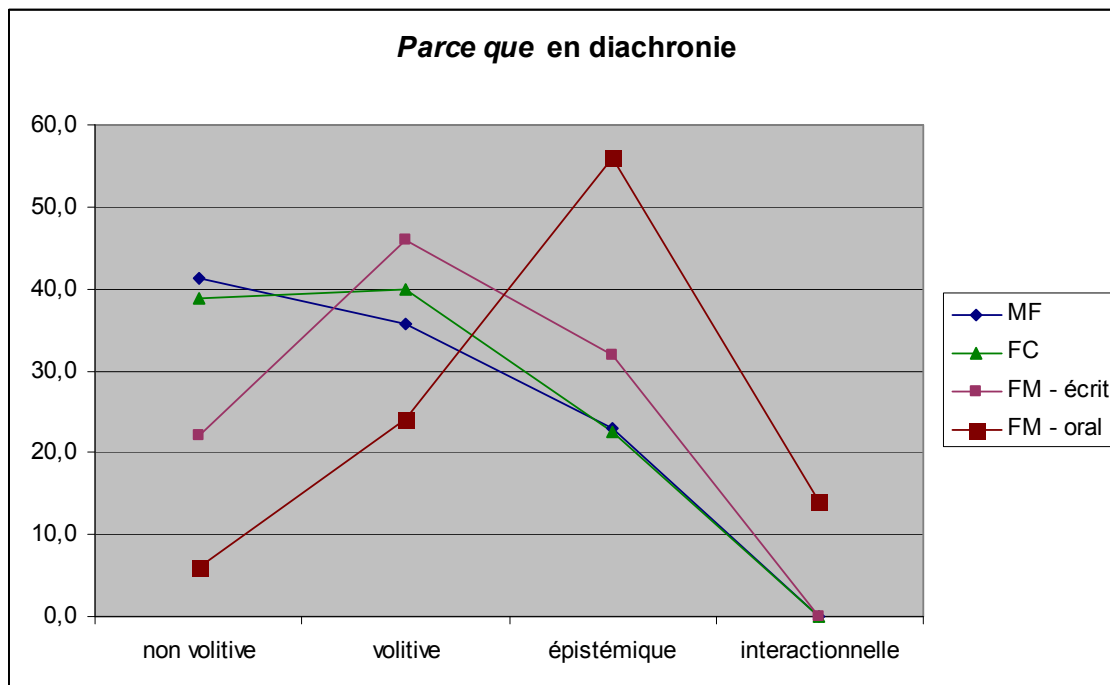
Nous présentons ici les résultats de notre étude sur corpus, qui porte (pour l'analyse sémantique) sur près de cinq cents occurrences de chaque marqueur, sélectionnées aléatoirement. Dans le corpus écrit, de l'ancien français (AF) au français moderne (FM), l'évolution sémantique de *car* est assez limitée, comme le suggère le graphique 2 ci-dessous.



Graphique 2 : évolution de *car* de l'ancien français au français moderne (écrit et oral)

Le processus de subjectification semble avoir continué pour *car*, avec une baisse des emplois de type « cause volitive » (plus objectifs) et une augmentation des emplois de type « cause épistémique » (plus subjectifs, cf. section précédente). La baisse des emplois de type « cause interactionnelle » – donc des emplois intersubjectifs – ne reflète pas selon nous une « désobjectification » : elle est liée, plutôt, à la disparition de *car* de la langue orale (où la part des emplois intersubjectifs est naturellement plus élevée). En effet, le comportement de *car* à l'écrit et à l'oral, en français moderne, est presque identique, comme le montre bien le graphique 2, la seule différence notable est la part plus élevée des emplois intersubjectifs, et la part plus faible, en conséquence, des autres emplois. Ces différences ne sont toutefois pas statistiquement significatives ($\chi^2 = 5.74$, $df = 3$ et $p = 0.14$ (N.S.)).

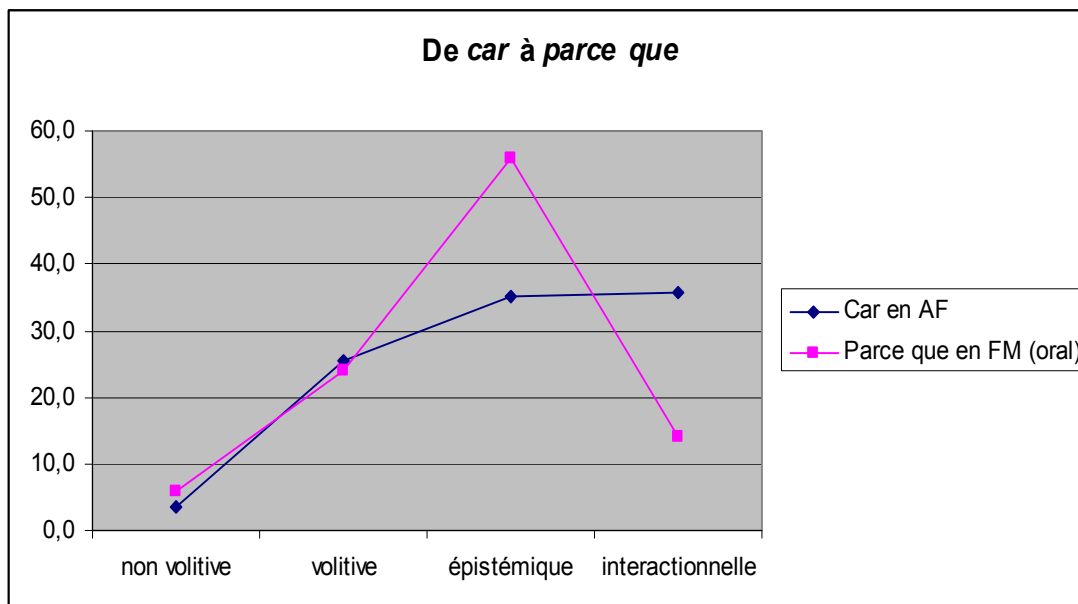
L'absence de contraste entre le *car* écrit et le *car* oral est particulièrement frappante si l'on compare les emplois de *parce que* à l'écrit et à l'oral. Le graphique 3 ci-dessous montre une évolution relativement régulière entre le moyen français (MF), le français classique (FC) et le français moderne (FM) écrit et oral, avec un processus de subjectification très net.



Graphique 3 : évolution de *parce que* entre moyen français, français classique et français moderne (écrit et oral)

Les emplois de cause non volitive baissent de manière très nette, tandis qu'augmentent les emplois de cause volitive puis ceux de cause épistémique, et que, enfin, des emplois intersubjectifs apparaissent à l'oral, en français moderne. Enfin, on remarque une différence très nette entre le français moderne écrit (avec une majorité d'emplois volitifs) et oral (majorité d'emplois épistémiques, et présence d'emplois interactionnels). Cette distinction entre l'oral et l'écrit pour *parce que* est statistiquement significative ($\chi^2 = 18.301$, $df = 3$ et $p < .0001$).

L'évolution de *parce que* laisse donc penser que cette conjonction tend déjà, à l'écrit, à se rapprocher des emplois de *car* en ancien français. Cependant, l'argument majeur pour le remplacement de *car* par *parce que* est ce qu'on peut observer à l'oral. Une comparaison de nos résultats en ancien français avec une étude de *car* et *parce que* sur un corpus de français oral (Simon & Degand 2007) montre en effet des similitudes frappantes entre l'emploi de *car* en ancien français et celui de *parce que* en français moderne, à l'oral (graphique 4).



Graphique 4 : *car* en AF, *parce que* en FM oral

Du point de vue de la subjectification, *parce que* semble avoir remplacé *car* en français moderne parlé – *parce que* restant cependant moins intersubjectif que *car* en ancien français. Les données de fréquence semblent bien confirmer qu'il y a eu remplacement : d'une part *car* a presque complètement disparu (avec 0,2 occurrences pour 10 000 mots en français moderne parlé), d'autre part *parce que* a en français moderne oral une fréquence comparable à *car* en ancien français. Statistiquement, la différence entre le *car* ancien et le *parce que* oral moderne est significative ($\chi^2 = 10,264$, $df = 3$ et $p < .05$). Ceci indique que *parce que* n'a pas (encore) entièrement remplacé *car*, en particulier dans le domaine interactionnel. Cette dernière différence pourrait bien sûr être liée à nos corpus. Il faut garder à l'esprit qu'il est délicat de comparer des corpus écrits et oraux de types bien différents ; en ancien français des écrits de type majoritairement littéraire (avec des textes historiques et juridiques en nombre inférieur) contre des conversations spontanées et des interviews en français moderne oral. Une solution pourrait être d'étudier des corpus intermédiaires, qu'il s'agisse d'écrit oralisé (internet, messages électroniques, SMS) ou d'oral formel (enregistrement de conférences, de discours politiques), afin d'observer l'interaction des normes.

5 Conclusion

A la lumière des faits présentés ci-dessus, il ne semble pas exagéré de parler, pour *car*, d'une « histoire mouvementée », comme nous l'avons fait en introduction. Son évolution implique l'intervention de divers processus – de lexicalisation et de grammaticalisation en latin, de grammaticalisation ainsi que de subjectification entre latin et gallo-roman, et de subjectification entre ancien français et français moderne. De l'ancien français au français moderne, malgré la poursuite de la subjectification, il semble que *car* ait constamment perdu en fréquence, selon un lent processus dont nous avons montré que tous les éléments sont réunis, en français moderne, pour qu'il aboutisse à la disparition totale de *car*, du moins dans la langue orale.

L'étude de l'évolution sémantique de *parce que* semble cependant exclure qu'elle ait pu servir de catalyseur à cette disparition : il y a apparemment remplacement de *car* par *parce que* en français moderne, mais ce n'est vrai qu'à l'oral, et ce doit donc être un phénomène relativement récent. Il faudrait cependant compléter ces données à l'aide de corpus oraux plus anciens.

Si l'on veut expliquer la disparition de *car* – plutôt que de se contenter de la décrire – une piste de recherche pourrait être l'étude plus systématique de l'évolution des modes d'expression de la cause, de l'ancien français au français moderne. On peut en effet exprimer la cause de manières très diverses : parataxe, participiales, groupes prépositionnels, etc. Il serait intéressant de voir lesquelles parmi ces structures ont vu leur fréquence augmenter au fil des siècles influençant ainsi le « système », au sens Saussurien, de la causalité en français.

Remerciements

Cette recherche est soutenue financièrement par un Pôle d'attraction interuniversitaire financé par le gouvernement fédéral belge sous le contrat PAI P6/44 « Grammaticalization and (inter)Subjectification ». La seconde auteure est chercheuse qualifiée du Fonds de Recherche Scientifique (FRS-FNRS) financé par la Communauté française de Belgique.

Références bibliographiques

- Bentolila, F. (1986). *Car* en français écrit. *La linguistique*, 22, 95-115.
- Degand, L. et Y. Bestgen. (2004). Connecteurs et analyses de corpus : de l'analyse manuelle à l'analyse automatisée. In: S. Porhiel et D. Klingler (eds.), *L'Unité texte*. Pleyben : Perspectives, 49-73.
- Degand, Liesbeth et H. Pander Maat (2003). A contrastive study of Dutch and French causal connectives on the Speaker Involvement Scale, A. Verhagen & J. van de Weijer (eds.) *Usage based approaches to Dutch*. Utrecht: LOT, 175-199.
- Ernout, A. et Fr. Thomas. (1964). *Syntaxe latine*, Paris : Klincksieck. 2^{ème} édition.
- Evers-Vermeul J. (2005). *The development of Dutch connectives; Change and acquisition as windows on form-function relations*. Thèse de doctorat. Utrecht : Utrecht University, UiL OTS.
- Fagard, B. (2006). *Evolution sémantique des prépositions dans les langues romanes : illustrations ou contre-exemples de la primauté du spatial ?* Thèse, Université Paris 7 / Università Roma 3.
- Francard, M., Geron, G., & Wilmet, R. (2002). La banque de données VALIBEL : des ressources textuelles orales pour l'étude du français en Wallonie et à Bruxelles. In : C. D. Pusch et W. Raible (dir.). *Romanistische Korpuslinguistik - Korpora und gesprochene Sprache / Romance Corpus Linguistics - Corpora and Spoken Language (= ScriptOralia; 126)*. Tübingen : Gunter Narr, 71-80.
- Frei, H. (1982 [1929]). *La grammaire des fautes*. Genève : Slatkine Reprints.
- Groupe A-I (1975). *Car, parce que, puisque*. *Revue Romane*, 10, 248-280.
- Heine, B. et T. Kuteva. (2002). *World lexicon of grammaticalization*. Cambridge : Cambridge University Press.
- Hopper, P. et E. Traugott. (1993). *Grammaticalization*. Cambridge : Cambridge University Press.
- Leumann, M., Hofmann, J.B. et A. Szantyr. (1965). *Lateinische Grammatik (Zweiter Band : Lateinische Syntax und Stilistik)*. Munich : Handbuch der Altertumswissenschaft.
- Pander Maat, H. et L. Degand (2001). Scaling causal relations and connectives in terms of Speaker Involvement. *Cognitive Linguistics*, 12/3, 211-245.
- Pander Maat, H. et Sanders, T. (2001). Subjectivity in causal connectives: An empirical study of language in use. *Cognitive Linguistics*, 12/3, 247-273.
- Pit, M. (2003). *How to express yourself with a causal connective. Subjectivity and causal connectives in Dutch, German and French*. Amsterdam / New York : Rodopi.
- Simon, A.C et L. Degand. (2007). Connecteurs de causalité, implication du locuteur et profils prosodiques. Le cas de *car* et de *parce que*. *Journal of French Language Studies*, 17, 323-341.

- Stukker, N. (2005). Causality marking across levels of language structure: A cognitive semantic analysis of causal verbs and causal connectives in Dutch. Dissertation doctorale, Université d'Utrecht [LOT Dissertation Series 118].
- Sweetser, E. (1990). *From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure*. Cambridge : Cambridge University Press.
- Traugott, E. (2003). From subjectification to intersubjectification. In: R. Hickey (ed.), *Motives for language change*. Cambridge: Cambridge University Press, 124-139.
- Traugott, E. et R. Dasher. (2002). *Regularity in Semantic Change*. Cambridge : Cambridge University Press.
- Väänänen, V. (1981 [1963]). *Introduction au latin vulgaire*. Paris : Klincksieck.
- Verhagen, A. (2000). Interpreting Usage: Construing the History of Dutch Causal Verbs. In M. Barlow et S. Kemmer (eds.), *Usage-Based Models of Language*, 261-286.
- Vogl, U. (2007). Het belang van conditionaliteit voor de ontwikkeling van temporeel naar causaal voegwoord. De geschiedenis van *dewijl, terwijl, weil* en *while*. *Nederlandse Taalkunde*, 12/1, 2-24.

¹ On retrouve une évolution similaire beaucoup plus tard, avec l'émergence des constructions *par ce que* « par ce fait que » > « à cause du fait que » en ancien français.

² Le remplacement de *car* par *parce que*, à l'oral, a été noté également par Frei (1982 : 229, cité par Bentolila 1986 : 96).

³ Notre corpus est composé des bases de données suivantes : BFM – Base de Français Médiéval [En ligne]. Lyon : UMR5191 ICAR / ENS-LSH, 2005, <http://bfm.ens-lsh.fr> ; DMF – Base du Dictionnaire de Moyen Français, UMR7118 ATILF / Nancy2, < <http://atilf.atilf.fr/dmf.htm> > ; Frantext, UMR7118 ATILF / Nancy2, <http://www.frantext.fr>. Pour le latin, nous avons eu recours au moteur de recherche Perseus, qui repose sur une base de données contenant des textes latins comprenant plus de 5 millions de mots.

Les corpus utilisés pour l'étude diachronique comprennent tous des textes assez variés typologiquement : textes littéraires, historiques et juridiques, en prose et en vers. Il faut noter cependant que, pour la période la plus ancienne, les textes disponibles sont en majorité littéraires.

⁴ Base Perseus, Tufts University, disponible en ligne : <http://www.perseus.tufts.edu/>.

⁵ <http://www.encyclopedia.cat/>.

⁶ Plus précisément Heine & Kuteva (2002 : 48, 171, 200, 210-211, 239, 246-247, 261, 291).

⁷ En moré, on trouve une postposition causale suivant cette même évolution.

⁸ L'existence et le statut de ces chaînes de grammaticalisation reste à vérifier. Nous les mentionnons dans une optique heuristique, non comme une donnée de fait.

⁹ La banque de données textuelles orales Valibel contient actuellement plus de 400 heures d'enregistrement transcrites en orthographe standard, correspondant à +/- 3,7 millions de mots. Environ 42% sont des entrevues portant sur des questions sociolinguistiques ; 30% sont des entrevues informelles (entre deux ou plusieurs locuteurs qui se connaissent) et 6% sont des conversations non sollicitées par un enquêteur. 6% relèvent des médias (interviews radiophoniques, débats, émissions d'information), 2,5% relèvent du secteur professionnel (réunions de travail) et un autre 2,5% du secteur scolaire (cours). Enfin, 11% sont constitués de tâches de lecture. Cette répartition évolue chaque année au gré des projets de recherche. Voir aussi Francard, Geron et Wilmet, 2002.

¹⁰ Tous emplois confondus.

¹¹ Si l'on observe plus particulièrement les dix dernières années (à partir de 1995), il y a 1482 occurrences de *car* pour 3 007 045 mots, soit une fréquence relative de 4,93 : la tendance à la baisse continue donc. Il nous faut noter ici que ces données incluent quelques occurrences d'homonymes de *car* (noms propres ou encore nom commun « autobus »), mais en quantité statistiquement insignifiante : 5 occurrences au 17^{ème}, 4 au 18^{ème}, 3 au 19^{ème}, 430 au 20^{ème}, 5 au 21^{ème} (soit moins de 0,1 occurrence pour 10 000 mots).

¹² Ce type de concept a d'ailleurs déjà été mis en œuvre pour la description de *car*, marqueur « d'actes de parole » (Groupe λ-1, 1975 : 254), en particulier pour opposer *puisque* et *car* à *parce que*, dont il est affirmé qu'il ne constitue

qu'un « opérateur » (ibid.). L'emploi de *parce que* à l'oral montre cependant que cette conjonction peut désormais avoir des emplois subjectifs, au même titre que *car*, comme le note Bentolila (1986 : 96, note 4).

¹³ Nous avons néanmoins montré que chaque connecteur encode un certain degré d'implication du locuteur qu'il apporte à son contexte d'énonciation influençant ainsi l'interprétation dans un sens plutôt objectivant ou plutôt subjectivant qui correspond à son emploi prototypique (voir aussi Stukker 2005 pour une interprétation similaire).

¹⁴ Nous ne pouvons pas nous prononcer sur la sémantique de *puisque* à l'oral, ce connecteur n'ayant à notre connaissance pas encore fait l'objet d'une telle étude.

¹⁵ Ces éléments constituent des *indices* de subjectification ; ils ne sont pas pour autant indispensables, comme le montre le fait que les exemples 17 et 18, pourtant hautement subjectifs, ne présentent pas de première ou deuxième personne.

¹⁶ L'emploi en tant que particule de renforcement de l'impératif (voir ex. 6) a peut-être facilité le développement de ces emplois intersubjectifs ; une étude typologique devrait néanmoins nous éclairer sur cette question puisque d'autres langues ont développé cet emploi intersubjectif pour des marqueurs qui n'ont pas connu cette fonction de renforcement de l'impératif.

¹⁷ L'étude sur corpus montre que les deux conjonctions ont un sémantisme très différent en ancien et moyen français. D'abord, *parce que* a des emplois causaux mais aussi de but (ex. i) et de moyen/manière (avec toutefois une possible inférence causale, comme dans l'ex. ii), tandis que *car*, comme nous l'avons vu plus haut, servait également au renforcement de l'impératif (ex. 6).

(i) Sunent mil grailles **por ço que** plus bel seit.

« Ils font sonner mille cloches afin que cela soit plus beau » (*Chanson de Roland*, v. 1004, 11^{ème} siècle)

(ii) En grant peine nus mist **Par ço quē** il mangat Ço que Eve li dunat Sur le defens de Dé

« Il nous mit dans une peine sans pareille en mangeant ce qu'Eve lui avait donné malgré l'interdiction de Dieu » (*Comput*, v. 533, 12^{ème} siècle)

¹⁸ Les tests de substitution reposant sur les intuitions linguistiques des locuteurs, il est hasardeux de vouloir en réaliser sur des états de langues anciens.