



L'emploi des connecteurs causaux français et finnois dans un corpus littéraire

MIKKO HAULE

L'auteur a soutenu sa thèse de doctorat *L'emploi des connecteurs causaux français et finnois dans un corpus littéraire* à la Faculté des Lettres de l'Université de Helsinki le 20 août 2024. Le rapporteur était le chercheur Benjamin Fagard (Lattice - CNRS, ENS - PSL & Paris Sorbonne Nouvelle) et le président de soutenance était le professeur Kaius Sinnemäki. La thèse de doctorat est disponible à l'adresse suivante : <http://urn.fi/URN:ISBN:978-951-51-9859-4>.

Mots-clés connecteur causal, relation causale, langue française, langue finnoise, subjectivité, marqueur d'oralité, analyse contrastive, analyse statistique, régression logistique à effets mixtes

La causalité peut être exprimée de différentes manières dans différentes langues. On peut par exemple utiliser des verbes, des noms ou des locutions prépositives qui expriment la cause. Dans ma thèse de doctorat, je me suis concentré sur l'emploi des connecteurs causaux français et finnois dans la langue littéraire. En réfléchissant aux sujets possibles pour une future thèse, mon attention a été attirée par des articles traitant des connecteurs causaux en français *parce que*, *car* et *puisque*. Dès lors, je me suis interrogé sur les facteurs pouvant expliquer le choix d'un locuteur ou d'un écrivain pour tel ou tel connecteur dans un certain contexte. De plus, je me suis demandé si les mêmes facteurs pouvaient être liés à l'utilisation de connecteurs causaux également dans ma langue maternelle, le finnois, qui, contrairement au français, est une langue finno-ougrienne.

J'ai donc décidé d'examiner l'emploi de trois connecteurs causaux français et finnois dans ma thèse. Les connecteurs français analysés sont ceux que je viens de mentionner : *parce que*, *car* et *puisque*. Les connecteurs finnois sont *koska*, *sillä* et *kun*. Le connecteur *kun* a également des sens autres que causaux, par exemple un sens temporel (Herlin 1998), mais dans ma recherche, je n'analyse que son emploi causal. Tous ces connecteurs sont fréquents et peuvent être utilisés dans des structures syntaxiquement similaires (v. p. ex. Havu & Hyvärinen 2013 : 169–174).

Les études antérieures portant sur les connecteurs en français et dans certaines autres langues ont montré statistiquement que de nombreux facteurs liés à la sémantique, à la pragmatique et au registre sont associés au choix du connecteur (p. ex. Degand & Pander Maat 2003, Pit 2003, Degand &

Fagard 2012, Zufferey 2012, 2014, Zufferey *et al.* 2018). En ce qui concerne les connecteurs en finnois, les études précédentes se sont concentrées sur leur analyse qualitative. Mon objectif était, premièrement, d'affiner la compréhension des similitudes et des différences entre les connecteurs français en examinant l'effet simultané de différents facteurs sur le choix du connecteur en français littéraire. Deuxièmement, il s'agissait de découvrir de la même manière les différences et les similitudes entre les connecteurs en finnois et, troisièmement, de savoir dans quelle mesure les connecteurs français et finnois se correspondaient.

Mon étude porte sur la langue contemporaine. Aussi bien le corpus français que le corpus finnois sont constitués de 30 ouvrages littéraires contemporains parus entre 1991 et 2015. Il s'agit principalement de romans, mais le corpus inclut également quelques recueils de nouvelles.

Certains de ces connecteurs peuvent être utilisés dans une grande variété de structures syntaxiques, tandis que dans de nombreux cas, la seule option en français serait *parce que* et en finnois *koska* (p. ex. Groupe λ -1 1975 : 250–254, Havu & Hyvärinen 2013 : 169–174). Comme les études antérieures ont déjà montré les cas où certains connecteurs ne pourraient pas être utilisés, j'ai délimité le sujet de ma thèse seulement aux cas où il serait syntaxiquement possible d'utiliser chacun de ces connecteurs. J'ai analysé uniquement les occurrences des connecteurs qui sont précédées d'un segment de conséquence et suivies d'un segment de cause, comme dans l'exemple : « Je suis sorti [CONNECTEUR] il faisait beau », en finnois : « Menin ulos [CONNECTEUR] oli kaunis ilma. »

J'ai utilisé un échantillon assez large, ce qui permet, entre autres, de prendre en compte la variation de la fréquence des connecteurs selon les ouvrages dans l'analyse statistique. J'ai pris un échantillon de 230 pages au maximum de chaque ouvrage français. Les fréquences des occurrences des connecteurs sont présentées dans le tableau 1. Au total, les occurrences des connecteurs en français analysées sont au nombre de 1450. *Parce que* et *car* sont de loin les connecteurs les plus fréquents, tandis que *puisque* est bien plus rare. Il n'y a que 162 occurrences de *puisque* dans les données, soit environ 11 % des occurrences des connecteurs en français. Quant au finnois, j'ai pris un échantillon dans lequel la somme des nombres des occurrences des connecteurs *koska* et *sillä* correspond à la somme des deux connecteurs

les plus fréquents en français, *parce que* et *car*. Le nombre des occurrences des connecteurs en finnois analysées est de 1836, mais leur nombre se distribue d'une manière bien plus égale qu'en français. Une différence essentielle ressort donc de mes données : en français, il y a deux connecteurs causaux très fréquents, tandis qu'en finnois, il y en a trois.

Tableau 1. Nombre des occurrences analysées.

<i>Parce que</i>	<i>Car</i>	<i>Puisque</i>	Total
667	621	162	1450
<i>Koska</i>	<i>Sillä</i>	<i>Kun</i>	Total
681	607	548	1836

Dans cette étude, j'ai utilisé des méthodes statistiques pour déterminer quels types de facteurs contextuels expliquent le choix de chaque connecteur. Ces facteurs analysés, que j'appelle *variables linguistiques*, sont au nombre de dix. Huit variables sont liées à l'expression de la subjectivité et deux à la représentation de l'oral. J'ai classé chaque occurrence de connecteur en fonction de ces variables. De plus, dans les données, les occurrences ont été catégorisées selon l'ouvrage dont elles proviennent. La méthode statistique utilisée est la régression logistique à effets mixtes (*mixed-effects logistic regression*) (v. p. ex. Schäfer 2020), qui permet d'examiner l'effet simultané des variables linguistiques sur le choix des connecteurs, tout en prenant en compte les différences entre les fréquences des connecteurs par ouvrage. Il est à noter que les ouvrages du corpus diffèrent considérablement les uns des autres quant au nombre de connecteurs différents. Certains ouvrages n'incluent aucune occurrence d'un connecteur donné, tandis que d'autres connecteurs sont très fréquents dans ces mêmes ouvrages.

Comme je viens de le constater, la notion de subjectivité est associée à la plupart des variables analysées. Le *type de relation causale* est l'une des variables les plus pertinentes dans cette étude. Les relations causales ont été

classées différemment dans les études précédentes (p. ex. Degand & Pander Maat 2003, Pit 2003, Simon & Degand 2007). J'ai utilisé principalement une catégorisation qui inclut dix types de relation causale.

Je n'entre pas ici dans les détails de la classification utilisée dans mon étude ; je me contente de donner des exemples de deux types de relation causale. Dans l'exemple 1, il s'agit d'une relation causale non volitive où le segment de cause exprime simplement la cause pour la conséquence qui est exprimée dans le premier segment, et aucune personne ne s'implique dans la relation :

1. L'ordinateur est cassé *parce qu'*il est tombé par terre.

En revanche, dans l'exemple 2, il s'agit d'une relation épistémique. Ce type de relation peut également être décrit comme une relation justificative. L'information « l'ordinateur est probablement tombé par terre » présentée dans le segment de conséquence n'est pas vraiment un fait, mais une conclusion que le locuteur tire sur la base de l'information présentée dans le segment de cause. Pour cette raison, la relation épistémique est considérée comme plus subjective que la relation non volitive dans l'exemple 1, où la conséquence est un phénomène objectivement observable.

2. L'ordinateur est probablement tombé par terre *parce qu'*il est cassé.

Une autre variable liée à la subjectivité est le *participant*, c'est-à-dire la personne impliquée dans la relation causale. Différents types de participants reflètent différents degrés sur l'échelle de subjectivité. À titre d'illustration, dans l'exemple 3, le participant, le « je », est à la première personne, tandis que dans l'exemple 4, le participant, Mme Michel, est à la troisième personne.

3. ...et je dis bonjour *car* il n'y a aucune raison de ne pas le faire. (*Week-end de chasse à la mère* : 200)
4. Mme Michel a appelé le SAMU *parce que* sa cheville commençait à ressembler à une pastèque [...] (*L'élégance du hérisson* : 80)

La première personne, c'est-à-dire le locuteur, peut être considérée comme plus subjective que la troisième personne, parce que le locuteur n'adopte pas le point de vue d'une autre personne dans la même mesure que son propre point de vue (Pander Maat & Degand 2001 : 215).

J'examine également la présence de différents *subjectivèmes*, c'est-à-dire de marqueurs de subjectivité, d'une part dans le segment de conséquence, et d'autre part, dans le segment de cause. Ces subjectivèmes comprennent, par exemple, des adverbes subjectifs tels que 'bien sûr', des adjectifs subjectifs tels que 'excellent' ou des verbes tels que 'aimer' ou 'détester'. La présence de subjectivèmes augmente le degré de l'expression de la subjectivité. J'ai classé chaque occurrence de connecteur selon qu'il y avait un ou plusieurs subjectivèmes dans le segment de conséquence ou que ces subjectivèmes étaient absents. Le même critère s'applique au segment de cause. Dans l'exemple 5, le segment de conséquence n'inclut aucun subjectivème, mais le segment de cause inclut le subjectivème « bon ».

5. Je te dis la guerre du Golfe *parce que* c'est un bon exemple. (*Je voudrais que quelqu'un m'attende quelque part* : 46)

Les occurrences ont également été classées selon le temps et le mode du verbe des segments de conséquence et de cause. En particulier, je distingue i) les temps passés de l'indicatif, ii) le présent et le futur de l'indicatif et iii) le conditionnel. Je considère que les temps passés de l'indicatif sont des temps plus objectifs que le présent, parce qu'ils permettent d'exprimer des événements plus éloignés de l'« ici et maintenant » du locuteur (v. p. ex. Pit 2003 : 146–151). J'attribue au conditionnel un degré de subjectivité plus élevé qu'à l'indicatif, parce que le conditionnel permet d'indiquer, entre autres, une certaine attitude du participant.

De plus, j'ai classé les occurrences des connecteurs en fonction du *statut informationnel* des segments de conséquence et de cause. Par statut informationnel, je fais référence à la différence entre une information nouvelle et une information connue. Cette dernière peut consister en ce qui a été préalablement mentionné ou elle peut relever de connaissances générales ou être évidente dans une situation communicative, tandis que

d'autres types d'informations sont nouvelles. Je considère qu'une information connue est plus objective qu'une information nouvelle, car elle est objectivée et factualisée parce qu'elle a déjà été mentionnée et n'est donc plus ouverte au débat (v. p. ex. Pit 2003 : 158–159).

Dans le segment de conséquence de l'exemple 6, l'information « je t'appelle » est une information connue, parce que, évidemment, l'interlocuteur, c'est-à-dire la personne qui a reçu l'appel, est au courant de cet appel. En revanche, l'information dans le segment de cause, « je voudrais te voir », est nouvelle pour l'interlocuteur.

6. — Je t'appelle *parce que* je voudrais te voir. (*Je voudrais que quelqu'un m'attende quelque part* : 106)

En plus de ces facteurs liés à la subjectivité, j'ai classé les occurrences des connecteurs selon deux variables liées à la représentation de l'oral. Premièrement, j'examine la question de savoir si les connecteurs apparaissent dans une séquence dialogique ou dans une séquence non dialogique. La séquence non dialogique correspond généralement à une séquence de narration.

Deuxièmement, j'examine la présence de marqueurs d'oralité dans le contexte du connecteur. Le terme « oralité » inclut plusieurs registres dont, par exemple, le registre familier ou populaire. L'exemple 7 ne contient pas de tels marqueurs, tandis que, dans l'exemple 8, je considère le mot « flic » comme un marqueur d'oralité lexical. Dans l'exemple 9, l'expression « t'es », utilisée au lieu de « tu es », est un marqueur d'oralité que je définis comme « phonétique ou grammatical ».

7. Les seuls motifs que j'aurais le droit d'invoquer — *car* j'aurais à expliquer les raisons pour lesquelles je quittais un poste aussi avantageux — seraient des arguments énoncés à la première personne du singulier. (*Stupeur et tremblements* : 165)

8. Bon, je raccroche *car* le flic me fait signe. (*Les Baltringues* : 185)

9. — Je le sais. *Parce que t'es différent. T'es un type bien, Markikette. (Le Livre des Baltimore : 264)*

Je considère que ce dernier type de marqueur d'oralité est plus fort que le marqueur lexical, parce que celui-ci ne viole pas les règles grammaticales ni les règles d'orthographe (v. p. ex. Rouayrenc 2010 : 236).

L'un des résultats principaux de cette étude est que les deux connecteurs français les plus fréquents, *parce que* et *car*, diffèrent l'un de l'autre d'une manière très similaire aux connecteurs *koska* et *sillä*. En comparant ces paires de connecteurs, on peut constater que l'emploi des connecteurs *parce que* et *koska* est favorisé par les contextes objectifs, c'est-à-dire les catégories de variables qui expriment un faible degré de subjectivité, comme les types de relation objectifs et le participant à la troisième personne. En outre, l'emploi de ces connecteurs est favorisé par les contextes « oraux », c'est-à-dire par les séquences dialogiques et la présence de marqueurs d'oralité. En revanche, les contextes subjectifs, par exemple le participant à la première personne et la présence de subjectivèmes, et les contextes « non oraux » favorisent le choix des connecteurs *car* et *sillä*. Dans la comparaison entre ces connecteurs, la variation du connecteur selon les ouvrages individuels est considérable et explique une plus grande partie de la variation que les variables linguistiques.

En ce qui concerne les connecteurs *puisque* et *kun*, ils se ressemblent principalement en ce que leur choix est favorisé par les informations connues dans le segment de cause. Cependant, l'emploi de *kun* est également très fortement favorisé par les séquences dialogiques et la représentation de l'oral.

Sur la base de ces résultats, les connecteurs français et finnois qui se ressemblent le plus sont *car* et *sillä*. Presque toutes les variables analysées ont un effet similaire sur le choix de ces connecteurs, favorisé par les contextes subjectifs et les contextes les moins oraux.

Les connecteurs *parce que* et *koska* se ressemblent également, mais dans une moindre mesure. En français, c'est surtout le connecteur *parce que* qui est favorisé par les contextes « oraux », tandis que dans le cas du finnois, ces types de contextes favorisent surtout l'emploi de *kun*. Dans des contextes particulièrement oraux, l'équivalent le plus naturel du connecteur *parce que*

serait donc *kun* et non *koska*. Le connecteur *kun* ressemble ainsi clairement d'une part à *parce que* et d'autre part à *puisque*.

Les résultats témoignent de l'importance d'examiner simultanément i) les facteurs liés à la subjectivité, ii) ceux liés à la représentation de l'oral et iii) les différences entre les ouvrages lors de l'analyse des connecteurs causaux dans la langue littéraire. **N**

MIKKO HAULE

UNIVERSITÉ DE HELSINKI

Bibliographie

- Degand, Liesbeth – Benjamin Fagard 2012. Competing connectives in the causal domain. French *car* and *parce que*. *Journal of Pragmatics* 44 (2) : 154–168. <https://doi.org/10.1016/j.pragma.2011.12.009>
- Degand, Liesbeth – Henk Pander Maat 2003. A contrastive study of Dutch and French causal connectives on the Speaker Involvement Scale. *Usage-based approaches to Dutch. Lexicon, grammar, discourse*, eds. Verhagen, Arie – Jeroen van de Weijer. Utrecht : LOT. 175–199.
- Groupe λ-1 1975. *Car, parce que, puisque*. *Revue romane* 10 (2) : 248–280.
- Havu, Eva – Irma Hyvärinen 2013. The expression of causal relations in Finnish, French, and German: The example of *koska*. *Comparing and Contrasting Syntactic Structures*, eds. Havu, Eva – Irma Hyvärinen. Mémoires de la Société Néophilologique de Helsinki 86. Helsinki : Uusfilologinen yhdistys. 163–198.
- Herlin, Ilona 1998. *Suomen kun*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Pander Maat, Henk – Liesbeth Degand 2001. Scaling causal relations and connectives in terms of speaker involvement. *Cognitive Linguistics* 12 (3) : 211–245.
- Pit, Mirna 2003. *How to express yourself with a causal connective. Subjectivity and causal connectives in Dutch, German and French*. Amsterdam – New York : Rodopi.
- Rouayrenc, Catherine 2010. *Le français oral. 2 - L'organisation et la réalisation de l'énoncé oral*. Paris : Belin.
- Schäfer, Roland 2020. Mixed-Effects Regression Modeling. *A Practical Handbook of Corpus Linguistics*, eds. Paquot, Magali – Stefan Th. Gries. Cham : Springer. 535–561.
- Simon, Anne Catherine – Liesbeth 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. <https://doi.org/10.1017/S095926950700302X>
- Zufferey, Sandrine 2012. “*Car, parce que, puisque*” revisited: Three empirical studies on French causal connectives. *Journal of Pragmatics* 44 : 138–153. <https://doi.org/10.1016/j.pragma.2011.09.018>
- Zufferey, Sandrine 2014. Givenness, procedural meaning and connectives. The case of French *puisque*. *Journal of Pragmatics* 62 : 121–135. <https://doi.org/10.1016/j.pragma.2013.09.022>
- Zufferey, Sandrine – Willem Mak – Sara Verbrugge – Ted Sanders 2018. Usage and processing of the French causal connectives ‘*car*’ and ‘*parce que*’. *Journal of French Language Studies* 28 (1) : 85–112. <https://doi.org/10.1017/S0959269517000084>

Corpus

- Barbery, Muriel 2022 [2006]. *L'Élé-
gance du hérisson*. Paris : Galli-
mard/Folio.
- Brisac, Geneviève 1996. *Week-end de
chasse à la mère*. S.l. : Éditions de
l'Olivier.
- Dicker, Joël 2017 [2015]. *Le Livre des
Baltimore*. Paris : De Fallois/Poche.
- Gavalda, Anna 2011 [1999]. *Je voudrais
que quelqu'un m'attende quelque
part*. Paris : J'ai lu.
- Nothomb, Amélie 1999. *Stupeur et
tremblements*. Paris : Albin Michel.
- Roubaudi, Ludovic 2002. *Les Baltrin-
gues*. Paris : Le dilettante.