

Le rôle de facteurs cognitifs dans l'omission du complémenteur *que* à Montréal

Yiming Liang (Université Paris Cité), Pascal Amsili (Université Sorbonne Nouvelle) et Heather Burnett (Université Paris Cité)

L'omission du complémenteur *que* (1) est une variable sociolinguistique classique en français canadien (Sankoff *et al.*, 1971 ; Sankoff, 1980 ; Warren, 1994, parmi d'autres). Malgré l'identification d'un grand nombre des facteurs explicatifs sociaux et linguistiques, un débat subsiste sur l'indépendance des effets phonologiques et syntaxiques (Connors, 1975 ; Martineau, 1985 ; Dion, 2003). Par ailleurs, si les facteurs cognitifs de l'omission du *that* en anglais ont fait l'objet de nombreux travaux, peu d'études les ont pris en compte pour le français. En particulier, Jaeger (2010) montre que la densité informationnelle (opérationnalisée comme la probabilité qu'un verbe enchaîne une complétive) a un effet prédominant sur l'omission du *that*, devant une dizaine d'autres facteurs. Cela soutient donc l'hypothèse de l'*Uniformité de la densité informationnelle*, qui prédit que, face à deux variantes, un locuteur préférera celle qui distribue l'information de façon plus uniforme sur une phrase (Levy & Jaeger, 2007 ; Jaeger, 2010). Cette hypothèse n'a pas encore été testée dans d'autres langues comme le français.

1. Bien je pense (que) c'est : c'est important ... (locuteur 128, Corpus Montréal 84)

L'objectif de cette étude est donc double : a) utiliser l'ensemble du corpus *Montréal 84* (Thibault & Vincent, 1990) pour vérifier les résultats d'études précédentes, en particulier ceux concernant l'effet du contexte phonologique à droite et du sujet de la complétive ; b) tester deux facteurs cognitifs (la fréquence du verbe matrice et sa probabilité d'enchaîner une complétive), pour voir si l'hypothèse de l'*Uniformité de la densité informationnelle* est aussi valable en français.

Grâce à une annotation semi-automatique du corpus, nous avons extrait 5818 complétives sur dix-sept verbes fréquents, et annoté dix facteurs sociaux, linguistiques et cognitifs pour chaque observation. À l'aide d'un modèle à effets mixtes (Johnson, 2009) contrôlant la variation intrinsèque des locuteurs, notre étude montre que, conformément à la majorité des études antérieures, les effets de l'éducation, de la profession, du contexte phonologique à droite et du sujet de la complétive sont significatifs. Notamment, la hiérarchie de la sonorité dans le contexte suivant *que* a l'effet le plus important sur le choix de prononcer *que*. Conformément à l'hypothèse de l'accessibilité (Ferreira & Dell, 2000), un sujet plus accessible comme *je* et *tu* est corrélé avec un taux d'omission plus élevé qu'un autre pronom, et un sujet nominal défavorise clairement l'omission. La régression séquentielle montre également que ces deux effets ne peuvent pas se réduire l'un à l'autre. En outre, il s'avère que les effets des deux facteurs cognitifs sont tous très significatifs ($p < 0,001$) : un verbe plus fréquent favorise l'omission du *que*, tandis qu'une complétive plus prédictible tend à apparaître sans *que*, ce second effet confirme que le lien entre la réduction syntaxique et la prédictibilité d'une structure existe aussi en français.

Notre étude fournit ainsi une mise à jour sur notre compréhension de l'omission de *que*, et montre comment les facteurs grammaticaux, sociaux et cognitifs interagissent pour affecter la variation dans une langue. Plus généralement, notre étude montre que les grands corpus sociolinguistiques des années 1970 et 1980 peuvent continuer à nourrir de nouvelles découvertes pour une compréhension plus fine de la variation, notamment sur ses aspects cognitifs, 50 ans après leur construction.

Références bibliographiques

- Connors, K. (1975). L'effacement de *que* : règle syntaxique. *Recherches linguistiques à Montréal*, 4, 17-33.
- Dion, N. (2003). L'effacement du *que* en français canadien : Une étude en temps réel. Mémoire de maîtrise, Université d'Ottawa.
- Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive psychology*, 40(4), 296-340.
- Jaeger, T. F. (2010). Redundancy and reduction : Speakers manage syntactic information density. *Cognitive psychology*, 61(1), 23-62.
- Johnson, D. E. (2009). Getting off the goldvarb standard : Introducing rbrul for mixed-effects variable rule analysis. *Language and linguistics compass*, 3(1), 359-383.
- Levy, R. P., & Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. In *Advances in neural information processing systems* (pp. 849-856). B. Schölkopf, J. Platt, T. Hoffman (eds.), Cambridge, MA : MIT Press.
- Martineau, F. (1985). *Elision variable de (que) dans le parler d'Ottawa-Hull*. Thèse de doctorat non publiée, Université d'Ottawa.
- Sankoff, G. (1980). *The social life of language*. University of Pennsylvania Press.
- Sankoff, G., Sarrasin, R., & Cedergren, H. (1971). Quelques considérations sur la distribution de la variable *que* dans le français de Montréal. In *Congrès de l'association-française pour l'avancement des sciences*. Sherbrooke.
- Thibault, P., & Vincent, D. (1990). Un corpus de français parlé : Montréal 84. *Université Laval, Québec*.
- Warren, J. (1994). Plus ça change, plus c'est pareil : The case of "que" in Montreal French. *Culture*, 14(2), 39-49.