

Approche computationnelle de la transparence sémantique

Marine Wauquier

19 novembre 2021

Morphology Reading Group

Salih Akin (2006)

Comment dériver un gentilé à partir d'un toponyme ? Les potentialités signifiantes de Seine-Maritime

- Article de sociolinguistique
- Tâche de production par les locuteurs
 - Non sans rappeler la linguistique expérimentale

- Rivalité affixale dans le contexte de la formation des gentilés
 - Relativement peu étudiée
 - pour le français : Eggert (2002), Plénat (2008), Roché et Plénat (2016)
 - pour l'espagnol : García Padrón et Morera Pérez (2015), Brizuela (2017), Chesnokova *et al.* (2021)
 - pour l'anglais : Roberts (2010)
 - pour le slovène : Kuzman et Fišer (2017)
 - Principalement des critères phonologiques

- Autres critères proposées par certaines études
 - Paramètres géographiques
 - Carte des suffixes significativement présents chez Plénat (2008)
 - Distribution 'cardinale' des suffixes chez Brizuela (2017)
 - Paramètres sociolinguistiques
 - Connotation chez García Padrón et Morera Pérez (2015), Chesnokova *et al.* (2021)
 - Facteur diastratique chez Brizuela (2017)

Paramètres géographiques

- Phénomène décrit mais pas évalué comme facteur prédictif
- Travail en cours avec J. Thuilier et D. Tribout
 - Le suffixe des gentils des villes voisines peut-elle aider à prédire le suffixe du toponyme d'une ville donnée ?
- Calcul de la distance (de trajet) entre paires de villes pour un échantillon de 1500 villes
 - Annotation complète pour 200 villes seulement, soit 235000 mesures
 - Régression logistique visant à prédire si dans une paire donnée les deux gentils portent le même suffixe ou non sur la base de la distance
 - Facteur significatif, confirmant que la probabilité d'avoir le même suffixe diminue à mesure que la distance augmente
 - Mais précision pas meilleure que le hasard (accuracy de 0.5, AUC de 0.5)

- Paramètre décrit de façon impressionniste
 - Description de gentilés individuels à partir de descriptions lexicographiques ou d'introspection par des locuteurs isolés, sans quantification réelle
- La valeur axiologique associé à un affixe (prestigieux, populaire...) participe-t-il de la sélection du suffixe ?
- Comment évaluer ce paramètre ?
 - Travail de Salih Akin (2006) comme une première piste

- Appel à proposition auprès des habitants de Seine-Maritime (Haute-Normandie) pour la dénomination des habitants du département (2005)
 - Recueil sur un site internet
 - 43% des 40 000 habitants ont participé (soit environ 17 200)
 - 614 propositions, dont
 - 143 'humoristiques' (*les marins se-noient*)
 - 163 basées sur un seul des éléments du toponyme (*Seinois*)
 - 308 basées sur les deux éléments du toponyme (*Seinomarín*)
 - Les 3 gentils les plus proposés ont fait l'objet d'une deuxième évaluation
 - Le plus fréquent était *Normands* (73 occurrences), bien devant *Seinomarín* (26 occurrences)

- Sur les 308 propositions analysées
 - 22 cas de dérivation simple (*Seine-maritimes*)
 - 13 cas de dérivation complexe (*Seinai-maritimes*)
 - 140 cas de dérivation allomorphique (*Séquano-marins*)
 - 20 cas de composition (*Seinemaritimes*)
 - 113 cas de permutation allomorphique (*Mariseinois*)
 - Fréquence des tokens, non des types
 - Par exemple, pour la dérivation complexe, il n'y a en réalité que 5 propositions différentes

- Ces chiffres viennent aussi avec des commentaires des locuteurs
 - Considérations esthétiques
 - *Seine-Maritains & Seine-maritaines* > "Voilà un nom qui sonnerait bien!"
 - Considérations 'cognitives'
 - *Les Marinseinois (es)* > "Pour retenir un nom, il doit être facile à prononcer..."
 - Considérations 'morphologiques'
 - *Séquano-marins* > "... étymologiquement correct..."
 - Considérations 'analogiques'
 - *Séquanomaritains* > "[...] on parle de séquanodyonisiens[...] et] d'altoséquanais..."
 - Considérations axiologiques
 - *Séquanomarins* > "...cela est précieux et pédant.."

- Pas de quantification de ces commentaires
 - Question de l'annotation et de l'automatisation de l'analyse
- Pas de croisement réel avec les propriétés morpho-phonologiques
 - Ces considérations favorisent-elles une construction ou un suffixe donné ? L'effet est-il lié à des propriétés phonologiques particulières ?
- Pas de contrôle sur "l'expérience"
 - Pas d'informations sur les locuteurs (âge, catégorie socio-professionnelle, éducation)
 - Un seul toponyme
 - Les données sont-elles quantitativement suffisantes pour être statistiquement significatives ?

- Akin, S. (2006). Comment dériver un gentilé à partir d'un toponyme? Les potentialités signifiantes de Seine-Maritime. *Cahiers de Sociolinguistique n 11*, «Noms propres, dynamiques identitaires et sociolinguistique», 63-80.
- Brizuela, S. L. E. (2017). 'Sufijos gentilicios en el español de México. rivalidad y restricciones de aplicabilidad'. *Anuario de letras. Lingüística y filología*, (5), 67–90.
- Chesnokova, O.S., Radović, Marija & Kotenyatkina, I.B. (2021). Spanish South American and Brazilian Demonyms: Morphosyntactic Structure and Axiological Values. *RUDN Journal of Language Studies, Semiotics and Semantics*, 12(3), 576—596.
doi:10.22363/2313-2299-2021-12-3-576-596

- Eggert, E. (2002). La dérivation toponymes-gentilés en français: mise en évidence des régularités utilisables dans le cadre d'un traitement automatique (Thèse de doctorat, Tours).
- Kuzman, T. & Fišer D. (2017). Corpus-Based Analysis of Demonyms in Slovene Twitter. In *Proceedings of the 5th Conference on CMC and Social Media Corpora for the Humanities (cmccorpora17)*, 30–34.
- Padrón, D. G., & Pérez, M. M. (2015). Gentilicios y lexicografía. *Onomázein*, (31), 81-98.
- Plénat, M. (2008). Quelques considérations sur la formation des gentilés. B. Fradin (éd.), *La raison morphologique. Hommage à la mémoire de Danièle Corbin*. John Benjamins, 155–174.

- Roberts, M. (2017). The semantics of demonyms in English. In Zhengdao Ye (Ed.) *The Semantics of Nouns* (pp. 205-220). Oxford Scholarship Online. doi:10.1093/oso/9780198736721.003.0008
- Roché, M. & Plénat M. (2016). De l'harmonie dans la construction des mots français. *Actes du CMLF 2016*.