Affix rivalry in French demonyms: an experimental approach

Marie Huygevelde *Ridvan Kayirici* *Olivier Bonami* *Barbara Hemforth*
Université Paris Cité, Laboratoire de linguistique formelle, CNRS

This study reports on two acceptability experiments on the influence of base phonology on speakers' preferences in the choice of a suffix forming a demonym (e.g. *nancéen*, 'from Nancy') from a toponym (e.g. *Nancy*). The results are broadly consistent with previous findings by Thuilier et al.'s (2023), but bring out interesting contrasts between tendencies in the established lexicon and speakers' preferences in novel formations.

1 Motivation

Ever since Aronoff (1976) coined the term, situations of rivalry, where multiple word formation processes are available to convey the same meaning, have been a major focus of attention for descriptive and theoretical morphology. Much progress in this area has been made possible by the systematic exploration of large lexical databases (e.g. Plag 1999; Lindsay & Aronoff 2013) and the application to these of statistical modelling (Baayen et al., 2013) and computational simulations (Arndt-Lappe, 2014).

One important limitation of this line of work is the inherent heterogeneity of the data found in the established lexicon, which contains words coined over centuries by speakers whose linguistic experience may have differed significantly—if anything because each new coinage may influence later formations. While some authors attempt to alleviate this problem by focusing on recent formations (Plag, 1999) or explicitly taking into account diachronic variation (Lindsay & Aronoff, 2013; Arndt-Lappe, 2014), a more direct (but less frequent) approach is to conduct behavioral experiments probing the preferences of speakers facing the task of producing, interpreting or judging rival formations (Anshen & Aronoff, 1981; Makarova, 2016).

This abstract reports on such behavioral experiments dedicated to the formation of demonyms in French. Demonyms are a particularly promising testbed for such a study: on the one hand, the semantic relationship between a toponym and its demonym is much more stable than what is found for other instances of word formation; hence we need not worry about sibylline decisions as to whether a particular pair of words instantiates the relevant morphosemantic contrast. On the other hand, rivalry is very prevalent, with at least four highly productive suffixes in French: *-ais* (see *Marseillais* from *Marseille*), *-ois* (*Lillois*, from *Lille*), *-éen* (*Nancéen*, from *Nancy*) and *-ien* (*Parisien*, from *Paris*). It is also very well documented, in no small part thanks to Thuilier et al.'s (2023) recent thorough study of more than 2,000 established toponyms. Thuilier et al.'s study serves as the starting point for the present research: our goal is to assess the extent to which speaker preferences in an experimental setting track the tendencies observed by Thuilier et al. in the established lexicon. We focus more specifically on the influence of phonological properties of the base on the choice of a suffix when speakers are faced with a novel, unknown toponymic base.

2 Methods

To explore the impact of the phonological makeup of toponyms in demonym formations, we ran two experiments investigating preferential choices with disyllabic toponymic bases with a final consonant (Experiment 1) or nasal vowel (Experiment 2). We limited our attention to the

four most productive suffixes *-ais*, *-ois*, *-éen* and *-ien* that have been highlighted in the literature (Eggert, 2002; Plénat, 2008; Thuilier et al., 2023)).

For the first experiment, we created 24 toponyms with a final segment representative of each of the 5 categories singled out by Thuilier et al..[1]. Our toponyms ended with the bilabial plosive /p/ (Nabope), the palatal approximant /j/(Naboye), the alveolar fricative /s/ (Nabosse), the post-alveolar fricative /ʃ/ (Naboche), and the nasal bilabial /m/ (Nabome) (standing for the 'other segment' category in Thuilier et al.'s study).

For the second experiment, we explored the impact of vowel backness on nasal last segments based on tongue positioning. We created 15 items with proper names manipulating the backness of both nuclei in a disyllabic word ending in a nasal vowel. Thus, our vocalic feature pairs consisted of /a/ and /o/ for the first position, /ɛ̃/ and /ɔ̃/ for the second position. We made use of the methodology proposed by Lohmann (2017) to control for the backness score of the base toponyms, we had thus vocalic nasal feature pairs with low /a/ /ɛ̃/, high /o/ /ɔ̃/, and medium /a/ /ɔ̃/ and /o/ /ɛ̃/ backness scores.

Our experiments took place online, using a local installation at LLF of Alex Drummond's IBEX software. Each experimental session started with written instructions, a short anonymized questionnaire, and a brief practice session.

During each trial, a base toponym was presented to the participants and they were asked to choose between four possible demonyms formed using each of the four main suffixes with a given toponym. Participants saw all conditions but each item in only one condition following a Latin square design. In total, an experimental session consisted of 24 trials with consonant-final toponyms, 15 trials with nasal vowel-final toponyms, and 11 trials of fillers based on real-world toponyms (i.e. Marseille). Items and fillers were presented in random order.

We recruited 71 participants from online academic and social networks. Two bilingual participants were excluded from our analyses (n = 69, mean age = 41). All participants took the experiment voluntarily with no compensation and gave their consent for the usage of their anonymized data for scientific purposes.

3 Results

All data were analyzed with Generalized linear mixed models (binomial family), in R (R version 4.1.1, glmer, lme4 version 1.1-31), with random intercepts for participants and item.[2]

Across the two experiments, the dependent variable was the choice between four categorically distributed nominal values that are constituted by the demonyms suffixes *-ais*, *-ois*, *-éen* and *-ien*. The phonological final segment constraints constituted the independent variable (5 conditions) for the consonantic demonym formations, and the backness feature as well as the syllabic position were the independent variables for the vocalic nasal demonym formation.

Figure 1 shows that all suffixes are used in all conditions but with clearly varying frequencies. We observed a preference for the suffix *-ien* for the bilabial plosive, alveolar fricative and palatal approximant conditions, while the post-alveolar fricative condition was more associated with the suffix *-ois*. For the consonantic nasal bilabial condition, we observed a tendency towards the suffixes *-éen* and *-ois* in a similarly weighted manner accounting for more than half of the preferences combined.

For statistical comparisons, the palatal fricative /ʃ/ served as the reference category. The suffix *-ois* was chosen significantly more often for the palatal fricative than for the bilabial plosive ($p < .001$) as well as the palatal approximant ($p < .001$), followed by /m/ ($p < .04$) and

[1] One item of each experiment had to be excluded for technical reasons

[2] Random slopes were not included because of convergence failures.

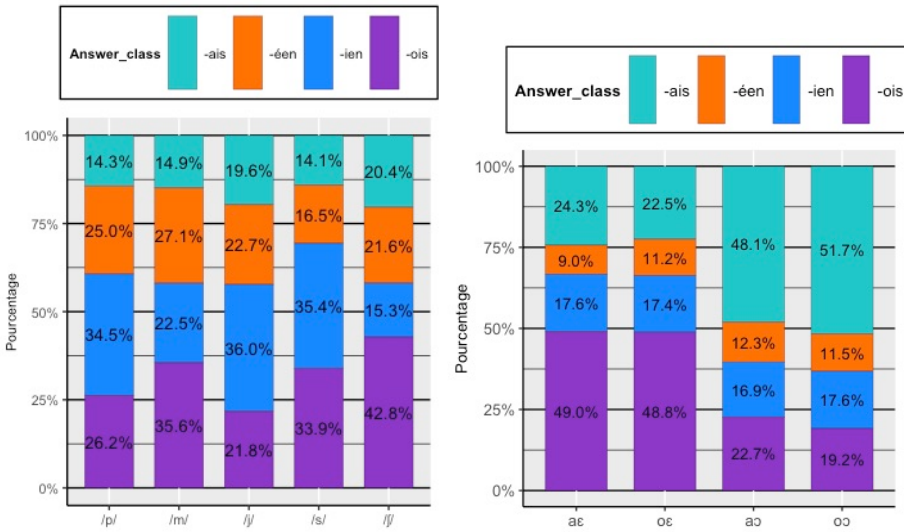

Figure 1: Suffix choices for Experiments 1 and 2

/s/ ($p < .01$). The suffix *-ien* was chosen least frequently for the palatal fricative /ʃ/ compared to all other phonological conditions: /s/ ($p < .001$), /j/ ($p < .001$) /p/ ($p < .001$), /m/ ($p < .02$). The suffix *-ais* was chosen more frequently in the palatal fricative than in the alveolar fricative ($p < .04$) and bilabial plosive ($p < .05$) conditions. For the suffix *-éen*, we did not observe any significant differences across the five conditions.

The second experiment showed significant preferences for the suffix *-ais* for the toponyms with a back last segment ($p < .001$), and for *-ois* for the toponyms with a front last segment ($p < .001$). We did not find any significant differences concerning the first syllable position or the backness score. We did not observe any significant effects for choices of *-éen* and *-ien*.

4 Discussion

Thuilier et al.'s study established an overall preference for the suffix *-ois* after consonants. In general, we observed similar results aligning with this tendency, in particular for the alveolar and post-alveolar fricative conditions.

For the alveolar fricative condition, Thuilier et al. documented 49.6% of use for *-ien*, and 37.6% for *-ois*. We found a similar tendency, with 35.4% for *-ien*, and 33.9% for *-ois*. These results might be indicative of a phonological motivation caused by the alveolar fricative feature, which is also captured in the corpora. For the post-alveolar fricative condition, *-ois* was chosen 42.8% of the time by the participants. While this made up most of their choices, the preference is less sharp in our experiment compared to its prevalence in the established lexicon where it reaches 69.5% in the "other fricative" condition.

For the bases ending with a bilabial plosive or a palatal approximant, we found contrasting results between data from experiments and corpora. Whereas 47.2% of *-ois* was found after plosives in the corpus, it accounted for only 26.2% of the choices in our experiment. Likewise, while *-ois* was the preferred option in the established lexicon after an approximant with 57.2% (vs. 20.5% for *-ien*), in our experiment *-ien* was preferred at 36% (only 21.8% for *-ois*).

In the nasal bilabial condition, we found both similarities and discrepancies with what Thuilier et al. observed in the established lexicon. Our participants showed the same preference for *-ois* documented by Thuilier et al. for the distribution of the suffixes with respect to nasal segments, although less strongly so (48.4% vs. 35.6%). When we consider the bilabial

nasal condition from 'other segment' perspective, where *-ois* rate drops to 27.9% (after suffixes *-ien* and *-ais* 32.3% and 28.2% respectively) in the established lexicon, we observe a slight discrepancy. This might be due to the fact that the 'other' lumps together a diverse set of segments in their study, while in our experiments, the consonant is specified as /m/. However, our results are surprising in another dimension. In the affix rivalry literature, it has been argued that dissimilative tendencies disfavor *-éen* and *-ien* after a nasal segment, be it a vowel or consonant. We found no evidence for this, as *-éen* and *-ien* made up half of the choices after /m/.

Thuilier et al.'s data suggest that the nasality of the vowel plays an important role in the choice of a suffix. Our findings suggest that demonym formations with nasal vowels might be strongly motivated by phonological constraints. Keeping in mind that our vocalic conditions involve also nasality, findings coming from our second experiment align with suggestions from the literature on vocalic qualities as indicated by Thuilier et al., and highlighted by Eggert, Plénat, and Roché & Plénat (2016): we see that the suffix *-ais* is disfavored after bases with last segment front vowel, whereas the suffix *-ois* disfavors bases with last segment back vowel.

One main conclusion of this study is that phonological constraints seem to play a distinct role for affix rivalry dynamics regarding French demonym formations. The discrepancies between observations from our controlled experiments and the results from the Thuilier et al. study might be due to factors beyond phonological motivations that may play a role for "real" demonyms but also to specifics of the experimental task.

References

Anshen, Frank & Mark Aronoff. 1981. Morphological productivity and phonological transparency. *Canadian Journal of Linguistics* 26. 63–72.

Arndt-Lappe, Sabine. 2014. Analogy in suffix rivalry: the case of English *-ity* and *-ness*. *English Language and Linguistics* 18. 497–548.

Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge: MIT Press.

Baayen, R. Harald, Anna Endresen, Laura A. Janda, Anastasia Makarova & Tore Nesset. 2013. Making choices in russian: pros and cons of statistical methods for rival forms. *Russian Linguistics* 37(3). 253–291.

Eggert, Elmar. 2002. *La dérivation toponymes-gentilés en français : mise en évidence des régularités utilisables dans le cadre d'un traitement automatique*: Tours, Phd thesis.

Lindsay, Mark & Mark Aronoff. 2013. Natural selection in self-organizing morphological systems. In Nabil Hathout, Fabio Montermini & Jesse Tseng (eds.), *Morphology in toulouse: Selected proceedings of décembrettes 7*, 133–153. Munich: Lincom Europa.

Lohmann, Arne. 2017. Phonological properties of word classes and directionality in conversion. *Word Structure* 10(2). 204–234.

Makarova, Anastasia. 2016. Variation in Russian verbal prefixes and psycholinguistic experiments. In Tanja Anstatt, Anja Gattnar & Christina Clasmeier (eds.), *Slavic languages in psycholinguistics*, 113–133. Tübingen: Narr Francke Attempto Verlag.

Plag, Ingo. 1999. *Morphological productivity*. Berlin: Mouton de Gruyter.

Plénat, Marc. 2008. Quelques considérations sur la formation des gentilés. In *La raison morphologique*, 155–174. John Benjamins.

Roché, Michel & Marc Plénat. 2016. De l'harmonie dans la construction des mots français. *SHS Web of Conferences* 27. 08003.

Thuilier, Juliette, Delphine Tribout & Marine Wauquier. 2023. Affixal rivalry in french demonym formation: The role of linguistic and non-linguistic parameters. *Word Structure* 16(1).