

Affix rivalry in French demonyms: an experimental approach

Rivalité affixale des gentils en français : une approche expérimentale

Marie Huygevelde

Université Paris Cité, Laboratoire de Linguistique Formelle, CNRS, Paris, France

mariehuygevelde@gmail.com

Ridvan Kayirici

Université de Lorraine, CNRS, IMoPA, F-54000 Nancy, France

ridvan.kayirici@cnrs.fr

Olivier Bonami

Université Paris Cité, Laboratoire de Linguistique Formelle, CNRS, Paris, France

olivier.bonami@u-paris.fr

Barbara Hemforth

Université Paris Cité, Laboratoire de Linguistique Formelle, CNRS, Paris, France

barbara.hemforth@linguist.univ-paris-diderot.fr

Abstract

As many other languages, French uses a variety of rival suffixes to form names of inhabitants, or demonyms, out of place names, or toponyms; most prominently the suffixes *-ais* (*Marseille* > *Marseillais*), *-ois* (*Lille* > *Lillois*), *-éen* (*Nancy* > *Nancéen*) and *-ien* (*Paris* > *Parisien*). Existing literature on the topic has focused on documenting phonological and geographical factors influencing the choice of affix on the basis of examination of the established lexicon. This paper reports on an experimental study probing the preferences of speakers in nonce formations, focusing on the influence of phonological factors. Experimental results provide evidence that speakers are sensitive to phonological properties of the base when coining a demonym, but differ in subtle and interesting ways from what can be concluded from the established lexicon.

Beyond studying the formation of French demonyms, this paper highlights the usefulness of experiments as complementary to the examination of the established lexicon in the study of morphological rivalry.

Keywords: experimental morphology, phonology, demonyms, affix rivalry

Résumé

Le français, comme beaucoup d'autres langues, utilise plusieurs suffixes rivaux pour former des noms d'habitants, ou gentils, à partir de noms de lieux, ou toponymes ; les plus fréquents sont *-ais* (*Marseille* > *Marseillais*), *-ois* (*Lille* > *Lillois*), *-éen* (*Nancy* > *Nancéen*) et *-ien* (*Paris* > *Parisien*). Les travaux sur le sujet ont documenté des contraintes phonologiques et géographiques qui influencent le choix de l'afixe, sur la base d'un examen du lexique établi. Cet article rend compte d'une étude expérimentale examinant les préférences des locuteurs et locutrices dans les formations nouvelles, en se concentrant sur l'influence de facteurs phonologiques. Les résultats expérimentaux suggèrent que les locuteurs et locutrices ont une sensibilité aux propriétés phonologiques de la base quand ils forment un gentilé, mais différent de manière subtile et intéressante des conclusions qui peuvent être atteintes en examinant le lexique existant.

Au-delà de la question de la formation des gentils en français, cet article met en lumière l'utilité de travaux expérimentaux comme complément à l'examen du lexique établi dans l'étude de la rivalité morphologique.

Mots-clés : morphologie expérimentale, phonologie, gentils, rivalité affixale

1. Introduction

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 modeling (Baayen, Endresen, Janda, Makarova & Nessel, 2013; Bonami & Thuilier, 2019) and computational simulations (Arndt-Lappe, 2014; Guzmán-Naranjo & Bonami, 2023).

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 (a.o. 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; Romaine, 1983; Makarova, 2016; Schirakowski, 2020; Copot & Bonami, 2024).

This paper reports on such behavioral experiments dedicated to the formation of demonyms (names of inhabitants) from toponyms (names of locations) 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 regular 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, as illustrated in Table 1. It is also very well documented, in no small part thanks to Thuilier, Tribout and Wauquier's (2023) recent thorough study of more than 2,000 established demonyms.

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. Section 2 of this paper outlines Thuilier *et al.*'s study, and presents a series of predictions that we aim to test experimentally. Section 3 describes the methods, and Section 4 the results of the experiment. We close in Section 5 with a discussion.

	Example	
Suffix	Toponym	Demonym
-ais	<i>Marseille</i>	<i>Marseillais</i>
-ois	<i>Lille</i>	<i>Lillois</i>
-éen	<i>Nancy</i>	<i>Nancéen</i>
-ien	<i>Paris</i>	<i>Parisien</i>

Table 1. The four most productive suffixes for forming demonyms in French.

2. Background and predictions

2.1. Thuilier *et al.*'s (2023) seminal study

In a recent paper, Thuilier *et al.* (2023) report on a statistical analysis of 2,218 French demonyms extracted from the ProLex database (Tran & Maurel, 2006) and ending in one of the 4 suffixes illustrated in Table 1. The main point of their study is to explore the extent to which the choice of a suffix for a demonym is influenced by two kinds of predictors: various morphological and phonological properties of the formation, and geographical properties of the toponym's referent (size of location and preferred suffixes in a geographical neighborhood). They explore this by training random forest classifiers to predict suffix choice from various combinations of predictors. This statistical method allows them to build a predictive model and to evaluate the relative importance of these factors in predicting the distribution of suffixes in the dataset. Overall model performance is assessed by examining classification accuracy for aggregated predictions after 10-fold cross-validation, and variable importance is assessed using the built-in function of the caret R package.

Thuilier *et al.*'s modeling efforts lead to three series of conclusions. First, a combination of 10 linguistic predictors leads to sizable classification success, with an accuracy over the whole dataset of 0.57, on a dataset where the larger class makes up 0.39 of the data. Second, the addition of geographical predictors leads to a more accurate model: despite a smaller dataset of 1,435 items for which geographical information was available, a second model taking into account geographical predictors reaches an accuracy of 0.62. Third, both models assign more or less the same relative importance to the linguistic predictors. Focusing on the first model, the following variables were found to play a role, in order of decreasing importance:

- (1) a. Whether the last segment of the base toponym is a front oral vowel (*Paris* /paʁi/ > *Parisien*).
- b. Whether the last segment of the base toponym is a nasal vowel (*Berlin* /berlɛ̃/ > *Berlinois*).
- c. The “backness score” of the base toponym, computed by averaging the backness of all vowels in the word (e.g., *Ivry* /ivʁi/ gets a score of 1, *Fitou* /fitu/ a score of 2, and *Toulouse* /tuluz/ a score of 3).
- d. The length of the base toponym (number of syllables).
- e. Whether the suffix attaches to exactly the base stem (*Toulouse* /tuluz/ > *Toulousain*) or an allomorph (*Sochaux* /soʃo/ > *Sochalien*).
- f. Whether the relationship between the base and demonym is opaque (*Châteaufort* /ʃatofɔʁ/ > *Castelfort-ain*).
- g. Whether the base toponym ends in an approximant (*Lille* /lil/ > *Lillois*).
- h. Whether the base toponym ends in an alveolar fricative (*Alsace* /alzas/ > *Alsacien*).
- i. Whether the base toponym ends in a nasal consonant (*Rennes* /ʁɛn/ > *Rennais*).
- j. Whether the base toponym ends in a fricative consonant that is not alveolar (*Orange* /oʁɑ̃ʒ/ > *Orangeois*).
- k. Whether the base ends in a plosive (*Sète* /set/ > *Sétois*).

As suggestive as these results are, there are limitations inherent in Thuilier *et al.*'s methods. First, because their study is based on examining the established lexicon, it does not probe directly the way rivalry is organized synchronically: the formations documented in their database were coined by speakers belonging to different linguistic communities at different times, some quite distant in the past. While synchronic organization certainly is influenced by statistical tendencies in the established lexicon, it is entirely possible that new formations do not track these tendencies exactly, and are influenced by other factors. Second, there is inherent complexity in the data that makes it challenging to examine the interplay of all relevant variables. For instance, Thuilier *et al.* set out to examine the role played by front vs. back vowels in the toponym. To do this, they use a holistic backness score for the word. This is well-motivated by the fact that toponyms vary in length, but precludes examining the contribution of the vowels in individual syllables and their possible interactions.

In this paper we address both limitations by relying on behavioral experiments where participants are asked to assess which suffixed demonym best fits a nonce toponymic base. This allows us, on the one hand, to probe directly linguistic intuitions, and, on the other hand, to manipulate the variables of interest, creating balanced datasets on which regression models can be deployed to study both the effects of individual predictors and their interactions. To keep the number of variables manipulated and hence the duration of the experiments manageable, we restrict attention to a subset of the phonological variables examined by Thuilier *et al.* The next section details what these variables are, and presents predictions based on the Thuilier *et al.* study that our experiments will put to the test.

2.2. Predictions

Among the phonological variables ranked as impactful by Thuillier *et al.* in determining demonym formations in the established lexicon, we focused our investigation on the following.

- (2) a. Various consonant types as last segments.
- b. Nasality.
- c. Vocalic quality of the last segment.
- d. The backness of vowels.

Nasality, vocalic quality and backness score were the most predictive phonological variables in Thuillier *et al.*'s study; it was hence obvious that we should take them into account. Regarding consonantal last segments as predictors, Thuillier *et al.* grouped segments into ad-hoc phonological classes which, as preliminary analysis suggested, would have predictive values: plosives, approximants, alveolar fricatives, other fricatives, and other segments, including all other consonants (e.g., nasals) as well as all vowels.

According to Thuillier *et al.*'s study, bases ending in a consonant generally favor the suffix *-ois* (52% of cases, vs. 25% for bases ending in a vowel), but this preference is modulated depending on the consonant. This suffix is the preferred choice after plosives (47% of the cases), approximants (57%), and even more strongly after all fricatives except alveolars (70%). Alveolar fricatives stand out in having a preference for *-ien* (50%), with the proportion of *-ois* dropping to 38%.

On variables related to nasality and vocalic quality of the last segments, the authors demonstrated that these two features modulated the proportion of demonym suffixes. The authors highlighted that final nasality disfavored suffixes *-éen* (1% vs. 8% in the overall dataset) and *-ien* (8% vs. 29%). In addition, they showed that back vowels favored the selection of the suffix *-ais*. In parallel, bases ending in a front vowel favor the suffixes *-ien* (47% vs. 29% in the overall dataset) and *-éen* (18% vs. 8%). Which suffixes would be favored in the case of toponyms ending by front vocalic nasals such as *-in* (/ɛ̃/)? To access more fine-grained observations, we intend to investigate the interaction of these two variables in demonym formations. Using nonce toponyms whose final nasal is either front or back, we should be able to assess whether nasality or vocalic quality plays a stronger role in influencing the choice of a suffix.

Finally, Thuillier *et al.* examined the influence of the overall backness of the vowels across the whole base. Leveraging methodology from Lohmann (2017), the authors attributed an overall backness score to toponymic bases by assigning each vowel a score of 1 (front) to

3 (back) and averaging over all syllables; for instance, *Aurillac* [oʁijak] gets a score of $(3 + 1 + 1)/3 = 1.7$. They then observed that bases with a high backness score favor *-ais*.

Overall, where the last segment influences preferences for a suffix, this is interpreted by Thuilier *et al.* (2023) as a dissimilation effect, in the spirit of Roché and Plénat (2016). According to these authors, there is a preference of choosing a suffix not leading to similar segments occurring at the end of the derivational stem and in the suffix, both in the feminine and in the masculine. For instance, the dispreference for *-éen* and *-ien* after nasals is hypothesized to be due to the fact that the suffixes contain a nasal themselves. Likewise, *-ois* would be dispreferred after alveolar fricatives because the feminine form of the suffix *-oise* ends in a segment of that class.

Under this perspective, we can suggest that dissimilative constraints in demonym formations relate directly to phonetic and acoustic realizations, how demonyms “sound” with regard to their toponyms. This concern was reported in Akin (2006) along with axiological considerations. Are preferences for demonym formation driven by the phonology of toponyms? Do speaker intuitions reflect tendencies and variables weight as reported in the established lexicon?

Linked to the direct experience of the speakers, we hypothesize that our experimental results should reflect the main tendencies identified by Thuilier *et al.* (2023). Nonce toponyms ending in a consonant should show an overall preference for the suffix *-ois* while different preferences should emerge when the last segment is an alveolar fricative. Concerning final nasal vowels, demonym suffix distribution should be impacted by the backness of the vowel. We expect that back nasals trigger more preferences in *-ais*. Regarding the average backness of vowels in toponyms, we predict an impact as long as it does not conflict with variables reported as more impactful, such as nasality of the last segment and its localized backness on the last syllable.

3. Methods

Our experiments are forced choice experiments where participants are asked to pick one of four suffixes as most appropriate to coin a demonym from a nonce toponym.

3.1. Design

To explore the impact of phonological variables on demonym formation, we ran two experiments investigating preferential choices on disyllabic toponymic bases with a final consonant (Experiment 1) or a nasal vowel (Experiment 2). To control for non-phonological

variables, we created as experimental items nonce toponyms that contrast in the phonological features of their last segment.

As Thuilier *et al.* (2023) show, the length of the toponym seems to play a role in demonym formation. For instance, longer toponyms (four syllables) present numerous attestations of demonyms in *-ien*, e.g., *mauritanien* from *Mauritanie*. We decided however to neutralize that variable and to focus on two-syllable toponyms, which is the most frequent case (52%) in Thuilier *et al.*'s dataset.

For each nonce toponym, we derived four corresponding demonyms and focused on 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).

3.1.1. Experiment 1 (Consonantal Last Segment)

For the first experiment on consonants, we decided to concentrate on five categories from Thuilier *et al.* classifying different toponymic endings based on their phonology, i.e. plosive, approximant, alveolar fricative, other fricative, nasal.¹ To keep our experimental design simple, rather than sampling from each of the sets, we selected one possible final consonant for each category among the ones referred in Thuilier *et al.* Regarding their phonology, we considered them as suitable representatives of their categories, covering across conditions various places of articulations: labial, alveolar, post-alveolar, palatal. We then created 24 sets of 5 nonce toponyms ending in each of the 5 crucial consonants (for a total of 120 disyllabic items), and chose an appropriate unambiguous orthography. This led us to add an orthographic *-e* in most consonant conditions, generally not pronounced word-finally in French but indicative of the pronunciation of the preceding consonant segment. In the case of the palatal approximant /j/, its orthographic representation was either *-ye* or *-i* depending on the adjacent vowel to preserve orthographic norms. This allowed us to ensure that the participants were including the final consonant segments in their phonetic and auditory representation of these toponyms. Table 2 presents the chosen consonants as well as a series of sample items.

¹ Thuilier *et al.* only compare directly the other four categories of consonants to a category of “other segment”, consisting of vowels (oral and nasal) and nasal consonants. As results for nasal consonants are reported elsewhere in the paper, we were able to complement the results of their study in the direction of a comparison of all five consonant categories.

Category	Chosen segment	Sample item
plosive	/p/	<i>Nabo^e</i>
approximant	/j/	<i>Naboⁱ</i>
alveolar fricative	/s/	<i>Nabosse</i>
other fricative	/ʃ/	<i>Naboche</i>
nasal	/m/	<i>Nabome</i>

Table 2. Final consonants used in Experiment 1.

3.1.2. Experiment 2 (Vocalic Last Segment)

The second experiment focused on words ending in a nasal vowel, and explored the impact of vowel backness on nasal last segments based on tongue positioning. Following Thuillier *et al.*'s observation that backness plays a role in demonym distribution, we created 15 disyllabic items manipulating the backness of both nuclei in our disyllabic toponyms ending in nasal vowels.

Thus, our vocalic feature pairs consisted of vowels /a/ and /o/ in the first syllable, and nasal vowels /ẽ/ and /õ/ in the second syllable. We controlled for the overall vocalic backness of our base toponyms making use of the methodology proposed by Lohmann (2017). We thus created 15 disyllabic items, characterized by pairings of nuclei with low /a/ and /ẽ/, high /o/ /õ/, and medium /a/ /õ/ and /o/ /ẽ/ backness scores. Table 3 shows some examples.

First vowel		Second vowel		Sample item
Category	Segment	Category	Segment	
oral front	/a/	nasal front	/ẽ/	<i>Fapin</i>
oral front	/a/	nasal back	/õ/	<i>Fapon</i>
oral back	/o/	nasal front	/ẽ/	<i>Fopin</i>
oral back	/o/	nasal back	/õ/	<i>Fopon</i>

Table 3. Vowel patterns for items in Experiment 2.

For this experiment, the independent variables were based on the position within the toponym (first syllable, second syllable) each with the conditions “Back” for higher backness scores and “Front” for lower backness scores.

3.2. Procedure

Our experiments took place online, using a local installation at Laboratoire de Linguistique Formelle (LLF, Paris) of Alex Drummond's IBEX software (maintained by Achille Falaise). Each experimental session started with written instructions, a short anonymized questionnaire, and a brief practice session. The instructions about the experiment were explicitly indicating to the participants that their general knowledge was not assessed. Facing names of cities that they probably would not know, we asked them to follow their intuition on which inhabitant names would be the most suitable option.

During each trial, participants would thus be presented a base toponym and asked to choose between four possible demonyms derived from the given toponym and associated with the four main suffixes *-ais*, *-éen*, *-ien*, *-ois* (see Figure 1). Items were distributed across lists such that participants saw each item in only one condition. 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*). The order of experimental items, toponyms and corresponding demonyms, as well as fillers were randomized individually for each participant.

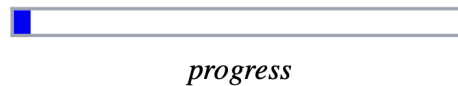


Figure 1. Screenshot of the experiment.

3.3. Participants

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, gave their consent for the usage of their anonymized data for scientific purposes with the possibility to withdraw within 7 days.

Over 69 participants, we obtained a total of 1,656 observations for the Experiment 1 on consonant last segments, and 1,035 observations for the Experiment 2 on nasal vocalic last segments.

4. Results

4.1. Experiment 1: Consonant final segments

Our first analysis focuses on the general distribution of suffixes across all conditions. We used a multinomial model using the `mblogit` function from the `mclogit` R package (Elff, 2024) to analyze our data, as well as `ggplot2` (Wickham, 2016) and `sjplot` (Lüdtke, 2024) for data visualization. The model included suffix choice as the dependent variable and the type of consonant final segment as predictor as well as random intercepts for participants and items. Including random slopes for participants or items led to convergence failure. We will therefore only report results from intercept models.

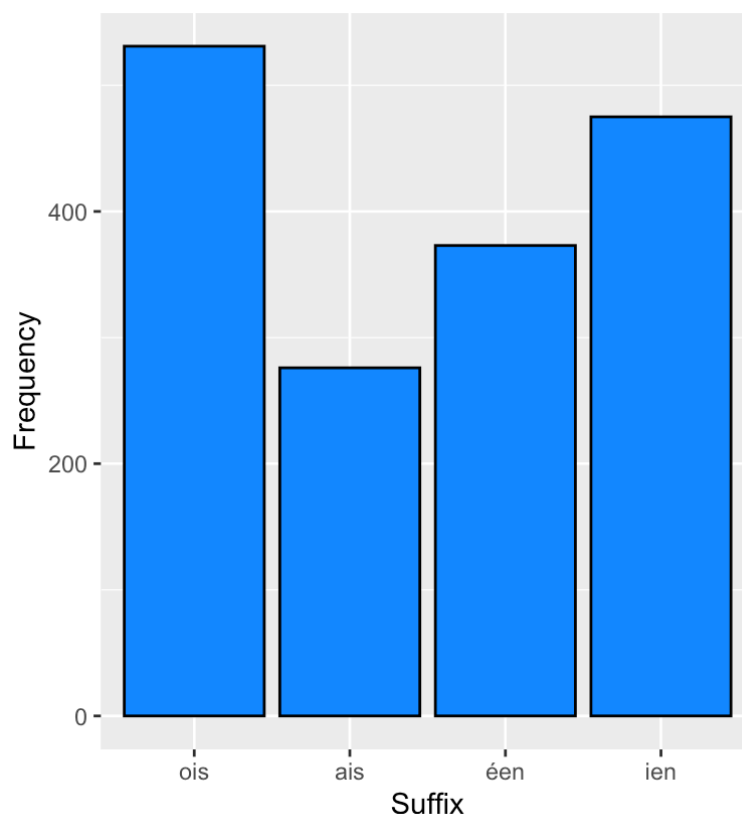


Figure 2. General frequency of suffixes after consonant final segments.

The most frequent choice of final consonant was *-ois* with 531 choices, closely followed by *-ien* (non-significant difference: Odds Ratios = .89, CI = 0.72 – 1.10, $p = .272$) with 475 choices.

The suffixes *-éen*, with 373 choices (Odds Ratios = .67, CI = 0.53 – 0.85, $p < .001$) and *-ais*, with 276 choices (Odds Ratios = .51, CI = 0.41 – 0.63, $p < .001$, see Figure 2) were selected significantly less often. These results are broadly in line with those of Thuillier *et al.* (2023), who found *-ois* to be the most frequent choice after consonants. One surprising result is the high proportion of *-éen*, which was the most dispreferred choice in Thuillier *et al.*'s study.

Our central question concerns the distribution of suffixes dependent on the type of consonants (Figure 3). We used *-ois* again as our reference category for the answer classes and /j/ as reference category for the final consonant (both the most frequent cases in Thuillier *et al.*). Based on their frequency, we take these categories as default cases. However, using different reference categories does not change the general picture.

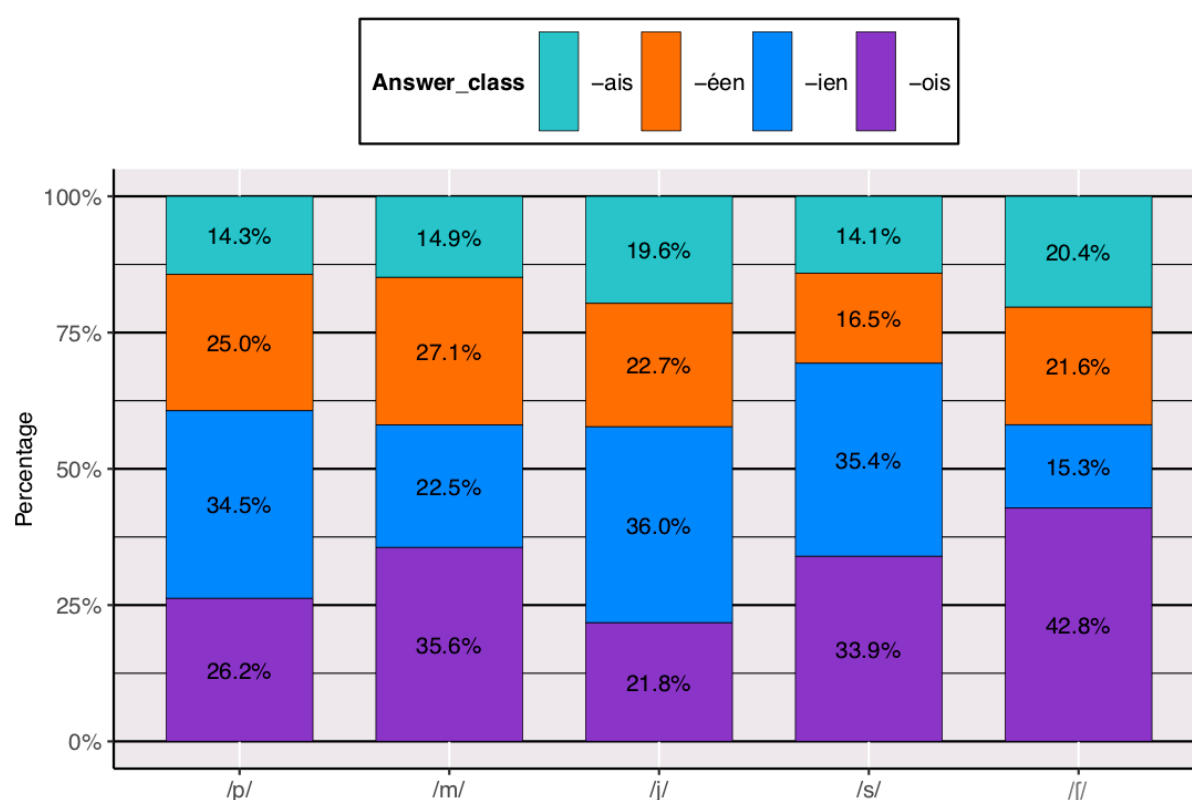


Figure 3. Percentage of choices across conditions.

Table 4 shows odds ratios across conditions for the different suffix choices as well as confidence intervals. All effects have to be interpreted in relation to the reference categories /j/ for final consonants and *-ois* for suffix choices, respectively. In the reference category /j/, the distribution of *-ais* and *-ois* is roughly balanced ($p = .634$). For all other final consonants, *-ais* is less frequent than *-ois*, leading to significant effects (/p/: $p = .001$, /m/: $p = .048$, /s/: $p = .001$, /ʃ/: $p = .002$). The suffix *-éen* is also about as frequent as *-ois* in the reference category /j/ ($p = .877$) but also for the final consonant /p/ ($p = .723$). It is, however, significantly less frequent than *-ois* for the final consonants /ʃ/ and /s/ ($p < .001$). The suffix

-ien was chosen significantly more often than *-ois* in the reference category /j/ ($p = .003$). The comparatively more balanced distribution for the final consonant /s/ leads to a significant effect ($p = .016$). The clear preference for *-ois* over *-ien* for the final consonants /ʃ/ and /m/ shows up as a significant effect as well ($p < .001$).

These results contrast with those of Thuilier *et al.*'s (2023) study in subtle ways. In their study, *-ois* is much more frequent than *-ais* across all final consonant categories, including approximants such as /j/, unlike what we found here. Likewise, *-éen* is strongly dispreferred after all consonant categories in their data, whereas we found it to be as likely as *-ois* after /j/ and /p/. As for *-ien*, Thuilier *et al.* found it to be dispreferred when compared to *-ois* in all categories except for alveolar fricatives such as /s/; our study reached a different conclusion on approximants such as /j/, which are the category showing a differential behavior (see discussion in Section 5).

	Suffix: <i>-ais</i>			Suffix: <i>-éen</i>			Suffix: <i>-ien</i>		
Predictors	Odds Ratios	Confidence Interval	p	Odds Ratios	Confidence Interval	p	Odds Ratios	Confidence Interval	p
ref. cat.: /j/	0.91	0.61 – 1.35	0.634	1.03	0.69 – 1.54	0.877	1.71	1.20 – 2.43	0.003
/ʃ/	0.50	0.32 – 0.78	0.002	0.45	0.29 – 0.70	<0.001	0.20	0.13 – 0.31	<0.001
/m/	0.45	0.28 – 0.73	0.001	0.70	0.45 – 1.09	0.116	0.37	0.24 – 0.56	<0.001
/p/	0.61	0.37 – 1.00	0.048	0.92	0.59 – 1.45	0.723	0.80	0.53 – 1.21	0.283
/s/	0.45	0.27 – 0.72	0.001	0.44	0.27 – 0.70	0.001	0.61	0.41 – 0.91	0.016

Table 4. Odds Ratios, 95% Confidence Intervals and p-values for suffix choices across conditions.

4.2. Experiment 2: Nasal vocalic last segments

As in Experiment 1, we used the suffix *-ois* as our reference category. For our first analysis, we used again the *mblogit* function from the *mclogit* package in R, with suffix choice as the dependent variable. Overall (Figure 4), we found that the suffixes *-ois* (360 choices) and *-ais* (380 choices) were largely preferred with no significant difference (Odds Ratios = 1.02, CI = 0.84 – 1.25, $p = .812$) after final nasal, while *-ien* (180) and *-éen* (114) were chosen much less frequently (*-ien* vs *-ois*: Odds Ratios = 0.48, CI = 0.39 – 0.59, $p < .001$; *-éen* vs *-ois*: Odds Ratios = 0.31, CI = 0.24 – 0.39, $p < .001$).

These results are in marked contrast with those of Thuilier *et al.* (2023), who found *-ais* to clearly be the preferred choice after nasal vowels.

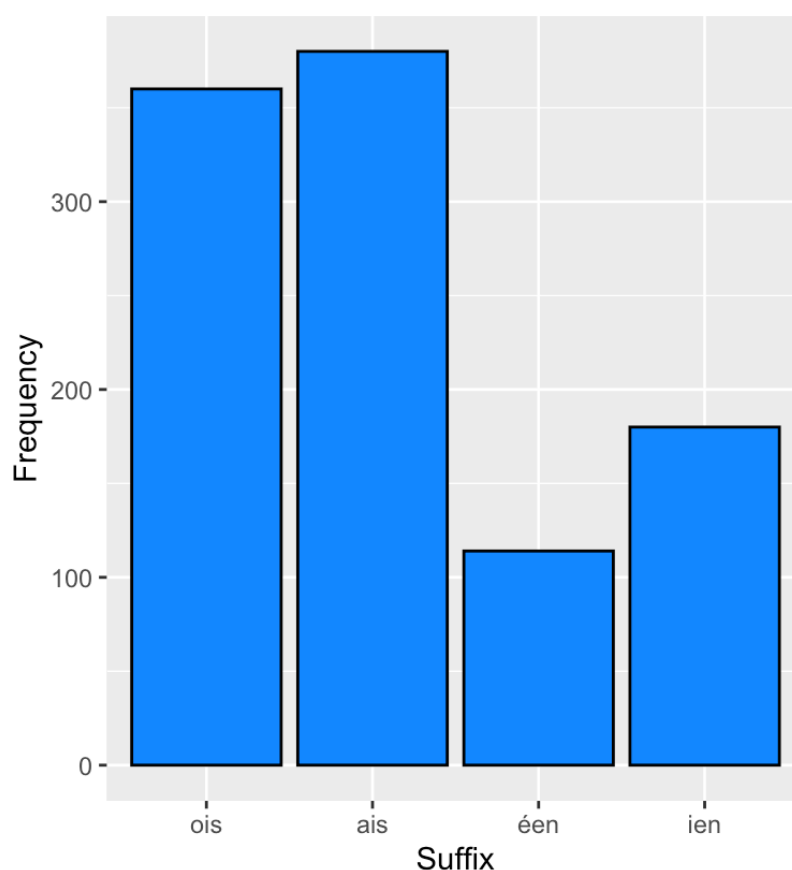


Figure 4. General frequency of suffixes after nasal segments.

Figure 5 shows the distribution of suffix choices across the backness conditions with both syllables coded as “Front” (FF), the first syllable coded as “Back” and the second syllable as “Front” (BF), the first syllable coded as “Front” and the second syllable as “Back” (FB), or both syllables coded as “Back” (BB). Given that the suffixes *-éen* and *-ien* were much less frequent than *-ois* and *-ais*, and moreover with very little variation across conditions, we decided to analyze only suffix choices with *-ois* and *-ais*. Since now the dependent variable only includes two possible choices for this model, we used the `glmer` function (family = binomial) from the `lme4`-package (Bates, Mächler, Bolker, & Walker, 2015) with suffix choice as the dependent variable and the mean-centered factors “First_syllable (Front, Back)” and “Second_syllable (Front, Back)” as predictors. “Front” was always coded as `-.5` and “Back” as `.5`. We also included participants and items as random intercepts. Models with random slopes for participants or items led to convergence failure so that the results we report reflect estimations from intercept models. P-values were estimated using the `lmerTest` package (Kuznetsova, Brockhoff, & Christensen, 2017).

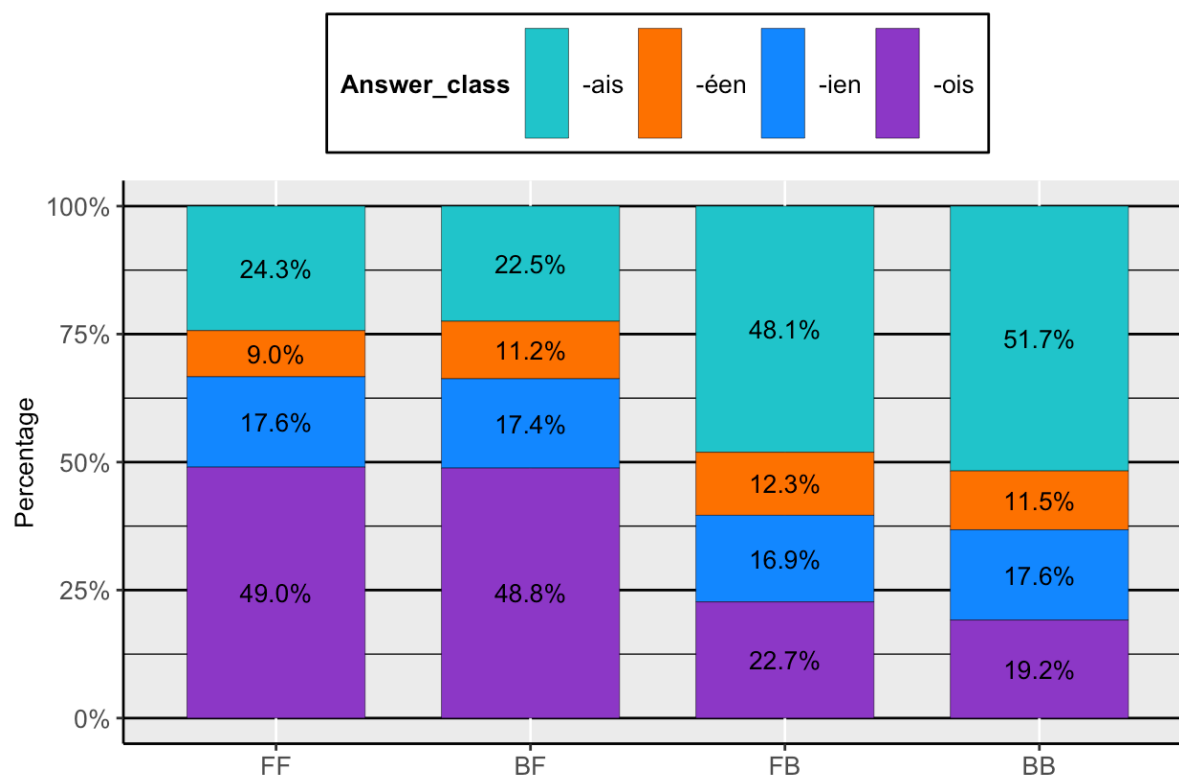


Figure 5. Percentage of choices across backness conditions.

Table 5 shows that only the backness score of the second syllable leads to a significant difference ($p < .001$) in suffix choice. Back vowels lead to a preference for the suffix *-ais*, while front vowels lead to a preference for *-ois*. No effect of the backness of the vowel in the first syllable was documented ($p = .330$), and the interaction between the backness of the two vowels was not significant either ($p = .626$).

These results undermine the hypothesis that the overall backness score of a word has an effect on the choice of suffix. If it did, then it would predict a statistical interaction with strongest effects on choices for the lowest (FF) and the highest (BB) backness scores and more moderate effects for the medium scores (FB, BF). However, this prediction was not confirmed. Our results suggest instead a local effect of the backness of the vowel adjacent to the suffix.

	<i>Odds Ratios</i>	<i>Confidence Interval</i>	<i>p</i>
(Intercept)	0.94	0.74 – 1.20	0.627
backness vowel 1	1.18	0.85 – 1.64	0.330
backness vowel 2	0.17	0.12 – 0.24	<0.001
interaction	1.18	0.61 – 2.29	0.626

Table 5. Odds Ratios, 95% Confidence Intervals and p-values for suffix choices between *-ois* and *-ais* depending on the backness scores of the first and second syllable.

5. Discussion

In the previous section, we identified tendencies in the way the phonological shape of a base toponym influences the choice of an affix to form a toponym. The overall results are in broad agreement with those of Thuilier *et al.* (2023) and prior literature, in confirming that such an influence seems to exist: both experiments documented relevant effects of the broad category of the final segment of the base (here consonant vs. nasal vowels), and more detailed effects of the point and mode of articulation of consonants, and the backness of nasal vowels. In this, our study usefully complements the state of the art: while prior literature has merely established that statistical imbalances exist in the established lexicon, we provide evidence that speakers have preferential judgements that are then likely to influence adoption of a new demonym.

Our results are also in line with Roché and Plénat's (2016) hypothesis that dissimilative constraints are one of the factors influencing the choice of an affix in the formation of toponyms. This is very clear in Experiment 2: suffixes containing a nasal vowel were found to be strongly dispreferred after a nasal-ending base; and a back vowel at the end of the base favors the use of a suffix with a front vowel, and vice versa. The results of Experiment 1, however, lead us to conclude that dissimilative tendencies either need to be refined or can sometimes be dampened by other factors. This is particularly clear for bases ending in the nasal consonant /m/: we would again expect the presence of a nasal in the base to disfavor the use of *-ien* and *-éen*, but no strong difference with other categories of final consonants was found.

The detailed comparison of our results with those of Thuilier *et al.* is a mixed bag. Sometimes the conclusions converge. For instance, in Experiment 1, the two studies agree that *-ais* and *-éen* are the dispreferred choices after consonants; while in Experiment 2, they agree that backness of final vowels influences the choice of suffixes. Sometimes they diverge. For instance, in Experiment 1, we found *-ien* to be the preferred choice after plosives, where Thuilier *et al.* found it to be *-ois*; and in Experiment 2, we found *-ois* and *-ais* to be equally likely after a nasal vowel, where Thuilier *et al.* found *-ais* to be preferred.

There are a variety of possible explanations for these discrepancies, which probably all play some role.

A most obvious one is the fact that the two studies examine samples with markedly different characteristics. First, our datasets do not document derivatives from toponyms of fewer or more than 2 syllables, or ending in an oral vowel. Second, in our study, each category of segments is reduced to a single representative (e.g., /p/ for all plosives, which also include

/b/, /t/, /d/, /k/ and /g/), possibly leading to other properties of that segment having an influence on results that are averaged out in Thuilier *et al.*'s study. Finally, the statistical distribution of the various categories is very different in the two studies; for instance, in Thuilier *et al.*'s study, 46% of consonant-final words end in an approximant, but only 9% belong to the category 'other fricative', whereas each category makes up exactly 20% of the data in our Experiment 1.

Such differences between the samples probably account for a portion of the discrepancies between our results, and diminish the interest of making a detailed comparison between the two studies. A project different from the one we conducted here would be to run an experiment based on a sample that has the same characteristics as Thuilier *et al.*'s. The design of such a sample would have been very costly, and would have come with its own limitations, as some interesting configurations of base properties are too rare in the established lexicon to be explored experimentally. For instance, Thuilier *et al.*'s dataset includes only 6 toponyms ending in /p/ out of a sample of 1,101 consonant final nouns, making it impossible to reach statistically sound conclusions on that configuration in a sample of that size; by comparison, the sample used in Experiment 1 contains 24 toponyms ending in /p/ out of a total of 120.

Be that as it may, differences between samples are very unlikely to be the only source of discrepancy between the two studies, since differences are still found even in parts of the samples that are more comparable. For instance, in Thuilier *et al.*'s data, *-ois* is preferred to *-ien* after all approximants (57% vs. 21%), including in the particular case of the approximant /j/ (41% vs. 21%). The use of /j/ in items with approximants hence cannot explain why we get different results.

Another likely cause of discrepancy between the two studies is the focus on formations accumulated over the centuries versus contemporary formations. It is entirely possible that the differential productivity of the rival affixes evolved over time, as well as the conditions influencing decisions in new coining events. Lexicon-based and experimental studies could be made more comparable in that dimension by conducting a new lexicon-based study centered on recent coinings; whether there is enough data for that to be realistic is an open question.

We could also question whether the type of location, variable in demonym formations in Thuilier *et al.*, could have played a role in the resulting discrepancies. In our experiment, we attempted to control for a potential effect of location types (e.g., cities, countries) in our instructions by not specifying the type of location of our nonce toponyms. Nonetheless, as pointed out by a reviewer, the fact that toponyms are not preceded by a definite article in our items suggests reference to a city rather than another type of location. However, the high

proportion of cities in Thuilier *et al.*'s data – reaching about 85% (1,884 over 2,218 toponyms) – suggests that the type of location cannot by itself explain all discrepancies in the results.

A final likely cause that we want to discuss is the inherent difference between attestations in a lexicon and a speaker's judgments in a specific task. The problem here is comparable to a familiar set of issues in empirical syntax and semantics when trying to combine results from corpus studies with acceptability judgment experiments: while there is a strong expectation of a monotonous relation between corpus frequency and acceptability (the more frequent a linguistic object, the more acceptable it is), the exact mathematical nature of that relation is not obvious, and generally not known. For instance, two constructions could be exactly as acceptable while one is more frequent than the other; and some acceptable configurations are vanishingly rare, to the point of not being attested at all (see for example Francis (2021) for cases of matches and mismatches between corpus frequency and acceptability judgments). In the same fashion, we do expect the prevalence of an affix in the established lexicon to be related to the extent to which speakers accept its use in nonce forms, but that relation can be complex, and depends on the details of the task at hand. As a case in point, the relatively high proportion of choices of *-éen* in our experiments (e.g., 16.5% after an alveolar fricative vs. 0.8% in the lexicon) is probably linked to the adoption of a forced-choice design, where participants are prompted to consider a rare option².

To conclude, this paper has showcased the fact that experiments provide complementary evidence on affix rivalry to quantitative lexical studies, using the domain of demonym formation in French as a test case. We submit that the general method adopted here

² A reviewer rightfully suggests that corpus-based measures of productivity might correlate more closely with preferences in an experimental production setting that type counts in the established lexicon. To check for this, we measured expanding productivity, that is, the proportion of nouns ending with each of the four suffixes among hapaxes in a corpus (Baayen, 2009). Expanding productivity is a measure of the relative contribution of different processes to neology, and hence the one measure of productivity directly relevant to this study. We used as a corpus years 2002 to 2007 of the newspaper *Le Monde* (roughly 77 million tokens). We relied on the fact that nominal demonyms, and only nominal demonyms, are capitalized in French orthography, to hone in on the relevant uses of the four suffixes, and exclude adjectives (e.g., *un bateau anglais* 'an English ship') as well as other (sometimes homographic) nouns such as language names (*l'anglais est une langue germanique* 'English is a Germanic language'). It turns out that three of the suffixes are very close in expanding productivity (*-ais*: $P^* = 0.00090$, *-ois*: $P^* = 0.00090$, *-ien*: $P^* = 0.00107$) while that of *-éen* is an order of magnitude lower ($P^* = 0.00008$). Hence the corpus-based measure associates closely with Thuilier *et al.*'s type counts, and does not change the picture.

can easily be redeployed for other cases of rivalry, and refined to take into account other types of conditioning factors.

References

- Akin, S. (2006). Comment dériver un gentilé à partir d'un toponyme ? Les potentialités signifiantes de Seine-Maritime. *Cahiers de sociolinguistique*, 11, 63-80. <https://doi.org/10.3917/csl.0601.0063>
- Anshen, F., & Aronoff, M. (1981). Morphological productivity and phonological transparency. *Canadian Journal of Linguistics*, 26, 63-72. <https://doi.org/10.1017/S0008413100023525>
- Arndt-Lappe, S. (2014). Analogy in suffix rivalry: the case of English *-ity* and *-ness*. *English Language and Linguistics*, 18, 497-548. <https://doi.org/10.1017/S136067431400015X>
- Aronoff, M. (1976). *Word formation in generative grammar*. MIT Press.
- Baayen, R. H. (2009). Corpus linguistics in morphology: Morphological productivity. In A. Lüdeling, & M. Kyto (Eds.), *Corpus Linguistics. An International Handbook* (pp. 899-919). De Gruyter Mouton.
- Baayen, R. H., Endresen, A., Janda, L. A., Makarova, A., & Nessel, T. (2013). Making choices in Russian: pros and cons of statistical methods for rival forms. *Russian Linguistics*, 37(3), 253-291. <http://dx.doi.org/10.1007/s11185-013-9118-6>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. <https://doi.org/10.18637/jss.v067.i01>
- Bonami, O., & Thuilier, J. (2019). A statistical approach to affix rivalry: French *-iser* and *-ifier*. *Word Structure*, 12(1), 4-41. <https://doi.org/10.3366/word.2018.0130>
- Copot, M., & Bonami, O. (2024). Baseless derivation: the behavioural reality of derivational paradigms. *Cognitive Linguistics*, 35(2), 221-250. <https://doi.org/10.1515/cog-2023-0018>
- Eggert, E. (2002). *La dérivation toponymes-gentilés en français. Mise en évidence des régularités utilisables dans le cadre d'un traitement automatique*. Doctoral dissertation, Université François Rabelais and Westfälische Wilhelms Universität.
- Elff, M. (2024). mclogit: Multinomial Logit Models, with or without Random Effects or Overdispersion. R package version 0.9.8, <<https://github.com/melff/mclogit>>.

- Francis, E. J. (2021). *Gradient Acceptability and Linguistic Theory*. Oxford University Press (online edn, Oxford Academic, 17 Feb. 2022). <https://doi.org/10.1093/oso/9780192898944.001.0001>
- Guzmán Naranjo, M., & Bonami, O. (2023). A distributional assessment of rivalry in word formation. *Word Structure*, 16(1), 86-113. <http://dx.doi.org/10.3366/word.2023.0222>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1-26. <https://doi.org/10.18637/jss.v082.i13>
- Lindsay, M., & Aronoff, M. (2013). Natural selection in self-organizing morphological systems. In N. Hathout, F. Montermini & J. Tseng (Eds.), *Morphology in Toulouse: Selected proceedings of Décembrettes 7* (pp. 133-153). Lincom Europa.
- Lohmann, A. (2017). Phonological properties of word classes and directionality in conversion. *Word structure*, 10, 204-234. <https://doi.org/10.3366/word.2017.0108>
- Lüdecke, D. (2024). *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.8.16, <<https://CRAN.R-project.org/package=sjPlot>>.
- Makarova, A. (2016). Variation in Russian verbal prefixes and psycholinguistic experiments. In T. Anstatt, A. Gattnar, & C. Clasmeier (Eds.), *Slavic languages in psycholinguistics* (pp. 113-133). Narr Francke Attempto Verlag.
- Plag, I. (1999). *Morphological productivity: Structural constraints in English derivation*. Mouton De Gruyter.
- Plénat, M. (2008). Quelques considérations sur la formation des gentils. In B. Fradin (Ed.), *La raison morphologique. Hommage à la mémoire de Danielle Corbin* (pp. 155-174). John Benjamins.
- Roché, M., & Plénat, M. (2016). De l'harmonie dans la construction des mots du français. In F. Neveu, G. Bergounioux, M.-H. Côté, J.-M. Fournier, L. Hriba & S. Prévost (Eds.), *Actes du cinquième Congrès mondial de linguistique française* (article 08003). EDP Sciences. <http://dx.doi.org/10.1051/shsconf/20162708003>
- Romaine, S. (1983). On the productivity of word formation rules and limits of variability in the lexicon. *Australian Journal of Linguistics*, 3(2), 177-200. <https://doi.org/10.1080/07268608308599308>

Schirakowski, B. (2020). (No) competition between deverbal nouns and nominalized infinitives in Spanish. *Borealis – An International Journal of Hispanic Linguistics*, 9(22), 257-283.
<http://dx.doi.org/10.7557/1.9.2.5215>

Thuilier, J., Tribout, D., & Wauquier, M. (2023). Affixal rivalry in french demonym formation: The role of linguistic and non-linguistic parameters. *Word Structure*, 16(1), 115-146.
<https://doi.org/10.3366/word.2023.0223>

Tran, M., & Maurel, D. (2006). Prolexbase. Un dictionnaire relationnel multilingue de noms propres. *Traitement Automatique des Langues*, 47(3), 115-139.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag.