Salvadori, J., Varvara, R., & Huyghe, R. (2024). Measuring affix rivalry as a gradient relationship. In A. Bagasheva, A. Nagano, & V. Renner (Eds.), *Competition in word-formation* (pp. 104–138). John Benjamins. https://doi.org/10.1075/la.284.04sal

Measuring affix rivalry as a gradient relationship

Justine Salvadori, Rossella Varvara & Richard Huyghe
University of Fribourg

Abstract

In many languages, affixes can be used to derive words with the same semantic types and can therefore compete in word formation. This paper discusses how to quantitatively assess the competition between derivational affixes based on their semantic similarity. Two possible measures of affix rivalry drawn from studies in ecology are examined: the Sørensen index, which considers the proportion of shared functions between rival affixes; and the Percentage similarity coefficient, which is based on the realization frequency of functions. Two complementary measures (Balanced richness and Balanced abundance) are also proposed to further analyze the semantic dissimilarity between rival affixes. Using the semantic competition between six French deverbal suffixes as a case study, we show how these four measures suit the quantification of affix rivalry and help capture different aspects of the phenomenon.

Keywords: affix rivalry, quantitative method, nominalization, deverbal suffix, French

1. Introduction

Affix rivalry occurs between affixes that have equivalent semantic functions and can therefore compete in the formation of derivatives. Rival morphological processes have been defined in the literature as “semantically identical” (Plag 1999: 227), as producing “identical results” (Fábregas 2010: 67) or as “correlated with a unique semantic content” (Fradin 2019: 68). However, semantic equivalence is not always considered a strict condition in research on affix rivalry, as the focus is often on the semantic differences that exist between rivals (see e.g., Martin 2010; Schulte 2015; Díaz-Negrillo 2017; Naccarato 2019; Aronoff 2020). According to Huyghe & Varvara (2023), affix rivalry without strict equivalence can be conceived in two ways. On the one hand, rival affixes can be assumed to have similar, but not necessarily identical, semantic functions. In particular, affixes may be regarded as rivals if they are used to derive words with the same coarse-grained semantic type but with finer semantic differences. For example, Nagano (2023) argues that *-ed* and *-y* compete in English to form adjectives that denote gradable properties, but differ in the description of closed-scale vs. totally open-scale properties. Accordingly, one can distinguish between “absolute” and “relative” rivalry, depending on whether semantic identity or similarity is observed. On the other hand, equivalence may be established only between some of the functions of polyfunctional affixes. For example, according to Lieber (2016), the English suffixes *-ation* and *-al* can both derive event (*conversation*, *portrayal*) and result (*coloration*, *acquittal*) nouns, but only the former can be used to derive instrument (*decoration*) and agent (*administration*) nouns. In such a configuration, rivalry can be defined as “total” or “partial” depending on whether it applies to all the semantic functions of two affixes or only to some of them (Huyghe & Wauquier 2021; Guzmán Naranjo & Bonami 2023).

It is debatable whether relative rivalry should be regarded as rivalry strictly speaking. Whether affixes with similar but non-identical functions do or do not compete in the formation of derivatives raises the onomasiological issue of whether concepts are clearly defined prior to naming them. Moreover, relative/absolute and partial/total forms of rivalry are not always clearly distinguished in the literature. Researchers investigating semantic differences between rival affixes do not necessarily make explicit whether these differences are related to distinct semantic functions or not. Apparent cases of relative rivalry may actually pertain to partial rivalry if it turns out that the differences observed between rival affixes depend on different semantic functions of polyfunctional affixes.

The possibility for rival affixes to not be strictly equivalent entails a gradient notion of morphological competition. Semantic differences observed between rival affixes can be more or less important, and affixes can be seen as more or less competing depending on how close they are semantically. This gradient nature of affix rivalry calls for an appropriate, i.e., quantified, assessment. Ideally, a coefficient of competition should be provided so that different situations of rivalry can be compared both within languages and cross-linguistically. In this paper, we explore measures of semantic similarity between polyfunctional affixes that can be used to assess their partial rivalry. We will present several possible indices of rivalry and their potential using simulated data, and then apply them to an actual case study of nominalizing suffixes in French.

The paper is organized as follows. In Section 2, we describe partial rivalry as a correlate of affix polyfunctionality and examine its variation. In Section 3, we introduce the measures that can be used to assess partial rivalry and analyze both semantic similarity and dissimilarity between rival affixes. In Section 4, we use these measures to investigate the rivalry between 6 deverbal suffixes in French, based on the analysis of a sample of 600 nominalizations, and we discuss the results observed.

2. Rivalry and polyfunctionality

Many non-semantic (e.g., phonological, morphological, syntactic, stylistic, sociolinguistic) factors can differentiate rival affixes and contribute to the resolution of affix rivalry. Existing differences are usually observed as tendencies across affix uses, and their importance and relative influence in the resolution of rivalry can be quantified (see e.g., Säily 2011; Baayen et al. 2013; Lindsay & Aronoff 2013; Arndt-Lappe 2014; Bonami & Thuilier 2019; Varvara 2020). However, affix rivalry being fundamentally defined by semantic equivalence (as opposed to other properties that possibly distinguish affixes), the degree of rivalry between two affixes can be primarily evaluated based on semantic properties.

As far as the semantic aspects of morphological competition are concerned, polyfunctionality plays an important role in the distinction of rival affixes. Not only do most affixes serve more than one semantic function, but two affixes also rarely have all their semantic functions in common. Various degrees of rivalry can be observed depending on the functions shared between rival affixes. More precisely, affix rivalry varies according to: (i) the proportion of functions shared between affixes, and (ii) the frequency of instantiation of shared functions among derivatives. We can illustrate this point with the example of the French agentive suffixes *-aire*, *-ant*, *-eur*, *-ien*, *-ier* and *-iste* (see e.g., Roché 2004; Lignon 2007; Schnedecker & Aleksandrova 2016; Huyghe & Wauquier 2021). All these suffixes have additional functions on top of that of forming agent nouns. They can be used to derive nouns that denote instruments (*aspirateur* 'vacuum cleaner'), beneficiaries (*légataire* 'beneficiary'), inhabitants (*Parisien* 'Parisian'), containers (*cendrier* 'ashtray') or partisans (*marxiste* 'Marxist'), among other semantic types. As indicated in Table 1, these additional functions are not shared between all agentive suffixes. Rivalry between pairs of agentive suffixes is always partial in that no two suffixes can be used to form derivatives with the same diversity of semantic types. However, Table 1 shows that the proportion of shared functions may vary between pairs of suffixes, leading to different degrees of rivalry. For example, -*ien* and -*iste*, which compete for 2 out of the 3 functions they serve, will be stronger rivals than -*aire* and -*eur*, which only compete for 1 out of the 4 functions they serve.

Table 1. Subset of semantic types realized by 6 suffixes in French

	Agent	**Instrument**	**Beneficiary**	**Inhabitant**	**Container**	**Partisan**
-aire	x	-	x	-	x	-
-ant	x	x	x	-	-	x
-eur	x	x	-	-	-	-
-ien	x	-	-	x	-	-
-ier	x	x	-	-	x	-
-iste	x	-	-	x	-	x

Further differences can be observed in the structure of shared functionality. A situation of partial rivalry can be symmetric or asymmetric depending on whether or not the proportion of shared functions is the same for competing affixes. In Table 1, rivalry is symmetric between -*aire* and -*iste* because both suffixes compete for one-third of their functions, but it is asymmetric between -*ien* and -*ier* because -*ien* competes with -*ier* for half of its functions whereas -*ier* competes with -*ien* for one-third of its functions. Shared functions may also vary in the way they intersect. Nestedness is observed if the functions of an affix A are a subset of the functions of an affix B, whereas overlap is observed if two affixes A and B have functions in common but also specific functions that are not covered by B and A, respectively (Plag 1999; Guzmán Naranjo & Bonami 2023). In Table 1, -*eur* is nested in -*ant* because all functions of -*eur* can be realized by -*ant* but not reciprocally, whereas -*aire* and -*ier* overlap because they share two functions but each of them is also associated with another function.

Possible variation in affix rivalry is not limited to the (non-)realization of semantic functions, but also depends on the frequency with which functions are realized. Rival affixes that share the same semantic functions can be considered as more or less competing depending on how frequently they are used to form words that instantiate these functions. Such a variation is illustrated with simulated data[1] for French agentive suffixes in Table 2. Based on these data, -*eur*/-*aire* and -*eur*/-*iste* have the same proportion of shared functions (1/4), but rivalry seems stronger in the case of -*eur*/-*iste* than in the case of -*eur*/-*aire* due to more similar proportions of derivatives with shared functions (70% and 65% vs. 70% and 40%). Similarly, pairs of rival affixes with the same asymmetry structure can be seen as more or less asymmetric depending on the proportion of derivatives that instantiate the distinctive functions. For example, although the functions of -*ien*/-*ier* and -*ien*/-*aire* have the same asymmetric distribution, with 1/2 and 2/3 non-competing functions for rival suffixes in both cases, the proportions of derivatives instantiating unshared functions are less balanced in the former case (65% and 40%) than in the latter (65% and 60%). As a consequence, -*ien*/-*ier* can be considered more asymmetric than -*ien*/-*aire*. Overlap also varies according to the realization frequency of semantic functions. It is higher in the case of -*iste*/-*eur* (with 65% and 70% of derivatives instantiating the overlapping function) than in the case of -*aire*/-*ien* (with 40% and 35% of derivatives instantiating the overlapping function), although the same proportions of functions overlap in both cases.

[1] The simulated data used in this study are generated manually to compare the different measures of rivalry we propose and to show how they can be impacted by specific differences in data sets.

Table 2. Simulated distribution of semantic types among derivatives with 6 suffixes in French (100 derivatives per suffix)

	Agent	Instrument	Beneficiary	Inhabitant	Container	Partisan
-aire	40	-	50	-	10	-
-ant	65	25	5	-	-	5
-eur	70	30	-	-	-	-
-ien	35	-	-	65	-	-
-ier	60	15	-	-	25	-
-iste	65	-	-	5	-	30

The differences observed in Table 1 and Table 2 show that situations of affix rivalry should be assessed with respect to both the functions that can or cannot be realized by competing affixes and the abundance of derivatives per affix and function.

3. Measures of rivalry

In this section, we discuss possible measurement of affix rivalry. We first review how morphological competition has been quantified in previous studies. We then introduce measures of similarity that can be used to assess rivalry relationships and illustrate their use and interpretation by means of simulated examples. We use similarity measures to approximate degrees of rivalry since similarity is a basic condition and a main factor of rivalry. Given that degrees of rivalry depend directly on semantic similarity, we assume that they can be evaluated through fine assessment of similarity relationships.

3.1 Quantifying competition

Quantitative approaches to rivalry have developed in recent years, mostly to investigate the discriminative properties of competing affixes. A variety of statistical methods have been used to determine the influence of structural and non-structural factors in the resolution of rivalry. For example, Arndt-Lappe (2014) has used analogical modeling to address the rivalry between *-ity* and *-ness* in English and to investigate suffix preferences related to the phonological properties and syntactic category of the base. Schirakowski (2020) has collected and analyzed experimental data to determine the role of argument structure in the competition between deverbal nouns and nominalized infinitives in Spanish. Using classifier ensembles, Thuilier et al. (2023) have examined whether phonological and morphological properties of base words, as well as extralinguistic properties such as geographical distance, motivate the selection between different rival suffixes used to form demonyms in French.

In some cases, the factors investigated include semantic properties and researchers have attempted to quantify the semantic distinctiveness of rival affixes. For instance, Naccarato (2019) has analyzed the rivalry between Russian (para)synthetic compounds ending in *-ec* and *-tel'* considering several formal and semantic properties. Her results show that *-ec* is favored to form compounds that denote animate entities, whereas *-tel'* is more frequently used to derive inanimate instrument nouns. Denistia et al. (2021) have used distributional semantics methods to examine the differences between the Indonesian prefixes *PE-* and *PEN-*. It appears that words formed with one prefix are more similar to each other than to words formed with the other prefix, and that *PE-* and *PEN-* are preferentially selected to derive agent and instrument nouns, respectively. Distributional semantics has also been used by Varvara et al. (2021) to investigate the differences between German nominal infinitives and suffixed nouns in *-ung*, showing that the former are more semantically transparent with respect to the base verb than the latter. Finally, Lieber & Plag (2022) have examined the differences between conversion and *-ing* nominalization in English with respect to various semantic features, including count/mass quantification, aspectuality and eventive/referential interpretation. Their findings reveal the existence of distinct semantic tendencies for each word formation process.

In all these studies, situations of morphological rivalry are taken for granted while investigation seeks to identify discriminative properties and evaluate their relative importance. Although results show that the differences observed between competing affixes are graded rather than categorical, the degree of morphological rivalry in itself is rarely addressed. Yet the need for a scalar evaluation of rivalry becomes evident when more than two affixes are examined, because of the variable rivalry observed between different pairs of competitors (Huyghe et al. 2023). A measure that describes degrees of rivalry would be required to precisely estimate the strength of a rivalry relationship.

Few studies have explored how to accurately measure morphological competition. Fernández-Domínguez (2017) has defined an index of competition that evaluates the prevalence of a derivative over its rivals. Given a pair (or a triplet) of lexemes derived from the same base but with different affixes, the index takes the total number of competing lexemes and their token frequency into consideration to describe their likelihood of use. It remains that this measure focuses on morphological doublets or triplets and is not meant to assess affix rivalry in itself. In a distributional semantics approach to suffix rivalry in French, Guzmán Naranjo & Bonami (2023) have approximated the semantics of a morphological process by means of the average difference between the distributional representations of derivatives and their bases. Although it is not presented as a coefficient of rivalry in the study, the measurable similarity between the representations of different word-formation processes could be considered a proxy for rivalry. Generally speaking, distributional methods have the advantage of being fully automated and easily applicable on a large scale and in different languages. However, their semantic accuracy can still be improved and representations based on average difference between bases and derivatives cannot disentangle affix polyfunctionality, nor allow for an in-depth analysis of partial rivalry. From a different methodological perspective, Salvadori & Huyghe (2023) have proposed to measure rivalry based on the semantic annotation of derivatives, while estimating the amount of shared and unshared functions among rival affixes. In this paper, we follow this latter approach and further explore similarity metrics that can represent the degree of rivalry between competing affixes.

3.2 Possible measures

Since affix rivalry is based on semantic equivalence, similarity measures can be used as a way to assess rivalry relationships. In this section, we present two similarity measures that take into account the proportion of shared functions between rival affixes and the frequency of realization of these functions, respectively. We then explore two complementary measures to further analyze the semantic dissimilarity between rival affixes.

3.2.1 Similarity coefficients

As noted in Section 2, the degree of rivalry between two affixes depends not only on the number of shared functions between them, but also on the frequency of instantiation of these functions among the derivatives formed with each affix. Two affixes with a certain number of shared functions can be viewed as more or less competing depending on whether they realize these functions at similar relative frequencies or not. To assess affix similarity in situations of partial rivalry, we can either focus on shared vs. unshared functions or take the distribution of derivatives into consideration, which determines the type of data and metric to be used. Two main groups of similarity measures can be distinguished in that respect:

- incidence-based measures, which depend on the number of distinct functions that are realized by rival affixes;
- abundance-based measures, which depend on the number of derivatives instantiating the different functions of rival affixes.

Table 1 exemplifies the type of data considered in incidence-based measures: for each suffix, it indicates which semantic functions are possibly observed or not. Table 2 exemplifies the type of data considered in abundance-based measures: for each suffix, it indicates how many derivatives in a given sample instantiate the different functions realized by the suffix.

Numerous measures have been proposed to assess the similarity or dissimilarity between organized structures, most notably in the field of ecology. Legendre & Legendre (2012), for example, list around 30 different similarity or dissimilarity indices used to assess the variation in species composition across geographical sites. In this paper, we select two standard measures to evaluate the similarity between rival affixes considering the following advantages: (i) they are very commonly used, especially in ecological studies; (ii) they are easy to compute; (iii) their scale is between 0 and 1, which makes the comparison between pairs easier; (iv) several complementary measures have already been proposed in the literature (see e.g., Baselga 2013; Legendre 2014); (v) the results can be used as a basis for other analyses (see e.g., Salvadori & Huyghe 2023 for a hierarchical clustering analysis).

For incidence-based measures, we use the similarity index proposed by Sørensen (1948), which can quantify how similar two affixes are according to the proportion of functions they share. The Sørensen similarity index (S) is calculated by means of the following formula:

$$S = \frac{2|A \cap B|}{|A| + |B|} \qquad \text{(equ. 1)}$$

where A is the set of functions of Affix α, B is the set of functions of Affix β, $A \cap B$ is the set of functions common to α and β, and $|X|$ the number of elements included in Set X. In the example given in Figure 1, α and β share two functions ($|A \cap B| = 2$), whereas they have 5 and 4 functions in total, respectively ($|A| = 5$, $|B| = 4$). Therefore, the Sørensen index for this pair will be (2*2)/(5+4) = 0.44.

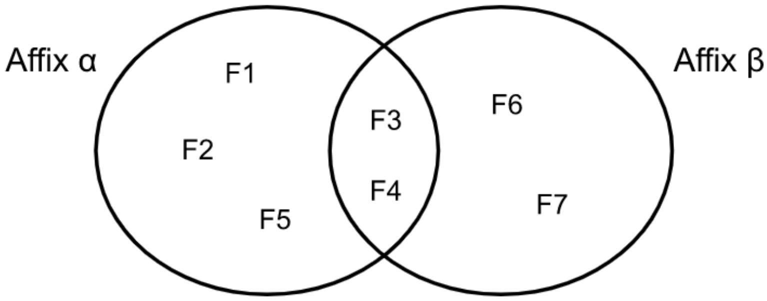

Figure 1. Functional overlap between two affixes

The Sørensen similarity index ranges from 0 to 1, where the maximum value of 1 indicates that two affixes share all of their functions (e.g., Pair 1 in Table 3), and a null value indicates that two affixes do not have any function in common (e.g., Pair 2 in Table 3). The higher the proportion of shared functions, the higher the Sørensen index. For example, in the simulated examples in Table 3, Affixes E and F both have 2/3 functions in common whereas E and G have 1/3 and 1/2 functions in common, respectively. Consequently, the Sørensen similarity index is higher in Pair 3 than in Pair 4. Interpreting this index as a measure of semantic rivalry, we can infer that Affix E competes more with Affix F than with Affix G.

Table 3. Simulated examples of rival affixes with presence/absence of functions and Sørensen index

		F1	F2	F3	F4	*S*
Pair 1	Affix A	x	x	x	x	1
	Affix B	x	x	x	x	
Pair 2	Affix C	x	x	-	-	0
	Affix D	-	-	x	x	
Pair 3	Affix E	x	x	x	-	0.67
	Affix F	x	x	-	x	
Pair 4	Affix E	x	x	x	-	0.4
	Affix G	x	-	-	x	

Since the Sørensen index takes only the presence or absence of functions into consideration, it will return the same value for two pairs of rival affixes with any number of derivatives per function, although the frequencies can be very different and influence similarity relationships. To weight the similarity index by the distribution of derivatives across functions, we turn to abundance-based measures. Here we use the Percentage similarity index (as a complement to the Percentage difference index proposed by Odum 1950), also commonly referred to as the "Bray-Curtis similarity", to quantify rivalry considering type frequencies. Percentage similarity (*PS*) is calculated by means of the following formula:

$$PS = \frac{2\sum_{i=1}^{p} \quad min(N_{i\alpha}, N_{i\beta})}{\sum_{i=1}^{p} \quad (N_{i\alpha} + N_{i\beta})} \qquad \text{(equ. 2)}$$

where $N_{i\alpha}$ is the number (i.e., the abundance) of derivatives with Affix α that realize Function i, $N_{i\beta}$ the number of derivatives with Affix β that realize Function i, p the total number of functions observed for α and β, and *min(a,b)* is the smaller of two numbers *a* and *b*. In Figure 2, the minimum number of derivatives formed with *α* and *β* is 0 for all unshared functions, 3 for F4 and 2 for F5, hence the numerator of *PS* is 2*(3+2), whereas the denominator is equal to the total number of derivatives formed with each affix (16+16). Therefore, *PS* for this pair is 10/32 = 0.31.

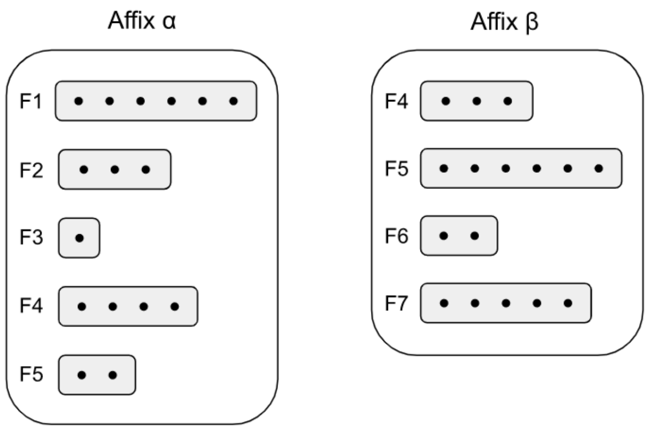

Figure 2. Abundance of derivatives per function for two affixes

Similarly to the Sørensen index, the *PS* coefficient ranges from 0 to 1, with higher values indicating stronger competition. With an equal number of shared and unshared functions, *PS* increases as the difference in numbers of derivatives per function decreases, as illustrated with simulated data in Table 4. While rival affixes in Pairs 5, 6 and 7 have the same proportion of shared functions and therefore the same Sørensen index, they vary considerably with respect to the numbers of derivatives per function. Derivatives formed with Affixes H and I are evenly distributed among the functions of each affix, whereas derivatives formed with Affixes J and K are unevenly distributed in favor of Function 1. The *PS* index varies accordingly, yielding higher values in Pair 5 than in Pair 6, and in Pair 6 than in Pair 7. These differences can be interpreted as follows: H competes more with I than with J, and more with J than with K.

Table 4. Simulated examples of rival affixes with equal *S* and different *PS* values

		F1	F2	F3	F4	*S*	*PS*
Pair 5	Affix H	30	30	30	30	0.86	0.75
	Affix I	40	40	40	0		
Pair 6	Affix H	30	30	30	30	0.86	0.58
	Affix J	80	20	20	0		
Pair 7	Affix H	30	30	30	30	0.86	0.33
	Affix K	110	5	5	0		

Based on the examples in Table 4, *PS* seems to provide a finer assessment of similarity than the Sørensen index. However, as a measure based on abundance only, *PS* is less sensitive to variation in the number of shared and unshared functions. We can imagine cases in which the Sørensen index varies while *PS* remains unchanged, as exemplified in Table 5. Although Pairs 8, 9 and 10 have the same ratio between the minimal number of derivatives instantiating shared functions and the total number of derivatives formed with rival affixes, the proportion of shared functions is different in the three cases (2/3 functions shared for both L and M in Pair 8, 2/3 and 1/2 functions shared for L and N in Pair 9, all functions shared for O and P in Pair 10). Consequently, Pairs 8, 9 and 10 have the same *PS* but different *S* indices.

Table 5. Simulated examples of rival affixes with equal *PS* and different *S* values

		F1	F2	F3	F4	*S*	*PS*
Pair 8	Affix L	20	20	20	0	0.67	0.67
	Affix M	20	20	0	20		
Pair 9	Affix L	20	20	20	0	0.80	0.67
	Affix N	40	20	0	0		
Pair 10	Affix O	10	10	20	20	1.00	0.67
	Affix P	20	20	10	10		

It appears that the Sørensen and Percentage similarity indices capture different aspects of similarity. The pros and cons of both incidence-based and abundance-based measures have been debated in ecological studies. The selection of an index to assess similarity relationships is often motivated by data availability and sampling methods. If robust information about abundance cannot be guaranteed, similarity is better measured using an incidence-based method (Baselga 2013), which does not necessarily affect empirical representations (Wilson 2012). Research purposes also influence the selection of a measure: an incidence-based metric can be used if one wants to give importance to rare elements (Liu et al. 2022), or more generally if one wants to focus on types rather than individuals (Anderson et al. 2011). On the other hand, abundance-based measures have the advantage of assessing similarity through individual occurrences, therefore allowing for quantitatively nuanced descriptions of similarity. Whenever possible, it seems interesting to use both incidence-based and abundance-based measures of similarity, comparing and combining the information they provide to achieve a detailed description of similarity relationships.

3.2.2 Complementary indices

The Sørensen and *PS* coefficients seem to be well suited for a general measurement of the similarity between derivational affixes. However, these measures could be complemented with additional information about dissimilarity structure (e.g., overlapping or nested rivalry, degree of asymmetry between rival affixes, size difference between rivals' niches). To further analyze the dissimilarity between competing affixes, we propose to quantify the balanced distribution of their unshared functions. We use both an incidence-based measure (Balanced richness) and an abundance-based measure (Balanced abundance) that depend on the number of unshared functions between rival affixes, the number of derivatives per unshared function and their distribution across affixes.

Balanced richness (*BR*) can be defined as the ratio between the numbers of unshared functions for two rival affixes. It is calculated from the following equation:

$$BR = \frac{min(|A \setminus B|, |B \setminus A|)}{max(|A \setminus B|, |B \setminus A|)} \quad \text{(equ. 3)}$$

where *A* is the set of functions of Affix *α*, *B* is the set of functions of Affix *β*, $X \setminus Y$ the relative complement of Set *Y* in Set *X* (i.e., the set of elements in *X* but not in *Y*), $|X|$ the number of elements included in Set *X*, *min(a,b)* the smaller of two numbers *a* and *b*, and *max(a,b)* the greater of two numbers *a* and *b*. In the example given in Figure 1, $|A \setminus B| = 3$, $|B \setminus A| = 2$, and therefore $BR = 0.67$. The higher *BR* is, the more evenly distributed are the unshared functions between two affixes. A *BR* of zero indicates that one rival is nested in the other (as long as similarity is not null), in which case dissimilarity between the rival affixes is due only to a difference in richness. In contrast, a *BR* of 1 indicates a symmetric overlap of functions between rival affixes (as long as similarity is not null) and that dissimilarity is due to an even distribution of unshared functions among rival affixes. In other words, *BR* is a measure of functional asymmetry in rivalry situations. The higher the coefficient is, the more a situation of partial rivalry tends towards functional symmetry.

The behavior of *BR* is illustrated with simulated data in Tables 6 and 7, where pairs of affixes have the same Sørensen similarity coefficient but differ in *BR*. All unshared functions in Pair 11 fall to one affix and Affix C is nested in Affix A, whereas no richness difference is observed in Pair 12 and Affixes E and F have the same number of unshared functions. In such cases, $BR = 0$ and $BR = 1$, respectively. By contrast, dissimilarity in Pairs 13 and 14 is caused partly by richness difference and partly by an unbalanced distribution of unshared functions. Accordingly, *BR* is neither equal to zero nor 1. Furthermore, the lower value of *BR* in Pair 13 indicates that the distribution of unshared functions is less balanced than in Pair 14, and that the situation of rivalry is more asymmetric in Pair 13 than in Pair 14.

Table 6. Simulated examples with equal *S* and different *BR* values

		F1	**F2**	**F3**	**F4**	***S***	***BR***
Pair 11	Affix A	x	x	x	x	0.67	0
	Affix C	x	x	-	-		
Pair 12	Affix E	x	x	x	-	0.67	1
	Affix F	x	x	-	x		

Table 7. Simulated examples with equal *S* and different *BR* values

		F1	**F2**	**F3**	**F4**	**F5**	**F6**	***S***	***BR***
Pair 13	Affix Q	x	x	x	x	x	-	0.29	0.25
	Affix R	x	-	-	-	-	x		
Pair 14	Affix S	x	x	x	-	-	-	0.29	0.67
	Affix T	x	-	-	x	x	x		

Next, Balanced abundance (*BA*) can be defined as the ratio between the numbers of derivatives with unshared functions for two rival affixes. It is calculated from the following equation:

$$BA = \frac{min(\sum_{j=1}^{q} N_{j\alpha}, \sum_{k=1}^{r} N_{k\beta})}{max(\sum_{j=1}^{q} N_{j\alpha}, \sum_{k=1}^{r} N_{k\beta})} \quad \text{(equ. 4)}$$

where $N_{j\alpha}$ is the number of derivatives with Affix α that realize the function j, $N_{k\beta}$ the number of derivatives with Affix β that realize the function k, q the total number of functions of α but not of β, r the total number of functions of β but not of α, *min(a,b)* the minimum of two values *a* and *b*, and *max(a,b)* the maximum of two values *a* and *b*. In the example given in Figure 2, $\sum_{j=1}^{q} N_{j\alpha} = 10$, $\sum_{k=1}^{r} N_{k\beta} = 7$, and therefore *BA* = 0.7. The higher *BA* is, the more similar are the numbers of derivatives instantiating unshared functions between two affixes. As in the case of *BR*, a *BA* of zero indicates that one rival is nested in the other (as long as similarity is not null). However, a *BA* of 1 differs from a *BR* of 1 because it indicates that the rival affixes form the same number of derivatives for unshared functions, regardless of symmetric overlap of functions.

The behavior of *BA* is illustrated with simulated data in Tables 8, 9 and 10. Rival affixes in Pairs 15, 16 and 17 have the same functional similarity and the same number of unshared functions, but the balance between the numbers of derivatives instantiating unshared functions is variable (with increasing disproportion from Pair 15 to Pair 17). It follows that, although rival affixes have the same *S*, *BR* and *PS* indices, they differ in *BA*, with higher coefficients in Pair 15 than in Pair 16, and in Pair 16 than in Pair 17. Dissimilarity in Pair 15 is due to unshared functions evenly distributed among rivals and with the same number of derivatives, whereas in Pairs 16 and 17 it is due to unshared functions evenly distributed among rivals but with different numbers of derivatives. In other words, only Pair 15 is fully symmetric, and Pair 17 is more asymmetric than Pair 16.

Table 8. Simulated examples of rival affixes with equal *S*, *BR*, *PS* and different *BA* values

		F1	F2	F3	F4	*S*	*BR*	*PS*	*BA*
Pair 15	Affix L	20	20	20	0	0.67	1	0.67	1
	Affix M	20	20	0	20				
Pair 16	Affix L	20	20	20	0	0.67	1	0.67	0.5
	Affix U	25	25	0	10				
Pair 17	Affix L	20	20	20	0	0.67	1	0.67	0.1
	Affix V	29	29	0	2				

Pairs 18 and 19 in Table 9 have the same incidence-based and *PS* measures, but in contrast to the pairs in Table 8, their *BR* coefficient is different from 1 because unshared functions are not evenly distributed among rival affixes. Due to variable proportions of derivatives realizing unshared functions, *BA* is very different between Pair 18 and Pair 19 and can be higher or lower than *BR* depending on the distribution of derivatives. The combination of *BR* and *BA* indicates that Pairs 18 and 19 have the same disproportion of unshared functions, but that unshared functions are associated with a fairly similar number of derivatives in Pair 18 (*BA* = 0.89), while functional asymmetry is enhanced by important differences in the number of derivatives realizing unshared functions in Pair 19 (*BA* = 0.09).

Table 9. Simulated examples of rival affixes with equal *S*, *BR*, *PS* and different *BA* values

		F1	F2	F3	F4	F5	F6	*S*	*BR*	*PS*	*BA*
Pair 18	Affix W	20	20	20	0	0	0	0.29	0.67	0.25	0.89
	Affix X	15	0	0	15	15	15				

Pair 19	Affix Y	56	2	2	0	0	0	0.29	0.67	0.25	0.09
	Affix X	15	0	0	15	15	15				

Finally, Pairs 20 and 21 in Table 10 illustrate the situation in which two pairs of rival affixes have the same total number of derivatives for unshared functions, but exhibit a different configuration of functional asymmetry. While both pairs have a *BA* of 1, their *BR* coefficient is different because unshared functions are less evenly distributed in Pair 21 than in Pair 20 and therefore Pair 21 is more asymmetric than Pair 20. This variation confirms the independence of *BR* and *BA* and shows how the two measures complement each other in the description of dissimilarity structures.

Table 10. Simulated examples with equal *S*, *PS*, *BA* and different *BR* values

		F1	**F2**	**F3**	**F4**	**F5**	**F6**	***S***	***BR***	***PS***	***BA***
Pair 20	Affix Z	10	30	30	0	0	0	0.29	0.67	0.14	1
	Affix AA	10	0	0	20	20	20				
Pair 21	Affix AB	10	15	15	15	15	0	0.29	0.25	0.14	1
	Affix AC	10	0	0	0	0	60				

4. Case study: Deverbal nominalizing suffixes in French

In this section, we apply similarity and dissimilarity measures to real linguistic material, viz. suffixes used to form deverbal nouns in French. We first explain how we collected the data (Section 4.1) and analyzed it semantically (Section 4.2) and then present the results[2] (Section 4.3).

[2] All analyses were performed in R version 4.2.1 (R Core Team, 2022).

4.1 Data collection

Two sets of rival suffixes were selected for this study based on the existing literature (see e.g., Dubois 1962; Thiele 1987; Huyghe & Wauquier 2021): 3 "eventive" suffixes (-*ade*, -*ment*, -*ure*) and 3 "agentive" suffixes (-*aire*, -*ant*, -*eur*). Note that the "eventive" and "agentive" labels are used as umbrella terms according to the semantic function both groups of suffixes supposedly realize most frequently. As mentioned in Section 2, most derivational processes are polyfunctional, and the 6 suffixes examined here should be no exception. It is well known that, in many languages, eventive suffixes also frequently allow for the formation of result nouns (see e.g., Jacquey 2006; Melloni 2011; Lieber 2016), whereas agentive suffixes are often used to form instrument-denoting derivatives (see e.g., Bauer 2002; Rainer 2015). As a corollary, these two sets of affixes may not necessarily be as distinct as they seem. It could be that suffixes that apparently belong to different groups share at least one function and could therefore be considered rivals. As a precaution, all 6 suffixes will consequently be analyzed together for the application of the different rivalry measures.

Morphological competition can only be investigated through lexical instantiation. Following the selection of affixes, we retrieved a random sample of 100 French deverbal nouns formed with each suffix from the French web corpus FRCOW16A (Schäfer & Bildhauer 2012; Schäfer 2015). We extracted all lemmatized forms tagged as verbs and nouns before automatically filtering verb-noun pairs that were formally related. Selected pairs were then randomly ordered for each of the 6 suffixes, and the first 100 semantically related pairs were retained for further analysis. In total, 600 nouns were thus collected. An example of a derivative selected for each suffix is reported in (1):

(1) a. *glisser* ‘slip’ + *-ade* → *glissade* ‘slip’
b. *licencier* ‘dismiss’ + *-ment* → *licenciement* ‘dismissal’
c. *graver* ‘engrave’ + *-ure* → *gravure* ‘engraving’
d. *signer* ‘sign’ + *-aire* → *signataire* ‘signatory’
e. *combattre* ‘fight’ + *-ant* → *combattant* ‘fighter’
f. *déménager* ‘move’ + *-eur* → *déménageur* ‘mover’

4.2 Semantic analysis

In order to evaluate the sample in a homogeneous way, all collected nouns were analyzed by a single annotator[3] according to 3 criteria: (i) the base verb, assuming that derivational processes apply to semantically specified items (Mel′čuk 1994; Fradin & Kerleroux 2003); (ii) the ontological type, which depends on the nature of the derivative’s referent (e.g., animate entity, event); (iii) and the relational type, which describes the semantic relation between the derivative and its morphological base (e.g., agent, result). For example, the noun *cambrioleur* ‘burglar’ is based on the verb *cambrioler* ‘burgle’; its ontological type is Animate as it denotes an animate entity; and its relational type is AGENT because it refers to the person doing the action of burgling. Ontological and relational types do not provide the same type of information and are not reducible to each other, as shown in examples (2-3). The same ontological type (Artefact) can be associated with different relational types (RESULT, INSTRUMENT, LOCATION) as illustrated in (2), whereas the same

[3] The same methodology was employed in a study involving two of the authors as annotators (Huyghe et al. 2023). The semantic analysis, which covered 501 meanings for 300 deverbal neologisms, was performed by 3 annotators in a double-blind fashion. Agreement scores were substantial, with an observed agreement of 83% (Cohen’s K = 0.77) for ontological types and 93% (Cohen’s K = 0.78) for relational types.

relational type (RESULT) can be associated with different ontological types (Artefact, State, Animate), as illustrated in (3).

(2) a. *bâtir* ‘build’ → *bâtiment* ‘building’
[Artefact-RESULT]
b. *désodoriser* ‘deodorize’ → *désodorisant* ‘deodorant’
[Artefact-INSTRUMENT]
c. *se promener* ‘promenade’ → *promenade* ‘promenade’
[Artefact-LOCATION]

(3) a. *bâtir* ‘build’ → *bâtiment* ‘building’
[Artefact-RESULT]
b. *énerver* ‘annoy’ → *énervement* ‘annoyance’
[State-RESULT]
c. *créer* ‘create’ → *créature* ‘creature’
[Animate-RESULT]

The ontological classification includes 14 types. They are identified on the basis of linguistics tests (see e.g., Godard & Jayez 1996; Flaux & Van de Velde 2000; Huyghe 2015) and are applied following a decision tree presented in Haas et al. (2023). For example, a noun is considered to denote an event if it can be employed as the subject of the verbs *avoir lieu* ‘take place’ or *se produire* ‘occur’, or as the object of the verbs *effectuer* ‘perform’, *procéder à* ‘proceed’ or *accomplir* ‘accomplish’. Some of the 14 ontological types are also combined to form 7 complex types in order to take into account monosemous derivatives with a hybrid semantic structure (see e.g., Cruse 1995; Pustejovsky 1995; Asher 2011; Murphy 2021). Such nouns are generally identified through co-predication, where predicates typical of different semantic types can be associated without any zeugma effect. In (4), for example, *distribuer* ‘distribute’ and *traduire* ‘translate’ apply simultaneously to *brochure* 'brochure' while referring to an artefact and a cognitive object, respectively. In addition, both simple and complex types can be assigned an additional Collective label, which is used to distinguish nouns that have plural reference in the singular form (5).

(4) Spior a décidé de distribuer une brochure traduite en plusieurs langues en France, en Grande-Bretagne, en Belgique, en Allemagne, en Italie et en Espagne. (*liberation.fr*)
'Spior decided to distribute a brochure translated into several languages in France, Great Britain, Belgium, Germany, Italy and Spain'

(5) *lotir* 'divide into plots' → *lotissement* 'housing estate' [Artefact.Collective]

The relational classification includes 17 semantic types adapted from the VerbNet (Kipper-Schuler 2005) and LIRICS (Petukhova and Bunt 2008) inventories of semantic roles. These relational types are distinguished on the basis of definitions presented in Salvadori & Huyghe (2023). For example, a noun is considered to refer to an agent if it denotes an entity that intentionally brings about the event denoted by the base verb (e.g., *cambrioler* 'burgle' → *cambrioleur* 'burglar') and an instrument if it denotes an entity that is used to perform the action denoted by the base verb (e.g., *charger* 'charge' → *chargeur* 'charger'). A transposition type is also added to the relational classification to account for cases in which the derivative denotes (roughly) the same eventuality as its base verb, as in *ruer* 'kick' → *ruade* 'kicking' (see e.g., ten Hacken 2015; Lieber 2015).

Ambiguity is known to be pervasive in nominalization and should be taken into account when examining morphological competition. For this study, we considered that a derivative was ambiguous if we could assign it more than one base verb,[4] one ontological type or one relational type. For example, two meanings of the noun *poseur* 'installer'/'poser' can be distinguished in (6) because they are derived from distinct meanings of *poser* 'install'/'pose'. Similarly, *logement* 'housing'/'house' can be considered ambiguous in (7) because it is associated with different ontological (Event, Artefact) and relational (TRANSPOSITION, LOCATION) types.

(6) a. *poser*$_1$ 'install' → *poseur*$_1$ 'installer'
[Animate-AGENT]
b. *poser*$_2$ 'pose' → *poseur*$_2$ 'poser'
[Animate-AGENT]

(7) a. *loger*$_1$ 'house'→ *logement*$_1$ 'housing'
[Event-TRANSPOSITION]
b. *loger*$_1$ 'house' → *logement*$_2$ 'house'
[Artefact-LOCATION]

Following the semantic annotation, 840 meanings were identified for the 600 nouns, which averages out to 1.4 meaning per noun.

4.3 Results

The following subsections present the results obtained for the 600 nouns. Section 4.3.1 provides general information about the realization of functions. We report on the results for the similarity measures based on

[4] Ambiguous verbs were identified through variation of argument structure, semantic role assignment or lexical aspect.

presence/absence and abundance data in Sections 4.3.2 and 4.3.3, and then compare them in Section 4.3.4.

4.3.1 General information

A total of 61 combined semantic types (i.e., that include an ontological type and a relational type) were identified in the dataset. To maximize the chances that they correspond to semantic functions and not to idiosyncratic or lexicalized meanings, we removed from the sample those that were observed only once per suffix, which corresponds to 58 lexical items and 24 combined semantic types. Analyses were therefore performed on 782 word meanings and 37 combined functions.

The most frequent functions in the sample are Animate-AGENT (24% of the 782 meanings), Event-TRANSPOSITION (19%) and Artefact-INSTRUMENT (9%), as exemplified in (8a-c). At the other end of the spectrum, 22% of the functions are realized only twice in the dataset, as in the case of Animate-THEME (8d), for example. On average, a function is instantiated by 21.1 nouns (*SD* = 40.2).

(8) a. *acquérir* 'buy' → *acquéreur* 'buyer'
[Animate-AGENT]
b. *se noyer* 'drown' → *noyade* 'drowning'
[Event-TRANSPOSITION]
c. *tonifier* 'invigorate' → *tonifiant* 'tonic'
[Artefact-INSTRUMENT]
d. *résider* 'reside' → *résident* 'resident'
[Animate-THEME]

As illustrated in Figure 3, the 6 suffixes are all polyfunctional. Important disparities can be observed between them with respect to the number of functions they realize and their frequency of realization. In our sample, *-ure* has the highest number of functions (16), followed by *-ant* (15), *-ade* (13), *-ment* (9), *-aire* (6) and *-eur* (4). These functions are unevenly served by the different suffixes, since the number of derivatives per function is highly variable. As a consequence, affix polyfunctionality can be accurately described by diversity measures that take into account both the number of functions and the distribution of derivatives across functions. Here we use the Hill-Shannon index (see Roswell et al. 2021), which can be interpreted as the number of functions a suffix would have if these functions were equally distributed among derivatives. According to this index, the most diverse suffix in our sample is *-ure* (*D* = 10.1), followed by *-ade* (6.8), *-ant* (6.5), *-ment* (4.6), *-aire* (3.7) and *-eur* (2.2).

It can be noted that both "eventive" and "agentive" suffixes do not necessarily conform to their labels. Although Event-TRANSPOSITION is the most frequent function for *-ade* (accounting for 43% of the derivatives), it is only observed for 21% of the nouns ending in *-ure*, which favor the denotation of concrete results, whether natural (16%) (as in 9a) or artefactual (19%) (as in 9b). As for *-ment*, it forms almost as many nouns denoting events with a stative facet (36%) (see 9c) as nouns denoting regular events (39%). Regarding agentive suffixes, *-eur*, *-ant* and *-aire* mostly form animate agents (66%, 47% and 46%, respectively), but they also frequently realize other functions: Artefact-INSTRUMENT in the case of *-eur* (29%) and *-ant* (20%) (see 10a-b); Animate-BENEFICIARY in the case of *-aire* (33%) (see 10c).

(9) a. *chier* 'shit' → *chiure* 'shit'
[Natural-RESULT]
b. *enjoliver* 'embellish' → *enjolivure* 'embellishment'
[Artefact-RESULT]
c. *apaiser* 'appease' → *apaisement* 'appeasement'
[Event.State-TRANSPOSITION]

(10) a. *défibriller* ‘defibrillate’ → *défibrillateur* ‘automatic defibrillator’
[Artefact-INSTRUMENT]
b. *décongestionner* ‘decongest’ → *décongestionnant* ‘decongestant’
[Artefact-INSTRUMENT]
c. *narrer* ‘narrate’ → *narrataire* ‘narratee’
[Animate-BENEFICIARY]

Overall, these results confirm what was presented in Section 2. Sketching the semantic profile of a suffix requires reckoning with its polyfunctionality, which varies according to the number of functions it serves as well as the frequency of realization of these functions. These two dimensions should be taken into account when evaluating rivalry situations.

	−ade	−ment	−ure	−aire	−ant	−eur
Animate−AGENT				48	48	92
Animate−BENEFICIARY				34	2	
Animate−EXPERIENCER					2	2
Animate−PATIENT					2	
Animate−PIVOT					4	
Animate−STIMULUS					2	
Animate−THEME					2	
Artefact−DESTINATION			3			
Artefact−INSTRUMENT		2	11		20	41
Artefact−LOCATION	3					
Artefact−PIVOT					2	
Artefact−RESULT	10		30			
Artefact−THEME					2	
Artefact.Cognitive−INSTRUMENT		4		7		
Artefact.Cognitive−RESULT			3	2		
Artefact.Collective−RESULT			2			
Cognitive−INSTRUMENT				3	2	4
Cognitive−PIVOT					3	
Cognitive−TOPIC	3					
Cognitive.Event−TRANSPOSITION	19	7	10	10		
Disease−CAUSE	2					
Domain−TRANSPOSITION			3			
Event−TRANSPOSITION	59	55	34			
Event.Collective−TRANSPOSITION	13					
Event.Financial−TRANSPOSITION		6				
Event.Natural−TRANSPOSITION			13			
Event.Phenomenon−TRANSPOSITION	7	11				
Event.State−TRANSPOSITION	11	51	9			
Natural−CAUSE			3		6	
Natural−LOCATION	2		2		2	
Natural−PATIENT					3	
Natural−RESULT	2		26			
Natural.Collective−THEME	2					
Property−CAUSE	3					
Property−MANNER		2	2			
Property−RESULT			5			
State−RESULT		4	3			

Scale: 0, 20, 40, 60, 80

Figure 3. Frequency of functions per suffix (darker shades indicate higher frequencies)

4.3.2 Incidence-based measures

To compute incidence-based measures of similarity and dissimilarity, we transformed the exact frequencies presented in Figure 3 into presence/absence data. The Sørensen similarity (*M* = 0.23, *SD* = 0.15) and Balanced richness (*M* = 0.45, *SD* = 0.25) scores calculated for each pair of suffixes are presented in Figure 4.

Several elements can be highlighted with respect to the similarity scores. As a general observation, their diversity supports the need to approach affix rivalry as a gradient phenomenon. Suffixes that are semantically close (e.g., *-aire*/*-eur*) do not necessarily share all their functions (e.g., Animate-BENEFICIARY is only observed for *-aire*), whereas suffixes that are semantically distant (e.g., *-ade*/*-ant*) may still have functions in common (e.g., Natural-LOCATION). As far as the distinction between "agentive" and "eventive" suffixes is concerned, there is no clear break between the similarity scores of the different pairs, although the first ranks do not seem to be random. Among the 15 possible pairs, the 6 most similar involve all the suffixes of the sample in the form of triplets distributed each time in 3 rival pairs, which seems to support the idea of 2 distinct semantic groups. That the distinction between "eventive" and "agentive" suffixes is tangible even when using an incidence-based measure is intriguing. As a reminder, the Sørensen coefficient gives the same importance to the most expected functions for each group (i.e., Event-TRANSPOSITION and Animate-AGENT, respectively) as to those that are intuitively more marginal (e.g., Animate-EXPERIENCER). Considering that all the studied suffixes are polyfunctional (see Section 4.3.1), these results show that "eventive" and "agentive" suffixes do not only compete for their most expected function, but also for others. For example, the 3 "eventive" suffixes all realize the Event.State-TRANSPOSITION function, whereas none of the "agentive" suffixes serves it. Conversely, the Cognitive-INSTRUMENT function is realized by the 3 "agentive" suffixes, but not by the "eventive" ones. From a systemic point of view, this can indicate that there are non-arbitrary associations between functions (i.e., that polyfunctionality is to some extent motivated) and that some of these associations are distinctive of general semantic groups. It remains that the “eventive” and “agentive” categories are not watertight. Some functions — such as Artefact-INSTRUMENT and Natural-LOCATION, for instance — are realized by both groups.

As presented in Section 3.2.2, the complementary measure of Balanced richness can be employed to analyze the dissimilarity between non-absolute rivals. Overall, the Balanced richness scores observed in our sample are quite variable. The pair *-ant*/*-ure* reaches a *BR* score of 0.92, tending towards symmetric functional overlap, whereas nestedness (*BR* = 0) can be observed in the case of *-ant*/*-eur*. The latter result indicates that all functions of *-eur* are realized by *-ant* but not reciprocally. By contrast, *-eur* realizes 2 functions (Artefact-INSTRUMENT and Animate-EXPERIENCER) that *-aire* does not serve, hence the higher *BR* score for the pair they form (0.5). Considering that *-ant*/*-eur* and *-aire*/*-eur* have roughly the same similarity score (*S* = 0.42 and *S* = 0.4, respectively), these results illustrate that the structure of dissimilarity does not necessarily depend on the degree of competition.

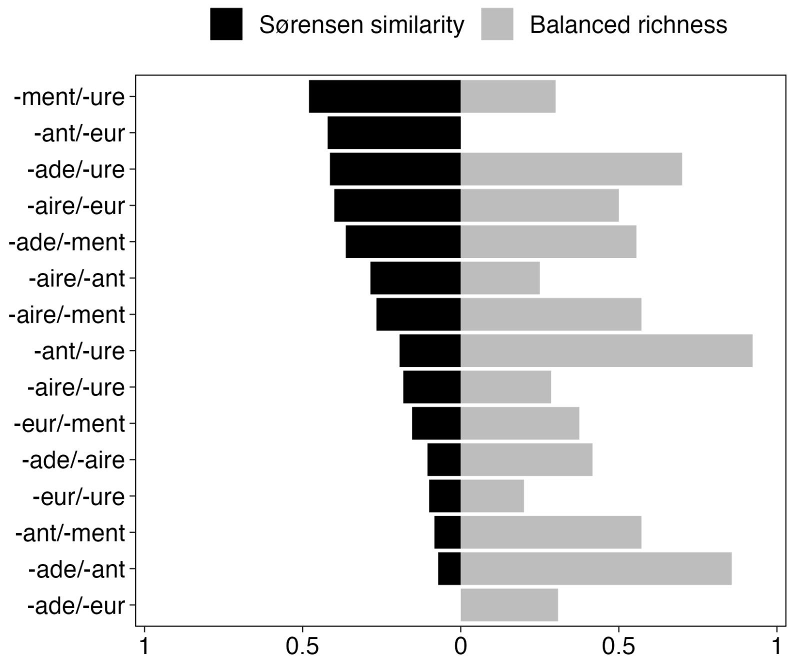

Figure 4. Scores for incidence-based measures (pairs of suffixes are ordered from top to bottom by decreasing similarity)

4.3.3 Abundance-based measures

The exact frequencies presented in Figure 3 were used to compute abundance-based measures of similarity and dissimilarity. Figure 5 presents the scores of Percentage similarity (*M* = 0.23, *SD* = 0.23) and Balanced abundance (*M* = 0.59, *SD* = 0.25) obtained for each pair of suffixes.

As far as similarity is concerned, most of the trends observed for the abundance data are analogous to those reported for the presence/absence data. On the one hand, a continuum of similarity is still present when frequencies are added to the calculations: almost all suffixes compete — even in very small proportions — and there are no perfect rivals in the sample either. On the other hand, suffixes belonging to the same semantic group (“eventive” or “agentive”) are also more similar to each other than to suffixes included in another group. Unlike what is reported for the incidence-based measures, however, an important difference is observed between the pairs composed of “eventive” or “agentive” suffixes and the pairs contrasting the two types of suffixes. A clear break in similarity scores is found between -*ment*/-*ure* (6th rank) and -*ant*/-*ure* (7th rank), as visible in Figure 5. Based on frequency, the homogeneity and distinctiveness of both “eventive” and “agentive” suffixes is much more salient than when considering only the presence or absence of functions. It appears that the general distinction between the two types of suffixes relies more on the frequent realization of identical functions than on the high proportion of shared functions. The intuition that two groups of competing suffixes can be distinguished is not based on the fact that suffixes within each group have many functions in common, but rather on the fact that most of their derivatives instantiate shared functions.

Just like Balanced richness scores, Balanced abundance scores range between the extremes. The pair -*ant*/-*eur* exhibits a nested pattern (as previously reported for the Sørensen index), hence the null *BA* value, whereas -*ade* and -*eur* obtain the highest score (*BA* = 0.98), as they form almost the same number of derivatives with unshared functions in our sample. Balanced abundance appears to be overall higher than Balanced richness, which indicates that more symmetric overlap is observed when frequencies are taken into consideration. As for the correspondence between Percentage similarity and Balanced abundance measures, they do not seem to be dependent on one another, which is in line with what can be observed for the Sørensen and the Balanced richness measures.

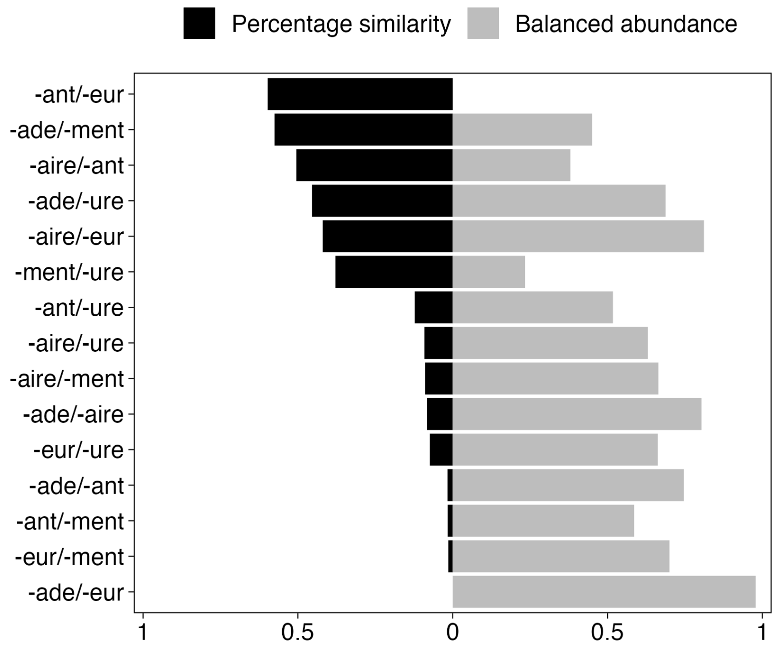

Figure 5. Scores for abundance-based measures (pairs of suffixes are ordered from top to bottom by decreasing similarity)

4.3.4 Comparison of the different measures

The relationship between incidence-based and abundance-based similarity measures can be further investigated. We expect the two to be related but also to capture different aspects of morphological competition, in accordance with what was presented in Section 3.2.1.

The mean Sørensen score (*M* = 0.23, *SD* = 0.15) is roughly equivalent to the mean Percentage abundance one (*M* = 0.23, *SD* = 0.23), although it is associated with less variance. As seen in the previous sections, the Sørensen scores are distributed continuously, whereas the Percentage similarity scores contrasting "eventive" and "agentive" suffixes are globally much lower than those of the pairs composed of the two types of suffixes. Accordingly, including frequencies in the analysis of rivalry situations can make differences between suffixes more salient. It remains that, as illustrated in Figure 6, there is a monotonic, positive relationship between the similarity scores obtained from presence/absence vs. abundance data: the higher the Sørensen score, the higher the Percentage similarity score, and vice versa. A Mantel test[5] confirms that there is indeed a significant correlation between the two similarity matrices (r = 0.868, p < 0.01). Suffix pairs that have a low to average Sørensen score (i.e., up to about 0.25) systematically present a very low Percentage similarity score (i.e., below 0.15), while those that have an average to high Sørensen score (i.e., above 0.30) generally obtain slightly higher Percentage similarity scores (i.e., above 0.35). Overall, the suffixes in our sample that have many functions in common tend to present a relatively similar distribution of derivatives across shared functions.

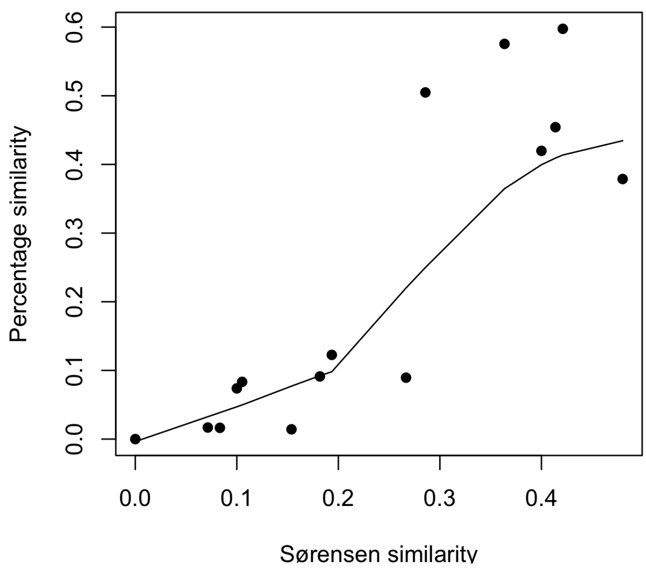

Figure 6. Relationship between the Sørensen and Percentage similarity measures (a polynomial regression line is added to illustrate the general trend)

[5] A classical correlation test should not be used when the assumption of independence is violated in the data, as is the case here: if the similarity score of a given pair of suffixes changes, other similarity scores are modified as well. The Mantel test consists in calculating an initial correlation between two matrices (we used the Spearman method), and then comparing it to many other correlation measures computed from the first matrix and from permutations of the second matrix.

Qualitative similarities and differences between the Sørensen and Percentage similarity measures can also be highlighted by comparing the pairs that are most (or least) similar in each case. Figure 7 shows the rankings obtained for the two measures. While the suffixes *-ment* and *-ure* are the most similar according to the Sørensen index, they lose 5 places in the ranking based on the Percentage similarity measure, meaning that, although they share a high number of functions, they realize them at very different frequencies. Opposite cases can also be observed. The pair *-aire/-ant*, for instance, is ranked only 6th according to the Sørensen coefficient, but gains 3 places in the ranking based on the Percentage similarity measure. The most noticeable differences are observed with pairs involving *-ment*, which can be explained by the high semantic concentration of *-ment* derivatives, as shown in Figure 3. Nouns ending in *-ment* instantiate mostly 2 out of the 9 functions observed for this suffix (Event-TRANSPOSITION and Event.State-TRANSPOSITION). This concentration affects similarity relationships depending on whether the prevalent functions are shared with rival suffixes (as in the case of *-ade*/*-ment*, rising from the 5th to the 2nd rank in Figure 7) or not (as in the case of *-eur*/*-ment*, dropping from the 10th to the 14th rank). Accordingly, the influence of semantic concentration on affix rivalry can be apprehended through the contrast between Percentage similarity and Sørensen indices.

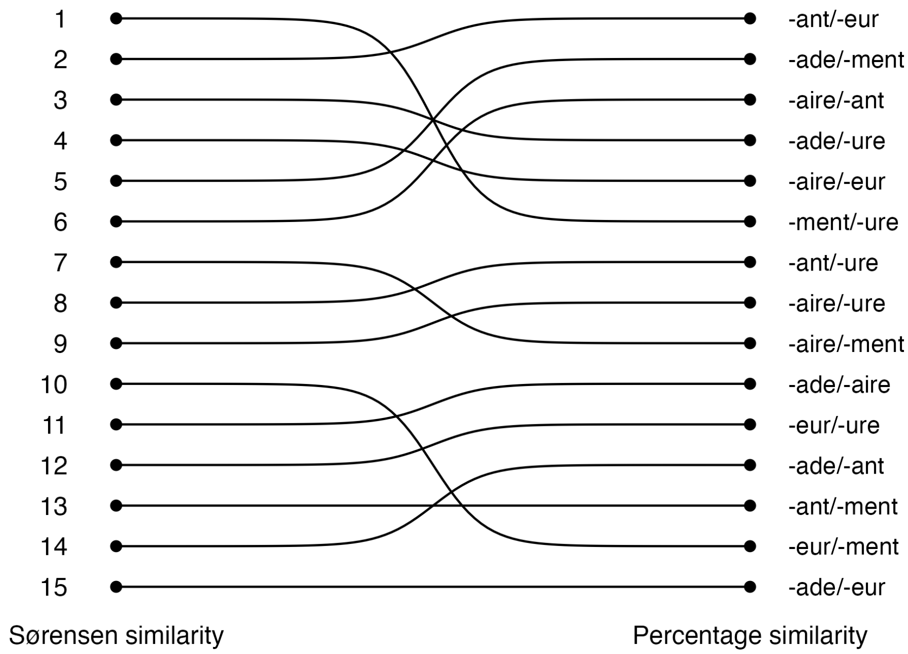

Figure 7. Ranking of the suffix pairs according to the Sørensen vs. Percentage similarity measures

The relationship between the Balanced richness and Balanced abundance measures can be examined as well. As can be seen in Figure 8, the two complementary indices are not correlated, which is confirmed by a Mantel test (p = 0.15). It follows that differences in ranking between dissimilarity measures are more important than between similarity measures, as shown in Figure 9. Many pairs of suffixes change rankings between Balanced richness and Balanced abundance, shifting from more symmetric to less symmetric rivalry (e.g., -*ant*/-*ure* and -*ant*/-*ment* losing 10 and 6 places, respectively) or the opposite (e.g., -*ade*/-*eur* and -*aire*/-*eur* gaining 9 and 5 places, respectively). The significant rank changes observed for -*ant* can be explained by the fact that this suffix has many exclusive functions that are infrequently realized. The important number of exclusive functions allows for counterbalancing with rival suffixes on the one hand, but their low abundance can cause imbalance in the number of derivatives with unshared functions on the other hand. For example, -*ant* and -*ure* have 12 and 13 functions that they do not share with each other (*BR* = 0.92). However, 10 out of the 12 exclusive functions of -*ant* (vs. 6 out of 13 exclusive functions of -*ure*) have only 2 or 3 realizations in our sample, which results in an important difference of abundance for unshared functions (*BA* = 0.52) — hence the shift in ranking observed in Figure 9. By contrast, the difference in ranking observed for pairs with -*eur* can be explained by the fact that -*eur* has very few exclusive functions, but that some of them are abundantly represented in the sample. This specificity favors asymmetric rivalry in terms of presence/absence of functions, but allows for possible compensation in terms of abundance of unshared functions. Such a situation can be observed in the case of -*eur* and -*aire*, which have 2 and 4 unshared functions, respectively (*BR* = 0.5), but since one exclusive function of -*eur* is realized by an important proportion of -*eur* derivatives, the abundance of unshared functions is comparable between the two suffixes (*BA* = 0.81). Overall, it appears that functional capacity and realization frequency have different influences on the dissimilarity structure of rival affixes, which can be precisely analyzed by comparing measures of Balanced richness and Balanced abundance.

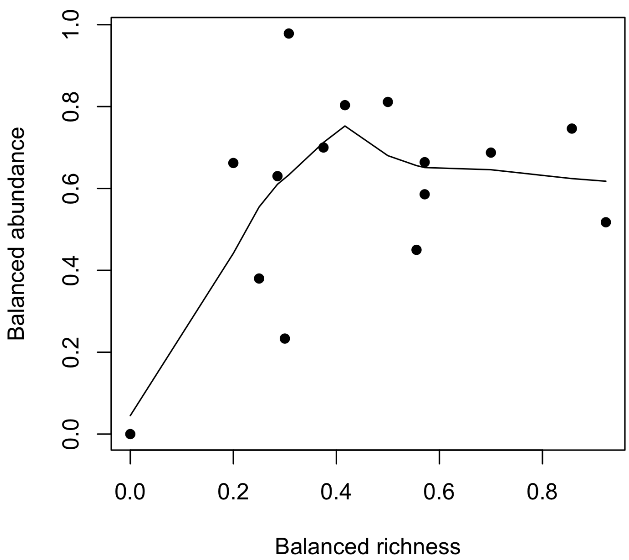

Figure 8. Relationship between the Balanced richness and Balanced abundance measures (a polynomial regression line is added to illustrate the general trend)

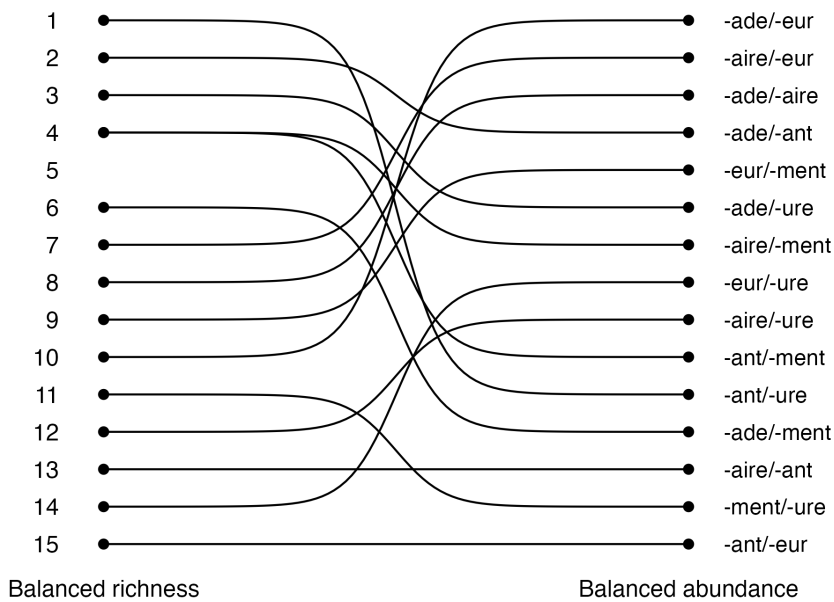

Figure 9. Ranking of the suffix pairs according to the Balanced richness vs. Balanced abundance measures

5. Conclusion

In this study, we explored measures to approach the degree of rivalry between competing affixes. Considering the graded aspects of rivalry related to affix polyfunctionality, we investigated both incidence- and abundance-based measures that depend on the number of semantic functions realized by rival affixes and the number of derivatives instantiating these functions, respectively. We introduced two possible coefficients of similarity between rival affixes and two complementary indices to analyze the dissimilarity between rivals. Through the analysis of simulated data, we presented the potential of these measures before applying them to a sample of French deverbal suffixes.

Incidence- and abundance-based measures represent different aspects of rivalry as they highlight different facets of similarity relationships, and they complement each other accordingly. Incidence-based measures allow a fine-grained investigation of functionality structures. Their interpretation can reveal principles of association between semantic functions that influence the proportion of functions shared between two affixes. As for abundance-based measures, they can weight functional rivalry by realization frequency and consequently shed a different light on the sharing of functions. The comparison between incidence- and abundance-based coefficients of similarity informs on the architecture of rivalries, and on the (in)congruence between the number of shared functions and the number of derivatives that instantiate these functions. In the sample of French deverbal suffixes we examined, incidence- and abundance-based similarities are highly correlated, but some idiosyncrasies can be observed, especially when suffixes share a moderate proportion of functions but form very few derivatives instantiating these functions.

Indices of balanced distribution for unshared functions provide further information about the dissimilarity between rival affixes. Nestedness, overlap and asymmetry between rivals can be quantified and described by means of specific measures with respect to both functional structure and realization frequency of unshared functions. In the case of French deverbal suffixes, no correlation is observed between incidence- and abundance-based measures of balanced distribution. In particular, a balanced distribution of distinctive derivatives is not necessarily associated with a balanced distribution of unshared functions. Two deverbal suffixes can form similar proportions of derivatives instantiating unshared functions independently of whether these functions are evenly distributed or not among the suffixes (i.e., independently of functional symmetry). Such a divergence confirms that a complete analysis of the dissimilarity between rival affixes requires to take both their functional richness and the distribution of their derivatives into consideration.

Overall, measuring affix rivalry offers the possibility to finely analyze rivalry situations and to compare them both within languages and cross-linguistically. The metrics we proposed in this paper should be considered a first step towards a comprehensive measurement of morphological competition. Further refinement is needed to take into consideration additional factors such as the productivity of word-formation processes, i.e., their capacity to produce new words (see e.g., Corbin 1987; Plag 1999; Bauer 2001; Fernández-Domínguez 2013). As measures of similarity or dissimilarity, the indices we presented are synchronic in nature and do not account for the diachronic evolution and change in productivity that can affect rivalry in the long run. They do not inform about the availability of an affix when coining new words at a given point in time. A certain degree of similarity between two affixes could be observed in the lexicon due to historical word formation, without the two affixes competing in contemporary lexical innovation. To improve the assessment of rivalry, similarity indices could be examined diachronically over different periods of time. They may also be combined with measures of potential or expanding productivity (Baayen 1993, 2009), which can account for the vitality of a word formation process and its contribution to the renewal of the lexicon.

Acknowledgements

We are grateful to the anonymous reviewer and the editors for their helpful comments on an earlier version of this paper. All remaining errors are our own. This work was supported by the Swiss National Science Foundation under Grant 100012_188782 ("The semantics of deverbal nouns in French").

Supplementary material

The dataset, the annotation guide and the statistical script used in this study are available at https://github.com/deverbal-nouns/affix-rivalry.

References

Anderson, Marti J., Crist, Thomas O., Chase, Jonathan M., Vellend, Mark, Inouye, Brian D., Freestone, Amy L., Sanders, Nathan J., Cornell, Howard V., Comita, Liza S., Davies, Kendi F., Harrison, Susan P., Kraft, Nathan J. B., Stegen, James C. & Swenson, Nathan G. 2011. Navigating the multiple meanings of β diversity: A roadmap for the practicing ecologist. *Ecology Letters* 14(1): 19-28. https://doi.org/10.1111/j.1461-0248.2010.01552.x10.1111/j.1461-0248.2010.01552.x

Arndt-Lappe, Sabine. 2014. Analogy in suffix rivalry: The case of English -*ity* and *-ness*. *English Language & Linguistics* 18(3): 497-548. https://doi.org/10.1017/S136067431400015X

Aronoff, Mark. 2020. *-less* and *-free*. In *Complex Words: Advances in Morphology*, Lívia Körtvélyessy & Pavol Štekauer (eds), 55-64. Cambridge: CUP.

Asher, Nicholas. 2011. *Lexical Meaning in Context: A Web of Words.* Cambridge: CUP.

Baayen, R. Harald. 1993. On frequency, transparency, and productivity. In *Yearbook of Morphology 1992*, Geert E. Booij & Jaap van Marle (eds), 181-208. Dordrecht: Kluwer.

Baayen, R. Harald. 2009. Corpus linguistics in morphology: Morphological productivity. In *Corpus Linguistics: An International Handbook*, Anke Lüdeling & Merja Kytö (eds), Vol. 2, 899-919. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110213881.2.899

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

Baselga, Andrés. 2013. Separating the two components of abundance-based dissimilarity: Balanced changes in abundance vs. abundance gradients. *Methods in Ecology and Evolution* 4(6): 552-557. https://doi.org/10.1111/2041-210X.12029

Bauer, Laurie. 2001. *Morphological Productivity*. Cambridge: CUP.

Bauer, Laurie. 2002. What you can do with derivational morphology. In *Morphology 2000: Selected Papers from the 9th Morphology Meeting, Vienna, 24–28 February 2000*, Sabrina Bendjaballah, Wolfgang U. Dressler, Oskar E. Pfeiffer & Maria D. Voeikova (eds), 37-48. Amsterdam: John Benjamins.

Bonami, Olivier & Thuilier, Juliette. 2019. A statistical approach to rivalry in lexeme formation: French *-iser* and *-ifier*. *Word Structure* 12(1): 4-41. https://doi.org/10.3366/word.2018.0130

Corbin, Danielle. 1987. *Morphologie dérivationnelle et structuration du lexique*. Tübingen: Niemeyer.

Cruse, D. Alan. 1995. Polysemy and related phenomena from a cognitive linguistic viewpoint. In *Computational Lexical Semantics*, Patrick Saint-Dizier & Evelyne Viegas (eds), 33-49. Cambridge: CUP.

Denistia, Karlina, Shafaei-Bajestan, Elnaz & Baayen, R. Harald. 2021. Exploring semantic differences between the Indonesian prefixes *PE-* and *PEN-* using a vector space model. *Corpus Linguistics and Linguistic Theory* 18(3): 573-598. https://doi.org/10.1515/cllt-2020-0023

Díaz-Negrillo, Ana. 2017. On the identification of competition in English derivational morphemes. The case of *-dom*, *-hood* and *-ship*. In *Competing Patterns in English Affixation*, Juan Santana-Lario & Salvador Valera (eds), 119-161. Lausanne: Peter Lang.

Dubois, Jean. 1962. *Étude sur la dérivation suffixale en français moderne et contemporain: essais d'interprétation des mouvements observés dans le domaine de la morphologie des mots construits.* Paris: Larousse.

Fábregas, Antonio. 2010. A syntactic account of affix rivalry in Spanish nominalizations. In *The Semantics of Nominalizations across Languages and Frameworks*, Artemis Alexiadou & Monika Rathert (eds), 67-92. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110245875.67

Fernández-Domínguez, Jesús. 2013. Morphological productivity measurement: Exploring qualitative versus quantitative approaches. *English Studies* 94(4): 422-447. https://doi.org/10.1080/0013838X.2013.780823

Fernández-Domínguez, Jesús. 2017. Methodological and procedural issues in the quantification of morphological competition. In *Competing Patterns in English Affixation*, Juan Santana-Lario & Salvador Valera (eds), 67-117. Lausanne: Peter Lang.

Flaux, Nelly & Van de Velde, Danièle. 2000. *Les noms en français: esquisse de classement.* Paris: Editions Ophrys.

Fradin, Bernard. 2019. Competition in derivation: What can we learn from French doublets in *-age* and *-ment*? In *Competition in Inflection and Derivation*, Franz Rainer, Francesco Gardani, Wolfgang U. Dressler & Hans Christian Luschützky (eds), 67-93. Cham: Springer.

Fradin, Bernard & Kerleroux, Françoise. 2003. Quelle base pour les procédés de la morphologie constructionnelle? *Silexicales* 3: 76-84.

Godard, Danièle & Jayez, Jacques. 1996. Types nominaux et anaphores : le cas des objets et des événements. In *Anaphores temporelles et (in-)cohérence* [Cahiers Chronos 1], Walter De Mulder, Liliane Tasmowski-De Ryck & Carl Vetters (eds), 41-58. Amsterdam: Rodopi.

Guzmán Naranjo, Matías & Bonami, Olivier. 2023. A distributional assessment of rivalry in word formation. *Word Structure* 16(1): 87-114. https://doi.org/10.3366/word.2023.0222

Haas, Pauline, Barque, Lucie, Huyghe, Richard & Tribout, Delphine. 2023. Pour une classification sémantique des noms en français appuyée sur des tests linguistiques. *Journal of French Language Studies* 33(1): 52-81. https://doi.org/10.1017/S0959269522000187

Huyghe, Richard. 2015. Les typologies nominales: présentation. *Langue française* 185(1): 5-27.

Huyghe, Richard & Wauquier, Marine. 2021. Distributional semantics insights on agentive suffix rivalry in French. *Word Structure* 14(3): 354-391. https://doi.org/10.3366/word.2021.0194

Huyghe, Richard, Lombard, Alizée, Salvadori, Justine & Schwab, Sandra. 2023. Semantic rivalry between French deverbal neologisms in *-age*, -

ion and -*ment*. In *The Semantics of Derivational Morphology. Theory, Methods, Evidence*, Sven Kotowski & Ingo Plag (eds), 125-158. Berlin: De Gruyter. https://doi.org/10.1515/9783111074917-006

Huyghe, Richard & Varvara, Rossella. 2023. Affix rivalry: Theoretical and methodological challenges. *Word Structure* 16(1): 1-23. https://doi.org/10.3366/word.2023.0218

Jacquey, Evelyne. 2006. Un cas de « polysémie logique » : modélisation de noms d'action en français ambigus entre processus et artefact. *TAL* 47(1): 137-166.

Kipper-Schuler, Karin. 2005. *VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon*. PhD dissertation, University of Pennsylvania.

Legendre, Pierre. 2014. Interpreting the replacement and richness difference components of beta diversity. *Global Ecology and Biogeography* 23(11): 1324-1334.

Legendre, Pierre & Legendre, Louis. 2012. *Numerical Ecology*. Third edition. Amsterdam: Elsevier.

Lieber, Rochelle. 2015. The semantics of transposition. *Morphology* 25(4): 353-369. https://doi.org/10.1007/ s11525-015-9261-4

Lieber, Rochelle. 2016. *English Nouns: The Ecology of Nominalization*. Cambridge: CUP.

Lieber, Rochelle & Plag, Ingo. 2022. The semantics of conversion nouns and -*ing* nominalizations: A quantitative and theoretical perspective. *Journal of Linguistics* 58(2): 307-343. https://doi.org/10.1017/S0022226721000311

Lignon, Stéphanie. 2007. Les noms de spécialistes en -*iste* et en -*ien* : le chimiste perturbé ou comment le physicien se réajuste. In *Perturbations et réajustements: langue et langage*, Béatrice Vaxelaire, Rudolph Sock, Georges Kleiber & Fabrice Marsac (eds), 287–296. Strasbourg: Presses Universitaires de l'Université Marc Bloch.

Lindsay, Mark & Aronoff, Mark. 2013. Natural selection in self-organizing morphological systems. In *Morphology in Toulouse. Selected Proceedings of Décembrettes 7 (Toulouse, 2-3 December 2010)*, Nabil Hathout, Fabio Montermini & Jesse Tseng (eds), 133-153. München: Lincom Europa.

Liu, Zhenyuan, Heino, Jani, Soininen, Janne, Zhou, Tingting, Wang, Weimin, Cui, Yongde, Chen, Yushun, Li, Zhengfei, Zhang, Junqian & Xie, Zhicai. 2022. Different responses of incidence-weighted and abundance-weighted multiple facets of macroinvertebrate beta diversity to urbanization in a subtropical river system. *Ecological Indicators* 143: 1-11. https://doi.org/10.1016/j.ecolind.2022.109357

Martin, Fabienne. 2010. The semantics of eventive suffixes in French. In *The Semantics of Nominalizations across Languages and Frameworks*, Artemis Alexiadou & Monika Rathert (eds), 109-140. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110226546.109

Mel′čuk, Igor A. 1994. *Cours de morphologie générale*. Montréal: Presses de l'Université de Montréal.

Melloni, Chiara. 2011. *Event and Result nominals*. Bern: Peter Lang.

Murphy, Elliott. 2021. *Linguistic Representation and Processing of Copredication.* PhD dissertation, University College London.

Naccarato, Chiara. 2019. Agentive (para)synthetic compounds in Russian: A quantitative study of rival constructions. *Morphology* 29(1): 1-30. https://doi.org/10.1007/s11525-018-9330-6

Nagano, Akiko. 2023. Affixal rivalry and its purely semantic resolution among English derived adjectives. *Journal of Linguistic*s 59(3): 499-530. https://doi.org/10.1017/S0022226722000147

Odum, Eugene P. 1950. Bird populations of the Highlands (North Carolina) Plateau in relation to plant succession and avian invasion. *Ecology* 31(4): 587-605. https://doi.org/10.2307/1931577

Petukhova, Volha & Bunt, Harry. 2008. LIRICS semantic role annotation: Design and evaluation of a set of data categories. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis & Daniel Tapias (eds), 39-45. Paris: European Language Resources Association.

Plag, Ingo. 1999. *Morphological Productivity: Structural Constraints in English Derivation*. Berlin: Mouton de Gruyter.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA: The MIT Press.

Rainer, Franz. 2015. Agent and instrument nouns. In *Word-Formation: An International Handbook of the Languages of Europe*, Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer, Vol. 2, 1304-1316. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110246278-030

Roché, Michel. 2004. Mot construit? Mot non construit? Quelques réflexions à partir des dérivés en *-ier(e)*. *Verbum* 26(4): 459-480.

Roswell, Michael, Dushoff, Jonathan & Winfree, Rachael. 2021. A conceptual guide to measuring species diversity. *Oikos* 130(3): 321-338. https://doi.org/10.1111/oik.07202

Salvadori, Justine & Huyghe, Richard. 2023. Affix polyfunctionality in French deverbal nominalizations. *Morphology* 33(1): 1-39. https://doi.org/10.1007/s11525-022-09401-4

Säily, Tanja. 2011. Variation in morphological productivity in the BNC: Sociolinguistic and methodological considerations. *Corpus Linguistics and Linguistic Theory* 7(1): 119-141. https://doi.org/10.1515/cllt.2011.006

Schäfer, Roland. 2015. Processing and querying large web corpora with the COW14 architecture. In *Proceedings of Challenges in the Management of Large Corpora 3 (CMLC-3)*, Piotr Bański, Hanno Biber, Evelyn

Breiteneder, Marc Kupietz, Harald Lüngen & Andreas Witt (eds), 28-34. Mannheim: Institut für Deutsche Sprache.

Schäfer, Roland & Bildhauer, Felix. 2012. Building large corpora from the Web using a new efficient tool chain. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds), 486-493. Paris: European Language Resources Association.

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

Schnedecker, Catherine & Aleksandrova, Angelina. 2016. Les noms d'humains en *-aire*: essai de classification. In *5e Congrès Mondial de Linguistique Française - CMLF 2016*, Franck Neveu, Gabriel Bergounioux, Marie-Hélène Côté, Jean-Michel Fournier, Linda Hriba & Sophie Prévost (eds), 12001. Paris: Institut de Linguistique Française. https://doi.org/10.1051/shsconf/20162712001

Schulte, Marion. 2015. Polysemy and synonymy in derivational affixation: A case study of the English suffixes *-age* and *-ery*. *Morphology* 25(4): 371-390. https://doi.org/10.1007/s11525-015-9271-2

Sørensen, Thorvald A. 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *Kongelige Danske Videnskabernes Selskabs Biologiske Skrifter* 5: 1-34.

ten Hacken, Pius. 2015. Transposition and the limits of word formation. In *Semantics of Complex Words*, Laurie Bauer, Lívia Körtvélyessy & Pavol Štekauer (eds), 187-216. Cham: Springer. https://doi.org/10.1007/978-3-319-14102-2_10

Thiele, Johannes. 1987. *La formation des mots en français moderne.* Montréal: Presses de l'Université de Montréal.

Thuilier, Juliette, Tribout, Delphine & Wauquier, Marine. 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

Varvara, Rossella. 2020. Constraints on nominalizations: Investigating the productivity domain of Italian *-mento* and *-zione*. *Zeitschrift für Wortbildung / Journal of Word Formation* 4(2): 78-99. https://doi.org/10.3726/zwjw.2020.02.05

Varvara, Rossella, Lapesa, Gabriella & Padó, Sebastian. 2021. Grounding semantic transparency in context: A distributional semantic study of German event nominalizations. *Morphology* 31(4): 409-446. https://doi.org/10.1007/s11525-021-09382-w

Wilson, J. Bastow. 2012. Species presence/absence sometimes represents a plant community as well as species abundances do, or better. *Journal of Vegetation Science* 23(6): 1013-1023. https://doi.org/10.1111/j.1654-1103.2012.01430.x