

Affix rivalry: Theoretical and methodological challenges

Richard Huyghe, Rossella Varvara

University of Fribourg

Abstract

Affix rivalry is a key element in the organization and evolution of derivational systems. Its study provides insight into word-formation semantics, morphological change, productivity, lexical variation, and the many-to-many relationship between form and meaning. Affix rivalry is characterized by its gradient nature and the multiplicity of factors that can be involved in its resolution. On the one hand, rival affixes can be regarded as more or less competing depending on their semantic similarity and on the proportion of semantic functions they have in common. On the other hand, the distinction between rival affixes can rely on different linguistic properties, including not only semantic, but also phonological, morphological, syntactic, stylistic, and sociolinguistic properties. Differences are observed as tendencies and have a variable influence on affix selection. Quantitative methods can be used to precisely assess degrees of rivalry and multifactorial resolution of competition. Based on the statistical analysis of large sets of derivatives, they can provide an accurate description of affix rivalry and help us better understand the mechanisms of competition in word formation.

Keywords: affix rivalry, quantitative methods, word formation, derivational semantics, data-driven approaches

1 Introduction

Competition between linguistic forms is an essential element in the organization of linguistic systems and in the development of natural languages. It can be observed in all domains of grammar and affects both individual speakers' production and general linguistic structures. By contributing to the delineation of grammatical functions, it plays an important role in the change, adaptation, emergence or extinction of linguistic items and constructions.

The morphological aspects of linguistic competition have been the object of constant interest in the history of linguistics, as described in detail by Gardani et al. (2019). Many studies in morphology have addressed the rivalry between word formation processes, especially with respect to affix selection. As a particular case of morphological competition, affix rivalry not only calls for refined analyses of affixes, but also provides a perspective on productivity,

morphological change, derivational semantics, sociolinguistic lexical variation, and the many-to-many relationship between form and meaning. In the existing literature, affix rivalry is often investigated theoretically, sometimes through large data sets, but not necessarily with the support of advanced quantitative methods. However, recent studies have explored the use of quantitative techniques to account for situations of affix rivalry, based on inferential statistics and computational methods (see Baayen et al. 2013, Arndt-Lappe 2014, Bonami & Thuilier 2019, Naccarato 2019, Varvara 2020, Huyghe & Wauquier 2021, Lieber & Plag 2022, among others). Quantitative approaches are particularly suitable for the description of affix rivalry, given both its inherent gradience and the multiplicity of factors that can be involved in its resolution.

On the one hand, rival affixes can be regarded as more or less competing depending on their semantic proximity and on the proportion of semantic functions they have in common. On the other hand, the distinction between rival affixes can rely on different linguistic properties, including not only semantic, but also phonological, morphological, syntactic, stylistic, and sociolinguistic properties. Differences are observed as tendencies and their respective influence on affix selection is variable, which requires appropriate evaluation. Quantitative methods can be used to precisely assess degrees of rivalry and multifactorial resolution of competition. By providing elements for gradient as opposed to categorical descriptions, they offer a means of investigating affix rivalry that is consistent with its proper nature and can ultimately help us better understand morphological competition.

This special issue aims to present quantitative approaches to affix rivalry, through a selection of studies that investigate different aspects of competition in derivation. These studies use a variety of techniques to analyze specific cases of rivalry in different languages, and illustrate the kind of results that can be achieved with quantitative methods. They also contribute to the theoretical reflection on the structures of affix rivalry and its multidimensional resolution. In this introductory article, we first discuss the definition of affix rivalry and the discriminative properties of rival affixes as they have been investigated in the abundant literature on morphological competition. Then we present how quantitative methods have been applied to the study of affix rivalry. Finally, we introduce the different contributions assembled in the special issue.

2 Defining affix rivalry

The definition of affix rivalry is rarely made explicit in morphological studies. However, it is not self-evident and may be subject to important variations, potentially hindering the identification of rival affixes. In this section, we present some of the issues related to the definition of affix rivalry and the possible distinction between different forms of rivalry. Note that although rivalry is observed and discussed for inflectional affixes (see e.g. Carstairs-McCarthy 1994, Thornton 2011, 2012, Merkuur et al. 2019, Stump 2019), it is mostly investigated in the case of derivational affixes, on which we focus here.

2.1 Semantic equivalence

In first approximation, affix rivalry can be defined as the relationship between two or more affixes that can be used to form the same semantic type of words. Such a definition raises a number of issues that require further discussion. Nevertheless, a consensual assumption is that affix rivalry should be defined on a semantic basis. While many structural or non-structural properties can be associated with affixes or derivational patterns, we did not find any study in the literature on rivalry that would consider as competing affixes that have some properties in common but not the semantic ones. For example, two affixes that are associated with the same phonological constraints or with the same stylistic marking are not considered as rival affixes unless the condition of semantic equivalence is satisfied. The reverse is not true.

The concept of semantic equivalence needs to be clarified when applied to affixes. It can have different implications depending on theoretical approaches to derivation, especially with respect to morphological input. If affixes are not considered as having a meaning in themselves but as phonological exponents of derivational patterns, or more generally as performing semantic operations that specify constraints on base selection, then equivalence should apply to all semantic elements involved in word-formation processes. Not only semantic properties of morphological output should be equivalent, but also those of the input, be it a single lexeme or a morphological family in a paradigmatic perspective. However, to our knowledge, very few studies take this as a necessary condition for affix rivalry. Equivalence is generally not required for input semantics, but only for output semantics, which appears by default as a sufficient condition for defining rival affixes across theoretical frameworks. What is usually considered though—and probably consistent with the definition of rival affixes as having the same output semantics—is that rival affixes operate on the same lexical class or classes of inputs and form the same lexical class of outputs.¹

Another issue to be discussed is the appropriate granularity to assess semantic equivalence. The identification of rival affixes depends on the precision with which semantic types are specified in word formation. Not only are semantic types less clearly defined than syntactic or inflectional classes, but the degree of semantic specification in derivational processes is also uncertain. It is known that derivational semantics (i.e. the semantic operations associated with derivational processes) is underspecified with respect to lexical semantics (i.e. the meaning of words as stored in the lexicon). Since the semantic properties of derivational processes can only be observed through their lexical realization, the level of abstraction at which derivational semantics properties can be inferred from the meaning of derivatives needs to be clearly defined. In theory, semantic equivalence between affixes should be established with the finest possible grain attached to derivational processes. However, the

¹ The question of whether semantic types are equivalent or not when realized by different lexical classes is beyond the scope of this paper. Nevertheless one could argue that lexical meanings are dependent on lexical classes (e.g. the lexical meaning of ‘property’ is not exactly the same when realized by nouns or adjectives, due to different relations to predication and reference). Consequently, the definition of rival affixes as forming the same semantic type of words would entail that these belong to the same lexical class.

precision of the semantic operations involved in derivation is still under debate, which can cause problems when identifying rival affixes based on semantic equivalence. For example, two suffixes in a given language could apparently compete in the formation of change-of-state verbs, but not be considered as strict competitors if they are actually used to form different subtypes of change-of-state verbs and if these subtypes are specified in derivational processes. In studies on rivalry, semantic equivalence is usually assessed with medium- or coarse-grained meanings, without addressing whether this is the appropriate granularity to account for derivational semantics. Further, the issue of semantic granularity can interfere with the question of how strict semantic equivalence should be in the definition of rival affixes and with the distinction of different forms of rivalry, as discussed in the next section.

2.2 Forms of rivalry

Semantic equivalence as a defining feature of affix rivalry can be questioned when reviewing studies on morphological competition. Whereas some authors explicitly refer to rival affixes as strictly equivalent semantically, others implicitly admit that they may to a certain extent differ in meaning. For example, Plag (1999: 227) defines rival morphological processes as “semantically identical”, Fábregas (2010: 67) considers affixes to be rivalling when they “give identical results”, and Fradin (2019: 68) asserts that morphological competition occurs when distinct exponents are “correlated with a unique semantic content”. On the contrary, Aronoff (1976: 38-39) claims that English rival suffixes *-ity* and *-ness* differ in semantic coherence and diversity, and many studies investigate possible semantic differences between rival affixes (see e.g. Martin 2010, Schulte 2015, Naccarato 2019, Nagano 2022). The initial definition of rival affixes as forming words of the same semantic type seems too narrow to account for all cases regarded as rivalling and needs to be adjusted accordingly.

There are two ways in which affixes can be considered rivals while not being strictly identical semantically. The first one is to assume that affixes with similar semantic functions can compete in the formation of derivatives. Rival affixes would not necessarily form the same semantic type of derivatives but could also be characterized by semantic resemblance between derivatives. From that perspective, affix rivalry can be defined as the relationship between two or more affixes that are used to derive words of identical or similar semantic types. As a corollary, one can distinguish between ‘relative’ and ‘absolute’ rivalry depending on whether semantic similarity or identity is observed. Some researchers may wish to maintain a categorical distinction between rival and non-rival affixes even in the case of relative rivalry. However, such a distinction raises the difficult question of how to define a threshold of similarity that separates rival from non-rival affixes. More consistently, the correlate of a definition in terms of semantic similarity should be the existence of a continuum of rivalry, ranging from no rivalry to absolute rivalry.

The second possibility that two rival affixes are not strictly equivalent depends on their inherent polyfunctionality. Many affixes are assigned multiple semantic functions, which adds a variable feature to affix rivalry. Rivalry can be characterized as ‘partial’ or ‘total’ depending on whether it occurs between all semantic functions of competing affixes, or only between

some of them (Huyghe & Wauquier 2021, Guzmán Naranjo & Bonami 2023). Affix rivalry can then be more explicitly defined as the relationship between two or more affixes that, in at least some of their uses, can form words of identical or similar semantic types.

Relative/absolute and partial/total forms of rivalry are logically compatible, but may not be easy to distinguish, mostly because of uncertainty about the semantic granularity of word-formation processes. Two competing affixes can be evaluated as relative or absolute rivals, or as partial or total rivals, according to the degree of semantic precision assigned to derivational patterns. If the semantic analysis is too fine-grained with respect to derivational semantics, then absolute or total rivals can be mistaken for relative or partial rivals. If the analysis is too coarse-grained, then relative or partial rivals can be mistaken for absolute or total rivals. Furthermore, when a small semantic difference is observed between two affixes, it may be unclear whether it reveals non-absolute rivalry or non-competing functions, possibly related to partial rivalry if the two affixes have other semantic functions in common. One may also wonder whether total absolute rival affixes are attested. The existence of affixes with strictly equivalent functions seems difficult to demonstrate. For example, Plag (1999, 2000) argues that verb-forming suffixes *-ize* and *-ify* in English compete for all of their output meanings, but cannot find attested examples of the more marginal similative meaning (i.e. 'do/act like x') with *-ify*. Note that evidence for the existence of total absolute rivals should be based on the analysis of all cases of derivation involving two rival affixes, whereas one recurring example of semantic divergence is sufficient to support relative or partial rivalry, which undoubtedly makes total absolute rivalry more difficult to establish.

The distinctions between relative/absolute and partial/total rivalry both imply a gradient notion of rivalry. On the one hand, relative rivals can be regarded as more or less competing depending on how similar their semantic outputs are. On the other hand, partial rivalry varies according to (i) the proportion of shared output meanings and (ii) the frequency of lexical realization of these meanings, i.e. how often affixes are used to derive words with competing semantic types. Overall, the semantic differences observed between affixes determine various degrees of rivalry that should be analyzed and quantified as such.

3 Resolution of rivalry

Under the traditional view, no two linguistic forms with the same function can persist in language, and in the long run, differences are expected to emerge between competing forms. Most studies on affix rivalry focus on how rival affixes are distinguished from each other synchronically or diachronically, and how situations of rivalry can be resolved accordingly. In that respect, affix rivalry is often used as an epistemological tool for approaching various aspects of affix distinctiveness and refining the analysis of affix properties. In this section, we present which properties can be discriminative and how the resolution of rivalry relates to diachrony and morphological productivity.

3.1 Discriminative properties

Rival affixes can be distinguished by structural and non-structural properties. Semantic differences can be observed between derivatives formed with competing affixes, attesting to relative or partial rivalry. For example, Cartoni et al. (2015) report that agentive suffixes in both French and Italian vary with respect to the type of agents denoted. Nagano (2022) argues that English adjectives ending in *-ed* describe properties that are gradable on a closed scale, whereas adjectives ending in *-y* are open-scaled. Some aspects of polyfunctionality can play an important role in the differentiation of semantic outputs. Schulte (2015) shows that although *-age* and *-ery* in English mostly derive words of the same semantic types, they differ in both frequency and interconnection of meanings. For instance, *-ery* forms relatively more location-denoting nouns and more polysemous nouns with both an action and a collective meaning than *-age*. Variation with respect to base selection or to the semantic relationship between bases and derivatives is also reported. According to Aronoff and Cho (2001), *-ship* differs from *-hood* in English in that it selects only stage-level nouns, whereas both stage-level and individual-level nouns can be used as a base with *-hood*. In their study on competition between deverbal nouns ending in *-ung* and nominalized infinitives in German, Varvara et al. (2021) find significant differences of semantic transparency between bases and derivatives, establishing that converted infinitives are more similar to their base than nouns suffixed with *-ung*.

Non-semantic properties can contribute to the distinction between rival affixes. A number of studies investigate possible phonological constraints on morphological bases depending on the affix. For example, Plag (1999) argues that verb-forming *-ize* and *-ify* can be considered as complementary suffixes in English, *-ize* being preferentially used with trochaic and dactylic bases and *-ify* with iambic and monosyllabic bases. Booij (2002) shows that Dutch agentive suffixes *-der* and *-aar* differ from the more common *-er* in that *-der* is selected by bases ending in /r/, and *-aar* by bases ending in a coronal sonorant consonant preceded by a schwa. Rival affixes can also find a morphological niche in which they can develop distinctive properties, as shown by Lindsay and Aronoff (2013) for English adjective-forming suffixes *-ic* and *-ical*. Although *-ical* is less productive than *-ic* in contemporary English, it is highly preferred over *-ic* with stems ending in *-olog-*, which can be seen as a reason for its persistence in the derivational system. Similarly, Missud and Villoing (2020) assert that *-age*, *-ion* and *-ment* in French are differently distributed with denominal and deadjectival verbs, and favoured by converted, suffixed and prefixed bases, respectively. Syntactic properties may also play a role in the distinction between competing processes, as illustrated by Schirakowski (2020) who explores differences between suffixed nominalizations and nominalized infinitives in Spanish, most notably with respect to argument realization. Schirakowski identifies a preference of suffixed deverbal nouns for the construction with internal arguments, and of nominalized infinitives for the construction with external arguments.

Competing affixes can also be distinguished based on stylistic or sociolinguistic factors. According to Lehrer (2000), weakly productive agentive suffixes *-ster*, *-eer* and *-eur* in English exhibit small differences in connotation, which may not be systematically realized but can be

salient in neologisms. The suffix *-ster* may be the most negatively connoted one, whereas *-eer* can convey a depreciatory judgement on proactive agents, and *-eur* is supposedly associated with skillfulness. Variation in frequency and productivity across registers has been scrutinized as well. For example, Guz (2009) examines the distribution of nouns ending in *-ity* and *-ness* in the *British National Corpus* and finds that *-ness* is preferred in fiction texts and *-ity* in academic prose, a result the author relates to differences in formality, technical meaning, and morphological composition. Some studies have investigated the potential influence of individual and social factors. As far as rivalry between *-ity* and *-ness* is concerned, Romaine (1983) demonstrates the existence of lexical preferences based on age and gender. Asking participants in an experiment their intuitions about the acceptability of possible words ending in *-ity* and/or *-ness*, she observes that younger speakers accept more nouns with both *-ness* and *-ity* than older speakers, and that men accept more *-ity* only and more *-ity* and *-ness* forms than women. As a complement, Säily (2011) shows that women tend to use *-ity* less productively in their writing than men, whereas no difference is observed for *-ness*. It also appears that gender interacts with socio-economic status in spoken language, with lower-class women using *-ness* with less lexical diversity than lower-class men, whereas no difference is observed for women in general.

Two important characteristics of discriminative properties should be mentioned. The first one is that differences are mostly observed as tendencies, through variation in realization frequency. Studies based on solid empirical ground rarely report on clear-cut differences between rival affixes, but on the propensity for certain uses. Affixes can be regarded as more or less distinguishable depending on how frequently they instantiate discriminative properties, and how representative of their specific uses these properties are. It follows that distinctive properties should be considered in a scalar approach and that the significance of the observed differences should be statistically tested. The second point is that the different kinds of discriminative properties are not mutually exclusive. A complete picture of rivalry resolution requires a holistic view on the possible differences between affixes. Some of them are examined together in the literature. For example, Bonami and Thuilier (2019) investigate phonological, morphological and semantic factors to account for the rivalry between *-iser* and *-ifier* in French. Varvara (2020) examines the competition between nominalizers *-mento* and *-zione* in Italian in light of morphological and syntactic properties of the verbal bases, as well as the frequency of bases and derivatives. In her study of agentive (para)synthetic compounds ending in *-ec* and *-tel'* in Russian, Naccarato (2019) examines the role of syntactic and semantic properties of bases and derivatives in the selection of the suffix, but also the distribution of derivatives across textual genres. Dressler et al. (2019) explore how both variation in morphosemantic transparency and speakers' age can affect the use of diminutives in German and Italian. A full assessment of the differentiation between rival affixes needs not only to identify distinctive properties but also to evaluate their relative importance in affix resolution.

In order to investigate the differences between competing affixes, many studies focus on doublets, i.e. (nearly) synonymous words derived from the same base but with different

affixes (see e.g. Gries 2001 for *-ic* and *-ical* in English, Rodrigues 2015 for *-da* and *-mento* in Portuguese, Fradin 2019 for *-age* and *-ment* in French, Aronoff 2020 for *-less* and *-free* in English, Radimský and Štichauer 2021 for *-mento* and *-zione* in Italian). Because they neutralize base differences, attested doublets seem like ideal minimal pairs to identify distinctive properties between competing derivational processes. Nevertheless, conclusions should be cautiously drawn from the observation of doublets. First, since lexical competition can generate idiosyncrasies that pertain to lexical semantics, but not to derivational semantics, only recurring differences between doublets can be exploited. Second, doublets may not reveal all existing differences between rival processes, precisely because of possible discriminative properties that do not license the formation of competing lexemes.

A general question about doublets is whether they are necessarily observed in case of affix rivalry. Morphological competition does not logically entail the existence of doublets in the lexicon. In the case of total absolute rivalry, pertinent morphological bases could be randomly selected by one of the rival affixes with no overlap due to lexical blocking. In the case of relative or partial rivalry, the contingency of onomasiological needs and lexicalization can make the existence of doublets accidental and to some extent unpredictable. However, in reality there do not seem to be cases of strong affix rivalry that do not generate doublets, and one may wonder whether, beyond lexical contingencies, there is a positive correlation between the degree of rivalry of two morphological processes and the number of doublets they produce—which as a corollary could be an indication of how competing two processes are.

3.2 Diachrony and productivity

Time is an important factor influencing affix rivalry, orthogonally to those mentioned in the previous section. The internal and external properties that possibly differentiate competing derivational processes can evolve through time, and relations of rivalry can vary accordingly. The diachronic evolution of affixes can lead to the emergence or resolution of situations of rivalry by determining changes in affix usage and functionality. For example, Uth (2010) observes that the suffix *-age* in French has been increasingly used since the 18th century to form nouns with an eventive meaning, which reinforces its rivalry with the suffix *-ment*, widely available since Old French to form event-denoting nouns. Rivalry situations can also result from the appearance of a new affix in a language. According to Hegedűs (2014), new affixes can be borrowed from other languages, possibly with the same etymology as that of a competing affix in the target language (e.g. French *-esque* borrowed in English to compete with *-ish*, while descending from the same Proto-Indo-European morpheme **-isko-*). They can also be internally formed, as in the case of the high degree prefix *hypra-* in French, presumably blended from *hyper-* and *supra-*, and spreading in the late 20th and early 21st centuries in competition with other suffixes such as *hyper-*, *ultra-*, *méga-*, *giga-* (Izert 2015).

The common expectation is that rival affixes will tend towards differentiation. Based on the assumption of linguistic economy, it is often argued that equivalent affixes will ultimately specialize or disappear, following Bréal's law of differentiation (1897). Examples of such

evolution can be found in language history. Hegedűs (2014) mentions the case of the Old English prefix *sam-* ‘half’ progressively replaced by *half-* in Middle English. Similarly, the French agentive suffix *-on* (as in *forgeron* ‘blacksmith’, *espion* ‘spy’, *piéton* ‘pedestrian’) was a close rival of *-eur* in Old and Middle French, but became unproductive and abandoned its agentive function in favour of *-eur*. Specialization over time can be observed in the case of the English suffixes *-dom* and *-ship* whose coexistence, according to Díaz Negrillo (2017), has been encouraged by the emergence of new senses (e.g. ‘realm’ and ‘skill’, respectively). Regarding the competition between *-ity* and *-ness* in English, Arndt-Lappe (2014) shows that, driven by analogy mechanisms, the two suffixes diachronically tend towards less variability in base selection and more consistency in preference patterns, with *-ity* prevailing over *-ness* in the selection of bases ending in *-able*, *-al*, *-ar*, *-ic*, and *-ive*.

It should be noted that the law of differentiation does not exclude that equivalent cases be observed in synchrony, as a transitory stage before rivalry resolution. Derivational systems may synchronically include cases of undistinguishable rivalry, as these could be destined to be resolved diachronically. Nevertheless, a strict application of the law of differentiation entails that complementary distribution should be ultimately achieved, which is rarely observed in reality. The fact that polyfunctional affixes have some distinctive functions seems sufficient to allow for their coexistence, regardless of whether they may have other functions in common. Evaluative morphology also provides examples of enduring situations of rivalry with hardly distinguishable properties, as in the case of pluractional verbal suffixes in French and Italian (Tovena & Kihm 2008, Amiot & Stosic 2014) or of high degree prefixes in French, which keep on developing over time with no clear sign of complementarity (Cartier & Huyghe 2021).

It seems that blocking, defined as the “nonoccurrence of a form due to the simple existence of another” (Aronoff 1976: 43), does not necessarily occur. Lexical blocking (i.e. blocking of a lexeme due to the existence of a synonym) is challenged by the existence of strictly synonymous doublets, reported for example by Bauer et al. (2013) in English, Fradin (2016) in French, and Radimský and Štichauer (2021) in Italian. As argued by Bauer (1983), blocking operates at the level of institutionalization, but not of coinage, and nonce words that are synonymous with an already existing lexeme may occur in individual realizations (see also Aronoff & Lindsay 2014). Pattern blocking, in which a morphological process takes precedence over an existing competing process (Rainer 1988, 2005), is not necessarily observed either. According to Rainer (2002), competition can be associated with free variation. In Spanish for example, the suffixes *-iano* and *-eano* can be indistinctly used with foreign names ending in *-e*, whereas bases ending in *-e* usually select *-eano*. In a similar vein, Bauer (2009: 193) comments on the coexistence of the prefixes *de-*, *dis-* and *un-* for 700 years in English without achieving complementary distribution and “without the blocking principle apparently having had a great deal of effect”.

It remains true that most competing affixes differ in terms of productivity, understood broadly as the potential of word-formation processes to coin neologisms (see Aronoff 1976, Baayen & Lieber 1991, Plag 1999, Bauer 2001, Gaeta & Ricca 2015, Dal & Namer 2017, among others).

It can be hypothesized that the blocking force of a word-formation process is positively correlated to its productivity, as shown by Maslen et al. (2004) in the case of inflectional competition. Furthermore, the diachronic evolution of rivalry situations is tightly related to productivity variation, although it is uncertain whether changes in productivity precede or follow competition. Van Marle (1988) argues that a decrease in productivity causes the emergence of a competing morphological process that will ultimately oust the original one, whereas Scherer (2015) claims that attested competition is the main factor of change in productivity. Most probably, both situations can occur and the relationship between productivity and competition is dynamic and system-dependent. Increase or decrease in productivity can be caused by other factors than competition (e.g. sociolinguistic variation), whereas the emergence of a process competing with an already existing one can be related to other factors than productivity (e.g. language contact). While there has been much debate on how to quantify productivity, it is widely considered as a measurable property, and a fine evaluation of diachronic productivity should be conducted on quantitative ground to allow for a better understanding of the mechanisms involved in morphological competition.

4 Quantitative research on affix rivalry

In the previous sections we have presented two scalar aspects of affix rivalry. The first one is related to semantic characterization as a definitional property of affix rivalry, considering that affixes compete more or less depending on the similarity degree, number and frequency of semantic functions they have in common. The second aspect pertains to the resolution of affix rivalry and the way competing affixes can be differentiated in their use (not only based on semantic properties). The analysis of discriminative properties among rival affixes usually reveals tendencies rather than clear-cut distinctions. Niches of affix use are fuzzily bounded and in continuous evolution from a diachronic point of view. In addition, the multiple discriminative features that can distinguish rivals (e.g. phonological, morphological, stylistic properties) jointly influence affix selection. In this context, quantitative methods appear to be well suited for the study of affix rivalry, since they are able to model gradient phenomena and to assess the relative influence of multiple factors. By means of statistical analysis, these methods determine if some observed differences are significant or only due to chance, and allow objective testing of research hypotheses.

In this section, we present an overview of the quantitative methods² that have been used in the study of affix rivalry, without claiming to be exhaustive. We first review the main sources and types of data used in quantitative research on morphological competition. Then we discuss some statistical methods and measures employed to analyze situations of affix rivalry.

² For a more detailed description of these methods, as well as for a general introduction to statistics, we refer the reader to the numerous handbooks of statistics for linguistic research that have been published in recent years (Oakes 1998, Baayen 2008, Levshina 2015, Desagulier 2017, Brezina 2018, Winter 2019, Gries 2021, among others).

4.1 Sources and types of quantitative data

Quantitative methods are data driven. Research questions can hardly be answered based on a few isolated observations, but require the collection and description of a number of items that can constitute a representative sample of the phenomenon under scrutiny. In the case of affix rivalry, samples of derivatives that can reflect the behaviour of competing affixes are examined, in order to find significant differences between them and to generalize findings. These samples are selected depending on theoretical perspectives. For example, some studies only consider differently affixed words derived from the same base, either to exclude confounding factors coming from the base or because they assume that true rivals correspond to doublets (Fradin 2019, Andreou & Lieber 2020, Radimský & Štichauer 2021). Others focus on neologisms insofar as they allow for an exploration of derivational semantics, while controlling for lexicalization effects that can affect the semantic properties of derivatives (Huyghe et al. 2023). Many researchers seek to maximize the number of data points to obtain more general and reliable results, considering large samples of derivatives issued from existing resources, such as corpora, dictionaries or other lexical databases.

Once a sample is defined, information about investigated features of derivatives can be retrieved from resources or specifically produced for research purposes. For instance, in their study of the rivalry between French suffixes *-iser* and *-ifier*, Bonami and Thuillier (2019) use the phonemic transcriptions provided by GLÀFF, a large scale inflectional lexicon for French (Hathout et al. 2014), to encode information about the stems selected by the two suffixes. Using a different strategy, Lieber and Plag (2022) examine more than 4,000 corpus tokens to investigate the rivalry between *-ing* nominalizations and conversion in English, and annotate them manually with respect to various semantic properties. The two competing morphological processes are interpreted as possible levels of the response variable, and the manually annotated properties are considered as possible predictors of that response variable.

Corpora are a major source of information in quantitative studies on affix rivalry. A wide variety of data can be directly or indirectly extracted from them, including linguistic information and metadata (e.g. author, date, medium, register). One variable often scrutinized is frequency. Token frequencies of derivatives or type frequencies of affixes can reveal properties of competing affixes. Frequency counts can be used in productivity measures and are therefore an important element in the assessment of rivalries and their evolution. Although various productivity measures have been proposed and discussed in the literature (see Aronoff 1976, Baayen 1993, 2009, Hay 2001, Gaeta & Ricca 2006, Fernández-Domínguez 2010, among others), they always involve at some point frequency counts of the items whose productivity is estimated. Frequency is also explored in combination with other potentially discriminative properties and interaction effects may be observed. For example, Varvara et al. (2021), while investigating the semantic differences between German *-ung* nouns and nominal infinitives, note that the latter include more transparent derivatives than the former, except for highly frequent *-ung* nouns, which show higher degree of semantic transparency.

In addition to frequency, corpora can provide co-occurrence counts between target words and contextual words in a specific span of text (e.g. in the same sentence). Co-occurrence data have been frequently used as an empirical basis for semantic description, from lexicography to modern computational linguistics. As summarized by Geeraerts (2010: 169), “words co-occurring with another one help to identify the properties of the word under scrutiny”. However, given the distribution of word frequencies, which according to Zipf’s law (1949) are inversely proportional to word ranks in frequency lists, raw co-occurrence frequencies are not maximally informative. In particular, very frequent words will tend to rank high in the list of collocates of many other words. A relevant statistical measure of co-occurrence should discern significant collocations from those due to the general distribution of word frequencies (see e.g. Evert 2005, Gries 2014, and Gries 2003 for an application to a case of affix rivalry).

The development of co-occurrence analysis in computational linguistics led to the emergence of complex models of contextual distribution called ‘distributional semantics models’, ‘vector space models’ or ‘word embeddings’,³ which can be viewed both as a statistical model of word meaning and a source of data for further analysis. Distributional semantics models approximate the meaning of a word based on its linguistic context. In the simplest models, a distributional representation is constructed for each target word in the form of a vector defined by co-occurrence values with the contexts considered. Word vectors can be computed based on corpus data and constitute the material for further quantitative analysis facilitated by the mathematical nature of vectors. The most common measure is the cosine similarity between two vectors, which is used to estimate the degree of semantic proximity between corresponding words. For instance, Denistia et al. (2021) explore the distributional vectors of words prefixed with *pe-* and *peN-* in Indonesian and examine their cosine similarity to highlight semantic differences between the two prefixes. Similarly, by applying cosine measure to average vectors of nouns ending with different suffixes, Huyghe and Wauquier (2021) assess degrees of rivalry between agentive suffixes in French. Vector dimensions, cosine similarity, and other quantitative information can be extracted or inferred from distributional semantics models to investigate morphological competition, and can be used as variables in statistical analyses such as regression or classification modelling.

Experimentation may be another way to collect information about rival affixes that can be used in statistical analysis. Various measurements can be elicited from experiments or surveys, such as acceptability judgements, reaction times in lexical decision tasks, eye-movements and fixation times in reading tasks, to mention a few. Although evidence from experiments is still rare in research on affix rivalry, some studies use elicitation and judgement tasks to explore discriminative properties of competing affixes. For example, Makarova (2016) investigates differences in productivity between the Russian attenuative prefixes *pri-* and *pod-* in an experimental study. In a cloze-test task, she asks informants to complete a word fragment in a stimulus sentence with the most appropriate prefix, and compares responses

³ See Lenci (2018) and Boleda (2020), among many others, for an introduction to these models.

with corpus attestations. In a similar vein, Schirakowski (2020) uses an acceptability judgement task to investigate the competition between deverbal nouns and nominalized infinitives in Spanish. Argument realization and semantic interpretation (episodic vs. generic event readings) are manipulated as independent variables to assess their role in rivalry resolution.

4.2 Statistical analysis

In addition to data collection, quantitative approaches fundamentally involve statistical analysis, which is used to determine the existence, significance and strength of relationships between variables of interest. Various statistical methods have been applied in studies on affix rivalry to determine whether competing affixes can be discriminated by some factors (e.g. phonological, morphological, semantic, sociolinguistic properties), to what extent the association of these factors can predict affix selection, and what is the contribution of each factor to the prediction. In this section, we review some hypothesis-testing methods based on monofactorial or multifactorial approaches.

Monofactorial methods are used to assess whether the values of one variable are correlated with those of another one. Different statistical tests are available and their selection is based on the type and distribution of variables. For example, chi-squared tests are used by Denistia and Baayen (2019) to determine if the Indonesian nominalizing prefixes *pe-* and *peN-* differ with respect to semantic types of derivatives. Similarly, Lieber and Plag (2022) apply chi-squared tests to evaluate whether the formation of converted vs. *-ing* nominalizations depends on each verbal and nominal property they examine, additionally reporting Cramér's V measures as an indicator of the strength of the relationship observed between variables. Varvara and Zamparelli (2019) conduct U-tests, as an alternative to t-tests for comparing means of non-normally distributed data, to evaluate argument concreteness as a possible discriminative factor between nominal infinitives and derived nouns in Italian.

Multifactorial methods investigate the effect of multiple independent variables and the possible ways in which these variables are related. They are adopted in the study of affix rivalry when examining the joint influence of various factors on affix selection. Regression models are among the most commonly used multifactorial methods. They have been fruitfully applied in linguistics (see e.g. Baayen & Moscoso del Prado Martín 2005, Bresnan et al. 2007) and recently in research on affix rivalry. Different kinds of regression models can be fitted according to the type of dependent variable considered, the most common ones in linguistic research being linear regression and logistic regression.⁴ The former is used when the response variable is numerical and when the relationship between variables is linear. For example, Varvara (2017) applies linear regression to assess the relationship between the frequency of two competing morphological processes in Italian and some base verb features. Logistic regression is used when the response variable is binary and categorical. In the case of

⁴ Further regression models can be used, such as ordinal or multinomial regression, poisson regression, and mixed-effects models (see e.g. Baayen 2008, Winter 2019, Gries 2021).

affix rivalry, it has been used to estimate the probability of selecting one affix over the other given the values of predictor variables. Bonami and Thuillier (2019), for instance, make use of logistic regression to investigate the competition between *-iser* and *-ifier* in French. They show that some phonological properties (length of the stem, final consonant) and some morphological properties (structure of the ascending morphological family) play a role in the distribution of the two suffixes.

Regression models can give rise to convergence problems with exceedingly sparse data or with collinear predictors. This may typically be the case with corpus data, because of the Zipfian distribution of frequencies, which consists in a large number of rare elements and a small number of very frequent ones. Alternative models can be selected in the family of tree-based methods that includes classification and regression trees (CARTs), conditional inference trees, and random or conditional inference forests (see Gries 2020, 2021 for a practical introduction to these methods). As in the case of logistic regression, CARTs and conditional inference trees can be used to predict affix selection given the distribution of rival affixes with respect to some independent variables. These models use recursive partitioning in order to find the values of the independent variables that best predict the response variable (in terms of accuracy or other statistical criteria). Random forests are based on multiple conditional inference trees, obtained through various randomizations of the data. These models can also provide a measure of variable importance that indicates the relative contribution of each independent variable to the prediction. Trees and forests have been used in the last decade in various areas of linguistics (see e.g. Tagliamonte & Baayen 2012, Dilts 2013, Hansen & Schneider 2013, Bernaisch et al. 2014), and most notably in studies on affix rivalry. Baayen et al. (2013) have compared different statistical models (logistic regression, trees, forests, and naive discriminative learning) while examining four different cases of rivalry in Russian. Makarova (2016) has analyzed experimental data on competing attenuative verbal prefixes in Russian by means of conditional inference trees and random forests. Similarly, Naccarato (2019) has investigated the rivalry between Russian (para)synthetic compounds ending in *-ec* and *-tel'* using conditional inference tree and random forest modelling.

Another example of multifactorial analysis that has been frequently used in the study of affix rivalry is analogical modelling. Analogical modelling is a supervised exemplar-based approach that predicts the behaviour of a new item on the basis of stored memory items (Skousen 1989, 1992, 1995, Eddington & Lonsdale 2007). Based on some given features, target items are compared to the elements previously stored in the model, in search of the most similar ones. In its linguistic implementation, analogical modelling relies on the assumption that the knowledge stored in the mental lexicon can serve as a basis to analogically produce information about new words. Arndt-Lappe (2014) has used analogical modelling to address the rivalry between English nominalizing suffixes *-ity* and *-ness* based on a corpus analysis of 2,700 neologisms from three different centuries. The analogical model predicts preference patterns for *-ity* and *-ness* based on the phonological properties of the two base-final syllables and on the syntactic category of the base. Plag et al. (2023) also use an analogical model to predict the distribution between conversion and *-ing* nominalizations in English. When

considering semantic properties (eventivity, quantification, and aspectual features), the model achieves an accuracy of 84%, which appears to be higher than the accuracy of a model that additionally considers phonological properties.

As a final note, it can be observed that monofactorial and multifactorial methods are mostly used to investigate the competition between pairs of affixes, which constitutes the prototypical case study on rivalry. However, these methods can be applied to more than two competing processes, i.e. with a non-binary dependent variable. The examination of more than one pair of rival affixes broadens the scope of research and, given the inherent scalarity of affix rivalry, raises the question of how to assess degrees of morphological competition. This question can be addressed using supervised methods such as the regression and classification methods described above, but also unsupervised methods such as clustering analysis, which consists in grouping similar objects of a dataset based on some input information. The uneven distribution of words with competing affixes in clusters based on semantic information can be viewed as an indication of various degrees of rivalry between different competing pairs of affixes (Huyghe et al. 2023). More broadly, a generalized measure of morphological competition could be useful to compare different pairs of rivals in a given language, as well as across languages. Fernández-Domínguez (2017) proposes an index of competition that evaluates the prevalence of a lexeme over its rivals, taking into consideration the number and frequency of rival lexemes. Such a measure could be extended to affixes through observations on their morphological series. In a different vein, Salvadori and Huyghe (to appear) propose to adapt the Bray-Curtis dissimilarity coefficient (Bray & Curtis 1957) to compute an index of affix rivalry based on the realization frequency of semantic functions among derivatives formed with competing affixes. Although research on measures of rivalry is still limited, it constitutes a challenging topic for future work on morphological competition.

5 Contributions

This special issue is dedicated to quantitative approaches to affix rivalry. It includes five contributions that investigate quantitative aspects of affix rivalry and provide new insights into morphological competition. The objective is threefold. First, the special issue presents empirical research on affix rivalry and brings together studies that carefully examine rival affixes and their possible discriminative properties in different languages (namely English, French and Russian). Second, the contributions explore methodological issues through the use of various quantitative methods to investigate affix rivalry. Based on large sets of data collected from corpora and lexical resources, they propose different statistical analyses of rivalry, allowing comparison of methods and showing the adaptability of quantitative techniques. Third, the studies help clarify some theoretical issues related to affix rivalry and contribute to our general understanding of the phenomenon. Overall, investigation through quantitative methods allows for a close look at the nature of affix rivalry and sheds light on the distribution of affix functions in the derivational system.

S. Arndt-Lappe examines the rivalry between *-ic* and *-ical* in English and investigates how differences in individual lexical knowledge can influence the pattern selection in affix rivalry. Using a computational analogical model based on phonological features, the author tests predictions for rival possible adjectives based on two different lexicons. The first one comprehends the full set of attested derivatives in the *British National Corpus*, thus simulating the mental lexicon of a speaker with an extended vocabulary, while the second one simulates a smaller lexicon, mainly based on the spoken language part of the corpus. Results show differences in prediction between the two lexicons. In addition to morphological niches previously observed for the less productive suffix *-ical* (e.g. preference for *-olog* bases), the model identifies niches specific to *-ic* that vary depending on the type of lexicon considered. These findings concern not only the psycholinguistic aspects of affix rivalry, but also the role of datasets size in computational modelling.

M. Aronoff proposes a theoretical discussion of morphological competition that focuses on its resolution and examines the ways in which it has been previously addressed. The author first reassesses the notion of blocking and then specifically reviews Kiparsky's elsewhere principle (traced back to Panini's grammar of Sanskrit), Yang's tolerance principle, and Gause's law of competitive exclusion. Based on documented cases of rivalry, he suggests that the outcome of morphological competition can be analyzed in evolutionary terms, with a notion of niche construction similarly usable in ecology and linguistics. It is argued that Gause's law of competitive exclusion is the fundamental principle that rules morphological competition and governs both blocking and elsewhere distribution, which consequently should not be considered as explanatory on their own. The tolerance principle is presented as a bottom-up complementary mechanism that can explain deviations from the general principle.

N. Bobkova and F. Montermini study the rivalry between three suffixes used in the formation of denominal adjectives in Russian (*-n-*, *-sk-*, and *-Ov-*), and examine the influence of various phonological, morphological, and semantic properties of their base lexemes in a multifactorial approach. They consider two different datasets: the first one including highly frequent derivatives extracted from the *Russian National Corpus*, and the second one including all doublets or triplets from the corpus. Based on a random forest classifier applied to the first dataset, they show that the semantic properties of the base (animacy in particular) and to a lesser extent segmental properties such as the last phoneme of the derivational stem are the best predictors of suffix discrimination. They then test the predictions of the model on the second dataset, finding that the general productivity of the suffixes and some properties of the base nouns can influence suffix selection probabilities.

M. Guzmán Naranjo and O. Bonami investigate how distributional semantics can be used to discriminate cases of affix rivalry in French. They represent the semantics of derivational processes by means of average difference vectors between bases and derivatives. First, they show that these average difference vectors allow for a clustering of French suffixes that is comparable to the taxonomies proposed by expert morphologists. Second, they use boosting tree classifiers to identify rival processes on the basis of their distributional representations,

assuming that rival processes will not be distinguishable and that classifiers will not perform better than by chance in their case. Results are consistent with previous knowledge on morphological competition between the processes investigated, but also show that different degrees of semantic discriminability among derivational processes can be identified, thus providing evidence for a gradient view of affix rivalry.

J. Thuilier, D. Tribout and M. Wauquier examine four rival suffixes used in the formation of demonyms in French (*-ois*, *-ais*, *-ien*, and *-éen*). Considering that the derivational processes used to form demonyms are semantically equivalent, they intend to determine whether phonological and morphological properties of the base toponyms, as well as extralinguistic properties such as geographical features, can motivate the selection of the different suffixes. Random forests are used to evaluate the competition between the four suffixes based on a sample of more than 2,000 French demonyms and their respective toponyms. Results highlight the multifactorial nature of the rivalry investigated, showing that various complementary factors are at play. Although the most important predictors are phonological features of toponyms, variables such as geographical features also contribute substantially to the accuracy of the prediction.

Acknowledgements

We would like to thank Andrea Sims for her helpful comments and suggestions on an earlier version of this paper. We are grateful to the authors for their participation and valuable contributions to the special issue, and to the editors of *Word Structure* for their support and assistance along the way. A special thanks goes to colleagues who accepted to take part in the reviewing process. This work was partially supported by the Swiss National Science Foundation under grant 10001F_188782.

References

- Amiot, Dany & Stosic, Dejan. 2014. When evaluative morphology, pluractionality and aspect get tangled up: a case study of French suffixed verbs. In Gavriilidou, Zoe & Revithiadou, Anthi. *Mélanges offerts à Anna Anastassiades-Syméonides à l'occasion de sa retraite*, 16-33. Kavala : Saita.
- Andreou, Marios, & Lieber, Rochelle. 2020. Aspectual and quantificational properties of deverbal conversion and-ing nominalizations: the power of context. *English Language & Linguistics* 24(2). 333-363.
- Arndt-Lappe, Sabine. 2014. Analogy in suffix rivalry: The case of English -ity and -ness. *English Language & Linguistics* 18(3). 497–548.
- Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge, MA: The MIT Press.
- Aronoff, Mark. 2020. -less and -free. In Körtvélyessy Livia & Štekauer Pavol (eds.), *Complex Words: Advances in Morphology*, 55-64. Cambridge: Cambridge University Press.

- Aronoff, Mark, & Lindsay, Mark. 2014. Productivity, blocking, and lexicalization. In Lieber, Rochelle & Štekauer Pavol (eds.), *The Oxford Handbook of Derivational Morphology*, 67-83. Oxford: Oxford University Press.
- Aronoff, Mark, & Cho, Sungeun. 2001. The semantics of *-ship* suffixation. *Linguistic Inquiry* 32(1). 167-173.
- Baayen, R. Harald. 1993. On frequency, transparency, and productivity. In Booij, Geert & van Marle, Jaap (eds.), *Yearbook of Morphology 1992*, 181-208. Dordrecht: Kluwer Academic Publishers.
- Baayen, R. Harald. 2008. *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Baayen, R. Harald. 2009. Corpus linguistics in morphology: Morphological productivity. In Lüdeling, Anke & Kytö, Merja (eds.), *Corpus Linguistics. An International Handbook*. Vol. 2, 899-919. Berlin: Mouton de Gruyter.
- Baayen, R. Harald, Endresen, Anna, Janda, Laura A., Makarova, Anastasia & Nessel, Tore. 2013. Making choices in Russian: pros and cons of statistical methods for rival affixes. *Russian Linguistics* 37(3). 253-291.
- Baayen, Harald, and Rochelle Lieber. 1991. Productivity and English derivation: A corpus-based study. *Linguistics* 29, 801-844.
- Baayen, Harald & Moscoso del Prado Martín, Fermin. 2005. Semantic density and past-tense formation in three Germanic languages. *Language* 81. 666–698.
- Bauer, Laurie. 1983. *English Word-Formation*. Cambridge: Cambridge University Press.
- Bauer, Laurie. 2001. *Morphological Productivity*. Cambridge: Cambridge University Press.
- Bauer, Laurie. 2009. Competition in English word-formation. In van Kemenade, Ans & Los, Bettelou (eds.), *The Handbook of the History of English*, 177–198. Malden, MA: Blackwell.
- Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2013. *The Oxford Reference Guide to English Morphology*. Oxford: Oxford University Press.
- Bernaisch, Tobias, Stefan Th. Gries, & Joybrato Mukherjee. 2014. The dative alternation in South Asian English(es): Modelling predictors and predicting prototypes. *English World- Wide* 35(1). 7–31.
- Boleda, Gemma. 2020. Distributional semantics and linguistic theory. *Annual Review of Linguistics* 6. 213–234.
- Bonami, Olivier & Thuilier, Juliette. 2019. A statistical approach to rivalry in lexeme formation: French *-iser* and *-ifier*. *Word Structure* 12(1). 4–41.
- Booij, Geert. 2002. *The morphology of Dutch*. Oxford: Oxford University Press.
- Bray, Roger J. and J. T. Curtis. 1957. An ordination of upland forest communities of southern Wisconsin. *Ecological Monographs* 27:4, 325-349.
- Bréal, Michel. 1897 [2005]. *Essai de sémantique*. Limoges: Lambert-Lucas.
- Bresnan, Joan & Cueni, Anna & Nikitina, Tatiana & Baayen, Harald. 2007. Predicting the dative alternation. In Bouma, Gerlof & Kramer, Irene & Zwarts, Joost (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Sciences.
- Brezina, Vaclav. 2018. *Statistics in corpus linguistics: A practical guide*. Cambridge: Cambridge University Press.

Carstairs-McCarthy, A. 1994. Inflection classes, gender, and the principle of contrast. *Language* 70:4, 737-788.

Cartier, Emmanuel & Huyghe, Richard. 2021. La concurrence affixale en diachronie: le cas des préfixes de haut degré en français. *Linx* 82.

Cartoni, Bruno, Fiammetta Namer, & Stéphanie Lignon. 2015. A cross-linguistic insight on agentive noun formation in Italian and French. *Carnets de Grammaire* 22, 81-98.

Dal, Georgette & Fiammetta Namer. 2017. Productivity. In Hippiusley, Andrew & Stump, Gregory (eds.), *The Cambridge Handbook of Morphology*, 70-90. Cambridge: Cambridge University Press.

Denistia, Karlina, and Baayen, R. Harald. 2019. The Indonesian prefixes PE- and PEN-: A study in productivity and allomorphy. *Morphology* 29. 385–407.

Denistia, Karlina, Shafaei-Bajestan, Elnaz, and Baayen, R. Harald. 2021. Exploring semantic differences between the Indonesian prefixes PE- and PEN- using a vector space model. *Corpus Linguistics and Linguistic Theory*, 1-26.

Desagulier, Guillaume. 2017. *Corpus linguistics and statistics with R*. Cham: Springer International Publishing.

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

Dilts, Philip. 2013. *Modelling phonetic reduction in a corpus of spoken English using random forests and mixed-effects regression*. Edmonton: University of Alberta. Unpublished Ph.D. dissertation.

Dressler, Wolfgang U. & Barbaresi, Lavinia Merlini & Schwaiger, Sonja & Ransmayr, Jutta & Sommer-Lolei, Sabine & Korecky-Kröll, Katharina. 2019. Rivalry and lack of blocking among Italian and German diminutives in adult and child language. In Rainer, Franz & Gardani, Francesco & Dressler, Wolfgang U. & Luschützky, Hans Christian (eds.), *Competition in Inflection and Word-Formation*, 123–143. Cham: Springer.

Eddington, David, and Lonsdale, Deryle. 2007. Analogical Modeling: An Update. In *Proceedings of the 19th European Summer School in Logic, Language, and Information*. Dublin: Trinity College.

Evert, Stefan. 2005. *The statistics of word cooccurrences : word pairs and collocations*. Ph. D. thesis, University of Stuttgart.

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

Fernández-Domínguez, Jesús. 2010. Productivity vs. lexicalization: Frequency-based hypotheses on word-formation. *Poznań Studies in Contemporary Linguistics* 46(2), 193-219.

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

- Fradin, Bernard. 2016. L'interprétation des nominalisations en N-*age* et N-*ment* en français. In: Buchi, E., Chauveau, J.-P., Pierrel, J.-M. (eds.), *Actes du XXVIIe congrès international de linguistique et philologie romane*, 53-66. Strasbourg : Société de Linguistique Romane / Eliphi.
- Fradin, Bernard. 2019. Competition in Derivation: What Can We Learn from French Doublets in -*age* and -*ment*? In: Rainer, Franz, Gardani, Francesco, Dressler, Wolfgang U. et Luschützky, Hans Christian (eds.), *Competition in Inflection and Derivation*, 67-93. Cham: Springer.
- Gaeta, Livio & Ricca, Davide. 2006. Productivity in Italian word formation: A variable-corpus approach. *Linguistics* 44(1). 57-89.
- Gaeta, Livio and Davide Ricca. 2015. Productivity. In Müller, Peter O., Ohnheiser, Ingeborg, Olsen, Susan, and Rainer, Franz (eds.), *Word-Formation: An International Handbook of the Languages of Europe*, 842-858. Berlin: De Gruyter.
- Gardani, Francesco & Rainer, Franz & Luschützky, Hans Christian. 2019. Competition in Morphology: A Historical Outline. In Rainer, Franz & Gardani, Francesco & Dressler, Wolfgang U. & Luschützky, Hans Christian (eds.), *Competition in Inflection and Word-Formation*, 3-36. Cham: Springer.
- Geeraerts, Dirk. 2010. *Theories of lexical semantics*. Oxford: Oxford University Press.
- Gries, Stefan Th. 2001. A corpus-linguistic analysis of English -*ic* vs. -*ical* adjectives. *ICAME Journal* 25. 65-108.
- Gries, Stefan Th. 2003. Testing the sub-test: a collocational-overlap analysis of English -*ic* and -*ical* adjectives. *International Journal of Corpus Linguistics* 8(1). 31-61.
- Gries, Stefan Th. 2014. Frequencies, probabilities, association measures in usage-/exemplar-based linguistics: some necessary clarifications. In Nikolas Gisborne & Willem Hollmann (eds.), *Theory and data in cognitive linguistics*, 15-48. Amsterdam & Philadelphia: John Benjamins.
- Gries, Stefan Th. 2020. On Classification Trees and Random Forests in Corpus Linguistics: Some Words of Caution and Suggestions for Improvement. *Corpus Linguistics and Linguistic Theory* 16 (3). 617-47.
- Gries, Stefan Th. 2021. *Statistics for Linguistics with R*. 3rd rev. & ext. Boston & Berlin: De Gruyter.
- Guz, Wojciech. 2009. English affixal nominalizations across language registers. *Poznań Studies in Contemporary Linguistics* 45.4: 461-485.
- Guzmán Naranjo, Matías & Bonami, Olivier. 2023. A distributional assessment of rivalry in word formation. *Word Structure* 16(1). 87-114.
- Hansen, Sandra & Schneider, Roman. 2013. Decision tree-based evaluation of genitive classification: An empirical study on CMC and text corpora. In Iryna Gurevych, Chris Biemann & Torsten Zesch (eds.), *Language processing and knowledge in the web*, 83-88. Berlin & New York: Springer.
- Hathout, Nabil & Sajous, Franck & Calderone, Basilio. 2014. GLÀFF, a large versatile French lexicon. In *Proceedings of the Conference on Language Resources and Evaluation (LREC 2014)*, 1007-1012.
- Hay, Jennifer. 2001. Lexical Frequency in Morphology : Is Everything Relative ? *Linguistics* 39. 1041-1070.
- Hegedüs, Irén. 2014. Affix doublets and affix rivalry in the history of English. *Argumentum* 10. 312-324.

- Huyghe, Richard & Wauquier, Marine. 2021. Distributional semantics insights on agentive suffix rivalry in French. *Word Structure* 14(3). 354–391.
- Huyghe, Richard & Lombard, Alizée & Salvadori, Justine & Schwab, Sandra. 2023. Semantic rivalry between French deverbal neologisms in *-age*, *-ion* and *-ment*. In Kotowski, Sven & Plag, Ingo (eds.), *The Semantics of Derivational Morphology*, 125-158. Berlin: De Gruyter.
- Izert, Małgorzata. 2015. *La construction préfixale de forte intensité en français contemporain*. Łask: Oficyna Wydawnicza Leksem.
- Lehrer, Adrienne. 2000. Are affixes signs? The semantic relationships of English derivational affixes. In Wolfgang U. Dressler, Oskar E. Pfeiffer, Markus A. Pöchtrager, John R. Rennison (eds.), *Morphological Analysis in Comparison*, 143-154. Amsterdam: John Benjamins.
- Lenci, Alessandro. 2018. Distributional models of word meaning. *Annual Review of Linguistics* 4, 151–171.
- Levshina, Natalia. 2015. *How to do linguistics with R. Data Exploration and Statistical Analysis*, Amsterdam-Philadelphia: John Benjamins.
- 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.
- Lindsay, Mark & Aronoff, Mark. 2013. Natural selection in self-organizing morphological systems. In Hathout, Nabil, Montermini, Fabio & Tseng, Jesse (eds.) *Morphology in Toulouse. Selected proceedings of Décembrettes 7*, 133-153. München: Lincom Europa
- Makarova, Anastasia. 2016. Variation in Russian verbal prefixes and psycholinguistic experiments. In Anstatt, Tanja, Gattnar, Anja, and Clasmeier, Christina (eds). *Slavic languages in psycholinguistics: Chances and challenges for empirical and experimental research*, 113-33. Tübingen: Narr Francke Attempto.
- Marle, J. van. 1988. On the Role of Semantics in Productivity Change. In Geert Booij and Jaap van Marie (eds.), *Yearbook of Morphology* 1. Dordrecht: Foris, 139–154.
- Martin, Fabienne. 2010. The semantics of eventive suffixes in French. In Alexiadou, Artemis & Rathert, Monika (eds.), *The Semantics of Nominalizations across Languages and Frameworks*, 109-140. Berlin: Mouton de Gruyter.
- Maslen, Robert J.C., Anna L. Theakson, Elena V.M. Lieven, and Michael Tomasello. 2004. A Dense Corpus Study of Past Tense and Plural Overregularization in English. *Journal of Speech, Language, and Hearing Research* 47(6): 1319–1333.
- Merkuur, Anne, Jan Don, Eric Hoekstra, and Arjen P. Versloot. 2019. Competition in Frisian Past Participles. In Rainer, Franz & Gardani, Francesco & Dressler, Wolfgang U. & Luschützky, Hans Christian (eds.), *Competition in Inflection and Word-Formation*, 195–222. Cham: Springer.
- Mikolov, Thomas, Sutskever, Ilya, Chen, Kai, Corrado, Greg, Dean, Jeffrey. 2013. Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th Conference on Advances in Neural Information Processing Systems*. 3111–19.
- Missud, Alice & Villoing, Florence. 2020. The morphology of rival *-ion*, *-age* and *-ment* selected verbal bases. *Lexique* 26, 29-52.
- Naccarato, Chiara. 2019. Agentive (para)synthetic compounds in Russian: a quantitative study of rival constructions. *Morphology* 29(1). 1–30.

Nagano, Akiko. 2022 Affixal rivalry and its purely semantic resolution among English derived adjectives. *Journal of Linguistics*, 1-32.

Oakes, Michael. 1998. *Statistics for corpus linguistics*. Edinburgh: Edinburgh University Press.

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

Plag, Ingo. 2000. On the mechanisms of morphological rivalry: A new look at competing verb-deriving affixes in English. In Bernhard Reitz & Sigrid Rieuwerts (eds.), *Anglistentag 1999 Mainz*, 63-76. Trier: Wissenschaftlicher Verlag Trier.

Plag, Ingo & Kawaletz, Lea & Arndt-Lappe, Sabine & Lieber, Rochelle. 2023. Analogical modeling of derivational semantics. In Kotowski, Sven & Plag, Ingo (eds.), *The Semantics of Derivational Morphology*, 87-124. Berlin: De Gruyter.

Podesva, Robert J., and Devyani Sharma (eds). 2013. *Research methods in linguistics*. Cambridge: Cambridge University Press.

Radimský, Jan, and Štichauer, Pavel. 2021. Nomina Actionis in the diachrony of Italian: a paradigm-based model of competition. *Lingue e linguaggio* 20 (1). 33-55.

Rainer, Franz. 1988. Towards a theory of blocking: the case of Italian and German quality nouns. In Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1*. Dordrecht: Foris, 155–185.

Rainer, Franz. 2002. Zum Problem der Suffixwahl im Bereich der Relationsadjektive zu spanischen Schriftstellernamen. In Dieter Kremer (ed.), *Onomastik. Akten des 18 Internationalen Kongresses für Namensforschung* (Trier, 12-17 April 1993), 115-127. Bd. V. Tübingen: Niemeyer.

Rainer, Franz. 2005. Constraints on Productivity. In: Štekauer, P., Lieber, R. (eds.), *Handbook of Word-Formation*, Dordrecht: Springer.

Rodrigues, Alexandra Soares. 2015. Semantic affix rivalry: the case of Portuguese nominalisers. In Jenny Audring, Jenny & Koutsoukos, Nikos & Masini, Francesca & Raffaelli, Ida (eds.) *MMM9 On-line Proceedings: Morphology and Semantics*, 134–137.

Romaine, Suzan. 1983. On the productivity of word formation rules and limits of variability in the lexicon. *Australian Journal of Linguistics* 3(2). 177–200.

Säily, Tanja. 2011. Variation in morphological productivity in the BNC: Sociolinguistic and methodological considerations. *Corpus Linguistics and Linguistic Theory* 7(1). 119–141.

Salvadori, Justine, and Huyghe, Richard (to appear). Affix polyfunctionality in French deverbal nominalizations. *Morphology*.

Scherer, Carmen 2015. Change in Productivity. In Müller, Peter O., Ohnheiser, Ingeborg, Olsen, Susan, and Rainer, Franz (eds.), *Word-Formation: An International Handbook of the Languages of Europe*. Vol. 3. 1781–1793. Berlin: de Gruyter.

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>.

Schulte, Marion. 2015. Polysemy and synonymy in derivational affixation—a case study of the English suffixes -age and -ery. *Morphology* 25(4). 371–390.

Skousen, Royal. 1989. *Analogical Modeling of Language*. Dordrecht: Kluwer.

- Skousen, Royal. 1992. *Analogy and Structure*. Dordrecht: Kluwer.
- Skousen, Royal. 1995. Analogy: A non-rule alternative to neural networks. *Rivista di Linguistica* 7(2). 213–231.
- Stump, G. (2019). An apparently noncanonical pattern of morphotactic competition. In Rainer, Franz & Gardani, Francesco & Dressler, Wolfgang U. & Luschützky, Hans Christian (eds.), *Competition in Inflection and Word-Formation*, 259–278. Cham: Springer.
- Tagliamonte, Sali A. & R. Harald Baayen. 2012. Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change* 24(2). 135–178.
- Thornton, Anna M. 2011. Overabundance (multiple forms realizing the same cell): a non-canonical phenomenon in Italian verb morphology. In Maiden, Martin, John Charles Smith, Maria Goldbach & Marc-Olivier Hinzelin (eds.) *Morphological Autonomy: Perspectives from Romance Inflectional Morphology*, 358–381. Oxford: Oxford University Press.
- Thornton, A. M. 2012. Reduction and maintenance of overabundance. A case study on Italian verb paradigms. *Word Structure* 5(2), 183–207.
- Tovena, Lucia & Kihm, Alain. 2008. Event internal pluractional verbs in some Romance languages. *Recherches Linguistiques de Vincennes* 37, 9–30.
- Uth, Melanie. 2010. The rivalry of the French nominalization suffixes -age and -ment from a diachronic perspective. In: Alexiadou, A., Rathert, M. (eds.), *The Syntax of Nominalizations across Languages and Frameworks*, (eds), 215–244. Berlin: De Gruyter.
- Varvara, Rossella. 2017. *Verbs as nouns: empirical investigations on event-denoting nominalizations*. PhD thesis, University of Trento.
- Varvara, Rossella. 2020. Constraints on nominalizations: investigating the productivity domain of Italian -mento and -zione. *Zeitschrift für Wortbildung /Journal of Word Formation* 2/2020, 78–99.
- 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.
- Varvara, Rossella, and Roberto Zamparelli. 2019. Competition Between Event-Denoting Deverbal Nouns and Nominal Infinitives in Italian. In Rainer, Franz, Gardani, Francesco, Dressler, Wolfgang U. and Luschützky, Hans Christian (eds.), *Competition in Inflection and Word-Formation*, 95–121. Cham: Springer.
- Winter, Bodo. 2019. *Statistics for linguists: An introduction using R*. New York: Routledge.
- Zipf, George K. 1949. *Human behaviour and the principle of least effort: An introduction to human ecology*. Cambridge, MA: Addison-Wesley.