How rivalrous are rivals? A distributional semantic approach to gradient competition in evaluative morphology

Ivan Lacić

Universitat Pompeu Fabra

ivan.lacic@upf.edu

Accepted for publication in *Word Structure*

Abstract

Affix rivalry is a form of morphological competition in which multiple affixes stand in a many-to-one relationship with a shared or closely related semantic function. Despite growing interest, evaluative morphology has remained largely peripheral to this line of research. Through a distributional semantic study of six Italian intensifying prefixes (*arci-*, *extra-*, *iper-*, *stra-*, *super-*, and *ultra-*), we argue that rivalry is gradient rather than binary: morphological processes compete to a greater or lesser degree depending on how much their semantic transformations overlap. We model both derivational outputs and the semantic transformations they induce as vectors in semantic space, and propose an information-theoretic operationalization of rivalry as the proportion of uncertainty about process identity that remains after observing the semantic transformation. Converging analyses show that the six prefixes are neither semantically equivalent nor categorically distinct: they occupy substantially overlapping regions of semantic space while retaining weak but detectable prefix-specific signatures. The analyses further show that classification accuracy and information-theoretic measures can diverge, underscoring the value of the latter for quantifying rivalry. Overall, rivalry among evaluative prefixes appears to share the gradient organization observed in other derivational domains, differing in degree rather than kind: competing within a narrow functional domain, intensifying prefixes sit toward the high-overlap end of the rivalry continuum, where rivals remain only marginally distinguishable.

Keywords: affix rivalry, evaluative morphology, intensification, distributional semantics, Italian

1 Introduction

Morphological systems rarely exhibit a strict one-to-one correspondence between form and meaning (Bybee 1985). More often than not, a single affix may serve multiple semantic functions, and a single function may be expressed by multiple affixes. This multiplicity of form-meaning mappings gives rise to *competition*, understood as "the fact that speakers routinely have to make a choice between alternative ways of realizing a certain concept" (Gardani et al. 2019, 4).

A particular subtype of morphological competition is *affix rivalry* (Aronoff 1976). Rival affixes are widespread across languages. In Italian, for instance, the suffixes *-mento*, *-zione*, and *-tura*, among others, compete in the creation of event-denoting nominalizations (e.g., *venerare* 'to worship' → *venerazione* 'worshiping'; *annegare* 'to drown' → *annegamento* 'drowning'; *spuntare* 'to trim' → *spuntatura* 'trim') (Varvara 2020). Similarly, in English, negative prefixes such as *in-*, *un-*, and *non-* are all available for the formation of negated adjectives (e.g., *accessible* → *inaccessible*, *unaccessible*, *non-accessible*) (Bauer et al. 2013).

While the prevailing view holds that affix rivalry should be defined on semantic grounds – that is, that some degree of semantic relatedness among rival affixes must be present – how this relatedness is conceptualized varies considerably. Some scholars adopt a strict interpretation, requiring rival affixes to yield derivatives with identical meanings (Plag 1999, Fradin 2019, Masini 2019, a.o.), thereby hinging on a categorical distinction between semantically identical and semantically distinct processes. Others allow for the possibility of semantic differentiation among rivals (Naccarato 2019, Aronoff 2020, Bonami & Guzmán Naranjo 2023, a.o.), conceiving rivalry as a situation in which two or more affixes, at least in one of their derivational patterns, participate in the formation of words associated with identical or closely related semantic functions. Partial functional overlap is thus considered a sufficient condition for rivalry. Under this view, rivalry is not a binary property but a gradient one: affixes can be more or less competing depending on how semantically close their outputs are, and semantic differences between rival forms can vary in magnitude and systematicity (Huyghe & Varvara 2023, Guzmán Naranjo & Bonami 2023). This gradient logic extends naturally from outputs to the derivational processes themselves: if two affixes induce near-identical semantic shifts from base to derivative, they are expected to be in stronger competition than if their shifts diverge systematically.

In recent years, affix rivalry has attracted renewed attention, and rival affixes have been examined across a variety of derivational domains, with much of the work focusing on how phonological, morphological, lexical, semantic, pragmatic, and/or stylistic factors shape the distribution of competing affixes (see, among many others, Plag 1999, Lindsay & Aronoff 2013, Arndt-Lappe 2014, Bonami & Thuilier 2019, Denistia & Baayen 2019, Huyghe et al. 2023, Thuilier et al. 2023). In addition, methodological studies such as Guzmán Naranjo & Bonami (2023) and Salvadori et al. (2024) have refined quantitative tools for measuring rivalry across broad sets of morphological processes.

Despite this growing body of work, evaluative morphology (EM) has remained largely peripheral to research on affix rivalry. When rivalry among evaluative affixes has been addressed, attention has focused primarily on diminutives (Dressler et al. 2019, Efthymiou 2024), for which "massive evidence for the coexistence of different productive diminutive suffixations on the same base" has been reported (Dressler et al. 2019, 129). Intensifying prefixation, by contrast, has received comparatively little attention: notable exceptions are Cartier & Huyghe (2021) and Kamber & Huyghe (2023), who investigated rivalry among French intensifying prefixes from a diachronic perspective. Synchronic studies of rivalry in prefixal intensification, however, remain rare. This scarcity is somewhat paradoxical, given that evaluative functions are frequently realized through multiple competing formal markers (Grandi 2015). It may, however, be linked to the inherent unpredictability of EM. Whereas much of the work on prototypical derivation has identified base-related properties that guide the distribution of rival affixes – under the expectation that competing processes will tend toward complementary specialization across different dimensions (Lindsay & Aronoff 2013) – EM often presents cases in which affixes continue to expand and overlap without stabilizing into complementary distributions (Cartier & Huyghe 2021). In addition, while quantitative modeling has become central in recent research on affix rivalry, studies of EM have largely remained descriptive, with limited engagement with statistical and distributional approaches. The intersection of quantitatively informed methods, evaluative morphology, and affix rivalry thus remains underexplored.

The present study addresses this gap through an empirically grounded investigation of rivalry among six Italian intensifying prefixes: *arci-*, *extra-*, *iper-*, *stra-*, *super-*, and *ultra-*. While each of these prefixes displays a range of semantic functions beyond INTENSIFICATION – for instance, *stra-* also derives forms with LOCATIVE SPATIAL (e.g., *stramurale* 'extramural'), LOCATIVE NON-SPATIAL (e.g., *stragiudiziale* 'extrajudicial'), and EXCESS (e.g., *strapagato* 'overpaid') meanings – INTENSIFICATION is the only function in which all six prefixes compete, thereby instantiating a case of *partial rivalry* (Guzmán Naranjo & Bonami 2023). Accordingly, this study focuses on a *many-to-one* relationship in which six forms converge on the function of INTENSIFICATION, while their other functions remain outside the scope of the analysis.

In their prototypical intensifying use, these prefixes attach to gradable adjectives to express a high or maximal degree of the property denoted by the base, often accompanied by pragmatic nuances (Dressler & Merlini Barbaresi 1994). Illustrative examples from `itWaC` (Baroni et al. 2009) are given in (1).

(1) a. *È semisconosciuto in Italia ma* ***arcinoto*** *negli Stati Uniti.*
be.IND.PRS.3SG semi.unknown in Italy but ARCI.known in.the States United
'It is relatively unknown in Italy but very well-known in the United States.'

b. *Un fisico mozzafiato, gambe* ***extra-lunghe*** *[...].*
a physique breathtaking legs EXTRA-long [...]
'A stunning physique, extra-long legs [...].'

c. *Credo sia una città molto caotica e* ***iper-moderna****.*
believe.IND.PRS.1SG be.SBJV.PRS.3SG a city very chaotic and IPER-modern
'I think it is a very chaotic and hyper-modern city.'

d. *Il palazzo Biscari è* ***stra-colmo*** *di gente.*
the palace Biscari be.IND.PRS.3SG STRA-full of people
'The Biscari palace is packed with people.'

e. *Le comunicazioni* ***superveloci*** *sono un'esigenza centrale.*
the communications SUPER.fast be.IND.PRS.3PL a.requirement central
'Superfast communications are a key requirement.'

f. *Ditemi di sì che mi farete* ***ultra-felice****.*
tell.IMP.2PL.1SG.DAT of yes that 1SG.ACC make.IND.FUT.2PL ULTRA-happy
'Tell me yes, you will make me extremely happy.'

As the examples show, the six prefixes are broadly interchangeable in their core intensifying function: substituting one for another in most of the sentences above yields a grammatical and semantically equivalent result. This functional equivalence makes them prima facie rivals in the sense defined above. While previous research on Italian intensifying prefixes has largely cataloged these forms (Montermini 2008), analyzed them in isolation (Napoli 2012, Micheli & Verdelli 2025), or compared them more descriptively (Calpestrati 2017), their competitive distribution has not yet been systematically addressed. It thus remains an open empirical question whether the outputs of these prefixes are truly semantically indistinguishable, or whether subtle but systematic differences persist beneath the surface equivalence – and if so, how these differences are distributed across the system.

Capturing such scalar distinctions calls for an analytical framework sensitive to degrees of semantic proximity among prefixes rather than one that reduces competition to categorical outcomes. To this end, we examine rivalry through a distributional semantic framework, modeling the semantic output of each prefix and the transformations it induces as vectors in a high-dimensional semantic space. In addition, we propose an information-theoretic operationalization of rivalry, quantifying the degree to which observing a semantic transformation reduces uncertainty about which process produced it. Under this formulation, rivalry is measured as a continuous quantity – the mutual information between semantic shifts and process identity – rather than a binary verdict, enabling graded comparisons both within and across prefix pairs.

Building on this perspective, the study pursues three interrelated empirical questions:

(i) Do rival prefixes occupy overlapping or distinct regions of distributional semantic space, thus providing evidence for stronger or weaker competition?

(ii) Do rival prefixes induce systematically different semantic shifts from base to derivative, revealing potential functional differentiation within the system?

(iii) To what extent can the identity of a prefix be inferred from the semantic shift it induces, and does this degree of distinguishability vary across specific pairs of rival prefixes?

The paper is organized as follows. The data are presented in §2, while §§3 and 4 report a series of analyses based on vectors of derivatives and vector offsets, respectively. Finally, §5 discusses the findings, their theoretical implications, and the limitations of the study, while §6 briefly concludes.

2 Data

The present study draws on two main resources: (i) a curated inventory of Italian intensified adjectives formed with the six prefixes under investigation and their bases, and (ii) word embeddings induced from corpus data.

2.1 Intensified derivatives

The focus of this study is six Italian intensifying prefixes: *arci-*, *extra-*, *iper-*, *stra-*, *super-*, and *ultra-*. Although these prefixes also combine with non-adjectival bases, the present analysis is restricted to adjectival ones, selected for two main reasons. First, as a lexical category whose semantics is particularly amenable to gradation, adjectives constitute a natural domain for the expression of INTENSIFICATION (Grandi 2017). Second, among the six examined prefixes, there is a marked preference for adjectival bases (Iacobini 2004, Montermini 2008).

The starting point for data collection was *InTens* (Lacić 2025), a manually curated dataset of intensified derivatives formed with the six prefixes under investigation. Constructed from the `itWaC` web corpus (Baroni et al. 2009) – to this day, the largest copyright-free, freely downloadable corpus of Italian – *InTens* comprises 48,069 tokens distributed across 3,683 types and provides an inventory of adjectival tokens whose prefixal element has been manually annotated for semantic function, distinguishing intensifying usage from other semantic functions. In the present study, we consider only those derivatives in which the prefix realizes an intensifying function.[1]

[1] To clarify that the restriction to intensifying uses does not entail an assumption that the derivatives included in the analysis are necessarily monosemous, we assessed the possibility that some derivatives may exhibit both intensified and non-intensified readings. To this end, we manually inspected a random sample of 300 tokens from 20 derivative types per prefix (with a frequency of at least 10 occurrences). No cases of genuine polysemy involving both intensified and non-intensified readings were found.

The dataset includes derivatives in three orthographic variants: monostring (e.g., *stra-bello* 'super nice'), hyphenated (*stra-bello*), and multi-string (*stra bello*).[2] Rather than extracting intensifying derivatives directly from raw corpus data, the present study relies on *InTens* as a controlled lexical filter, ensuring that all sampled types correspond to bona fide intensifying uses of the prefix. From this resource, all types annotated as having an intensifying function were matched against the vocabulary of the trained embedding model, described in §2.2 below.

2.2 Distributional vectors

Distributional semantics, a mainstream methodological paradigm in contemporary linguistics (Boleda 2020, Lenci & Sahlgren 2023), rests on the Distributional Hypothesis (Harris 1954, Firth 1957), according to which words that occur in similar linguistic contexts tend to have similar meanings. Words are accordingly represented as vectors encoding their contextual distributions, and the semantic relatedness of two words is typically measured as the cosine similarity between their vectors, with higher values indicating greater contextual and, by extension, semantic overlap. Since Mikolov & Chen, et al. (2013), prediction-based neural models such as `word2vec`, which learn dense, low-dimensional representations known as *word embeddings*, have become standard. Building on Marelli & Baroni (2015), a growing body of research has applied distributional methods to the study of word formation (Lapesa et al. 2018, Stupak & Baayen 2022, Kotowski & Schäfer 2023, Schäfer 2025, Huyghe & Varvara 2025, among others), demonstrating that such representations provide a powerful means of modeling the semantics of morphological processes at scale.

In the present study, distributional vectors were induced from the previously introduced `itWaC` corpus. Since the focus lies on derivation, the aim was to abstract away from distributional differences among inflected forms of the same lexeme and instead capture how lexemes relate to one another. Accordingly, the entire corpus was reduced to lexeme representations by merging all tokens sharing the same lemma and part-of-speech tag – as provided with the corpus – into a single token prior to training. Moreover, to accommodate the orthographic variability typical of these forms, the three orthographic variants of a given derivative were merged into one token, so that their occurrences were pooled into a single distributional context set and a single vector was learned from them directly regardless of spelling conventions. Using the `gensim` (Řehůřek & Sojka 2010) implementation of the

[2] The presence of hyphenated and, especially, multi-string spellings raises questions concerning the boundary between prefixal formations, first members of compounds, and syntactically independent intensifying modifiers, which is arguably not categorical for these elements. We adopt the term *prefix* because it is the established designation for these formatives in the literature on (Italian) intensification (Montermini 2008, Calpestrati 2017, Grandi 2017, Kamber & Huyghe 2023, Micheli & Verdelli 2025, a.o.), which provides the comparative context for the present study. Crucially, however, the analysis does not depend on the grammatical status of the forms under examination: its object is rivalry among intensifying processes, a question that arises regardless of whether a given realization is best analyzed as a bound prefix, a compound member, or a free modifier. Accordingly, orthographic realization is not used as a criterion for grammatical status, and monostring, hyphenated, and multi-string spellings are treated as variants of the same intensifying construction.

`word2vec` algorithm, we trained 12 models, varying the architecture (continuous bag-of-words, CBOW, vs. Skip-gram), the context window size (2 vs. 5 tokens), and the dimensionality (200, 300, or 500), with five training epochs and, given the low overall frequency of intensified derivatives, a minimum frequency threshold of 10. Evaluated against the Italian versions (Leviant & Reichart 2015) of SimLex-999 (Hill et al. 2015) and WordSim-353 (Finkelstein et al. 2002), the Skip-gram model with a 5-token window and 500 dimensions provided the best overall performance across the benchmarks and was selected for all subsequent analyses (see the supplementary materials for full evaluation results).

Intersecting the filtered intensified *InTens* derivatives and their bases with the vocabulary of the selected embedding model yielded 401 base-derivative pairs represented in the vector space. Their distribution across prefixes is shown in Table 1. It is worth noting that multiple prefixes frequently combine with the same base adjective: 95 of the 274 unique bases in the 401-pair sample (34.7%) occur with more than one prefix, with some bases (e.g., *famoso* 'famous', *fino* 'fine, thin', *potente* 'powerful') combining with up to four different intensifiers.

Table 1: Distribution of intensified types in the vector space

arci-	*extra-*	*iper-*	*stra-*	*super-*	*ultra-*	SUM
8	14	88	33	145	113	401

The dataset is considerably imbalanced across prefixes. Rather than sub-sampling to the smallest common size – which would reduce the data to only eight types per prefix – we adopted a maximalist approach, retaining all available types to capture the full semantic spectrum of each category and avoid distortions introduced by artificial balancing (cf. Huyghe & Wauquier 2021). Two prefixes are nonetheless severely underrepresented in the embedding vocabulary: *arci-* ($n = 8$ types) and *extra-* ($n = 14$ types). For these classes, estimates that depend on distributional density or centroid stability are inherently less reliable than for the four better-attested prefixes, and their results should therefore be interpreted with appropriate caution. Nevertheless, *arci-* and *extra-* are retained in the analyses of derivative vectors and offset centroids (§3 and §4.1), as their behavior is itself theoretically informative. They are, however, excluded from the instance-level classification and information-theoretic analyses (§4.2), which require more data per class than these two prefixes provide.

3 Insights from vectors of derivatives

With this data in place, we first turn to vectors of prefixed adjectives, allowing us to examine rivalry in terms of morphological outputs. This section develops the empirical analysis in three steps. First, §3.1 examines the extent to which intensified derivatives retain the distributional meaning of their bases across prefixes. Second, §3.2 investigates the internal structure of derivative space, assessing the cohesion and permeability of each prefix-specific

subclass. Third, §3.3 tests whether derivative vectors encode a detectable prefix-specific signal despite the expected semantic convergence among rival prefixes. Together, these analyses provide a multi-level characterization of intensified derivatives in the embedding space and establish the empirical baseline for the offset-based investigation presented in §4.

3.1 Base-derivative semantic relatedness

To quantify semantic relatedness between a base and its derivative, we compute cosine similarity for each base-derivative pair. As noted in §2.2, higher cosine values indicate stronger semantic retention, that is, the derivative remains distributionally close to its base. Figure 1 presents the distribution of cosine similarity values across the six prefixes.[3]

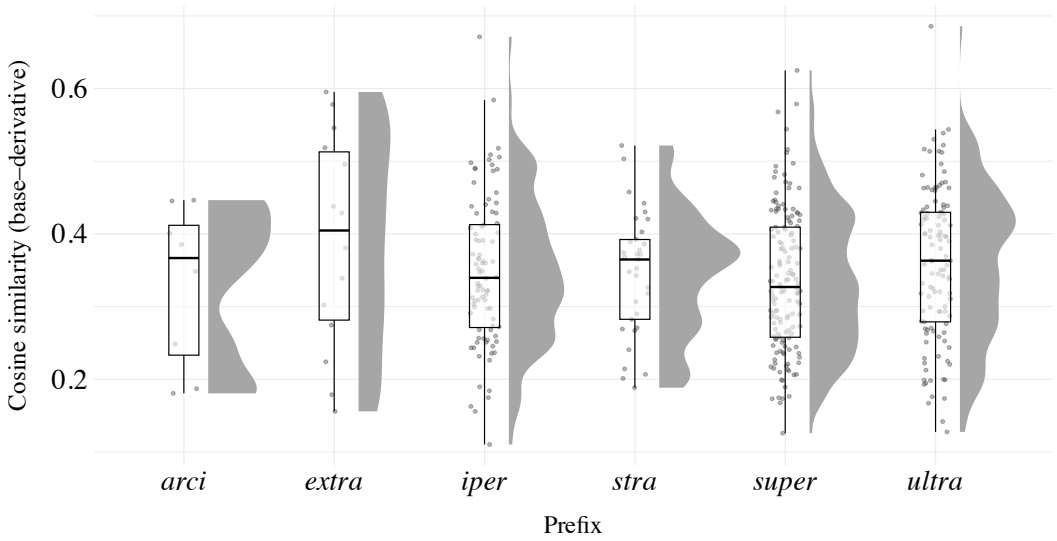

Figure 1: Distribution of cosine similarities for base-derivative pairs. For each prefix, the plot shows (left to right) a box-and-whisker plot summarizing the median and interquartile range, jittered data points for individual base-derivative pairs, and a kernel density (violin) plot of the same distribution. The irregular, multi-modal violins for *arci* ($n = 8$) and *extra* ($n = 14$) reflect kernel density estimation on small samples and should therefore be interpreted with caution.

Descriptively, median cosine similarity scores are relatively close, ranging from *super*'s lowest value of 0.327 to *extra*'s highest of 0.405, suggesting broadly similar degrees of base retention across prefixes. However, because multiple prefixes frequently combine with the same base adjective, observations sharing a base are not statistically independent. To account for this dependency structure, we fit a linear mixed-effects model predicting cosine similarity from prefix identity, with base identity included as a random intercept. This allows

[3] For ease of reading, the hyphen indicating prefix status is omitted in running text throughout the empirical analyses (§§3–4), e.g., *super* instead of *super-*.

the model to distinguish systematic differences between prefixes from variation attributable to the particular lexical bases they combine with.

The random-effect variance for base ($\sigma^2 = 0.007$) exceeds the residual variance ($\sigma^2 = 0.003$), indicating that differences between bases account for a substantial share of the variation in cosine similarity. An omnibus model comparison reveals only a marginal improvement in fit when prefix is added ($\chi^2(5) = 10.56$, $p = .061$), suggesting that prefix identity contributes relatively little explanatory power once differences between bases are taken into account. At the level of individual prefixes, none shows a significant increase in base retention relative to the reference level *super* – selected because it is the most frequent prefix in the sample – after correction for multiple comparisons. This is consistent with the omnibus test's marginal result. Estimated marginal means likewise show substantial overlap in adjusted cosine similarity across prefixes: *super* and *iper*, for instance, have nearly identical adjusted values, while *extra* shows the highest estimated base retention, although this difference is not statistically significant.

Despite this inferential overlap, qualitative inspection of the highest-scoring derivatives reveals characteristic tendencies in the types of adjectives that remain distributionally closest to their bases under each prefix. Table 2 lists the ten derivatives per prefix (eight for *arci*) with the highest base-derivative cosine similarity.

The prefix *extra* shows the highest median cosine similarity (0.405) and primarily intensifies adjectives denoting physical properties, often within a commercial domain (*fondente* 'dark', *fresco* 'fresh', *duro* 'hard', *lungo* 'long', *fino* 'fine'). Following with a median of 0.367, *arci* combines mostly with adjectives expressing epistemic states (*noto* 'known', *conosciuto* 'known') and emotional conditions (*stufo* 'fed up', *contento* 'pleased'). Close behind, *stra* (0.365) favors adjectives with emotional or emphatic connotations (*meritato* 'deserved', *figo* 'cool', *mitico* 'legendary', *maledetto* 'cursed'). The fourth highest-scoring prefix is *ultra* (0.363), which straddles two distinct domains: technical contexts signaling advanced functionality or heightened quality (*regolabile* 'adjustable', *portatile* 'portable', *moderno* 'modern'), and political or ideological ones (*liberista* 'free-market', *nazionalista* 'nationalist', *conservatore* 'conservative', *reazionario* 'reactionary'). Next, *iper* (0.340) similarly gravitates toward scientific and technical registers (*calorico* 'caloric', *proteico* 'protein-rich', *reattivo* 'reactive'), with considerable overlap with *ultra* across both political and technical domains. Finally, *super* displays the lowest median cosine similarity of 0.327, most frequently intensifying visually striking or functionally enhanced properties (*sexy*, *attillato* 'tight-fitting', *palestrato* 'muscular', *blindato* 'armored', *accessoriato* 'equipped'), while also appearing in descriptive and technical contexts of enhanced quality (*grandangolare* 'wide-angle', *panoramico* 'panoramic', *calorico* 'caloric'). This prefix-specific selectivity aligns with cross-linguistic observations on intensive prefixation raised by Rainer (2015, 1344), who notes that among pan-European intensive prefixes, "each has its particular, synchronically unmotivated preferences concerning potential bases (*extra-* for commercial, *ultra-* for political terms, etc.)".

Table 2: Ten highest-ranking derivatives for each prefix, ordered by cosine similarity between derivative and base. The median similarity across *all* derivatives for that prefix is given for reference above each column, alongside the total number of derivatives (n) available for that prefix. Additionally, the cosine similarity for each individual base-derivative pair is given in parentheses. Only eight derivatives are available for *arci*.

arci Median: 0.367 ($n = 8$)	***extra*** Median: 0.405 ($n = 14$)	***iper*** Median: 0.340 ($n = 88$)
noto 'known' (.45)	*fondente* 'dark' (.58)	*liberista* 'free-market' (.67)
stufo 'fed up' (.45)	*colto* 'cultured' (.50)	*calorico* 'caloric' (.58)
contento 'pleased' (.40)	*lucido* 'shiny' (.44)	*proteico* 'protein-rich' (.52)
famoso 'famous' (.39)	*fresco* 'fresh' (.43)	*accessoriato* 'equipped' (.51)
convinto 'convinced' (.35)	*duro* 'hard' (.38)	*cromatico* 'chromatic' (.50)
sicuro 'sure' (.25)	*lungo* 'long' (.34)	*consumistico* 'consumerist' (.49)
nuovo 'new' (.19)	*disponibile* 'available' (.30)	*nazionalista* 'nationalist' (.49)
conosciuto 'known' (.18)	*interessante* 'interesting' (.27)	*lipidico* 'lipidic' (.49)
	fino 'fine' (.22)	*colorato* 'colored' (.49)
	vecchio 'old' (.18)	*reattivo* 'reactive' (.47)
stra Median: 0.365 ($n = 33$)	***super*** Median: 0.327 ($n = 145$)	***ultra*** Median: 0.363 ($n = 113$)
meritato 'deserved' (.52)	*sexy* (.63)	*liberista* 'free-market' (.69)
colmo 'full' (.50)	*attillato* 'tight-fitting' (.58)	*nazionalista* 'nationalist' (.54)
figo 'cool' (.46)	*grandangolare* 'wide-angle' (.57)	*regolabile* 'adjustable' (.54)
mitico 'legendary' (.43)	*palestrato* 'muscular' (.54)	*portatile* 'portable' (.52)
contento 'pleased' (.42)	*blindato* 'armored' (.52)	*chic* (.52)
potente 'powerful' (.42)	*panoramico* 'panoramic' (.51)	*conservatore* 'conservative' (.51)
maturo 'mature' (.40)	*sfruttato* 'exploited' (.50)	*reazionario* 'reactionary' (.51)
convinto 'convinced' (.39)	*accessoriato* 'equipped' (.49)	*morbido* 'soft' (.49)
maledetto 'cursed' (.39)	*affollato* 'crowded' (.49)	*moderno* 'modern' (.49)
carico 'loaded' (.39)	*calorico* 'caloric' (.48)	*liberale* 'liberal' (.48)

3.2 Derivative vector density and centroid neighborhood

In addition to the relationship between bases and derivatives, it is instructive to examine the structure of the derivative space, as the degree to which derivatives formed with the same prefix cluster together provides an indication of each subclass's semantic cohesion, that is, the extent to which members of a prefix-based subclass share similar semantic properties and occupy a relatively compact region of the semantic space (cf. Huyghe & Wauquier 2021).

Figure 2 presents the distribution of pairwise cosine similarities among intensified derivatives formed with the same prefix.[4] The rationale behind this measure is straightforward:

[4]For each prefix, all unique pairs are considered. For instance, *arci* with eight derivatives yields $\frac{8\times(8-1)}{2} = 28$ pairs. While the number of pairwise comparisons scales quadratically with the number of derivatives, potentially inflating the dataset for more frequent prefixes, the medians used here provide a comparable summary statistic across prefixes regardless of the number of pairs.

the more semantically coherent a subclass of intensified derivatives is, the more tightly its members should cluster in vector space.

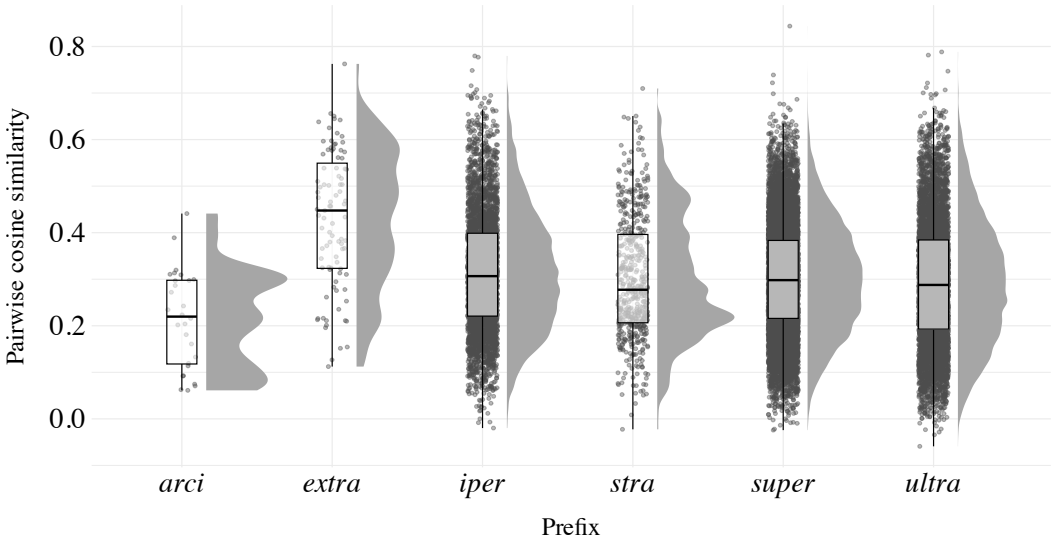

Figure 2: Distribution of pairwise cosine similarity among derivatives within each prefix, reflecting the internal semantic density of each subclass in the vector space. Higher values indicate tighter clustering among derivatives sharing the same prefix.

The distributions reveal substantial differences in internal semantic density. *Extra* forms the most tightly clustered derivative subclass (Median = 0.448), followed by *iper* (0.307), while the remaining prefixes cluster in an intermediate range, with *super* (0.298) and *ultra* (0.288) close to each other, and *stra* slightly more dispersed (0.277). On the other hand, *arci* stands out as the least dense subclass (0.220), suggesting the greatest semantic dispersion among its members. Interquartile ranges (IQR) reveal additional differences in within-prefix variability: *extra* shows the broadest spread (IQR = 0.226), whereas *super* displays the narrowest distribution (0.167), with the remaining prefixes clustering around IQR values between 0.178 and 0.191.

Since pairwise similarities are not independent – each derivative participates in multiple pairs – we model pairwise proximity values using a crossed random-effects linear mixed model, with random intercepts for both members of each pair ($\text{Proximity}_{ij} \sim \text{Prefix} + (1 \mid \text{word}_i) + (1 \mid \text{word}_j)$), and *super* as the reference level. This allows us to distinguish systematic differences in internal semantic density between prefixes from variation attributable to the particular derivatives being compared. The random-effect variances for word_i ($\sigma^2 = 0.003$) and word_j ($\sigma^2 = 0.003$) indicate substantial pairwise dependence, which is appropriately accounted for by the model, although residual variation remains the dominant component ($\sigma^2 = 0.010$). An omnibus likelihood-ratio test shows that prefix

identity significantly improves model fit ($\chi^2(5) = 226.7$, $p < .001$), indicating systematic differences in semantic density across prefixes once pairwise dependence is controlled.

Relative to *super*, *extra* exhibits markedly higher internal density ($\beta = 0.107$, $p < .001$), while *iper* and *stra* also show modest increases ($\beta = 0.034$, $p < .001$; $\beta = 0.023$, $p = .003$). By contrast, *arci* is significantly less dense than *super* ($\beta = -0.146$, $p < .001$), while *ultra* shows a small but significant decrease ($\beta = -0.007$, $p = .018$). Estimated marginal means (EMMs), which account for the dependency structure, confirm the same ordering: *extra* has the highest adjusted density (EMM = 0.411, 95% CI [0.378, 0.444]), whereas *arci* has the lowest (EMM = 0.159, 95% CI [0.115, 0.203]). The remaining prefixes form an intermediate group (*iper* = 0.338; *stra* = 0.327; *super* = 0.305; *ultra* = 0.298).

While the density analysis reveals how tightly derivatives formed with a given prefix cluster together, it does not tell us where these clusters are located in semantic space or how similar the corresponding processes are to one another. To address this complementary question, we draw on Wauquier et al. (2020) and represent each derivational process by its centroid, calculated as the average vector of all derivatives formed with the same prefix. The centroid thus captures the central tendency of a prefix's derivative distribution, providing an abstract representation of its typical semantic profile while abstracting away from variation among individual derivatives.

Table 3 presents the cosine similarities between the derivative centroids of the six prefixes. Higher values indicate greater semantic overlap between the corresponding processes.

Table 3: Cosine similarity between derivative centroids for each prefix

	arci	*extra*	*iper*	*stra*	*super*	*ultra*
arci	1.000					
extra	0.460	1.000				
iper	0.590	0.662	1.000			
stra	0.760	0.575	0.728	1.000		
super	0.663	0.712	0.853	0.842	1.000	
ultra	0.613	0.708	0.907	0.739	0.914	1.000

Prefix pairs involving *super* display exceptionally high cosine similarity, notably *super–ultra* (0.914) and *super–iper* (0.853), paralleled by *iper–ultra* (0.907). These values indicate a tight alignment of centroid profiles, suggesting that derivatives formed with these three prefixes occupy largely overlapping semantic space. By contrast, *arci* derivatives emerge as the most semantically distinct within the prefixal system, consistently recording the lowest cosine similarities with derivative centroids of all other prefixes, with a mean inter-prefix cosine similarity of 0.617. Even its highest value, observed with *stra* (0.760), remains well below the tight *super*/*ultra*/*iper* cluster, suggesting that *arci* occupies a more specialized subdomain rather than participating in the core competitive space. In addition, *extra* derivatives show moderate to high similarities with those of other prefixes (ranging from 0.460 with *arci* to 0.712 with *super*), placing it in a relatively independent but still central position. Finally, the *stra* derivative centroid occupies an intermediate role: it aligns strongly with *super* (0.842)

and shows solid similarity with *ultra* (0.739), *iper* (0.728), and *arci* (0.760), but is less similar to *extra* (0.575). This suggests that *stra* participates in multiple subclusters within the intensifying space.

Following the quantification of similarity between the derivative centroids, it is instructive to complement these findings with a qualitative examination. However, since centroids are abstract entities, they do not correspond to the vector representation of any actual lexeme. To explore the semantic content captured by each centroid, it is thus necessary to examine its nearest neighbors, as they can be seen as representatives of the morphosemantic class denoted by the centroid (Huyghe & Wauquier 2020).

The choice of the number of nearest neighbors to examine is inherently arbitrary, as no universally established threshold exists. Here, we consider the 100 nearest neighbors of each prefix-specific derivative centroid – a number broad enough to allow for generalization yet constrained enough to enable a detailed qualitative assessment of the neighborhood. In doing so, we restrict our analysis to adjectives.[5] The full lists of the 100 nearest neighbors of each centroid are provided in the supplementary materials; for reference, the 20 nearest neighbors are given in Appendix A (Tables A1–A6).

Because the derivative sets associated with different prefixes partially overlap and occupy nearby regions of the embedding space, the nearest neighbors of a prefix-specific centroid are not expected to consist exclusively of derivatives bearing the same prefix. Table 4 reports the composition of the 100-nearest-neighbor sets, broken down into: (i) adjectival derivatives formed with the same intensifying prefix as the centroid; (ii) adjectival derivatives formed with the other five prefixes under investigation; and (iii) adjectives without an intensifying prefix.

Table 4: Composition of the 100 nearest adjectival neighbors of each derivative centroid

	Same-prefix derivatives	Other-prefix derivatives	Non-intensifying-prefix adjectives
arci	8	71	21
extra	13	41	46
iper	37	63	0
stra	23	69	8
super	57	43	0
ultra	40	60	0

[5] A brief inspection of the 100-type distributional neighborhoods without word-class restriction reveals noteworthy patterns. For *iper*, *super*, and *ultra*, no non-adjectival elements are found, suggesting their strong semantic affinity for adjectives even when unconstrained by part-of-speech considerations. By contrast, for *arci* and *stra*, several high-frequency verbs such as *dire* 'say', *credere* 'believe', and *pensare* 'think' appear among the nearest neighbors. Particularly interesting is the behavior of *extra*: its 100-type distributional neighborhood includes 31 nouns and verbs, all of which are culinary terms – e.g., *vanillina* 'vanilla flavoring', *julienne* 'French cut', *confettura* 'jam', and *margarinare* 'mix with margarine'. We hypothesize that this phenomenon stems primarily from the highly frequent adjective *extravergine* 'extra virgin', used almost exclusively in the collocation *olio extravergine* 'extra virgin olive oil', thereby pulling *extra*-derivatives toward the culinary region of the semantic space.

Clear differences emerge in neighborhood homogeneity. The centroids of *iper*, *super*, and *ultra* form fully "closed" neighborhoods, composed exclusively of adjectives formed with intensifying prefixes, with same- vs. other-prefix splits of 37/63, 57/43, and 40/60, respectively, indicating compact and cohesive regions of semantic space. By contrast, *arci*, *extra*, and *stra* centroids display more heterogeneous profiles – *arci* contains only 8 same-prefix derivatives alongside 21 adjectives without an intensifying prefix, *extra* shows 13 same-prefix derivatives alongside 46 non-intensifying-prefix items, and *stra* includes 23 same-prefix derivatives and 8 non-intensifying-prefix words – suggesting greater permeability to the broader adjectival space.

Because prefixes differ strongly in type frequency, and because a centroid built from n_p derivatives is itself somewhat biased toward those same derivatives – an unavoidable property of averaging, more pronounced for smaller classes – it is useful to compare each prefix's observed same-prefix count to an empirical null obtained by repeatedly ($B = 1{,}000$) sampling n_p random adjectives, computing their centroid the same way, and counting how many reappear among its own top-100 nearest neighbors. This procedure helps to isolate genuine prefix-driven clustering from averaging artifacts. As shown in Table 5, five of the six prefixes – *extra*, *iper*, *stra*, *super*, and *ultra* – show significantly more same-prefix neighbors than this null predicts (Holm-corrected $p < .001$ for all five), though the size of the effect varies considerably, from *extra* ($z = 2.4$) to *super* ($z = 18.0$). *Arci*, however, is statistically indistinguishable from chance ($p = .780$): its heterogeneous neighborhood reflects an absence of detectable cohesion, though with $n = 8$ the test's power to detect a genuine but modest effect is limited (SD = 0.64 under the null).

Table 5: Observed vs. expected (Monte Carlo null, $B = 1{,}000$) same-prefix neighbor counts among each centroid's top-100 nearest neighbors

	Obs	Exp (null)	p_{Holm}
arci	8	7.70	.78
extra	13	10.22	$< .001$
iper	37	6.84	$< .001$
stra	23	7.64	$< .001$
super	57	8.44	$< .001$
ultra	40	7.39	$< .001$

Considerable overlap also exists across neighborhoods, with certain derivatives acting as nearest neighbors of multiple centroids. Table 6 reports the number of shared items among the sets of the 100 nearest adjectival neighbors for each centroid pair. The greatest overlap is observed between the centroids of *arci* and *stra* (71 shared neighbors), followed by *super* and *ultra* (59), *iper* and *ultra* (55), and *stra* and *super* (53). These values indicate that the corresponding centroids occupy largely overlapping regions of the intensification space. Importantly, the high *arci*–*stra* overlap remains stable when varying the neighborhood size (e.g., top 300 or top 500 neighbors), suggesting that it is not an artifact of rank truncation

but reflects a genuine concentration of both prefixes within the same dense region of the intensification space.

Table 6: Overlap between the 100 nearest adjectival neighbors of each derivative centroid

	arci	*extra*	*iper*	*stra*	*super*	*ultra*
arci	x					
extra	13	x				
iper	17	15	x			
stra	71	18	18	x		
super	37	37	27	53	x	
ultra	24	36	55	30	59	x

Turning to a qualitative inspection of the 100 closest neighbors, we use the neighborhoods as an exploratory tool to identify potential semantic affinities associated with each prefix. The following observations should therefore be understood as suggestive patterns emerging from the distributional neighborhoods rather than as exhaustive semantic profiles.

Starting with *arci*, its neighbors – mostly derivatives formed with adjectives such as *stufo* 'fed up', *contento* 'pleased', and *convinto* 'convinced' – suggest an association with emotional intensity and conviction, while items like *noto* 'known', *famoso* 'famous', and *rivale* 'rival' point to notions of reputation or familiarity. These examples suggest an affinity between *arci* and semantic domains involving personality traits, emotional states, and reputation.

Extra shows a strong connection to culinary terminology, with neighbors like *brut*, *fino* 'fine', and *fondente* 'dark' (as in *cioccolato fondente* 'dark chocolate'). Terms such as *forte* 'strong', *vergine* 'virgin', *lucido* 'shiny', and *fresco* 'fresh' further highlight its association with enhanced product qualities. From this perspective, *extra* is primarily associated with notions of high quality, often linked to the physical and sensory properties of food, as well as to the commercial domain more broadly.

As for *iper*, adjectives like *espressivo* 'expressive', *spettacolare* 'spectacular', and *potente* 'powerful' point to *iper*'s focus on functional or performative intensification, whereas words such as *industriale* 'industrial', *lipidico* 'lipidic', and *cromatico* 'chromatic' suggest an affinity for industrial qualities and technical terminology.

Examining *stra*, we observe a connection to popularity, sociability, and positive personality traits, evident in words like *simpatico* 'likeable', *contento* 'pleased', and *famoso* 'famous'. Additionally, an inclination toward strong emotions and social attributes, as well as a more informal, everyday use of intensification, is reflected in the presence of terms such as *figo* 'cool', *maledetto* 'cursed', and *felice* 'happy'. Overall, the observed neighbors suggest that *stra* is primarily associated with social traits and hyperbolic emotions.

Next, *super* emerges as highly versatile, applying both to physical qualities and to positive attributes, making its use broader than that of the aforementioned prefixes. Words

like *divertente* 'entertaining', *romantico* 'romantic', and *affidabile* 'reliable' indicate a general intensification of positive traits, while terms like *duro* 'hard', *forte* 'strong', and *resistente* 'resistant' connect *super* to physical robustness and resilience.

Finally, *ultra* is associated with attributes denoting professionalism, such as *preciso* 'precise', *professionale* 'professional', and *determinato* 'determined'. In addition, terms such as *raffinato* 'refined' and *competitivo* 'competitive' emphasize sophistication and high standards. Consequently, the observed neighborhood of *ultra* suggests a strong association with technical, professional, and high-standard qualities, partially overlapping with *iper* but with a more high-end connotation.

Taken together, these qualitative observations suggest that, despite their shared intensifying function, the six prefixes occupy partially differentiated regions of the semantic space. The observed tendencies range from more emotional and colloquial associations (*arci*, *stra*) to more technical and domain-specific associations (*iper*, *ultra*, *super*), while *extra* displays a particularly distinctive association with food-related and commercial contexts.

3.3 Prefix discrimination in derivative space

The preceding analyses have shown that derivative sets differ in internal semantic density and that several prefix centroids are highly similar, pointing to substantial semantic overlap among rival prefixes. This raises a complementary question: *to what extent can intensified derivatives nevertheless be discriminated by prefix in distributional space?* If prefixes exert at least partially distinct semantic effects on their outputs, derivatives sharing the same prefix should retain some prefix-specific signal in semantic space. Conversely, if rivalry entails genuine semantic convergence, prefix-based discrimination should be limited.

We address this question using linear discriminant analysis (LDA), a supervised classification method that identifies linear combinations of the original variables that maximize separation between predefined groups. LDA therefore provides a quantitative assessment of how much information about prefix identity can be recovered from the original 500-dimensional vectors (cf. Stupak & Baayen 2022, Plag et al. 2025). Specifically, we apply LDA to PCA-transformed data. Principal component analysis (PCA) first transforms the original 500-dimensional space into a smaller set of mutually uncorrelated dimensions (principal components), ordered according to the amount of variation they explain. LDA then identifies combinations of these components that best distinguish the different prefix classes. In this way, the analysis asks whether the semantic information encoded in the derivative vectors contains a systematic signal of which prefix produced a given derivative.[6]

LDA was trained on progressively larger subsets of principal components (5, 10, 20, 35, 50, 75, 100, 200, 300, and 400), with performance evaluated using leave-one-out cross-validation (LOO-CV). The first 10 components account for 34.5% of the total variance,

[6] A technically inclined reader may note that LDA is a *generative* classifier: it fits an explicit probability distribution to each class. Given the strong class imbalance, however, the assumption of a shared covariance structure across classes is questionable in a high-dimensional space. We therefore do not interpret the fitted distributional parameters and rely instead on cross-validated classification accuracy, which does not require the generative assumptions to be fully satisfied.

increasing to 52.1% by PC25 and 68.1% by PC50, with substantial variation remaining in the higher components. Classification accuracy reaches its maximum with 20 components (59.4%), while 10 components already yield 57.4%. Both figures substantially exceed the majority-class baseline of 36.2%, obtained by always predicting *super*. Thus, although prefix identity is far from perfectly recoverable, the derivative vectors contain a clear, above-baseline signal of prefix identity.

Interestingly, adding further components does not improve performance. Accuracy declines steadily beyond 20 components, reaching 16.9% when all 400 components are included. The decline in performance as more components are introduced is consistent with a bias-variance trade-off in linear classification with a limited sample size. Although the PCA components are mutually orthogonal, estimating class boundaries in an increasingly high-dimensional space introduces more parameters and can therefore increase estimation variance and reduce generalization to unseen observations. The results consequently suggest that much of the prefix-related signal is concentrated in a relatively low-dimensional part of the semantic space, with around 20 components providing the best balance between retaining relevant information and avoiding unnecessary complexity.

The LDA findings thus indicate that prefix identity leaves a detectable but limited distributional signature: it can be predicted from the semantic representations at rates above the majority baseline, albeit far from perfectly. The extent of discrimination is likely constrained by multiple factors. First, distributional representations of low-frequency words are known to be more noise-prone, and the frequency threshold (frequency ≥ 10) used here is relatively low. Second, multiple prefixes frequently attach to the same base, and since derivative vectors encode substantial base semantics, base-sharing derivatives tend to cluster regardless of which prefix they carry. In such cases, the prefix-specific signal is partially obscured by base-level similarity – a limitation that motivates the vector-offset approach taken in §4. Ultimately, and perhaps most importantly, the moderate yet non-trivial classification ceiling is coherent with the notion of gradient rivalry. The six prefixes examined here are semantically convergent by design: they compete in expressing near-equivalent intensifying meanings. In such a system, substantial overlap in distributional space is expected. Consequently, the inability of LDA to achieve categorical separation should not be read as methodological failure, but rather as empirical evidence of semantic convergence among the examined processes.

4 Insights from vector offsets

The preceding analysis showed that derivative vectors do not form categorically distinct prefix-based groups, with *iper*, *super*, and *ultra* occupying substantially overlapping regions of semantic space. As noted above, this pattern may partly reflect the influence of base semantics: derivative vectors inherit much of their semantic structure from their bases, so groupings may primarily reflect base similarity rather than prefix-specific effects. The present analysis therefore asks whether the same pattern persists when base semantics are factored out – and, if so, whether the signal becomes sharper.

To isolate the semantic contribution of prefixation more directly, we adopt a vector-offset approach. Since distributional models do not provide representations for affixes, this contribution is commonly approximated through vector offsets, that is, the difference between the embedding of a derivative and that of its base (e.g., $\vec{\textit{strabello}} - \vec{\textit{bello}}$ as one instance of the shift induced by *stra*). This approach builds on the observation that embeddings encode relations between words as approximately constant vector differences (Mikolov & Yih, et al. 2013): base-derivative pairs linked by the same process are therefore expected to exhibit similar offsets (Bonami & Paperno 2018).

In what follows, we first examine the similarity between the average semantic transformations associated with each prefix, that is, their average vector offsets (§4.1). We then assess the extent to which prefix identity can be recovered from individual offsets using a supervised classifier (§4.2).

4.1 Offset centroid similarity

Individual vector offsets are inherently noisy due to the idiosyncratic properties of the lexeme pairs from which they are derived, which inevitably weakens their predictive quality (Mickus et al. 2019, Bonami & Guzmán Naranjo 2023). Accordingly, we first abstract away from individual offsets and focus on the aggregate behavior of each derivational pattern. Specifically, for each prefix, we compute the average vector offset – that is, the offset centroid – by averaging all offsets associated with that prefix. This approach assumes that word pairs related by the same morphological process exhibit a broadly consistent semantic shift, such that the resulting average vector captures the general direction of that shift.

Averaging vector offsets nevertheless involves an important limitation: it does not account for potential polyfunctionality of morphological processes. If a given process realizes multiple distinct semantic functions, the corresponding offsets may cluster into separate regions of the vector space. In such cases, the average vector would not accurately represent any individual transformation, but rather an abstracted direction that may obscure meaningful internal structure. In the present study, however, this concern is substantially mitigated by the design of the dataset. As highlighted in §2.1, the derivatives analyzed here are selected specifically for their intensifying interpretation, such that the semantic contribution of each prefix is restricted to a relatively homogeneous functional domain. While minor contextual nuances may remain, the major axis of variation is shared INTENSIFICATION rather than qualitatively distinct semantic functions.

We then assess the similarity between the offset centroids using cosine similarity. Table 7 reports the resulting pairwise similarities.

The cluster structure observed in the derivative-centroid analysis (see Table 3) is not only replicated but sharpened in offset space. The two similarity matrices are strongly correlated (Mantel test, Spearman's $\rho = 0.91$, $p = .0014$, exact test over all 720 permutations of prefix labels), confirming that the relative similarities among prefixes are largely preserved once base semantics are factored out. Within this shared structure, *super*, *ultra*, and *iper* remain the most tightly grouped prefixes, with pairwise similarities ranging from 0.758 to 0.829, indicating substantial overlap in the direction of the semantic transformations they induce.

Table 7: Pairwise cosine similarities between vector offset centroids

	arci	*extra*	*iper*	*stra*	*super*	*ultra*
arci	1.000					
extra	0.277	1.000				
iper	0.308	0.463	1.000			
stra	0.487	0.442	0.506	1.000		
super	0.395	0.539	0.758	0.630	1.000	
ultra	0.366	0.592	0.767	0.546	0.829	1.000

Notably, the *super–ultra* similarity (0.829) is the highest in the entire matrix, suggesting that, once base semantics are factored out, these two prefixes are nearly indistinguishable in terms of the shift they induce. This convergence is theoretically meaningful: it suggests that the two prefixes do not merely happen to attach to similar adjectival bases, but that they impose genuinely similar transformations regardless of what base they apply to.

The peripheral positions of *arci* and *extra* are likewise preserved under the offset representation. *Arci* remains the most distinctive prefix, showing consistently low similarities with the other prefixes, particularly *extra* (0.277) and *iper* (0.308), consistent with a more specialized semantic profile. On the other hand, *extra* maintains moderate similarity with the core cluster while remaining least similar to *arci*, suggesting a partially independent but not wholly peripheral position. In addition, *stra* continues to occupy a genuinely intermediate position, with similarities that place it closer to the core cluster than to the periphery without fully belonging to either.[7]

Overall, the replication of the cluster structure in offset space suggests that the close grouping of *iper*, *super*, and *ultra* cannot be attributed to shared preferences for particular bases. Rather, these prefixes exhibit substantial similarity in the semantic transformations they induce, even after the contribution of the base adjective has been factored out.

4.2 Recoverability of prefix identity in offset space

The preceding analyses established a consistent structural pattern. At the level of average vectors, the six prefixes occupy only partially distinct regions of semantic space: derivative centroids show substantial overlap, whereas offset centroids reveal a more compact cluster of *iper*, *super*, and *ultra*, with the remaining prefixes occupying more peripheral positions. For rivalry, however, centroid discrimination is only an indirect diagnostic: average offset vectors necessarily smooth over variation among individual offsets, so well-separated averages do not guarantee that the corresponding instances are themselves distinguishable. This final empirical section therefore shifts the focus from aggregate geometry to instance-level

[7] The tight clustering of *super*, *ultra*, and *iper* is statistically robust: all pairwise similarities among these three prefixes, as well as between each of them and *stra*, remain significant after Holm correction ($p_{\text{Holm}} \leq .002$, permutation test, $B = 10{,}000$). By contrast, the *arci–stra* similarity shows no directional tendency whatsoever ($p_{\text{Holm}} = .725$), suggesting that the observed similarity is not distinguishable from the null distribution at the centroid level.

recoverability, i.e., the extent to which the offset of a single derivative allows us to infer which prefix produced it. Because *arci* and *extra* are represented by only 8 and 14 derivatives, respectively, their sample sizes are too small for the cross-validated classification analyses to yield reliable estimates. We therefore restrict the analyses in this section to the four better-attested prefixes: *iper*, *stra*, *super*, and *ultra*.

4.2.1 Multiclass recoverability and information-theoretic discrimination

In their methodologically oriented distributional study of rivalry in French, Guzmán Naranjo & Bonami (2023) treat classification as a binary test of "success" or "failure" under the logic of rivalry: if processes are genuine rivals, the semantic shifts they induce should be largely indistinguishable. In distributional terms, this predicts that vector offsets – representing the shift in semantic space – should not carry process-specific signatures. If, however, a classifier can assign offsets to their originating prefix above chance, this indicates that the semantic shifts retain distinct, recoverable structure, thereby calling into question the assumption of rivalry among the processes considered. The authors acknowledge that this framing does not fully capture what is arguably a scalar phenomenon, as their confusion matrices reveal a gradient of discriminability that resists reduction to a binary rival/non-rival verdict, and they endorse a gradient view of rivalry as better supported by the data. Their classifier for French deadjectival nouns in *-té*, *-ité*, *-itude*, and *-erie* illustrates this tension: overall accuracy (86.6%) barely exceeds the majority-class baseline (83.1%), which under a strict categorical criterion would classify all four as rivals – yet the underlying confusion matrix shows *-té* almost never correctly distinguished from *-ité*, while *-ité* and *-erie* are correctly told apart roughly half the time, with *-itude* falling in between. A single binary verdict thus conceals at least three distinguishable degrees of discriminability within one nominally "rival" set.

The present section builds on these insights, asking not only whether prefix identity is recoverable from individual offsets, but also to *what degree* – and whether the gradient structure of recoverability maps onto the similarity relations established by the centroid analysis. We report two related but distinct analyses. The first, presented in this section, is a single multiclass classifier trained to predict prefix identity from vector offsets. Its predictions are evaluated in two ways: using standard classification metrics and by reinterpreting its calibrated output probabilities through an information-theoretic lens. The information-theoretic measures are not a separate model, nor an additional input to the classifier; they are a different summary of the same classifier's predictions. A second, related analysis, presented in §4.2.2, trains new, separate binary classifiers for each pair of prefixes, since isolating a single pair's rivalry requires removing the other prefixes from the classification problem. Together, the analyses address three increasingly specific questions: (i) whether prefix identity is recoverable from individual offsets at all, addressed by the global measures; (ii) if so, which prefixes contribute most to this recoverable signal, addressed by a per-prefix decomposition; and (iii) which pairs of prefixes are hardest to tell apart, addressed by the pairwise analysis.

We report both standard classification metrics and information-theoretic measures because they capture different aspects of recoverability. Our operationalization of rivalry is

itself a statement about *uncertainty reduction*, which accuracy-based metrics were not designed to quantify. Namely, accuracy and F1-score depend only on the classifier's discrete class assignments; they therefore distinguish neither the confidence of a prediction nor the degree of uncertainty among alternative classes. The information-theoretic measures, by contrast, take the classifier's full predicted probability distribution into account and are therefore sensitive to these differences. As we show below, this distinction is empirically consequential: two of our four examined prefixes show opposite rankings under the two types of measures, indicating that the information-theoretic analysis captures aspects of the recoverable structure that standard classification metrics alone do not.

Before introducing the formal measures, it is useful to clarify the underlying intuition. The information-theoretic measures we use build on two core notions, entropy and mutual information, the latter being defined in terms of the former, so we begin with entropy. Entropy quantifies uncertainty about a categorical outcome – here, which prefix produced a given derivative – before any additional evidence is considered. If the four prefixes were equally frequent, entropy would be maximal ($\log_2 4 = 2$ bits); if one prefix accounted for nearly all derivatives, entropy would approach zero, since the outcome would be almost perfectly predictable. Consider, for example, the derivative *straconvinto* 'very convinced': knowing only that one of the four prefixes was attached to *convinto* leaves substantial uncertainty about which prefix was selected. Mutual information, in turn, quantifies how much this uncertainty is reduced by additional evidence – here, how much observing the vector offset of a derivative reduces uncertainty about which prefix produced it. If offsets contain no information about prefix identity, mutual information is zero; if they uniquely determine prefix identity, mutual information equals the entropy of the prefix distribution, as all uncertainty is resolved. Accordingly, rivalry can be formulated as an information-theoretic question: *how much information about prefix identity remains in distributional semantic space once base semantics has been factored out?* This formulation naturally accommodates gradient rivalry: prefix identity may be recoverable to some degree while substantial uncertainty – and hence semantic overlap between rival processes – remains.

To estimate this recoverable information, we train an `XGBoost` (eXtreme Gradient Boosting) classifier (Chen & Guestrin 2016). As a non-linear ensemble of boosted decision trees, `XGBoost` can capture complex interactions and local structure in high-dimensional data without assuming linear separability. The classifier is trained directly on the original 500-dimensional vector offsets, allowing it to exploit information distributed across the full semantic space.

Because a single base can combine with several prefixes, a classifier could in principle exploit base-specific regularities rather than prefix-specific ones, so that its performance would not reflect the prefixes alone. To rule this out, the classifier is always tested on derivatives whose bases it has not encountered during training: all derivatives sharing a base are assigned to the same cross-validation fold. Within this constraint, evaluation follows a nested ten-fold cross-validation design, with five-fold cross-validation used within each outer training set for hyperparameter tuning. Each derivative's prediction thus comes from a classifier trained without that derivative or any other derivative sharing its base, and these out-of-fold

predictions form the basis of all reported measures. Finally, to ensure that class imbalance does not skew performance, inverse-frequency case weights are applied during training, balancing the influence of each prefix without altering the underlying dataset.

With the classifier's out-of-fold predictions in place, we can now define the information-theoretic measures formally, all of them in bits. Mutual information is estimated in three steps: we quantify the uncertainty about the prefix before the offset is observed, then the uncertainty that remains after it is observed, and finally the difference between the two. The first is given by the marginal entropy of the prefix distribution,

$$H(\text{Prefix}) = -\sum_{y} p(y) \log_2 p(y),$$

where $p(y)$ is the relative frequency of prefix y in the dataset. The second – the uncertainty remaining once the offset is known – cannot be computed directly, since the true probability of each prefix given an offset is unknown; we therefore estimate it from the classifier's predictions. For each derivative i, with offset vector $\mathbf{x}_i$ and actual prefix y_i, let $\hat{p}(y_i \mid \mathbf{x}_i)$ be the calibrated probability the classifier assigns to the correct prefix. The average negative log-probability assigned to the correct prefix, known as the cross-entropy of the predictions,

$$\widehat{CE}(\text{Prefix} \mid \text{Offsets}) = -\frac{1}{N} \sum_{i=1}^{N} \log_2 \hat{p}(y_i \mid \mathbf{x}_i),$$

is low if the classifier confidently assigns high probability to the correct prefix, and approaches the marginal entropy if its predictions are no better than the prefix frequencies alone. The difference between the two,

$$\hat{I}(\text{Offsets}; \text{Prefix}) = H(\text{Prefix}) - \widehat{CE}(\text{Prefix} \mid \text{Offsets}),$$

therefore measures how much the vector offsets reduce uncertainty about prefix identity. Because a classifier can only extract information that is present in the offsets, this is a conservative estimate: it captures the information that is *recoverable* from the offsets by the classifier, which may be less than the total information they contain.[8] We refer to this quantity as recoverable mutual information, or simply mutual information.

Figure 3 presents the classifier's performance for the four prefixes. The most salient pattern is a pronounced asymmetric pull exerted by *super*: it is the single most frequent prediction for both *stra* derivatives (69.7% of cases) and *ultra* derivatives (46.9%), in the latter case exceeding *ultra*'s own correct-classification rate (40.7%). For *iper*, predictions are split nearly evenly across *iper* itself (34.1%), *super* (30.7%), and *ultra* (35.2%). Thus, rather than behaving simply as one member of a mutually confusable cluster, *super* acts as a distributional attractor, drawing in the majority of *stra* derivatives, a plurality of *ultra* derivatives, and nearly a third of *iper* derivatives.

[8] Formally, the expected cross-entropy of any predictive distribution is at least as large as the true conditional entropy, so the difference estimates a lower bound on the mutual information between offsets and prefix identity; this formulation is standard in information-theoretic probing of linguistic representations (Pimentel et al. 2020, a.o.). The probabilities were calibrated using one-vs-rest Platt scaling, with the resulting values renormalized to sum to one.

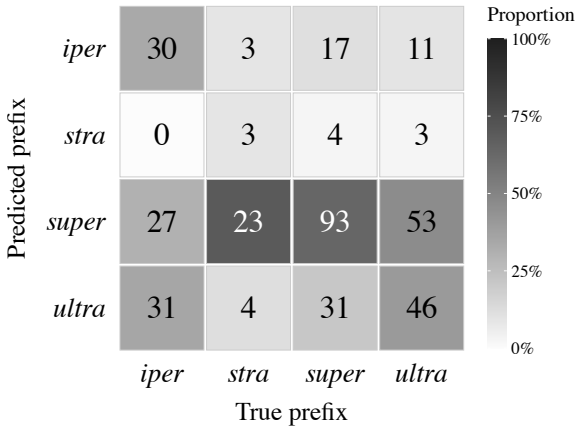

Figure 3: `XGBoost` confusion matrix for vector offsets. Columns represent the true prefix and rows the predicted prefix; cells on the diagonal (top left to bottom right) correspond to correct classifications. Numbers indicate counts. Cell shading indicates the proportion of instances of each true prefix assigned to each predicted prefix: the darker the cell, the more often that prefix is predicted for instances of the column class.

To contextualize the classifier's predictive performance, we establish two baseline accuracies. A naive strategy that always predicts the most frequent class – the no information rate (NIR) – yields an accuracy of 0.383. A chance-level baseline based on random guessing weighted by class frequencies yields 0.297.[9]

The classifier achieves an overall accuracy of 0.454 (95% CI [0.404, 0.504]), significantly exceeding the NIR baseline ($p = .003$) and indicating that offsets contain statistically reliable information about prefix identity. At the same time, Cohen's κ of 0.190 indicates only modest agreement beyond chance after accounting for class imbalance. The classifier therefore captures systematic prefix-related structure, but a substantial portion of prefix identity remains unresolved.

The information-theoretic analysis makes the extent of this recoverability more explicit. Relative to the marginal entropy of the prefix distribution (1.85 bits), the classifier's predictions reduce uncertainty by 0.18 bits (95% CI [0.11, 0.24]), our estimate of the recoverable mutual information between offsets and prefix. In relative terms, the classifier thus recovers approximately 10% of the uncertainty about prefix identity, leaving the remaining 90% unresolved. This captures the central pattern of the analysis: above-chance classification and rivalry are not mutually exclusive. The prefixes are neither perfectly indistinguishable nor cleanly discriminable: they occupy substantially overlapping regions of semantic transformation space while retaining weak but detectable prefix-specific signatures.

[9] Rather than always predicting the same class, this baseline assigns each prediction at random, with the probability of guessing a given prefix equal to its frequency in the data. Its expected accuracy is the sum of each class's squared proportion. This is always less than or equal to the NIR baseline, which concentrates all predictions on the single most frequent class rather than spreading them proportionally.

Per-prefix metrics in Table 8 reveal that classification performance is uneven across the four well-attested prefixes. *Super* achieves the highest F1-score (0.545), driven primarily by relatively high recall (0.641). *Ultra* (F1 = 0.409) and *iper* (F1 = 0.403) show intermediate and very similar performance, whereas *stra* is much less successfully classified (F1 = 0.140). This ranking, however, reflects only the classifier's final class assignments. As the next analysis shows, a different picture emerges when we take into account the classifier's full predicted probability distribution rather than only its single best prediction.

Table 8: Classification metrics for the `XGBoost` classifier

	Precision	Recall	F1-score
iper	0.492	0.341	0.403
stra	0.300	0.091	0.140
super	0.474	0.641	0.545
ultra	0.411	0.407	0.409

The global mutual information reported above shows how much information about prefix identity is recoverable overall, but not how this information is distributed across prefixes. We therefore decompose the analysis by prefix using a one-vs-rest formulation. For each prefix y, we define a binary random variable Z_y indicating whether a derivative is formed with y, and compute, in the same way as above,

$$\hat{I}(\text{Offsets}; Z_y) = H(Z_y) - \widehat{CE}(Z_y \mid \text{Offsets}),$$

which measures how much observing an offset reduces uncertainty about whether a derivative is formed with prefix y. Because raw values in bits are not directly comparable across prefixes of different frequency, we take the normalized mutual information,

$$\frac{\hat{I}(\text{Offsets}; Z_y)}{H(Z_y)},$$

as our primary measure for comparing recoverability across prefixes; it expresses the proportion of uncertainty about membership in prefix y that is resolved by the offsets.[10]

The results reveal an uneven and, in some respects, counterintuitive pattern. By the normalized measure, *stra* ranks highest (0.130, 95% CI [0.049, 0.192]), followed by *iper* (0.110, 95% CI [0.054, 0.162]), with *super* (0.055, 95% CI [0.006, 0.100]) and *ultra* (0.054, 95% CI [0.008, 0.097]) at the lower end.[11] The confidence intervals overlap considerably, so this ranking should be read as a tendency rather than a firm ordering. Strikingly, however, it

[10] These per-prefix quantities are not additive and do not sum to the global $\hat{I}(\text{Offsets}; \text{Prefix})$; they are reported for comparative purposes only.

[11] In raw bits, the order of the top two reverses (*iper*: 0.086 bits; *stra*: 0.055 bits), reflecting *iper*'s larger class size; since raw values are not comparable across prefixes of different frequency, we rank by the normalized measure throughout.

nearly reverses the F1 ranking: *super*, the best-classified prefix, is among the least recoverable, whereas *stra*, the worst-classified, ranks highest. Because *super* is also the most frequent prefix in the sample, we checked whether this reversal merely reflects residual class imbalance surviving the inverse-frequency training weights, recomputing the measure on 2,000 class-balanced resamples of the out-of-fold predictions (33 derivatives per prefix, matching *stra*'s sample size). The pattern holds: *stra* retains the highest mean (0.136, 95% CI [0.089, 0.191]), followed by *iper* (0.094, 95% CI [0.038, 0.164]), while *super* drops to the bottom of the ranking (0.033, 95% CI [0.006, 0.074]), just below *ultra* (0.049, 95% CI [0.011, 0.104]). The intervals for *super* and *ultra* overlap substantially, but the interval for *stra* does not overlap with that of *super*, indicating that the contrast between the two ends of the ranking is robust to class imbalance.

Why is the best-classified prefix among the least recoverable? The answer lies in the classifier's confidence. Of the 234 held-out derivatives not formed with *super*, 103 (44.0%) are nonetheless classified as *super*. In other words, more non-*super* derivatives are misclassified as *super* than are correctly assigned to any other single class. The *stra* derivative *strasimpatico* 'very likeable', for instance, is assigned only a 2.9% probability of belonging to *stra*, compared with 80.7% for *super*. Because F1-score reduces the classifier's probabilities to a single class decision, it registers these misclassifications as false positives but not how confidently they are made: a derivative assigned to *super* with a probability of 0.51 counts the same as one assigned with 0.99. Mutual information, computed from the probabilities themselves, penalizes these confident errors heavily – and this is precisely the distinction that a reading of the confusion matrix alone, in the spirit of Guzmán Naranjo & Bonami (2023), cannot draw.

Stra shows the mirror image. Its recall is very low – true *stra* derivatives are seldom assigned the *stra* label – but non-*stra* derivatives are almost never mistaken for *stra* either (specificity: 0.980). Because one-vs-rest mutual information is computed from the calibrated *stra*-probability assigned to *every* derivative, not only those where *stra* is the correct label, this near-perfect handling of the many non-*stra* cases contributes substantially to the measure. Where *super* is penalized by mutual information for being confidently wrong on negatives, *stra* is rewarded for being reliably right on them – and it is this asymmetry, invisible to F1-score in both directions, that explains why *stra* performs poorly by that measure yet emerges as the most informative prefix under normalized mutual information.

These recoverability estimates partially mirror, but do not simply reproduce, the centroid similarity structure reported in §4.1. At the centroid level, *iper*, *super*, and *ultra* form a tightly clustered core, with pairwise similarities ranging from 0.758 to 0.829, substantially higher than those involving the remaining prefixes. This centroid-level proximity, however, does not translate uniformly into instance-level indistinguishability: *iper*, despite belonging to this cluster, retains the second-highest mutual information, whereas *super* – the most central and most frequent member of the cluster – is among the least recoverable. Centroid proximity and instance-level recoverability therefore capture distinct properties: two prefixes may have similar average transformations while differing in how consistently those transformations are realized across individual base-derivative pairs.

4.2.2 Pairwise rivalry as normalized information overlap

The per-prefix analysis above shows how recoverable each prefix is against all others at once, but not *which* competitors a given prefix is confused with. If *iper* is only moderately recoverable, is that because it resembles *super*, *ultra*, *stra*, or all three equally? Rivalry, however, should also be examined as an inherently dyadic relation between specific pairs of processes, and a single four-way analysis blurs these pairwise relationships together. We therefore turn to distinguishability one pair at a time, estimating it separately for each of the $\binom{4}{2} = 6$ pairs among the four well-attested prefixes.

For each pair (i, j), we restrict the dataset to derivatives formed with prefixes i and j and train a new binary classifier – one that need only tell the two prefixes in that pair apart – using the same nested cross-validation design as the multiclass classifier above. From its calibrated out-of-fold probabilities, we estimate the mutual information between the offset and prefix identity for that pair and normalize it by the pair's marginal entropy. This normalized mutual information provides a measure of how informative the offsets are about prefix identity, ranging from 0 (offsets carry no information distinguishing the two prefixes) to 1 (offsets fully determine which of the two produced a derivative). To express this in terms of rivalry rather than distinguishability, we define the pairwise rivalry index R_{ij} as its complement, that is, one minus the normalized mutual information. Thus, $R_{ij} = 1$ indicates complete indistinguishability (100% remaining uncertainty, maximal rivalry), whereas $R_{ij} = 0$ indicates fully discriminable semantic transformations (0% remaining uncertainty, no rivalry). Concretely, the lowest value we observe – *iper–stra* at $R = 0.808$ – means that, even for the most discriminable pair in the dataset, 81% of the uncertainty about which of the two prefixes produced a given derivative remains unresolved by the classifier after observing its offset. Table 9 presents the resulting pairwise rivalry matrix.

Table 9: Pairwise rivalry strength (R_{ij}) for the four well-attested prefixes. Higher values indicate greater rivalry, i.e., greater indistinguishability of the two prefixes' semantic transformations.

	iper	*stra*	*super*	*ultra*
iper	1.000			
stra	0.808	1.000		
super	0.904	0.940	1.000	
ultra	0.926	0.868	0.947	1.000

Rivalry is uniformly high across all six pairs, ranging from 0.808 to 0.947: even the most discriminable pairs remain largely overlapping, confirming at the pairwise level the broad semantic convergence found throughout. Within this compressed range, however, consistent structure emerges. The two least rival (most discriminable) pairs both involve *stra*: *iper–stra* ($R = 0.808$) and *stra–ultra* (0.868). The most rival (least discriminable) pair is *super–ultra* (0.947), the two prefixes whose average transformations were already the most similar at the centroid level.

That last observation raises a natural question: does pairwise rivalry, measured derivative by derivative, track the centroid-level similarity of the prefixes' average transformations (see Table 7)? In other words, are prefixes whose typical semantic shift points in a similar direction also the ones hardest to tell apart case by case? We tested this by correlating the six R_{ij} values against the corresponding offset-centroid cosine similarities. The two are positively associated (Spearman's $\rho = 0.83$; exact Mantel test over all 24 permutations of prefix labels, $p = .083$) – prefixes with more similar average shifts do tend to be harder to distinguish individually – although the association does not reach significance. With only four prefixes, the test has very limited power, so this result should be interpreted as suggestive: it is consistent with a link between average semantic direction and instance-level distinguishability, but does not provide strong evidence for such a relationship.

Accordingly, rather than merging the centroid-based and information-theoretic measures into a single composite rivalry score, we keep them as two complementary views of gradient rivalry. What both views converge on is that the intensifying prefixes examined here are neither fully equivalent nor cleanly discriminable: they occupy substantially overlapping regions of semantic transformation space, and rivalry among them is a graded, pair-specific property rather than an all-or-nothing verdict. *Super*, *ultra*, and *iper* form the geometric core of the system, with the most tightly aligned average transformations (§4.1); yet instance-level recoverability does not follow this clustering straightforwardly, since *super* – the most central and most frequent member of that core – is among the *least* recoverable prefixes once class frequency is controlled for (§4.2.1). *Stra*, conversely, is the worst-classified prefix by standard metrics yet ranks highest by mutual information, and forms the most discriminable pairs of any prefix examined. Taken together, these results caution against reading instance-level distinguishability directly off either centroid geometry or classification accuracy alone: each reflects a genuine but partial facet of a more graded underlying structure.

5 Discussion

This study set out to investigate the nature of rivalry in evaluative morphology, and to advance a broader argument about morphological competition itself: that rivalry is inherently gradient rather than binary, with competing processes being more or less rival according to the degree to which their semantic transformations overlap. Evaluative morphology, despite growing interest in affix rivalry, has remained largely peripheral to this line of research, and offers a demanding test case precisely because its members are so semantically convergent. We pursued this argument through a distributional semantic study of six Italian intensifying prefixes – *arci-*, *extra-*, *iper-*, *stra-*, *super-*, and *ultra-* – asking whether and to what degree their outputs and transformations are distinguishable in semantic space. Three interconnected empirical questions guided the investigation: (i) whether the prefixes occupy distinct or overlapping regions of distributional space; (ii) whether they induce systematically different semantic shifts from base to derivative; and (iii) how much of prefix identity is recoverable from the transformation alone. The following sections summarize the answers

the analyses provide to these questions (§5.1), discuss their theoretical and methodological implications (§5.2), and address the limitations of the study (§5.3).

5.1 Summary of findings

Regarding the first question – whether the prefixes occupy distinct or overlapping regions of distributional space – the examination of derivative vectors in §3 established that adjectives formed with the six prefixes are broadly co-distributed in semantic space, reflecting the shared intensifying function that motivates treating these prefixes as rivals in the first place. At the same time, meaningful variation was detected beneath this surface equivalence. Derivatives formed with *extra-* cluster most tightly and show the highest, though not significantly higher, base retention, reflecting a semantically cohesive but domain-specific class centered on food-related and commercial properties. *Arci-* derivatives, by contrast, are the most dispersed, with low internal density and the greatest semantic distance from all other prefixes at the centroid level, pointing to a more specialized affective and epistemic profile. The remaining four prefixes – *iper-*, *stra-*, *super-*, and *ultra-* – form a broad intermediate band, with *iper-*, *super-*, and *ultra-* emerging as a particularly coherent core cluster on the basis of high inter-centroid cosine similarities, together with strong neighborhood overlap between *ultra* and the other two.

Regarding the second question – whether the prefixes induce systematically different semantic shifts – the vector-offset analysis in §4, which factors out base semantics and isolates the semantic contribution of each prefix, both replicated and sharpened this picture. The tight clustering of *iper-*, *super-*, and *ultra-* is preserved and even more pronounced in offset space – most strikingly, the *super–ultra* offset similarity (0.829) is the highest in the entire matrix – confirming that these three prefixes not only attach to similar bases but impose highly similar semantic transformations regardless of the adjective they apply to. Qualitatively, the most plausible characterization is that *iper-*, *super-*, and *ultra-* converge on a comparatively "pure", register-neutral scalar intensification: they raise the degree of the base property without importing the affective and colloquial connotations that mark *arci-* and *stra-*, or the domain-specific commercial and culinary anchoring of *extra-*. This finding suggests that the semantic convergence of these prefixes is not an artifact of shared selectional preferences but reflects a structural property of the processes themselves.

Regarding the third question – how much of prefix identity is recoverable from the transformation alone – the information-theoretic analysis, restricted to the four best-attested prefixes (*iper-*, *stra-*, *super-*, *ultra-*; §4.2), provided the most direct quantification of rivalry. Globally, only around 10% of the uncertainty about prefix identity is recoverable from the vector offsets, leaving approximately 90% unresolved even after the semantic transformation is observed. This result simultaneously confirms the empirical reality of rivalry and its fundamentally graded character. The classifier exceeds both baselines, establishing that the offsets carry genuine prefix-specific structure, but the modest Cohen's κ (0.190) makes clear how weak that structure is.

The most instructive finding, however, emerged from a divergence between the measures. *Super-* achieves the highest F1-score of the four prefixes (see Table 8), yet ranks among

the lowest by mutual information, with the lowest mean once class frequency is controlled for. This reversal can be traced to a concrete mechanism: the classifier confidently, and wrongly, defaults to *super-* for a large share of the other prefixes' derivatives, inflating its classification performance without reflecting any genuinely distinctive semantic signature. *Stra-* shows the opposite profile, poorly classified by standard metrics yet ranking highest by mutual information, because its offsets provide a dependable basis for ruling *stra-* out even when they rarely single it out. The contrast between the two prefixes survives class-balanced resampling. The lesson generalizes beyond these two prefixes: the prefix a classifier labels most successfully is not necessarily the one whose semantic contribution is most recoverable.

Finally, the pairwise rivalry index R_{ij}, computed for the six pairs among these four prefixes, addressed the second part of the third question, revealing that competition is not uniform across pairs. Rivalry values are consistently high, confirming the overall pattern of semantic convergence, but structured variation is detectable. The most discriminable pair is *iper–stra* ($R = 0.808$, the lowest rivalry value in the matrix), consistent with their divergent qualitative profiles: technical and functional vs. colloquial and affective. At the other extreme, *super–ultra* ($R = 0.947$) and *iper–ultra* ($R = 0.926$) are the least discriminable, reinforcing the picture of near-equivalent transformations at the core of the system. A positive association was observed between centroid-level similarity and pairwise rivalry (Spearman's $\rho = 0.83$; exact Mantel test, $p = .083$), suggesting that prefixes whose average transformations point in similar directions also tend to be harder to distinguish derivative by derivative; with only four prefixes, however, this association can be reported as suggestive but not established.

5.2 Theoretical and methodological implications

These findings speak directly to the debate between categorical and gradient conceptions of rivalry introduced in §1. While the contribution is in the first instance empirical – a quantitative characterization of one evaluative morphological system – it adjudicates between the two conceptions for the case at hand, and we offer it as support for the broader gradient view. A strict categorical view of rivalry requires semantic identity between rival processes, and the present results support no such verdict for any pair. Even *super-* and *ultra-*, the most tightly grouped prefixes, do not approach complete identity: their offset-centroid similarity is 0.829, the highest observed in the matrix, but still well below the value of 1.0 corresponding to identical directional representations. Yet the results resist equally the opposite verdict, that these prefixes are semantically distinct in any strong sense: the overwhelming majority of prefix uncertainty remains unresolved by the semantic transformation, and, in the derivative-level distributional analyses, semantic space is heavily shared across all six prefixes. What emerges is precisely the gradient picture argued for by a growing body of research (Guzmán Naranjo & Bonami 2023, Salvadori et al. 2024, a.o.): rivalry is a matter of degree rather than kind, with different pairs occupying different positions along the continuum.

What the present study contributes beyond prior work is threefold. First, it brings evaluative morphology – a domain that has remained peripheral to (quantitative) research on affix rivalry – into contact with the toolkit of distributional and information-theoretic methods, demonstrating that these methods are viable and informative even when applied to a semantically highly convergent system. Second, the information-theoretic operationalization of rivalry proposed here offers a methodological advance over the classifier paradigm of Guzmán Naranjo & Bonami (2023) that is not merely conceptual: by framing rivalry as the proportion of prefix uncertainty left unresolved by the semantic transformation, it recovers structure that classification accuracy alone actively obscures, as the *super-*/*stra-* reversal demonstrates directly. Third, the pairwise rivalry index R_{ij} provides a pair-level metric of competitive strength that complements both centroid-based similarity and global entropy measures, capturing the structurally important insight that rivalry is not only a property of a system as a whole but also of specific process pairs within it. Together, these contributions establish a replicable analytical pipeline that can be applied to rivalry systems in other derivational domains and/or languages.

The qualitative profiles of the six prefixes also shed light on what differentiates rivals within a largely overlapping space. The distinction between *arci-* and *stra-* on the one hand, and *iper-*, *super-*, and *ultra-* on the other, is visible in the semantic character of their distributional neighbors: *arci-* and *stra-* show a marked affinity for affective, colloquial, and evaluative adjectives (e.g., the shared neighbors *figo* 'cool', *fottuto* 'fucking', *simpatico* 'likeable'), whereas the latter cluster gravitates toward domain-neutral or technical intensification. *Extra-* occupies a genuinely distinct niche, anchored in commercial and culinary contexts in a way that no other prefix replicates. This pattern is consistent with the semantic contribution shared by *iper-*, *super-*, and *ultra-* identified in the offset-space analysis above: not a substantive meaning beyond degree itself, but a maximally general increase in intensity, unmarked for the affective or domain-specific associations that characterize the remaining prefixes. These preferences align with Rainer's previously cited cross-linguistic observations on intensive prefixes. We also note that the affective and colloquial character of these neighbor sets is suggestive of a difference in register between the two groups of prefixes. This interpretation, however, goes beyond what our distributional measures establish directly: our analyses characterize the semantic domains a prefix associates with, not the stylistic register of its uses, and the inference from the former to the latter would require independent, register-annotated evidence to confirm. We therefore offer it as a hypothesis rather than a finding: if register does prove to be an axis of differentiation among these rivals, then accounts of rivalry attending only to semantic content may underestimate the extent to which competing affixes are functionally distinguishable.

Finally, our findings invite a broader reflection on how competition in evaluative morphology relates to rivalry in derivation more generally. In the present data, the structure of competition does not appear to differ qualitatively from that observed in other, non-evaluative derivational domains: we find gradient rather than categorical rivalry, variable degrees of semantic overlap across pairs, and prefix-specific domain niches. These patterns broadly resemble those reported for rival processes in "prototypical" derivation, such as the

cross-linguistic patterns observed for agent nouns (cf. Naccarato 2019, Huyghe & Wauquier 2021). What may distinguish the present case, then, is a matter of *degree* rather than kind. The intensifying prefixes examined here are particularly convergent in semantic terms: all six realize essentially the same function ('very ADJ'), and even the most separable pair leaves the overwhelming majority of prefix-identity uncertainty unresolved. Whereas rival derivational processes may be differentiated partly through semantic specialization (e.g., ontological distinctions between action types, in the case of French agent nouns studied by Huyghe & Wauquier 2021), the intensifying prefixes compete within a much narrower functional domain. This may push them toward the high-overlap end of the rivalry continuum and, in turn, make non-semantic dimensions of differentiation – such as productivity, register, and sociolinguistic factors – more prominent. Whether evaluative morphology is systematically characterized by stronger semantic convergence than other derivational domains is a question that the present case study cannot settle. It does, however, suggest a concrete comparative hypothesis: evaluative and non-evaluative rivalry may share the same underlying gradient organization while differing in the extent to which semantic overlap constrains the competing processes. The framework developed here provides a way to test this hypothesis by applying the same measures across different domains of morphological rivalry.

5.3 Limitations and future directions

First, and most fundamentally, the present measures establish semantic overlap between processes, which is an essential component of rivalry but does not, strictly speaking, constitute rivalry itself. Namely, two processes are genuine rivals only insofar as speakers face a choice between them when coining a new form. The analyses reported here speak to how similar the semantic effects of two prefixes are, but not to whether, or how often, either prefix is actually selected over the other in real coining events. Nor do the present measures capture change over time. All analyses are synchronic, drawing on a single snapshot of the derivational lexicon, and thus cannot establish whether a prefix's availability for new coinages – and, consequently, its status as an active rival, rather than a historically productive but currently receding process – has changed over time. A more comprehensive account would therefore extend the present measures diachronically, tracking whether semantic convergence or divergence between prefixes changes over time, and complement them with independent measures of productivity, capturing the vitality of each process rather than only the semantic profile of its existing output.

Second, and most directly related to the present dataset, the pronounced class imbalance – *arci-* being represented by only 8 types and *extra-* by 14 – meant that these two prefixes could not be included in the instance-level classification and entropy analyses of §4.2, which require enough held-out data per class to be evaluated reliably. This leaves the question of *arci-* and *extra-*'s instance-level recoverability genuinely open, rather than merely uncertain: at the centroid level, both remain the most peripheral prefixes in the system, but whether this reflects a stable property of their semantic contribution or is itself an artifact of data sparsity cannot be determined from the present data. Larger, more balanced datasets will be needed to extend the instance-level analysis to these two prefixes with any reliability.

Third, this study focused on adjectival bases and the shared function of INTENSIFICATION, leaving open how rivalry patterns might differ across other base categories or the broader functional range of these prefixes. As Nagano et al. (2024) observe, even when rival forms coexist along the onomasiological axis, variation along the semasiological one – the range of additional meanings a prefix may encode – can suffice to ensure coexistence, allowing each form to find its niche. Applying the vector-offset and entropy-based measures to non-intensifying functions of the same prefixes would make it possible to ask whether rivalry is uniformly stronger or weaker across functions, or whether specific pairs are rivals in one domain but not another – a question that the present framework is well-positioned to answer.

Finally, and more ambitiously, the present work – like most research on morphological competition – is confined to instances of rivalry within a single process type. As Körtvélyessy (2026) argues, this fails to capture the full range of alternatives available in real-life naming acts. If rivalry ultimately stems from the choices speakers make when coining new lexemes, it is unlikely that they consciously select among a closed set of synonymous processes. Rather, speakers aim to convey a particular concept and draw on whatever lexical and morphological resources most effectively achieve that communicative goal. In the case of INTENSIFICATION, evidence for this broader competitive landscape is provided by Grandi (2017), who, in a questionnaire study on the perceived synonymity of Italian intensifiers, found that the elative suffix *-issimo*, the adverb *troppo* 'too', and constructions such as *da morire* 'to die for' are judged synonymous with *stra-* – used there as representative of intensifying prefixation – by a substantial portion of respondents. Extending the analysis to encompass such alternatives would yield a more holistic understanding of linguistic competition in the evaluative domain.

6 Conclusion

This study set out to investigate rivalry in evaluative morphology through a distributional semantic study of six Italian intensifying prefixes, and to advance a broader argument: that rivalry is gradient rather than binary, with competing processes being more or less rival according to the degree to which their semantic transformations overlap. The analyses support this view. The prefixes examined here are neither categorically equivalent nor meaningfully distinct: they are graded rivals, occupying substantially overlapping regions of semantic transformation space while retaining weak but real distributional signatures – signatures whose strength, as the divergence between classification accuracy and mutual information shows, depends as much on how recoverability is measured as on which prefix is being measured. In this respect, rivalry in evaluative morphology does not appear to differ in kind from rivalry in other derivational domains, but in degree: competing within a narrow functional domain, intensifying prefixes sit toward the high-overlap end of the rivalry continuum. More generally, the study shows that distributional-semantic methods and information-theoretic operationalizations of rivalry offer a principled framework for investigating morphological competition, one that can be applied to other languages and

derivational domains. Embracing the structured indistinguishability they reveal – rather than forcing a binary verdict – is, we argue, precisely what contemporary theories of morphological rivalry require.

Data availability

The data and analysis scripts supporting the results reported in this paper are openly available at: https://osf.io/qe3m7/.

Acknowledgments

Martin Schäfer, an anonymous reviewer, and the editors of *Word Structure* are gratefully acknowledged for their constructive and valuable feedback throughout the revision process.

Abbreviations

1 = first person; 2 = second person; 3 = third person; ACC = accusative; ADJ = adjective; DAT = dative; FUT = future; IMP = imperative; IND = indicative; PL = plural; PRS = present; SBJV = subjunctive; SG = singular. In the glosses, the intensifying prefixes are labeled by their own form in small capitals (e.g., ARCI, SUPER).

Appendix A

Tables A1–A6 list the 20 nearest neighbors of the derivative centroid for each prefix, along with the cosine similarity between each derivative and the centroid. For brevity, and because all derivatives can be approximately glossed as 'very ADJ', only the translation of the base adjective is provided.

Table A1: 20 nearest neighbors of the *arci* derivative centroid, with cosine similarity values between derivatives and centroid

arci	
arcistufo 'fed up' (.62)	*strabello* 'beautiful' (.52)
arcicontento 'pleased' (.60)	*supertrash* 'trash' (.52)
arcinoto 'known' (.60)	*supercattolico* 'Catholic' (.52)
arcisicuro 'sure' (.58)	*ultranoto* 'known' (.52)
supercoccolato 'pampered' (.57)	*superappassionato* 'passionate' (.52)
arciconvinto 'convinced' (.57)	*ultrafamoso* 'famous' (.52)
stramitico 'legendary' (.55)	*supermuscoloso* 'muscular' (.51)
superbuono 'good' (.54)	*supergasato* 'excited' (.51)
straconvinto 'convinced' (.54)	*strasimpatico* 'likeable' (.51)
arcifamoso 'famous' (.53)	*stragonfio* 'swollen' (.51)

Table A2: 20 nearest neighbors of the *extra* derivative centroid, with cosine similarity values between derivatives and centroid

extra	
extravecchio 'old' (.78)	*supermorbido* 'soft' (.68)
extrafino 'fine' (.78)	*extralucido* 'shiny' (.68)
extradry 'dry' (.78)	*extrafresco* 'fresh' (.68)
extrabrut 'brut' (.77)	*superenergetico* 'energetic' (.66)
extrafondente 'dark' (.76)	*superconcentrato* 'concentrated' (.65)
extraduro 'hard' (.75)	*supertrasparente* 'transparent' (.65)
iperespressivo 'expressive' (.73)	*stramaturo* 'mature' (.65)
ultramorbido 'soft' (.72)	*iperconcentrato* 'concentrated' (.65)
superfino 'fine' (.71)	*ultraraffinato* 'refined' (.64)
extralight 'light' (.70)	*ultrabrillante* 'shiny' (.63)

Table A3: 20 nearest neighbors of the *iper* derivative centroid, with cosine similarity values between derivatives and centroid

iper	
ipersuscettibile 'susceptible' (.74)	*ipernormale* 'normal' (.68)
iperreattivo 'reactive' (.72)	*iperorganizzato* 'organized' (.68)
stracollaudato 'tested' (.71)	*iperespresso* 'expressed' (.67)
iperindustriale 'industrial' (.71)	*iperconsumistico* 'consumerist' (.67)
iperburocratico 'bureaucratic' (.71)	*superorganico* 'organic' (.67)
superpersonale 'personal' (.71)	*iperselettivo* 'selective' (.67)
ultrasintetico 'synthetic' (.70)	*ipercontrollato* 'controlled' (.67)
iperespressivo 'expressive' (.69)	*iperdivergente* 'divergent' (.67)
iperpotente 'powerful' (.69)	*ipermaggioritario* 'majority' (.66)
iperspettacolare 'spectacular' (.69)	*ipersofisticato* 'sophisticated' (.66)

Table A4: 20 nearest neighbors of the *stra* derivative centroid, with cosine similarity values between derivatives and centroid

stra	
supercoccolato 'pampered' (.70)	*straconosciuto* 'known' (.64)
stramitico 'legendary' (.70)	*ultrafamoso* 'famous' (.64)
straconvinto 'convinced' (.66)	*strafamoso* 'famous' (.63)
supermuscoloso 'muscular' (.66)	*strafelice* 'happy' (.63)
strasimpatico 'likeable' (.66)	*superdivertente* 'entertaining' (.63)
stracontento 'pleased' (.65)	*superbuono* 'good' (.63)
strafigo 'cool' (.65)	*strafico* 'cool' (.63)
strabello 'beautiful' (.65)	*supergasato* 'excited' (.63)
stragonfio 'swollen' (.65)	*ultracarico* 'loaded' (.63)
supertrash 'trash' (.64)	*ultranoto* 'known' (.62)

Table A5: 20 nearest neighbors of the *super* derivative centroid, with cosine similarity values between derivatives and centroid

super	
ultraridotto 'reduced' (.72)	*ultraregolabile* 'adjustable' (.68)
superdivertente 'entertaining' (.72)	*superconfortevole* 'comfortable' (.68)
supermuscoloso 'muscular' (.71)	*ultracompetitivo* 'competitive' (.68)
stracollaudato 'tested' (.71)	*superbuono* 'good' (.67)
supercompleto 'complete' (.70)	*supercoccolato* 'pampered' (.67)
superaffidabile 'trustworthy' (.70)	*superduro* 'hard' (.67)
superfreddo 'cold' (.69)	*stramitico* 'legendary' (.66)
superfamoso 'famous' (.69)	*ipereconomico* 'economical' (.66)
superpotenziato 'enhanced' (.69)	*ultrarigido* 'rigid' (.66)
supergrande 'big' (.68)	*ultravivace* 'lively' (.66)

Table A6: 20 nearest neighbors of the *ultra* derivative centroid, with cosine similarity values between derivatives and centroid

ultra	
ultraridotto 'reduced' (.74)	*superaffidabile* 'trustworthy' (.69)
stracollaudato 'tested' (.72)	*ipersofisticato* 'sophisticated' (.69)
ultraregolabile 'adjustable' (.71)	*superdivertente* 'entertaining' (.68)
superduro 'hard' (.70)	*superstrutturato* 'structured' (.68)
ultraelevato 'elevated' (.70)	*ultravivace* 'lively' (.68)
supersottile 'thin' (.70)	*ultraspettacolare* 'spectacular' (.68)
ultrarigido 'rigid' (.69)	*ultraluminoso* 'luminous' (.68)
ultrapreciso 'precise' (.69)	*ultraprofessionale* 'professional' (.67)
ultrabrillante 'bright' (.69)	*superfreddo* 'cold' (.67)
iperpotente 'powerful' (.69)	*superpreciso* 'precise' (.67)

References

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

Aronoff, Mark. 2020. *-less* and *-free*. In Körtvélyessy, Lívia & Štekauer, Pavol (eds.), *Complex words: Advances in morphology*, 55–64. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108780643.004.

Baroni, Marco & Bernardini, Silvia & Ferraresi, Adriano & Zanchetta, Eros. 2009. The WaCky wide web: A collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation* 43. 209–226. https://doi.org/10.1007/s10579-009-9081-4.

Bauer, Laurie & Lieber, Rochelle & Plag, Ingo. 2013. Negatives. In *The Oxford reference guide to English morphology*, 354–384. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198747062.003.0017.

Boleda, Gemma. 2020. Distributional semantics and linguistic theory. *Annual Review of Linguistics* 6(1). 213–234. https://doi.org/10.1146/annurev-linguistics-011619-030303.

Bonami, Olivier & Guzmán Naranjo, Matías. 2023. Distributional evidence for derivational paradigms. In Kotowski, Sven & Plag, Ingo (eds.), *The semantics of derivational morphology: Theory, methods, evidence*, 219–258. Berlin: De Gruyter. https://doi.org/10.1515/9783111074917-008.

Bonami, Olivier & Paperno, Denis. 2018. Inflection vs. derivation in a distributional vector space. *Lingue e linguaggio* 17(2). 173–196. https://doi.org/10.1418/91864.

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.

Bybee, Joan. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.9.

Calpestrati, Nicolò. 2017. Intensification strategies in German and Italian written language. The case of *prefissi intensivi* or *Fremdpräfixe*. In Napoli, Maria & Ravetto, Miriam (eds.), *Exploring intensification. Synchronic, diachronic and cross-linguistic perspectives*, 305–326. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.189.16cal.

Cartier, Emmanuel & Huyghe, Richard. 2021. La concurrence affixale en diachronie: le cas des préfixes de haut degré en français. *Linx* 82. https://doi.org/10.4000/linx.8078.

Chen, Tianqi & Guestrin, Carlos. 2016. XGBoost: A Scalable Tree Boosting System. In Krishnapuram, Balaji & Shah, Mohak & Smola, Alexander J. & Aggarwal, Charu & Shen, Dou & Rastogi, Rajeev (eds.), *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. New York: Association for Computing Machinery. https://doi.org/10.1145/2939672.2939785.

Denistia, Karlina & Baayen, R. Harald. 2019. The Indonesian prefixes *PE-* and *PEN-*: A study in productivity and allomorphy. *Morphology* 29(3). 385–407. https://doi.org/10.1007/s11525-019-09340-7.

Dressler, Wolfgang U. & Merlini Barbaresi, Lavinia. 1994. *Morphopragmatics*. Berlin: De Gruyter. https://doi.org/10.1515/9783110877052.

Dressler, Wolfgang U. & Merlini Barbaresi, Lavinia & 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. https://doi.org/10.1007/978-3-030-02550-2_5.

Efthymiou, Angeliki. 2024. Diminutive formation in Modern Greek: Variation and competition. In Bagasheva, Alexandra & Nagano, Akiko & Renner, Vincent (eds.), *Competition in word-formation*, 176–205. Amsterdam: John Benjamins. https://doi.org/10.1075/la.284.06eft.

Finkelstein, Lev & Gabrilovich, Evgeniy & Matias, Yossi & Rivlin, Ehud & Solan, Zach & Wolfman, Gadi & Ruppin, Eytan. 2002. Placing search in context: The concept revisited. *ACM Transactions on Information Systems* 20(1). 116–131. https://doi.org/10.1145/503104.503110.

Firth, John R. 1957. A synopsis of linguistic theory, 1930–1955. In *Studies in linguistic analysis*, 1–32. Oxford: Basil Blackwell.

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. & Luschützky, Hans Christian (eds.), *Competition in inflection and word-formation*, 67–93. Cham: Springer. https://doi.org/10.1007/978-3-030-02550-2_3.

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. https://doi.org/10.1007/978-3-030-02550-2_1.

Grandi, Nicola. 2015. The place of evaluation within morphology. In Grandi, Nicola & Körtvélyessy, Lívia (eds.), *Edinburgh handbook of evaluative morphology*, 74–90. Edinburgh: Edinburgh University Press. https://doi.org/10.1515/9780748681754-010.

Grandi, Nicola. 2017. Intensification processes in Italian: A survey. In Napoli, Maria & Ravetto, Miriam (eds.), *Exploring intensification. Synchronic, diachronic and cross-linguistic perspectives*, 55–77. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.189.04gra.

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.

Harris, Zelig S. 1954. Distributional structure. *Word* 10. 146–162.

Hill, Felix & Reichart, Roi & Korhonen, Anna. 2015. SimLex-999: Evaluating semantic models with (genuine) similarity estimation. *Computational Linguistics* 41(4). 665–695. https://doi.org/10.1162/COLI_a_00237.

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: Theory, methods, evidence*, 143–176. 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.

Huyghe, Richard & Varvara, Rossella. 2025. Semantic granularity in derivation. *Linguistics Vanguard* 11(1). 73–85. https://doi.org/10.1515/lingvan-2024-0186.

Huyghe, Richard & Wauquier, Marine. 2020. What's in an agent? A distributional semantics approach to agent nouns in French. *Morphology* 30. 185–218. https://doi.org/10.1007/s11525-020-09366-2.

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.

Iacobini, Claudio. 2004. Prefissazione. In Grossman, Maria & Rainer, Franz (eds.), *La formazione delle parole in italiano*, 97–163. Berlin: De Gruyter.

Kamber, Julie & Huyghe, Richard. 2023. Concurrence morphologique et préfixation de haut degré: le cas de *mega-*, *giga-* et *hypra-*. *Travaux de linguistique* 86. 59–83. https://doi.org/10.3917/tl.086.0059.

Körtvélyessy, Lívia. 2026. Word-formation at the crossroads of productivity, creativity, and structural complexity. *Linguistics* 64(1). 137–163. https://doi.org/10.1515/ling-2024-0038.

Kotowski, Sven & Schäfer, Martin. 2023. Quantifying semantic relatedness across base verbs and derivatives: English *out*-prefixation. In Kotowski, Sven & Plag, Ingo (eds.), *The semantics of derivational morphology: Theory, methods, evidence*, 177–218. Berlin: De Gruyter. https://doi.org/https://doi.org/10.1515/9783111074917-007.

Lacić, Ivan. 2025. InTens – A dataset of Italian intensified derivatives. Description and application in a productivity study. In Huyghe, Richard & Prudent, Megan & Salvadori, Justine (eds.), *Proceedings of the Fifth International Workshop on Resources and Tools for Derivational Morphology*, 73–86. Fribourg: University of Fribourg.

Lapesa, Gabriella & Kawaletz, Lea & Plag, Ingo & Andreou, Marios & Kisselew, Max & Padó, Sebastian. 2018. Disambiguation of newly derived nominalizations in context: A Distributional Semantics approach. *Word Structure* 11(3). 277–312. https://doi.org/10.3366/word.2018.0131.

Lenci, Alessandro & Sahlgren, Magnus. 2023. *Distributional semantics*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9780511783692.

Leviant, Ira & Reichart, Roi. 2015. Separated by an un-common language: Towards judgment language informed vector space modeling. *arXiv preprint*. https://doi.org/10.48550/arXiv.1508.00106.

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

Marelli, Marco & Baroni, Marco. 2015. Affixation in semantic space: Modeling morpheme meanings with compositional distributional semantics. *Psychological Review* 122(3). 485–515. https://doi.org/10.1037/a0039267.

Masini, Francesca. 2019. Competition between morphological words and multiword expressions. In Rainer, Franz & Gardani, Francesco & Dressler, Wolfgang U. & Luschützky, Hans Christian (eds.), *Competition in inflection and word-formation*, 281–305. Cham: Springer. https://doi.org/10.1007/978-3-030-02550-2_11.

Micheli, M. Silvia & Verdelli, Martina. 2025. Old and new intensifying prefixes in Italian: The case of *uber-*. *Zeitschrift für Wortbildung/Journal of Word Formation* 9(2). 28–62. https://doi.org/10.21248/zwjw.2025.2.139.

Mickus, Timothee & Bonami, Olivier & Paperno, Denis. 2019. Distributional effects of gender contrasts across categories. In Jarosz, Gaja & Nelson, Max & O'Connor, Brendan & Pater, Joe (eds.), *Proceedings of 2nd Meeting of the Society for Computation in Linguistics*, 174–184. Amherst: Society for Computation in Linguistics. https://doi.org/10.7275/g11b-3s25.

Mikolov, Tomas & Chen, Kai & Corrado, Greg & Dean, Jeffrey. 2013. Efficient estimation of word representations in vector space. *arXiv preprint*. https://doi.org/10.48550/arXiv.1301.3781.

Mikolov, Tomas & Yih, Wen-tau & Zweig, Geoffrey. 2013. Linguistic regularities in continuous space word representations. In Vanderwende, Lucy & Daumé III, Hal & Kirchhoff, Katrin (eds.), *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 746–751. Atlanta: Association for Computational Linguistics.

Montermini, Fabio. 2008. *Il lato sinistro della morfologia: la prefissazione in italiano e nelle lingue del mondo*. Milano: FrancoAngeli.

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 & Bagasheva, Alexandra & Renner, Vincent. 2024. Towards a competition-based word-formation theory. In Bagasheva, Alexandra & Nagano, Akiko & Renner, Vincent (eds.), *Competition in word-formation*, 1–31. Amsterdam: John Benjamins. https://doi.org/10.1075/la.284.01nag.

Napoli, Maria. 2012. Uno stra-prefisso: l'evoluzione di *stra-* nella storia dell'italiano. *Rivista Italiana di Linguistica e Dialettologia* 14. 89–112.

Pimentel, Tiago & Valvoda, Josef & Hall Maudslay, Rowan & Zmigrod, Ran & Williams, Adina & Cotterell, Ryan. 2020. Information-theoretic probing for linguistic structure. In Jurafsky, Dan & Chai, Joyce & Schluter, Natalie & Tetreault, Joel (eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 4609–4622. Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.420.

Plag, Ingo. 1999. *Morphological productivity: Structural constraints in English derivation*. Berlin: De Gruyter. https://doi.org/10.1515/9783110802863.

Plag, Ingo & Heitmeier, Maria & Domahs, Frank. 2025. Morpho-phonology is not independent of semantics: The case of German nominal number marking. *The Mental Lexicon* 20(2–3). 226–259. https://doi.org/https://doi.org/10.1075/ml.24008.pla.

Rainer, Franz. 2015. Intensification. In Müller, Peter O. & Ohnheiser, Ingeborg & Olsen, Susan & Rainer, Franz (eds.), *Word-formation: An international handbook of the languages of Europe.* Vol. 2, 1340–1351. Berlin: De Gruyter. https://doi.org/10.1515/9783110246278-033.

Řehůřek, Radim & Sojka, Petr. 2010. Software framework for topic modelling with large corpora. In Calzolari, Nicoletta & Choukri, Khalid & Baegaard, Bente & Mariani, Joseph & Odijk, Jan & Piperidis, Stelios & Rosner, Mike & Tapias, Daniel (eds.), *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, 45–50. Valletta: European Language Resources Association.

Salvadori, Justine & Varvara, Rossella & Huyghe, Richard. 2024. Measuring affix rivalry as a gradient relationship. In Bagasheva, Alexandra & Nagano, Akiko & Renner, Vincent (eds.), *Competition in word-formation*, 104–138. Amsterdam: John Benjamins. https://doi.org/10.1075/la.284.04sal.

Schäfer, Martin. 2025. The role of meaning in the rivalry of *-ity* and *-ness*: Evidence from distributional semantics. *English Language & Linguistics* 29(4). 805–850. https://doi.org/10.1017/S1360674324000443.

Stupak, Inna V. & Baayen, R. Harald. 2022. An inquiry into the semantic transparency and productivity of German particle verbs and derivational affixation. *The Mental Lexicon* 17(3). 422–457. https://doi.org/10.1075/ml.22012.stu.

Thuilier, Juliette & Tribout, Delphine & Wauquier, Marine. 2023. Affix 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.

Wauquier, Marine & Hathout, Nabil & Fabre, Cécile. 2020. Contributions of distributional semantics to the semantic study of French morphologically derived agent nouns. In Audring, Jenny & Koutsoukos, Nikos & Manouilidou, Christina (eds.), *Proceedings of the 12th Mediterranean Morphology Meeting. "Rules, patterns, schemas and analogy"*, 111–121. Patras: University of Patras. https://doi.org/10.26220/mmm.3254.