

RESEARCH ARTICLE

The linguistic factors of semantic transparency: Evidence from verb-to-noun derivation in French

Rossella Varvara^{1,2}  and Richard Huyghe³ 

¹Università di Torino, Corso Svizzera 185, 10149 Torino, Italy

²Università di Pavia, Corso Strada Nuova 65, 27100 Pavia, Italy

³Université de Fribourg, Avenue de Beauregard 13, CH-1700 Fribourg, Switzerland

Corresponding author: Rossella Varvara; Email: rossella.varvara@unipv.it

Received: 20 July 2024; **Revised:** 22 December 2024; **Accepted:** 06 January 2025

Keywords: semantic transparency; derivation; nominalization; distributional semantics; productivity; frequency

Abstract

Semantic transparency is usually defined as the extent to which the lexical meaning of a morphologically complex word can be inferred from its structure and constituents. Recent studies have emphasized the need to distinguish two aspects of transparency: relatedness (i.e. the degree to which the meaning of lexical constituents is retained in that of a complex word) and compositionality (i.e. the degree to which the meaning of a complex word is determined by the meaning of its constituents and the way they are combined). In this paper, we investigate the influence of a variety of linguistic factors on both relatedness and compositionality. Our objective is twofold, as we seek to (i) determine more precisely the impact of lexical and morphological properties on transparency and (ii) better understand the distinction between relatedness and compositionality based on their respective determinants. The study focuses on deverbal nouns in French and estimates relatedness and compositionality based on human judgments and computational methods. The results indicate that the frequency and ambiguity of bases and derivatives, as well as the productivity and polyfunctionality of nominalizing suffixes, have different effects on relatedness and compositionality. They confirm the relevance of the distinction between the two aspects of transparency.

1. Introduction

The semantic transparency of a morphologically complex word can be defined as the extent to which its lexical meaning can be inferred from its structure and constituents. Semantic transparency is a scalar property, and both derived and compound words range from fully transparent (e.g. *enjoyment*, *mountaintop*) to fully opaque (e.g. *department*, *ladybird*), with variable loss of morphological motivation in case of opacity. Semantic transparency has long been a central topic in psycholinguistic studies focused on the representation and processing of morphologically complex words (Marslen-Wilson et al., 1994; Rastle et al., 2000; Schreuder & Baayen, 1995; a.o.). The transparency of compound words, in particular, has

attracted a lot of attention and has been investigated as a factor that possibly influences their semantic processing (Frisson et al., 2008; Ji et al., 2011; Juhasz, 2007; Libben et al., 2003; Sandra, 1990; Zwitserlood, 1994; a.o.). In recent years, research has evolved to encompass more theoretical aspects, raising or revisiting issues about the determinants of semantic transparency, based on computational methods that have provided new ways to assess transparency (Günther et al., 2020; Lombard et al., 2022; Marelli & Baroni, 2015; Stupak & Baayen, 2022; Varvara et al., 2021; a.o.).

Variations in the operationalization and measurement of semantic transparency have shown that different aspects of transparency can be explored. In particular, relatedness, as the degree to which the meaning of lexical constituents is retained in that of a complex word, can be distinguished from compositionality, as the degree to which the meaning of a complex word is determined by its morphological structure (Bell & Schäfer, 2016; Gagné et al., 2016; Günther & Marelli, 2019). The distinction between these two aspects of transparency has often been overlooked, which may explain some of the conflicting results observed in the studies on the semantic transparency of complex words. It calls for further investigation into the factors determining each aspect of transparency and into the influence of each aspect on the processing of complex words.

In this paper, we focus on how linguistic properties that are presumably correlated with transparency, such as word frequency and morphological productivity, specifically influence relatedness and compositionality. Using verb-to-noun derivation in French as a case study, we analyze a sample of 500 deverbal nouns formed with 10 suffixes, for which we collected measures of relatedness based on human ratings and measures of compositionality based on distributional data. Our objective is twofold, as we seek to (i) determine more precisely the impact of lexical and morphological factors on transparency and (ii) better understand the distinction and the relationship between relatedness and compositionality based on their respective determinants. More broadly, the study aims to provide an analytical account of semantic transparency in derivation while improving our understanding of the semantic processes involved in word formation.

The paper is structured as follows. In [Section 2](#), we present the background of the study, focusing on the characterization and measurement of the different aspects of transparency. In [Section 3](#), we describe the methods used to collect data on verb-noun pairs in French. In [Section 4](#), we present the results of the study with regard to relatedness and compositionality and their comparison. In [Section 5](#), we discuss these results and their theoretical implications.

2. Background

In this section, we explore the distinction between relatedness and compositionality as different aspects of semantic transparency, drawing on previous studies. We present the different methods that have been used to measure the semantic transparency of complex words. We then discuss the lexical and morphological properties previously identified as potential factors influencing transparency.

2.1. The different aspects of semantic transparency

Definitions of semantic transparency in linguistic and psycholinguistic studies on complex words are not consensual and can vary significantly. Some authors focus on the role of the

constituents and restrict transparency to the semantic relatedness between a complex word and its constituents. This view is prevalent in studies on compound words. For instance, Zwitserlood (1994, 344) considers a compound to be semantically transparent if it is ‘synchronically related to the meaning of its composite words’. Similarly, Sandra (1990, 550) states that semantic transparency ‘refers to the relationship between compound and constituent meanings’ and explicitly distinguishes it from compositionality which ‘refers to the possibility of determining the whole-word meaning from the constituent meanings’. According to Libben et al. (2020, 340), the term *semantic transparency* ‘most often refers to the extent to which the constituents of compounds maintain their whole word meaning within the compound structure’, and it is therefore ‘related to, but not identical to the notion of semantic compositionality’. However, other authors equate transparency with compositionality, especially in studies on derivation. According to Marslen-Wilson et al. (1994, 5), ‘a morphologically complex word is semantically transparent if its meaning is synchronically compositional’. In a similar vein, Plag (2003, 46) considers derivatives to be transparent if their ‘meaning is predictable on the basis of the word-formation rule according to which they have been formed’.

Variation in definitions and the distinct role of relatedness and compositionality have been pointed out by Bell & Schäfer (2016), Gagné et al. (2016), Günther et al. (2020), among others. On the one hand, relatedness, also referred to as ‘retention’, ‘literality’ or ‘similarity’, depends on the preservation of meaning between base words and complex words. For example, *toothbrush* contrasts with *honeymoon* because it retains the meanings of its two constituent lexemes, while *honeymoon* does not. Similarly, *enjoyment* contrasts with *department* in that it can be semantically linked to the meaning of the verb *enjoy*, whereas *department* is not synchronically related to *depart*.

Compositionality, on the other hand, is based on the fact that the meaning of complex words is determined not only by the meaning of their lexical constituents but also by the semantic operations involved in word formation. Complex words can be seen as more or less transparent depending on both the retention of their constituent meanings and the semantic coherence with their morphological structure. This coherence can be affected by factors such as lexicalization and semantic change. Morphological processes can also be associated with multiple semantic operations that are not equally frequent, which influences the predictability of the meaning of complex words. In compounding, different relations between the constituent words can be expressed with variable likelihood. For example, *drumstick* and *breadstick* share the same compound head, but instantiate different relations between the constituents (one of use and the other of ingredient), and one of these relations may be more frequent in *X-stick* compounds than the other. Accordingly, the meaning of a compound word can be more or less predictable, and considered more or less transparent, depending on the likelihood of the relationship expressed by a compounding pattern (Schäfer, 2018).

The necessity of distinguishing between relatedness and compositionality is made evident by the fact that they are not logically equivalent. Complex words can be related to their base words without their meaning being necessarily compositional. In compounding, a complex word can retain the meaning of its constituents, but its lexical meaning may still be unpredictable based on the combination of these constituents. According to Günther & Marelli (2019), for example, ‘the meaning of *sandman* is obviously related to both *sand* and *man*, but it is not compositional since key aspects of this meaning – about being mythological, and the part about using a special kind of sand to make children fall asleep – would not be predicted from the combination of the constituents if a speaker was not familiar with the

actual meaning of the compound.’ The same holds in derivation, where the semantic pattern associated with an affix may not be verified even if a derivative retains the meaning of its base. For example, the French noun *intercalaire* ‘insert’ is clearly related to the verb *intercaler* ‘insert’, but it does not denote an agent or a beneficiary, which are arguably the two main meanings associated with the deverbal suffixation in *-aire* in French – as in *plagiaire* ‘plagiarist’ derived from *plagier* ‘plagiarise’ and *destinataire* ‘addressee’ derived from *destiner* ‘address’.

The relationship between relatedness and compositionality is asymmetric. In its standard definition,¹ compositionality entails relatedness, but the reverse is not true. Although necessary, relatedness is not a sufficient condition for compositionality since complex words may retain the meaning of their bases without being compositional.

As noted by Gagné et al. (2016), the lack of distinction between the different aspects of transparency affects the comparability of results across studies and may explain some of the conflicting results reported. A clear differentiation between the various aspects of transparency seems necessary when investigating both the linguistic determinants of transparency and the influence of transparency on the processing of complex words.

2.2. *Measuring semantic transparency*

Different methods have been used to assess the semantic transparency of complex words, and numerous measures of transparency have been proposed in the literature. In most studies, these measures are not explicitly linked to relatedness or compositionality, but only to transparency in general. However, they seem to capture different aspects of transparency. A key distinction can be made based on the source of the evaluation, as transparency measures can be obtained through human judgments or computational models.

Human judgments can be collected from expert or nonexpert judgments, and they usually address relatedness. Expert judgments are mostly based on categorical classifications. Compounds, for instance, have been classified into four categories depending on whether they do or do not retain the meaning of each of their constituents: both constituents can be transparent in the meaning of the compound (TT, *bedroom*); or only the second constituent is transparent, whereas the first one is opaque (OT, *strawberry*); or only the first constituent is transparent, whereas the second one is opaque (TO, *jailbird*); or both constituents can be opaque (OO, *hogwash*). Such a classification has been adopted as a preliminary step in the selection of linguistic materials in different studies (e.g. Gagné et al., 2016; Libben et al., 2003). However, categorical classifications are suboptimal with respect to the scalar nature of transparency, and the limited number of experts involved makes the assessment strongly dependent on individual variation. To overcome these difficulties, many studies use rating scales (e.g. a Likert scale from 1 to 7) and collect nonexpert judgments in behavioral experiments, which can involve large numbers of individuals.

In most cases, nonexpert ratings focus on the semantic relatedness between complex words and their bases, although the prompts used to collect these ratings vary as much as the terms used to define transparency. A common question has been to rate how strongly ‘related’ the meaning of a constituent word is to a compound (Juhász, 2007; Kim et al., 2019;

¹ See, for instance, the formulation of the compositionality principle by Hoeksema (2000, 851): ‘the meaning of a complex expression is determined by the meanings of its parts and the operations performed on those parts’.

Zwitserlood, 1994) or to a derived word (Marelli & Baroni, 2015). Gagné et al. (2019) also asked how much a constituent word (e.g. *flower*) ‘retains’ its meaning in a compound (e.g. *flowerbed*). Lombard et al. (2022) instructed participants to rate how ‘close’ the meanings of derived and base words are. Reddy et al. (2011), although interested in the compositionality of compounds, asked participants to rate how ‘literal’ compounds are and how ‘literally’ constituent words are used in compounds. In a few studies on compound words (e.g. Gagné et al., 2019; Marelli & Luzzatti, 2012), participants were additionally asked to evaluate the predictability of compounds based on the meaning of their constituents, therefore approaching compositionality rather than just relatedness. In a similar vein, Hay (2001) investigated the decomposability of derived words (which can be viewed as the counterpart to compositionality) and asked participants to rate how ‘complex’ a word is, while defining a complex word as one ‘which can be broken down into smaller, meaningful, units’ (Hay, 2001, 1048).

In parallel with human judgments, computational methods have been used to evaluate the semantic transparency of complex words. Distributional semantic models (DSMs) in particular have been employed to measure transparency and address relatedness or compositionality. DSMs are computational models that represent the distribution of words in a corpus by means of numerical vectors (see Lenci, 2018 for an introduction). Based on the Distributional Hypothesis (Harris, 1954), according to which the distribution of a word depends on its meaning, distributional vectors are assumed to represent the semantic properties of words. Originally, DSMs were based on cooccurrence vectors extracted from large corpora (so called ‘count-based models’, Baroni et al., 2014), but they evolved rapidly with the use of neural network architectures. ‘Predictive models’ have been developed, in which word vectors (also called ‘word embeddings’) do not correspond to cooccurrence counts but to the weight used by the models to predict a word given its contexts (or to predict the context of a given word). Widely used neural models are Word2vec (Mikolov et al., 2013), GloVe (Pennington et al., 2014) or FastText (Bojanowski et al., 2017), to mention a few. The field of word embeddings is constantly evolving, now focused on large language models such as BERT, GPT and LLaMA families, among many others. These models have been used not only in NLP pipelines but also in quantitative and empirical research in linguistics (see Boleda, 2020 for an overview). In particular, DSMs have been exploited to measure the semantic transparency of complex words.

The simplest distributional measure used in studies on transparency is the cosine of the angle formed by two word vectors, which indicates the semantic similarity between two words. Cosine similarity between complex words and their bases has been examined in studies on the transparency of compounds (Reddy et al., 2011) and derived words (Kotowski & Schäfer, 2023; Lombard et al., 2022). An alternative measure is the correlation between two word vectors (e.g. between the vectors of a base and a derivative) as proposed by Stupak & Baayen (2022). Cosine and correlation measures evaluate word similarity and can approximate relatedness, but they are not strictly speaking measures of relatedness since the semantic similarity between words may depend on additional factors. For example, two nouns derived from the same verb may be transparently related to that verb and still differ in their semantic similarity to it, if one denotes the same eventuality as the verb (e.g. an eventive noun) and the other denotes a participant in that eventuality (e.g. an agentive noun).

Cosine and correlation measures are calculated directly from word vectors extracted from DSMs. However, by applying linear-algebraic operations to vectors, it is also

possible to obtain compositional measures. These measures involve the computation of a hypothetical vector for a complex word based on the vectors of its constituents. For example, the hypothetical vector of a derived noun such as *achievement* can be obtained from the vector of the base verb *achieve* and the vector of the suffix *-ment*. Although distributional representations for affixes are not immediately available in DSMs, various strategies can be used to extract them. One possible method is to average all the vectors of the words formed with a given affix, as proposed, for example, by Lazaridou et al. (2013) and Wauquier et al. (2020). Another option is to represent the derivational pattern through a shift vector, computed as the average of the differences between the vectors of the words formed with an affix and those of the base words (Guzmán Naranjo & Bonami, 2023; Kisselew et al., 2015; Padó et al., 2016). Yet another solution, used, for example, by Marelli & Baroni (2015), is to treat affixes as linear functions that operate on stem vectors to generate output vectors representing derived words. In this approach, affixes are modeled as coefficient matrices, which are applied to word vectors through matrix multiplication.

When a vector representation of an affix is obtained, it can be added to the base vector to form a purely compositional vector of the derived word. Composition can be achieved through various algebraic operations commonly used to analyze the meaning of phrases and sentences, such as addition, multiplication or weighted addition (see Guevara, 2010; Landauer & Dumais, 1997; Mitchell & Lapata, 2010; Zanzotto et al., 2010). To assess the semantic transparency of a derived word, the predicted vector is compared to the actual vector of the word – for instance, through cosine similarity – and the difference between the hypothetical and the actual vector can be used as a measure of (non)compositionality. Compositional vectors have been used to study derived words, but also compounds, by creating distributional representations for compound heads and modifiers and then comparing the actual compound vectors to the predicted ones (Günther & Marelli, 2019).

2.3. *Linguistic factors*

The role of semantic transparency in the representation and processing of complex words has been extensively investigated in psycholinguistic research. However, semantic transparency is also related to various linguistic factors that may determine its impact on mental processes. Our aim in the present study is to investigate the influence of various linguistic properties on semantic transparency, viewed alternatively as relatedness or compositionality. In this section, we review the main factors that have been discussed in theoretical linguistics as determinants or correlates of transparency: word frequency, lexical ambiguity, date of attestation of words and morphological productivity.

2.3.1. *Word frequency*

It has been widely assumed, at least since the work of Pagliuca (1976), that frequency and semantic transparency are negatively correlated. Pagliuca observed that in a sample of English prefixed words, the most frequent ones were the least likely to have predictable meanings, and he presumed that high frequency causes prefixed words to drift semantically from their base. Bybee (1985) shared the same intuition, assuming a direct relationship between frequency and ‘lexical split’, defined as the diachronic process by which a complex word becomes opaque with respect to its base. A possible explanation for this relationship is

that frequency facilitates the storage and direct access to words, whose meanings are not necessarily processed compositionally and can therefore become opaque (see Baayen & Lieber, 1997).

Hay (2001) empirically tested this claim, considering not only the frequency of derived words but also that of their bases. She argued that relative frequency (i.e. the ratio between the frequency of a derivative and that of its base), rather than absolute frequency, is negatively correlated with semantic transparency. She first conducted a survey in which 16 students had to evaluate the complexity of derived words, considered as their decomposability. Derivatives that are more frequent than their base were rated as less decomposable or, in other words, less transparent. In the second part of the study, Hay investigated semantic transparency through dictionary definitions, considering complex words whose base is mentioned in their definition to be transparent, or opaque in the opposite case, and found a negative correlation between relative frequency and transparency. However, the study had certain methodological limitations: first, the definition of complexity in the questionnaire could be misleading and not easily interpretable as compositionality (as confirmed by the fact that some participants interpreted it in the opposite direction); second, the use of dictionary definitions to assess transparency brought a binary measure of transparency rather than a continuous one.

Noting the lack of evidence for the relationship between frequency and transparency, Johnson et al. (2023) tested it using a distributional measure of semantic transparency. They first observed an unexpected positive correlation between the frequency of derivatives and their transparency, but this result was nuanced by the fact that frequency accounted for almost none of the variance observed in the statistical analysis. When further controlling for the lexical ambiguity of derivatives, Johnson and colleagues observed an interaction effect on transparency: derivatives that are more frequent than expected (with respect to the frequency of the base) become increasingly opaque as their number of meanings increases. This result was confirmed with human judgments about semantic similarity in base-derivative pairs. The effect of frequency on similarity ratings could not be observed for derivatives with few meanings, as opposed to moderately or highly ambiguous words. It was concluded that low semantic transparency is mostly observed for derivatives that are both highly frequent and highly ambiguous.

2.3.2. Lexical ambiguity

Johnson et al. (2023) examined the role of lexical ambiguity when investigating the relationship between frequency and transparency. The ambiguity of constituent and complex words can be expected to influence transparency because of semantic selection and extension. On the one hand, complex words may be related to only some of the meanings of ambiguous bases. On the other hand, complex words may have multiple meanings, created through metaphor or metonymy, that are no longer related to the base words. Reddy et al. (2011) tried to control for ambiguity in the experiment they conducted on compound words by randomly selecting five occurrences of a target word from a corpus and asking participants to rate compositionality for the most frequent sense found among these occurrences. Padó et al. (2016) examined the degree of ambiguity of base words and found a negative effect on the predictability of derivatives, as evaluated through distributional measures. The ambiguity of bases and complex words can affect the assessment of transparency, both in human ratings

and computational estimates, since evaluating the semantic relationship between morphologically related words is easier for monosemous than for ambiguous words.

2.3.3. Age of complex words

Words with a long history are likely to have undergone semantic change or acquired new senses, which in the case of morphologically complex words can impact the semantic relationship with constituent words. Although speakers may lack knowledge of the historical evolution of words, they have to consider the effects of semantic shifts when assessing transparency in contemporary usage. Accordingly, it can be hypothesized that the age of a complex word affects its semantic transparency and that the older a derivative or a compound is (i.e. the farther in time its attestation date is), the less transparent it is. Lombard et al. (2022) tested the influence of word age on the semantic opacity of derived words in French. They found a significant effect of word age in speakers' judgments about the semantic proximity between bases and derivatives, but not in their distributional similarity. To our knowledge, the correlation between word age and transparency has not yet been tested for compound words.

2.3.4. Morphological productivity

So far, we have reviewed word-specific properties as potential determinants of semantic transparency. However, variation in transparency may also be caused by properties of morphological processes. Productivity in particular, as the capacity of a morphological process to produce new words, is often assumed to correlate positively with transparency (Schäfer, 2018, 72). Aronoff (1976, 39), based on a study by Zimmer (1966), argued that the semantic coherence of a word-formation process is linked to its productivity. Similarly, Baayen (1993, 199) and Plag (2003, 177) claimed that semantic transparency is essential for a word-formation process to be productive. The positive effect of productivity on transparency has also been highlighted by Bell & Schäfer (2016) in their research on noun-noun compounds in English. A more nuanced position has been defended by Stupak & Baayen (2022) in a study on transparency and productivity in particle verbs and affixed words in German. Stupak and Baayen used various distributional measures of semantic transparency, as well as two measures of productivity (realized productivity V and potential productivity P ; see Baayen, 2009). For both particle verbs and affixed words, they found a positive correlation between V and the similarity of semantic vectors for bases and derivatives, but no correlation with P . However, a positive correlation was observed between P and the similarity of particles to particle verbs. These results suggest that not all aspects of productivity are equally related to semantic transparency and that the relevant aspects may depend on the morphological process involved.

3. Method

In this section, we describe the lexical materials used in our study, the measures used to estimate relatedness and compositionality, and the information collected about the linguistic properties tested as potential factors of transparency. We provide details on the sample of complex words analyzed and on the methodology used to operationalize the different variables.

3.1. Linguistic materials

Our study is based on the analysis of 500 verb-noun pairs in French. We selected 10 suffixes known to form deverbal nouns (-ade, -age, -ance, -erie, -ette, -eur, -ion, -ment, -oir, -ure), and for each of them, we considered 50 verb-noun pairs. Knowing from previous studies (e.g. Marelli & Baroni, 2015) that weakly related pairs are less frequent than transparent ones, we preselected a set of French deverbal nouns identified as weakly transparent by Lombard et al. (2022) to increase the variance of transparency scores. More precisely, we selected the verb-noun pairs in Lombard et al.'s experiment that obtained an average transparency score below 4 on a scale of 7 in judgments collected from 309 French speakers. Thirty-seven verb-noun pairs in which the noun is formed with one of the 10 suffixes examined were thus selected.

The rest of the pairs were randomly extracted from the FRCOW corpus, which is a French web-crawled corpus containing 10.8 billion tokens (Schäfer, 2015; Schäfer & Bildhauer, 2012). Nouns ending with one of the 10 suffixes and for which a verb is also attested in the corpus were automatically extracted. Formal pairings between verbs and nouns were made by subtracting the affixal form to the nouns, adding verbal inflectional affixes to the resulting base (including possible allomorphs), and searching for resulting forms in the lemmatized corpus. We filtered the lists under the condition that both candidate verbs and candidate nouns have a frequency greater than or equal to 100 in FRCOW, both to increase the probability that the words will be known by speakers and to ensure that a sufficient amount of data will be available to generate reliable distributional representations. For each suffix, we randomly ordered the candidate pairs detected and selected the first 50 relevant verb-noun pairs per suffix. The selection criterion required that verbs and nouns be related at least diachronically, which was verified using two lexicographic resources: *Trésor de la Langue Française Informatisé* and *Robert Historique de la Langue Française*. Table 1 presents examples of verb-noun pairs selected for the 10 suffixes examined.

3.2. Relatedness ratings

The set of 500 verb-noun pairs was divided into 10 random samples of approximately 50 pairs each, composed for optimal distribution with respect to suffix diversity, frequency of both verbs and nouns, and origin of the selected items. Three to four opaque items from the previous study were present in each sample and completed with nouns extracted from FRCOW. In addition, derivatives with identical base verbs were distributed in different samples. Each sample formed the material of an independent survey in which participants were asked to rate the semantic relatedness between a base verb and a derived noun on a scale from 0 (for no relatedness at all) to 10 (for full relatedness).² At any time, participants could declare that they did not know one or both words presented to them and could then skip the question. In the analysis, we used average ratings per base-derivative pair as the measure of semantic relatedness.

² The exact instructions in French were 'Selon vous, sur une échelle de 0 à 10, dans quelle mesure le sens du verbe X se retrouve-t-il dans le sens du nom Y? (0 = Pas du tout / 10 = Complètement)'. English translation is 'On a scale from 0 to 10, to what extent is the meaning of verb X found in the meaning of noun Y? (0 = Not at all / 10 = Completely)'.

Table 1. Examples of verb-noun pairs selected for each suffix

Suffix	Example
-ade	<i>fusiller</i> ‘shoot’ - <i>fusillade</i> ‘gunfire’
-age	<i>partir</i> ‘leave’ - <i>partage</i> ‘sharing’
-ance	<i>ordonner</i> ‘order’ - <i>ordonnance</i> ‘prescription’
-erie	<i>railler</i> ‘mock’ - <i>raillerie</i> ‘mockery’
-ette	<i>pousser</i> ‘push’ - <i>poussette</i> ‘pushchair’
-eur	<i>traiter</i> ‘treat’ - <i>traiteur</i> ‘caterer’
-ion	<i>confesser</i> ‘confess’ - <i>confession</i> ‘confession’
-ment	<i>châtier</i> ‘punish’ - <i>châtiment</i> ‘punishment’
-oir	<i>baver</i> ‘drool’ - <i>bavoir</i> ‘bib’
-ure	<i>pointer</i> ‘point’ - <i>pointure</i> ‘shoe size’

Participants were French native speakers recruited on the Prolific crowdsourcing platform. They received financial compensation for answering the surveys. Participants who answered the questions too quickly to provide accurate evaluations (less than two seconds per question on average), as well as participants who systematically provided transparency scores greater than or equal to 9 for opaque items, were discarded and replaced by other participants. Answers from 50 participants per survey were finally collected. Participants were allowed to participate in multiple surveys if they wished. A total of 212 different persons participated in the study, aged 19 to 73 ($M = 31.7$, $SD = 10.5$), including 118 men, 92 women, and two who preferred not to declare.

3.3. Distributional measures of compositionality

In order to compute a distributional measure of semantic compositionality, we built a vector space model based on the Word2vec algorithm (Mikolov et al., 2013), using the Gensim library (Řehůřek & Sojka, 2010). The model was trained on the FRCOW corpus, considering lemmatized and POS-tagged words with a minimum frequency of 50.³

Using the same technique as Kisselew et al. (2015), Padó et al. (2016), and Guzmán Naranjo & Bonami (2023), we first generated distributional representations for the 10 selected suffixes by averaging offset vectors between bases and derivatives across a sample of 50 verb-noun pairs per suffix. These pairs were randomly selected from the list of verb-noun pairs automatically extracted from FRCOW and checked manually. As an additional requirement, we considered pairs that were not included in the experimental materials in order to avoid overfitting. However, for two suffixes (-ette and -ade), we could not find enough pairs available and therefore completed the list with items from the experimental set (15 pairs for -ette and 36 pairs for -ade).⁴ We then summed the base word

³ We used a CBOW model, with `min_count = 50`, `sample = 6e-5`, and number of epochs = 5 as the additional parameters.

⁴ To assess the impact of including these experimental items in the sample used to create affix representations, we conducted an experiment using offset vectors built from the exact set of pairs used in the study. The results were comparable, as the compositionality scores derived from the two suffix representations showed an almost perfect correlation (Pearson’s $\rho = .99$).

vectors with the affix representations in order to obtain predicted vectors for the derived words. We used a simple additive model, which has demonstrated effective performance in previous studies. Finally, the cosine similarity between predicted and observed derivative vectors was calculated and used as a measure of semantic compositionality.

We further experimented with two different variants of the distributional measure. In the first variant, we used affix representations from offset vectors based on the same verb-noun pairs as the ones used in the experimental material. The objective was to assess the impact of using different sets of base-derivative pairs on the reliability of suffix representations since additional pairs could not be systematically used for all 10 suffixes. In the second variant, we considered the cosine similarity between bases and derivatives as a possible measure of semantic transparency. We observed an almost perfect correlation between the similarity measures computed with offset vectors based on different samples of verb-noun pairs (Kendall's $\tau = .99$) and a high correlation between these measures and the base-derivative cosine similarity measures (Kendall's $\tau = .85$).

3.4. Operationalizing linguistic factors

The aim of our study is to examine the influence of various linguistic factors on semantic transparency, specifically investigating potential differences between relatedness and compositionality. The factors under examination include those previously studied in the literature on semantic transparency: word frequency, lexical ambiguity, word attestation date and morphological productivity. We also added morphological polyfunctionality, defined as the number of semantic functions associated with a morphological process (Salvadori & Huyghe, 2023). The semantic functions of a process correspond to the semantic types of derivatives it produces – for instance, agent, instrument, experiencer, theme, etc., in the case of the suffix *-er* in English (Panther & Thornburg, 2002; Plag, 2003; Ryder, 1999). Note that polyfunctionality, as a property of morphological processes, is different from lexical ambiguity, which is a property of individual words. The distinction is made evident by the fact that not all complex words formed through a particular process realize all of its semantic functions. Morphological polyfunctionality may correlate negatively with transparency if a high number of functions associated with a process increases the likelihood of instantiating only certain functions in complex words, thereby affecting full compositionality.

The frequency of both base verbs and derived nouns was extracted from the FRCOW corpus. Noun frequency ranges from 102 to 1,012,113 (with a median value of 3,749), and verb frequency ranges from 103 to 8,341,228 (with a median value of 25,838). Although the influence of absolute and relative frequency on semantic transparency has been discussed in the literature, we initially focused on absolute frequency to investigate the roles of base and derivative frequency separately – also considering that their relationship and combined influence could be explored in the statistical analysis. Nevertheless, we conducted additional analyses to specifically test the effect of relative frequency, as will be discussed in Section 5.

The ambiguity of both base verbs and derived nouns was approximated by the number of senses listed for each of them in the online dictionary *Wiktionnaire*.⁵ We assigned the default value of 1 to words absent from the dictionary. The maximum number of senses listed is

⁵ <https://fr.wiktionary.org/>

19 for the nouns included in our sample (with a median value of 2), and 51 for the verbs (with a median value of 4).

The date of attestation of the derived nouns was determined based on the Google Ngram 2012 French dataset.⁶ Following Bonami & Thuilier (2019), we considered the middle of the first sequence of 10 years with at least one attestation per year as the date of initial attestation. To complete the information for the five words that were not attested in this dataset, we checked the Google Ngram resource from 2019⁷ and consulted the *Trésor de la Langue Française informatisé*.⁸ Finally, the dates selected for initial attestation ranged from 1567 to 1985, with 1718 as the median.

Measures of productivity have been widely discussed in the literature (see Aronoff, 1976; Baayen, 1993; Bauer, 2001; Fernández-Domínguez, 2013; a.o.). The two most frequently used measures are ‘realized productivity’ V , as the number of word types attested for a given morphological process, and ‘potential productivity’ P , as the ratio between the number of hapax legomena and the total number of tokens found for a morphological process in a corpus (Baayen, 2009). In this study, we primarily focused on P as the most general measure of productivity, considering that V mostly represents past productivity and does not account for the potential use of a morphological process to create new words in synchrony. However, additional results and analyses involving V will be reported and examined in the discussion. A well-known issue with P is that it heavily depends on the size of the sample considered (as observed already by Baayen, 1992, but see also Gaeta & Ricca, 2002, 2006). To overcome this issue, we used a modified version of the potential productivity measure. By employing LNRE statistical models (*Large Number of Rare Events*, Baayen, 2001), specifically a finite Zipf-Mandelbrot model (fZM, Evert, 2004), we computed P based on an arbitrary sample size⁹ for each suffix examined in the study.

The polyfunctionality of suffixes was estimated by taking into consideration not only the number of different semantic functions associated with a suffix, but also the frequency with which these functions are realized among the derivatives formed with the suffix. Following Varvara et al. (2022), we used the Hill-Simpson diversity index (Hill, 1973; Jost, 2006; Roswell et al., 2021) to account for both the number and frequency of functions in polyfunctionality measures. This index is equivalent to the inverse of the traditional Simpson index (Simpson, 1949) and is calculated from Equation 1, where S is the total number of functions of a suffix, and p_i the number of word types realizing a function i divided by the total number of types per suffix.¹⁰

$$^2D = \frac{1}{\sum_{i=1}^S (p_i)^2} \quad (1)$$

The measure was computed based on the semantic analysis of a sample of 3,260 deverbal nouns ending with the 10 selected suffixes. The meaning of the nouns was described through

⁶ <https://storage.googleapis.com/books/ngrams/books/datasetsv3.html>

⁷ We did not consider the 2019 dataset from the beginning because it is not available for download. The 2019 data have been consulted through the Google Ngram viewer: <https://books.google.com/ngrams/>.

⁸ <http://atilf.atilf.fr/tlf.htm>

⁹ The productivity scores were computed using the zipfR package (Baroni & Evert, 2014; Evert & Baroni, 2007) in R (R Core Team, 2024).

¹⁰ Hill-Simpson indices were computed using the MeanRarity package (Roswell & Dushoff, 2023) in R (R Core Team, 2024).

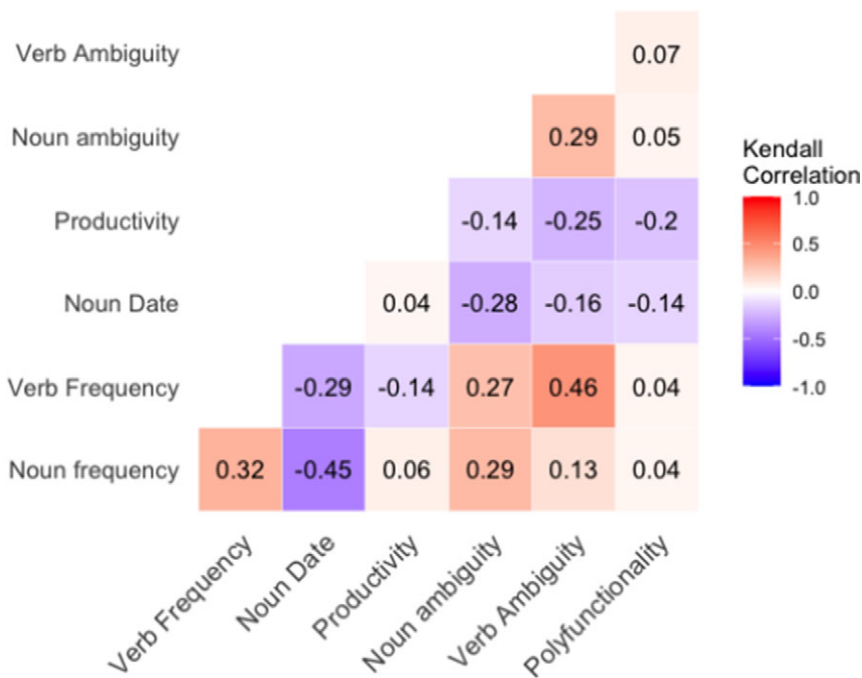


Figure 1. Significant correlations between linguistic factors

a combination of their ontological type (i.e. depending on the nature of the referents) and their relational type (i.e. depending on the semantic relation with the base verb). The semantic annotation was performed by six experts, with substantial inter-annotator agreement observed across subsamples (Cohen’s $\kappa = .69$).¹¹ A total of 5,346 different senses were identified in the whole sample, from which the number of distinct semantic functions per suffix and their realization frequency were inferred.

The different linguistic properties under study are either not significantly correlated, either significantly correlated, in which case the correlations are weak or moderate. Pairwise correlation coefficients are reported in Figure 1. The highest correlations are found between ambiguity and frequency in base verbs, and between date of attestation and frequency in derived nouns. The first correlation supports the idea that words with multiple senses are generally more frequent than words with fewer senses. The second suggests that older complex nouns (still present in contemporary corpora) tend to occur more frequently than newer ones, most likely because prolonged usage gives words more opportunities to appear and expand their meaning.

4. Results

In this section, we present and analyze data on the semantic transparency of deverbal nouns, collected through experimental and computational methods. We begin by describing the data

¹¹ For further information on the annotation scheme, we refer the reader to the annotation guidelines provided by Salvadori et al. (2024).

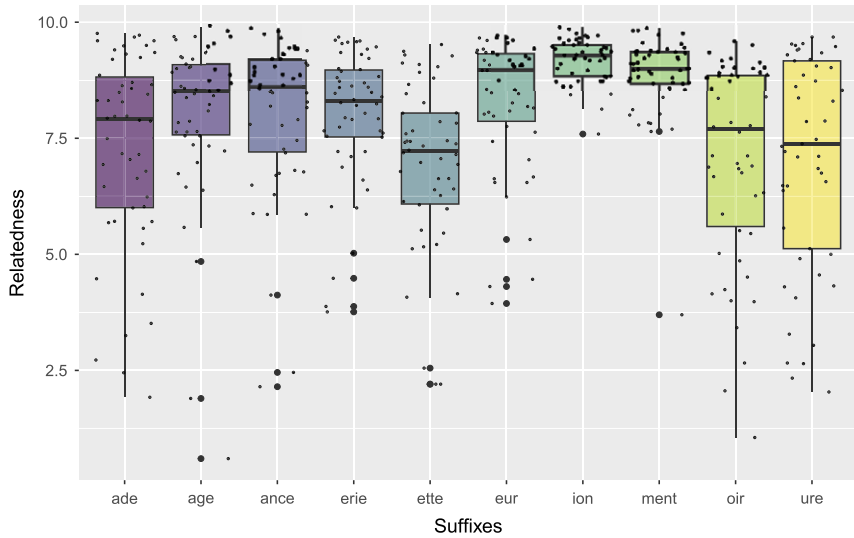
from human ratings and distributional measures and then examine the influence of linguistic factors in both cases. Finally, we compare the results obtained for relatedness and compositionality in a contrastive analysis.

4.1. General information

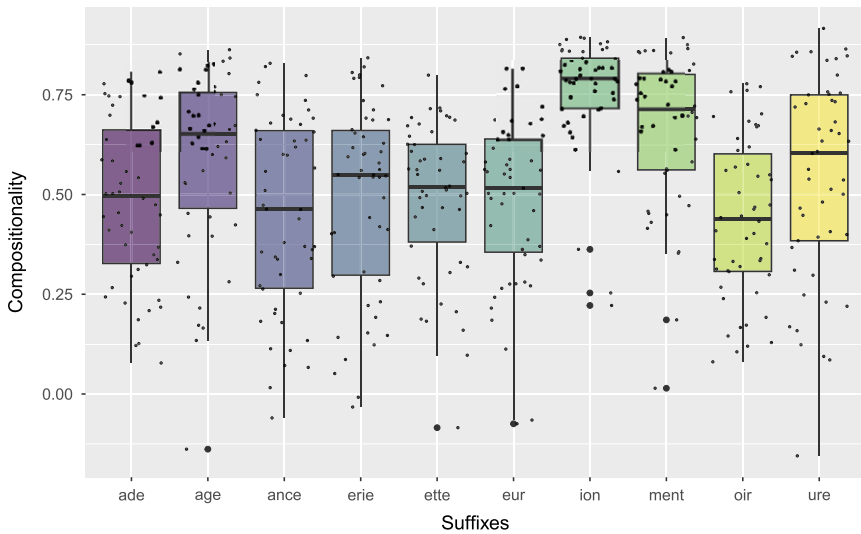
Relatedness ratings (averaged per verb-noun pair) range from 0.60 to 9.90 ($M = 7.86$, $SD = 1.85$), whereas distributional measures of compositionality range from -0.16 to 0.92 ($M = 0.54$, $SD = 0.23$). Important differences can be observed among the 10 selected suffixes, as can be seen in Table 2 and in Figure 2. As far as relatedness is concerned, the suffixes *-ion* and *-ment* show the highest average ratings, as well as the lowest variance, whereas at the other end of the scale, *-oir* and *-ure* have the lowest average ratings, as well as the highest variance. Similar observations can be made with respect to compositionality, with nouns ending in *-ion* and *-ment* showing on average the highest similarity between predicted and observed vectors and the lowest variance in similarity scores, and nouns ending in *-oir* and *-ance* showing the lowest average similarity, although not the highest variance. Overall, the variation among suffixes is consistent between relatedness and compositionality, but some noticeable discrepancies can be observed, as in the case of the suffix *-eur*, which stands 3rd in the relatedness ranking, but 8th in the compositionality ranking.

Table 2. Relatedness and compositionality measures per suffix

Suffix	Transparency measure	Mean	SD	Median	Min.	Max.
<i>-ade</i>	Relatedness	7.26	2.11	7.91	1.92	9.76
	Compositionality	0.49	0.21	0.50	0.08	0.81
<i>-age</i>	Relatedness	8.05	1.79	8.52	0.60	9.90
	Compositionality	0.59	0.23	0.65	-0.14	0.86
<i>-ance</i>	Relatedness	8.00	1.73	8.61	2.15	9.84
	Compositionality	0.45	0.24	0.46	-0.06	0.83
<i>-erie</i>	Relatedness	7.97	1.46	8.30	3.76	9.68
	Compositionality	0.49	0.24	0.55	-0.03	0.84
<i>-ette</i>	Relatedness	6.94	1.79	7.22	2.20	9.52
	Compositionality	0.49	0.19	0.52	-0.08	0.80
<i>-eur</i>	Relatedness	8.28	1.46	8.96	3.94	9.70
	Compositionality	0.48	0.21	0.52	-0.07	0.81
<i>-ion</i>	Relatedness	9.17	0.47	9.29	7.59	9.88
	Compositionality	0.76	0.15	0.79	0.22	0.89
<i>-ment</i>	Relatedness	8.83	0.94	8.99	3.70	9.85
	Compositionality	0.66	0.19	0.71	0.01	0.89
<i>-oir</i>	Relatedness	7.05	2.19	7.69	1.05	9.58
	Compositionality	0.44	0.20	0.44	0.08	0.78
<i>-ure</i>	Relatedness	7.03	2.33	7.38	2.04	9.68
	Compositionality	0.54	0.26	0.60	-0.15	0.92



(a) Distribution of relatedness ratings per suffix



(b) Distribution of compositionality scores per suffix

Figure 2. Distribution of transparency measures per suffix.

4.2. *The effect of linguistic factors on semantic relatedness and compositionality*

We analyzed the relationship between linguistic properties and semantic transparency by means of regression analysis. We performed two distinct analyses, with either relatedness or compositionality as the response variable. Depending on the nature of the variables and the distribution of the observed values, we used different types of mixed-effect regression models. A beta regression model was used to analyze relatedness given the discontinuous rating scale and the negative skewness of the observed values; that is, values were mostly

Table 3. Results of a mixed-effect beta regression with relatedness as the response variable. Significant *p*-values (at *p* < .05) are indicated in bold

	Estimate	Std. Error	z-value	<i>p</i> -value
(Intercept)	1.37	0.08	16.37	< . .001
Verb frequency	−0.04	0.05	−0.71	.475
Verb ambiguity	−0.18	0.05	−3.51	< . .001
Noun frequency	0.15	0.05	2.85	.004
Noun ambiguity	−0.15	0.05	−3.40	< . .001
Noun date of attestation	0.04	0.05	0.74	.462
Productivity	0.28	0.09	3.07	.002
Polyfunctionality	0.10	0.07	1.34	.180
V fq:V ambiguity	−0.04	0.03	−1.38	.166
N fq:N ambiguity	−0.04	0.04	−1.21	.227
Prod:Polyfunct	0.10	0.11	0.91	.361

distributed at the upper end of the scale, which prevented linear regression assumptions to be met, especially with respect to the normal distribution of residuals. Since for each item we collected ratings from different participants, and since participants rated 50 nouns each, we included in the model by-participant random variables, as well as by-item and by-suffix random intercepts. In the case of compositionality, the observed values were continuous and did not exhibit any noticeable skewness. The distribution of the residuals from a linear model was sufficiently close to normal to allow the use of a mixed-effect linear regression with a by-suffix random intercept.

All predictor variables (frequency, ambiguity, word attestation date, productivity and polyfunctionality) were centered on their mean, and frequency and ambiguity values were additionally log-transformed. Three possible interactions were tested, depending on lexical classes and morphological features. Specifically, we investigated whether lexical ambiguity interacts with frequency for both nouns and verbs, and whether productivity interacts with polyfunctionality for suffixes. Following Barr et al. (2013), random-effects structures were kept maximal as long as they were supported by the data. Multicollinearity was assessed using the variance inflation factor (VIF).¹²

The regression model used to predict relatedness¹³ included a by-participant random slope for noun ambiguity and did not show any multicollinearity issue.¹⁴ The results of the analysis are reported in Table 3. Significant effects can be observed for verb ambiguity, noun ambiguity, noun frequency and suffix productivity, as plotted in Figure 3. As expected, lexical ambiguity influences negatively relatedness: the more ambiguous the base verb or the derived noun is, the less verb-noun pairs seem to be related. Contrary to our expectations, noun frequency is positively associated with relatedness: the more frequent a derived noun

¹² All analyses were performed using the lme4 (Bates et al., 2015) and glmmTMB (Brooks et al., 2017) libraries in R (R Core Team 2024).

¹³ To avoid computations with zeros, we transformed the relatedness variable by adding to each value 0.001 and then rescaling it in order to have a maximum value of 10.

¹⁴ The VIF values for the predictor variables are the following: verb frequency = 2.42; noun frequency = 2.16; verb ambiguity = 1.99; noun ambiguity = 1.45; word date = 2.10; polyfunctionality = 1.71; productivity = 1.98.

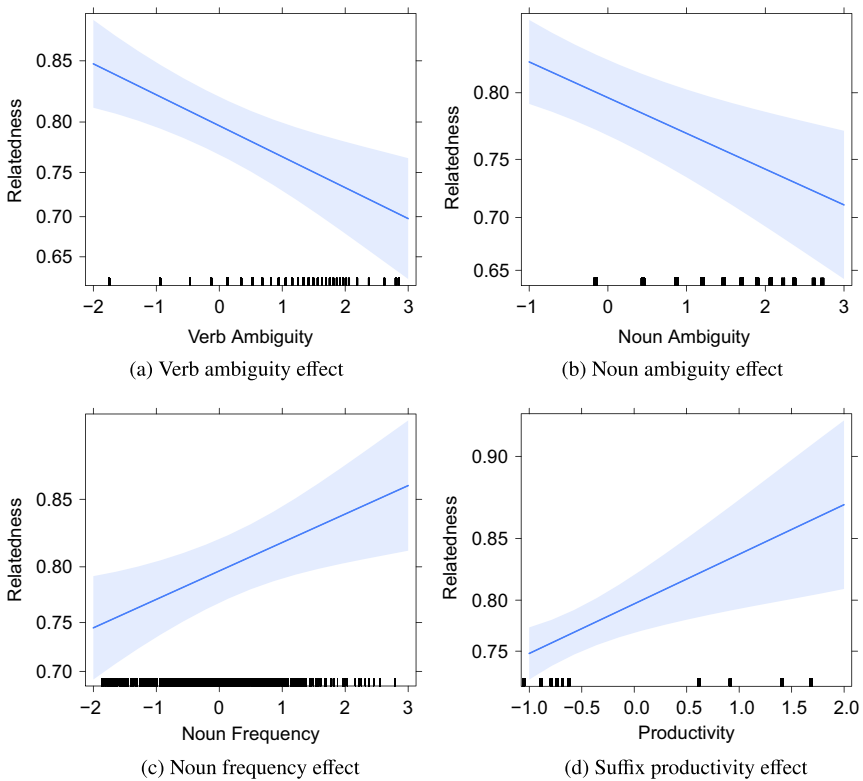


Figure 3. Effect plots for relatedness.

is, the more related it is to the base verb. Finally, a positive relationship is observed between morphological productivity and relatedness: the more productive a suffix is, the more closely the nouns it forms are related to their base.

The results of the regression model used to predict compositionality¹⁵ are presented in Table 4. Significant effects are observed for verb frequency, verb ambiguity and the interaction between productivity and polyfunctionality, as plotted in Figure 4. Verb frequency and verb ambiguity negatively affect compositionality: the more frequent or ambiguous a base verb is, the less compositional the derived nouns are. An enhancing interaction effect between productivity and polyfunctionality can be observed: more productive suffixes derive more compositional nouns as their polyfunctionality increases. As far as polyfunctionality is concerned, the trend observed does not go in the expected direction, except in the case of weakly productive suffixes.

4.3. The relationship between relatedness and compositionality

As reported in the previous section, relatedness and compositionality are influenced differently by linguistic factors. This contrast confirms that relatedness and compositionality are

¹⁵ The model did not show any multicollinearity issues. The VIF values for the predictor variables are the following: verb frequency = 2.26; noun frequency = 2.10; verb ambiguity = 1.98; noun ambiguity = 1.42; word date = 2.06; polyfunctionality = 1.69; productivity = 1.88.

Table 4. Results of a mixed-effect linear regression with compositionality as the response variable. Significant *p*-values (at $p < .05$) are indicated in bold

	Estimate	Std. Error	df	<i>t</i> -value	<i>p</i> -value
(Intercept)	0.61	0.02	6.63	33.41	< . .001
Verb frequency	−0.07	0.01	390.77	−5.27	< . .001
Verb ambiguity	−0.03	0.01	368.20	−2.05	.041
Noun frequency	0.02	0.01	437.55	1.588	.113
Noun ambiguity	−0.02	0.01	390.67	−1.76	.080
Noun date of attestation	0.01	0.01	264.98	1.03	.303
Productivity	0.13	0.02	6.45	6.17	< . .001
Polyfunctionality	0.07	0.02	5.82	4.16	.006
V fq:V ambiguity	−0.01	0.01	487.48	−1.23	.218
N fq:N ambiguity	−0.001	0.01	486.95	−0.14	.889
Prod:Polyfunct	0.11	0.02	5.41	4.48	.005

not equivalent and need to be distinguished when investigating semantic transparency. However, relatedness and compositionality are not fully independent, as can be seen in Figure 5, and a moderate positive correlation can be observed between them (Kendall's $\tau = .42$). As stated in Section 2.1, compositionality entails relatedness, but not reciprocally. Accordingly, cases of high compositionality and low relatedness tend to be rarer than the reverse, as indicated by the local polynomial regression curve (LOESS curve) in Figure 5.

If relatedness is subsumed by compositionality, we may wonder what determines the part of compositionality that is not explained by relatedness (i.e. which factors specifically influence compositionality independently of relatedness). To investigate these factors, we repeated the regression analysis for compositionality while adding relatedness ratings to the linguistic predictors previously examined.¹⁶ The model estimates are presented in Table 5. As expected, relatedness is a significant positive predictor of compositionality – the closer a derivative is to its base, the more likely it is to be compositional. When comparing regression models with and without relatedness as predictor, the main difference concerns verb ambiguity and polyfunctionality since they are no longer significant in the model that includes relatedness. In other words, verb ambiguity and polyfunctionality account for part of the variance in compositionality that is also explained by relatedness. By contrast, the effects of verb frequency and suffix productivity are significant in both models, hence they are predictors of compositionality that are not covered by relatedness.

5. Discussion

The results of our study empirically confirm that relatedness and compositionality are distinct aspects of semantic transparency. As measured by human ratings and distributional

¹⁶Note that this procedure is equivalent to analyzing the predictors of residualized compositionality (i.e. considering as the response variable the residuals of a model in which compositionality is predicted by relatedness).

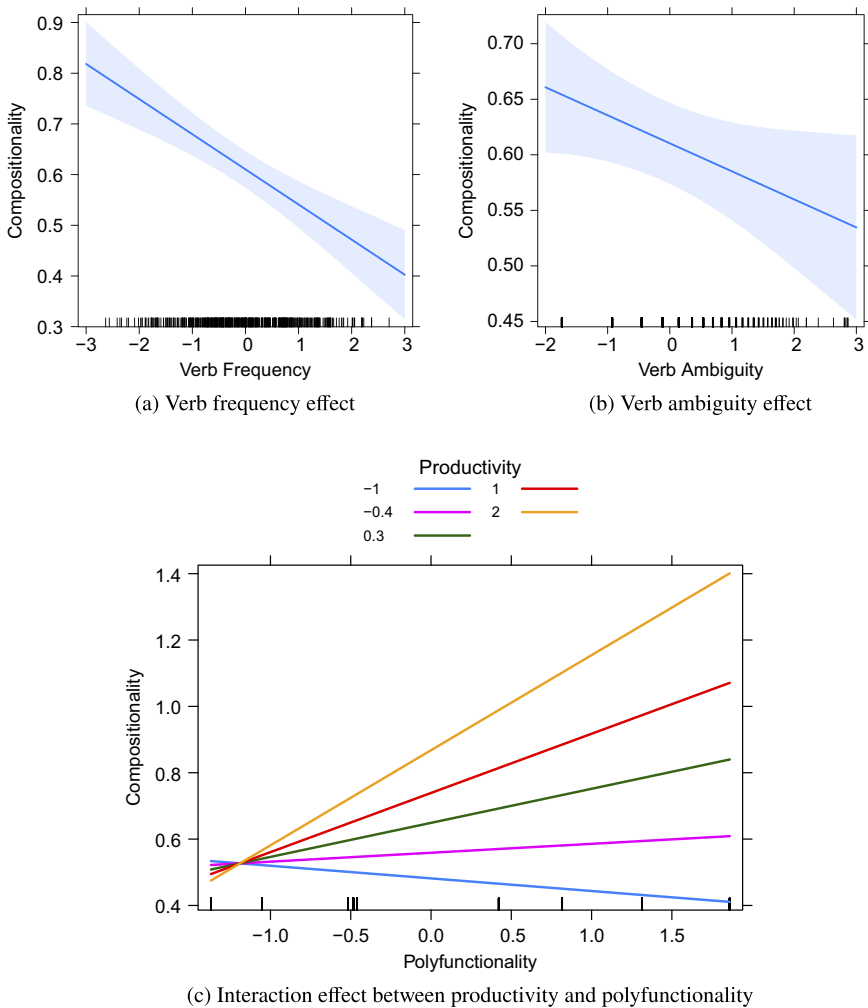


Figure 4. Effect plots for compositionality.

methods, relatedness and compositionality are not strongly correlated and are differently influenced by linguistic factors (as summarized in Figure 6). In this section, we discuss the findings for each factor examined, starting with lexical properties (frequency, ambiguity and attestation date of complex words) and then addressing morphological properties (productivity and polyfunctionality).

5.1. Lexical factors

The widespread assumption that the frequency of complex words correlates negatively with transparency was not verified in our results, neither for compositionality nor for relatedness. A negative correlation between frequency and transparency could be attributed to multiple reasons, such as lexical familiarity and direct lexical access (which allow disconnection from

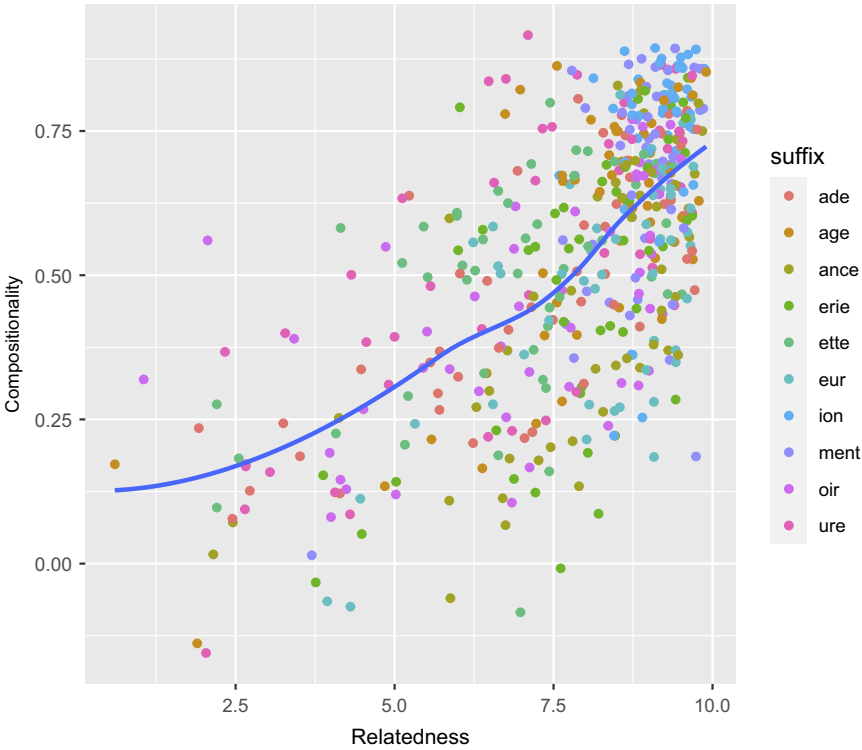


Figure 5. Scatterplot of relatedness and compositionality scores with fitted polynomial regression line.

Table 5. Results of a mixed-effect linear regression with compositionality as the response variable and including relatedness in the predictor variables. Significant *p*-values (at *p* < .05) are indicated in bold

	Estimate	Std. Error	df	<i>t</i> -value	<i>p</i> -value
(Intercept)	0.59	0.27	6.03	21.71	< . .001
Verb frequency	−0.06	0.01	485.15	−5.64	< . .001
Verb ambiguity	−0.01	0.01	482.55	−0.75	.456
Noun frequency	0.004	0.01	486.75	0.42	.675
Noun ambiguity	0.002	0.01	486.73	0.26	.791
Noun date of attestation	0.003	0.01	470.69	0.36	.720
Productivity	0.08	0.3	6.10	2.61	.040
Polyfunctionality	0.05	0.03	5.86	1.99	.094
Relatedness	0.19	0.01	486.03	15.88	< . .001
V fq:V ambiguity	−0.002	0.01	483.06	−0.25	.802
N fq:N ambiguity	0.003	0.01	482.73	0.44	.663
Prod:Polyfunct	0.09	0.04	5.69	2.28	.065

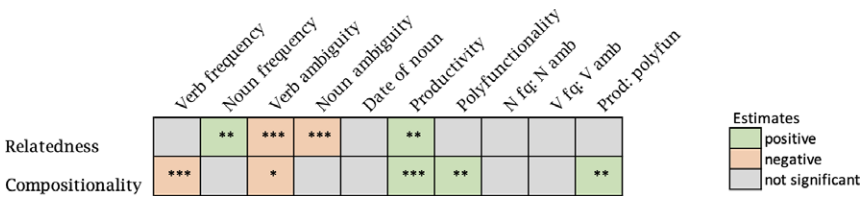


Figure 6. Significance of independent variables in regression models predicting relatedness and compositionality. Asterisk notation indicates statistical significance: $p < .05$ (*), $p < .01$ (**), and $p < .001$ (***).

the base), but also to the relationship between frequency and ambiguity. Based on the meaning-frequency law (Zipf, 1945), it can be assumed that frequent complex words have developed additional meanings over time and that these meanings are not necessarily related to the base. Accordingly, the hypothesized effect of frequency may be confounded with other factors, such as word age and lexical ambiguity. By including frequency, attestation date and ambiguity in the same analysis, we were able to examine the influence of frequency independently of potentially related factors. It was found that frequency per se does not necessarily impact transparency, especially when compared to lexical ambiguity.

More precisely, our results show that the frequency of derivatives does not affect compositionality, and unexpectedly, it is positively correlated with relatedness.¹⁷ In the case of relatedness ratings, the positive influence of frequency may be caused by greater familiarity with lexical items, which is known to affect human judgments – with less informed judgments generally eliciting lower similarity or relatedness ratings. The only negative correlation we identified was between base frequency and compositionality, suggesting that frequent verbs do not follow standard nominalization processes. This effect is specific to compositionality, since it is not observed with relatedness and persists in predictive models of compositionality that include relatedness as an independent variable. Accordingly, the correlation between base frequency and compositionality is determined by factors other than lack of relatedness. These factors may be linked to additional properties such as morphological family size and number of lexical competitors, if it appears that frequent verbs have many derivatives, or cause their derivatives to deviate from strict compositionality in order to be distinguished from near synonyms.

As discussed in Section 2.3.1, an alternative perspective advocated by Hay (2001) suggests that relative frequency has a greater impact on transparency than absolute frequency. We tested this claim by substituting absolute frequencies with the ratio between derivative and base frequency in our statistical analyses. A significant effect was observed on both relatedness and compositionality,¹⁸ but with an unexpected positive influence of relative frequency: the more frequent a derived noun is compared to its base verb, the more transparent it is. This effect may be caused by distinct factors depending on the aspect of transparency considered, as it may reflect the negative influence of base frequency on

¹⁷ This result echoes the findings of Johnson et al. (2023), although the positive correlation they reported was not based on human ratings but on a computational evaluation of transparency.

¹⁸ The inclusion of relative frequency in the regression models did not affect the other factors in the analysis of relatedness, but it turned noun ambiguity and noun attestation date into significant predictors in the analysis of compositionality. The complete estimates of the two regression models are provided in the Appendix section.

compositionality, and the positive influence of derivative frequency on relatedness. The idea that relative frequency is more influential than absolute frequency should therefore be nuanced, as the role of relative frequency can be differently affected by variations in absolute frequency. Overall, our results indicate that the effect of frequency on transparency is more complex than previously assumed and that it should be analyzed separately for relatedness and compositionality while controlling for related lexical factors.

The lexical ambiguity of both bases and derivatives emerged as a significant predictor of transparency with the anticipated effect: ambiguity always influences transparency negatively. This negative correlation can be explained by partial relatedness since not all senses of ambiguous bases are necessarily present in their derivatives, nor are all senses of ambiguous derivatives necessarily linked to their morphological bases. The more meanings bases and derivatives have, the higher the probability of partial relatedness and the greater the loss of transparency. Interestingly, the results show that semantic opacity is not influenced only by the ambiguity of the derivative but also by that of the base. A notable finding is that the ambiguity of derivatives is only marginally significant in compositionality measures ($p = .080$). It seems that ambiguity in derivatives has minimal impact on compositionality and that the loss of relatedness caused by derivative ambiguity has less influence on compositionality than that caused by base ambiguity. Further exploration of the role of ambiguity on transparency would require information about the frequency of each word sense, which may play a role both in relatedness ratings (depending on the pregnancy of the different word senses in speakers' minds) and in compositionality measures (especially when aggregating the different senses of a word in a single distributional representation).

No significant effect of the attestation date of the derivatives could be observed in the regression analyses predicting relatedness and compositionality (except in the model using relative frequency as a predictor of compositionality). This result contrasts with the findings of Lombard et al. (2022), and the difference may be explained by the fact that Lombard and colleagues examined the influence of word age without considering related factors. As seen in Figure 1, the attestation date of deverbal nouns is moderately correlated with their frequency and ambiguity, as well as with the frequency of their bases. Analyzing the combined influence of these factors may inhibit the apparent impact of word age if it mostly affects transparency through frequency and ambiguity. Note that the influence of word age is still overlooked in empirical studies on transparency, although the role of lexicalization on morphological demotivation has been pointed by many authors. A more comprehensive investigation into diachrony would be necessary, incorporating extensive data and wide ranges of attestation dates, before firm conclusions can be drawn about its influence on transparency.

5.2. *Morphological factors*

The two morphological properties we investigated as possible factors of transparency are productivity and polyfunctionality. The positive correlation usually assumed between productivity and transparency was confirmed for both relatedness and compositionality, although in the case of compositionality, an interaction with polyfunctionality was additionally observed. Nouns formed with productive deverbal suffixes tend to be not only clearly related to their bases but also consistent with the meaning conveyed by the suffixes. Generally speaking, transparency can be seen as a necessary condition for productivity since productivity relies on the compositionality of newly derived words to facilitate semantic

processing and effective communication. It can be noted that productivity remains a significant predictor in the regression analysis of compositionality that includes relatedness as an independent variable, which indicates that the effect of productivity on compositionality is not limited to that already involved in relatedness.

The influence of productivity has been discussed by Stupak & Baayen (2022), who found the distributional similarity between base and complex words in German to be positively correlated with realized productivity V , but not with potential productivity P . Our observations were based on potential productivity, but we additionally tested realized productivity by substituting it for potential productivity in the statistical models. Comparable effects were observed since V was also found to positively influence both relatedness and compositionality.¹⁹ The difference between our results and those presented by Stupak & Baayen (2022) is attributable to the method we used to compute potential productivity, which was based on arbitrary sample size (see Section 3.4). While it is often argued that V and P measures of productivity do not converge, our extrapolated version of P was strongly correlated with V (Kendall's $\tau = .96$). Note that regression models including a standard version of P show no significant effect of that variable on either relatedness nor compositionality, which is consistent with the observations of Stupak and Baayen.

As mentioned above, an interaction was observed between productivity and polyfunctionality in the analysis of compositionality. The influence of suffix polyfunctionality on compositionality, as opposed to relatedness, can be explained by the fact that compositionality is inherently dependent on the properties of derivational processes. Incidentally, these results confirm that morphological polyfunctionality is different from lexical ambiguity since polyfunctionality and ambiguity behave differently as predictors of transparency. It remains that the effect of polyfunctionality is not as anticipated, as in many cases, polyfunctionality enhances the effect of productivity. In particular, polyfunctionality has a positive impact on compositionality for nouns formed with moderately or highly productive suffixes, whereas it is a factor of opacity for nouns formed with the least productive suffixes. It seems that suffixes that are available for coining new words are more transparent if they are polyfunctional or, to put it differently, that a high number of semantic functions for productive suffixes is a factor of transparency. Productive nominalizing suffixes with multiple functions are more likely to be coherent in derivation than those with fewer functions, as if the semantic diversity of affixes was a condition for the semantic consistency of derivatives. However, more observations are needed to validate these results on a larger scale since only 10 suffixes were examined in the present study.

Overall, it appears that derivational processes may differ in terms of semantic transparency, both with respect to compositionality and relatedness, and that their potential for transparency is related to some of their inherent properties, such as productivity and polyfunctionality. Building on these findings, additional properties of morphological patterns (e.g. lexical class or semantic type of inputs and outputs) could be explored to provide a comprehensive understanding of the morphological factors that determine transparency.

¹⁹ The only difference observed between the models with P and V concerns the interaction effect with polyfunctionality in the prediction of compositionality, which became only marginally significant when replacing P with V . The detailed results of the regression analyses with realized productivity are provided in the Appendix section.

6. Conclusion

In this study, we have investigated the influence of various linguistic factors on the semantic transparency of nouns derived from verbs in French. We have distinguished two aspects of transparency: (i) relatedness as the degree to which the meaning of base words is retained in that of complex words and (ii) compositionality, as the degree to which the meaning of complex words is determined by the meaning of their constituents and the way they are combined. These aspects have been assessed through human judgments and computational methods, and their dependence on linguistic factors has been evaluated statistically.

The results of the study show that relatedness and compositionality can be differently influenced by lexical and morphological factors. Certain properties, such as base ambiguity and morphological productivity in the case of verb-to-noun derivation, can affect both relatedness and compositionally. However, the two aspects of transparency can also have distinct determinants. In the case of French deverbal nouns, relatedness is specifically influenced by derivative frequency and ambiguity, whereas compositionality is determined by base frequency and affix polyfunctionality.

It appears that previously identified factors of transparency are not necessarily influential or that their influence is more complex than usually assumed. Frequency in particular is not always negatively correlated with semantic transparency, and the frequency of base words can have a greater impact on transparency than that of complex words. As for morphological productivity, it may interact with other properties such as the polyfunctionality of word-formation processes. Additional factors related to morphological and lexical networks may also be considered, such as morphological family size or number of synonyms for complex words. Crucially, our study underscores the need to examine potential factors of transparency in combined analyses, in order to disentangle their respective influences, as their actual importance may be obscured by related factors.

The fact that relatedness and compositionality have different linguistic determinants confirms the necessity of distinguishing them when investigating the semantic transparency of complex words. The methods used to measure transparency, whether they are based on speakers' judgments or computational techniques, should specify which aspect of transparency is assessed. Further, the analysis of semantic transparency might be refined by providing a more detailed description of both relatedness and compositionality. Relatedness can be precisely evaluated depending on the number and frequency of word senses for both base and complex words, and the partial vs. total mapping of base and complex word senses. A thorough distinction can also be drawn, not only between compositional and non-compositional meanings, but also between complex words whose lexical meaning is strictly equivalent to the compositional meaning and those that include additional features on top of the compositional meaning. Such refinements could be reflected in the way transparency is assessed, and composite measures could be used to account for the different dimensions of semantic transparency.

Acknowledgments. The authors wish to thank three anonymous reviewers, as well as the audience of the 21st International Morphology Meeting held in Wien, for their feedback and comments.

Funding statement. This research was supported by the Swiss National Science Foundation under grant 10001F 188782 ('The semantics of deverbal nouns in French').

Data availability statement. The datasets, computational codes, and statistical scripts used in this study are available at <https://osf.io/md8kj/>

References

- Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge, MA: The MIT Press.
- Baayen, Harald. 1992. Quantitative aspects of morphological productivity. In *Yearbook of morphology* 1991, 109–149. Dordrecht: Springer.
- Baayen, Harald. 1993. On frequency, transparency and productivity. In *Yearbook of morphology* 1992, 181–208. Dordrecht: Springer.
- Baayen, Harald. 2001. *Word frequency distributions*. Dordrecht: Kluwer Academic Publishers.
- Baayen, Harald. 2009. Corpus linguistics in morphology: morphological productivity. In Anke Lüdeling & Merja Kyto (eds.), *Corpus linguistics. An international handbook*, 900–919. Berlin: De Gruyter.
- Baayen, Harald & Rochelle Lieber. 1997. Word frequency distributions and lexical semantics. *Computers and the Humanities*. 30. 281–291.
- Baroni, Marco, Georgiana Dinu & Germán Kruszewski. 2014. Don't count, predict! a systematic comparison of context-counting vs. context-predicting semantic vectors. In *Proceedings of the 52nd annual meeting of the association for computational linguistics (volume 1: Long papers)*, 238–247.
- Baroni, Marco & Stefan Evert. 2014. The zipfr package for lexical statistics: A tutorial introduction. <http://mirrors.nic.cz/R/web/packages/zipfr/vignettes/zipfr-tutorial.pdf>.
- Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. 67(1). 1–48.
- Bauer, Laurie. 2001. *Morphological productivity*. Cambridge: Cambridge University Press.
- Bell, Melanie & Martin Schäfer. 2016. Modelling semantic transparency. *Morphology*. 26(2). 157–199.
- Bojanowski, Piotr, Edouard Grave, Armand Joulin & Tomas Mikolov. 2017. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*. 5. 135–146.
- Boleda, Gemma. 2020. Distributional semantics and linguistic theory. *Annual Review of Linguistics*. 6(1). 213–234.
- Bonami, Olivier & Juliette Thuilier. 2019. A statistical approach to rivalry in lexeme formation: French *-iser* and *-ifier*. *Word Structure*. 12(1). 4–41.
- Brooks, Mollie E., Kasper Kristensen, Koen J. van Benthem, Arni Magnusson, Casper W. Berg, Anders Nielsen, Hans J. Skaug, Martin Maechler & Benjamin M. Bolker. 2017. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*. 9(2). 378–400.
- Bybee, Joan. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins Publishing.
- Evert, Stefan. 2004. A simple Inre model for random character sequences. In *7èmes journées internationales d'analyse statistique des données textuelles, Louvain-la-Neuve, Belgium*, vol. 2004, 411–422.
- Evert, Stefan & Marco Baroni. 2007. zipfr: Word frequency distributions in R. In *Proceedings of the 45th annual meeting of the ACL*, 29–32.
- Fernández-Domínguez, Jesus. 2013. Morphological productivity measurement: Exploring qualitative versus quantitative approaches. *English Studies*. 94(4). 422–447.
- Frisson, Steven, Elizabeth Niswander-Klement & Alexander Pollatsek. 2008. The role of semantic transparency in the processing of English compound words. *British Journal of Psychology*. 99(1). 87–107.
- Gaeta, Livio & Davide Ricca. 2002. Corpora testuali e produttività morfologica: i nomi d'azione italiani in due annate della Stampa. In Roland Bauer & Hans Goebel (eds.), *Parallela IX*, 223–249. Wilhelmsheld: Egert.
- Gaeta, Livio & Davide Ricca. 2006. Productivity in Italian word formation: a variable-corpus approach. *Linguistics*. 44(1). 57–89.
- Gagné, Christina L., Thomas L. Spalding & Kelly A. Nisbet. 2016. Processing English compounds: Investigating semantic transparency. *SKASE Journal of Theoretical Linguistics*. 13(2). 2–22.
- Gagné, Christina L., Thomas L. Spalding & Daniel Schmidtke. 2019. LADEC: The large database of English compounds. *Behavior Research Methods*. 51(5). 2152–2179.
- Guevara, Emiliano. 2010. A regression model of adjective-noun compositionality in distributional semantics. In Roberto Basili & Marco Pennacchiotti (eds.), *Proceedings of the 2010 workshop on GEometrical models of natural language semantics*, 33–37. Uppsala, Sweden: Association for Computational Linguistics. <https://aclanthology.org/W10-2805>.
- Günther, Fritz & Marco Marelli. 2019. Enter sandman: compound processing and semantic transparency in a compositional perspective. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 45(10). 1872–1882.

- Günther, Fritz, Marco Marelli & Jens Bölte. 2020. Semantic transparency effects in German compounds: a large dataset and multiple-task investigation. *Behavior Research Methods*. 52(3). 1208–1224.
- Guzmán Naranjo, Matías & Olivier Bonami. 2023. A distributional assessment of rivalry in word formation. *Word Structure*. 16(1). 87–114.
- Harris, Zellig S. 1954. Distributional structure. *Word*. 10. 146–162.
- Hay, Jennifer. 2001. Lexical frequency in morphology: is everything relative? *Linguistics*. 39(6). 1041–1070.
- Hill, Mark O. 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology*. 54(2). 427–432.
- Hocksema, Jack. 2000. Compositionality of meaning. In Geert Booij, Christian Lehmann & Joachim Mugdan (eds.), *Morphology: an international handbook on inflection and word-formation*, 851–857. Berlin: De Gruyter.
- Ji, Hongbo, Christina L. Gagné & Thomas L. Spalding. 2011. Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds. *Journal of Memory and Language*. 65(4). 406–430.
- Johnson, Martha Booker, Micha Elsner & Andrea D Sims. 2023. High frequency derived words have low semantic transparency mostly only if they are polysemous. Paper presented at the ISMO conference, Nancy.
- Jost, Lou. 2006. Entropy and diversity. *Oikos*. 113(2). 363–375.
- Juhász, Barbara J. 2007. The influence of semantic transparency on eye movements during English compound word recognition. In Roger P. G. Van Gompel, Martin H. Fischer, Wayne S. Murray & Robin L. Hill (eds.), *Eye movements: a window on mind and brain*, 373–389. Oxford: Elsevier.
- Kim, Say Young, Melvin J. Yap & Winston D. Goh. 2019. The role of semantic transparency in visual word recognition of compound words: A megastudy approach. *Behavior Research Methods*. 51(6). 2722–2732.
- Kisselew, Max, Sebastian Padó, Alexis Palmer & Jan Šnajder. 2015. Obtaining a better understanding of distributional models of German derivational morphology. In Matthew Purver, Mehrnoosh Sadzadeh & Matthew Stone (eds.), *Proceedings of the 11th international conference on computational semantics*, 58–63. London, UK: Association for Computational Linguistics. <https://aclanthology.org/W15-0108>.
- Kotowski, Sven & Martin Schäfer. 2023. Quantifying semantic relatedness across base verbs and derivatives: English *out-* prefixation. In Sven Kotowski & Ingo Plag (eds.), *The semantics of derivational morphology. Theory, methods, evidence*, 177–218. Berlin: De Gruyter.
- Landauer, Thomas K. & Susan T. Dumais. 1997. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*. 104(2). 211–240.
- Lazaridou, Angeliki, Marco Marelli, Roberto Zamparelli & Marco Baroni. 2013. Compositionally derived representations of morphologically complex words in distributional semantics. In Hinrich Schuetze, Pascale Fung & Massimo Poesio (eds.), *Proceedings of the 51st annual meeting of the association for computational linguistics (volume 1: Long papers)*, 1517–1526. Sofia, Bulgaria: Association for Computational Linguistics. <https://aclanthology.org/P13-1149>.
- Lenci, Alessandro. 2018. Distributional models of word meaning. *Annual Review of Linguistics*. 4(1). 151–171.
- Libben, Gary, Christina L. Gagné & Wolfgang U. Dressler. 2020. The representation and processing of compound words. In Vito Pirrelli, Ingo Plag & Wolfgang U. Dressler (eds.), *Word knowledge and word usage. a cross-disciplinary guide to the mental lexicon*, 336–352. Berlin: De Gruyter.
- Libben, Gary, Martha Gibson, Yeo Bom Yoon & Dominiek Sandra. 2003. Compound fracture: the role of semantic transparency and morphological headedness. *Brain and Language*. 84(1). 50–64.
- Lombard, Alizée, Marine Wauquier, Cécile Fabre, Nabil Hathout, Lydia-Mai Ho-Dac & Richard Huyghe. 2022. Evaluating morphosemantic demotivation through experimental and distributional methods. *Linguisticae investigationes*. 45(1). 83–115.
- Marelli, Marco & Marco Baroni. 2015. Affixation in semantic space: modeling morpheme meanings with compositional distributional semantics. *Psychological Review*. 122(3). 485.
- Marelli, Marco & Claudio Luzzatti. 2012. Frequency effects in the processing of Italian nominal compounds: modulation of headedness and semantic transparency. *Journal of Memory and Language*. 66(4). 644–664.
- Marslen-Wilson, William, Lorraine K. Tyler, Rachel Wakslar & Lianne Older. 1994. Morphology and meaning in the English mental lexicon. *Psychological Review*. 101(1). 3–33.
- Mikolov, Tomas, Kai Chen, Greg Corrado & Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
- Mitchell, Jeff & Mirella Lapata. 2010. Composition in distributional models of semantics. *Cognitive Science*. 34(8). 1388–1429.
- Padó, Sebastian, Aurélie Herbelot, Max Kisselew & Jan Šnajder. 2016. Predictability of distributional semantics in derivational word formation. In Yuji Matsumoto & Rashmi Prasad (eds.), *Proceedings of COLING 2016, the*

- 26th international conference on computational linguistics: Technical papers, 1285–1296. Osaka, Japan: The COLING 2016 Organizing Committee. <https://aclanthology.org/C16-1122>.
- Pagliuca, William. 1976. PRE-fixing. Manuscript, SUNY Buffalo.
- Panther, Klaus-Uwe & Linda L. Thornburg. 2002. The roles of metaphor and metonymy in English -er nominals. In René Dirven & Ralf Pörings (eds.), *Metaphor and metonymy in comparison and contrast*, 279–322. Berlin, New York: De Gruyter Mouton.
- Pennington, Jeffrey, Richard Socher & Christopher Manning. 2014. GloVe: Global vectors for word representation. In Alessandro Moschitti, Bo Pang & Walter Daelemans (eds.), *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 1532–1543. Doha, Qatar: Association for Computational Linguistics. <https://aclanthology.org/D14-1162>.
- Plag, Ingo. 2003. *Word-formation in English*. Cambridge: Cambridge University Press.
- R Core Team. 2024. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Rastle, Kathleen, Matt H. Davis, William D. Marslen-Wilson & Lorraine K. Tyler. 2000. Morphological and semantic effects in visual word recognition: a time-course study. *Language and Cognitive Processes*. 15(4–5). 507–537.
- Reddy, Siva, Diana McCarthy & Suresh Manandhar. 2011. An empirical study on compositionality in compound nouns. In Haifeng Wang & David Yarowsky (eds.), *Proceedings of 5th international joint conference on natural language processing*, 210–218. Chiang Mai, Thailand: Asian Federation of Natural Language Processing.
- Řehůřek, Radim & Petr Sojka. 2010. Software framework for topic modelling with large corpora. In *Proceedings of the LREC 2010 workshop on new challenges for NLP frameworks*, 45–50. Valletta, Malta: ELRA. <http://is.muni.cz/publication/884893/en>.
- Roswell, Michael & Joanathan Dushoff. 2023. *Meanrarity: Hill diversity estimation and visualization*. <https://github.com/mikeroswell/MeanRarity>. R package version 0.0.1.0005.
- Roswell, Michael, Jonathan Dushoff & Rachael Winfree. 2021. A conceptual guide to measuring species diversity. *Oikos*. 130(3). 321–338.
- Ryder, Mary Ellen. 1999. Bankers and blue-chippers: an account of -er formations in Present-day English. *English Language and Linguistics*. 3(2). 269–297.
- Salvadori, Justine, Lucie Barque, Pauline Haas, Richard Huyghe, Alizée Lombard, Matthieu Monney, Sandra Schwab, Delphine Tribout, Rossella Varvara & Marine Wauquier. 2024. SONDE, the semantics of nouns derived from verbs in French: annotation guidelines. https://github.com/semantics-deverbal-nouns/annotation-guide/blob/main/annotation_guide_1.3.pdf.
- Salvadori, Justine & Richard Huyghe. 2023. Affix polyfunctionality in french deverbal nominalizations. *Morphology*. 33(1). 1–39.
- Sandra, Dominiek. 1990. On the representation and processing of compound words: automatic access to constituent morphemes does not occur. *The Quarterly Journal of Experimental Psychology*. 42(3). 529–567.
- Schäfer, Martin. 2018. *The semantic transparency of English compound nouns*. Berlin: Language Science Press.
- Schäfer, Roland. 2015. Processing and querying large web corpora with the COW14 architecture. In Piotr Bański, Hanno Biber, Evelyn Breiteneder, Marc Kupietz, Harald Lungen & Andreas Witt (eds.), *Proceedings of challenges in the management of large corpora 3 (cmcl-3)*, UCREL Lancaster: IDS.
- Schäfer, Roland & Felix Bildhauer. 2012. Building large corpora from the web using a new efficient tool chain. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the eighth international conference on language resources and evaluation (LREC'12)*, 486–493. Istanbul, Turkey: European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2012/pdf/834_Paper.pdf.
- Schreuder, Robert & Harald Baayen. 1995. Modeling morphological processing. In Laurie B. Feldman (ed.), *Morphological aspects of language processing*, 131–154. Hillsdale, NJ: Lawrence Erlbaum.
- Simpson, Edward H. 1949. Measurement of diversity. *Nature*. 163. 688.
- Stupak, Inna & Harald Baayen. 2022. An inquiry into the semantic transparency and productivity of german particle verbs and derivational affixation. *The Mental Lexicon*. 17(3). 422–457.
- Varvara, Rossella, Gabriella Lapesa & Sebastian Padó. 2021. Grounding semantic transparency in context: A distributional semantic study on german event nominalizations. *Morphology*. 31. 409–446.
- Varvara, Rossella, Justine Salvadori & Richard Huyghe. 2022. Annotating complex words to investigate the semantics of derivational processes. In Harry Bunt (ed.), *Proceedings of the 18th joint ACL-ISO workshop on interoperable semantic annotation within LREC2022*, 133–141. Marseille, France: ELRA.

Wauquier, Marine, Nabil Hathout & Cécile Fabre. 2020. Semantic discrimination of technicality in french nominalizations. *Zeitschrift für Wortbildung/Journal of Word Formation*. 4(2). 100–119.

Zanzotto, Fabio Massimo, Ioannis Korkontzelos, Francesca Fallucchi & Suresh Manandhar. 2010. Estimating linear models for compositional distributional semantics. In Chu-Ren Huang & Dan Jurafsky (eds.), *Proceedings of the 23rd international conference on computational linguistics (Coling 2010)*, 1263–1271. Beijing, China: Coling 2010 Organizing Committee. <https://aclanthology.org/C10-1142>.

Zimmer, Karl E. 1966. Affixal negation in english and other languages: an investigation of restricted productivity. *Language*. 42(1). 134–142.

Zipf, George Kingsley. 1945. The meaning-frequency relationship of words. *The Journal of General Psychology*. 33(2). 251–256.

Zwitserslood, Pienie. 1994. The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes*. 9(3). 341–368.

Appendix

Table A1. Results of a mixed-effect beta regression with relatedness as the response variable and including relative frequency in the predictor variables. Significant p-values (at $p < .05$) are indicated in bold.

	Estimate	Std. Error	z –value	p-value
(Intercept)	1.31	0.08	15.90	< . 001
Verb ambiguity	–0.15	0.05	–3.13	.002
Noun ambiguity	–0.15	0.04	–3.41	< . 001
Relative frequency	0.09	0.04	2.13	.033
Noun date of attestation	–0.02	0.04	–0.40	.692
Productivity	0.27	0.09	2.89	.004
Polyfunctionality	0.09	0.08	1.19	.236
Prod:Polyfunct	0.09	0.11	0.85	.394

Table A2. Results of a mixed-effect linear regression with compositionality as the response variable and including relative frequency in the predictor variables. Significant p-values (at $p < .05$) are indicated in bold.

	Estimate	Std. Error	df	t –value	p-value
(Intercept)	0.60	0.02	5.62	27.54	< . 001
Verb ambiguity	–0.04	0.01	488.80	–3.52	< . 001
Noun ambiguity	–0.03	0.01	461.13	–2.65	.008
Relative frequency	0.05	0.01	490.84	4.30	< . 001
Noun date of attestation	0.03	0.01	284.98	2.55	.011
Productivity	0.12	0.03	6.39	4.72	.003
Polyfunctionality	0.07	0.02	6.04	3.18	.019
Prod:Polyfunct	0.11	0.03	5.64	3.48	.014

Table A3. Results of a mixed-effect beta regression with relatedness as the response variable and including realized productivity in the predictor variables. Significant p -values (at $p < .05$) are indicated in bold.

	Estimate	Std. Error	z-value	p-value
(Intercept)	1.31	0.07	18.02	< . .001
Verb frequency	−0.05	0.05	−0.94	.350
Verb ambiguity	−0.17	0.05	−3.37	< . .001
Noun frequency	0.15	0.05	2.79	.005
Noun ambiguity	−0.15	0.05	−3.33	< . .001
Noun date of attestation	0.04	0.05	0.72	.472
Realized productivity	0.21	0.09	2.28	.023
Polyfunctionality	0.01	0.06	0.14	.889
V fq:V ambiguity	−0.04	0.03	−1.31	.190
N fq:N ambiguity	−0.04	0.04	−1.27	.205
Realized prod:Polyfunct	0.03	0.12	0.26	.798

Table A4. Results of a mixed-effect linear regression with compositionality as the response variable and including realized productivity in the predictor variables. Significant p -values (at $p < .05$) are indicated in bold.

	Estimate	Std. Error	df	t -value	p-value
(Intercept)	0.58	0.02	6.57	25.04	< . .001
Verb frequency	−0.07	0.01	473.88	−5.28	< . .001
Verb ambiguity	−0.03	0.01	482.06	−2.12	.035
Noun frequency	0.02	0.01	488.00	1.73	.084
Noun ambiguity	−0.02	0.01	481.48	−1.94	.053
Noun date of attestation	0.01	0.01	451.88	0.90	.369
Realized productivity	0.10	0.03	6.45	3.33	.014
Polyfunctionality	0.02	0.02	5.98	1.21	.272
V fq:V ambiguity	−0.01	0.01	485.14	−1.16	.248
N fq:N ambiguity	−0.00	0.01	484.86	−0.22	.827
Realized prod:Polyfunct	0.10	0.04	5.92	2.32	.060