Lucie Barque

Université Sorbonne Paris
Nord & Laboratoire de Linguistique Formelle

[iD] https://orcid.org/0000-0002-1806-848X

Is metonymy really more regular than metaphor? A first investigation into the French Lexicon

Abstract

Regular polysemy is considered more intrinsically linked to metonymy than to metaphor in both linguistic and psycholinguistic literature. Yet, on the one hand, metonymy exhibits varying degrees of regularity, and on the other hand, metaphor is also subject to polysemic regularity. This paper aims to assess the extent to which metonymy and metaphor exhibit regularity in the lexicon. Based on the analysis of lexicographic descriptions of a sample of nearly 3,000 French nouns, this study compares the regularity of polysemy patterns—defined as recurring sense alternations—within the two types of figures. While the findings confirm theoretical assumptions by identifying more regular metonymic patterns than regular metaphoric ones, they reveal, on average, no significant difference between metonymy and metaphor in terms of regularity. The resulting dataset is publicly available to allow further exploration of how regular polysemy emerges in the lexicon and its correlation with lexical figures.

Keywords

Regular polysemy, metonymy, metaphor, French lexicon

1. Introduction

Since Jurij Apresjan's foundational paper on regular polysemy (1974), both linguistic and psycholinguistic literature have consistently associated metonymy with regularity and metaphor with irregularity. Theoretical studies on regular

polysemy have primarily focused on highly regular metonymic alternations, such as the animal-meat or the container-content relations (Nunberg & Zaenen, 1997; Pustejovsky, 1995, among others). In contrast, metaphor has received relatively little attention in discussions of regular polysemy, despite also exhibiting patterns of regularity (e.g., Copestake & Briscoe, 1995; Vicente & Falkum, 2017).

In this paper, we aim to investigate to what extent metonymy is more regularly observed than metaphor in the lexicon, both in terms of number of sense alternations and in terms of number of polysemous words associated with each of them. To do so, we analyse a large set of sense alternations observed within a broad sample of French nouns and provide regularity scores for metonymy and metaphor patterns. In addition to empirically testing the theoretical assumption of a strong correlation between figure and regularity, this study aims to advance the understanding of the various factors that contribute to the regularity of sense alternations. It also seeks to provide data that can be used in psycholinguistic studies, for instance, to examine the distinct effects of figure and regularity on the cognitive processing of polysemous words and their mental representation[1].

The remainder of the paper is organized as follows. Section 2 provides a brief review of previous research on the relationship between lexical figures and regular polysemy, as well as past attempts to assess scalar regularity using lexical resources or corpora. Section 3 details the methodology used for data selection and analysis. Section 4 presents the main findings, which confirm that metaphor also plays a central role in the formation of regular polysemy, although generating fewer patterns than metonymy. However, the proportion of polysemous words it produces is comparable to that of metonymic patterns.

2. Related work

2.1. Regular polysemy and lexical figures

According to Apresjan (1974), a sense extension exhibited by a word can be considered regular if at least one other word within the same semantic class undergoes the same type of alternation. Both theoretical and empirical studies in linguistics tend to associate regular polysemy with metonymy (Nunberg & Zaenen,

---

[1]   The data are available at https://osf.io/txvhw/.

1997; Pustejovsky, 1995; Dölling, 2020). For instance, among the 50 nominal regular polysemy patterns listed in Apresjan (1974), 47 are metonymic. Similarly, Srinivasan and Rabagliati (2015) examine the cross-linguistic occurrence of regular polysemy by testing 27 patterns, 24 of which are metonymic. Psycholinguistic studies on the processing and mental representation of ambiguous words also suggest a correlation between figurative meaning and regularity: metonymy is seen as involving more closely related senses than metaphor (e.g., Klepousniotou et al., 2008; Lopukhina et al., 2018; Yurchenko, 2020) and more closely related senses tend to exhibit greater regularity (e.g., Xu et al., 2020).

The connection between figure and regularity can be partly explained by the structural characteristics of the two figures. Metonymy relies on a contiguity relationship between the referent of the derived sense and the referent of the source sense, which results in a referent closely tied to the one denoted by the source. In contrast, metaphor is based on an analogy between two referents, leading to the creation of a referent distinct from the original. This difference generally means that two senses linked by metonymy are more closely related than those linked by metaphor, especially when contiguity involves referential inclusion—i.e., when the referent of the derived sense is part of the referent of the source sense (e.g., in the metonymy "animal > meat"). As a result, the target sense is generally more predictable. Thus, interpreting or producing a metonymic sense, which is generally closer to its source meaning and more predictable than a metaphorical sense, requires less cognitive effort, which helps explain why metonymy is more regularly used.

2.2. Identifying polysemy patterns

The use of structured large-scale lexical resources has enabled the automatic detection of lemmas exhibiting similar sense alternations. This approach, primarily applied to English WordNet data (Buitelaar, 1998; Peters, 2006, a.o.), has led to inventories of polysemy patterns with varying degrees of completeness and precision. For example, Buitelaar (2000) leveraged the synset hierarchy in WordNet to identify 1,341 sense clusters that potentially represent regular polysemy across 3,336 English nouns. For French, Barque et al., (2018) identified around a hundred metonymy and metaphor patterns from two lexical resources. While these studies have provided initial data on regular alternations in the lexicon, the results—extracted automatically—have not been fully validated through manual annotation, nor were they obtained from a systematic sampling of the lexicon.

Two lexical resources hold promise for a more precise identification of polysemy patterns, as they provide hand-crafted annotations of relations between word senses. The RL-fr (*Réseau Lexical du Français*) is a French lexical network in which nodes correspond to semantically specified lexical units, and arcs represent paradigmatic or syntagmatic relations. One particular paradigmatic relation is co-polysemy (Polguère, 2018). The RL-fr is still under development and currently contains 9,659 co-polysemy relations, which can be leveraged to infer regular patterns of sense alternation (Polguère, 2022). As for English, Chain-Net is a recently released database in which relations between 22,178 senses of 6,500 words from Princeton WordNet 3.0 (Fellbaum, 1998) have been annotated for metonymy, metaphor, or homonymy (Maudslay et al., 2024). However, this second set of studies has not yet undertaken a comprehensive inventory of regular alternations.

Our study aims to bridge this gap by offering a rigorous identification and classification of regular polysemy patterns attested in the French lexicon.

2.3. Assessing polysemy patterns regularity

Calculating the degree of regularity of a polysemy pattern requires, at least, evaluating the proportion of words that truly fit the pattern in relation to the total number of words that could potentially do so. It also requires information about the frequency of these words and, ideally, the distribution of their meanings. The more polysemous words a pattern generates in the language relative to its generative potential, the more regular it is. This regularity increases if the words belonging to the pattern are commonly used and if the polysemous words it generates have balanced meanings. For example, nouns like *bœuf* 'beef' and *autruche* 'ostrich' should not contribute equally to calculating the regularity of the "animal>meat" pattern, as *bœuf* is more frequent than *autruche*, and the meanings 'animal' and 'meat' are more balanced for the former than for the latter. Accordingly, at least four variables should be considered when determining the regularity of a polysemy pattern A→B:

1. The number of words attested with sense A,
2. The number of words attested with both senses A and B,
3. The frequency of these words in the corpus,
4. The relative frequency of senses A and B for polysemous words.

Lombard et al., (2023) proposed several measures that account for a varying number of parameters depending on the availability of the data. These measures yield regularity scores that range from 0 (indicating irregularity) to 1 (indicating

systematicity), enabling comparisons between patterns within the same language and across different languages. Since the information related to parameter 4 is unavailable for a large portion of the lexicon, both for French and other languages, we will use the following formula, which relies on the first three parameters to compute the regularity degree of a pattern $P_{A \to B}$. In this formula, n represents the number of polysemous words instantiating P, N denotes the number of words having source meaning A in the dataset, and f corresponds to the word's frequency. Frequencies are log-transformed to minimize discrepancies in absolute frequency between words.

$$Reg(P_{A \to B}) = \frac{\sum_{i=1}^{n} \log(f_i)}{\sum_{j=1}^{N} \log(f_j)}$$

2.4. This study

For practical reasons, we restrict our analysis to sense alternations involving meanings that do not belong to the same coarse semantic class, such as the animal–person metaphor (e.g., *mouton* 'sheep/follower') or the animal–meat metonymy (e.g., *mouton* 'sheep/mutton'). By contrast, polysemic extensions within the same semantic class, such as the person-person metaphor (e.g., *boucher* 'butcher/killer'), are not considered here. Nevertheless, it could be the case that one of the two figures of speech under study is more likely than the other to generate new meanings within the same semantic class as the source meaning. This suggests that our account of the predominance of metonymy in regular polysemy should ultimately be complemented by an analysis of such within-class extensions.

The study is based on the analysis of lexicographic data to infer properties of regular polysemy in the lexicon. We know that lexicographic data offer an imperfect view on the lexicon of a given language, defined as the theorical sum of words and word senses known and/or used by the speakers of that language. This is particularly problematic for assessing polysemy regularity, since the more a polysemy is regular, the more the meaning it produces is predictable (from the source meaning), so the less it is probable for these senses to be systematically consigned in the dictionary. For instance, in theory, any noun that primarily refers to a cooking container can also be used denote the quantity of food it holds (e.g., *she ate two plates of rice*). Another well-known limitation of lexicographic data is that they do not provide information on sense frequency, however a crucial parameter for assessing the regularity of polysemy patterns.

Despite these limitations, our study will give a first empirical assessment of how metonymy and metaphor contribute to regular polysemy in the French nominal lexicon. It also provides a large set of manually curated data for conducting psycholinguistic experiments on the intricate relationships between lexical figures and regularity.

3. Data

In this section, we first expose the methodology used to select a sample of nouns representing common French and to extract regular sense alternation based on their lexicographic description (Section 3.1). Then we detail how we manually analysed these data in order to identify regular sense extensions as well as their instances in the selected data (Section 3.2).

3.1. Data selection

3.1.1. A semantically annotated version of the French Wiktionary

As noted above, most previous studies that aimed to identify polysemy patterns from lexicographic data used WordNet (Buitelaar, 1998; Peters, 2006, a.o.). Since French lacks WordNet-like resources that meet the two conditions of coverage and quality, we leverage data from the *SuperWikt-fr* lexicon, a recently released version of the French Wiktionary in which every nominal sense has been automatically annotated with its general semantic class (Angleraud et al., 2025). The resource is broad coverage, with more than 300,000 annotated nominal senses, and the quality of the automatic semantic classification allows for an extraction of sense alternation patterns that we will be further analysed manually. The semantic classification was performed using supervised classifiers trained and tested on a large set of manually curated data. The classification reached about 85% precision, with performance varying across semantic categories. The easiest senses to classify were unsurprisingly those having concrete referents, such as Person (F-score: 96%), Body (F-score: 85%) or Artifact (F-score: 85%), while the most challenging cases mostly correspond to abstract meanings, such as Cognition (66%) or State (62%). Moreover, the use of supersenses as semantic tags (e.g., Person, Act, State) allows for a certain interoperability with other studies investigating regular polysemy with WordNet-like resources for other languages.

3.1.2. A representative sample of familiar French nouns

The objective of this study is to examine to what extent metonymy exhibits greater regularity than metaphor in the French lexicon. For this purpose, we need to define a representative sample of French that is large enough to capture most of the recurring sense alternations while being of reasonable size to allow for a detailed manual analysis. To achieve this, we leverage data from *Lexique 3*, a resource that provides information on words found in a large French corpus, including both textual frequency and subjective familiarity estimates (New et al., 2004). More precisely, we focus on the 5,989 nouns reported in *Lexique 3* as being known by 100% of respondents. This constraint ensures that the analysed set consists of words generally familiar to adult French speakers. It should be noted, however, that this criterion does not specify which senses of these words are commonly known—an issue to which we return below.

Since our interest lies in regular polysemy, we exclude from this set the 1,579 nouns that are monosemous, i.e., associated with a single sense in *SuperWikt-fr*. Among the remaining 4,410 multisemous nouns, 1,494 are mono-categorial, meaning all their different senses belong to the same broad semantic class (or supersense). Because our analysis is restricted to sense alternations that involve a shift in semantic category, these nouns are also set aside. Table 1 presents the distribution of the 2,916 remaining nouns according to the number of distinct supersenses their meanings have been automatically associated within the *SuperWikt-fr*.

Table 1.

Repartition of the 2,926 nouns of the dataset according to their ambiguity level (supersense granularity)

Ambiguity level	2	3	4	5	6	7	8	9	10	11	
Nb of nouns	1,510	750	353	165	68	42	16	5	4	3	2,916

The distribution of the 8,455 supersense instances in the dataset is given in Table 2, in descending order. The table provides a raw characterisation of each supersense label.

Table 2

Distribution of supersenses in the dataset

Supersense	Definition	Nb
Artifact	Manufactured physical object (includes devices, buildings, constructed place, etc)	1,350
Act	Action caused by an agent	1,074
Cognition	Mental object / informational object and other products of mental activity (includes fields of knowledge, doctrines and methods)	950
Person	Humans and imaginary characters with prominent human attributes	768
Attribute	Constitutive property of an entity or of a situation	464
State	Non-constitutive property of an entity, resulting states, or state of affairs	346
Animal	Non-human living being	338
Object	Natural physical object (includes microorganisms and natural places)	329
Food	Natural food or prepared food	319
Event	Dynamic situation without an agent (natural event, accidental event)	295
Body	Body or body part or organ of animals or humans	276
Institution	All types of organizations, created with a purpose	248
Substance	Physical material (liquid, solid, or gas) plus chemical compounds, whether manufactured or not	220
Plant	Entity of the plant kingdom or a part of this entity	215
Artifact*Cognition	Document, informational medium	213
Quantity	Abstract numerical object (includes numbers, units of measurement, proportions, etc.)	181
Act*Cognition	Communication act	150
Possession	Abstract entities of the type intangible assets or financial assets (includes money, intellectual property, financial assets)	145
GroupxPerson	Group of people	127
Communication	Linguistic objects, languages, communication medium (e.g., radio, internet)	123
Feeling	Transitory psychological or physiological state	117
Time	Time period or time reference	107

<div align="right">Table 2 (Continuation)</div>

Supersense	Definition	Nb
Phenomenon	Entity observed at a given moment, subject to a form of temporality but not occurring (natural forces, sensory manifestations)	100
		8,455

3.1.3. Extracting candidates for polysemy patterns

The algorithm for automatically extracting groups of nouns with similar supersense alternations operates in two steps. First, following the order in which senses are described, we group together those that belong to the same supersense. For instance, the noun *punaise* '*bug', 'wretch', 'drawing pin*' has four senses in *Wiktionary*, with senses 2 and 3 falling under the same semantic category, *Person*, and thus being grouped together (1).

(1) *punaise* 'bug', 'wretch', 'drawing pin'
Animal [1. Nom désignant de nombreuses sortes d'insectes, souvent de forme plate, qui émettent une odeur désagréable lorsqu'ils sont effrayés ou tués. 'A name referring to many types of insects, often flat in shape, that emit an unpleasant odor when frightened or crushed.']
Person [2. (Par analogie) (Péjoratif) Personne petite, malingre, peu imposante. '(By analogy) (Pejorative) A small, scrawny, unimposing person.', 3. (Familier) (Par extension) Personne malveillante. '(Informal) (By extension) A malicious person.']
Artifact [4. (Par analogie de forme) Petit clou court, à tête plate et large, dont on se sert pour fixer au mur ou sur une planche des feuilles de papier, des images, des affiches, etc. '(By analogy of shape) A short, small nail with a flat, wide head, used to fasten sheets of paper, pictures, posters, etc., to a wall or a board.']

Then, we examine the relationship between the first supersense and each of the subsequent supersenses. For a noun whose senses are grouped into *n* distinct supersenses, *n-1* supersense alternations are extracted. Returning to *punaise* in (1), only the semantic alternations of type Animal-Person and Animal-Artifact are analysed. This approach calls for several comments. In the specific case of *punaise*, it makes sense not to compare the two Person-type senses with the Artifact sense, as they are not directly related. However, it is often the case that a derived sense does not stem directly from the primary meaning of the word but

rather from another derived sense. It would therefore have been more satisfactory to analyse all possible ordered combinations of distinct supersenses. However, the reality of the amount of data to be analysed leads us to limit ourselves to this first type of extension, leaving the other cases for future study. Furthermore, we assume that the first listed sense corresponds to the primary meaning, from which all other senses derive, either directly or indirectly. We are aware that this proxy has its limitations, especially concerning morphologically complex words, which will be discussed later.

Following this methodology, we identified a total of 354 recurring supersense alternations, with the number of nouns per pattern ranging from 2 to 184. Table 3 presents the 15 most frequently observed supersense alternations in the selected data, each illustrated by a noun with at least two senses annotated with these supersenses.

Table 3.

Most frequent supersense alternations automatically extracted from the dataset

Sense alternation	Nouns	Example
Artifact – Cognition	184	*tube* 'tube / hit'
Act – Artifact	178	*chargement* 'loading / cargo'
Artifact – Act	137	*équerre* 'square / L-sit'
Person – Artifact	128	*cycliste* 'cyclist / cycling shorts'
Act – Event	118	*arrêt* 'stop something / something stops'
Act – Cognition	110	*révélation* 'disclosure'
Animal – Person	108	*mouton* 'sheep'
Artifact – Body	91	*trompe* 'horn / trunk'
Artifact – Object	89	*bouton* 'button / pimple'
Person – Animal	72	*ouvrière* 'worker / worker bee'
Cognition – Act	72	*signature* 'signature / signing'
Animal – Artifact	71	*souris* 'mouse'
Act – State	69	*isolement* 'isolation'
Act – Attribute	62	*organisation* 'organization / set up'
Body – Artifact	61	*pied* 'foot / (table) leg'

3.2. Data analysis

Not every supersense alternation automatically extracted from the dataset instantiates a given polysemy pattern. The analysis therefore involves identifying which ones do and which ones do not.

3.2.1. Discarding unknown meanings

The first step of the analysis is to determine whether the annotator—the author of the paper in this case—is familiar with the meanings associated with a given supersense pair. We know that the selected nouns are well-known, given the way they were selected from *Lexique 3* (cf. section 3.1.2), but *Wiktionary* provides a detailed account of polysemy, including many specialized meanings. Since analysing relationships between senses that are at least partially unknown introduces significant uncertainty, such senses must be excluded from our analysis. The examples below illustrate cases where the author is not familiar with the two senses described in (2b) and (3b). Linguistic labels are often included in definitions (in brackets, such as [mécanique]), which could potentially be used as indicators for automatically identifying rare meanings. However, since these labels are not consistently applied to mark specialized or uncommon senses, as seen in (2b), and can also appear for common meanings, as shown in (3a), this step was carried out manually.

   (2) *polka* 'polka', 'wheelbarrow'
  a. Danse de salon d'origine tchèque ou polonaise.
    (A ballroom dance of Czech or Polish origin.)
  b. Brouette utilisée dans les champignonnières.
    (A wheelbarrow used in mushroom farms)

   (3) *rouage* 'component / toll'
  a. [mécanique] Chacune des pièces qui concourent au fonctionnement d'une machine.
    ([mechanics] Each component that contributes to the operation of a machine)
  b. [histoire, Droit d'Ancien Régime] Péage pour circuler sur les routes.
    ([History, Old Regime Law] Toll for travelling on roads)

As explained in the previous section, analysed supersense alternations occur between groups of senses (e.g., the *Person* supersense associated with two definitions in example (1)). A supersense alternation *A–B* associated with

a given noun is discarded if all senses of *A* or all senses of *B* are unknown to the annotator.

3.2.2. Identifying semantic classification errors

For each noun exhibiting a supersense alternation A–B, we must first verify whether the two senses involved indeed correspond to semantic type A and B respectively. Since nominal senses have been automatically tagged based on their lexicographic description (cf. section 3.1.1), some errors are inevitably present in the data. The two examples below illustrate such false positives. In (4), the two senses of *boulanger* 'baker' does not instantiate a Person-Artifact supersense alternation, as the target meaning has been misclassified—likely because the noun *tenant* used in its definition can also refer to an artifact. In (5), the source meaning of *nounours* 'teddy bear' has been classified as Animal instead of Artifact, probably due to the fact that the lexicographic definition uses *ours* (bear) as hypernym rather than explicitly stating that the noun refers to a toy.

(4) *boulanger* (baker)
a. Artisan dont le métier est de fabriquer ou de vendre le pain. → Person
   (Someone who bakes and sells bread)
b. Tenant d'une boulangerie. → ~~Artifact~~ [Person]
   (The owner of a bakery)

(5) *nounours* (teddy bear)
a. Ours en peluche. → ~~Animal~~ [Artifact]
   (Teddy bear, litt. Plush bear)
b. Person très caline. → Person
   (Cuddly person)

Once corrected, these pairs of meanings are either filtered out—if both meanings belong to the same broad class, as in *boulanger* 'baker' – or reclassified into the appropriate supersense alternation group of supersense alternation, as with *nounours* 'teddy bear' in the Artifact-Person alternation.

3.2.3. Focusing on regular metonymical and metaphorical relations

The next step is to determine whether two meanings of a given noun can be considered directly related through a regular metonymic or metaphorical link. First, we must establish whether a given sense alternation represents a case of metonymy, metaphor, or another type of relationship. Then, for metonymies and metaphors, we need to assess whether they follow a regular pattern. In line with

Apresjan (1974), we consider a sense extension to be regular if it is instantiated by at least two polysemous words. For example, the metonymic shift from a noun denoting an attribute to a noun denoting a person characterized by that attribute is regular, since it applies, among others, to *nullité* 'worthlessness / a nobody' and *laideur* 'ugliness / unattractive person'.

Since our analysis focuses on pairs of senses involving a supersense shift, the data should not include cases of specialization or generalization. The four possible values describing the semantic link between meanings A and B are as follows:

1. Homonymy: the two meanings are not semantically related. For example, the two meanings of the noun *palais* 'palate / palace' have no semantic connection.

> (6) *palais* 'palate / palace'
> a. Paroi supérieure qui sépare la fosse nasale de la bouche. → Body
>   (Upper wall / roof that separates the nasal cavity from the mouth)
> b. Vaste demeure urbaine d'un souverain, d'un prince, d'un chef d'État […].
>   → Artifact
>   (A large urban residence of a sovereign, a prince, a head of state […])

2. Metaphor: the two meanings are related through a metaphoric extension, where the referent of B can be seen as similar to the referent of A. In (7), the target meaning of *girouette* refers to a person whose behaviour resembles that of a weather vane.

> (7) *girouette* 'weather vane / fickle person'
> a. Pièce de fer-blanc ou d'autre matière mince et légère, ordinairement en forme de flèche empennée ou de coq, qu'on pose sur un pivot en un lieu élevé, de manière qu'elle tourne au gré du vent et qu'elle en indique la direction. → Artifact
>   (A piece of tinplate or another thin and lightweight material, usually in the shape of a feathered arrow or a rooster, placed on a pivot in a high location so that it turns with the wind and indicates its direction.)
> b. Personne qui change souvent d'avis, de sentiment, de parti. → Person
>   (A person who often changes their mind, feelings, or allegiance.)

3. Metonymy: The two meanings are related through a metonymic extension, where the referent of B is in a contiguity relationship with the referent of A. For example, *taxi* can refer to a type of vehicle or to the person who drives that vehicle.

> (8) *taxi* 'taxi / taxi driver'
>   a. Véhicule automobile terrestre privé, conduit par un chauffeur, destiné au transport payant de passagers et de leurs bagages, de porte à porte. → Artifact
>   (A private land motor vehicle, driven by a chauffeur, intended for the paid transportation of passengers and their luggage, door to door.)
>   b. Chauffeur de taxi. → Person
>   (Taxi driver)

The examples in (7) and (8) are prototypical instances of metaphor and metonymy respectively. However, it is common for a meaning relationship to involve both figures (e.g., Goossens, 2003). For example, *araignée* 'spider / spider strainer' refers to a kitchen tool that is not similar to the animal itself, but rather to the web the spider constructs. Cases where an analogy is built upon a contiguity relationship are classified as metaphor. Conversely, cases where a contiguity extension is built upon an analogy relationship are classified as metonymy.

4. Undecidable: There is a relationship between the two meanings, unlike in cases of homonymy, but this relationship is undecidable. Such indeterminacy typically arises in cases of morphological co-derivation, where meanings A and B, both resulting from a polyfunctional morphological process, interact with semantic extensions (e.g., Salvadori, 2024). For instance, the French suffix *-ier* can derive nouns denoting both persons (e.g., *rentier* 'male annuitant', *aventurière* 'female explorer') and artifacts (*dentier* 'dentures', *gouttière* 'gutter'). This polyfunctionality often applies to a single formal base, leading to nominal ambiguity. A clear example is *cuisinière* 'cook / kitchen stove', which can refer to both a person and an artifact. The two senses of *cuisinière* are likely derived in parallel from the same root (*cuisine* 'kitchen'). However, theoretically, one could posit a contiguity relationship (the kitchen stove is what the cook typically uses to cook) or an analogy relationship (the kitchen stove has the same function as the cooker), both of which are attested for simplex nouns. Cases like this are classified as undecidable.

(9) *cuisinière* 'cook / kitchen stove'
  a. Celle qui cuisine, qui prépare, qui cuit la nourriture → Person
     (A woman who cooks, prepares, and bakes food)
  b. [Électroménager] Fourneau de cuisine servant à chauffer ou faire cuire les aliments, souvent muni d'éléments chauffants sur sa surface de travail. → Artifact
     ([Appliance] A kitchen stove used for heating or cooking food, often equipped with heating elements on its surface)

4. Results

The analysis protocol described in the previous section was manually applied to a reduced set of 60 supersense alternations, each associated with at least 20 candidate nouns (mean = 50), resulting in a total of 2,960 instances of supersense alternations to be analysed. The results of this analysis are displayed in Table 4.

Table 4.
Distribution of the 2,992 analysed instances of supersense alternations

Discarded supersense alternations	Because of unknown meaning	1,327
	Because of the absence of supersense shift	86
Classified supersense alternations	Homonymy	97
	Metaphor	638
	Metonymy	728
	Undecidable	84
		2,960

4.1. Discarded sense alternations

During the first step of the analysis, 1,327 instances of supersense alternation (45%) were discarded because at least one sense in the pair was unknown to the author (cf. Section 3.2.1). Table 5 lists the 60 supersense alternations along with the percentage of instances discarded due to unknown meanings. This high proportion of rare meanings can be attributed to Wiktionary's extensive coverage and its collaborative nature, where some contributors may focus on specialized domains. For instance, the dataset includes 142 senses related to different types

of charges on a heraldic shield. The Animal-Cognition supersense alternation, which has the highest rate of discarded instances due to unknown target meanings (98%), primarily involves such specialized senses, as illustrated in (10) with *crapaud* (toad).

(10) *crapaud* 'toad'
   a. Batracien qui ressemble à la grenouille, aux pattes plus courtes, aux mœurs plus terrestres et dont le corps est couvert de pustules venimeuses. → Animal
   (A batrachian resembling a frog, with shorter legs, more terrestrial habits, and a body covered in venomous pustules)
   b. [héraldique] Meuble représentant l'animal du même nom dans les armoiries. Il est généralement représenté dans sa forme naturelle (non stylisée). → Cognition
   (A heraldic charge representing the animal of the same name. It is usually depicted in its natural (non-stylized) form)

Table 5 shows that supersense alternations highly vary on that respect. For instance, only 18% of the Animal-Person sense alternations (first column, third line) were filtered out. By comparison, 70% of the Person-Animal supersense alternation (third column, 13th line) were discarded, primarily because many of the animals belong to unfamiliar sub-species, such as *républicain* (republican), which refers to a species of passerine birds that live in communities in large suspended nests.

Table 5.

Percentage of discarded instances of supersense alternations due to unknown meanings

Plant-Food	11	Feeling-Act	35	Person-Artifact	48
Act-Feeling	17	Food-Act	36	State-Act	48
Animal-Person	18	Institution-Artifact	36	Body-Person	50
Act-Institution	21	Act-Cognition	37	Artifact-Cognition	51
Act-Artifact*Cognition	22	Artifact-Quantity	37	Object-Artifact	53
Animal-Food	25	Plant-Artifact	39	Person-Cognition	55
Cognition-Person	26	Attribute-Artifact	39	Body-Cognition	55
Act-Possession	27	Act-GroupxPerson	40	Artifact-Body	56
Food-Person	27	Artifact-Act	42	Artifact-Object	56

Table 5 (Continuation)

Attribute-Act	28	Artifact-State	42	Animal-Artifact	59
Artifact-GroupxPerson	29	Quantity-Artifact	42	Act-Object	64
Cognition-Act	30	Act-Person	42	Act-Attribute	65
Attribute-Person	31	Body-Artifact	43	Person-Animal	70
Artifact-Possession	31	Cognition-Artifact	43	Food-Artifact	70
Artifact-Person	32	Substance-Artifact	43	Person-Food	70
Attribute-Cognition	32	Object-Cognition	43	Body-Plant	76
Artifact-Attribute	32	Artifact-Food	44	Person-State	78
Act-Artifact	33	Animal-Substance	45	Artifact-Plant	82
Artifact-Institution	34	Act-State	47	Artifact-Animal	93
Cognition-Institution	35	Food-Body	47	Animal-Cognition	98

Additionally, 86 instances of supersense alternations (3%) were discarded due to semantic classification errors. After correction, the pair of meanings no longer imply a supersense shift (cf. example (4)).

4.2. Classified sense alternations

The analysis of the remaining 1,547 supersense alternations is presented in the lower part of Table 4. Among these alternations, 5% are instances of homonymy, 41% corresponds to metaphors, 47% to metonymies and 5% were classified as undecidable. Of the 1,366 pairs involving metaphor or metonymy, 78% were found to follow a regular pattern as defined in the previous section. This result provides a first assessment of the prevalence of regular polysemy among cases of (raw) sense alternations in a representative set of the lexicon.

Tables 6 and 7 present the full list of metonymic and metaphoric polysemy patterns that were extracted from the data, along with the number of their attestations (n) and the number of words that could theoretically instance each pattern (N). For example, in Table 6 (line 1), of the 138 nouns of the dataset that primarily refer to an animal, 59 exhibit a metaphorical meaning referring to a person. The regularity score of the pattern (Reg) was calculated using the formula provided in section 2.3, with word frequency estimates from *Lexique 3* (New et al., 2004).

Table 6.

Metaphorical patterns extracted from the dataset, ranked in descending order of regularity score. Column "n" indicates the number of words that enter the pattern A–B. Column "N" indicates the number of words whose first meaning is of type A.

Metaphorical patterns	n	N	Reg	Example
1. Animal-Person	59	138	0.48	*mouton* 'sheep'
2. Body-Artifact	25	73	0.39	*pied* 'foot / (table) leg'
3. Plant-Artifact	12	50	0.37	*feuille* 'leaf / sheet of paper'
4. Institution-Artifact	3	20	0.21	*prison* 'prison / cage'
5. Food-Artifact	12	78	0.20	*banane* 'banana / bum bag'
6. Food-Body	11	78	0.19	*miche* 'bun / cheek or boob'
7. Animal-Artifact	19	138	0.18	*souris* 'mouse'
8. Food-Person	12	78	0.18	*légume* 'vegetable'
9. Artifact-Cognition	63	504	0.15	*bagage* 'luggage / background'
10. Food-Act	8	78	0.14	*pêche* 'peach / punch'
11. Person-Artifact	29	247	0.10	*explorateur* 'explorer'
12. Person-Animal	15	247	0.09	*ouvrière* 'worker / worker bee'
13. Artifact-Person	36	504	0.07	*bulldozer* 'bulldozer'
14. Artifact-Body	32	504	0.07	*disque* 'disk'
15. Artifact-Object	28	504	0.07	*cuvette* 'basin'
16. Object-Artifact	5	65	0.07	*bloc* 'block'
17. Body-Plant	4	73	0.07	*pied* 'foot / (vine) plant'
18. Cognition-Artifact	7	137	0.06	*trapèze* 'trapezoid / trapeze'
19. Plant-Food	3	50	0.06	*tomate* 'tomato plant / Pastis&grenadine'
20. Artifact-Act	24	504	0.05	*planche* 'board / floating on your back'
21. Object-Cognition	2	65	0.05	*croûte* 'scab / tacky painting'
22. Body-Person	3	73	0.05	*membre* 'limb / member'
23. Artifact-Food	15	504	0.04	*chausson* 'slipper / (apple) turnover'
24. Act-Feeling	9	400	0.03	*torture* 'torture / agony'
25. Cognition-Act	3	137	0.02	*épopée* 'epic / epic journey'
26. Animal-Food	2	138	0.02	*souris* 'mouse / lamb chank'
27. Body-Cognition	2	73	0.02	*ossature* 'skeleton / structure'
28. Artifact-Plant	6	504	0.01	*aiguille* 'needle'

Table 6 (Continuation)

Metaphorical patterns	n	N	Reg	Example
29. Artifact-Attribute	3	504	0.01	*façade* 'front / appearance'
30. Artifact-Quantity	5	504	0.01	*plafond* 'ceiling'
31. Artifact-GroupxPerson	4	504	0.01	*brochette* 'skewer / bunch'
32. Artifact-State	4	504	0.01	*impasse* 'dead end / deadlock'
33. Person-Food	2	247	0.01	*religieuse* 'nun / religieuse'
Mean	**14**	**249**	**0.1**	

Table 7.

Metonymic patterns extracted from the dataset, ranked in descending order of regularity score. Column "n" indicates the number of words that enter the pattern A–B. Column "N" indicates the number of words whose first meaning is of type A.

Metonymy patterns	n	N	Reg	Example
1. Feeling-Act	12	20	0.71	*plaisir* 'pleasure'
2. Substance-Artifact	20	46	0.43	*verre* 'glass'
3. Act-Artifact	109	400	0.27	*armement* 'arming / weapons'
4. Plant-Food	17	50	0.25	*tomate* 'tomato tree / tomato'
5. Quantity-Artifact	5	21	0.22	*décimètre* 'decimeter / 10-centimeter ruler'
6. Animal-Food	19	138	0.18	*canard* 'duck'
7. Attribute-Act	26	118	0.17	*politesse* 'courtesy'
8. Institution-Artifact	5	20	0.17	*menuiserie* 'carpentry workshop / woodwork'
9. Act-Cognition	49	400	0.16	*essai* 'attempt / essay'
10. State-Act	5	27	0.15	*équilibre* 'balance / handstand'
11. Body-Artifact	7	73	0.13	*jambe* 'leg / pantleg'
12. Body-Person	6	73	0.13	*cerveau* 'brain'
13. Cognition-Act	15	137	0.10	*disco* 'disco music / dance'
14. Act-Institution	35	400	0.10	*administration* 'administration / administrators'
15. Attribute-Cognition	12	118	0.10	*absurdité* 'absurdity / nonsense'

Table 7 (Continuation)

Metonymy patterns	n	N	Reg	Example
16. Attribute-Person	11	118	0.10	*laideur* 'ugliness / unattractive person'
17. Animal-Substance	11	138	0.10	*crocodile* 'crocodile / crocodile leather'
18. Plant-Artifact	4	50	0.09	*gazon* 'grass / lawn'
19. Food-Act	6	78	0.09	*méchoui* 'mechoui'
20. Artifact-Act	39	504	0.08	*barre* '(high) bar'
21. Act-Artifact*Cognition	27	400	0.07	*convocation* 'summons / summons letter'
22. Attribute-Artifact	7	118	0.06	*résistance* 'resistance / resistor'
23. Object-Artifact	2	65	0.05	*terrain* 'ground / building plot'
24. Act-Possession	19	400	0.05	*emprunt* 'loan'
25. Act-GroupxPerson	13	400	0.05	*défense* 'defense / defenders'
26. Artifact-Person	18	504	0.04	*violon* 'violin / violinist'
27. Artifact-Institution	13	504	0.04	*bibliothèque* 'bookshelf / library'
28. Artifact-GroupxPerson	11	504	0.04	*cabinet* 'office / cabinet'
29. Person-Artifact	7	247	0.03	*ballerine* 'ballerina / ballet slippers'
30. Act-State	19	400	0.03	*isolement* 'isolation'
31. Person-State	4	247	0.03	*gendarme* 'military police officer / status'
32. Act-Person	8	400	0.03	*apparition* 'appearance / apparition'
33. Act-Object	8	400	0.02	*projection* 'projection / ejection'
34. Act-Feeling	6	400	0.02	*excitation* 'excitement'
35. Cognition-Institution	5	137	0.02	*pharmacie* 'pharmacy'
36. Artifact-Cognition	3	504	0.01	*carillon* 'carillon / chime'
37. Artifact-Body	3	504	0.01	*maillot* 'swimsuit / pubic area'
38. Artifact-Object	4	504	0.01	*fond* 'bottom / seabed'
39. Act-Attribute	6	400	0.01	*séduction* 'seduction / appeal'
40. Cognition-Artifact	2	137	0.01	*mécanique* 'mecanics / mechanism'
41. Artifact-Food	4	504	0.01	*raclette* 'scraper / raclette'
42. Artifact-Quantity	6	504	0.01	*cuillère* 'spoon / spoonful'
43. Artifact-Possession	4	504	0.01	*enveloppe* 'envelope / cash envelope'
Mean	**14**	**270**	**0.1**	

First, considering the number of supersense alternations associated with polysemy patterns, we observe that metonymies (43 supersense alternations) occur more frequently than metaphors (33 supersense alternations), thus empirically supporting the claim that metonymy is more characteristic of regular polysemy than metaphor. However, when examining the regularity scores, the results reveal no significant difference in the mean scores between metonymy and metaphor (0.1 in both cases). This unexpected outcome warrants at least two important observations.

The first one concerns the semantic granularity used to describe regular polysemy. Supersenses were selected because they have been used to describe large sets of lexical data across several languages. However, supersenses are not uniform in terms of semantic granularity: some correspond to broader semantic classes (e.g., Act, Artifact), while others are more specific (e.g., Animal, Feeling). As a result, some supersense alternations encompass multiple patterns, while others do not. For instance, the Act → Artifact metonymy is linked to several patterns, depending on the role of the referent of the artifact sense within the situation denoted by the act sense. These patterns include:

- Act → Artifact, instrument of the action
  (e.g., *équipement* 'equipping/equipment')
- Act → Artifact, location where the action takes place
  (e.g., *arrêt* '*stopping*/stop')
- Act → Artifact, result of the action
  (e.g., *tatouage* 'tattoing/tattoo')

This fixed semantic granularity to define the list of polysemy patterns in the dataset has a clear impact on their estimated degree of regularity, which will need further refinement in future analyses. For instance, the Artifact-Quantity metonymy (Table 7, line 42) appears to have a low degree of regularity (0.01) because it is evaluated across a broad set of 504 nouns primarily referring to manufactured objects. However, the observed regular sense extension is more specifically associated with the Container → Quantity pattern (e.g., *cuillère* 'spoon / spoonful', *seau* 'bucket / bucketful'). In contrast, metaphors appear to involve less specific source meanings.

The second observation concerns the properties of the target meaning in metonymy compared to metaphor. In metonymy, target meanings tend to be generally more frequent and less marked than in metaphor, where a significant proportion of meanings created through analogical extension belong to familiar language. This aspect will also require further exploration in future analyses.

Despite these two observations suggesting that the method used to calculate regularity scores may underestimate the regularity of metonymy patterns compared to metaphor ones, it remains true that, on average, regular metonymy does not tend to produce significantly more polysemy than regular metaphor. This result should at least prompt us to reassess with greater caution the assumed strong correlation between lexical figure and polysemy regularity.

Conclusion

This paper provides a first quantitative assessment of how much more regular metonymy is compared to metaphor. The study is based on a carefully defined, representative sample of nominal sense alternations in French. The results show that metonymy is significantly more regular than metaphor in terms of the number of alternation patterns observed in the analysed sample but that, on average, these patterns produce comparable proportions of polysemic words. The greater propensity of metonymy for regularity is thus nuanced by these results, which call for a measured evaluation of the correlation between figure and degree of regularity. We identify several promising avenues for further investigation to enhance the assessment of regular polysemy. From a methodological perspective, we have highlighted the limitation of relying on judgments from a single annotator. The next step is to collect multiple annotations to measure inter-annotator agreement—both on word sense familiarity, to determine which sense alternations are perceived as natural by speakers, and on identifying the underlying figure behind these alternations, to assess how easily speakers recognize semantic relationships between known senses of the same word. Furthermore, because supersenses provide a level of granularity that is often too coarse, they do not allow for a sufficiently precise analysis of polysemy patterns. Building on the descriptions provided in this study, a more fine-grained analysis will therefore be developed in future work to obtain more precise regularity scores. Finally, we have highlighted the complex relationship between polysemy and morphological derivation, leaving open the question of whether stronger regularities stem from this connection.

References

Apresjan, Ju. (1974). Regular polysemy. *Linguistics*, *14*, 5–32.

Angleraud, N., Barque, L., & Candito, M. (2025). Annotating the French Wiktionary with supersenses for large scale lexical analysis: a use case to assess form-meaning relationships within the nominal lexicon. In *Proceedings of the 31st International Conference on Computational Linguistics* (pp. 5321–5332). https://aclanthology.org/2025.coling-main.356/.

Barque, L., Haas, P., & Huyghe, R. (2018). Polysémie régulière et néologie sémantique: constitution d'une ressource pour l'étude des sens nouveaux. *Neologica: revue internationale de la néologie*, *12*, 91–108. https://hal.science/hal-03823528v2/file/Neologica_polyReg.pdf.

Buitelaar, P. (1998). *CoreLex: systematic polysemy and underspecification*. Phd Thesis, Brandeis University, Waltham.

Buitelaar, P. (2000). Reducing lexical semantic complexity with systematic polysemous classes and underspecification. In *NAACL-ANLP 2000 Workshop: Syntactic and Semantic Complexity in Natural Language Processing Systems*, *Vol. 1* (pp. 14–19). https://aclanthology.org/W00-0103.pdf.

Copestake, A., & Briscoe, T. (1995). Semi-productive polysemy and sense extension. *Journal of semantics*, *12*(1), 15-67. https://www.researchgate.net/profile/Ted-Briscoe/publication/2609962_Semi-Productive_Polysemy_and_Sense_Extension/links/0912f51370d570a560000000/Semi-Productive-Polysemy-and-Sense-Extension.pdf?_sg%5B0%5D=started_experiment_milestone&origin=journalDetail&_rtd=e30%3D.

Dölling, J. (2020). Systematic polysemy. In L. Matthewson, C. Meier, H. Rullmann, T. E. Zimmermann & D. Gutzmann (Eds), *The Wiley Blackwell Companion to Semantics*. Wiley Online Library, Hoboken.

Falkum, I., & Vicente, A. (2015). Polysemy: Current perspectives and approaches. *Lingua*, 157, 1–16. https://philarchive.org/archive/FALPCP.

Fellbaum, C. (Ed.). (1998). *WordNet: An electronic lexical database*. MIT press, Cambridge.

Goossens, L. (2003) 'Metaphtonymy: The interaction of metaphor and metonymy in expressions for linguistic action'. In R. Dirven & R. Pörings (Eds), *Metaphor and metonymy in comparison and contrast*. (pp. 349–377). Mouton de Gruyter, Berlin, New York.

Klepousniotou, E., Titone, D., & Romero, C. (2008). Making sense of word senses: the comprehension of polysemy depends on sense overlap. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1534–1543.

Lombard, A., Huyghe, R., Barque, L., & Gras, D. (2023). Regular polysemy and novel word-sense identification. *The Mental Lexicon*, *18*(1), 94–119.

Lopukhina, A., Laurinavichyute, A., Lopukhin, K., & Dragoy, O. (2018). The mental representation of polysemy across word classes. *Frontiers in psychology*, *9*, 192.

Maudslay, R. H., Teufel, S., Bond, F., & Pustejovsky, J. (2024). ChainNet: Structured Metaphor and Metonymy. In WordNet. *Proceedings of LREC-COLING 2024* (pp. 2984–2996). ELRA and ICCL, Torino. https://arxiv.org/pdf/2403.20308.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2 : A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 516–524. https://link.springer.com/content/pdf/10.3758/BF03195598.pdf.

Nunberg, G., & Zaenen, A. (1997). La polysémie systématique dans la description lexicale. *Langue française*, 12–23.

Peters, W. (2006). In Search for More Knowledge: Regular Polysemy and Knowledge Acquisition. In P. Sojka, K.-S. Choi, C. Fellbaum & P. Vossen (Eds.), *Proceedings of GWC 2006* (pp. 245–251). Masaryk University, Brno.

Polguère, A. (2018). A lexicographic approach to the study of copolysemy relations. *Russian Journal of Linguistics*, *22*(4), 788–820.

Polguère, A. (2022). Fishing out regular polysemy with the Gener (generic) relation. *Lexique*, 27–54. https://www.peren-revues.fr/lexique/856.

Pustejovsky, J. (1995). *The Generative Lexicon*, MIT Press, Cambridge.

Salvadori, J. (2024). *L'ambiguïté des noms déverbaux en français. Une étude quantitative du sens construit*. Phd Thesis, University of Fribourg, Fribourg.

Srinivasan, M., & Rabagliati, H. (2015). How concepts and conventions structure the lexicon: Cross-linguistic evidence from polysemy. *Lingua*, *157*, 124–152. https://escholarship.org/content/qt69z8f7qf/qt69z8f7qf_noSplash_426afa3377f6db5afe571954392fa2ff.pdf.

Vicente, A., & Falkum, I. L. (2017). Polysemy. In: M. Aronoff (Ed.), *Oxford Research Encyclopaedia of Linguistics*. Oxford University Press, Oxford.

Xu, Y., Duong, K., Malt, B. C., Jiang, S., & Srinivasan, M. (2020). Conceptual relations predict colexification across languages. *Cognition*, 201, 104–280.

Yurchenko, A., Lopukhina, A., & Dragoy, O. (2020). Metaphor is between metonymy and homonymy: Evidence from event-related potentials. *Frontiers in Psychology*, *11*, 2113.