

Metonymy is more multilingual than metaphor: Analysing tropes using ChainNet and the Open Multilingual Wordnet

Francis Bond 

Dept. of Asian Studies; Sinofon Project
Palacký University
bond@ieee.org

Rowan Hall Maudslay

Dept. of Computer Science & Tech.
University of Cambridge
rh635@cam.ac.uk

Abstract

The senses of a word are often systematically related to each other, either by metaphor or metonymy. Because speakers of different languages share the same basic cognitive common ground, it is possible that the same metaphors and metonyms will appear across different languages. In this paper, we investigate the extent to which English metaphors, metonyms, and homonyms are evidenced across different languages. To achieve this analysis we use ChainNet and the Open Multilingual Wordnet. ChainNet provides detailed annotations of figurative sense relations such as metaphor and metonymy in English, while the Open Multilingual Wordnet aligns multilingual synsets across more than 30 languages. We find that metonyms are more universal than metaphors, and that both metaphors and metonyms are much more universal than homonyms. Further work is needed to determine which metaphors or metonyms are more universal than others.

1 Introduction

Words exhibit multiple senses that are semantically related. This phenomenon is known as **polysemy**. Polysemy can be decomposed into two distinct categories: metaphor and metonymy (Jakobson, 1956). A **metaphor** is a language usage which frames one thing in terms of another which is analogically similar. As an example, consider these two usages of the word *death*:

- (1) a. News of her *death* moved him deeply.
b. The *death* of Rome was now inevitable.

In (1a), the word *death* is used literally to refer to the end of a life, while in (1b) it is used metaphorically to refer to the end of an empire's hegemony.

A **metonym**, on the other hand, is a language usage in which a word stands in for another meaning based on some understood association or contiguity. Senses which are metonymically related are often very similar, but are of different semantic types

(Pustejovsky, 1995). For example, consider these two usages of the word *apple*:

- (2) a. The *apple* was delicious.
b. He watered the *apple* in the garden.

In (2a) the word *apple* refers to a fruit, while in (2b) the word *apple* refers to a tree which bears this fruit. These senses are metonymically related because they refer to different semantic types (fruit and tree), but they share a close association.

Metaphor and metonymy are collectively known as **tropes**. Tropes are productive processes, and it is therefore possible for the same tropes to occur in different languages. For example, the metaphor in example (1) also occurs in Catalan (with the word *mort*) and in Slovene (*smrt*), and the metonym in example (2) also occurs in French (*pomme*) and in Chinese (苹果).

In contrast to polysemy, words can also exhibit senses which are semantically unrelated. This is known as **homonymy**. Consider these two usages of the word *bat*:

- (3) a. I hit the ball with my *bat*.
b. I could hear a *bat* in the darkness.

In (3a), *bat* refers to a wooden implement used in ball games, while in (3b) *bat* refers to a nocturnal mammal that uses echolocation. In this case, homonymy has arisen because these senses have distinct etymological origins: the sense of *bat* in (3a) comes from Old French, while the sense of *bat* in (3b) has a Scandinavian origin. Homonymy is widely thought to be a coincidence of a language's development, and it is therefore unlikely that a particular homonym will appear in different languages which have developed independently. Indeed, the two senses of *bat* in example (3) are realised using different wordforms in many languages, including Chinese, French, Croatian, Indonesian, and so on.

In this paper, we investigate to which extent the metaphors, metonyms, and homonyms in English

appear in other languages. Because metaphor and metonym are productive, it is impossible to enumerate every metaphor or metonym that exists in a language (e.g. Black, 1962, 1977). For this reason, we choose to focus specifically on conventionalised metaphors and metonyms. These are the metaphors and metonyms that are widely used by a language community, including those in examples (1) and (2). To analyse whether conventional metaphors and metonyms in English appear across languages, we exploit two lexical resources. The first resource is the Open Multilingual Wordnet (OMW: Bond and Foster, 2013), which is a multilingual lexicon in which different languages are aligned using the same inventory of word senses. The second resource is ChainNet (Maudslay et al., 2024), which identifies metaphors, metonyms, and homonyms in the English component of the OMW.

The remainder of this paper is organized as follows. In Section 2, we describe the lexical resources used in our study. In Section 3, we detail the methodology we employed to analyse the presence of metaphors, metonyms, and homonyms across languages. We present the numerical results, highlighting the stronger cross-linguistic presence of metonymy compared to metaphor. In Section 4, we look at how we can use the synonyms and translations to further specify the metaphor and metonymy links. In Section 5, we discuss the implications of our findings, including challenges related to wordnet construction and the potential role of large language models (LLMs) in future studies. Finally, in Section 6, we summarise our conclusions and propose directions for future research, particularly regarding the application of our findings to sense-tagged corpora and the systematic annotation of figurative tropes across languages.

2 Lexical resources

Effective cross-linguistic analysis depends on reliable lexical resources. In this section, we detail the two key resources which form the backbone of our study: ChainNet and the Open Multilingual Wordnet. We additionally briefly introduce UniMet and the Italian Metaphor Database, which are existing resources that are closely-related to our work.

2.1 The Open Multilingual Wordnet

A **wordnet** (Miller, 1995) is a type of lexicon with two special properties. The first is that multiple dif-

ferent wordforms can be associated with the same concept, if those wordforms are synonymous. For example, in the Princeton English WordNet (PWN: Fellbaum, 1998) the words *dessert*, *sweet*, and *afters* are all associated with the concept glossed by “a dish served as the last course of a meal”. For this reason, in a wordnet a concept is known as a **synset** (synonym set). The second property is that synsets and senses in a wordnet are linked to each other by different semantic and lexical relations. For example, in PWN the synset given above is connected to the synset which has the gloss, “part of a meal served at one time”, by the “is a” relation.

The OMW is a collection of wordnets for different languages. These languages are linked through the collaborative interlingual index (CILI; Bond et al., 2016), which is a language-neutral list of concepts. The combined wordnets include English (Fellbaum, 1998), Albanian (Ruci, 2008), Arabic (Elkateb et al., 2006), Chinese (Huang et al., 2010), Danish (Pedersen et al., 2009), Finnish (Lindén and Carlson., 2010), French (Sagot and Fišer, 2008), Hebrew (Ordan and Winter, 2007), Indonesian and Malaysian (Nuril Hiranfa et al., 2011), Italian (Pianta et al., 2002; Toral et al., 2010), Japanese (Isahara et al., 2008), Norwegian Bokmål and Norwegian Nynorsk (Lars Nygaard, personal communication 2012), Persian (Montazery and Faili, 2010), Portuguese (de Paiva and Rademaker, 2012), Polish (Piasecki et al., 2009), Thai (Thoongsup et al., 2009), and Basque, Catalan, Galician and Spanish (Gonzalez-Agirre et al., 2012). We used version 1.4 of the OMW from <https://github.com/omwn/omw-data>, accessed through the python **wn** module (Goodman and Bond, 2021).

2.2 ChainNet

The OMW identifies the senses of words in different languages, but it does not identify if and how these senses are related. Consider the nominal senses of the English word *tear* in the Open English Wordnet (OEWn: McCrae et al., 2019):

- tear*₁ [*teardrop*] a drop of the clear salty saline solution secreted by the lacrimal glands, e.g. “his story brought tears to her eyes”
- tear*₂ [*rip, rent, snag, split*] an opening made forcibly as by pulling apart
- tear*₃ [*bust, binge, bout*] an occasion for excessive eating or drinking
- tear*₄ the act of tearing, e.g. “he took the manuscript in both hands and gave it a mighty tear”

The senses of *tear* exhibit metaphor, metonymy, and homonymy, but this structure is not identified in the OEWN. ChainNet addresses this by identifying how senses are related to one another. In ChainNet, every nominal sense is either a prototype, or is linked to another sense by metaphor or metonymy. Homonyms are implicitly those senses which are not connected to each other, directly or indirectly.

An example of ChainNet annotation for the word *tear* is shown in Figure 1. The sense *tear*₂ (a hole made by ripping) is a prototypical sense, which is extended by metonymy to *tear*₄ (the act of ripping) and by metaphor to *tear*₃ (a binge). The sense *tear*₁ is a separate prototype, which is disconnected from the other senses, indicating that it is a homonym. ChainNet is also annotated with “feature transformations” (not shown here), where specific characteristics of a sense are either retained, lost, or altered when the sense is extended through metaphor; we do not use these feature transformations in this paper.

ChainNet was created using manual annotation. Of the 15,234 polysemous nouns in the OEWN, 6,500 were annotated, for a total coverage of 22,178 senses. By far the majority of words with more than three senses have been annotated (>90%). This makes ChainNet the first dataset to systematically capture inter-sense relations at scale. The data and tagging guidelines are available at <https://github.com/rowanhm/ChainNet>.

2.3 UniMet: Universal Metonymy

Khishigsuren et al. (2022) have created a resource with metonymy tropes using wordnet concepts from the Universal Knowledge Core (UKC: Giunchiglia et al., 2017, 2018, 2023), another wordnet-like multilingual lexicon. Khishigsuren et al. identified metonyms in a top down manner. More specifically, they first identified 26 metonymy patterns based on pairs of UKC domains, such as **body part** and **person**. These domain pairs were then used to extract synset pairs, such as “the upper part of the human body” → “a person who is in charge”. A total of 51,000 candidate synset pairs were filtered to 4,900 pairs. Every possible lexicalisation of these pairs from the UKC was extracting, yielding a total of 20,000 instances of metonymy from 189 languages. All parts of speech were included, and therefore the cases of metonymy in this data include pairs of senses of different parts of speech (e.g. noun–verb pairs).

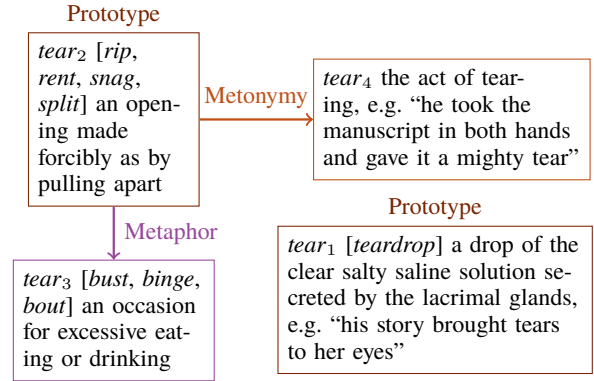


Figure 1: Complete annotation for *tear* in ChainNet

Cases are also included that consist of different wordforms: Khishigsuren et al. distinguish between what they call “morphological metonymy”, in which a pair consists of different wordforms that are morphologically related (such as French *garde–garder*) and “lexicalised metonymy”, which are pairs in which both wordforms are the same.

Considering only lexicalized metonymy between nouns in English, there are 3,298 pairs in Khishigsuren et al.’s data. By contrast, there are 6,116 metonymy pairs in ChainNet, and 7,521 metaphors. Only 265 pairs are shared by both resources, although for these 96 are in different directions in the two resources.

There are 20 metonymy pairs in UniMet which are classified as examples of metaphor in ChainNet. Of these, 9/20 are cases of generalisation, for example a shift from the definition of the word *kale* as “a hardy cabbage with coarse curly leaves that do not form a head” → any “coarse curly-leaved cabbage”, or equivalently the shift of *enamel* from “a colored glassy compound fused to the surface of metal or glass or pottery” → “any smooth glossy coating”. Generalisation was considered a subtype of metaphor in the ChainNet annotation guidelines, and is not commonly considered to be a type of metonymy. Of the remaining 11 cases, six appear to be mistakes with UniMet (i.e. clear cases of metaphor) while five appear to be mistakes with ChainNet (i.e. clear cases of metonymy). Examples of mistakes with UniMet include a pair of definitions for *muscle*, “one of the contractile organs of the body” → “a bully employed as a thug or bodyguard”, as well as a pair of definitions for *mouth*, “the opening through which food is taken in and vocalizations emerge” → “a spokesperson” (the latter being evoked in sentences such as “the *mouth* of the organisation”). Of the five mistakes

in ChainNet, one is an example of specialisation (*essence*: “any substance possessing to a high degree the predominant properties of a plant or drug or other natural product from which it is extracted” → “a toiletry that emits and diffuses a fragrant odor”), three are for artefacts made from a specific material (e.g. *brass*: “an alloy of copper and zinc” → “a memorial made of brass”), and one is a plant–food alternation (*broccoli*: “plant with dense clusters of tight green flower buds” → “branched green undeveloped flower heads”).

2.4 The Italian Metaphor Database

Alonge (2006) described the construction of a database of Italian Metaphors built through a study of a corpus and linked to the Italian Wordnet, but we have not found any release of the data.

3 Analysing tropes across languages

Having outlined our primary lexical resources, we now investigate how English metaphors and metonymy are used cross-lingually. We start off by investigating how likely it is that two different senses of an English word (in the OEWn) will both be translated using a single word in different languages (in the OMW). We conduct this investigation for three different types of sense pairs, based on ChainNet annotation: those which are linked directly linked by metaphor, those which are directly linked by metonymy, and those which are not directly linked. Our hypothesis is that sense pairs which are linked are more often translated using the same word.

As an example, consider the English word *head*, which can refer to either a body part or metaphorically to a person in charge. This is also true in Japanese (頭 *atama*), Italian (*capo*), and many other languages. In English the word *head* is also used to count animals, as in the sentence “200 *head* of cattle”. This metonymic extension of *head* is also a possibility in Japanese (頭 *tou*) and Italian (*capo*). However, not every extension will be shared by all languages. For example, in English a third common extension of the word *head* is a metaphor to refer to the main part of a grammatical constituent, as in the sentence “the *head* of an NP is N”. This extension does not exist in Japanese, where the main part of a grammatical constituent is translated as 主要語 (*shuyougo* “main element”) or 主辞 (*shuji* “main appendant”). We show the relevant senses of *head* from ChainNet in Figure 2.

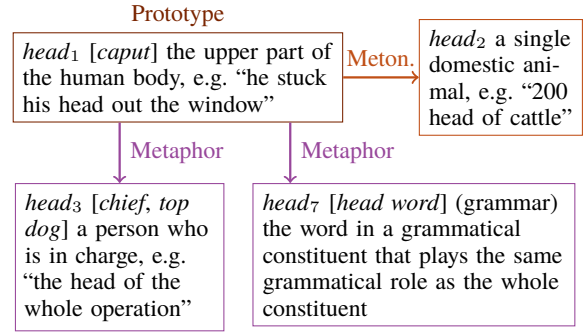


Figure 2: ChainNet-style annotation for *head*

When we evaluate whether tropes appear across different language, it is important to test across multiple wordnets produced by different groups using different methodologies. One of the reasons for this is because we need to consider the possibility that a wordnet has been constructed by naïvely translating an English wordnet. This could result in the situation where the senses of an English word also occurred together in another language, even if speakers of that language did not use these senses in that way.

To evaluate whether another language has the same tropes as English, we use the following procedure:

1. For every English word $w \in \mathcal{W}$ that has been annotated in ChainNet:

For each pair of senses s_1, s_2 from w :

- Look up all the translations of each sense, yielding two sets of words, $\mathcal{W}_1, \mathcal{W}_2 \subset \mathcal{W}$.
- Measure the overlap between $\mathcal{W}_1$ and $\mathcal{W}_2$ using the Jaccard index, which is defined as size of the intersection of two sets over the size of their union:

$$\text{jaccard}(\mathcal{W}_1, \mathcal{W}_2) = \frac{|\mathcal{W}_1 \cup \mathcal{W}_2|}{|\mathcal{W}_1 \cap \mathcal{W}_2|}$$

- Store the overlap and note if the senses are linked (and if so how they are linked).

2. Compute the average overlap considering only the pairs of senses which are unlinked, only the pairs sense which are linked by metaphor, and only the pairs of senses which are linked by metonymy.
3. Normalise the overlap scores by dividing them by the average overlap of all sense pairs.

Language	Code	Unlinked	Metaphor	Metonymy	All	Translated
Albanian	sq	0.74	1.27	2.38	0.009	822
Basque	eu	0.82	1.30	1.80	0.102	5,562
Bulgarian	bg	0.73	1.54	2.11	0.007	380
Catalan	ca	0.83	1.23	1.84	0.101	6,000
Croatian	hr	0.81	1.22	1.98	0.051	3,576
Danish	da	0.74	1.46	2.09	0.007	332
Dutch	nl	0.79	1.38	1.90	0.032	2,991
Finnish	fi	0.82	1.35	1.76	0.106	7,971
French	fr	0.94	1.10	1.25	0.294	18,975
Galician	gl	0.70	2.13	1.55	0.004	271
Greek	el	0.72	1.30	2.45	0.025	1274
Hebrew	he	0.69	1.58	2.31	0.011	472
Icelandic	is	0.73	1.51	2.13	0.005	388
Indonesian	id	0.87	1.21	1.58	0.123	9,803
Italian (IWN)	it	0.79	1.46	1.80	0.018	968
Italian (MWN)	it	0.80	1.39	1.81	0.065	4,614
Japanese	ja	0.89	1.11	1.55	0.088	8,130
Lithuanian	lt	0.63	1.46	2.86	0.010	692
Mandarin Chinese	cmn	0.74	1.57	1.97	0.016	1,347
Norwegian Bokmål	nb	0.73	1.49	2.12	0.008	339
Norwegian Nynorsk	nn	0.73	1.52	2.10	0.008	340
Polish	pl	0.75	1.43	2.07	0.024	1,326
Portuguese	pt	0.80	1.26	1.96	0.083	5,414
Romanian	ro	0.83	1.30	1.71	0.108	6,429
Slovak	sk	0.73	1.35	2.31	0.025	2,067
Slovenian	sl	0.81	1.41	1.74	0.111	6,243
Spanish	es	0.85	1.17	1.79	0.095	5,915
Standard Arabic	arb	0.74	1.28	2.38	0.016	1,337
Standard Malay	zsm	0.87	1.22	1.60	0.126	10,272
Swedish	sv	0.70	1.71	2.10	0.012	417
Thai	th	0.80	1.48	1.71	0.038	2,336
Mean		0.78	1.39	1.96	0.056	3,774.3

Table 1: Differences in the translation overlap by language

Results are shown in Table 1. For Italian, there are two results, which are computed from the Ita-WordNet (IWN: Toral et al., 2010) and the Multi-WordNet (MWN: Pianta et al., 2002) respectively. Our hypothesis clearly holds: over all languages, unlinked senses share the fewest translations (0.78 overlap), while metaphors share more (1.39) and metonyms share the most of all (1.96).

The wordnets in OMW are of vastly different sizes, so the number of sense pairs that have a translation (“Translated” in Table 1) varies from as few as 271 for Galician to as many as 18,975 for French. The average score of all those sense pairs that have a translation (“All” in Table 1)

is also wildly different, ranging from 0.004 for Galician to 0.294 for French. Only Galician has the score for metaphor (2.13) larger the score for metonymy (1.55), which we hypothesise is due to data sparsity, rather than some language specific property: we would expect it to behave much like Portuguese (which has a metaphor/metonym overlap of 1.26/1.96).

Several of the wordnets are made semi-automatically, by translating the English wordnet and then correcting the translations (French, Indonesian, Romanian, Slovenien, Standard Malay). In these wordnets, a high proportion of senses with translations were identified (all >0.1, while

Sense	Mandarin	Japanese	Finnish	Italian	Portuguese
<i>cherry</i> ₁ (wood)	櫻桃木	桜, 桜材	kirsikkapuumetsä	ciliegio	cerejeira
<i>cherry</i> ₂ (tree)	櫻桃树	桜, 櫻	kirsikkapuu	ciliegio	cerejeira
<i>cherry</i> ₃ (fruit)	櫻桃	櫻桃, 桜ん坊	kirsikka	cerasa , ciliegia	ginja , cereja
<i>cherry</i> ₄ (colour)	櫻桃红	桜ん坊色	kirsikanpunainen ²	ciliegia	cereja

Table 2: Translations of the senses of *cherry* in other languages

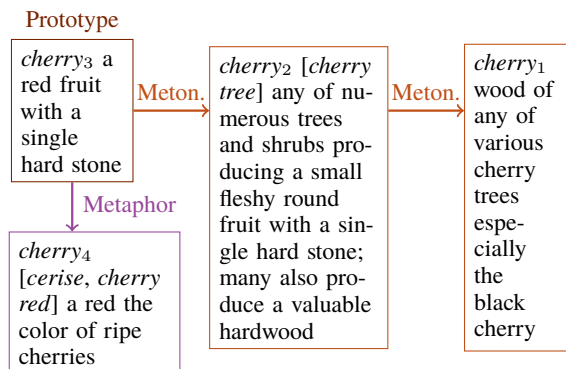


Figure 3: Complete annotation for *cherry* in ChainNet

the average is only 0.056), suggesting that the construction method that is used for a wordnet affects the overlap scores.

In order to measure perfectly how well tropes carry over between languages, we would need to mark metonymy and metaphor systematically for each language, and make sure all synsets have all relevant translations. Even so, it is striking how uniform our results are, even for very different resources and different languages. Without exception, senses linked by tropes are more likely to have an identical translation than those senses which are not linked. With one exception, metonymy is more likely to have an identical translation than metaphor. We therefore conclude that the same metonyms and metaphors appear in different languages, to varying degrees.

4 Using translations to specify tropes

Even when a pair of senses is not translated with the same word, it is often translated by a derivationally-related word (like morphological metonymy, §2.3). As an example, consider the four senses of the English word *cherry*, which are shown in Figure 3. The word *cherry* has a single prototype, *cherry*₃ (fruit). The sense is extended by metaphor for *cherry*₄ (colour), and by metonymy for *cherry*₂ (tree); the cherry tree sense is then itself extended by metonymy by *cherry*₁ (wood). Example translations of these senses are

shown in Table 2. We added some translations which were missing from the OMW; those that we added are shown underlined. In each of the shown languages, different wordforms are used for the different senses. However, there is some relation between these wordforms. In Mandarin, the word for the tree, wood, and colour is the same as the word for the fruit but with the word for tree, lumber, or colour attached to the end.¹ Japanese has a special word for the cherry fruit, 桜ん坊 (*sakuranbō* “cherry”), with the etymological origin probably being 桜 (*sakura* “cherry tree”) + の (*no* “of”) + 坊 (*bō* “monk”), because the cherry fruit resembles the shaved head of a monk. The word sense can be built compositionally, by adding 材 (*zai* “lumber”) to 桜 (*sakura* “cherry tree”). Finnish also adds a suffix word for tree (*puu*) and other compounds.² Finally, Italian and Portuguese both describe the fruit and colour using the same wordform, and the tree and wood another wordform. In both cases, there is a clear relation between these two wordforms (*ciliegia* vs. *ciliegio* for Italian and *cereja* vs. *cerejeira* for Portuguese).

The *cherry* example demonstrates that even if the translations are not identical, sometimes different translations still share some semantic link. We can address this by analysing other synonyms in the same language, or translations in another language. For example, the metonymically-linked senses *cherry*₃ (fruit) and *cherry*₂ (tree) both share the same wordform in English, *cherry*. However, *cherry*₂ also has the synonym, *cherry tree*. The difference between them (+tree), can be used to classify this metonym as specifically a metonym which connects something to a tree. If we do this for all tropes, we find the same patterns repeated. Consider, for example, the first four senses of *chestnut*

¹This is also possible for English, and indeed the OEWn contains the synonym *cherry tree* for *cherry*₂ and *cherry red* for *cherry*₄ (but not *cherry wood* for *cherry*₁).

²We think there may be a mistranslation for the cherry wood: *kirsikkapuumetsä* means “cherry forest” which could also be “cherry wood”. This would suggest a problem with the wordnet.

in the OEWN:

<i>chestnut</i> ₁	wood of any of various chestnut trees of the genus <i>Castanea</i>
<i>chestnut</i> ₂	[<i>chestnut tree</i>] any of several attractive deciduous trees yellow-brown in autumn; yield a hard wood and edible nuts in a prickly bur
<i>chestnut</i> ₃	edible nut of any of various chestnut trees of the genus <i>Castanea</i>
<i>chestnut</i> ₄	the brown color of chestnuts

These senses each correspond to one of the four senses of *cherry*, and they show the same relation patterns of metonymy (*chestnut*₃ → *chestnut*₂ → *chestnut*₁) and metaphor (*chestnut*₃ → *chestnut*₄). Additionally, they have the same differences in synonym wordforms between *chestnut*₃ and *chestnut*₂ (+tree). When we look across the entire dataset, we find this and many other patterns repeating.

One interesting difference we have found in these patterns is that the direction in ChainNet is different for the same relation, depending on which parts of a tree are culturally important. For trees like cherry, chestnut and walnut, which are most widely used for their fruit, the metonymy chain goes: fruit → tree → wood. However, for trees used mainly for their wood, such as teak or mahogany, the metonymy goes from wood → tree.

The data and scripts to recreate these findings are available at <https://github.com/bond-lab/chainnet-xling>.

4.1 Different languages mark different things

One point of interest in Table 2 is that different languages mark different alternation patterns. For example, Japanese explicitly marks the colour sense with 色 (*iro* “colour”), Mandarin and Finnish explicitly mark the tree sense, and Mandarin and Japanese explicitly mark the wood sense. By building up a collection of these rules, it would be possible to predict how a language would accommodate new senses, and to specify which patterns are marked or unmarked in a particular language. For example, Japanese marks fields of study with 学, but English does not mark fields of study: the word *history* refers to both the past, and to the study of the past. We can therefore use information from Japanese to subtype the unmarked relation in English, for pairs like *accounting*₃ (defined as “the occupation of maintaining and auditing records and preparing financial reports for a business”) and *accounting*₂ (“a system that provides quantitative information about finances”), which in Japanese are

会計 and 会計学 respectively. The same is true for the sense *history*₅ (“all that is remembered of the past as preserved in writing; a body of knowledge”), which in Japanese is 歴史 or 史, and *history*₃ (“the discipline that records and interprets past events involving human beings”), which in Japanese is 歴史学 or 史学.

To facilitate investigation into derivation marking, we have created a database that link pairs of concepts (using CILI) to differences between them. This database is linked to the OMW so that the information is accessible in any wordnet, and makes it possible to search based on concepts, tropes, words, or differences. This database makes it easy to find, for example, all metaphors connected with the CILI concept i77824 (defined as “a red fruit with a single hard stone”), or to lookup all tropes linked by +tree. This effectively makes it possible to identify subregularities in WordNet senses. For example, a search for metonyms with -spot identifies all senses of the form “a playing card or domino or die whose upward face shows *n* pips”, where *n* is *one, four, five, six, seven, eight, nine, and ten*.³ Many of these subregularities are already captured in the wordnet hierarchy: all of these concepts are all children of the {spot, pip} synset, defined as “a mark on a die or on a playing card (shape depending on the suit)”.⁴

In future work, we intend to use the difference links and tropes to identify regular patterns, and then to use these patterns to generate all possible exemplars. Using this method, it would be possible to identify missing items in the OMW. However, as language is not completely regular, this will require manual checking of the results.

4.2 Limitations

While the findings point to significant cross-linguistic patterns, several limitations need to be addressed to refine the analysis. Currently we only look for prefixes and suffixes. Because of this, more complicated morphology or substitution (e.g. 树/木 “tree”/“wood” in Chinese) will be missed. In future work, we may consider other patterns. Moreover, some of the difference links we find are just spelling variants (e.g. +te, evidenced in *toilet/toilette* or *sextet/sextette*). These can be explicitly marked as spelling variants in the GWA LMF format for wordnets (McCrae et al., 2021),

³Strangely, the OEWN is missing *two-spot* and *three-spot*.

⁴The exception is *one-spot*, which is likely the result of an annotation error in the OEWN.

but this has not yet been done systematically for English. We intend to highlight these as suggested improvements to the upstream wordnets when we find them.

5 Discussion and future work

The findings of this study have far-reaching implications for our understanding of figurative language across languages. By demonstrating that metonymy is more widely shared across languages than metaphor, we show a measurable difference between the relations. This indicates that metonymy, as a form of meaning extension based on association and contiguity, is more grounded in universal cognitive processes, possibly related to human categorization and object-referencing behaviours. In contrast, metaphor, which often requires analogical reasoning and cultural framing, appears to be more language-specific or culturally nuanced. These findings fit well with existing discussion in metaphor research. For example, [Lakoff and Johnson \(1980, p. 39\)](#) argue that “the grounding of metonymic concepts is in general more obvious than in the case with metaphorical concepts, since [metonymy] usually involves direct physical or causal associations”.

A trope could be shared across multiple languages for several different reasons. One possibility is that the same trope arose independently in different languages, because speakers of different languages share some basic cognitive common ground, and the trope builds on some common cultural context. For example, the metaphor for *death* in example (1) and the metonym for *apple* in example (2) both appear in multiple languages: these tropes could have arisen independently in different speaking communities, because death is a universal part of the human experience, and apples have been grown around Eurasia since antiquity. Alternatively, it is also possible that the same trope appears in multiple languages because it was imported into one language from another. For example, the presence of the *death* metaphor in Catalan and Slovene could be because they are both Indo-European languages, which share a common ancestor. However, even if the same metaphor entered different languages from a common origin, the fact that the metaphor persists today is still of note. Many historical homonyms have presumably not had the same staying power, as otherwise we would expect the “unlinked” scores in Table 1 to

be higher for languages which are more closely related in terms of their typology.

At a practical level, our results provide valuable insights for constructing multilingual lexical resources. The fact that metonymy exhibits more cross-linguistic consistency suggests that resources like wordnets can benefit from prioritizing metonymic links when constructing multilingual synsets. Future research could explore how languages with different cultural and historical backgrounds handle metonymy and metaphor differently, in order to enrich linguistic databases with finer-grained semantic distinctions. We would also like to conduct a detailed analysis of possible sense extension along the lines of [Peters \(2003\)](#) and [Alonge and Lönneker \(2004\)](#). Looking at a small subset of metonymical relations in three languages from EuroWordnet (Dutch, English and Spanish), [Peters \(2003\)](#) investigated how metonymy can be used to suggest missing senses, while [Alonge and Lönneker \(2004\)](#) did the same for metaphors. It should also be noted that in many cases a relationship could be described as metaphor or metonymy, depending on the perspective of the analyst and whether they are highlighting contiguity or similarity ([Steen, 2005, p. 5](#)).

Finally, we would like to add the metaphor/metonymy links to the WordNet LMF as new sense-level links. If this were done, it would then be possible to add the ChainNet links to the Open English WordNet (OEWN). ChainNet’s detailed annotations of metaphor and metonymy offer a structured view of inter-sense relationships that is currently missing from OEWN. By integrating these links directly into OEWN, researchers and developers could benefit from a unified resource in which figurative sense extensions are directly accessible. This would make OEWN not only a more comprehensive lexical resource but also an essential tool for studying lexical polysemy and figurative language at scale. In addition, it would make it possible to keep the ChainNet data synchronised with the OEWN as new synsets and senses are added or deleted.

6 Conclusions

In this study, we have demonstrated that metonymy is more consistently multilingual than metaphor across over 30 languages. This finding underscores the importance of recognizing metonymy as a fundamental cognitive process that transcends cul-

tural and linguistic boundaries, whereas metaphor seems to be more contextually bound to individual languages and cultures. Our use of ChainNet and the Open Multilingual Wordnet has allowed us to systematically explore these patterns and provide new insights into the structure of polysemy.

The results point to a greater need for linguistic resources that reflect these differences in how figurative language operates cross-linguistically. As linguistic researchers, we should continue to probe how and why certain figurative relations hold across languages while others do not. Future research should expand the dataset and explore these phenomena across more languages, ultimately deepening our understanding of the interplay between language, cognition, and culture.

In order to continue research such as this, it is essential to continue expanding wordnets and enhancing their searchability. Expanding these resources and linking them to sense-tagged corpora will unlock even deeper insights into the nature of polysemy and figurative language across cultures.

References

- Antonietta Alonge. 2006. [The Italian metaphor database](#). In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*, Genoa, Italy. European Language Resources Association (ELRA).
- Antonietta Alonge and Birte Lönneker. 2004. [Metaphors in wordnets: From theory to practice](#). In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*, Lisbon, Portugal. European Language Resources Association (ELRA).
- Max Black. 1962. *Models and Metaphor*. Cornell University Press.
- Max Black. 1977. More about metaphor. *Dialectica*, pages 431–457.
- Francis Bond and Ryan Foster. 2013. [Linking and extending an open multilingual wordnet](#). In *51st Annual Meeting of the Association for Computational Linguistics: ACL-2013*, pages 1352–1362, Sofia.
- Francis Bond, Piek Vossen, John McCrae, and Christiane Fellbaum. 2016. CILI: The collaborative interlingual index. In *Proceedings of the 8th Global Wordnet Conference (GWC 2016)*, pages 50–57.
- Valeria de Paiva and Alexandre Rademaker. 2012. Revisiting a Brazilian wordnet. In *Proceedings of the 6th Global WordNet Conference (GWC 2012)*, Matsue.
- Sabri Elkateb, William Black, Horacio Rodríguez, Musa Alkhalifa, Piek Vossen, Adam Pease, and Christiane Fellbaum. 2006. Building a wordnet for Arabic. In *Proceedings of The fifth international conference on Language Resources and Evaluation (LREC 2006)*.
- Christine Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press.
- Fausto Giunchiglia, Khuyagbaatar Batsuren, and Abed Alhakim Freihat. 2018. One world-seven thousand languages (best paper award, third place). In *International Conference on Computational Linguistics and Intelligent Text Processing*, pages 220–235. Springer.
- Fausto Giunchiglia, Khuyagbaatar Batsuren, and Gabor Bella. 2017. [Understanding and exploiting language diversity](#). In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, pages 4009–4017.
- Fausto Giunchiglia, Gábor Bella, Nandu C. Nair, Yang Chi, and Hao Xu. 2023. [Representing interlingual meaning in lexical databases](#). *Artificial Intelligence Review*, page 11053–11069.
- Aitor Gonzalez-Agirre, Egoitz Laparra, and German Rigau. 2012. Multilingual central repository version 3.0: upgrading a very large lexical knowledge base. In *Proceedings of the 6th Global WordNet Conference (GWC 2012)*, Matsue.
- Michael Wayne Goodman and Francis Bond. 2021. Intrinsically interlingual: The Wn Python library for wordnets. In *11th International Global Wordnet Conference (GWC2021)*.
- Chu-Ren Huang, Shu-Kai Hsieh, Jia-Fei Hong, Yun-Zhu Chen, I-Li Su, Yong-Xiang Chen, and Sheng-Wei Huang. 2010. Chinese wordnet: Design and implementation of a cross-lingual knowledge processing infrastructure. *Journal of Chinese Information Processing*, 24(2):14–23. (in Chinese).
- Hitoshi Isahara, Francis Bond, Kiyotaka Uchimoto, Masao Utiyama, and Kyoko Kanzaki. 2008. Development of the Japanese WordNet. In *Sixth International conference on Language Resources and Evaluation (LREC 2008)*, Marrakech.
- Paul Jaccard. 1901. Étude comparative de la distribution florale dans une portion des alpes et des jura. *Bulletin de la Société vaudoise des sciences naturelles*, 37:547–579.
- Roman Jakobson. 1956. The metaphoric and metonymic poles. In Roman Jakobson and Morris Halle, editors, *Fundamentals of Language*. Mouton de Gruyter.
- Temuulen Khishigsuren, Gábor Bella, Thomas Brochhagen, Daariimaa Marav, Fausto Giunchiglia, and Khuyagbaatar Batsuren. 2022. [Metonymy as a universal cognitive phenomenon: Evidence from multilingual lexicons](#).

- George Lakoff and Mark Johnson. 1980. *Metaphors we live by*. University of Chicago Press.
- Krister Lindén and Lauri Carlson. 2010. Finnwordnet — wordnet påfinska via översättning. *LexicoNordica — Nordic Journal of Lexicography*, 17:119–140. In Swedish with an English abstract.
- Rowan Hall Maudslay, Simone Teufel, Francis Bond, and James Pustejovsky. 2024. [ChainNet: Structured metaphor and metonymy in WordNet](#). In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 2984–2996, Torino, Italy. ELRA and ICCL.
- John P. McCrae, Michael Wayne Goodman, Francis Bond, Alexandre Rademaker, Ewa Rudnicka, and Luís Morgado da Costa. 2021. The global wordnet formats: Updates for 2020. In *11th International Global Wordnet Conference (GWC2021)*.
- John P. McCrae, Alexandre Rademaker, Francis Bond, Ewa Rudnicka, and Christiane Fellbaum. 2019. [English WordNet 2019 —an open-source wordnet for English](#). In *Proceedings of the 11th Global Wordnet Conference (GWC 2019)*.
- George A. Miller. 1995. [WordNet: A lexical database for English](#). *Commun. ACM*, 38(11):39–41.
- Nurri Hirfana Mohamed Noor, Suerya Sapuan, and Francis Bond. 2011. Creating the open Wordnet Bahasa. In *Proceedings of the 25th Pacific Asia Conference on Language, Information and Computation (PACLIC 25)*, pages 258–267, Singapore.
- Mortaza Montazery and Heshaam Faili. 2010. Automatic Persian wordnet construction. In *23rd International conference on computational linguistics*, pages 846–850.
- Noam Ordan and Shuly Wintner. 2007. Hebrew wordnet: a test case of aligning lexical databases across languages. *International Journal of Translation*, 19(1):39–58.
- Bolette Sandford Pedersen, Sanni Nimb, Jørg Asmussen, Nicolai Hartvig Sørensen, Lars Trap-Jensen, and Henrik Lorentzen. 2009. DanNet — the challenge of compiling a wordnet for Danish by reusing a monolingual dictionary. *Language Resources and Evaluation*, 43(3):269–299.
- Wim Peters. 2003. [Metonymy as a cross-lingual phenomenon](#). In *Proceedings of the ACL 2003 Workshop on the Lexicon and Figurative Language*, pages 1–9, Sapporo, Japan. Association for Computational Linguistics.
- Emanuele Pianta, Luisa Bentivogli, and Christian Giardi. 2002. Multiwordnet: Developing an aligned multilingual database. In *Proceedings of the First International Conference on Global WordNet*, pages 293–302, Mysore, India.
- Maciej Piasecki, Stan Szpakowicz, and Bartosz Broda. 2009. [A Wordnet from the Ground Up](#). Wrocław University of Technology Press. (ISBN 978-83-7493-476-3).
- James Pustejovsky. 1995. *The Generative Lexicon*. MIT Press, Cambridge, MA.
- Ervin Ruci. 2008. On the current state of Albanet and related applications. Technical report, University of Vlora. (<http://fjalnet.com/technicalreportalbanet.pdf>).
- Benoît Sagot and Darja Fišer. 2008. Building a free French wordnet from multilingual resources. In *Proceedings of the Sixth International Language Resources and Evaluation (LREC’08)*, Marrakech, Morocco.
- Gerard Steen. 2005. [Metonymy goes cognitive-linguistic](#). *Style*, 39(1):1–11.
- Sareewan Thoongsup, Thatsanee Charoenporn, Kergrit Robkop, Tan Sinthurahat, Chumpol Mokarat, Virach Sornlertlamvanich, and Hitoshi Isahara. 2009. Thai wordnet construction. In *Proceedings of The 7th Workshop on Asian Language Resources (ALR7), Joint conference of the 47th Annual Meeting of the Association for Computational Linguistics (ACL) and the 4th International Joint Conference on Natural Language Processing (IJCNLP)*, Suntec, Singapore.
- Antonio Toral, Stefania Bracale, Monica Monachini, and Claudia Soria. 2010. Rejuvenating the italian wordnet: upgrading, standardising, extending. In *Proceedings of the 5th International Conference of the Global WordNet Association (GWC-2010)*, Mumbai.