JPC1: Metaphor and Metonymy in English, explored through ChainNet and Wordnet

Francis Bond
bond@ieee.org

December 25, 2025

1 Introduction

I look at the tropes (metaphors and metonyms) identified in ChainNet (Maudslay et al., 2024), and use wordnet to tell us something more about them.

The original discussion asked for some different tasks:

- import chainnet to (English) wordnet
- project to other languages
- visualise
  I have some rough code (chainnet-viz) that I will share
- compare to corpus
- look at sentiment
- look for errors
  - find exceptions to patterns
  - evaluate and fix
  - mainly direction, but can be type

At first, I decided to go in a different direction, as there were some unexpected difficulties in importing chainnet into wordnet (the `wn` module did not yet support the `metaphor` and `metonym` links, neither did the `wn-edit` module).

But `wn` version 0.14 supports `metaphor` and `metonym`, so I wrote a new `wn_edit` module and imported ChainNet to the English Wordnet, both for the OMW English Wordnet, and an updated Open English Wordnet.

2 Background

We will use wordnet to quantitatively test some of the hypothesis about Metaphor and Metonymy. Metaphor and metonymy are treated in cognitive linguistics as conceptual mechanisms, not merely stylistic devices (Kövecses, 2010, Ch 12). In Kövecses' framework, metaphor involves understanding one conceptual domain in terms of another, typically mapping structure from a more concrete source domain onto a more abstract target domain. These mappings are partial and systematic, highlighting selected aspects of the target and motivating families of related linguistic expressions (e.g. TIME IS MONEY, ARGUMENT IS WAR). The primary cognitive function of metaphor is thus conceptualization, providing structured ways of reasoning about abstract experience.

By contrast, metonymy operates within a single conceptual domain or idealized cognitive model, where one entity provides mental access to another via relations of contiguity or functional association rather than similarity (Kövecses, 2010). Typical metonymic relations include `part{whole`, `producer{product`, `place{institution`, and `container{contained` (e.g. `the White House announced`, `drink a glass`). Metonymy primarily serves referential and attentional functions, enabling efficient access to salient aspects of a conceptual structure without invoking cross-domain analogy.

The crucial distinction in Kövecses' account therefore lies in both conceptual scope and function: metaphor involves *cross-domain mappings* for understanding, while metonymy involves *within-domain mappings* for reference. Although the two frequently interact in natural language, maintaining this distinction is essential for analyzing systematic patterns of meaning extension and lexical polysemy. Metaphor involves two concepts distant in our conceptual system, while metonymy relates two elements closely related to each other in the concept space (Kövecses, 2010, p370).

Resources

ChainNet is a structured lexical resource that explicitly models metaphorical and metonymic sense extensions within WordNet (Maudslay et al., 2024). It annotates senses in Princeton WordNet 3.0 with typed relations that capture systematic patterns of semantic extension, allowing metaphor and metonymy to be analyzed as graph-structured phenomena rather than isolated examples. By linking related senses into chains, ChainNet enables both qualitative linguistic analysis and quantitative investigation of the structural properties of meaning extension, including cross-linguistic comparison when combined with multilingual WordNet resources.

Princeton WordNet 3.0 is a large, manually curated lexical database of English in which words are grouped into synsets representing distinct concepts and interconnected by typed semantic relations such as hypernymy, meronymy, and antonymy (Fellbaum, 1998). It provides a widely adopted conceptual backbone for lexical semantics, serving as the reference ontology for numerous NLP applications and as the foundation for many multilingual WordNets. In this work, WordNet 3.0 functions as the base semantic network over which metaphorical and metonymic extensions are modeled.

The wn Python module is a modern, extensible interface for working programmatically with WordNet data and related lexical resources (Goodman and Bond, 2021). It provides unified access to multiple WordNet versions and languages, supports querying of synsets, senses, and semantic relations, and allows users to integrate external annotations and extensions. Its design facilitates reproducible experimentation and large-scale analysis, making it well suited for computational studies of lexical semantics and for integrating resources such as ChainNet into broader processing pipelines.

The wn_edit module[1] extends the 'wn' module with capabilities for creating, modifying, and exporting WordNet lexicons. It works directly with 'wn.lmf' data structures, ensuring full compatibility with the 'wn' ecosystem.

3 Approach

For every trope, we look at either the src and target nodes, or some difference between them. The code is in `analyse-tropes.py`

In order to test whether metaphor (and metonymy) typically map from a more concrete source to a more abstract target, we investigate three measures of abstractness (Depth, Abstractness and Synonym Density), defined below.

In order to test if metaphors are more likely to cross domains, we look at the topic. We use the path distance in the wordnet graph to measure distance between source and target. We also investigate the difference in depth between source and target.

For the nodes we look at:

- Depth (distance to root)
- Abstractness
- Inclusiveness ToDo

[1] https://github.com/bond-lab/wn_edit

- Synonymy density (number of synonyms)
- Topic (lexicographer file)

Depth is just the maximum depth from the root (`synset.min_depth()`). Synonymy density is the number of synonyms the concept has (`len(ss1.lemmas())`). Topic is the lexicographer file. Abstractness is calculated as suggested by Mensa et al. (2018) 0 if the synset is a hyponym of ***physical entity***, 1 otherwise.

We should also calculate inclusiveness (Iliev and Axelrod, 2017, p 719). This is defined as: if a concept c has n descendents (both direct and indirect) in a tree that has a total of N nodes, then:

$$\text{inclusiveness}(c) = log((n+1)/N) \tag{1}$$

For the differences we look at:

- Distance between nodes (`path_distance`)
- Difference in depth between nodes (source depth - target depth)

These are calculated for each node or pair of nodes, as appropriate.

Between tropes (Metaphor vs Metonymy) we measure for significance using the Mann-Whitney U test. It tests whether values from one group tend to be larger or smaller than values from the other group. Within trope (source vs target) we use Wilcoxon signed-rank tests. It compares paired observations (which are statistically dependent), testing whether the median difference between paired values is zero.

4 Results

Statistical Methods

Between-Trope Comparisons We compared metaphor and metonymy using the Mann-Whitney U test, a non-parametric test appropriate for comparing two independent samples without assuming normal distributions. Given the large sample sizes ($n_{metaphor} = 7500$, $n_{metonymy} = 6116$), this test provides robust results even with skewed distributions.

Within-Trope Comparisons To compare source and target synsets within each trope type, we used the Wilcoxon signed-rank test, a non-parametric paired test that does not assume normal distributions. This test is appropriate because each trope instance provides a natural pairing of source and target synsets.

Table 1: Descriptive Statistics for Metaphor and Metonymy

Measure	Metaphor			Metonymy		
	Mean	SD	IQR	Mean	SD	IQR
Source Depth	7.14	2.00	2.00	6.95	2.01	2.00
Target Depth	7.14	1.94	2.00	7.20	1.96	2.00
Source Synonyms	2.29	1.67	2.00	2.39	1.89	2.00
Target Synonyms	2.54	2.30	2.00	2.41	1.84	2.00
Source Abstractness	0.50	0.50	1.00	0.61	0.49	1.00
Target Abstractness	0.56	0.50	1.00	0.62	0.49	1.00
Path Distance	0.12	0.07	0.05	0.10	0.06	0.04
Depth Difference	0.01	2.31	2.00	-0.24	2.47	3.00
Abstract Difference	-0.06	0.48	0.00	-0.00	0.49	0.00
N	7500			6116		

Table 2: Statistical Comparison: Metaphor vs Metonymy

Measure	Mann-Whitney U	Cohen's d	Effect Size
Source Depth	$p < .001$	0.095	negligible
Target Depth	$p = 0.02$	-0.031	negligible
Source Synonyms	$p = 0.008$	-0.059	negligible
Target Synonyms	$p = 0.63$	0.060	negligible
Source Abstractness	$p < .001$	-0.233	small
Target Abstractness	$p < .001$	-0.116	negligible
Path Distance	$p < .001$	0.210	small
Depth Difference	$p < .001$	0.105	negligible
Abstract Difference	$p < .001$	-0.120	negligible

Table 3: Statistical Comparison: Source vs Target within Trope Types

Comparison	Mean Diff	Wilcoxon	Cohen's d	Effect Size
Metaphor Depth	0.01	$p = 0.98$	0.003	negligible
Metaphor Synonyms	-0.25	$p < .001$	-0.090	negligible
Metaphor Abstractness	-0.06	$p < .001$	-0.131	negligible
Metonymy Depth	-0.24	$p < .001$	-0.099	negligible
Metonymy Synonyms	-0.02	$p = 0.48$	-0.007	negligible
Metonymy Abstractness	-0.00	$p = 0.48$	-0.009	negligible

Effect Sizes We report Cohen's d as a standardized measure of effect size, interpreted as: negligible ($|d| < 0.2$), small ($0.2 \leq |d| < 0.5$), medium ($0.5 \leq |d| < 0.8$), or large ($|d| \geq 0.8$). For between-trope comparisons, positive values indicate metaphor has higher values; for within-trope comparisons, positive values indicate source has higher values than target.

Key Findings: Between Tropes Statistically significant differences ($p < .001$) between metaphor and metonymy:

- **Source Depth**: Metaphor shows higher values (Cohen's $d = 0.095$, negligible effect).
- **Source Abstractness**: Metaphor shows lower values (Cohen's $d = -0.233$, small effect).
- **Target Abstractness**: Metaphor shows lower values (Cohen's $d = -0.116$, negligible effect).
- **Path Distance**: Metaphor shows higher values (Cohen's $d = 0.210$, small effect).
- **Depth Difference**: Metaphor shows higher values (Cohen's $d = 0.105$, negligible effect).
- **Abstract Difference**: Metaphor shows lower values (Cohen's $d = -0.120$, negligible effect).

Key Findings: Source vs Target Statistically significant differences ($p < .001$) between source and target:

- **Metonymy Depth**: Source shows lower values than target (mean diff = -0.24, Cohen's $d = -0.099$, negligible effect).
- **Metaphor Synonyms**: Source shows lower values than target (mean diff = -0.25, Cohen's $d = -0.090$, negligible effect).
- **Metaphor Abstractness**: Source shows lower values than target (mean diff = -0.06, Cohen's $d = -0.131$, negligible effect).

Table 4: Most Frequent Topic Transitions in Metaphor (Top 15)

Source Topic	Target Topic	Count
noun.artifact	noun.artifact	805
noun.person	noun.person	501
noun.act	noun.act	406
noun.communication	noun.communication	291
noun.attribute	noun.attribute	206
noun.group	noun.group	144
noun.artifact	noun.communication	107
noun.cognition	noun.cognition	99
noun.animal	noun.animal	96
noun.animal	noun.person	89
noun.act	noun.event	85
noun.act	noun.communication	85
noun.quantity	noun.quantity	85
noun.state	noun.state	81
noun.artifact	noun.act	73

Table 5: Most Frequent Topic Transitions in Metonymy (Top 15)

Source Topic	Target Topic	Count
noun.artifact	noun.artifact	321
noun.act	noun.act	241
noun.communication	noun.communication	184
noun.person	noun.person	181
noun.act	noun.communication	126
noun.food	noun.plant	113
noun.artifact	noun.act	112
noun.cognition	noun.cognition	102
noun.act	noun.artifact	102
noun.attribute	noun.attribute	98
noun.plant	noun.plant	97
noun.communication	noun.act	89
noun.group	noun.group	79
noun.act	noun.attribute	77
noun.attribute	noun.act	75

Importing tropes to Wordnet

We used the new `wn_edit` module to create a version of the English wordnet that contains the metaphor and metonymy links.

The script `enhance.py`[2] does the following:

1. Load an English Wordnet (the default is `omw-en:1.4`)
2. Loads and parses the chainnet data (`chainnet_{metaphor|metonymy}.json`)
3. Creates an index from the sensekey to the sense id
4. Adds the metaphor and metonym links and their reverses to the wordnet (`metaphor`, `has_metaphor`, `metonym`, `has_metonym`)
5. Saves it with a new version (the default is `omw-en:1.4.cn`)

The new wordnet can now be used to investigate tropes:

```
>>> import wn
>>> wn.add('omw-en:1.4_cn.xml')
Skipping omw-en:1.4.cn (OMW English Wordnet based on WordNet 3.0 with tropes from ChainNet); already added
>>>
>>> ewn=wn.Wordnet(lexicon='omw-en:1.4.cn')
>>>
>>> for s in ewn.senses('chestnut'):
...   print(s, s.relations())
...
Sense('omw-en-chestnut-00373209-s') {'derivation': [Sense('omw-en-chestnut-04972350-n')]}
Sense('omw-en-chestnut-12262905-n') {'has_metonym': [Sense('omw-en-chestnut-12262553-n')]}
Sense('omw-en-chestnut-12262553-n') {'metonym': [Sense('omw-en-chestnut-12262905-n')],
                                     'has_metonym': [Sense('omw-en-chestnut-07772274-n')]}
Sense('omw-en-chestnut-07772274-n') {'metaphor': [Sense('omw-en-chestnut-02468504-n')],
                                      'metonym': [Sense('omw-en-chestnut-12262553-n'),
                                                  Sense('omw-en-chestnut-04972350-n')]}
Sense('omw-en-chestnut-04972350-n') {'derivation': [Sense('omw-en-chestnut-00373209-s')],
                                     'has_metonym': [Sense('omw-en-chestnut-07772274-n')],
                                     'metonym': [Sense('omw-en-chestnut-02388735-n')]}
Sense('omw-en-chestnut-02468504-n') {'has_metaphor': [Sense('omw-en-chestnut-07772274-n')]}
Sense('omw-en-chestnut-02388735-n') {'has_metonym': [Sense('omw-en-chestnut-04972350-n')]}
>>>
```

5 Discussion

Comparing Metaphor and Metonymy, we see small but signifigant differences in the abstractness of the source (where the source of metapors is less abstract (more concrete) than the source of metonyms), shown in Figure 2. A similar difference is found between the length of the path between the source and target and in the distance between them (Figure 1), as predicted by Kövecses.

[2] Available at https://github.com/fcbond/ChainNet

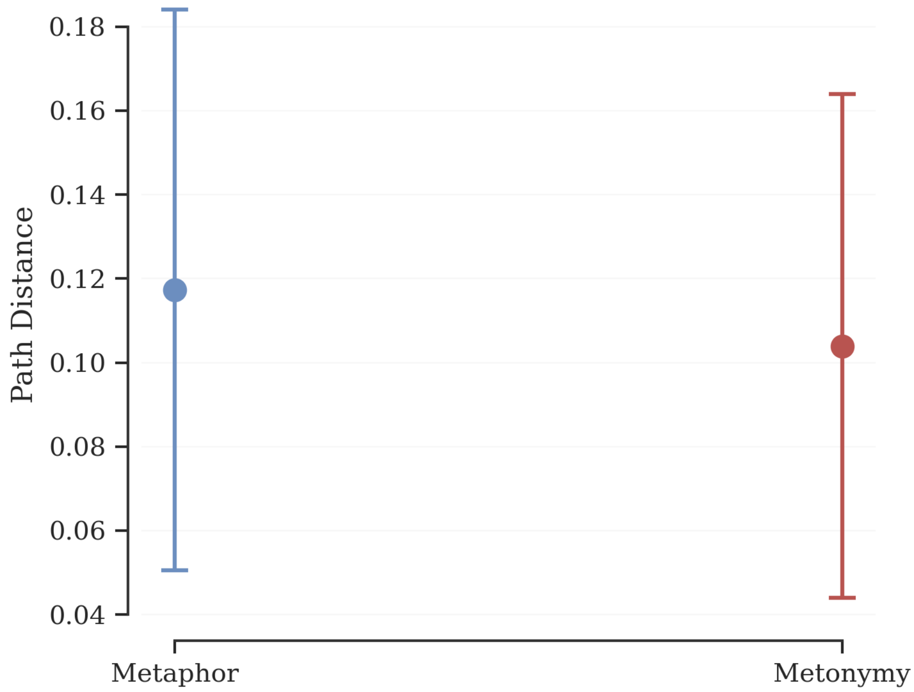


Figure 1: Path Distance Comparison

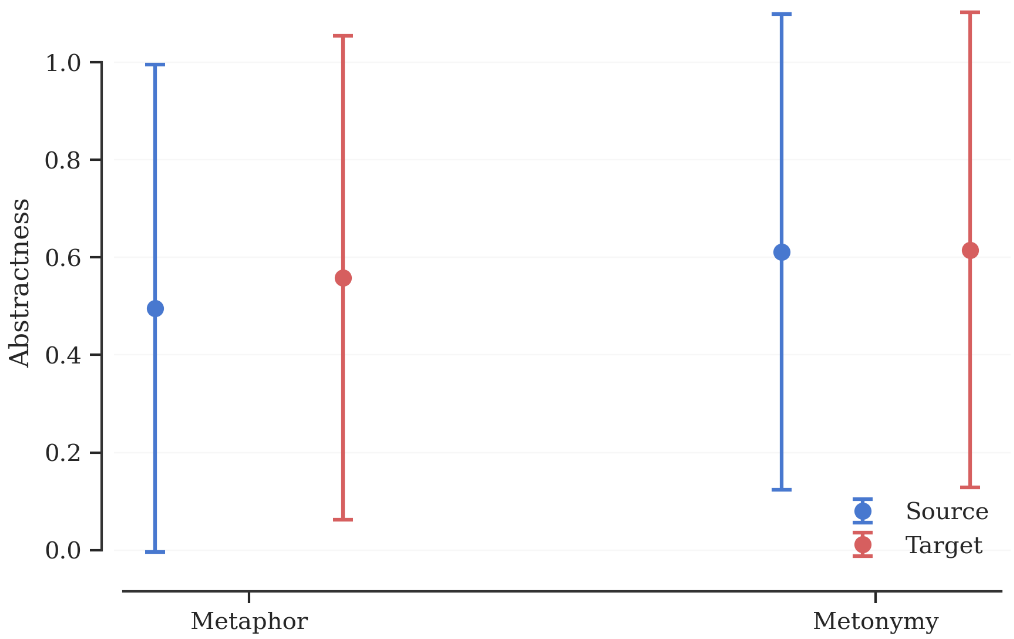


Figure 2: Abstractness Distance Comparison

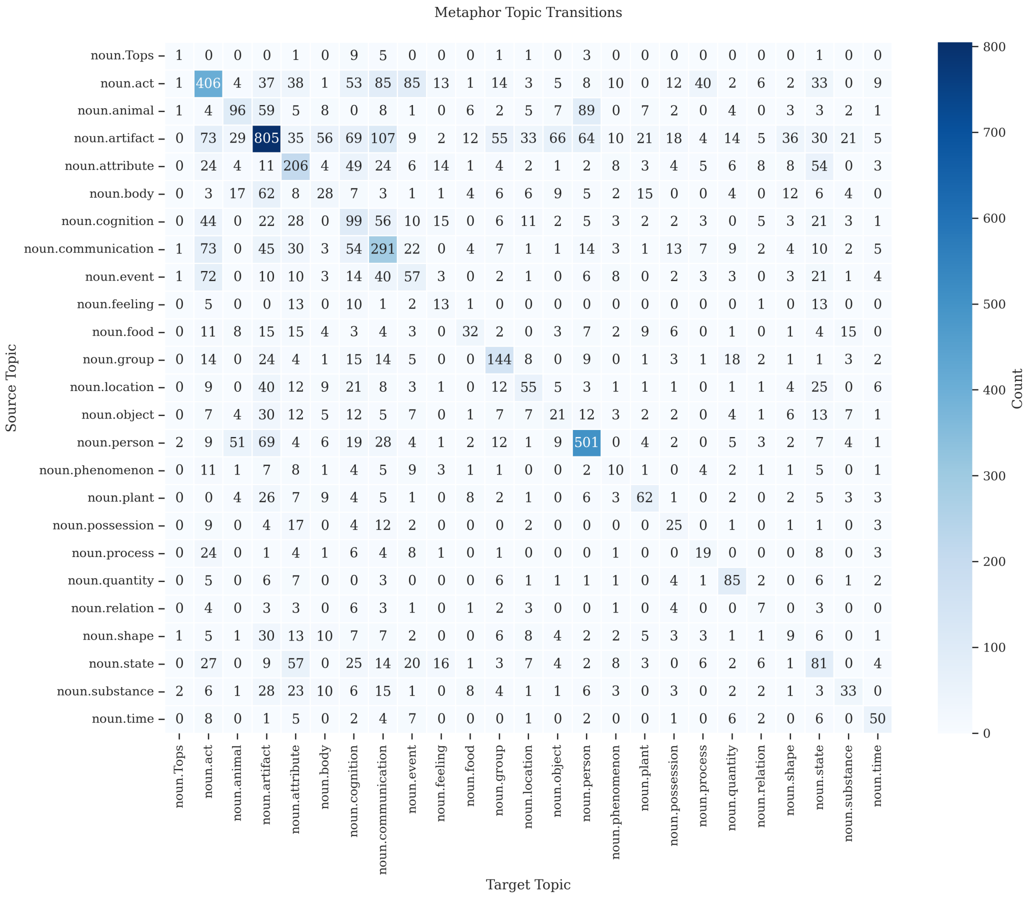
There are also significant differences between source depth, target abstractness and depth distance, but all very small.

Somewhat surprisingly, all the differences within tropes were very small (Table 3). There are statistically significant differences between abstractedness and synonym density for metaphors (source is more concrete and has fewer synonyms) and depth for metonymy (source is closer to the root than target).

When we look at the topics (Figure 3), we find that metaphors are more likely to share the same topic between source and target than metaphors, the opposite of what Kövecses predicted for domains. This could be because wordnet topics are different from domains, in future work we will also look at domains, and consider why this should be the case for topics.

There are several other measures it would be nice to look at: inclusiveness (describe in Section 3), the domain from eXtended WordNet Domains (González et al., 2012) and sentiment.

The wordnet merge script is has not been extensively tested, but works with the latest versions of the Open Multilingual Wordnet and the Open English Wordnet. It uses `wn.compat.sensekey` to look up senses with sense



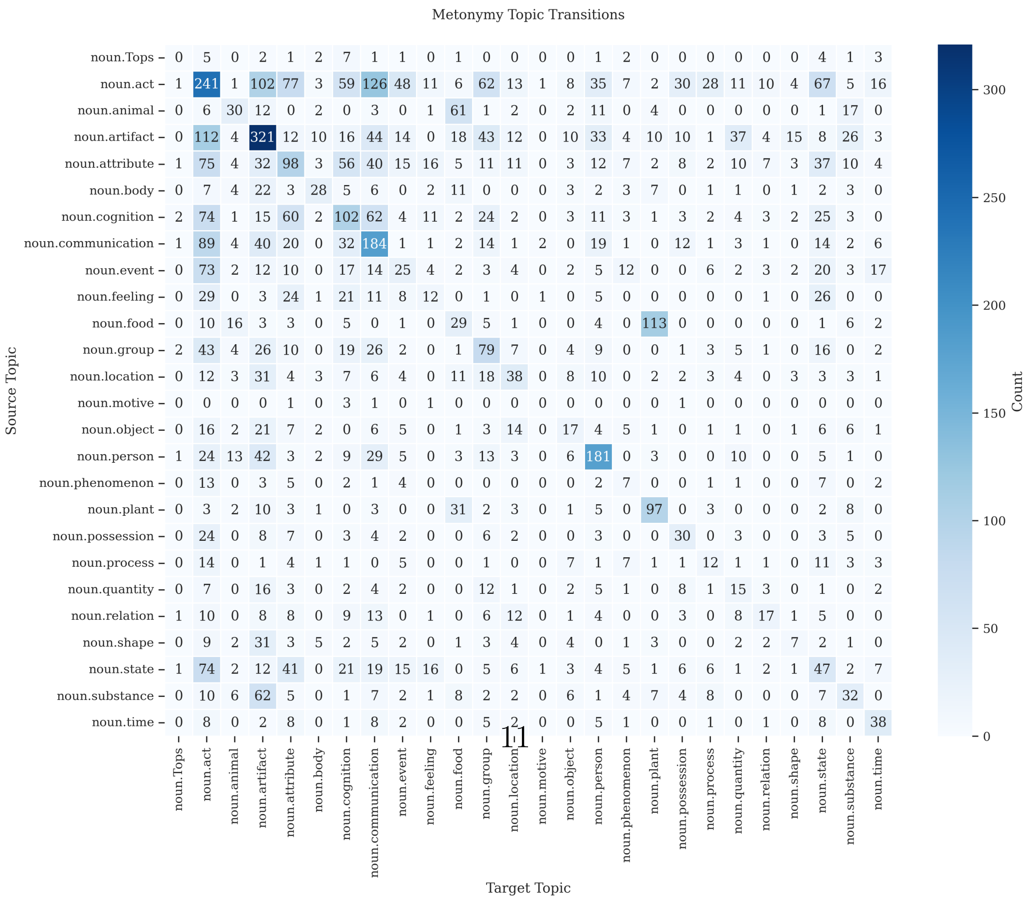


Figure 3: Topic heatmaps for metaphor and metonymy

keys. 66 relations that were in `omw-en:1.4` are no longer in `oewn:2024`.[3]

6 Conclusions

Comparing Metaphor and Metonymy, we are able to show quantitatively, small but signifigant differences in the abstractness of the source (where the source of metaphors is less abstract (more concrete) than the source of metonyms). A similar difference is found between the length of the path between the source and target and in the distance between them as predicted by Kövecses.

In future work we wish to also compare inclusiveness, which may be a better measure than abstractness, domains and sentiment.

References

Fellbaum, C., editor (1998). *WordNet: An Electronic Lexical Database*. MIT Press.

González, A., Rigau, G., and Castillo, M. (2012). A graph-based method to improve wordnet domains. In *Proceedings of the 13th International Conference on Computational Linguistics and Intelligent Text Processing - Volume Part I*, CICLing'12, page 17–28, Berlin, Heidelberg. Springer-Verlag.

Goodman, M. W. and Bond, F. (2021). Intrinsically interlingual: The Wn Python library for wordnets. In *11th International Global Wordnet Conference (GWC2021)*.

Iliev, R. and Axelrod, R. (2017). The paradox of abstraction: Precision versus concreteness. *Journal of Psycholinguist Research*, 46:715–729.

Kövecses, Z. (2010). *Metaphor: A Practical Introduction*. Oxford University Press, Oxford, 2 edition.

Maudslay, R. H., Teufel, S., Bond, F., and Pustejovsky, J. (2024). ChainNet: Structured metaphor and metonymy in WordNet. In Calzolari, N., Kan, M.-Y., Hoste, V., Lenci, A., Sakti, S., and Xue, N., editors, *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.

[3] For example fro *horseman, hacker, jihad, annulet* and *charade*.

Mensa, E., Porporato, A., and Radicioni, D. (2018). Annotating concept abstractness by common-sense knowledge. In *XVIIth International Conference of the Italian Association for Artificial Intelligence*, pages 415–428.