# Figurative, Polysemous, Conventional: Designing a Dataset of Regular Metaphor

**Anna Temerko, Pablo Gamallo, Marcos Garcia**
CiTIUS – Research Centre in Intelligent Technologies, Universidade de Santiago de Compostela
Santiago de Compostela, Spain
a.temerko@usc.es, {pablo.gamallo, marcos.garcia.gonzalez}@usc.gal

**Abstract**

Metaphor, a figure of speech and a cognitive device, offers a powerful way to explain one conceptual domain in terms of another. Particularly successful metaphorical mappings conventionalize through frequent use and lose their creative quality. They become sense extensions of polysemous words. Our dataset project captures such metaphors with ten regular polysemy patterns that manifest repetitively in the meaning structures of English words. Regular metaphor, unlike its counterpart regular metonymy, has not previously received a dedicated dataset, and we intend to close this gap. The dataset under construction features naturalistic sentences extracted from a general language corpus and is manually annotated with sense labels for metaphorically extended polysemes. Its intended use is to support linguistic, cognitive and computational investigations into patterns of meaning in polysemy, while accounting for its complexity, regularity, continuity and heterogeneity. We see neural language models as an excellent experimental ground for such research because they are able to show both distributional (continuous) and symbolic (discrete) behavior in language processing and representation. In this paper, we reflect on how these systems tally.

**Keywords:** metaphor, polysemy, large language models

## 1. Introduction

Metaphor is a figure of speech that involves mappings between two conceptual domains based on analogy or similarity between the source domain and the target domain (Lakoff and Johnson, 1980). Through frequent use and entrenchment (Grady et al., 1999; Langacker, 2017), metaphors conventionalize, thus moving away from the creative pole of the continuum (Cardillo et al., 2012). Some metaphors also lose their perceived figurative quality in this process and are processed similarly to literal expressions (Cardillo et al., 2010, 2017; Lai et al., 2009). These shifts are what contributes to a conventionalized, entrenched metaphor becoming a sense extension of a polysemous word (Egg and Kordoni, 2022; Reijnierse et al., 2018; Steen et al., 2010b). Some metaphoric mappings are so salient that they repeat themselves over and over across lexical units of a semantic domain, thus forming regular polysemy patterns (Lombard et al., 2023, 2024; Srinivasan and Rabagliati, 2015).

Metaphor and polysemy present a double challenge for natural language processing: polysemy, a language phenomenon that implies multiplicity of word's senses, causes ambiguity (Haber and Poesio, 2024), whereas metaphor, describing one concept in terms of another, requires semantic cross-domain integration (Momen et al., 2026). Both phenomena increase processing difficulties for humans and machines and both are ubiquitous: most lexical words show varying degrees of polysemy (Durkin and Manning, 1989; Haber and Poesio, 2024; Zipf, 1945), and metaphors are pervasive even in scientific and formal domains (Steen et al., 2010a).

This work-in-progress contribution presents a novel dataset that aims to capture the regularization of metaphor with 10 polysemy patterns of English. Its intended use is to support linguistic and cognitive analyses of regular metaphor and polysemy using computational linguistic methods. In particular, we envisage experimentation with neural language models (LMs) including distributional semantic modeling, contextualized representation analysis, clustering, sense induction methods and other types of exploration and probing.

Polysemy and figurativeness cover phenomena that evade attempts to represent them in exclusively discreet or exclusively continuous ways. Recent work on polysemy shows how lexical meaning is continuous in sense individuation, sense relatedness, productivity and regularity of polysemy patterns. At the same time, grouping these continuous phenomena into symbolic, discrete-like categories can serve as a meaningful generalization when observing human language processing and building theories of language (Boleda, 2025; Futrell and Mahowald, 2025; Li, 2024; Trott and Bergen, 2023).

In this light, neural language models can serve as a suitable tool to approach such heterogeneous, fluid language phenomena. They can flexibly offer both symbolic and distributed (continuous) representations, as argued in recent position papers by Boleda (2025), Futrell and Mahowald (2025), and Grindrod (2024) and demonstrated specifically on polysemy material by Li and Armstrong (2024) and Li and Joanisse (2021). Research into linguistic capabilities of LMs shows that they are able to



*Proceedings of the Workshop on Structured Linguistic Data and Evaluation*, pages 202–211

sufficiently encode relevant linguistics information (Garí Soler and Apidianaki, 2021; Li and Joanisse, 2021; Wiedemann et al., 2019). The architectural design and the emergent linguistic capabilities of LMs make them a suitable experimental ground to study language in its complexity and heterogeneity. Our rich and systematic language resource provides the means to do so.

## 2. Background and Related Work

Applying the framework of regular polysemy to metaphor is not supported by many linguistic approaches to polysemy. The term was first introduced and defined by Apresjan (1974, p. 16): "Polysemy of the word A with the meanings $a_i$ and $a_j$ is called regular if [...] there exists at least one other word B with the meanings $b_i$ and $b_j$, which are semantically distinguished from each other in exactly the same way as $a_i$ and $a_j$ [...]". Apresjan (1974) underlines that the typical sense extension device for this polysemy type is metonymy, while metaphorization is a source of irregular senses. Unlike metaphor that maps two disjuncted domains, metonymy operates on domain contiguity. The studies that cite and build upon Apresjan's work (1974) adopt this idea. Pustejovsky's Generative Lexicon tradition (Pustejovsky, 1991) and computational linguistic work relying on it exclude metaphor from regular polysemy investigation as well.

In contrast, a number of works in lexicography (Nimb and Pedersen, 2000; Polguère, 2018, 2022), lexicology (Berri, 2020), cognitive and psycholinguistics (Lombard et al., 2023, 2024; Srinivasan and Rabagliati, 2015), and intersecting computational domains (Barque and Chaumartin, 2009; Freihat et al., 2013; Peters and Peters, 2000) treat patterns of metaphor as regular, on par with metonymies. In particular, the psycholinguistic studies of Lombard et al. (2023) and Lombard et al. (2024) develop metrics of polysemy pattern regularity based on data from Wordnet (for English) and expert lexicological expertise (for French) showing that patterns of metaphor can be as regular as those of metonymy, with varying regularity degrees.

The concept of regular polysemy is important in the ongoing discussion on meaning representation in the human mind. Recently, hybrid approaches dominate, combining the sense enumeration approach and one meaning representation (see Haber and Poesio, 2024 for an overview). Li (2024) develops a symbolic-continuous polysemy theory based on previous experiments with neural language models (Li and Armstrong, 2024; Li and Joanisse, 2021), where regular polysemy patterns create entrenched meaning modulations that are continuous in sense individuation but allow grouping into emerging discrete-like sense clusters. Psycholinguistic evidence delivers similar accounts (Haber and Poesio, 2020; Rabagliati and Snedeker, 2013). Involving regular metaphors into the experimentation of this kind could give us a more complete picture of how patterns of meaning form. Unfortunately, the exclusion of metaphor from regular polysemy theories, and consequently from NLP experimentation that adopts these theories (if any at all), is partly a reason for the absence of a sizable corpus that could help to systematically compare two of the most important regular sense extension devices (metonymy and metaphor) on a scale. The present work contributes to the field by developing such a resource.

## 3. Dataset Design and Methodology

**Scope.** We treat conventional metaphor mappings that form patterns across the vocabulary, as per the above definition of regular polysemy (Apresjan, 1974). Conventionalized metaphors usually appear as senses of polysemous words in the dictionaries (Egg and Kordoni, 2022; Reijnierse et al., 2018; Steen et al., 2010a). If we can observe the same mappings in the sense structures in a series of words, we can consider it a regular pattern.

**Source of patterns, words and senses.** We borrow metaphorical patterns and their lexical realizations from previous research in regular polysemy, in particular Lombard et al. (2024). The authors examine the distribution of 15 polysemy patterns in the English lexicon by extracting lexical material from WordNet (Fellbaum, 1998; Miller, 1995), manually validating it and applying quantitative measures to assess the degree of regularity of the patterns. These measures are based on the number of words that fit the pattern, the number of words that could potentially do so and the frequency of target words in a corpus. From 15 patterns, we discarded those that were extremely irregular and would not warrant sufficient material for experimentation (less than 10 words identified as fitting the pattern), as well as those that involve highly specialized, terminological senses usually unknown to general language users. The remaining 10 patterns still fall under a wide range of regularity degrees, with some of them approached previously in psycholinguistics (Srinivasan and Rabagliati, 2015) and computational lexicology (Barque and Chaumartin, 2009).

The patterns covered at the current stage of the project are ANIMAL CRY - COMMUNICATION and NATURAL PHENOMENON - WAY OF OCCURRING. We examine both noun and verb realizations when possible. The sense mappings covered by these patterns are exemplified in Table 1 for words *roar* and *deluge*.[1] Dictionaries are used to verify that sense descriptions of selected words exhibit the intended

[1]The complete inventory of words and senses under

| Pattern | Word | PoS | Literal | Metaphoric |
|---|---|---|---|---|
| ANIMAL CRY - COMMUNICATION | roar | V | (of a lion or other large wild animal) utter a full, deep, prolonged cry | (of a person or crowd) utter a loud, deep, prolonged sound, typically from anger, pain, or excitement; utter or express in a loud tone; (of a crowd) encourage (someone) to do something by loud shouts or cheering; laugh loudly |
| ANIMAL CRY - COMMUNICATION | roar | N | a full, deep, prolonged cry uttered by a lion or other large wild animal | a loud, deep sound uttered by a person or crowd, generally as an expression of pain, anger, or approval; a loud outburst of laughter |
| NATURAL PHENOMENON - WAY OF OCCURRING | deluge | V | overwhelm with a flood | to overwhelm with a large number or amount |
| NATURAL PHENOMENON - WAY OF OCCURRING | deluge | N | a severe flood; a heavy fall of rain | a great quantity of something arriving at the same time |

Table 1: Sample of sense inventory. The table presents senses (*Literal* and *Metaphoric*) covered by two exemplified *Patterns* realized by words *roar* and *deluge*. Sense descriptions derived from OED.

| Pattern | Word | PoS | S. | Sentence |
|---|---|---|---|---|
| ANIMAL CRY - COMMUNICATION | roar | V | L | The lion has **roared**: who will not fear? |
| ANIMAL CRY - COMMUNICATION | roar | V | M | He **roared** a great shout of laughter. |
| ANIMAL CRY - COMMUNICATION | roar | N | L | We ate in the London Zoo and our meals were made interesting by the chatter of monkeys and the **roar** of lions in the background. |
| ANIMAL CRY - COMMUNICATION | roar | N | M | Against the **roar** of rage he waved for silence. |
| ANIMAL CRY - COMMUNICATION | bleat | V | L | The sheep were pushing into the fold, stumbling and **bleating**. |
| ANIMAL CRY - COMMUNICATION | bleat | V | M | He **bleats** that his critics are talking the country down |
| ANIMAL CRY - COMMUNICATION | bleat | N | L | The only sounds to be heard were the sheep's teeth tearing grass and their low, rumbling **bleats**. |
| ANIMAL CRY - COMMUNICATION | bleat | N | M | They're trying to buy me off in hopes I'll bow to their idiotic new arrangements without a **bleat**. |
| NATURAL PHENOMENON - WAY OF OCCURRING | torrent | N | L | This produced a **torrent** of words but little effective action. |
| NATURAL PHENOMENON - WAY OF OCCURRING | torrent | N | M | A quiet river on a summer's day may be a raging **torrent** in February. |
| NATURAL PHENOMENON - WAY OF OCCURRING | deluge | V | L | The lake of liquid peat had burst through into the workings beneath it and was **deluging** into the colliery. |
| NATURAL PHENOMENON - WAY OF OCCURRING | deluge | V | M | Awards were **deluged** on him, as were titles, praise and eulogies in the national press. |
| NATURAL PHENOMENON - WAY OF OCCURRING | deluge | N | L | One **deluge** won't end the drought. |
| NATURAL PHENOMENON - WAY OF OCCURRING | deluge | N | M | Afterwards viewers sent a **deluge** of complaints about how boring the show was and asking them never to return to the Hacienda. |

Table 2: Dataset sample. Column *Pattern* contains the names of regular poylysemy patterns; *Word* contains the words instantiating the pattern; *PoS* marks either a verb (*V*) or a noun (*N*); *S.* (sense) labels literal (*L*) and metaphorical (*M*) senses; *Sentence* contains extracted sentences.

metahporical pattern (Oxford English Dictionary — OED and American Heritage Dictionary — AHD)[2]. Reimann and Scheffler (2024) experimentally show that using sense descriptions from dictionaries is more reliable than human scoring on metaphor conventionalization. Such subjective annotation correlates with word frequency (Do Dinh et al., 2018; Momen et al., 2026) because human annotators tend to assign higher novelty scores to rare words. As a result, each individual annotator's exposure to less frequent vocabulary mingles with assessment of metaphor novelty. Less frequent words, however, can have perfectly conventionalized metaphoric senses listed in the dictionaries, while relatively common words may be used creatively. Our dictionary-based approach, even though time and effort-consuming, helps to avoid this confound.

**Source of sentences.** We present metaphorically motivated polysemes in their full sentence context to observe their naturalistic use in texts.

these patterns, as well as the list of patterns that remain to be annotated can be consulted in Appendices A and B.

[2]OED: oed.com, AHD: ahdictionary.com

Linguistic experiments often use handcrafted sentences that are tailored to the narrow purposes of each study (e.g. Cardillo et al., 2010, 2017; Lai et al., 2009). While offering a general high quality and controlled conditions of a manually created resource, these can potentially introduce bias. Momen et al. (2026) evaluate a series of language models on various pre-existing conventional metaphor datasets and find that results from synthetic data differ from the extracted data, as manually created sentences are intentional in sense distinction by lexical means. With this in mind, we opted for sentence extraction from a general language corpus (British National Corpus — BNC) (BNC Consortium, 2007) that represents naturalistic language use as opposed to synthetic data. In cases where we cannot draw enough material from BNC, we revert to Wikipedia dump [3] and Web corpora (Goldhahn et al., 2012).

**Size.** We selected 10 regular metaphor patterns and 10 word instantiations per pattern. As we aim for the explorations with LLMs, we draw at least 50 sentences for each word sense (the base sense and the derived metaphorical sense), so we have enough material for sensible experimentation and evaluation (see Table 2 for a sample of our dataset). To study regular metonymy and regular metaphor in comparison, we structure our dataset in a way that makes it compatible with the regular metonymy dataset by Alonso et al. (2013) and its version expanded and adapted by Li and Armstrong (2024).

## 4. Annotation

**Sense annotation.** We annotate the extracted sentences to identify the senses in which the target words are used (literal base sense and derived metaphorical sense of the pattern). Suitable sentences unambiguously evoke the sense of interest, without needing to revert to any wider context. Vague sentences that do not fulfill this condition are discarded. We plan to engage crowdsourced annotators to validate the initial sense annotation delivered by a language professional.

**Frequency annotation**. We sense-annotate all extracted sentences, whether or not they are suitable to be included in the dataset, in order to collect statistics on sense frequency. Usually, word frequency is controlled for in the experimentation with conventional metaphors, but, for the reasons explained above, it may not be an optimal indicator. In addition, due the unnatural, morphologically unmotivated tokenization of infrequent words, neural language models too show different performance on metaphors with low and high word frequency (Neidlein et al., 2020).

[3] dumps.wikimedia.org

A corpus-derived metric of sense frequency, on the other hand, can open interesting experimental venues by showing in what sense the word is used more often — literal or metaphoric. Klepousniotou et al. (2008) show that very conventionalized metaphoric senses become more dominant than the literal senses. Corpus-derived sense dominance can potentially be a good predictor of metaphor conventionalization, the relation that we aim to investigate in the future work. Unfortunately, existing resources on word sense frequency are scarce and have a small coverage (e.g. Bennett et al., 2016; Liu et al., 2024). One of our contributions will be an estimate of the sense frequency of each target word in the scope of the patterns and as distributed in BNC.

**Annotation difficulties.** For highly polysemous words with closely related, overlapping senses, identifying the base sense of a metaphoric extension can be challenging. We generally relied on the standard Metaphor Identification Procedure (Pragglejaz Group, 2007), but more extensive lexicological analyses were often needed. When complexity of the sense structure and the degree of word's polysemy are high, metaphoric sense annotation becomes a very time-consuming task.

Unlike for metonymy, sense annotation for metaphor is not complicated by underspecification, when several senses can be active at the same time (Alonso et al., 2013; Frisson, 2009). Metonymy operates on conceptual domain contiguity, but metaphor implies domain disjunction: although senses are related, they are conceptually far from each other and easier to disambiguate.

## 5. Discussion

The sense annotation procedure implies assigning a discrete, individual sense to a word in context. In practice, however, we had to deal with closely related, overlapping senses that were not easy to delineate. As can be seen with the example of *roar* in Table 1, such senses have been grouped together based on their similarity. The case of metaphor, in particular, brings another difficulty with it: varying underlying motivations for sense extension are grouped in the same sense category. For instance, the metaphorically motivated sense of the word *ripple* describes a manner in which events occur or unfold based on the similarity with disturbed water. It some cases, it maps dynamic visual imagery on another image (*a thing resembling a ripple or ripples in appearance or movement*), but in other, it connects it to sound (*a gentle rising and falling sound that spreads through a group of people*) and psychological states (*a particular feeling or effect that spreads through someone or something*) (sense descriptions from OED).

Consequently, even though the resulting annotation suggests discrete, uniform meaning classes, there is variation and continuity within them. This continuity can be encoded in language model representations and revealed with probing experiments by observing the internal activations in their relation to the variability of meanings in metaphorical patterns. Despite recent linguistic arguments for continuity in polysemy, the standard NLP experiments rather align with the earliest view of mental lexicon organization — sense enumeration approach (Katz and Fodor, 1963) — and highly discrete, often binary categories (polysemous vs. monosemous, regular vs. irregular, figurative vs. literal). Recent contextual language models offer the means to change this paradigm.

## 6. Conclusions and Future Work

This work-in-progress contribution presents a novel dataset framing metaphor as a regular sense extension of a polysemous word. Although rarely approached from this perspective, regular metaphor can contribute to the theoretical, cognitive and computational linguistic investigation of the lexical meaning structure and representation, as well as conventionalization of figurative speech. We also contribute with data on sense frequency which is of particular importance for polysemous words but largely unavailable.

We plan to use the dataset in the experimentation with neural language models. The distributional, continuous nature of these systems makes them a useful device to study polysemy and figurativeness in their complexity and heterogeneity. We continue the work on the dataset aiming to cover 10 metaphoric polysemy patterns with expert linguistic labeling, as well as subsequent crowd-sourced annotation for additional validation.

## 7. Limitations

Several aspects of dataset design and annotation can potentially introduce bias into the data. Manually picked sentences that very clearly realize the target sense do not faithfully represent real language use where ambiguity and vagueness are ubiquitous. This might limit our capacity to generalize insights derived from our data to language in general. A similar concern is related to the words that have been discarded from the analysis because of the difficulty to identify their base sense, or cases where metonymy and metaphor are intertwined. Discarding uncomfortable cases that do not fit into a theoretical framework distorts the picture by simplifying the complex and diverse linguistic landscape we are dealing with.

## Acknowledgments

This work was funded by MCIU/AEI/10.13039/501100011033 (grants with references PID2021-128811OA-I00, PRE2022-102762, CNS2024-154902, PID2024-161928OB-I00, and AIA2025-163322-C62), and by the Galician Government (ERDF 2024-2027: Call ED431G 2023/04, and ED431B 2025/16).

## 8. Bibliographical References

Héctor Martínez Alonso, Bolette Sandford Pedersen, and Núria Bel. 2013. Annotation of regular polysemy and underspecification. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 725–730.

Jurij D. Apresjan. 1974. Regular polysemy. *Linguistics*, 12(142):5–32.

Lucie Barque and François-Régis Chaumartin. 2009. Regular polysemy in WordNet. *Journal for language technology and computational linguistics*, 24(2):5–18.

Marina Berri. 2020. Polisemia regular por metáfora. *Cuad. lingüíst. hisp.*, (35):19–36.

Gemma Boleda. 2025. LLMs as a synthesis between symbolic and distributed approaches to language. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 9365–9379, Suzhou, China. Association for Computational Linguistics.

Eileen R Cardillo, Christine E Watson, Gwenda L Schmidt, Alexander Kranjec, and Anjan Chatterjee. 2012. From novel to familiar: tuning the brain for metaphors. *Neuroimage*, 59(4):3212–3221.

Erik-Lân Do Dinh, Hannah Wieland, and Iryna Gurevych. 2018. Weeding out conventionalized metaphors: A corpus of novel metaphor annotations. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 1412–1424, Stroudsburg, PA, USA. Association for Computational Linguistics.

Kevin Durkin and Jocelyn Manning. 1989. Polysemy and the subjective lexicon: Semantic relatedness and the salience of intraword senses. *Journal of Psycholinguistic Research*, 18:577–612.

Markus Egg and Valia Kordoni. 2022. Metaphor annotation for German. In *Proceedings of the Thirteenth Language Resources and Evaluation Con-*

*ference*, pages 2556–2562, Marseille, France. European Language Resources Association.

Abed Alhakim Freihat, Fausto Giunchiglia, and Biswanath Dutta. 2013. Regular polysemy in wordnet and pattern based approach. *International Journal On Advances in Intelligent Systems*, 6:199–212.

Steven Frisson. 2009. Semantic underspecification in language processing. *Language and Linguistics Compass*, 3(1):111–127.

Richard Futrell and Kyle Mahowald. 2025. How linguistics learned to stop worrying and love the language models. *arXiv [cs.CL]*.

Aina Garí Soler and Marianna Apidianaki. 2021. Let's play mono-poly: BERT can reveal words' polysemy level and partitionability into senses. *Transactions of the Association for Computational Linguistics*, 9:825–844.

Joseph Grady, Todd Oakley, and Seana Coulson. 1999. *Blending and Metaphor*, Metaphor in Cognitive Linguistics, pages 101–124. John Benjamins Publishing Company, New York, Chichester. ISBN 9789027236814.

Jumbly Grindrod. 2024. Transformers, contextualism, and polysemy. *arXiv [cs.CL]*.

Janosch Haber and Massimo Poesio. 2020. Assessing polyseme sense similarity through co-predication acceptability and contextualised embedding distance. In *Proceedings of the Ninth Joint Conference on Lexical and Computational Semantics*, pages 114–124.

Janosch Haber and Massimo Poesio. 2024. Polysemy — evidence from linguistics, behavioral science, and contextualized language models. *Comput. Linguist. Assoc. Comput. Linguist.*, 50(1):1–67.

Jerrold Katz and Jerry Fodor. 1963. The structure of a semantic theory. *Language*, 39:170–210.

Ekaterini Klepousniotou, Debra Titone, and Carolina Romero. 2008. Making sense of word senses: the comprehension of polysemy depends on sense overlap. *J. Exp. Psychol. Learn. Mem. Cogn.*, 34(6):1534–1543.

Vicky Tzuyin Lai, Tim Curran, and Lise Menn. 2009. Comprehending conventional and novel metaphors: an ERP study. *Brain Res.*, 1284:145–155.

George Lakoff and Mark Johnson. 1980. *Metaphors We Live By*. University of Chicago Press, Chicago, IL.

Ronald W Langacker. 2017. Entrenchment in cognitive grammar. In *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge*, pages 39–56. American Psychological Association, Washington.

Jiangtian Li. 2024. Semantic minimalism and the continuous nature of polysemy. *Mind Lang.*

Jiangtian Li and Blair C Armstrong. 2024. Probing the representational structure of regular polysemy via sense analogy questions: Insights from contextual word vectors. *Cogn. Sci.*, 48(3):e13416.

Jiangtian Li and Marc F Joanisse. 2021. Word senses as clusters of meaning modulations: A computational model of polysemy. *Cogn. Sci.*, 45(4):e12955.

Alizée Lombard, Richard Huyghe, Lucie Barque, and Doriane Gras. 2023. Regular polysemy and novel word-sense identification. *The Mental Lexicon*, 18(1):94–119.

Alizée Lombard, Anastasia Ulicheva, Maria Korochkina, and Kathy Rastle. 2024. The regularity of polysemy patterns in the mind: Computational and experimental data. *Glossa Psycholinguistics*, 3(1).

Omar Momen, Emilie Sitter, Berenike Herrmann, and Sina Zarrieß. 2026. Surprisal and metaphor novelty: Moderate correlations and divergent scaling effects. *arXiv [cs.CL]*.

Arthur Neidlein, Philip Wiesenbach, and Katja Markert. 2020. An analysis of language models for metaphor recognition. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 3722–3736, Stroudsburg, PA, USA. International Committee on Computational Linguistics.

Sanni Nimb and Bolette Sanford Pedersen. 2000. Treating metaphoric senses in a danish computational lexicon –different cases of regular polysemy. In *Proceedings of the 9th EURALEX International Congress*, pages 679–691, Stuttgart, Germany. Institut für Maschinelle Sprachverarbeitung.

Wim Peters and Ivonne Peters. 2000. Lexicalised systematic polysemy in WordNet. In *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC'00)*, Athens, Greece. European Language Resources Association (ELRA).

Alain Polguère. 2018. A lexicographic approach to the study of copolysemy relations. *Russ. J. Linguist.*, 22(4):788–820.

Alain Polguère. 2022. Fishing out regular polysemy with the gener (generic) relation. *Lexique*, (31).

Pragglejaz Group. 2007. MIP: A method for identifying metaphorically used words in discourse. *Metaphor Symb.*, 22(1):1–39.

James Pustejovsky. 1991. The Generative Lexicon. *Computational Linguistics*, 17(4):409–441.

Hugh Rabagliati and Jesse Snedeker. 2013. The truth about chickens and bats: ambiguity avoidance distinguishes types of polysemy: Ambiguity avoidance distinguishes types of polysemy. *Psychol. Sci.*, 24(7):1354–1360.

W Gudrun Reijnierse, Christian Burgers, Tina Krennmayr, and Gerard J Steen. 2018. DMIP: A method for identifying potentially deliberate metaphor in language use. *Corpus Pragmat.*, 2(2):129–147.

Sebastian Reimann and Tatjana Scheffler. 2024. When is a metaphor actually novel? annotating metaphor novelty in the context of automatic metaphor detection. In *Proceedings of the 18th Linguistic Annotation Workshop (LAW-XVIII)*, pages 87–97, St. Julians, Malta. Association for Computational Linguistics.

Mahesh Srinivasan and Hugh Rabagliati. 2015. How concepts and conventions structure the lexicon: Cross-linguistic evidence from polysemy. *Lingua*, 157:124–152.

Gerard J Steen, Aletta G Dorst, J Berenike Herrmann, Anna A Kaal, and Tina Krennmayr. 2010a. Metaphor in usage. *Cogn. Linguist.*, 21(4):765–796.

G.J. Steen, A.G. Dorst, J.B. Herrmann, A.A. Kaal, T. Krennmayr, and T. Pasma. 2010b. *A method for linguistic metaphor identification. From MIP to MIPVU.* Number 14 in Converging Evidence in Language and Communication Research. John Benjamins.

Sean Trott and Benjamin Bergen. 2023. Word meaning is both categorical and continuous. *Psychol. Rev.*, 130(5):1239–1261.

Gregor Wiedemann, Steffen Remus, Avi Chawla, and Chris Biemann. 2019. Does bert make any sense? interpretable word sense disambiguation with contextualized embeddings. In *Proceedings of the 15th Conference on Natural Language Processing (KONVENS 2019)*.

George Kingsley Zipf. 1945. The meaning-frequency relationship of words. *The Journal of general psychology*, 33:251–6.

## 9. Language Resource References

Bennett, Andrew and Baldwin, Timothy and Lau, Jey Han and McCarthy, Diana and Bond, Francis. 2016. *LexSemTm: A Semantic Dataset Based on All-words Unsupervised Sense Distribution Learning*. Association for Computational Linguistics. PID https://aclanthology.org/P16-1143/.

BNC Consortium. 2007. *The British National Corpus, version 3 (BNC XML Edition)*. Oxford Text Archive. PID http://hdl.handle.net/20.500.14106/2554.

Cardillo, Eileen R and Schmidt, Gwenda L and Kranjec, Alexander and Chatterjee, Anjan. 2010. *Stimulus design is an obstacle course: 560 matched literal and metaphorical sentences for testing neural hypotheses about metaphor*. Springer Science and Business Media LLC. PID https://doi.org/10.3758/BRM.42.3.651.

Cardillo, Eileen R and Watson, Christine and Chatterjee, Anjan. 2017. *Stimulus needs are a moving target: 240 additional matched literal and metaphorical sentences for testing neural hypotheses about metaphor*. Springer. PID https://doi.org/10.3758/s13428-016-0717-1.

Fellbaum, Christiane. 1998. *WordNet: An Electronic Lexical Database*. The MIT Press. PID https://doi.org/10.7551/mitpress/7287.001.0001.

Goldhahn, Dirk and Eckart, Thomas and Quasthoff, Uwe. 2012. *Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages*. European Language Resources Association (ELRA). PID http://www.lrec-conf.org/proceedings/lrec2012/pdf/327_Paper.pdf.

Liu, Lei and Gong, Tongxi and Shi, Jianjun and Guo, Yi. 2024. *A high-frequency sense list*. Frontiers Media SA. PID https://doi.org/10.3389/fpsyg.2024.1430060.

Miller, George A. 1995. *WordNet: a lexical database for English*. Association for Computing Machinery. PID https://doi.org/10.1145/219717.219748.

# Appendices

## A. Pattern List

Below we list polysemy patterns covered by the present study.

1. ANIMAL - ARTIFACT
2. ANIMAL - PERSON
3. ANIMAL CRY - COMMUNICATION
4. ARTIFACT - MESSAGE
5. BODY PART - OBJECT PART
6. NATURAL EVENT - WAY OF OCCURRING
7. NATURAL EVENT - SOCIAL EVENT
8. PERSON - ANIMAL
9. PERSON - ARTIFACT
10. PHYSICAL PROPERTY - PSYCHOLOGICAL PROPERTY

## B. Sense Inventories

Two of the patterns listed in the Appendix A have been annotated: ANIMAL CRY - COMMUNICATION and NATURAL PHENOMENON - WAY OF OCCURRING. Tables 3 and 4 present sense inventories for these patterns as covered by the annotated words. The sense descriptions are mainly derived from the OED. In certain cases, we draw additional material from AHD:

- when OED offers a circular definition that reuses the term being defined (or a morphologically related word) in the meaning description. For instance, the metaphorical sense of the verb *shower* is described by OED using its nominal counterpart:

  > *(of a mass of small things) fall or be thrown in a **shower**: bits of broken glass **showered** over me.*

- in a similar situation to the above, but when a synonym is used as a sense description, especially if this synonym is one of the target lexemes under study. For example, the literal sense of the verb *inundate* is simply defined as

  > *flood: the islands may be the first to be **inundated** as sea levels rise*.

- when OED merges the base literal sense with the derived metaphorical sense in a single sense description, as, for instance, in the case of the verb *bellow*:

  > *(of a person or animal) emit a deep loud roar, typically in pain or anger: he **bellowed** in agony.*

- when AHD lists more pattern-related senses than OED does, which is also reflected in the corpus annotation. This is the case for such words as *howl*, *riffle* and *torrent*.

| Word | PoS | Literal | Metaphorical | Source |
|---|---|---|---|---|
| bark | V | (of a dog, fox, or seal) give a bark | (of a person) make a sound resembling a bark; utter (a command or question) abruptly or aggressively; call out in order to sell or advertise something | OED |
| | N | the sharp explosive cry of a dog, fox, or seal | a sound resembling a bark, typically one made by someone laughing or coughing | OED |
| bellow | V | To make the deep roaring sound characteristic of a bull. | sing (a song) loudly and tunelessly; shout something with a deep loud roar | OED |
| | N | The roar of a large animal, such as a bull. | A very loud utterance or other sound. | AHD |
| bleat | V | (of a sheep, goat, or calf) make a characteristic weak, wavering cry | speak or complain in a weak, querulous, or foolish way | OED |
| | N | the weak, wavering cry made by a sheep, goat, or calf | a person's weak or plaintive cry; a complaint | OED |
| bray | V | (of a donkey or mule) utter a bray | (of a person) speak or laugh loudly and harshly | OED |
| | N | the loud, harsh cry of a donkey or mule | a sound, voice, or laugh resembling a bray | OED |
| cackle | V | (of a bird, especially a hen or goose) give a raucous clucking cry | laugh in a loud, harsh way; talk at length without acting on what is said | OED |
| | N | the raucous clucking cry of a bird such as a hen or a goose | a loud, harsh laugh | OED |
| chirp | V | (of a small bird or an insect) make a short, sharp, high-pitched sound | (of a person) say something in a lively and cheerful way | OED |
| | N | A short, high-pitched sound, such as the one a small bird or insect makes | To make a short, high-pitched sound. | AHD |
| growl | V | (of an animal, especially a dog) make a low guttural sound in the throat | (of a person) say something in a low grating voice, typically in a hostile or angry | OED |
| | N | a low guttural sound made in the throat by a hostile dog or other animal | low guttural sound or utterance made by a person, especially to express hostility or anger | OED |
| grunt | V | (of an animal, especially a pig) make a low, short guttural sound | (of a person) make a low inarticulate sound, typically to express effort or indicate assent | OED |
| | N | a low, short guttural sound made by an animal or a person | a low, short guttural sound made by an animal or a person | OED |
| howl | V | make a howling sound | weep and cry out loudly; to laugh heartily | OED+AHD |
| | N | a long, doleful cry uttered by an animal such as a dog or wolf. | a loud cry of pain, fear, anger, or amusement; | OED |
| roar | V | (of a lion or other large wild animal) utter a full, deep, prolonged cry | (of a person or crowd) utter a loud, deep, prolonged sound, typically from anger, pain, or excitement; utter or express in a loud tone; (of a crowd) encourage (someone) to do something by loud shouts or cheering; laugh loudly | OED |
| | N | a full, deep, prolonged cry uttered by a lion or other large wild animal | a loud, deep sound uttered by a person or crowd, generally as an expression of pain, anger, or approval; a loud outburst of laughter | OED |

Table 3: Sense inventory for the pattern ANIMAL CRY - COMMUNICATION. The table presents senses (*Literal* and *Metaphoric*) covered by each *Word*, as well as the *Source* of the definitions. *PoS* marks either a verb (*V*) or a noun (*N*).

| Word | PoS | Literal | Metaphoric | Source |
|---|---|---|---|---|
| backwash | N | the motion of receding waves; a backward current created by an object moving through water or air | the unpleasant after-effects of an event | OED |
| deluge | V | overwhelm with a flood | to overwhelm with a large number or amount | OED |
| | N | a severe flood; a heavy fall of rain | a great quantity of something arriving at the same time | OED |
| flood | V | cover or submerge (an area) with water in a flood; become covered or submerged by a flood; become swollen and overflow (its banks) | arrive in overwhelming amounts or quantities; completely fill or suffuse; overwhelm with large amounts or quantities | OED |
| | N | an overflow of a large amount of water beyond its normal limits, especially over what is normally dry land; the inflow of the tide | an outpouring of tears; an overwhelming quantity of things or people happening or appearing at the same time | OED |
| inundation | N | flooding | an overwhelming abundance of people or things | OED |
| inundate | V | to cover with water, especially floodwaters. | overwhelm (someone) with things or people to be dealt with | OED+AHD |
| rain | V | rain falls; (of the sky, the clouds, etc.) send down rain | fall or cause to fall in large or overwhelming quantities; | OED |
| | N | the condensed moisture of the atmosphere falling visibly in separate drops; falls of rain | a large or overwhelming quantity of things that fall or descend | OED |
| riffle | V | to flow in rough waves or become choppy, as water. | to thumb through (the pages of a book, for example); to shuffle cards; | OED |
| | N | a rocky or shallow part of a stream or river where the water flows brokenly; a stretch of choppy water caused by such a shoal or sandbar; a rapid; a wave or ripple in such water. | an act or sound of riffling through something; The act or an instance of shuffling cards; disturb the surface of; ruffle: | OED+AHD |
| ripple | V | (of water) form or flow with a series of small waves on the surface; cause (the surface of water) to form small waves | move in a way resembling a series of small waves; (of a sound or feeling) spread through a person, group, or place | OED |
| | N | a small wave or series of waves on the surface of water, especially as caused by a slight breeze or an object dropping into it; | a thing resembling a ripple or ripples in appearance or movement; a gentle rising and falling sound that spreads through a group of people; a particular feeling or effect that spreads through someone or something | OED |
| shower | V | - | to throw or cause small things or pieces to fall over; throw (a number of things) all at once towards someone; give someone a great number of (things); give a great number of things to (someone) | OED+AHD |
| | N | a brief and usually light fall of rain, hail, sleet, or snow | a mass of small things falling or moving at the same time; a large number of things happening or given at the same time | OED |
| torrent | N | a strong and fast-moving stream of water or other liquid; a heavy downpour; a deluge. | an overwhelming outpouring of (something, typically words) | OED+AHD |
| wave | V | - | move to and from with a swaying motion while remaining fixed to one point | OED |
| | N | a long body of water curling into an arched form and breaking on the shore; a ridge of water between two depressions in open water | a shape regarded as resembling a breaking wave; a sudden occurrence of or increase in a phenomenon, feeling, or emotion | OED |
| roar | V | (of a lion or other large wild animal) utter a full, deep, prolonged cry | (of a person or crowd) utter a loud, deep, prolonged sound, typically from anger, pain, or excitement; utter or express in a loud tone; (of a crowd) encourage (someone) to do something by loud shouts or cheering; laugh loudly | OED |
| roar | N | a full, deep, prolonged cry uttered by a lion or other large wild animal | a loud, deep sound uttered by a person or crowd, generally as an expression of pain, anger, or approval; a loud outburst of laughter | OED |

Table 4: Sense inventory for the pattern NATURAL PHENOMENON - WAY OF OCCURRING. The table presents senses (*Literal* and *Metaphoric*) covered by each *Word*, as well as the *Source* of the definitions. *PoS* marks either a verb (*V*) or a noun (*N*).