

Methods in Lexical Semantics: Metaphor and Meaning

Francis Bond ♠ (and Rowan Hall Maudslay ♦)

♠ Departments of Asian Studies and Linguistics and Sinofon Project
Palacký University Olomouc

♦ CultureLab
Université PSL

bond@ieee.org

LOT Winter School 2026



Faculty
of Arts

Metaphor and Metonymy

- Metonymy and metaphor are two fundamental tropes (figures of speech)
- Both play crucial roles in the way we understand and use language
- **Metonymy** involves the substitution of one term for another to which it is closely related
 - ▶ *I ate a chestnut*
 - ▶ *I pruned the chestnut [tree]*
 - ▶ *This table is made from chestnut [wood]*
 - ▶ Metonymy works by contiguity or association between concepts, usually by some shared attribute, or by a part-whole relation
- **Metaphor** involves understanding one concept in terms of another, often unrelated, concept.
 - ▶ *This horse has a chestnut on its leg* “a small horny bump”
 - ▶ Metaphors create new concepts in the target domain by mapping elements from the source domain
- They are crucial in understanding novel uses of words



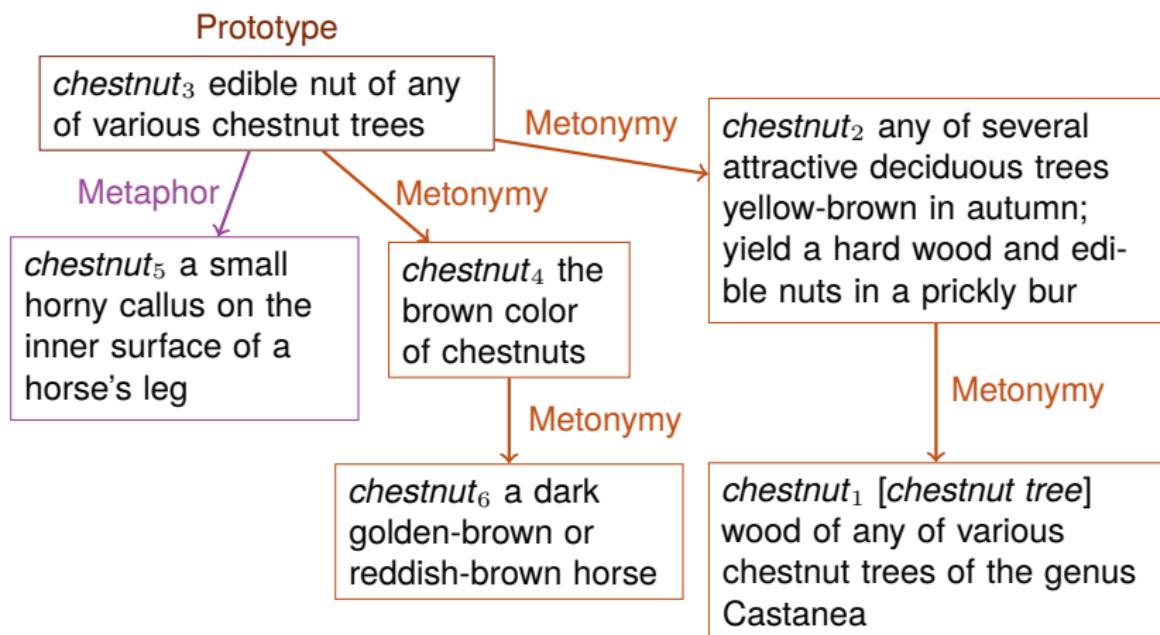
- ChainNet is an attempt to comprehensively model these tropes
(Maudslay et al., 2024; Bond and Maudslay, 2025)
 - ▶ All nouns (word + synset combinations) in the Princeton Wordnet (Fellbaum, 1998) with three or more senses* were annotated
 - ▶ Every sense linked to another sense, or treated as a homonym (unrelated)
 - ▶ Most annotated by 1 annotator, some by 2 annotators
 - Inter-annotator agreement 70% (88% for same prototype)
 - Intra-annotator agreement 81% (92% for same prototype)
 - ▶ Originally grew out of work in determining homonyms — which senses of a word are related

* almost, and some with two senses. All should eventually be annotated

** if anyone knows where the Italian Metaphor Database is, please let us know



chestnut



Can we see how English metaphors are used cross-lingually?

- For example, are translations of linked pairs likely to be translated with the same word
 - ▶ If so then metaphor/metonymy holds cross language
 - ... or this may indicate issues with automatic construction of wordnets
- For example *head* “body part” is metaphorically extended to “person in charge” in English and this is also true in Japanese (頭), Italian (*capo*) and many other languages
- But not all metaphors are shared: *head* “main part of a grammatical constituent” is not 頭 in Japanese
- ChainNet is made for English, but we can expect many of the tropes to work with other languages
- And we have many wordnets linked to English in the Open Multilingual Wordnet



Measure the translation overlap with OMW

- For every English word that has been annotated
 - ▶ For each pair of English senses
 - ▶ Look up all translations of each sense
 - ▶ measure the overlap (we use Jaccard similarity)
 - ▶ Store and note if the senses are linked or not
- So for example *head* has 3 senses (really many more)
 - ① head, caput
ヘッド, 頭, 頭部
 - ② head, chief, top dog
大頭, 主任者, 御頭, 頭領, 頭
 - ③ head, head word
主要語
- Overlap
 - ▶ 1-2: 0.125
 - ▶ 2-3: 0
 - ▶ 1-3: 0



We do this for all the lexicons in OMW-1.4 I

Lang	Unlinked	Metaphor	Metonymy	Non-Zero Instances
sv	0.70	1.71	2.10	417
he	0.69	1.58	2.31	472
eu	0.82	1.30	1.80	5,562
ja	0.89	1.11	1.55	8,130
ro	0.83	1.30	1.71	6,429
sq	0.74	1.27	2.38	822
da	0.74	1.46	2.09	332
lt	0.63	1.46	2.86	692

- **unlinked** has no link
- **metaphor** is linked by metaphor
- **metonymy** is linked by metonymy

Normalize by dividing by the average distance for all senses.



We do this for all the lexicons in OMW-1.4 II

Lang	Unlinked	Metaphor	Metonymy	Non-Zero Instances
es	0.85	1.17	1.79	5,915
hr	0.81	1.22	1.98	3,576
fr	0.94	1.10	1.25	18,975
nb	0.73	1.49	2.12	339
bg	0.73	1.54	2.11	380
gl	0.70	2.13	1.55	271
cmn	0.74	1.57	1.97	1,347
it	0.80	1.39	1.81	4,614
arb	0.74	1.28	2.38	1,337
pl	0.75	1.43	2.07	1,326
el	0.72	1.30	2.45	1,274
iwn	0.79	1.46	1.80	968
nn	0.73	1.52	2.10	340
id	0.87	1.21	1.58	9,803

We do this for all the lexicons in OMW-1.4 III

Lang	Unlinked	Metaphor	Metonymy	Non-Zero Instances
is	0.73	1.51	2.13	388
ca	0.83	1.23	1.84	6,000
sk	0.73	1.35	2.31	2,067
fi	0.82	1.35	1.76	7,971
nl	0.79	1.38	1.90	2,991
sl	0.81	1.41	1.74	6,243
th	0.80	1.48	1.71	2,336
pt	0.80	1.26	1.96	5,414
zsm	0.87	1.22	1.60	10,272
Mean	0.78	1.39	1.96	3,774.3

Unlinked share fewer translations, metaphor shares more and metonymy shares much more.

Metonymy is more likely to hold cross linguistically!

Discussion

- The wordnets are of vastly different sizes: the number of sense pairs that have a translation varies from 271 (Galician) to 18,975 (French).
- Only Galician has the score for metaphor (2.13) larger the score for metonymy (1.55), probably due to data sparsity. We would expect it to behave much like Portuguese (1.26/1.96).
- 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 for all senses have all relevant translations, ...
- 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



Translations of the senses of *cherry*

Sense	Mandarin	Japanese	Finnish	Italian
<i>cherry</i> ₁	樱桃木	桜, 桜材	kirsikkapuumetsä	ciliegio
<i>cherry</i> ₂	樱桃树	桜, 櫻	kirsikkapuu	ciliegio
<i>cherry</i> ₃	樱桃	樱桃, 桜ん	kirsikka	cerasa, ciliegia
<i>cherry</i> ₄	樱桃红	桜ん	kirsikanpunainen	ciliegia, rosso ciliegia

- When we look at the translations, we see that the metonymy that is unmarked in English, is marked in other languages
- Even within English, other synonyms have information



Can we use other lemmas to classify links

- Synsets have sets of synonyms, so the other senses may include useful information
- Consider the synsets with a sense of chestnut
 - s1 chestnut “edible nut of any of various chestnut trees of the genus *Castanea*” (NATURAL_BODY)
 - s2 chestnut, chestnut tree “any of several attractive deciduous trees yellow-brown in autumn; yield a hard wood and edible nuts in a prickly bur” (PLANT)
 - s3 chestnut “wood of any of various chestnut trees of the genus *Castanea*” (CHEMICAL!)
- These are annotated as s1 → s2 → s3 (metonymy)
- If we look at all the senses
 - ▶ s1 → s2 has a diff: “+ tree”
 - ▶ s2 → s3 has a diff: “- tree”
- These patterns repeat ...

Cherry → Cherry Tree → Cherry

They can be found across languages

- s1 → s2: kastanje → kastanjetræ (da)
- s1 → s2: chestnut → chestnut tree (en)
- s1 → s2: marron → marronnier (fr)
- s1 → s2: kastanja → kastanjapuu (fi)
- s1 → s2: kastanje → kastanjeboom (nl)
- s1 → s2: castană → castan (ro)
- s1 → s2: castanăcomestibilă → castan (ro)



Ongoing Research to Find Patterns

- We have made a searchable database where the edges are decorated by the morphological differences
 - ▶ So we can see all pairs linked by `en:+ tree` or `da:+træ`
 - ▶ The nodes link to synsets, so we can look up supertypes: we can say this link is of type **FOOD:PLANT** or **PLANT IS FOOD** in the cognitive linguistic style
 - ▶ In this way we can group the tropes together in families
 - ▶ Similar to **Stoyanova et al. (2025)**
- Other people have done this top down (**Khishigsuren et al., 2022**, e.g start at **FOOD:PLANT** and looked for examples)
- With ChainNet we can go bottom-up and find many more patterns
 - ▶ We can use multi-lingual features
 - ▶ We need to investigate methods of clustering
 - ▶ We need to find a good way of evaluating
- We would like to also map to existing Conceptual Metaphors



The Patterns can be used Generatively

- If we see a word that is of type **FOOD**, then we can expect that there may be a kind of plant using the same word
 - ▶ Useful for analyzing texts that otherwise don't make sense
 - ▶ Can be used to analyse humour and word play



BBC Spaghetti Tree hoax (1957)



Faculty
of Arts

Metaphor in Large Language Models

- Because these tropes are so established, LLMs capture the patterns well, and can generalise to novel examples, at least for languages with a lot of online data
- E.g. For Spanish, **Puraivan et al. (2024)** found ChatGPT could differentiate literal from metaphorical senses with an accuracy of 85-88% (on a small dataset)
 - ▶ This is high enough to cause **automation blindness**
 - ▶ It is very hard for a human to spot mistakes
- We expect that this accuracy will go down for languages with less data and for less common senses, but have not yet confirmed this
- There are also interesting questions of how much metaphor interpretation in multilingual models is effected by other languages
The DanNet group have found metaphors in Danish not found in English are less likely to be recognized
- So we feel it is important to keep working on producing reliable manually annotated data



Evaluating LLM-Generated Explanations of Metaphors (Pedersen et al., 2025)

- Metaphors have culturally distinct imagery and semantics (e.g., agricultural or national norms).
- Standard benchmarks (often English-centric) may hide biases against medium-resource languages.
- LLMs may transfer English knowledge inappropriately to other languages.

Models evaluated:

- **ChatGPT-4o mini** via web/API (multilingual, but English-heavy training)
- **Llama 3.1 405B**, large open model (multilingual but likewise English-biased).

Tested on a culturally sensitive evaluation dataset of Danish metaphors & model explanations.



Dataset and Experimental Design

Metaphor dataset: 150 Danish expressions

- 75 single-word metaphors + 75 multiword idioms
- Divided into:
 - ▶ **Culture-specific** Danish metaphors (e.g., farming/nautical, social norms)
 - ▶ **Cross-cultural** ones shared with English
- Validated by native speakers and bilingual informants

Prompting setup:

- Four prompt variants:
 - ▶ Isolation vs. context
 - ▶ Danish vs. English prompt language
- Each model generated 600 explanations (all combinations)

Evaluation:

- Human expert judges explanations on a 4-point qualitative scale



Faculty
of Arts

Key Findings and Discussion

Main results:

- Both models explain **cross-cultural metaphors** far better than culture-specific ones
- Models show a strong **English bias** — higher quality explanations are produced when metaphors overlap with English
- Culture-specific sentiment and nuance are often lost in translation

Implications:

- Bias arises not only from data distribution but also from *erroneous transfer* of English metaphorical knowledge.
- Culturally grounded metaphor understanding remains a challenge for current multilingual LLMs
- Highlights risks for linguistic diversity and under-resourced languages in AI



LLMs and metaphor: a success and a failure

(1) *Han skød hendes argumenter ned.*

Danish

He shot down her arguments.

- Correctly identifies metaphorical usage
- Maps **physical attack** → **argumentation**
- Produces fluent, accurate explanation

(1) *Han har rent mel i posen.*

Danish

He has nothing to hide (Lit: He has clean flour in the bag.)

- Plausible but wrong English-based metaphor
“This metaphor refers to being prepared or well-supplied.”
- Over-literal explanation
“Clean flour symbolizes purity of intentions” but without connecting it to social trust / transparency
- Misses social–cultural grounding



Faculty
of Arts

Future Work

- We are interested in looking at the relation between in-context uses and tropes
Using sense-tagged corpora we can see
 - ▶ If basic uses or extended uses are more common
 - ▶ Do similar tropes cluster?
 - ▶ Do different genres use different tropes?
- We would like to do large scale annotation on another language (probably start with Japanese)
- We are looking in detail at a small sample: currently investigating **ANIMAL:PERSON** metaphors
 - ▶ *Be a lamb and give me a hand, ...*
 - ▶ In Chinese, Czech, English, French, Indonesian, Japanese, Javanese, Korean, Slovak and Vietnamese — it would be good to include some sign
- We would like to investigate similarities and differences in images,

...



Faculty
of Arts

Summary

- The combination of detailed annotation of meaning and cross-lingual links allow us to explore differences in meaning representation
 - ▶ Metonymy is based on association and contiguity
more universal
 - ▶ Metaphor often requires analogical reasoning and cultural framing
more culture and language specific
- We don't know if the tropes are independently discovered, come from a common ancestor, or are borrowed (the same problem as with cognates in etymology)
- We need to add to wordnet and make searchable
 - ▶ Add as sense-level relations, not etymology, contra **Khan et al. (2020)**
- LLMs may capture these regularities, but are hard to explore with



Acknowledgements

- Thanks to Bolette Pederson and the DanNet group for some very useful discussions
- ChainNet is primarily the work of Rowan Hall-Maudslay
r.hallmaudslay@gmail.com
<https://github.com/rowanhm/ChainNet>
- The Open Multilingual Wordnet brings together wordnets made by many projects, we thank them all
- Data and code for the cross-lingual analysis is available at
<https://github.com/bond-lab/chainnet-xling>
(still changing a lot)



- The OMW is a collection of wordnets for different languages.
- These languages are linked through the collaborative interlingual index (CILI; **Bond et al., 2016**), a language-neutral list of concepts that is linked to by the various wordnets.
- 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 Wintner, 2007**), Indonesian and Malaysian (**Nurril Hirfana 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**).
- Data is available from <https://github.com/omwn/omw-data>



References I

Francis Bond and Rowan Hall Maudslay. 2025. Metonymy is more multilingual than metaphor: Analysing tropes using ChainNet and the open multilingual wordnet. In *13th International Global Wordnet Conference (GWC2025)*. Pavia.

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)*.

References II

Christine Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press.

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).



References III

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.

Fahad Khan, Laurent Romary, Ana Salgado, Jack Bowers, Mohamed Khemakhem, and Toma Tasovac. 2020. Modelling etymology in LMF/TEI: The grande dicionário houaiss da língua portuguesa dictionary as a use case. In Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 3172–3180. European Language Resources Association, Marseille, France. URL <https://aclanthology.org/2020.lrec-1.388/>.



References IV

Temuulen Khishigsuren, Gábor Bella, Thomas Brochhagen, Daariimaa Marav, Fausto Giunchiglia, and Khuyagbaatar Batsuren. 2022. Metonymy as a universal cognitive phenomenon: Evidence from multilingual lexicons. URL osf.io/preprints/psyarxiv/f6yjt.

Krister Lindén and Lauri Carlson. 2010. Finnwordnet — wordnet på finska 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 Nicoletta Calzolari, Min-Yen Kan, Veronique Hoste, Alessandro Lenci, Sakriani Sakti, and Nianwen Xue, editors, *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 2984–2996. ELRA and ICCL, Torino, Italy. URL <https://aclanthology.org/2024.lrec-main.266>.



References V

Nurri Hifana 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.



References VI

Bolette S. Pedersen, Nathalie Sørensen, Sanni Nimb, Dorte Haltrup Hansen, Sussi Olsen, and Ali Al-Laith. 2025. Evaluating LLM-generated explanations of metaphors – a culture-sensitive study of Danish. In Richard Johansson and Sara Stymne, editors, *Proceedings of the Joint 25th Nordic Conference on Computational Linguistics and 11th Baltic Conference on Human Language Technologies (NoDaLiDa/Baltic-HLT 2025)*, pages 470–479. University of Tartu Library, Tallinn, Estonia. URL <https://aclanthology.org/2025.nodalida-1.51/>.

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.



References VII

Emanuele Pianta, Luisa Bentivogli, and Christian Girardi. 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*. Wroclaw University of Technology Press. URL http://www.plwordnet.pwr.wroc.pl/main/content/files/publications/A_Wordnet_from_the_Ground_Up.pdf, (ISBN 978-83-7493-476-3).

E. Puraivan, I. Renau, and N. Riquelme. 2024. Metaphor identification and interpretation in corpora with ChatGPT. *SN Computer Science*, 5:976.

Ervin Ruci. 2008. On the current state of Albanet and related applications. Technical report, University of Vlora. (<http://fjalnet.com/technicalreportalbanet.pdf>).



References VIII

Benoît Sagot and Darja Fišer. 2008. Building a free French wordnet from multilingual resources. In European Language Resources Association (ELRA), editor, *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*. Marrakech, Morocco.

Ivelina Stoyanova, Verginica Barbu Mititelu, Svetlozara Leseva, and Gianina Iordachioaia. 2025. Expanding and enhancing derivational and morphosemantic relations in princeton wordnet. In *GWC 2025*.

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*



References IX

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.



Faculty
of Arts