How Universal is Metonymy? Results from a Large-Scale Multilingual Analysis

Temuulen Khishigsuren[1], **Gábor Bella**[2], **Thomas Brochhagen**[3], **Daariimaa Marav**[1], **Fausto Giunchiglia**[2], **Khuyagbaatar Batsuren**[1]
[1]National University of Mongolia [2]University of Trento [3]Universitat Pompeu Fabra
kh.temulen@gmail.com; khuyagbaatar@num.edu.mn

1 Introduction

Several works from cognitive linguistics claim that systematic metonymy is universal across human languages (Barcelona et al., 2003; Brdar and Brdar-Szabó, 2003; Croft, 2002; Gibbs et al., 1994; Kövecses and Radden, 1998; Lakoff and Johnson, 2008; Panther and Radden, 1999). However, cross-linguistic surveys on the phenomenon have so far been limited to a small number of well-studied languages such as English. An important reason for this limitation is that current methodologies in cross-linguistic semantic analysis require serious involvement of language experts (Brdar-Szabó and Brdar, 2003a,b, 2012) or native speakers (Kamei and Wakao, 1992; Slabakova et al., 2013; Srinivasan and Rabagliati, 2015) or simply not suitable for metonymy studies such as the elicitation techniques (Koptjevskaja-Tamm et al., 2015).

On the other hand, the recent trend of exploiting digitally available lexical resources makes large-scale semantic studies feasible; e.g., the study of the emotion domain in 2474 languages (Jackson et al., 2019). This method is especially suitable for systematic metonymy as it is lexically encoded. Therefore, we used a lexico-semantic content of multilingual lexical databases to build a large-scale metonymy corpus that covers 26 metonymy patterns and 20 thousand metonymy instances (word pairs) in 189 languages belonging to 69 genera. Due to the broad linguistic coverage, our results considerably strengthen the stance on metonymy as a universal phenomenon. This new, freely available, online corpus of metonymy examples categorized by patterns is also reusable for future studies.

2 Methods

Among the various kinds of resources—databases, dictionaries, corpora—that were available to us, *multilingual lexical databases* were suitable to semi-automatically build a large multilingual corpus of metonymies. The explicit representation of words, their meanings, and their domains in multiple languages, as well as the presence of a cross-lingual alignment of meanings and domains enable us to extract metonymy in an efficient, partially automated manner.

Our database of choice is the Universal Knowledge Core (UKC)[1] (Giunchiglia et al., 2017), due to its wide linguistic, lexical, and conceptual coverage (120 thousand word meanings, 2 million words in 1127 languages). Metonymy patterns are straightforward to model through the three-layered *domain–concept–lexicon* architecture of the UKC. The concept layer represents *supra-lingual meanings* as a hierarchy of concepts based on the standard lexicographic broader–narrower relationship. The domain layer of the UKC provides a simple semantic categorization of concepts into domains such as *Animal*. The lexical layer, finally, consists of a separate *lexicon* for each language, each one lexicalizing the supra-lingual concept layer.

Metonymy corpus extraction process consisted of three steps. First, from the metonymy patterns mentioned in the literature, we selected a subset for which the UKC provides data. Second, through automatic extraction and expert validation, we identified metonymically related concepts. Third, based on the definitive set of metonymically related concept pairs, we automatically extracted lexicalizations for all languages in the database.

3 Results

Table 1 reports the statistics of the metonymy corpus[2] extracted semi-automatically from the database. Overall, 4,951 concept pairs were annotated as metonymically related, and the corresponding 20,095 metonymy instances were retrieved in 189 languages from the database. These 189 lan-

[1]http://ukc.datascientia.eu/

[2]The metonymy corpus is freely accessible as a stand-alone resource at https://github.com/kbatsuren/UniMet.

Proceedings of the 4th Workshop on Computational Typology and Multilingual NLP (SIGTYP 2022), pages 96 - 98

Table 1: Metonymy corpus statistics

Metonymy pattern	Illustrative example	Met. concepts	Met. instances	Langs	Families	Genera
Substance for Artifact	He filled the *glass* with water.	390	1,775	110	24	46
Fruit for Plant	The gardener watered the *lemon*.	408	3,330	114	24	43
Instrument for Action	She *combed* her hair.	617	2,083	95	24	40
Community for Place	He traveled to the *country*.	87	735	97	22	38
Plant for Food	*Broccoli* is delicious.	318	1,539	80	19	34
Animal for Meat	The *chicken* is tasty.	156	746	85	19	33
Action for Result	My thumb has a deep *cut*.	729	1,559	77	17	32
Object for Action	They are well *dressed*.	546	1,653	79	17	31
Substance for Action	I *milked* cows by hand.	242	942	78	17	31
Emotion for Cause	You are my *joy*.	104	405	64	14	28
State for Causal agent	He was a *success*.	160	645	59	15	27
Food for Action	They had *breakfasted* so early.	51	210	55	13	27
Building for People	*Church* sang a song.	71	384	63	15	26
Possessed for Possessor	She married *power*.	190	713	60	14	26
Agent for Action	The sheep will be *butchered*.	232	655	53	14	26
Product for Content	The *book* is interesting.	46	546	60	14	25
Body part for Person	I saw many new *faces* today.	156	442	61	12	25
Action for Food	They provided a *drink* at the party.	16	90	47	14	22
Animal for Fur	She likes to wear *mink*.	51	271	35	14	20
Container for Contained	He drank half of the *bottle*.	88	461	42	10	19
Event for People	*Party* went crazy.	61	257	39	11	18
Action for Object	A *lift* fell to the bottom of its shaft.	91	153	41	12	17
Action for Agent	You may be a *help* later.	57	207	34	12	17
Time for Action	We *honeymooned* in Bali.	25	108	36	11	17
Food for Event	*Dinner* took longer than usual.	9	132	36	11	17
Action for Time	My *shift* is over this morning.	50	54	21	8	14
Total		4,951	20,095	189	34	69

guages belong to 34 different families (phyla) and 69 genera. They are also geographically stratified (Figure 1). To the best of our knowledge, these results provide the widest linguistic coverage so far on metonymy (the broadest prior study we are aware of is by Hilpert (2007) that reported 39 phylogenetically different languages using *eye* to refer to *vision*).

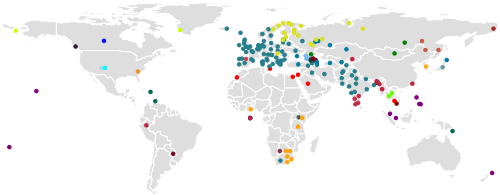

Figure 1: The presence of metonymy in world's languages (the same colors indicate the same family).

Based on the number of genera, the most universal pattern is SUBSTANCE FOR ARTIFACT for which we found 110 languages from 46 genera. FRUIT FOR PLANT and INSTRUMENT FOR ACTION are also very widely attested patterns for each we found 43 and 40 genera, respectively. Nevertheless, even the least widely covered patterns are attested across phylogenetically diverse languages from around the world. For example, the least diverse pattern, ACTION FOR TIME, is still attested in 21 languages from 14 genera and eight families from Africa, East Asia, the Pacific, Europe, and the Middle East. This result suggests that diverse societies use ACTION FOR TIME metonymies.

On the conceptual level, even specific concept pairs appear to be universal: for instance, 49 languages from 22 genera use the concept 'pear' to refer to its plant name. The English examples of COMMUNITY FOR PLACE and INSTRUMENT FOR ACTION patterns in Table 1 are attested in 69 and 39 other languages from 34 and 22 genera, respectively.

4 Conclusions

Metonymy is regarded by most linguists as a universal phenomenon. However, the field data backing up claims of universality has not been large enough so far to provide conclusive evidence. We introduce a large-scale analysis of metonymy based on a lexical corpus of over 20 thousand metonymy instances from 189 languages and 69 genera. No prior study, to our knowledge, is based on linguistic coverage as broad as ours. Drawing on corpus analysis, evidence of universality is found at three levels: systematic metonymy in general, particular metonymy patterns, and specific metonymy concepts.

References

Antonio Barcelona et al. 2003. Names: A metonymic "return ticket" in five languages. *Jezikoslovlje*, 4(1):11–41.

Mario Brdar and Rita Brdar-Szabó. 2003. in english, croatian and hungarian. *Metonymy and pragmatic inferencing*, 113:241.

Rita Brdar-Szabó and Mario Brdar. 2003a. The manner for activity metonymy across domains and languages. *Jezikoslovlje*, 4(1):43–69.

Rita Brdar-Szabó and Mario Brdar. 2003b. Referential metonymy across languages: What can cognitive linguistics and contrastive li nguistics learn from each other? *International Journal of English Studies*, 3(2):85–106.

Rita Brdar-Szabó and Mario Brdar. 2012. The problem of data in the cognitive linguistic research on metonymy: a cross-linguistic perspective. *Language Sciences*, 34(6):728–745.

William Croft. 2002. The role of domains in the interpretation of metaphors and metonymies. *Metaphor and metonymy in comparison and contrast*, pages 161–205.

Raymond W Gibbs, Raymond W Gibbs, and Jr Gibbs. 1994. *The poetics of mind: Figurative thought, language, and understanding*. Cambridge University Press.

Fausto Giunchiglia, Khuyagbaatar Batsuren, and Gabor Bella. 2017. Understanding and exploiting language diversity. In *IJCAI*, pages 4009–4017.

Martin Hilpert. 2007. Chained metonymies in lexicon and grammar. *Aspects of meaning construction*, pages 77–98.

Joshua Conrad Jackson, Joseph Watts, Teague R Henry, Johann-Mattis List, Robert Forkel, Peter J Mucha, Simon J Greenhill, Russell D Gray, and Kristen A Lindquist. 2019. Emotion semantics show both cultural variation and universal structure. *Science*, 366(6472):1517–1522.

Shin-ichiro Kamei and Takahiro Wakao. 1992. Metonymy; reassessment, survey of acceptability, and its treatment in a machine translation system. In *30th Annual Meeting of the Association for Computational Linguistics*, pages 309–311.

Maria Koptjevskaja-Tamm, Ekaterina Rakhilina, and Martine Vanhove. 2015. The semantics of lexical typology. In *The Routledge handbook of semantics*, pages 450–470. Routledge.

Zoltán Kövecses and Günter Radden. 1998. Metonymy: Developing a cognitive linguistic view. *Cognitive linguistics*, 9(1):37–77.

George Lakoff and Mark Johnson. 2008. *Metaphors we live by*. University of Chicago press.

Klaus-Uwe Panther and Günter Radden. 1999. *Metonymy in language and thought*, volume 4. John Benjamins Publishing.

Roumyana Slabakova, Jennifer Cabrelli Amaro, and Sang Kyun Kang. 2013. Regular and novel metonymy in native korean, spanish, and english: Experimental evidence for various acceptability. *Metaphor and Symbol*, 28(4):275–293.

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