Colexification patterns in the lexicon of physical actions: Synchrony and diachrony

Giorgos Amoiridis[1] & Thanasis Georgakopoulos[2]
[1,2]Aristotle University of Thessaloniki, [1]gamoirid@gmail.com, [2]ageorgakopoulos@enl.auth.gr

Keywords: diachronic colexifications, meaning change, polysemy, cognitive mechanisms

This study aims to add to the body of knowledge pertaining to the evolution of word meaning by examining synchronic crosslinguistic data in parallel with diachronic data from Ancient Greek. Specifically, we construct a synchronic colexification network of concepts from the physical actions domain with a twofold goal: (1) to identify diachronic pathways in the Ancient Greek lexicon that will enrich the synchronic network; and (2) to examine the degree to which cross-linguistic colexification networks and within-language polysemy patterns are interrelated. The specific hypothesis we test is that high network degree for a concept in a colexification network translates into high frequency of the lexemes expressing this concept in a language. In line with relevant research (e.g., Youn et al., 2016; Jackson, 2019), we utilize quantitative computational means to gather and analyze data from two types of resources: (a) dictionaries and (b) the CLICS[3] (Rzymski et al., 2019) database containing meaning associations in 3156 language varieties. We start with four Swadesh concepts (Swadesh, 1952) from the physical actions domain, i.e., SPLIT, SCRATCH, RUB, and WIPE, and employ the protocol developed by Georgakopoulos & Polis (2021) to create binary matrices that include all those concepts that colexify with the four initial concepts. We have compiled two such matrices which are the basis for the synchronic network: a CLICS[3] dataset and a 22-language dictionary-based dataset. Each of those is run through a Python script whose result is imported into Gephi (https://gephi.org/) where we carry out visualizations and draw network metrics. In a final step, we integrate the diachronic dimension into this synchronic network by investigating meaning extensions in the Ancient Greek lexicon from three periods, i.e., Homeric (8th c. BCE), Classical (6th-4th c. BCE), and Post-Classical (3rd c. BCE-3rd c. CE). This investigation is coupled with the identification of the cognitive motivations behind these extensions. Regarding our first research question, we can identify certain semantic changes in the domain in question. For instance, we find that in Homeric Greek the verb *smḗkhō* expresses the meaning WIPE (see [1]), but it only extends to the meaning WASH after Homer (see [2]). WASH is derived from WIPE meaning through metonymy, as washing generally involves the removal of dirt and grime from an object or surface using water.

1.	*ek*	*kefalês*	*d'*	***ésmēkhen***	*halòs*
	ELAT	head(F):GEN.SG	PTCL	**wipe:PST.3SG**	sea(F):GEN.SG
khnóon		*atrugétoio*			
flake(M):ACC.SG		unharvested:GEN.SG.F			

'He **wiped** the unharvested sea's [salt] flakes off his head' (Homer, *Odyssey* 6.226; 8th c. BCE)

2.	*kaì*	*mēdèn*	*autòn*	*proseksergázesthai*	*all'*
	CONJ	INDEF	DEM.ACC.SG.M	do:INF.PASS	CONJ

állous	*kaì*	*katakhéein*	*kaì*	***smḗkhein***
DEM.ACC.PL.M	CONJ	pour:INF	CONJ	**wash:INF**

'[The person taking the bath] should do nothing but others [must] pour water and **wash** him' (Hippocrates, *De diaeta in morbis acutis* 18; 5th-4th c. BCE)

Regarding the second question, we rely on *degree centrality*, a network metric which is considered a strong indicator of node importance in a graph (Newman, 2010). Given that polysemy and frequency have been found to be correlated, with more frequent words having more meanings (e.g., Bond *et al.*, 2019), we expect that high network degree for a concept in a colexification network will translate into high frequency of the lexemes expressing this concept in a language. To sum up, we aim to demonstrate that colexification networks can serve as a starting point for uncovering pathways of meaning evolution and can also be used to test certain hypotheses within languages. This parallel investigation promises to shed additional light on how polysemy patterns evolve over time and on the way a semantic field is structured.

References

Bond, Francis, Arkadiusz Janz, Marek Maziarz & Ewa Rudnicka. 2019. Testing Zipf's meaning-frequency law with wordnets as sense inventories. *10th Global Wordnet Conference. GWC 2019, Wroclaw, Poland: Global Wordnet Association*. 342–352. https://aclanthology.org/2019.gwc-1.44 (10 December, 2022).

Georgakopoulos, Thanasis, & Stéphane Polis, S. 2021. Lexical diachronic semantic maps: Mapping the evolution of time-related lexemes. *Journal of Historical Linguistics* 11(3). 367-420. https://doi.org/10.1075/jhl.19018.geo

Jackson, C. Joshua, Joseph Watts, Teague R. Henry, Johann-Mattis List, Robert Forkel, Peter J. Mucha, Simon J. Greenhill, Russel D. Gray & Kristen A. Lindquist. 2019. Emotion semantics show both cultural variation and universal structure. *Science* 366(6472). 1517-1522. https://doi.org/10.1126/science.aaw8160

Newman, Mark. 2010. *Networks: An Introduction*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199206650.001.0001

Rzymski, Christoph and Tiago Tresoldi et al. 2019. The Database of Cross-Linguistic Colexifications, reproducible analysis of cross-linguistic polysemies. https://doi.org/10.1038/s41597-019-0341-x

Swadesh, Morris. 1952. Lexico-Statistic Dating of Prehistoric Ethnic Contacts: With Special Reference to North American Indians and Eskimos. *American Philosophical Society* 96(4). 452-463.

Youn, Hyejin, Logan Sutton, Eric Smith, Christofer Moore, Jon F. Wilkins, Ian Maddieson, William Croft & Tanmoy Bhattacharya. 2016. On the universal structure of human lexical semantics, *Proceedings of the National Academy of Sciences of the United States of America* 113(7). 1766-1771. https://doi.org/10.1073/pnas.1520752113