

Lexis and Tradition: Variation in the vocabulary of Buddhist Literature

Ligeia Lugli (King's College London, UK)

This poster reports on the initial phase of a British Academy-funded project on lexico-semantic variation in Buddhist Sanskrit.

It outlines the overall scope and aims of the project, which addresses three sets of questions:

(1) Is there a discernible pattern in the distribution of near-synonyms across Buddhist texts and traditions? If there is, to what extent may this point to different communities of practice, and thus provide a clue to self-identity within Buddhist traditions?

(2) To what extent inter-sectarian dialogue is likely to have affected lexical choices and led to a degree of terminological harmonisation across different schools?

(3) To what extent doctrinal and conceptual evolution drove lexico-semantic innovation? In particular, to what extent might the constraints of orthodoxy have induced semantic change in the religio-philosophical vocabulary by promoting the re-contextualisation of traditional wording patterns in new doctrinal frameworks?

To answer these questions a diachronic corpus of Buddhist Sanskrit texts is being compiled. Within this corpus, lexico-semantic variation will be analysed using the methods for diachronic onomasiology described by Geeraerts (especially 2002 and 2012). The research will focus on two thematically related conceptual fields, NAMING and CONCEPTUALISING. To track semantic variation within these fields, the corpus is being semantically tagged using a word-sense annotation system derived from the taxonomy of the *Historical Thesaurus of English* and adapted to Sanskrit (Lugli 2015).

The poster also reports on the progress made in the development of a diachronic corpus of Buddhist Sanskrit. It describes the challenges that Buddhist Sanskrit materials pose for corpus linguistics and how these challenges have been addressed within this project, focussing especially on the difficulties related to the automated tokenisations and lemmatisation of Buddhist Sanskrit, the problems of corpus design that arise from the absence of a clear relative chronology for Indic Buddhist texts, and the task of adapting an existing semantic tagsets to cultural-specific concepts of South Asian Buddhism.

Finally, the poster presents preliminary findings on a Buddhist Sanskrit keyword, *saṃjñā*, and summarises how corpus data about this word problematise current assumptions and translation practices in the field of Buddhist Studies.

References

- Geeraerts, D. (2002). The scope of diachronic onomasiology. In Vilmos Agel, Andreas Gardt, Ulrike Hass-Zumkehr & T. Roelcke (Eds.) *Das Wort. Seine strukturelle und kulturelle Dimension. Festschrift fuer Oskar Reichmann zum 65. Geburtstag* (pp. 29–44). Tuebingen: Niemeyer.
- Geeraerts, D., Gevaert, C. & Speelman, D. (2012). How anger rose: Hypothesis testing in diachronic semantics. In Kathryn Allan and Justyna Robinson (Eds.) *Current Methods in historical semantics* (pp. 109–131). Berlin: de Gruyter.
- Kay, Ch., Roberts J., Samuels, M., Wotherspoon, I., & Alexander M. (Eds.). (2016). *The Historical Thesaurus of English*, version 4.2. Glasgow: University of Glasgow. <http://historicalthesaurus.arts.gla.ac.uk/>.
- Lugli, L. 2015. Mapping meaning across time and cultures: innovations in Sanskrit lexicography. In L. Li, J. McKeown, & L. Liu (Eds.) *Words Dictionaries and Corpora: Proceedings of the 9th International Conference of ASIALEX* (pp. 514-523). Hong Kong: The Hong Kong Polytechnic University.