

Network Analysis as a Means of Assessing Translatability of Emotion Words

Katie Hoemann, Northeastern University, khoemann@gmail.com
Margherita de Luca, Sapienza University of Rome, margherita.deluca@uniroma1.it
Lisa Feldman Barrett, Northeastern University, l.barrett@northeastern.edu

Keywords: culture, concepts, corpus linguistics, equivalence, alignment

Emotion words, as labels for emotion concepts, communicate culturally held understandings about experiences, values, and goals. Cross-cultural differences in emotion concepts are well-established (e.g., Russell, 1991), and indeed this domain shows poor semantic alignment across languages (Thompson et al, 2018). Purportedly ‘untranslatable’ emotion words are particularly salient examples of this phenomenon: Tagalog speakers have a word, “gigil”, for “when something is so cute that you want to squeeze it”; second language learners of Dutch have a difficult time appreciating the nuances of “gezelligheid”; and Italian does not have an easy translation for English “excitement”. But are these words actually untranslatable, or does translatability perhaps exist on a continuum? Previous work has employed a variety of methods for comparing word meaning across languages (Ogarkova, 2016), including feature rating, free association, and metaphor analysis. However, most of these methods are focused on describing the nature of differences in meaning, or comparing entire semantic fields, rather than measuring the (dis)similarity of individual words.

In this project, we use free association data from the Small World of Words study (De Deyne et al, 2008) to construct semantic networks for ‘untranslatable’ Dutch words (e.g., “gezellig”) and their possible English translations. Network analyses of free association data have been used to model lexical knowledge and semantic growth, but have not yet been used to compare meaning across languages. Moreover, by using closely-related languages, we are able generate networks that are maximally comparable (Majid et al, 2015). We populated semantic spaces for Dutch and English target words using semantic relations as well as back translation protocols. We then characterized these semantic spaces in terms of their network metrics. For each Dutch to English comparison, we selected networks with equivalent metrics and evaluated the similarity of target words’ behavior within their respective networks using these same metrics (e.g., degree, betweenness centrality, clustering coefficient). Using this data-driven approach, we found no (single) English word matches for Dutch targets.

Semantic networks can also be constructed based on distributional semantic models (DSMs; e.g., Utsumi, 2015). To contextualize our results from the free association networks, we present DSM networks based on lexical neighborhood data from Continuous Bag-of-Words (CBOW) models trained for Dutch and English, which have been shown to accurately predict behavioral data in both languages (Mandera et al, 2017). These DSM networks allow us to compare words based on both paradigmatic and syntagmatic relationships they maintain within their respective languages, thereby providing insight into aspects of word meaning that may be ignored by other methods. Taken together, our results illustrate a novel means of understanding word meaning and assessing translatability. In our discussion, we evaluate our network analyses against other existing measures of cross-linguistic word similarity, including data from forthcoming property generation, feature rating, and scenario elicitation studies.

References

- De Deyne, S., & Storms, G. (2008). Word associations: Norms for 1,424 Dutch words in a continuous task. *Behavior research methods*, 40(1), 198-205.
- Majid, A., Jordan, F., & Dunn, M. (2015). Semantic systems in closely related languages. *Language Sciences*, (49), 1-18.
- Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, 92, 57-78.
- Ogarkova, A. (2016). Translatability of emotions. In H. L. Meiselman (Ed.), *Emotion Measurement* (pp. 575-599). Woodhead Publishing.
- Russell, J. A. (1991). Culture and the categorization of emotions. *Psychological bulletin*, 110(3), 426.
- Thompson, B., Roberts, S., & Lupyan, G. (2018, July). Quantifying semantic alignment across languages. Research presented at the 40th annual meeting of the Cognitive Science Society, Madison, WI.
- Utsumi, A. (2015). A complex network approach to distributional semantic models. *PloS one*, 10(8), e0136277.