

---

## Semantic entropy measures and the semantic transparency of noun noun compounds

---

Melanie J. Bell  
*Anglia Ruskin University*  
[melanie.bell@anglia.ac.uk](mailto:melanie.bell@anglia.ac.uk)

Martin Schäfer  
*Universität Jena*  
[post@martinschaefer.info](mailto:post@martinschaefer.info)

Recently, two different entropy measures based on the relational structure of English compounds have been used in studies of semantic transparency and lexical decision times. Pham & Baayen (2013) show that the entropy of semantic relations in the modifier family is negatively correlated with semantic transparency. Schmidtke et al. (2015) find that the relational entropy for individual compounds is correlated with lexical decision time. However, neither study takes the ambiguity of the compound constituents into account. Our contribution addresses this gap. In a model of semantic transparency, we show that the relational entropy of the head family interacts with a second entropy measure, namely synset entropy: uncertainty about the reading of a given constituent in terms of homonymy and polysemy.

We used the publically available compound dataset from Bell & Schäfer (2016), which is annotated with semantic relation and WordNet synset (sense) of the constituents. We calculated the entropy of the probability distributions of the synsets and semantic relations for all modifier and head constituent families and modelled the semantic transparency ratings collected by Reddy et al. (2011). Our final model of compound transparency shows an interaction between the two entropy measures based on the head constituent families. When the synset entropy is low, perceived transparency is unaffected by relational entropy; when the synset entropy is high, perceived transparency is negatively correlated with relational entropy, mirroring the finding by Pham & Baayen (2013) for modifier families.

These findings suggest that relation entropy is not negatively correlated with semantic transparency across the board. If the reading of the head is easily predictable (low synset entropy), it makes little difference to perceived transparency whether the relation entropy is high or low. But with increasing synset entropy, the effect of relation entropy becomes increasingly pronounced. When the reading of the head is uncertain, the compound appears more transparent if the relation is easily predictable across all possible senses than if it is not. One way of interpreting this result is that in cases of high synset entropy, low relation entropy masks constituent ambiguity, while high relation entropy necessitates greater activation of different readings to arrive at even a ‘gist’ interpretation, thereby increasing processing effort and lowering perceived transparency.

**References:** • Bell, M. J. & Schäfer, M. (2016): Modelling semantic transparency. *Morphology* 26(2), 157–199. • Pham, H. & Baayen, R. H. (2013). Semantic relations and compound transparency: A regression study in CARIN theory. *Psihologija* 46 (4), 455–478. • Reddy, S., McCarthy, D. & Manandhar, S. (2011). An Empirical Study on Compositionality in Compound Nouns . In: *Proceedings of the 5th IJCNLP 2011*. • Schmidtke, D., Kuperman, V., Gagné, C.L. & Spalding, Th. L. (2016). Competition between conceptual relations affects compound recognition: the role of entropy. *Psychon Bull Rev* 23(2), 556–570.