Modeling lexical processing with linear mappings

Harald Baayen

Baayen et al. (2019) proposed a computational model for the mental lexicon

which approximates comprehension and production with linear mappings between

high-dimensional representations of form and meaning. In my presentation, I

will discuss three case studies that illustrate the new opportunities that come

with this approach.

The first case study, carried out in collaboration with Susanne Gahl

(Berkeley), addresses the spoken duration of English homophones. Gahl (2008)

had previously reported that low-frequency homophones have longer durations

than their high-frequency counterparts. Using the DLM with empirical word

embeddings to represent words' semantics, we have been able to show that the

extent to which a words' form is supported by its semantics is a strong

determinant of its spoken word duration.

The second case study addresses the question of how to represent the meanings

of complex words. My collaborator Elnaz Shafaei-Bajestan focused on noun

plurals in English. Upon closer inspection of the word embeddings of singulars

and plurals, it turns out that the change from a singular embedding to a plural

embedding varies with the semantic class to which a given lexeme belongs. The

original conceptualization of pluralization proposed in Baayen et al. (2019),

which assumed that the same general 'plural vector' is to be added to a

singular's vector to obtain the corresponding plural, clearly is too

simplistic.

The third case study addresses the issue of lexical learning. The DLM provides

two ways for estimating linear mappings between form vectors and meaning

vectors. One way is to make use of multivariate multiple regression, the other

uses the learning rule of Widrow and Hoff (1960). My collaborators Maria

Heitmeier and Yu-Ying Chuang used the DLM with incremental learning to model

the lexical decision latencies of the British Lexicon Project (Keuleers et al.,

2012). Prediction accuracy for reaction times increases substantially when predictors were used that were derived from a DLM that updates its connection weights as the experiment unfolds.

References

Baayen, R. H., Chuang, Y-Y., Shafaei-Bajestan, E., and Blevins, J.P. (2019).

The discriminative lexicon: {A} unified computational model for the lexicon and

lexical processing in comprehension and production grounded not in

(de)composition but in linear discriminative learning, Complexity,

https://www.hindawi.com/journals/complexity/2019/4895891/.

Chuang, Y-Y. and Baayen, R.H. (2021). Discriminative learning and the lexicon:

NDL and LDL. Oxford Research Encyclopedia of Linguistics,

Gahl, S. (2008). Time and Thyme Are not Homophones: The Effect of Lemma

Frequency on Word Durations in Spontaneous Speech, Language 84, 474-496.

Keuleers, E., Lacey, P., Rastle, K. and Brysbaert, M. (2012). The British

Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic

English words. Behavior Research Methods 44, 287-304.

Widrow, B. and Hoff, M. E. (1960). Adaptive switching circuits, 1960 WESCON

Convention Record Part IV, 96-104.