Sensorimotor distance: A fully grounded and efficient measure of semantic similarity

Cai Wingfield
c.wingfield@lancaster.ac.uk

Louise Connell
l.connell@lancaster.ac.uk

Lancaster University

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 682848

Introduction

- Computational modelling and experimental design across the cognitive sciences rely on measures of **semantic similarity between concepts**.
- We present a new measure of **sensorimotor distance** between concepts.
- Unlike other measures, sensorimotor distance is **fully grounded** in sensorimotor experience.
- Here we investigate how sensorimotor distance explains human similarity judgements.
- We further present an online tool for computing sensorimotor similarity for nearly **800 million pairs of concepts**.

A fully grounded measure

- Sensorimotor distance is based on the **Lancaster Sensorimotor Norms** (Lynott et al., 2020).
- **Six perceptual dimensions**: vision, touch, audition, olfaction, taste and interoception.
- **Five action-effector dimensions**: Hand/arm, leg/foot, torso, mouth/throat and head (excluding mouth/throat).
- Norms have ratings for nearly 40,000 concepts.
- This allows comparisons of concepts based on their multidimensional sensorimotor experience profiles.

Sensorimotor distance

Computing sensorimotor distance

- Sensorimotor distance between two concepts is computed using their 11-dimensional sensorimotor rating vectors.
- We recommend cosine distance, which we present here.
- We also tested correlation, Euclidean and Minkowski-3 distances; fits to human data weren't as good.

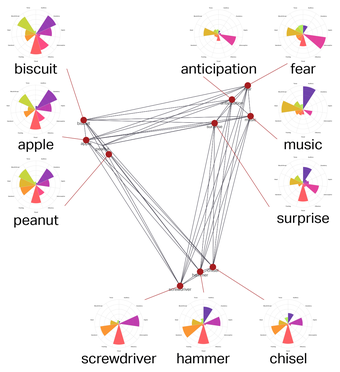
Visualising relationships in sensorimotor space

- By computing pairwise distances between concepts, we visualise their **multidimensional arrangement** in sensorimotor space.
- Sensorimotor information applies to both concrete and abstract categories (Lynott et al., 2020), and reveals semantic structure amongst collections of concepts.

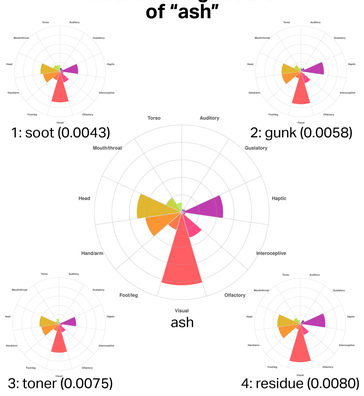
Capturing detailed conceptual relationships

- Sensorimotor ratings encode surprisingly nuanced information about concepts and their relationships.
- For example, this is revealed by searching for **nearest neighbours**.

Visualising a multidimensional arrangement



Nearest neighbours of "ash"



Sensorimotor distance explains human similarity judgements

Comparison to other measures of semantic similarity

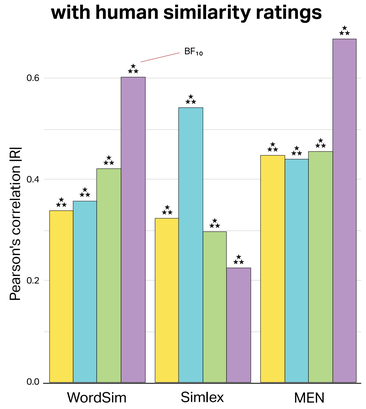
- We modelled human similarity judgement datasets using sensorimotor distance alongside three other traditional measures.

Measure	Relies on
WordNet (Jiang & Conrath, 1997)	Distance in a taxonomic database.
Feature overlap (Buchanan et al., 2019)	Counting shared semantic features from a norming study
LSA (Landauer & Dumais, 1997)	Linguistic-distributional similarity of words.

- We modelled three datasets of human similarity judgements.
- Sensorimotor distance was a good predictor of similarity judgements, comparable with other measures.
- No single predictor was preferred for all datasets.

Dataset	N (pairs)
WordSim (Finkelstein et al., 2002)	353
Simlex (Finkelstein et al., 2002)	999
MEN (Bruni et al., 2014)	3,000

Correlations with human similarity ratings

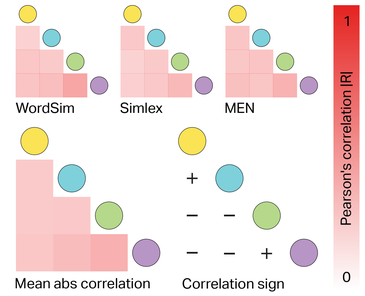


Correlations between predictors

- Sensorimotor distance was somewhat correlated with other measures (|R| = .212–.300).
- Other measures were somewhat correlated with each other (|R| = .224–.377).
- Patterns of correlations consistent over datasets.

WordNet distance · Feature overlap · LSA · Sensorimotor distance · BF > 100 p < 0.0001 · Preferred predictor

Between-predictor correlations



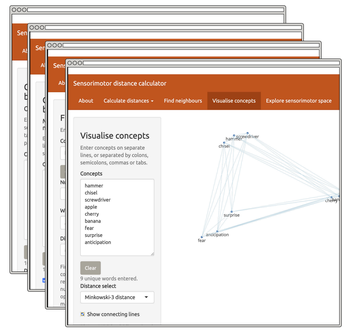
A web app for sensorimotor distance calculations and visualisations

- **We present a tool for computing and visualising sensorimotor distance.**
- Coverage of nearly 800 million concept pairs.
- Free and open source.

Access the web app

Modes of operation

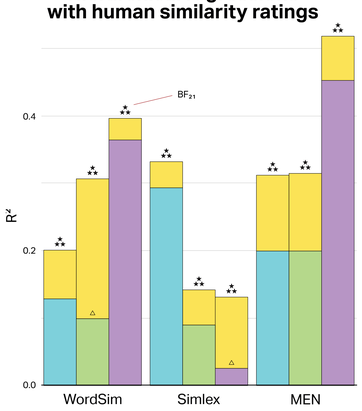
- Compute distances: one-to-one, one-to-many, many-to-many.



- Visualise arrangements of select concepts.
- Find nearest neighbours.

Sensorimotor distance explains unique variance

Hierarchical regressions with human similarity ratings



Sensorimotor distance explains variance missed by other measures

- We entered sensorimotor distance into hierarchical regressions with each of the other predictors.
- For all datasets, sensorimotor distance explained **substantial additional variance** over the other predictor.
- In two cases, sensorimotor distance was the better predictor.

Sensorimotor distance consistently explains variance

- We entered all four predictors into a single hierarchical regression.
- No single predictor was consistently preferred over all datasets.
- Sensorimotor distance was consistently the second best, and was **included in every best model**.
- BF-inclusions for sensorimotor distance were consistently high (log BF = 9.7–47.9).

Dataset (Total R^2)	**Predictors** (log BF-inclusion)				
	1 (Best)	**2**	**3**	**4 (Weakest)**	**Not a predictor**
WordSim 48.9%	LSA 33.4	Sensorimotor 10.6	WordNet 6.1	Feature overlap 4.1	
Simlex 33.8%	WordNet 82.3	Sensorimotor 9.7	Feature overlap 5.1		LSA -0.3
MEN 60.4%	LSA 262.4	Sensorimotor 47.9	Feature overlap 32.0	WordNet 20.0	

Findings: Sensorimotor distance…

1. **…is a fully grounded measure of semantic similarity.**
2. **…captures a consistent portion of behavioural variance, comparable with other measures.**
3. **…captures unique information missed by other measures.**
4. **…provides a useful tool for experimental design and computational modelling.**

References

Lancaster Sensorimotor Norms: Lynott, D., et al. (2020). *Behavior Research Methods*. doi:10.3758/s13428-019-01316-z
Simlex: Hill, F., et al. (2016). *Computational Linguistics*. doi:10.1162/COLI_a_00237
WordSim: Finkelstein, L., et al. (2002). *ACM Transactions on Information Systems*. doi:10.1145/503104.503110
MEN: Bruni, E., et al. (2014). *Journal of Artificial Intelligence Results*. doi:10.1613/jair.4135
LSA: Landauer, T. K. & Dumais, S. T. (1997). *Psychological Review*.
WordNet: Jiang, J. J. & Conrath, D. W. (1997). *arXiv:cmp-lg/9709008*
Feature overlap: Buchanan, E. M., et al. (2019). *Behavior Research Methods*.
SPoSE: Hebart, M. N., et al. (2020) *Nature Human Behavior*.
RSA: Kriegeskorte, N., et al. (2008) *Frontiers in Systems Neuroscience*.

Download poster PDF