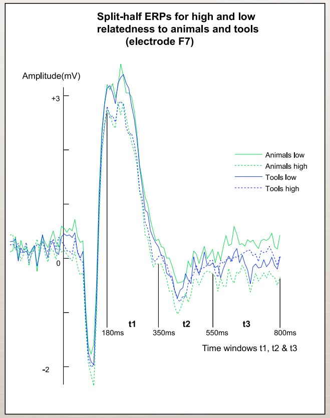
Exploring the High Dimensional Semantic Space in the Brain

Sverker Sikström(1) & Petter Kallioinen(1, 2)

(1) Lund University Cognitive Science, (2) Lund University Institution for Psychology
sverker.sikstrom@lucs.lu.se, petter.kallioinen@lucs.lu.se

Figure 1



Semantic Spaces and Neural Semantics

Finding how semantic representations are organized in the brain is a central task for cognitive neuroscience. The main paradigm for neuro-imaging is subtracting brain activity evoked from two types of stimuli, chosen prior to the experiment. We propose an alternative approach, utilizing Semantic spaces, generated from co-occurrence in huge text corpora (see box). Semantic spaces offer quantification of semantic relations and demonstrate automatic extraction of meaning.

The semantic space approach enable:

1. Flexible extracting of variables, including data driven approaches, and stimuli from other studies

2. Correlation instead of subtraction between conditions, and use of semantic covariates.

3. Comparisons of psychological variables in semantic space.

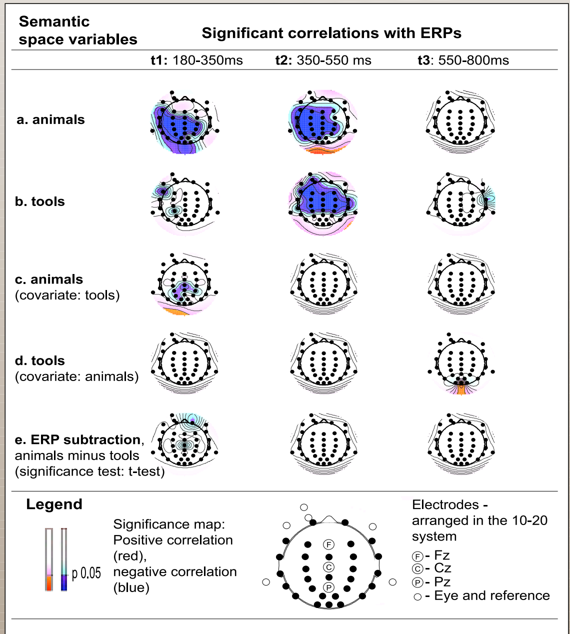
4. Exhaustive search of the semantic space

5. Empirical grounding - replicable semantic relations based in actual language use

Investigation below of semantic relatedness to animals and tools exemplify 1, 2 and 5

ERPs Correlate with Animal and Tool Relatedness

Figure 2



Animals and tools are two of the most studied categories in research on neural semantics. Theorists have suggested that (1) words are grounded in the perception action systems (Barsalou, 2008) and that (2) representation of animal-concepts involve visual areas (Martin, 2007) while tool-concepts involve motor areas.

Previous studies has focused on category membership, (i.e. "animal-dog"). Here we investigate relatedness to the category (i.e. "animal-zoo"). This is motivated by interest in typicality and overlap of microfeatures as semantic organization schemes (Patterson et al, 2007).

We use semantic space to measure distance between

(1) a set of words (presented in am earlier memory experiment where ERPs where recorded), and

(2) category centroids, average representation in semantic space for animal and tool words (from McRae et al, 2005).

The distances between words (1) and centroids (2) were correlated with ERP amplitude for each electrode. Word length, word frequency and bigram frequency were covariates.

Results

Figure 1 show ERPs for high and low, animal- and tool- relateness. Maps show significant correlations (**figure 2**). Animal and tool relatedness show both overlapping and distinct correlations (**figure 2, a & b**). Use of the other category as covariate yields unique correlations (**figure 2, c &d**). The subtraction method shows very little significant difference (**figure 2, e**).

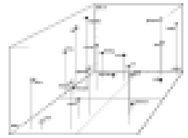
Discussion

We find effects with the semantic space approach where the subtraction method fail and demonstrate the benefits from the introduction: Flexibility (**1.**), the incorporation of arbitrary words from previous literature. Correlational approach and use of covariates (**2.**). The positive results imply empirical groundedness (**5.**).

The significance maps fit the general notion of visual processing (posterior and lateral) as important for animals, more frontal motor activations for tools, and a left hemisphere domination for words. The early latencies of the correlations is evidence for early semantic processing (Hauk et al, 2006).

The different effects of two separate relatedness gradients in the same set of words, fit the view that semantic organization involve overlapping microfeatures, rather than distinct categories.

What is a Semantic Space?

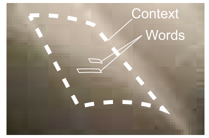


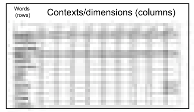
Example of a 3-dimensional semantic space

A model where semantic relatedness is expressed as distance in a space! We use a high dimensional (100-300) space with representations of 15000 words, created with LSA, Latent Semantic Analysis.

Create a semantic space

1. Select a corpora (natural text), define contexts and words.

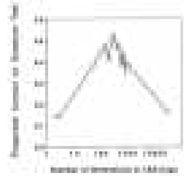




2. Create a co-occurrence matrix. Count how many times each word appear in each context.

3. Project to fewer dimensions, and

4. test with a synonym test. Does synonyms generally have shorter distance than other words?



An introduction to semantic spaces:
Landauer et al (1997). *A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge.* Psychological Review vol. 104 (2) pp. 211-24

References

Barsalou (2008). *Grounded cognition.* Annual review of psychology vol. 59 pp. 617-45

Hauk et al (2006). *The time course of visual word recognition as revealed by linear regression analysis of ERP data.* Neuroimage vol. 30 (4) pp. 1383-400

McRae et al (2005). *Semantic feature production norms for a large set of living and nonliving things.* Behavior research methods vol. 37 (4) pp. 547-59

Martin (2007). *The representation of object concepts in the brain.* Annual review of psychology vol. 58 pp. 25-45

Patterson (2007). *The reign of typicality in semantic memory.* Philos Trans R Soc Lond, B, Biol Sci vol. 362 (1481) pp. 813-21