Contextual Characteristics of Concrete and Abstract Words

Diego Frassinelli, Daniela Naumann & Sabine Schulte im Walde
frassinelli@ims.uni-stuttgart.de
Institut für Maschinelle Sprachverarbeitung, University of Stuttgart

Take Home Message
In comparing concrete and abstract words, Distributional Semantics is not completely in line with Grounding Theories of cognition
- √ concrete words
- X abstract words

We show that:
- Distributionally similar words have a similar range of concreteness scores (Study 1)
- Target and context words share similar concreteness scores both at type (Study 2) and token level (Study 3)
- Concrete targets have more variable contexts (concreteness) than abstract targets (Study 3)

Concrete vs. Abstract Words

Grounded Theory of Cognition → Both concrete and abstract words should co-occur with concrete words:
- Concrete: grounded in the sensory motor system
- Abstract: mapped to concrete concepts

Distributional Hypothesis → Represent the meaning of a word in terms of its linguistic context
- Targets as vectors of co-occurrences
- Cognitively plausible representation

Materials
- 9,241 target nouns
- 9,241 context nouns (dimensions)
- Measures:
  - Co-occurrences: ENCOW14A corpus with symmetric window ±20
  - Concreteness scores: 1 - 5 (Brysbaert et al., 2014)

Study 1: Concreteness in Distributionally Similar Words
Do distributionally similar words have also similar concreteness scores?

Study 2: Concreteness of Context Words - Type Level
What is the nature of the contexts of concrete and abstract words?

Study 3: Concreteness of Context Words - Token Level
What is the nature of the contexts of concrete and abstract words when considering their frequency?

Concrete words → mainly concrete contexts
Abstract words → mainly abstract contexts

Concrete and Abstract Words

Grounded Theory of Cognition → Both concrete and abstract words should co-occur with concrete words:
- Concrete: grounded in the sensory motor system
- Abstract: mapped to concrete concepts

Distributional Hypothesis → Represent the meaning of a word in terms of its linguistic context
- Targets as vectors of co-occurrences
- Cognitively plausible representation

Materials
- 9,241 target nouns
- 9,241 context nouns (dimensions)
- Measures:
  - Co-occurrences: ENCOW14A corpus with symmetric window ±20
  - Concreteness scores: 1 - 5 (Brysbaert et al., 2014)

Study 1: Concreteness in Distributionally Similar Words
Do distributionally similar words have also similar concreteness scores?

Study 2: Concreteness of Context Words - Type Level
What is the nature of the contexts of concrete and abstract words?

Study 3: Concreteness of Context Words - Token Level
What is the nature of the contexts of concrete and abstract words when considering their frequency?

Concrete words → mainly concrete contexts
Abstract words → mainly abstract contexts

Concrete and Abstract Words

Grounded Theory of Cognition → Both concrete and abstract words should co-occur with concrete words:
- Concrete: grounded in the sensory motor system
- Abstract: mapped to concrete concepts

Distributional Hypothesis → Represent the meaning of a word in terms of its linguistic context
- Targets as vectors of co-occurrences
- Cognitively plausible representation

Materials
- 9,241 target nouns
- 9,241 context nouns (dimensions)
- Measures:
  - Co-occurrences: ENCOW14A corpus with symmetric window ±20
  - Concreteness scores: 1 - 5 (Brysbaert et al., 2014)