

Computerlinguistik

E07: Bedeutung

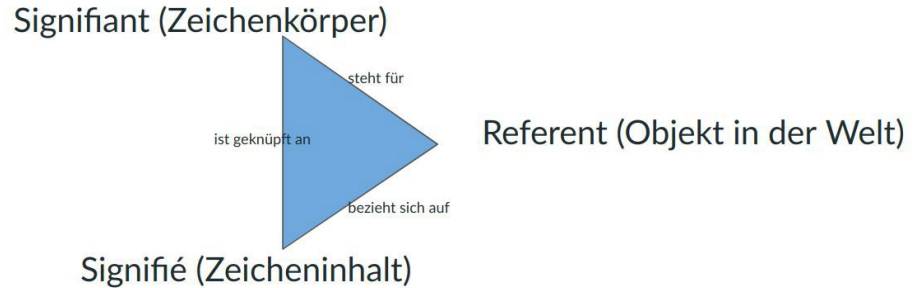
Semantik

- Einführung
- Wortsemantik vs. Satzsemantik
- Merkmaltheorie
- Prototypentheorie
- Semantische Relationen
- Ausblick: Word Embeddings

	Zeichenebene	Wortebene	Satzebene
Ausdrucksseite (signifiant)	Phonetik / Phonologie	Morphologie	Syntax
Inhaltsseite (signifié)		Wortsemantik	Satzsemantik

Semantik - Einführung

- Sprachliches Zeichen
→ Semiotisches Dreieck



- Mentales Lexikon: Teil des Langzeitgedächtnisses
- Konzepte: Bausteine unseres Wissens, basieren auf Erfahrungen
 - Kategorienkonzepte (Types): Repräsentanten von Informationen über Klassen - Relationsbeziehung: Äquivalenz
 - Partikularkonzepte (Token): Repräsentanten von Informationen über Einzelnes - Relationsbeziehung: Identität

Wortsemantik vs. Satzsemantik

➤ Lexikalische Semantik (oder Wortsemantik):

- Kontextunabhängige Bedeutungen von Wörtern
- Relationen zwischen Bedeutungen (Sinnrelationen)

➤ Satzsemantik:

- Prinzip der Kompositionalität:

$$\begin{array}{r} \text{Wortbedeutung} \\ + \quad \text{Art der Zusammensetzung} \\ \hline = \quad \text{Satzbedeutung} \end{array}$$

Merkmalsemantik

- **Merkmalthypothese:** Bedeutungen sind nicht ganzheitlich, sondern aus elementaren Inhaltselementen zusammengesetzt. Diese Inhaltselemente nennt man *semantische Merkmale* oder *Seme*.
- **Merkmalbündel:** Innere Struktur von Bedeutungen
- **Semantische Opposition:** Gegenüberstellung distinktiver Bedeutungsmerkmale
- Ermittlung wesentlicher Merkmale: Kriterium der Notwendigkeit
- Probleme der Merkmalthypothese: Relationale Merkmale - Verben - kognitive Plausibilität

Prototypensemantik

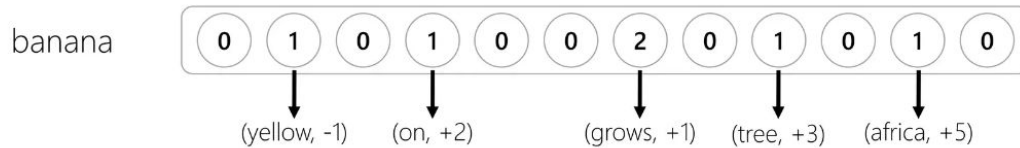
- **Prototypenhypothese:** Prototypen sind mentale Repräsentationen von Kategorien, die nicht klar umgrenzt werden können. Kontinuum der Kategoriezugehörigkeit.
- Entstehung von Prototypen: Frequenz und Relevanz
- Problem: Mentale Bilder oder abstrakte Repräsentationseinheiten
- Alternative oder Ergänzung zur Merkmalsemantik?

Semantische Relationen

- **Synonymie:** Bedeutungsgleichheit (gleiche denotative Merkmale, ggfs. unterschiedliche konnotative Merkmale)
- **Referenzidentität:** Gleicher Referent, ggfs. unterschiedliche Bedeutungen
- **Ambiguität:** Unterschiedliche Bedeutungen einer Wortform (Polysemie vs. Homonymie; Homophonie vs. Homographie)
- **Hyponymie** (Unterordnung) vs. **Hyperonymie** (Überordnung) vs. **Kohyponymie** (Nebenordnung)
- **Kontradiktion** (keine Zwischenstufen, nicht steigerbar) vs. **Antonymie** (Zwischenstufen, steigerbar)

Ausblick: Embeddings – Vektorrepräsentation

- John Firth (1957): “You shall know a word by the company it keeps.”
- Zusammenstellung von Kontexten, in denen Wörter stehen
- Numerische Repräsentation in einem hochdimensionalen Vektorraum



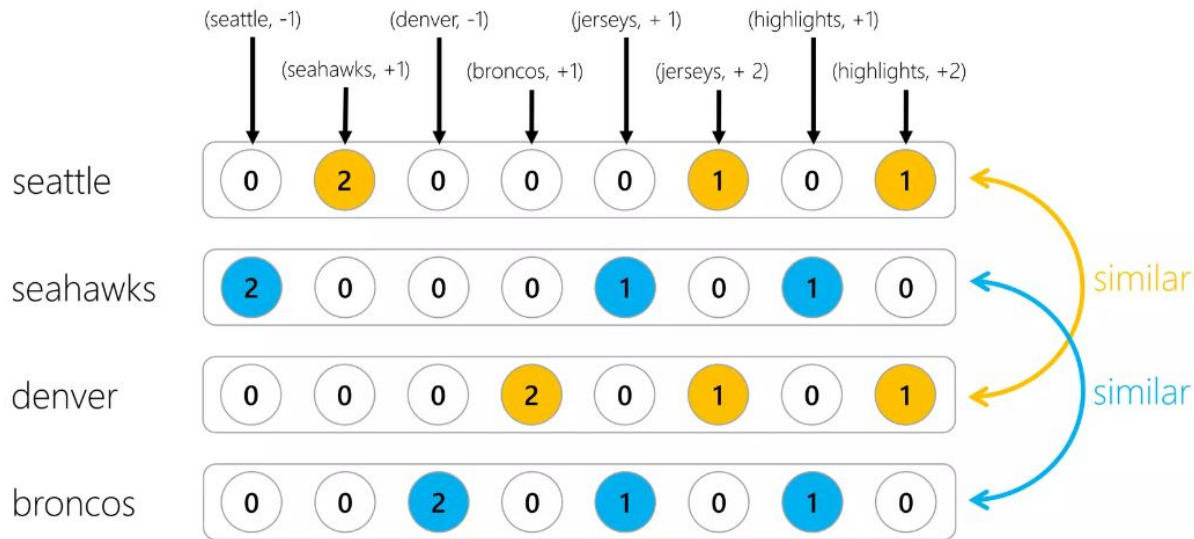
The vector can correspond to **neighboring word context**.

e.g., "yellow banana grows on trees in africa"

-1 0 +1 +2 +3 +4 +5

Aus (der insgesamt sehr übersichtlichen Präsentation von Bhaskar Mitra):
<https://www.slideshare.net/BhaskarMitra3/a-simple-introduction-to-word-embeddings>

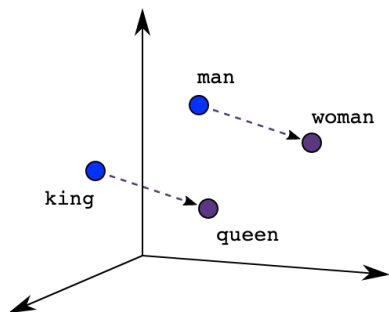
Ausblick: Embeddings – Vektorähnlichkeit



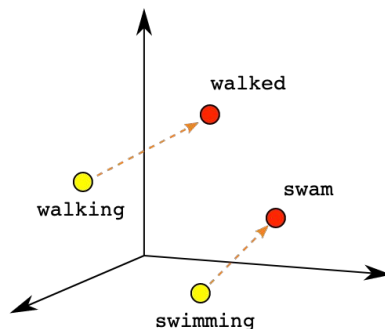
Auch aus (der insgesamt sehr übersichtlichen Präsentation von Bhaskar Mitra):
<https://www.slideshare.net/BhaskarMitra3/a-simple-introduction-to-word-embeddings>

Ausblick: Embeddings – Relationen

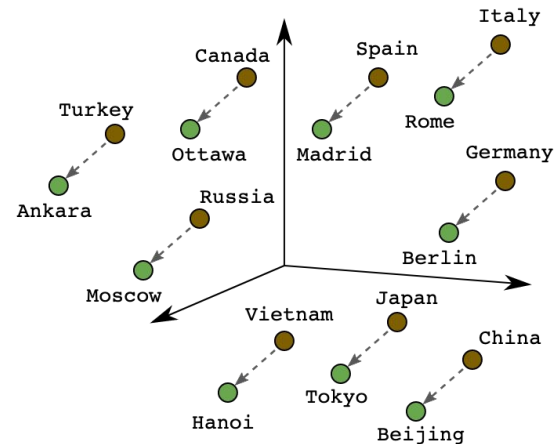
- Relationen zwischen Bedeutungen von Wörtern lassen sich über Vektoroperationen “errechnen”



Male-Female



Verb Tense



Country-Capital

Literatur / Hausaufgabe

➤ **Zur Nachbereitung:**

- Lesen Sie: Schwarz/Chur (1996) Kap. 1 - 2.2 (S. 13-60)
 - Spielen Sie gerne etwas mit Word Embeddings herum (siehe letzte Folie, das Konzept ist zentral für alle Bereiche der modernen Sprachverarbeitung)
- Die Texte (bzw. Links) finden Sie im Ilias-Seminarordner.