Analysis of Language Resources

First Lecture

Hiroyuki Akama

What is Language ?

- Stored data of written documents or spoken words (in repositories, archives and memories) ?

- Modality of communication that serves functions of the body (effectors, neural networks, brain) ?

- Collective ←→ Individual

- Dynamic ←→ Static

- And many forms of dualism.

Sign of Language: Diversity

- Fundamental element of language
- Where does the diversity come from ?
- Compared with the gene information, the diversity of language is excessive. → Arbitrariness
- But the identification task (for example, translation) is considered as being easy or possible. → Motivation
- Profound contradiction

Two Conflicting Viewpoints

- **Motivation**
- There is a tight, natural and necessary relationship between the sign of language and the real object that it represents (Plato's Cratylus).
- Etymology: The meanings of signs are bound to their sources.

- **Arbitrariness**
- The sign of language is contingently determined without reflecting nature or reality.
- The meanings of signs are never fixed and stable.
- There is nothing but *difference* in language (Saussure).

Two Concepts of Motivation

- SIMPLE: There is a remarkable similarity between a word and its referent

- COMPLEX: Language is grounded in our perception, physical exercise and bodily (sensory motor) experience.
  → Embodiment theory

Two Concepts of Arbitrariness

- SIMPLE: There is NO logical relationship between the sound and the meaning of a word.

- COMPLEX: Language is a system of interdependent terms in which the value of each term results solely from the simultaneous presence of the others. (Saussure; Maruyama) →Amodal theory

Today's Opposition in Linguistics

- **Based on Motivation**
- **Embodiment theory**, Perceptual Symbol theory (Barsalou, Glenberg, Richardson,...)
- Thinking the communication at the brain level
- Analogous bodily reaction to the referent of a word
- physical experience bound to signs
- Mirror System

- **Based on Arbitrariness**
- **Amodal theory** (Fodor, Kintsh, Pylyshyn,...)
- Revival of the Saussurian Linguistics (Modern Linguistics)
- Data manipulation in statistical linguistics, corpus linguistics and computational linguistics
- Free transposition of signs
- Association, Calculation

Embodiment Theory : Example

- Relationship between the meaning comprehension of words and the space recognition
- To Measure the reaction time of participants' judging if the words with the latent meaning of direction such as upward or downward match with the picture appearing in the upside or downside
- The promotion effect ←→ Interference
- In recognizing the meaning of the words, to clarify the relation between the body and the language, specifically between movement of the body and perception of the brain.

Amodal Theory: Example

- To Simulate the human's thought process using the lexical co-occurrence data by a multivariate analysis (Vector Space Model)

- The Latent Semantic Analysis (LSA) using Singular Value Decomposition (SVD), for example, permits us to presume the genre of a book even though its title does not contain any keyword indispensable for determining the attribution (Landauer, Foltz, Laham).

- But it cannot simulate the affordance (Symbol-grounding Problem) (Glenberg, Robertson)

Summary

- Two main viewpoints of linguistics.
- Amodal theory ←→Multimodal theory
- Lexical co-occurrence ←→Embodiment
- Logic, Sign, Computation ←→ Biology, Body, Perception, Physical exercise
- Is there no way to mediate viewpoints in linguistics?