Development of Diachronic Corpus

Project by the National Institute for Japanese Language and Linguistics, Japan, NINJAL: 2009–13, 4 year project.

Main purpose: Study of Japanese language
(sub) purpose: Study of Japanese (classic) literature

Parole

- corpus
- static
- structure of language
- visible
- tangible

Langue

- language
- complex system
- sequential reader pattern prediction
- visualization
- dynamic presentation of diachronic language change

Figure 1: Corpus and Description, Langue and Parole: The nature of language is dynamic and always changing while the phenomena of language might be static. We should consider the dynamic change of language as a component comprised of various elements. The feature of language we usually observe is a complex system and tangled with wide-ranging elements.

A case study: use of SAKURA (cherry blossoms) in Mt. Yoshino → Kokinshū (ca. 905) vs Shinkokinshū (1205)

Sakura (桜) and Yoshino (吉野), a place name in Nara prefecture
← Kokinshū (ca. 905)
Shinkokinshū (1205) → during 300 years differences.

Future Task

- To define linguistic units suitable for each era
- To develop a dictionary for machine analysis
  → it allows us syntagmatic and paradigmatic analysis

Conclusion

- Addressed basic concepts and framework of diachronic corpus
- Illustrated the serial comparison model for historical analysis
  → Lexical differences between any two groups of texts