

# Problems of Automatic Translation of Japanese Dialogues into German.

Melanie Siegel

Deutsches Forschungszentrum fuer Künstliche Intelligenz GmbH  
Stuhlsatzenhausweg 3, D-66123 Saarbrücken \*

A machine translation system can only perform part of the tasks a human translator does. An empirical study (Siegel et al. 1993) shows which tasks are in question: the translator must, for example, deal with syntactically "incorrect" sentences, consider the interrelationship between the participants, control the dialogue through asserting or omitting information, restructure utterances and answer questions. One of the central problems of translation Japanese dialogues into German is translation mismatches: Information that is not explicit in the source language Japanese is required for the generation of the target language German (Kameyama 1986, Siegel/Metzing 1994).

Firstly Japanese zero pronominals have to be resolved. They are frequent in Japanese spoken language, and every machine translation system that concerns Japanese has to deal with them (cf. Kameyama 1985). Other translation mismatches are number, definiteness and missing reference. There are almost no articles and no number markings in Japanese. The German translation of Japanese utterances in our experiment frequently contain referential expressions that were not included in the source (Siegel 1994).

Solution methods for zero pronominals mostly solve only one aspect of the problem, as for example zero pronominals with antecedents in the immediate preceding sentence. In contrast to these approaches we follow the 'sublanguage'-approach and show how different solution methods can be combined to solve different sorts of zero pronominals in Japanese dialogues:

- Centering approaches deal with zero pronominals with antecedents in the preceding dialogue text. They can be combined with simple semantic restrictions. But in contrast to such general approaches most of the examples in our data material need domain-specific restrictions.
- Lexical-pragmatic restrictions can solve most of those zero pronominal that refer to hearer or speaker. They base on information about honorifics, evaluation, mental processes, emotions and wishes.

---

\* tel.:(0681) 302-5288, e-mail: siegel@dfki.uni-sb.de

- A lot of the examples can be solved with information on idiomatic expressions in the source and target language.

The translation of zero pronominals itself however is part of the transfer. It is not sufficient to simply substitute a zero pronominal by a referent before transfer. CLE-a\_terms (Alshawi 1992) are adequate representations for zero pronominal and thus suitable to be used as an input structure for the transfer module. They are underspecified as the information on pronominalization is kept and restricted as it is possible to restringate the (zero) pronominal.

Information on number and definiteness must be found partly in the source of the Japanese sentence, partly in the discourse structure and partly in lexical information of the German nouns. Therefore the search for this information must be placed in the center of a machine translation system, the transfer. Mismatches of reference also have to be solved in the transfer.

If the stating of restrictions for zero pronominals and of transfer is preceded by an empirical study, it becomes obvious that domain-specific restrictions and rules are predominant to general ones.

## References

- [1] Alshawi, Hiyan (1992): The Core Language Engine. Cambridge, Mass.
- [2] Kameyama, Megumi (1985): Zero Anaphora: The Case of Japanese. Michigan.
- [3] Kameyama, Megumi (1996): A Property-sharing Constraint in Centering. In: Proceedings of the 24th Annual Meeting of the Association of Computational Linguistics, New York: 200-206.
- [4] Metzger, D., Siegel, M. (1994): Zero Pronoun Processing: Some Requirements for a VERBMOBIL System. VM-Memo 46. Bielefeld.
- [5] Siegel, Melanie (1994): Definitheit und Numerus. Anforderungen an den Transfer Japanisch-Englisch. VM-Memo 56. Bielefeld.
- [6] Siegel, M., Metzger, D. (1994): Nullpronomina und die Organisation von Wissensquellen für den Transfer Japanisch-Englisch. VM-Memo 12. Bielefeld.
- [7] Siegel, M., Kuroda, H., Kudo, E. (1993): Dialogdaten: Terminplanung japanisch-deutsch. Arbeitsberichte Computerlinguistik 2-93. Universität Bielefeld.