

A SPOKEN LANGUAGE TRANSLATION SYSTEM : ASURA

Gen-ichiro Kikui, Mark Seligman, Toshiyuki Takezawa, Tsuyoshi Morimoto,
Masami Suzuki, Kenji Kita, Masaaki Nagata, Toshihisa Tashiro, Herbert Tropic†
Shigeki Sagayama, Jun-ichi Takami, Kazumi Ohkura, Akira Kurematsu

ATR Interpreting Telephony Research Laboratories

Seika-cho, Soraku-gun, Kyoto, 619-02, JAPAN

†Siemens AG, Corporate Research, ZFE ST SN 74

Otto Hahn-Ring 6 D-8000 Munich 83, GERMANY

Abstract

This videotape introduces a speech-to-speech machine translation system, which translates Japanese speech into English and German speech. Emphasis here is on the speech recognition and automatic translation. The tape also shows an international joint experiment enabling real-time conversations using speech translation systems.

1 Distinctive Features of the System

Our spoken language translation system, called ASURA (Advanced Speech Understanding and Rendering System of ATR), consists of three basic components: speech recognition, translation and speech synthesis. The video focuses on speech recognition and translation. Speech synthesis is explained in [Sagisaka *et al.*, 1992].

1.1 Speech Recognition

(1)HMM-LR Speech Recognition is developed and employed combining allophonic HMMs (Hidden Markov Models) and a generalized LR parser extended to handle phoneme context [Nagai *et al.*, 1992].

(2)HMMs are generated automatically from statistical data by our Successive State Splitting (SSS) algorithm [Takami and Sagayama, 1992].

(3) A speaker adaptation technique allows accurate speaker-independent recognition.

1.2 Language Translation

(1) Expressions peculiar to spoken language can be accepted and translated properly into the target languages [Kume *et al.*, 1989][Dohsaka, 1990] [Hasegawa, 1990].

(2) Linguistic knowledge sources for translation are defined declaratively while still maintaining efficiency [Nagata, 1992].

(3) Declarative grammars are handled efficiently using a new unification algorithm [Tomabechi, 1991] and rule application strategy [Kikui, 1992].

1.3 Integration of Speech Recognition and Language Translation

For efficiency, linguistic constraints are distributed throughout speech recognition and language translation

[Takezawa *et al.*, 1991]. Speech recognizer produces sentential candidates, some of which may be ruled out by the translation component.

References

- [Dohsaka, 1990] Kohji Dohsaka. Identifying the Reference of Zero-Pronouns in Japanese Based on Pragmatic Constraint Interpretation. In *Proceedings of ECAI-90*, 1990.
- [Hasegawa, 1990] Toshiro Hasegawa. Rule Application Control Method in a Lexicon-driven Transfer Model of a Dialogue Translation System. In *Proceedings of ECAI-90*, 1990.
- [Kikui, 1992] Gen-ichiro Kikui. Feature Structure based Semantic Head Driven Generation. In *Proceedings of Coling-92*, 1992.
- [Kume *et al.*, 1989] Masako Kume, Gayle K. Sato, and Kei Yoshimoto. A Descriptive Framework for Translating Speaker's Meaning. In *Proceedings of European Chapter of ACL-89*, 1989.
- [Nagai *et al.*, 1992] Akito Nagai, Jun-ichi Takami, and Shigeki Sagayama. The SSS-LR Continuous Speech Recognition System: Integrating SSS-Derived Allophone Models and A Phoneme-Context-Dependent LR Parser. In *Proceedings of ICSLP-92*, 1992.
- [Nagata, 1992] Masaaki Nagata. An Empirical Study on Rule Granularity and Unification Interleaving Toward an Efficient Unification-based Parsing System. In *Proceedings of Coling-92*, 1992.
- [Sagisaka *et al.*, 1992] Yoshinori Sagisaka, Nobuyoshi Kaiki, Naoto Iwahashi, and Katsuhiko Mimura. ATR ν -talk Speech Synthesis System. In *Proceedings of ICSLP-92*, 1992.
- [Takami and Sagayama, 1992] Jun-ichi Takami and Shigeki Sagayama. A Successive State Splitting Algorithm for Efficient Allophone Modeling. In *Proceedings of ICASSP-92*, 1992.
- [Takezawa *et al.*, 1991] Toshiyuki Takezawa, Kenji Kita, Junko Hosaka, and Tsuyoshi Morimoto. Linguistic Constraints for Continuous Speech Recognition in Goal-Directed Dialogues. In *Proceedings of ICASSP-91*, 1991.
- [Tomabechi, 1991] Hideto Tomabechi. Quasi-Destructive Graph Unification. In *Proceedings of ACL-91*, 1991.