

Multilingual Conversational System Research

9807-11

Progress Report: January 1, 2000—June 30, 2000

James Glass and Stephanie Seneff

Project Overview

The long-term goals of this research project are to foster collaboration between MIT and NTT speech and language researchers and to develop language-independent approaches to speech understanding and generation. We will initiate this effort by developing the necessary human language technologies that will enable us to port our conversational interfaces from English to Japanese. The Jupiter weather information system will be used as the basis of this porting process. This work will involve the close collaboration with NTT researchers both in Japan and at MIT.

Progress Through June 2000

During this period, we have continued our data collection efforts from NTT employees, and have been using the data to improve the capabilities of the Mokusei system. We have begun transcribing these data for system evaluation and re-training. At this time, Mokusei is able to answer appropriately to approximately 60% of user queries.

One of the biggest changes we have made to the system has been the incorporation of a new version of our language generation component, called GENESIS-II. As we described in our previous progress report, we have completely redesigned this component in order to address several limitations which we observed in our multilingual research. Over the past six months, we have spent a considerable amount of effort upgrading the Mokusei domain to GENESIS-II, and have made significant improvements. Although there are still areas which need more work (e.g., winter forecasts), the current language generation capability is much more natural sounding than its predecessor.

One of the issues we have been troubled by is a lack of consistency among the recognition vocabulary, the parsing grammar, the back-end geography tables, and the generation vocabulary. The inconsistencies have been a source of understanding errors. We have developed scripts to help identify and rectify these inconsistencies. As part of this process, we have used Hiragana as an intermediate representation, so that we can more easily detect mis-spellings, and have consistent word representations. For this reason, we also plan to develop a Hiragana transcription tool, which we believe our transcribers will be able to use more easily, and whose output can be converted to a consistent format for the Mokusei system.

Research Plan for the Next Six Months

In the next six months, we plan to continue improving all system capabilities, and to slowly ramp up our data collection efforts. Motivated by initial user feedback, we intend to spend a substantial effort to improve the quality of the Mokusei speech output. Our plan is to develop and incorporate a Japanese version of our ENVOICE corpus-based concatenative speech synthesizer. When combined with a domain-dependent corpus, this synthesizer is capable of concatenating variable-length units (e.g., phrases, words, sub-word units) to produce very natural sounding speech. We hope to investigate an appropriate set of sub-word units for Japanese, and then design, and record a corpus appropriate for the Mokusei domain. It is also possible that we can make use of an existing NTT weather corpus for this purpose.