

PROGRESS REPORT FOR DARPA SLS PROGRAM AT DRAGON SYSTEMS, INC.¹

Janet Baker and Larry Gillick

Dragon Systems, Inc.
320 Nevada Street
Newton, Massachusetts 02160
DRAGON@A.ISI.EDU
TEL: (617) 965-5200
FAX: (617) 527-0372

INTRODUCTION

This report provides a brief summary of recent work and describes anticipated future work in continuous speech recognition research at Dragon Systems under the auspices of DARPA's Spoken Language Systems project.

REVIEW OF CURRENT WORK

Dragon Systems has continued to develop its continuous speech recognition capability on personal computers, with a strong focus on achieving real-time or near real-time performance in that environment. We are now able to run recognition using vocabularies of up to 5000 words, although the larger vocabularies do slow down the recognition substantially; therefore more work still needs to be done to improve the computational efficiency of our algorithms.

The recognizer now has an interface which allows for on-line adaptation. This interface constitutes a beginning step in the development of a "DragonDictate"-like error correction facility for continuous speech, one that would allow a user to dictate text, correcting errors as he proceeds, with the system gradually improving its models by using the feedback from the user's error corrections.

In recent months Dragon has turned its attention to Resource Management, and this has resulted in a new focus on the special characteristics of this particular task; of course, Dragon continues to place a strong emphasis on improving overall recognition accuracy in a way that will benefit the general run of speech recognition applications. In order to use the digitized speech supplied on CD-Rom the first step was to write software that would closely emulate the old signal processing done on our standard hardware, so that we would have a baseline performance assessment that could be used to evaluate the new, more computationally demanding signal

processing algorithms that we plan to implement in the near future. We then went through several cycles of development on our adaptation and training algorithms in the course of gradually improving our performance on the RM1 development test data, and are still in a period of rapid development. To enhance our ability to perform experiments with the large quantity of speech data that has been supplied, we ported our recognizer to the IBM RS-6000 workstation. We have a collection of these machines and, now that they have been networked together, it is possible to do experiments on many speakers simultaneously. The recognizer has also been ported to the Apple Macintosh computer.

With our focus on improving the accuracy of the recognizer has come a concern with the sources of error. To enhance our ability to study our sources of error, we have developed a diagnostic program known as ERRSPEC, which displays segmented spectrograms of utterances and models, together with a variety of revealing plots that highlight where the recognition algorithm has gone wrong.

FUTURE PLANS

Dragon plans to continue its work on the Resource Management Task and to begin work on ATIS in the coming months. A primary goal will continue to be improvements in our overall accuracy, but with the additional aim of moving to a more broadly based speaker independent modeling strategy, one that is likely to be based on mixture distributions (at the PEL level, at the PIC level, and at the word level). Investigation of alternative signal processing algorithms will also be a high priority, as we move away from the computational constraints that our old hardware had placed on us. Post processing strategies based on the outputs from the N-Best algorithm will also be explored.

1. This work was sponsored by the Defense Advanced Research Projects Agency and was monitored by the Space and Naval Warfare Systems Command under Contract N000-39-86-C-0307.