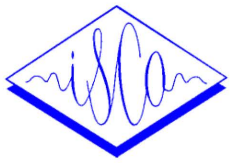

European Conference on
Speech Technology
Edinburgh, Scotland, UK
September 1987

LARGE VOCABULARY NATURAL LANGUAGE SPEECH RECOGNITION IN SOFTWARE

James K. Baker and Janet M. Baker

ABSTRACT

This presentation provides a description and report of Dragon Systems large vocabulary natural language, speaker-dependent isolated word recognition systems. Based on stochastic processing, these systems are implemented primarily in software running on a personal computer (PC) or workstation. The only processor used in addition to the host PC microprocessor is a simple 2 MHZ 8-bit processor to assist in the real-time data acquisition of the speech signal. Additional components on a sparsely populated audio board, include an A/D converter or CODEC with some TTL logic for dynamic range/gain control.

A brief chronology of test performance is reported. All speech training and test data were recorded in typical office environments, using an inexpensive hand-held omni-directional cassette-recorder type mike.

Although these tasks can be processed both off and on-line, the on-line facility provides a graceful interactive user interface. If more than 1 word is deemed probable by the recognition process, then a short rank-ordered "top choice" menu is offered to the user in-line. If the top word is correct, confirmation is implicit by continued speech or delay. If the correct recognition is ranked 2 or more, the user may quickly select it, with a single-keystroke, for example, and continue. New words or phrases can easily be introduced automatically in a few seconds during the dictation. Full voice console and dictionary facilities run concurrently for immediate user access. An application interface enables most MS.DOS software including popular word processors, spreadsheets, and database programs, to accept both voice and keyboard entries flexibly.

PERFORMANCE TEST SUMMARY

Date	Speaker Gender	Active Vocab.	Perplexity	Concurrent Language Model	PC Processor	Response Time (avg.)	Total# Test Words	% correct	% top 5
Task: General Prose Dictation									
7/84	F	1277	1277	No	4.5MHZ8088	22sec	268	87.7%	100%
Task: Engineering Documentation Dictation									
11/85	M	2000	not calculated	Yes	6MHZ80286	1.5sec	3050	92.9%	97.9%
Task: Natural Language Data Base Inquiries									
9/86	F	868	868	No	8MHZ80286	near real-time	846	92.4%	98.6%

Dragon Systems, Inc., Chapel Bridge Park, 90 Bridge Street, Newton, MA 02158, U.S.A.;
Tel. (617) 965-5200, Fax. (617) 527-0372