

AD-A125 134

SPEECH RECOGNITION: ACOUSTIC PHONETIC AND LEXICAL  
KNOWLEDGE REPRESENTATION(U) MASSACHUSETTS INST OF TECH  
CAMBRIDGE RESEARCH LAB OF ELECTRONICS V W ZUE

1/1

UNCLASSIFIED

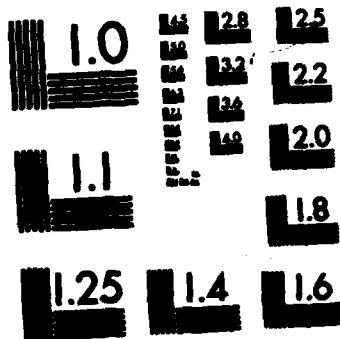
22 FEB 83 N00014-82-K-0727

.F/G 5/8

NL



END  
FILMED  
+



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

12

AD A125134

STATUS REPORT

SPEECH RECOGNITION: ACOUSTIC PHONETIC  
& LEXICAL KNOWLEDGE REPRESENTATION

OFFICE OF NAVAL RESEARCH  
DEPARTMENT OF THE NAVY

Contract Number N00014-82-K-0727

covering the periods

July 1, 1982 - September 30, 1982  
October 1, 1982 - December 31, 1982

Submitted by  
Victor W. Zue

February 22, 1983

DTIC  
ELECTE  
MAR 2 1983  
S D D

DTIC FILE COPY

MASSACHUSETTS INSTITUTE OF TECHNOLOGY  
Research Laboratory of Electronics  
Cambridge, Massachusetts 02139

DISTRIBUTION STATEMENT A  
Approved for public release  
Distribution Unlimited

88 02 028 170

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE   |  | READ INSTRUCTIONS<br>BEFORE COMPLETING FORM   |
|---|--|---|
| 1. REPORT NUMBER  | 2. GOVT ACCESSION NO.<br><i>AD-A125134</i> | 3. RECIPIENT'S CATALOG NUMBER   |
| 4. TITLE (and Subtitle)<br>SPEECH RECOGNITION: ACOUSTIC PHONETIC<br>AND LEXICAL KNOWLEDGE REPRESENTATION  |  | 5. TYPE OF REPORT & PERIOD COVERED<br>Status Report<br>07/01/82 - 09/30/82<br>10/01/82 - 12/31/82 |
| 7. AUTHOR(s)<br>Victor W. Zue   |  | 6. PERFORMING ORG. REPORT NUMBER  |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Research Laboratory of Electronics<br>Massachusetts Institute of Technology<br>Cambridge, Massachusetts 02139  |  | 8. CONTRACT OR GRANT NUMBER(s)<br>N00014-82-K-0727  |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Office of Naval Research - Dept. of the<br>Navy - 800 North Quincy Street<br>Arlington, Virginia 22217   |  | 10. PROGRAM ELEMENT, PROJECT, TASK<br>AREA & WORK UNIT NUMBERS<br>NR-049-542                      |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)   |  | 12. REPORT DATE<br>February 22, 1983  |
|   |  | 13. NUMBER OF PAGES<br>5  |
|   |  | 15. SECURITY CLASS. (of this report)<br>Unclassified  |
|   |  | 15a. DECLASSIFICATION/DOWNGRADING<br>SCHEDULE   |
| 16. DISTRIBUTION STATEMENT (of this Report)<br>Approved for public release; distribution unlimited  |  |   |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)  |  |   |
| 18. SUPPLEMENTARY NOTES   |  |   |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number)<br>A large vocabulary and a continuous speech recognition system, which are acoustic-phonetically based phontactic constraints acoustic cues to word boundaries.<br><i>→ the purpose of this program is to</i>   |  |   |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number)<br>Develop a speech data base facility under which the acoustic characteristics of speech sounds in various contexts can be studied conveniently; investigate the phonological properties of a large lexicon of, say 10,000 words, and determine to what extent the phonotactic constraints can be utilized in speech recognition; study the acoustic cues that are used to mark work boundaries; develop a test bed in the form of a large-vocabulary, IWR system to study the interactions of acoustic, phonetic and lexical knowledge; and develop a limited continuous speech recognition system with the goal of recognizing any English word from its spelling in order to assess the in- |  |   |

20. Interactions of higher-level knowledge sources.

|                           |                                     |
|---------------------------|-------------------------------------|
| <b>Accession For</b>      |                                     |
| NTIS GRA&I                | <input checked="" type="checkbox"/> |
| DTIC TAB                  | <input type="checkbox"/>            |
| Unannounced               | <input type="checkbox"/>            |
| Justification             |                                     |
| By                        |                                     |
| Distribution/             |                                     |
| <b>Availability Codes</b> |                                     |
| Dist                      | Avail and/or<br>Special             |
| A                         |                                     |



## PROGRESS REPORT

Contract Number N00014-83-K-0727

Periods Covered: 07/01/82 - 09/30/82  
10/01/82 - 12/31/82

---

Implementation of the essential parts of the SPIRE speech research facility has been completed. Speech signals can now be digitized, analyzed and displayed conveniently. We have now digitized approximately thirty minutes of speech spoken by seven different speakers. Phonetic transcriptions for some of these words and sentences have also been entered. In addition, SPIREX, a data base facility to examine the statistical properties of speech sounds, has also been implemented on the Lisp machine.

A study of the structural constraints on the sound patterns in American English has also been completed. With the help of an on-line dictionary, we were able to determine the various ways with which words can be accessed without detailed phonetic information. The 20,000 word Merriam-Webster Pocket Dictionary has been carefully scrutinized in terms of the phonemic spellings and stress markers.

In order to design the corpus to be used to study the acoustic cues to word boundaries, the Harvard phonetically balanced sentences were carefully examined. Our initial conclusion is that the distribution of consonants with words and across word boundaries are sufficiently rich that the Harvard sentences can be used as the initial corpus. Data for the pilot study were re-examined using the SPIREX facility for confirmation.

A preliminary version of the "front-end" portion of the large vocabulary isolates word recognition system has been implemented. The system can now obtain, from the digitized waveform, a set of acoustic parameters. From these acoustic parameters, the speech signal is segmented into acoustic regions. A set of re-write rules is then applied to transfer the acoustic segments into a string (or lattice) of broad phonetic labels. This string is used to access a 20,000 word lexicon, and the result is a set of words satisfying the broad phonetic description. So far we have tested approximately 100 words spoken by two male speakers, and the results are promising. Very often, this procedure will result in significant reduction (down to <1%) of the lexical candidates.

A preliminary version of the continuous digit recognition system has been implemented. The system is intended to utilize acoustic phonetic knowledge, and the knowledge is implemented in a modular fashion so that it can easily be extended to other tasks. The current implementation will produce a broad phonetic description of the signal, which can then be used to propose the underlying digit strings. Due to the strong constraints on the task, we expect this initial broad classification will very often produce a unique digit string.

Page 2  
PROGRESS REPORT  
Contract Number N00014-83-K-0727

A variety of system software has been written on the Lisp machine. This includes device drivers for the Digital Sound A/D-D/A converter and the Versatek Printer Plotter and a critical-band digital fitter bank for spectrum analysis.

DISTRIBUTION LIST

|   | <u>DODAAD Code</u> |      |
|---|--------------------|------|
| Director<br>Advanced Research Project Agency<br>1400 Wilson Boulevard<br>Arlington, Virginia 22209<br>Attn: Program Manager                                       | HX1241             | (2)  |
| Associate Director for Mathematical<br>and Physical Sciences Research Program<br>Office of Naval Research<br>800 North Quincy Street<br>Arlington, Virginia 22217 | N00014             | (3)  |
| Administrative Contracting Officer<br>E19-628<br>Massachusetts Institute of Technology<br>Cambridge, Massachusetts 02139  |                    | (1)  |
| Director<br>Naval Research Laboratory<br>Washington, D.C. 20375<br>Attn: Code 2627  | N00173             | (6)  |
| Defense Technical Information Center<br>Bldg. 5, Cameron Station<br>Alexandria, Virginia 22314  | S47031             | (12) |
| TACTEC<br>Battelle Memorial Institute<br>505 King Avenue<br>Columbus, Ohio 43201  | 79986              | (1)  |