





And the second second

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

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

88 02 028 170

AD A 1

34

251

ş

FILE COPY

DISTRIBUTION STATEMENT A Approved for public releases Distribution Unlimited

ECURITY CLASSIFICATION OF	THIS PAGE (When Date E	Entered)	
REPORT DO	REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
I. REPORT NUMBER		2. GOVT ACCESSIO	NO. 3. RECIPIENT'S CATALOG NUMBER
•		AD-A 125 13	Y
1. TITLE (and Sublitle)		5. TYPE OF REPORT & PERIOD COVERED Status Report	
SPEECH RECOGNITION: ACOUSTIC PHONETIC AND LEXICAL KNOWLEDGE REPRESENTATION		07/01/82 - 09/30/82	
		NTATION	10/01/82 - 12/31/82 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s)	•		8. CONTRACT OR GRANT NUMBER(a)
Victor W. Zue			N00014-82-K-0727
			nol
PERFORMING ORGANIZATION	NAME AND ADDRESS		10. PROGRAM EL SMENT, PROJECT, TASK
Massachusetts Ins	titute of Tec	chnology	NR 049-542
Cambridge, Massac	husetts 02139) ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~	
1. CONTROLLING OFFICE NAM	E AND ADDRESS	· • •	
UITICE OF Naval R	esearch - Dep Ouincy Street	pt. of the	February 22, 1983
Arlington, Virgin	ia 22217	-	13. NUMBER OF PAGES
14. MONITORING AGENCY NAME	& ADDRESS(II dillerent	from Controlling Offi	ce) 15. SECURITY CLASS. (of this report)
			Unclassified
			15. DECLASSIEICATION/DOWNGRADING
			SCHEDULE
6. DISTRIBUTION STATEMENT Approved for publ	(of this Report) ic release; c	listribu+io	n unlimited
6. DISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT	(of this Report) ic release; c	listribu+io	n unlimited
6. DISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT	(of this Report) ic release; c	listribu+io	n unlimited
6. DISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT	(of this Report) ic release; c	listribu+io	n unlimited
6. DISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT 8. SUPPLEMENTARY NOTES	(of this Report) ic release; c	listribu+io	n unlimited
6. DISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT 8. SUPPLEMENTARY NOTES	(of this Report) ic release; c	listribu+io	n unlimited
6. DISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT 8. SUPPLEMENTARY NOTES	(of this Report) ic release; c	listribu+io	n unlimited
6. DISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT 8. SUPPLEMENTARY NOTES	(of this Report) ic release; c	listribu+io	in unlimited
6. DISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT 8. SUPPLEMENTARY NOTES 9. KEY WORDS (Continue on rove	(of this Report) iC release; c (of the abstract entered in 	listribu+io	n unlimited nt trom Report) mber)
6. DISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT 8. SUPPLEMENTARY NOTES 9. KEY WORDS (Continue on rove A large vocabulary	(of this Report) iC release; (fof the abstract entered if ree elde if necessary and y and a conti	listribu+io Block 20, 11 dillere Identify by block nu NuOus speed	n unlimited nt from Report) mber) ch recognition system,
BISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT 8. SUPPLEMENTARY NOTES 9. KEY WORDS (Continue on reveal A large vocabulary which are acoustic	(of this Report) ic release; c (of the obstract entered in resc elde if necessary and y and a conti c-phoneticall word boundard	listribu+io Block 20, 11 different Identify by block nu nuous speed y based pho	n unlimited nt trom Report) mbor) ch recognition system, ontactic constraints
6. DISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT 8. SUPPLEMENTARY NOTES 9. KEY WORDS (Continue on rover A large vocabulary which are acoustic acoustic cues to a	(of this Report) ic release; c for the abstract entered if for the abstract entered if rate side if necessary and y and a conti c-phoneticall word boundari	listribu+io Block 20, 11 different Identify by block nu nuous speed y based pho es.	n unlimited nt from Report) mber) ch recognition system, ontactic constraints M
 DISTRIBUTION STATEMENT Approved for publ Approved for publ DISTRIBUTION STATEMENT (DISTRIBUTION STATEMENT (SUPPLEMENTARY NOTES SUPPLEMENTARY NOTES KEY WORDS (Continue on revealed a constitution on revealed a constitution on the second constitution of the second constitution on the second constitution on the second constitution on the second constitution of the second constitution of the second constitution of the second constitution on the second constitution on the second constitution of the second constitution on the second constitution of the second constitution of	(of this Report) ic release; c for the abstract entered in for the abstract entered in and a conti c-phoneticall word boundari Se of Ho	Identify by block nu nuous speed y based pho es.	n unlimited In unlimited In the second sec
a. DISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT a. SUPPLEMENTARY NOTES b. KEY WORDS (Continue on revea A large vocabulary which are acoustic acoustic cues to the <i>file puipo</i> b. ABSTRACT (Continue on revea Continue on revea A large vocabulary which are acoustic acoustic cues to the <i>file puipo</i> b. ABSTRACT (Continue on revea Facility under which	(of this Report) ic release; c (of the abstract entered if raw elde if necessary and y and a conti c-phoneticall word boundari <u>Se of Ho</u> the accustic of the	Istribution Block 20, 11 different Identify by block num nuous speed y based photes.	n unlimited nt trom Report) mber) ch recognition system, ontactic constraints Man is to new is to new is to
 DISTRIBUTION STATEMENT (Approved for publ DISTRIBUTION STATEMENT (DISTRIBUTION STATEMENT (SUPPLEMENTARY NOTES SUPPLEMENTARY NOTES SUPPLEMENTARY NOTES Auge vocabulary Alarge vocabulary Which are acoustic acoustic cues to s Mich are acoustic Contexts can be studie 	(of this Report) ic release; c fof the abstract entered in for the abstract entered in for the abstract entered in and a conting c-phoneticall word boundari se side if necessary and the acoustic chi- ed conveniently	Istribu+io Block 20, 11 different Identify by block nu nuous speed y based pho es.	n unlimited nt trom Report) mber) ch recognition system, ontactic constraints May is to new is to page 1 property of speech data base of speech sounds in various the phonological properties of
 DISTRIBUTION STATEMENT Approved for publ Approved for publ 7. DISTRIBUTION STATEMENT (8. SUPPLEMENTARY NOTES 9. KEY WORDS (Continue on rover A large vocabulary which are acoustic acoustic cues to a 7. DISTRACT (Continue on rover facility under which contexts can be studio a large lexicon of, s 	(of this Report) ic release; c (of the abstract entered in (of the abst	Istribution Block 20, 11 dillere Identify by block num nuous speed y based pho es. Identify by block num aracteristics ; investigate , and determi	n unlimited nt trom Report) mber) ch recognition system, ontactic constraints
DISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT 8. SUPPLEMENTARY NOTES 9. KEY WORDS (Continue on rover A large vocabulary which are acoustic acoustic cues to s	(of this Report) ic release; of for the abstract entered in for the abstract entered in for the abstract entered in and a conting c-phoneticall word boundari <u>Se of H</u> as side if necessary and the acoustic chi- ed conveniently ay 10,000 words a utilized in g	Istribution Block 20, 11 dillered Identify by block num nuous speed y based pho- es. <u>S prof</u> Identify by block num aracteristics ; investigate , and determing peech recogni	n unlimited nt from Report) the from Report) the recognition system, ontactic constraints
DISTRIBUTION STATEMENT Approved for publ Approved for publ . DISTRIBUTION STATEMENT . DISTRIBUTION STATEMENT . SUPPLEMENTARY NOTES . SUPPLEMENTARY NOTES . SUPPLEMENTARY NOTES . SUPPLEMENTARY NOTES . AUXILIARY WORDS (Continue on rover A large vocabulary which are acoustic acoustic cues to s . ADDE DUI DO . ADSTRACT (Continue on rover Cacility under which contexts can be studie a large lexicon of, s tic constraints can be that are used to mark Cocabulary. IWR system	(of this Report) ic release; ((of the obstract entered in (of the obst	Istribution Block 20, 11 dillere Identify by block num nuous speed y based pho es. <u>S prof</u> identify by block num aracteristics ; investigate , and determin peech recogni s; develop a interactions	n unlimited n unlimited nt tram Report) the recognition system, ontactic constraints
DISTRIBUTION STATEMENT Approved for publ Approved for publ . DISTRIBUTION STATEMENT . DISTRIBUTION STATEMENT . SUPPLEMENTARY NOTES . SUPPLEMENTARY NOTES . SUPPLEMENTARY NOTES . AUTOR Continue on revea A large vocabulary which are acoustic acoustic cues to a <u>Mic puipo</u> . ABSTRACT (Continue on revea Contexts can be studie a large lexicon of, s tic constraints can be that are used to mark cocabulary, IWR system cal knowledge; and der . ABSTRACT (Continue on revea . AB	(of this Report) ic release; c for the abstract entered in for the abstract entered in resc elde if necessary and y and a conti c-phoneticall word boundari so side if necessary and the acoustic chi ed conveniently ay 10,000 words e utilized in sy work boundaries in to study the velop a limited	Identify by block nu nuous speed y based pho es. <u>Sock 20, 11 differen</u> nuous speed y based pho es. <u>Social</u> identify by block num aracteristics ; investigate , and determi peech recogni s; develop a interactions continuous s	n unlimited n unlimited nt from Report) three Report) three Reports three Re
6. DISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT 8. SUPPLEMENTARY NOTES 8. SUPPLEMENTARY NOTES 9. KEY WORDS (Continue on rover A large vocabulary which are acoustic acoustic cues to re- <i>The puipo</i> 9. ABSTRACT (Continue on rover Cacility under which contexts can be studie a large lexicon of, s tic constraints can be that are used to mark cocabulary, IWR system cal knowledge; and der coal of recognizing and	(of this Report) ic release; of (of the abstract entered in (of the abstract entered in and a contic c-phoneticall word boundari <u>se side if necessary</u> and the acoustic chi ed conveniently ay 10,000 words a utilized in sy work boundaries in to study the velop a limited ny English word	Identify by block nu nuous speed y based pho es. <u>Speed pho</u> es. <u>Speed pho</u> es. <u>s</u> investigate , and determi peech recogni s; develop a interactions continuous speed pased pho es.	n unlimited n unlimited nt tran Report) ch recognition system, ontactic constraints
DISTRIBUTION STATEMENT Approved for publ 7. DISTRIBUTION STATEMENT 8. SUPPLEMENTARY NOTES 8. SUPPLEMENTARY NOTES 9. KEY WORDS (Continue on rover A large vocabulary which are acoustic acoustic cues to s <u>The purpo</u> 0. ABSTRACT (Continue on rover facility under which contexts can be studie a large lexicon of, s tic constraints can be that are used to mark rocabulary, IWR system cal knowledge; and der coal of recognizing an D 1 JAN 73 1473 EDITION	(of this Report) ic release; c for the abstract entered in res elde if necessary and y and a contic c-phoneticall word boundari <u>Sc of M</u> , se side if necessary and the acoustic chi ed conveniently ay 10,000 words e utilized in sy work boundaries in to study the velop a limited ny English word of 'Nov \$\$ 15 ossols	Identify by block nu nuous speed y based pho es. <u>Sock 20, 11 differen</u> nuous speed y based pho es. <u>Social</u> identify by block num aracteristics ; investigate , and determi peech recogni s; develop a interactions continuous s from its spe	n unlimited nt from Report) mber) ch recognition system, ontactic constraints

1933 AUGUAR AMPARAS MANAGAM

1

13 54 61 64 19

20. teractions of higher-level knowledge sources.

Accession For NTIS GRAAI X DTIC TAB Unannounced Justification By_ Distribution/ Availability Codes Avail and/or Dist Special

All the Carton

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 Merriman-Webster Pocket Dictionary has been carefully scrutinized in termsof 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

111111111111

ENVIRONM S

Ø

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

,

Sec. Carrie

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