

AD 717624



Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

TM-4652/001/00

INTRODUCTION AND OVERVIEW OF THE
VICENS-REDDY SPEECH RECOGNITION SYSTEM

4 December 1970

DDC
RECEIVED
FEB 8 1971
C

This document is being reproduced
for public release and sale; its
distribution is unlimited.

TECHNICAL MEMORANDUM

(TM Series)

The work reported herein was supported by the Advanced Research Projects Agency of the Department of Defense under Contract DAHC15-67-C-0149, ARPA Order No. 1327, Amendment No. 3, Program Code No. 1D30, and 1P10.

INTRODUCTION AND OVERVIEW OF THE
VICENS-REDDY SPEECH RECOGNITION SYSTEM

by

Iris Kameny

H. Barry Ritea

4 December 1970

SYSTEM

DEVELOPMENT

CORPORATION

2500 COLORADO AVE.

SANTA MONICA

CALIFORNIA

90406

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied of the Advanced Research Projects Agency or the U. S. Government.



Distribution of this document is unlimited.

4 December 1970

i
(Page ii blank)

System Development Corporation
TM-4652/001/00

ABSTRACT

This document provides an introduction and overview for the document series TM-4652, Analysis and Development of the Vicens-Reddy Speech Recognition System.

1. INTRODUCTION

For about a year, SDC has been involved in a program of development of voice communication with the computer. This has included studies of both automatic speech recognition and speech synthesis. Initial investigations have indicated that one of the most successful speech recognition systems was developed by Pierre Vicens (under the direction of Raj Reddy) on the PDP-10 computer at Stanford University. It was decided to implement this system on the IBM 360/67 at SDC to provide a base for further development.

This particular system is unique in the sense that it approaches the problem of speech recognition as a whole, rather than treating particular aspects of the problem as in previous attempts. For example, where earlier systems treated only segmentation of speech into phoneme groups, or detected phonemes in a given context, the Vicens-Reddy system processes the incoming speech signal, applies heuristics to segment the signal and to identify phoneme-like units and then uses the total phonemic pattern to recognize an entry in the lexicon.

The Vicens-Reddy system can be divided into six parts: (1) hardware preprocessing, (2) software preprocessing, (3) segmentation, (4) recognition, (5) lexicon development, and (6) lexicon usage. Preprocessing involves digitizing and normalizing the raw voice input. The output of the preprocessing is an array, called the Q-matrix, whose rows are the amplitudes and zero-crossing counts of the raw voice input over 10 ms. time periods.* Segmentation applies heuristics to combine the minimal segments of the above Q-matrix into larger transitional or sustained segments of a new array, called the P-matrix. Recognition applies heuristics to the P-matrix in classifying sustained segments into phoneme-like groups (fricative, vowel, stop, nasal, consonant or burst) and produces the R-matrix or feature matrix. The lexicon development involves the addition of the feature matrix from the previous recognition process into the lexicon. In lexicon usage, the feature matrix from the recognition process is used to achieve the best match in the lexicon.

* A 10 ms. time period will be referred to as a minimal segment throughout.

4 December 1970

2
(Last page)

System Development Corporation
TM-4652/001/00

Some amount of speech research at SDC will be based on this program. Such research requires detailed explanations of the various heuristics used in the program. However, the available documentation is not complete. This series of documents will begin with a detailed description of the Vicens-Reddy system, along with explanations for a large number of previously unexplained heuristics used in the program. Later documents will describe various modifications to the system.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification at title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) System Development Corporation Santa Monica, California		2a. REPORT SECURITY CLASSIFICATION Unclassified	
		2b. GROUP	
3. REPORT TITLE Introduction and Overview of the Vicens-Reddy Speech Recognition System			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Technical Report--July 1970 - December 1970			
5. AUTHOR(S) (First name, middle initial, last name) Iris Kameny H. Barry Ritea			
6. REPORT DATE 4 December 1970		7a. TOTAL NO. OF PAGES 2	7b. NO. OF REFS 0
8a. CONTRACT OR GRANT NO. DAHC15-67-C-0149		8a. ORIGINATOR'S REPORT NUMBER(S) TM-4652/001/00	
8. PROJECT NO. ARPA Order #1327, Amendment #3, Program c. Code #1D30, and 1P10 d.		8b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
10. DISTRIBUTION STATEMENT Distribution of this document is unlimited.			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY	
13. ABSTRACT This document provides an introduction and overview for the document series TM-4652, Analysis and Development of the Vicens-Reddy Speech Recognition System.			

DD FORM 1473
1 NOV 65

UNCLASSIFIED

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT