

SPEECH SYNTHESIS UTILIZING
MICROCOMPUTER CONTROL

BY

JOSEPH N. UZEL

B.S.A.E., Virginia Polytechnic Institute, 1972

B.S.E., Florida Technological University, 1977

RESEARCH REPORT

Submitted in partial fulfillment of the requirements
for the degree of Master of Science: Engineering
in the graduate studies program of
the College of Engineering
of Florida Technological University at Orlando, Florida

Fall Quarter
1978

SPEECH SYNTHESIS UTILIZING
MICROCOMPUTER CONTROL

JOSEPH N. UZEL

ABSTRACT

This report explores the subject of speech synthesis. Information given includes a brief explanation of speech production in man, an historical view of speech synthesis, and four types of electronic synthesizers in use today.

Also included is a brief presentation on phonetics, the study of speech sounds. An understanding of this subject is necessary to see how a synthesizer must produce certain sounds, and how these sounds are put together to create words.

Finally a description of a limited text speech synthesizer is presented. This system allows the user to enter English text via a keyboard and have it output in spoken form.

The future of speech synthesis appears to be very bright. This report also gives some possible applications of verbal computer communication.

Approved by:

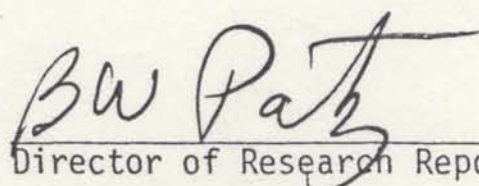

Director of Research Report

TABLE OF CONTENTS

List of Tables iv

List of Illustrations v

Introduction 1

 I. Human Speech Production 2

 II. A History of Speech Synthesis 5

III. Phonetics 9

 IV. Electronic Speech Synthesis 14

 Adaptive Differential Pulse Code Modulation 14

 Linear Predictive Coding 17

 Formant Synthesis 20

 Text Synthesis 30

 V. Microcomputer Controlled Text Synthesizer 33

 Hardware 33

 Software 37

 Results 43

 VI. Conclusions 47

Appendix A

 Text Synthesizer Vocabulary Assembly Listing 49

Appendix B

 Text Synthesizer Program Assembly Listing 56

References Cited 61

LIST OF TABLES

1. IPA Phonetic Symbols	10
2. Phoneme Classes	13
3. Comparison of Data Rates for Storage of Speech Vocabularies	16
4. Steady State English Vowel Formants	23
5. Fricative Spectra	30
6. Votrax Phonetic Codes	36
7. ASCII Code Conversion Chart	40

LIST OF ILLUSTRATIONS

1. The Human Vocal System	3
2. Block Diagram of the VODER	7
3. Block Diagram of a Computer Voice Response System	15
4. Block Diagram of Adaptive Differential Coding of Speech . . .	18
5. Waveforms Coded by 3-Bit (a) DPCM and (b) ADPCM	19
6. Linear Predictive Coding Used in the TMC0280 Single Chip Synthesizer	21
7. Formant Synthesizer Block Diagrams (a) Series (b) Parallel . .	22
8. Relationships Between the First Two Formants of Steady State English Vowels	25
9. Formant Sweeps for the English Diphthongs	26
10. Stop Consonant Patterns	27
11. Parallel Formant Circuit Used in the A1 Cybernetics Systems Model 1000 Speech Synthesizer	29
12. Block Diagram of a Text Synthesizer	31
13. Synthesizer Circuit of the VOX-1 Board	35
14. Partial Results of the "L" Command	41
15. Illustration of Program Features	44

INTRODUCTION

For centuries man has been fascinated with the concept of artificially produced speech. Mechanical analogs of the vocal system produced speech sounds 200 years ago; however, it was the recent invention of the digital computer which made practical speech synthesis feasible.

The purpose of this report is to explore the subject of speech synthesis. Questions to be answered are: (1) How is speech produced in man? (2) What approaches have been taken in the past to synthesize speech? (3) What methods are presently used to electronically produce artificial speech?

In addition to answering the above questions this report also presents a description of a limited text speech synthesizer which utilizes a Votrax speech module and is controlled by a 6800 microcomputer. The control software allows the user to enter English text via a keyboard or to store coded text in a memory table and then have the text output in spoken form by the machine.

Some applications of speech synthesis include talking typewriters, verbal response learning systems, aircraft flight control, and telephone information retrieval systems.

CHAPTER 1

HUMAN SPEECH PRODUCTION

Human speech production is a physiological process which is understood fairly well. Figure 1 shows the human vocal tract and those components which make speech possible. The vocal tract is a non uniform acoustic tube, 16 to 18 centimeters in length, which extends from the glottis to the lips, and varies in shape as a function of time. The major anatomical components causing this time-varying change are the lips, tongue, jaw, and the velum. The velum is a flap which couples the nasal tract to the vocal tract through a trap-door type of action. The nasal cavity is about 12 centimeters long and has an approximate volume of 60 cubic centimeters(Flanagan 1972a).

The vocal system can produce three basic types of sound: voiced, fricative, and plosive. Voiced sounds, such as the vowels, are produced by the vibration of the vocal chords due to air released from the lungs. These vibrations interrupt the airflow and generate a series of sharp pulses that excite the vocal tract.

Fricative sounds(s,sh,f,th) are generated by forcing air through a constricted vocal tract at a high velocity which causes turbulence. Plosives(p,t,k) are produced by closing the vocal tract completely with the lips or tongue, allowing a pressure buildup, and then abruptly opening the closure.

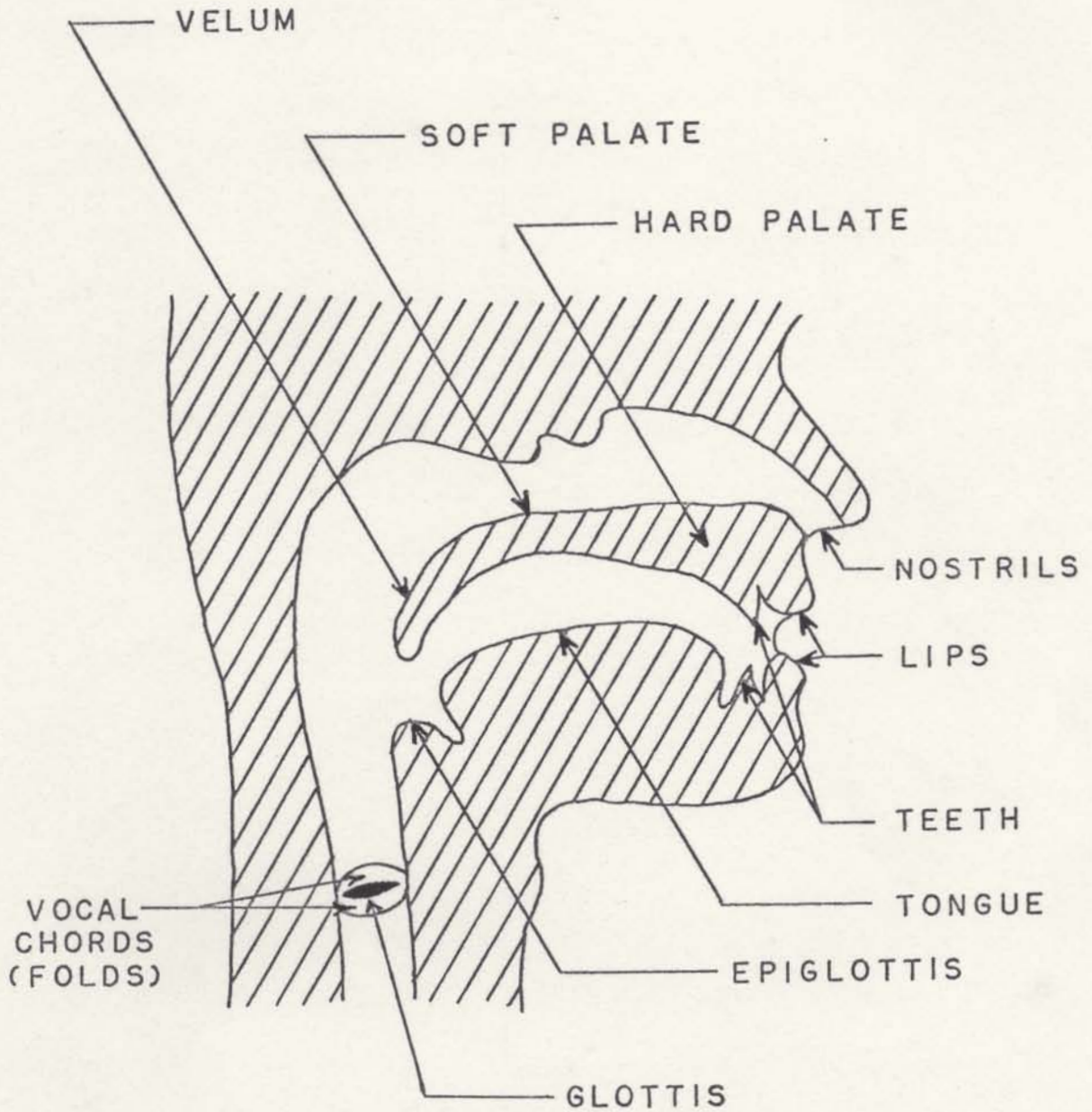


Fig. 1. The Human Vocal System

