R T I

RESEARCH TRIANGLE INSTITUTE

(NASA-CR-170438) DEVELOPMENT OF A SPEECH AUTOCUER Quarterly Report, 1 Oct. - 31 Dec. 1980 (Research Triangle Inst., Research Triangle) 17 p HC A02/MF A01 CSCL 17B

N82-32569

Unclas

G3/32 33564

DEVELOPMENT OF A SPEECH AUTOCUER

Autocuer Quarterly October 1, 1980 to December 31, 1980

NASA Contract No. NAS5-25832

RTI/1878/00-05Q

Prepared for
National Aeronautics and Space Administration
Goddard Space Flight Center
Greenbelt, Maryland 20771

Robert L. Beadles
Paul N. Kizakevich
Dewey T. Lawson
Michael L. McCartney

RESEARCH TRIANGLE PARK, NORTH CAROLINA 27709

TABLE OF CONTENTS

1. INTRODUCTION

This is the fifth quarterly progress report for the autocuer project. The autocuer project is supported by National Aeronautics and Space Administration Contract No. NAS5-25832, from Goddard Space Flight Center. The purpose of the autocuer project is to fabricate and field test a wearable, visually based prosthesis for the deaf based upon the proven method for removing lipreading ambiguity known as Cued Speech.

The reporting period covered herein is October 1 through December 31, 1980. Project activities during the period are summarized in the pages following.

2. PROJECT MANAGEMENT ACTIVITIES

The autocuer project review was held at RTI on November 6, 1980. In addition to RTI personnel, these people attended:

Gallaudet College
- R. Orin Cornett

NASA
- Donald S. Friedman
- Michael Peck
- James Hogue

Stanford University
- John G. Linvill

Telesensory Systems, Inc.
- James Bliss
- Steve Brugler
- Jared Bernstein

Veterans Administration
- Frank Coombs

A project review was held on November 17 at Gallaudet College to provide an update to interested parties from government agencies other than NASA and the Veterans Administration.

Of particular importance to the project was the approval by the RTI Board of Directors on December 15 of the exclusive license agreement and contract with TSI for the manufacture of the field test autocuer and the commercialization of the autocuer following the field test. Details on this and other commercialization activities are given in Section 6.

3. SOFTWARE DEVELOPMENT

3.1 Linear Prediction and Speech Database

During this quarter we have used our own additions to the speech database to complete an evaluation of analyzer phoneme identification errors vis-a-vis the BB&N database labeling. The results have influenced both the analyzer design and our procedures for database label assignment. In the process we also have refined our techniques for maximizing the project benefits from relatively small additions of labeled speech data.

In the light of this work we have planned one final addition to the database for use both in the final stages of software design and in prototype evaluation and testing before production of the field test units. This addition will be a subset of the sentences presently in the database (to facilitate comparisons with previous analyses); recorded with the autocuer design microphone mounted in the design templepiece; in a realistic acoustic environment (reverberation, ambient noise, etc.); spoken by a female subject who routinely communicates with her profoundly deaf child via lipreading and manual cues. The sentences will be embedded in a context designed to facilitate natural reading. The sample sentences will be digitized linearly and also digitally encoded using the codec that is part of the autocuer design.

As the designs of various parts of the linear prediction analysis have reached their final forms we have converted them to fixed point modules within the database analyzer and documented them for rapid conversion to microprocessor object code. Several factorization schemes for speeding up the initial autocorrelation calculation of linear prediction coefficients have been evaluated. The most promising of these are being tested for their effect on fixed point coefficient precision and stability.

4. HARDWARE DEVELOPMENT

An engineering evaluation sample of the NSC800 CMOS microprocessor was received on November 5, indicating that its revised development schedule will probably be met. RTI plans for orderly integration of the NSC800 into the field test autocuer design are described below.

4.1 Microcomputer Development

During this quarter, we concluded the development of tools necessary for the autocuer hardware and software development. The Pro-Log/Mostek STD Bus was selected as a foundation for the implementation and testing of autocuer hardware subsystems.

A STD Bus system was assembled which consists of (1) a Z80 (NSC800 compatible) microcomputer module, (2) 48K RAM memory, (3) dual single-density floppy disks, and (4) three serial interfaces for console, line printer and NOVA database link. In addition, we interfaced a general-purpose circuit breadboard to the STD Bus for the initial test of autocuer subsystem designs. Each subsystem will then be fabricated using wire-wrap techniques on STD Bus modules inserted in the development system for evaluation, then made available for autocuer software development and verification.

To support the STD Bus development system, we acquired the CP/M Version 2.2 operating system and adapted it to our memory and input/output configuration. Although the CP/M adaptation should have taken one manweek, we encountered a variety of CP/M system failures which could not be easily resolved. After repeated examination of our CP/M system software modifications involving an additional two manweeks, we could not identify any source of error. We then conducted a systematic check of all hardware components and identified the problem as a faulty motherboard in the STD Bus cardcage. With this problem resolved, we now have a 48K CP/M operating system executing on the STD Bus development system.

4.2 Display Development

Problems with aluminum adherence to the display substrate were traced to surface preparation and to insufficient aluminum thickness. Both were corrected by having NASA/Goddard prepare a set of 20 substrates.

Using the same surface preparation materials, we have aluminized several sets of high curvature lenses and have experimented with protective coatings of SiO. Ideal protection is reputed to be provided by 250 nm of SiO, but with our vacuum system, the lens temperature becomes high enough to blister the aluminum coating. Thinner coatings, <100 nM, appear to offer adequate scratch and oxidation protection with less heating and less coloration.

A working laboratory prototype of the display subsystem, including microphone, was presented at the November review meeting. Modifications to this prototype are continuing, to make it both cosmetically more acceptable and simpler to manufacture.

Available plastic eyeglass frames are not sufficiently uniform to allow adequate control over lens-to-display-substrate spacing and alignment when the lens is mounted in the frame. Consequently, we will make the lens and substrate a module which will then mount in the eyeglass frame. A substrate-to-cable connection is being evaluated, which uses both flexible connectors (Tecknit zebra strips) and flexible circuitry (Sheldahl copper-clad polyester and polyimide films).

American Optical Company of Raleigh no longer does any lens grinding, so we have lost our primary local supplier of high-curvature lenses. An alternate shop, White Optical of Durham, is making two additional pairs, but is reluctant to make any more (efficient grinding of these lenses requires special setups and special tools). Also, the reflective coatings we have put on the high-curvature lenses are not uniform and are usually too thick, thickness monitor notwithstanding. Both of these problems can be solved by employing an optical specialty house for the grinding and coating. Having experienced the lens manufacturing and coating problems ourselves, we can now write a specification for this task that will enable the optical house to perform the task rapidly and efficiently.

4.3 Speech Preprocessor

Directional microphones (Knowles) suitable for mounting in eyeglass templepieces are being evaluated and used to record additional speakers for the database. Low-power analog amplifiers and filters have been designed and fabricated. These circuits were employed in the speech analysis demonstration at the November review committee meeting.

The Mostek 5116 Codec has also been evaluated as a compression A/D converter. This device with its support circuitry has been configured to provide a 10 kHz sample rate for a 100-5000 Hz input bandwidth. The dynamic range is essentially equivalent to that of a 12-bit linear A/D converter. Our configuration routes the 8-bit encoded serial data output from the Codec to a CMOS UART where it is converted to 8-bit parallel form for storage in RAM. Total power consumption of this speech preprocessor, including Codec and UART, is 45 mW.

5. TRAINING AND TESTING ACTIVITIES

To provide baseline data and to expose the student subjects at Gallaudet to material containing errors for the first time, a randomized list of the 500-word trained vocabulary was analyzed at RTI using the zero-crossings based analyzer. Concurrently, retraining of the Gallaudet student subjects (six) was continued on videotaped correctly-cued material. Subjects were then retested on correctly-cued words. They had been tested on words in February '80 scoring an average of 85.5% on correctly-cued material. The new test results (October '80) showed that they were at the same level of reading the correctly-cued word output (average of 85%) after 15 hours of retraining. The initial level had been reached after original training of about 45 hours. Pretesting of uncued material before training yielded average scoes of 49%.

Some training was done on correctly-cued sentence materials. Based on an initial pretest on sentence material given at the start of the project (Fall '79), a test on correctly-cued sentence material was given. This activity was terminated before all subjects had been tested because of the arrival of the speech-analyzer output. RTI speech-analyzer output was incorporated into videotapes at Gallaudet on 11/17/80, and training and testing with this material was begun. A 200-word pretest was given (100 words per session) on uncued lipreading of words from the basic 500-word list used previously, but with random order. For results, see Table I.

Additional speech-analyzer output was received from RTI immediately after this pretesting. This made it possible to complete a 100-word speech-analyzer-cued test before the end of the semester. Raw scores are shown in Table I. Training and testing activities were terminated for the semester on 12/11/80.

We have only begun the detailed analysis of the results of the test administered just before the end of the semester. Overall, uncued scores averaged 58.1%, while the speech-analyzer cued scores averaged 62.0%, about what we expected for this level of analyzer accuracy (~55% right, 35% wrong, 10% missing).*

The effects of cue errors when lipreading ambiguity is high are as expected for words in isolation. These examples are illustrative:

- The word tall was cued as if it were call. Five of the six subjects recorded it as call, one as car.
- The word made was cued as make. Four subjects recorded it as make, two as may.
- The word too was cued as tool. Four subjects recorded tool, two recorded too.
- The word day was cued as tay. All subjects recorded a word with t in it, either tay or stay.

In all these cases, there is little to no evidence on the lips to help make a correct response. Thus a cue error in this setting usually causes a word error. What the effect of an equivalent error would be for connected speech--where context is available--must await the field test.

While waiting for new speech-analyzer based cued output to resume training and testing of the returning Gallaudet student subjects (five), we will complete the evaluation of the effects of analyzer errors for this test. Also, the design for the field test will be revised. The implications of error rate and the kind of training needed to enable the subjects to cope with errors will be taken into account in the field test design.

*Since the analyzer was optimized for connected speech, its accuracy is lower on isolated words. Thus these results represent a worst-case baseline.

TABLE I. TESTING RESULTS

	PRETEST UNCUED									POST TEST CUED (Speech Analyzer)		
	A1 - 100			A101 - 200			A1 - 200			B401 - 500		
Subject	√	X	%	√	X	%	√	X	%	√	X	%
1	60	40	60	56	44	56	116	84	58	56	44	56
2	72	28	72	57	43	57	129	71	64.5	65	35	65
3	68	32	68	56	44	56	124	76	62	71	29	71
4	59	41	59	50	50	50	19	91	54.5	56	44	56
5	51	49	51	48	·2	48	99	101	49.5	52	48	52
6	64	36	64	58	42	58	122	78	61	72	28	72
AVERAGES		62			54			58			62	

6. COMMERCIALIZATION

All the remaining steps to securing Telesensory Systems, Inc. as the manufacturer to commercialize the autocuer were accomplished during the reporting period.

On October 17, the autocuer system and the new display patent applications were submitted to the United States Patent Office. These applications are the basis for the exclusive license agreement with TSI.

In mid-November, the TSI board of directors approved the final drafts of the exclusive license agreement and the contract. After being signed there, the documents were returned to RTI, then sent to Gallaudet where they were signed on December 6. As the final step to the commercialization agreement, they were approved by the RTI board of directors on December 15 and signed by the designated RTI representative.

Preparation of the RTI petitions to NASA for waiver of rights to the two inventions was deferred until we receive a legal opinion on the effects (if any) of the new Patent Act (signed into law on December 12, 1980) on preparation of the petitions. We expect to receive that opinion early in January.

7. PROJECT MILESTONE AND COST REPORT

Approval of the request for a six-month project extension was received on October 15. To reflect the new schedule, the milestone list was jointly revised by Goddard and RTI on October 21. The revised milestone list is shown in Table II. Following the November 6 project review, the dates for milestones 17 and 18 were changed from those set on October 21 to those shown in Table II. This was done to given TSI a full six months for fabrication of the field test units per the terms of the TSI-RTI-Gallaudet contract. The following milestone events are noted:

Milestone 1--Met on December 15 when decision was made to use Mostek 5116 Codec as the preprocessor.

Milestone 22--Real-time cue generation from connected speech was demonstrated at project review on November 6.

Milestone 27--Draft field test plan was completed on 9/30/80, discussed at the November 6 review.

Milestones 35A and 36--TSI representatives attended the November 6 project review. Final signature on three-party license agreement/ contract was secured on December 15, 1980.

Project cost status as of December 31 is shown in Figure 1 following.

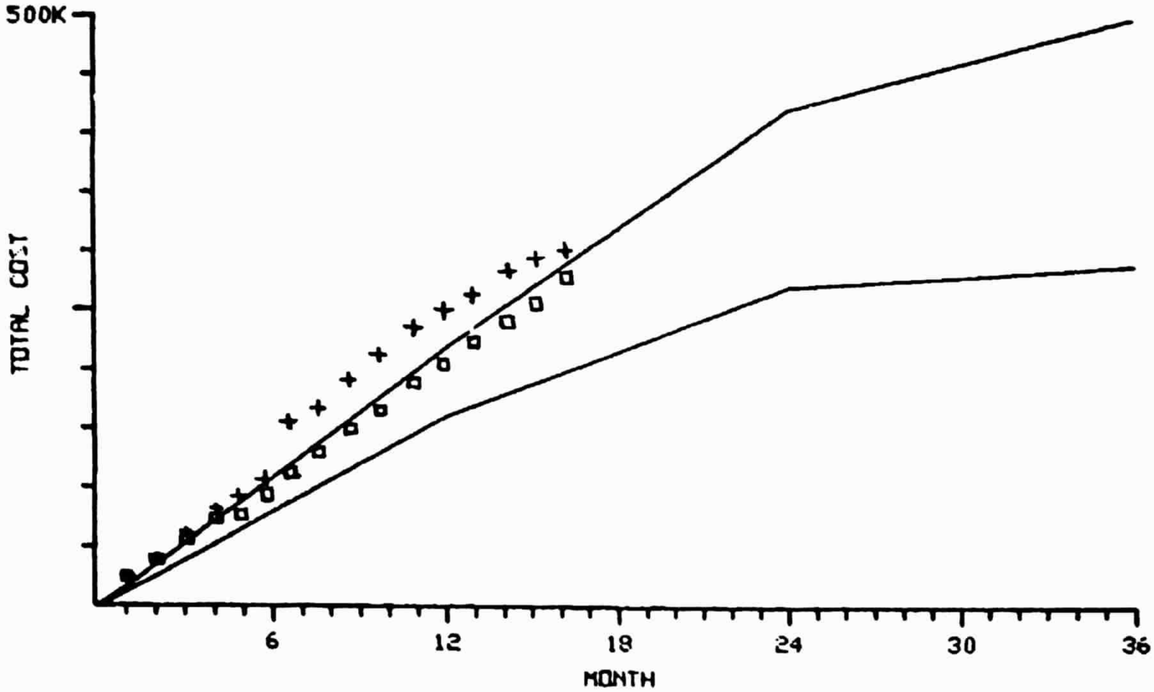
TABLE II. AUTOCUER PROJECT MILESTONES

Milestone	Event Description	Date
1	Complete software and hardware studies of preprocessor options.	1/1/81
2	Submit final preprocessor specifications to GSFC. GO/NO-GO decision on use of GSFC parts.	2/1/81
3	Complete comparative study of analog vs. digital approaches to pitch tracker implementation.	12/15/79
4	Select and finalize pitch tracker design.	1/15/80
5	Select optical techniques for field test prototypes.	5/1/80
6	Decision to use (omit) telemetry in display system.	12/1/79
7	Complete power consumption studies.	2/1/80
8	Finalize display system design and complete laboratory prototype.	7/1/80
8A	Make or Buy decision on display subsystem.	7/15/80
9	Subcontract for field test prototype display systems.	9/1/80
10	Delivery of field test prototype display systems.	6/1/81
11	Complete comparative analysis of commercially available microprocessors with PASCAL software.	8/1/80
12	Select IC chip set for autocuer field test unit.	8/1/80
13	Complete electrical and mechanical designs.	3/15/81
14	Complete breadboard, wirewrap prototype.	4/15/81
15	Complete all functional tests and run diagnostic program to verify electrical design.	5/1/81
16	Finalize field test unit mechanical and electrical designs, order ICs, release purchase order for circuit boards and packages.	5/15/81
17	Installation of board-level components and board-level hardware and software tests complete. Delivery of GSFC-fabricated circuits and units.	8/1/81
18	Final assembly of RTI-produced field test units complete. Two working units delivered to TSI.	8/15/81

AUTOCUER PROJECT MILESTONES (Con.)

Milestone	Event Description	Date
19	Complete functional and operational tests.	12/1/81
20	Complete general-purpose speech analysis software tools.	9/15/79
21	Analyzer sampling rate reduced to approximately 10 kHZ.	7/1/80
22	Cue generation in real time on connected speech. Project GO/NO-GO decision.	10/15/80
23	Complete autocuer diagnostic software.	5/15/81
24	EPROM programming hardware and software functional.	6/1/81
25	Operational test software complete.	10/15/81
26	Release units to field test subjects.	2/15/82
27	Draft field test plan complete.	9/30/80
28	Final field test plan complete.	2/15/81
29	FDA approval of field test plan.	6/15/81
30	Complete field test subject pretraining.	1/15/82
31	Field test subject pretest.	2/15/82
32	Field test subject midpoint test.	8/15/82
33	Field test subject endpoint test.	2/15/83
34	Request for proposal.	11/15/79
35	Selection of manufacturer.	5/1/80
35A	Signed contract with manufacturer	11/1/80
36	Manufacturer begins participation in hardware design.	11/1/80
37	Complete delivery of components and subassemblies to manufacturer.	8/15/81
38	Commercialization target date.	2/15/83
39	Year 1 end GO/NO-GO decision.	1/1/81
40	Year 2 end GO/NO-GO decision.	1/1/82

EXPENDITURE SUMMARY OF RTI PROJECT 42U-1878
TITLE: AUTOCUER
SPONSOR: NASA AND THE VA (NASA CONTRACT NO. NAS5-25832)
START DATE: AUGUST 20, 1979



NOTE THAT THE SOLID LINE INDICATES AN EVEN ALLOCATION OF CONTRACTED FUNDS OVER TIME; THE SQUARES INDICATE ACTUAL LABOR EXPENDITURES AT MONTHLY INTERVALS; THE PLUS SIGNS INDICATE TOTAL CONTRACT COST AT MONTHLY INTERVALS.

Figure 1. Autocuer project cost summary.