DOCUMENT RESUME

ED 189 624 CS 205 704

TITLE	Automated Dictionaries, Reading and Writing: Chairman's Report of a Conference on Educational Uses of Word Processors with Dictionaries (New York, New York, June 14-15, 1979).
INSTITUTION	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Dec 79
NOTE	31p.
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Computer Assisted Instruction: *Dictionaries: Educational Research: Elementary Secondary Education: *English Instruction: Futures (of Society): *Instructional Materials: Lexicography: Linguistics: Technological Advancement
IDENTIFIERS	*Word Processing

ABSTRACT

This booklet summarizes a discussion of the opportunities provided by computerized information-handling technology to improve student achievement in reading and writing. The first section discusses the development and educational use of an automated dictionary (AD) that would allow a student to designate a word by typing it and receive information about it, either aurally by electronically generated speech or visually on a cathode ray tube screen. Also discussed are the possibilities of incorporating an AD in a word processor and the larger contexts in which these systems would function. The second section outlines the technical implementation of computer hardware and software in the classroom. The third section describes the kinds of scientific research in lexicography, linguistics, psychology, and education that will be required to realize the full potential of such systems. The original proposal made to the National Institute of Education suggesting the feasibility of ADs is appended. (AEA)

AUTOMATED DICTIONARIES, READING AND WRITING

Chairman's Report
of a
Conference on
Educational Uses of
Word Processors with Dictionaries

June 14-15, 1979

George A. Miller
Chairman

U.S. Department of Health, Education, and Welfare
Patricia Roberts Harris, Secretary
Mary F. Berry, Assistant Secretary for Education

National Institute of Education
Michael Timpane, Acting Director

December 1979

The New Information-Handling Technology*

We have entered a revolutionary age in the technology of information handling The microelectronic revolution, extensively documented in the September 1977 issue of *Scientific American*, is making possible a roughly tenfold decrease each five years in the cost of the integrated circuits that are at the heart of contemporary digital computers. There are available for individual use, at relatively low prices, sophisticated calculators and computers and means of generating images on video tubes:

Handheld calculators providing the four arithmetic functions and square root now sell for about $10, and the cost of calculators providing trigonometric and logarithmic functions has declined to the same level. Calculators able to handle 400 step programs are now at the $100 level.

Handheld instructional devices providing drill and games in arithmetic ($15 25) and in spelling, with a simulated human voice ($65), are selling well.

Personal computers which include a TV tube display for letters, numbers, and graphics; keyboard, processor, a sizeable memory, and weigh less than 50 pounds, are now available for as little as $600. Accessories include word-processing devices allowing flexible construction and editing of written text and production of typed copy. These computers are comparable in capacity to computers costing hundreds of thousands of dollars a decade ago, and are only the first of many that will become available at decreasing prices or with much greater capability at the same price. Incorporation of microcomputers into TV sets could make substantial computer power available in the home to each child.

A *videodisc* system coming onto the market for $700 provides the following, in conjunction with a regular TV set:

- 54,000 separate frames of full color picture image, or alphanumeric or computer information, on one side of a disc similar in size and cost to an LP record. The 54,000 frames correspond to 30 minutes of video at 30 frames per second, any part of which can be played at regular or reduced speed.
- Random access (dialed in more expensive models) to each of the frames, which are individually numbered and can be viewed individually for any length of time desired and may contain one quarter of a page of easily legible text (or a full page with special high resolution video tubes).

Intelligent videodisc systems under development include a computer which allows controlled sequencing of frames, which may be based on student responses, as well as the possibility of transfering computer programs from the disc to the computer. Video discs can store 10 billion bits of information (*Encyclopedia Britannica* contains 2 billion bits; a human chromosome has a capacity of about 20 billion bits, and the human brain perhaps 10,000 billion). This immense storage capacity could probably include on a single disc all the computer courseware ever published.

*Reprinted from *Testing, Teaching and Learning, Chairmen's Report of a Conference on Research on Testing,* National Institute of Education, October, 1979.

FOREWORD

The National Institute of Education (NIE), the primary federal agency for educational research and development, sponsored the conference reported on here with the aim of exploring opportunities provided by the new information-handling technology (see box) to improve student achievement in reading and writing.

Factors contributing to this opportunity include the sharply declining cost of information-handling technology, including word processors; improved knowledge and technical capacity for handling strings of characters and words with computer logic, for comparing them with a lexicon of the English language and drawing inferences on their syntax and meaning; and modern information-processing models of human performance and learning.

The specific stimulus for this conference was a paper prepared by George A. Miller at an NIE-sponsored conference on Testing, Teaching and Learning, August 17-26, 1978. It is reprinted here as an appendix to this report.

Arthur S. Melmed
Office of Special Studies
National Institute of Education

CONTENTS

FOREWORD iii

INTRODUCTION 1

GOALS 3
Automated Dictionary 3
Word Processor with Dictionary 3
Commercial WP/D systems 4
Educational WP/D systems 4

TECHNICAL IMPLEMENTATION 5
Hardware 5
Software 7
Lexicographic aspects 7
Linguistic aspects 10
Psychological aspects 10
Educational aspects 12

SCIENTIFIC BASE 17
Lexicographic Research 17
Linguistic Research 18
Psychological Research 18
Educational Research 19
Comment on Research 19

RECOMMENDATIONS 21

REFERENCES 23

APPENDIX 25

INTRODUCTION

In August 1978, George Miller proposed to the National Institute of Education (NIE) that automated dictionaries (see Appendix) might serve a variety of educational purposes. In March 1979, Arthur Melmed asked Raj Reddy to undertake an analysis of the feasibility and cost of such a device. The preliminary analysis, prepared by M.S. Fox, D. Bebel, and A. Parker of Carnegie-Mellon University revealed that retrieval systems incorporating modern computer technology could easily make automated dictionaries widely available. On June 14-15, NIE sponsored a small conference in New York at the Rockefeller University to explore further the educational value of automated dictionaries. George Miller chaired the meeting. Present were:

Mark Aronoff, State University of New York at Stony Brook
John Parker Damon, McCarthy-Towne School, Acton, Mass.
Lawrence T. Frase, Bell Laboratories
John M. Mays, National Institute of Education
Arthur Melmed, National Institute of Education
George A. Miller, The Rockefeller University
Seymour Papert, Massachusetts Institute of Technology
Raj Reddy, Carnegie-Mellon University
Sylvia Scribner, Center for Applied Linguistics

The group discussed a wide range of topics. Subsequently, Miller prepared a document summarizing the sense of the meeting. That preliminary version was circulated to the participants, whose additions and corrections are reflected in the present report.

GOALS

Advances in computer technology have reduced the cost of powerful microprocessors until small businesses, schools, and even private individuals can now afford to purchase them. As one facet of its responsibility to evaluate educational implications of new technology, NIE is interested in possible applications of these new systems to education in general. One aspect of this general interest is the possible application of such systems in the teaching of communicative skills based on what Miller had called automated dictionaries.

Automated Dictionary

The simplest system that would qualify as an automated dictionary (AD) would allow a user to designate a word (e.g., by typing it) and receive information about it (e.g., its dictionary entry), either auditorily (e.g., by electronically generated speech) or visually (e.g., on a cathode ray tube). Although the original purpose of the conference was to advise NIE concerning practical goals for the development and educational use of AD systems, the group also discussed larger contexts in which such systems might function.

Word Processor with Dictionary

Seymour Papert persuaded the group that, if the ultimate aim is to facilitate the acquisition of communicative skills (particularly reading and writing), the most useful device to give students is a word processor (see box). For an adult writer, revising and editing a draft text is an important part of the process of writing. Children seldom have that experience. Their first draft is almost always their final text; in a pencil-paper-eraser technology, the sheer mechanical messiness of it makes revision difficult. A word processor could remove an obstacle to one of the central skills of writing. Editing is simple on a good word processor and the product is legible, attractive, and easily stored, scored, or duplicated. Moreover, inasmuch as lexical knowledge is a component of writing and reading, the acquisition of lexical knowledge should not be artifically separated from the mastery of the more general skills.

The focus of discussion broadened, therefore, to include the possibility of incorporating an AD in a word processor. Two uses of word processors with dictionaries (WP/D) were considered.

Commercial WP/D Systems

Raj Reddy suggested that, given a choice between some existing word processor and the same system plus lexical information, most people would prefer the WP/D system. If his evaluation is correct, manufacturers of word processors could probably be persuaded to make their products more attractive by the addition of a lexical retrieval system, and some of the development costs of an educationally useful device would be borne by commercial users.

Moreover, once the lexical feature became available on business machines, less expensive versions would probably be offered to hobbyists for home computers. Conference participants familiar with the history of computer technology were confident that, if WP/D is commercially successful, it will eventually become available to educators if they are simply willing to wait for it.

Educational WP/D Systems

It is unlikely that a WP/D system designed for offices would be optimal for classrooms. The size of vocabulary and the detail of lexical entries might be very different in an educational system. Moreover, an educational version should be light, sturdy, and attractive to children, its operation should not disrupt other classroom activities, and it should offer a variety of language games that would be inappropriate in most offices. Some participants in the conference argued that the appropriate classroom system is a general purpose computer: word processing (with or without lexical resources) is but one of many educational uses of these versatile machines. Discussion at the conference, however, was confined to AD and WP/D systems.

An appropriate goal for NIE would be to establish a continuing program of evaluation of educational AD and WP/D systems in appropriate social and educational contexts. In order to effect such a program it will be necessary to encourage the design, construction, and programming of AD and WP/D systems, to support research intended to establish pedagogical principles to be incorporated in them, and to support studies of their use in representative educational environments. Active coordination by NIE will be required, since three independent disciplines will have to cooperate: computer science, lexicography, and psychology.

TECHNICAL IMPLEMENTATION

In order to assess the educational value of AD or WP/D systems it is necessary to have one or more of them to experiment with. Such systems can be implemented with existing hardware if we are willing to make whatever compromises are necessary in the data base and the programming in order to get them operating. It is recommended that such instantiations of the concept be undertaken as soon as possible, but with full realization that improved models will evolve as experience accumulates and as further research can support technical improvements in both the hardware and software.

Hardware

Although WP/D systems open up different educational applications from an AD alone, the two systems pose similar technical problems—in both, the central problem is how lexical information is to be stored and retrieved. The major difference between them seems to be that a handheld AD is feasible at the present time, whereas a similarly small and self-contained WP/D probably is not. Since the problems are so similar, the following discussion of hardware is limited to the implementation of an experimental AD system.

One way to realize an AD would employ a videodisc. A videodisc has 54,000 frames; each frame could contain one or several dictionary entries (e.g., lists of homographs); a videodisc playback under computer control could display any frame in less than 2 sec. Videodisc systems are currently available for experimental purposes—all that is required in order to create an experimental system is for someone to prepare and photograph enough entries to make the videodisc.

Videodiscs might also be useful for graphic displays related to dictionary entries: for the kinds of illustrations used in dictionaries, or for maps, portraits, pictures of famous places, or artwork relevant to some particular cirriculum. Retrieval could be triggered by the same signal that triggers retrieval of written information. Illustrations could be displayed on a color TV spatially adjacent to the AD or WP/D display. This arrangement would allow experimental comparisons of AD or WP/D systems with and without correlated graphic displays.

Greater flexibility in the size of dictionary is possible with Read-Only-

Memory (ROM), Random-Access Memory (RAM), an external device (standard or floppy disk), or, eventually, electro magnetic bubble memory. Fox, Bebel, and Parker analyzed three categories of AD (handheld, stand-alone, terminals) and estimated the present costs to build one prototype model of these systems (between $2,000 and $36,000, depending on size of dictionary, category of device, type of memory, number of terminals).

Those estimates do not include the costs of software, economies of scale, or anticipated reductions in the cost of memory. Fox, Bebel, and Parker estimated that, with volume production, a 3000-word handheld AD would cost $2,000 at present and a 10,000-word AD would cost $5,000. In 1988 the costs (in 1979 dollars) would be $200 and $300, respectively; in 1988 a 30,000-word handheld AD would cost $700. These estimates are probably high. Cost reductions are possible because the electronic components (CPU and ROM circuits) are expected to decrease in price much more rapidly than the nonelectronic hardware (keyboard, power supply, display), and because large memory is a dominant aspect of AD systems. These economies also suggest that an AD could be added to a commercial word processor without increasing its cost significantly.

Even at these lower prices, however, handheld AD systems will not complete with printed dictionaries: an unabridged dictionary offers half a million entries for less than $100. Smaller dictionaries are useful to young children, however, and other benefits than lexical information should be offered by AD systems. The challenge, therefore, is to process lexical information in an AD system in active, useful ways that will be attractive to students and teachers.

With appropriate software, any of the lexical games suggested by Miller could be incorporated. The games under-utilize the CPU capabilities, which suggests that more complex functions could be performed, e.g., text editing, sentence parsing.

Input and output (I/O) devices should be designed to facilitate easy, natural interaction between child and machine. We expect that progress in achieving interaction will require considerable imagination and trial-and-error. As a starting point, however, light pens, or overlays sensitive to the pressure of a finger, might be used instead of keyboards. Acoustic I/O should also be considered: although electronic speech recognition is not yet reliable enough for general use, it will be; electronically produced speech is currently intelligible and can be substituted for written phonetic symbols. Speech processing requires memory, of course, and so reduces correspondingly the storage available for other lexical information. Until circuits are perfected and extremely large memories become extremely cheap, therefore, speech I/O is probably impractical. However, the possibility that a walkie-talkie (handheld radio transponder) may eventually become the preferred I/O device for an AD system should not be lightly dismissed.

Although a number of interesting hardware problems are apparent, hardware is not a limiting factor in the creation of AD or WP/D systems. Cost will limit the market temporarily, but by the time satisfactory solutions have been found to many software problems, appropriately large memories should be available at very reasonable costs.

Software

Programs of instructions to control computer hardware are called software. We shall use the term broadly, however, to include also the various assumptions that underlie the choice of processes to be implemented.

Lexicographic Aspects

Dictionaries are published in a range of sizes, where size (and price) depends on the number and length of lexical entries. Customers might prefer all the information in an unabridged dictionary, but they usually settle for a cheaper, more convenient college or pocket edition. Instead of *Funk and Wagnalls New Standard* or *Merriam-Webster's Third New International* (about 450,000 entries in 3000 pages, where an entry may include many different senses and subsenses), most people buy the *American Heritage Dictionary* or the *Merriam-Webster's New Collegiate Dictionary* (about 150,000 entries), or an even smaller pocket dictionary (about 50,000). Still smaller dictionaries are published for beginning students. In an alphabetical listing, of course, the more entries there are the longer it takes to find the entry you want; then you must search through the alternative senses for the one most appropriate in the given context.

Physical size and the amount of information available are also correlated for handheld AD systems, although the correlation is reduced somewhat by the CPU and display, which need not change in size as the number of memory circuits changes. Since the density of ROM circuits is expected to double every 18 mos. over the next decade, the correlation between physical size and information available should be attenuated even further in the future. The correlation between retrieval time and information available will also hold in automated systems, but retrieval should be fast enough that even the longest delays would not be bothersome. In AD or WP/D systems where several terminals timeshare a CPU and lexical memory is centralized, the physical size of a user's device is already independent of the amount of information available, and retrieval delays depend on the timesharing system, not on the number of lexical entries.

With the traditional size information delay correlations suspended, we need to reconsider how many lexical entries are desired or desirable. "The more the better" is a natural answer. The words people want to look up are usually the rare, unfamiliar ones that can be found only in large dictionaries. It is a lexicographic fact, however, that entries in large dictionaries are more sophisticated than are entries in small dictionaries, both because the entries in small dictionaries are shorter and because simpler words are used in the definitions: it is a criterion of lexicographic excellence not to use a sense of any word in a definition unless that sense is itself defined in that dictionary. If we were to opt for an unabridged dictionary, therefore, it would not be intelligible to many children. But if we were to opt for a small dictionary with simple definitions, it would be of little use to many adults. The educational level of the user is an important parameter in automated systems, just as it always has been in printed dictionaries.

Given unlimited memory, the simple solution would be to include both—adding 50,000 redundant but short entries to a data base that already has 450,000 sophisticated entries is a relatively modest increment. In a system with two dictionaries, a command to DEFINE a word might first produce a simple version, if it existed; a command to EXPAND the definition could lead to the sophisticated version. Experimental assessment should provide data as to which users requested expanded definitions for which words.

In an educational WP/D system, it would be instructive to provide RAM circuits adequate to hold a few thousand entries that students could use to write their own dictionaries. (It should be expected that students would discover a variety of other uses for such a retrieval system.) this facility might also be useful in commercial WP/D systems to accomodate the specialized vocabularies of different occupational and professional groups.

Lexical entries characteristically give spelling, part of speech, pronunciation, some etymology, a listing of alternative senses with usage notes for some. Entries may also include illustrations, or contrast a group of near synonyms. Extensive abbreviation is often used, and much thought has been given to typographic conventions. The question arises whether all this information should be retained, or whether entries should be recast in a form more convenient for computer processing, or more suitable to particular educational purposes.

If the goal is to evaluate prototype educational WP/D systems, arguments to begin with the traditional form of dictionary entry are reasonably persuasive. Whereas typing a dictionary is an expensive task, machine-readable dictionaries already exist (Sherman, 1979). An elaborate character generator for a CRT display is not particularly expensive, and until we have had some experience with these systems we have have little idea what educational purposes we might hope to achieve by modifying the way entries are written. A novel type of dictionary entry may emerge from the development of these systems, but that should not head our list of problems initially. On the other hand, the use of existing dictionaries raises problems of copyright; can permission be obtained to use copyrighted materials for experimental purposes?

Existing dictionaries list many homographs. 'Palm,' for example, takes three entries; one for the tree, one for the part of the hand, and one for the verb. Moreover, each entry has several senses, and some senses have subsenses. If a user merely indicated an interest in 'palm,' the device would have to display all three entries. The more entries that must be displayed, the larger the display screen should be, and the longer it takes to locate the particular sense or subsense appropriate in the given context. Little is know about the ability of small children to make such selections accurately.

It would be economical, both of display size and of time, if the AD could perform some of those selections for the user, In order to do so, however, the device must have additional input information. For example, if a student were reading a story in which 'palm' occurred on line 17, he might type PALM17 and see displayed the particular sense that the author or a teacher had selected in advance as appropriate in the context of line 17. Alternatively, if the device is used to display a story that the student wants to read, the machine could scan the context of 'palm' for other worlds that could be used to determine which sense seemed most appropriate. The feasibility and value of such refinements is a proper question for empirical investigation.

For writing with a WP/D system, some kind of thesaurus organization would probably be valuable (Sedelow, 1976). *Roget's International Thesaurus*, for example, provides a conceptual hierarchy six levels deep. At the sixth level are groups of words sharing a given sense or concept. The internal structure of the thesaurus can be represented abstractly as a matrix, where rows represent words and columns represent lexical concepts (Bryan, 1973); cells of the matrix are 1 if the given word can express that concept, and 0 otherwise. Thus, a row gives the concepts a word can express and a column gives the different words that express the same concept (that occur together at the sixth level of the thesaurus). From such matrices it is possible to construct paths between semantically related groups of words. Given a thesaurus in this form, an author in search of a better word might ask, not for definitions, but for near synonyms of a word already thought of; those could be displayed in a menu format; touching one of them would produce a similar display of *its* near synonyms, and so on until an apposite word was found. It should be possible to key a thesaurus to a dictionary, so that if an unfamiliar word appeared in the course of a thesaurus search, its appropriate sense from the dictionary could be made available.

Synonymy is not the only interesting relation between words. The value of incorporating other relations e.g., antonyms, superordinate and subordinate words, free associations, rhymes, word frequencies should be explored empirically. Precisely how this kind of information should be collected, how it should be represented in a lexical device, and what commands should be available for using it are all matters for further study. The first systems to be tested will adopt the simplest and most convenient form, but the results should establish a baseline against which any advantages of future additions or refinements can be measured.

Although dictionaries and thesauruses were the primary focus of the conference, the data base conception need not be limited to words. On-line data bases of dates, geographical indexes, and companions to literature, poetry, and the theatre, provide a means of exploring deeper questions in various subject matter areas. The definition of "revolution," for instances, comes alive and remains fixed in memory when the student can easily go on to explore the French, American and Russian revolutions. These expanded mental contexts support further learnings. Studies of the conditions and strategies that impel student explorations of complex knowledge will provide important data on intellectual development and on the effective structural characteristics of different subject matter areas.

It should be recognized, however, that the inclusion of all these kinds of information in a unified database will pose a substantial development problem, since millions of bytes of data will be involved, and since it must be retrievable both alphabetically and taxonomically (i.e., via related concepts). The data base management system will have to be carefully designed in advance, for modifications and improvements will be very difficult to make in the middle of the task.

Linguistic Aspects

At the conference it was noted by several participants that adding a syntactic parser to a word processing system might also have educational applications. In particular, Lawrence Frase spoke of "text processing systems" able to operate intelligently on phrases, sentences, or even longer chunks of text. Thus, the concept of a text processor with dictionary (TP/D) was invented, but its instantiation and evaluation was not set as a present goal for NIE. Rather, a TP/D system might be thought of as the distant goal toward which work can be expected to move in the future.

Psychological Aspects

In his proposal for an educational AD system, Miller (Appendix) assumed that, in addition to displaying definitions, a variety of games and puzzles should be provided to make the use of the system enjoyable to children.

- Spelling tests, where words are written at dictation (either dictation by a teacher, or by the system itself) and scored by the system. Texas Instruments is already marketing a version of this called Speak & Spell.
- A string of letters is displayed and the student judges whether they form a word. The word recognition task has been widely used by psychologists to explore how lexical knowledge is organized in a person's memory (Miller, 1978a).
- Anagrams, where a string of letters is displayed (perhaps in the form of a familiar word) and the student makes as many words as possible using those letters. Children's methods of solving anagrams have been found to provide a useful index of their linguistic and cognitive development (Schniederman, Reber, and Hainline, 1978).
- A definition is displayed and the student tries to guess what the word is. This procedure can induce the tantalizing tip-of-the-tongue experience (Brown and McNeill, 1966).
- A picture is displayed and the student names it. The latency and errors are known to be inversely related to the familiarity of the object (Oldfield and Wingfield, 1965).
- Two words are displayed and the student tries to find the shortest path between them through the thesaurus (Quillian, 1968).

These are merely examples. A wide variety of word games have been compiled by William and Mary Morris, and many of them could be adapted to AD or WP/D systems. In many cases, accuracy and latency information can be (unobtrusively) collected and compared with experimental results published in the psycholinguistic literature.

Such games and puzzles have as their goal to make a child interested in words. A focus on individual words, however, can interfere with the thought processes involved in reading or writing. Conference participants speculated that the reason they do not use a dictionary every time they are uncertain of a word's meaning is not that it takes time to find the dictionary and turn to the right page; the real reason is that using a dictionary presupposes a preoccupation with words, rather than with the message that the words were intended to convey. It is only when a text becomes indecipherable that they turn to a dictionary for help. If this introspective appraisal is correct, automating dictionary lookup will not automatically increase dictionary use during reading or writing. The possibility remains, however, that lexical information would be more useful while studying than while reading for pleasure, and more useful while editing than while composing a first draft. In any case, it would be instructive to collect all the words that any particular child had called for. Such lists could be incorporated into games or puzzles for that child, or they could signal a need for teacher intervention.

Several developmental psycholinguists have undertaken research on the growth of vocabulary in children. Most of this work has involved preschool children, since the learning process goes so rapidly that it is difficult to keep track of what older children know. Early research concentrated on vocabulary size and rate of growth; during the past decade more interest has been expressed in the underlying structure of lexical knowledge, i.e., the relations between words that reflect a child's conceptual knowledge. Carey (1978) describes the learning process as having two phases: (1) a rapid phase in which a new word is identified and assigned to some general semantic domain, followed by (2) a slower working out of the specific distinctions among words assigned to the same domain. Almost nothing definite is known concerning conditions that would facilitate learning of either type, although it is plausible to suppose that they might be different. At the present time, therefore, the main contribution to the present project that might be expected from this branch of psychology is a number of research workers interested in pursuing the topic.

Educational Aspects

Obviously, any system that is introduced into the schools should be matched to the interests and abilities of the children who will use it. Developmental psychologists have devoted some attention to age categories for toys, and manufacturers of electronic games that are played on the family television set have also given some thought to this question. But most of the experience we have in making computers accessible to children has been based on mathematics (see, e.g., Papert, 1979). Papert has found that powerful computers combined with programming languages that are well conceived and easy to use can smooth a child's entrance into a computer culture. In some of his exploratory studies, a WP was integrated with a mathematical component in several ways; the two can be mutually supportive in the acquisition of skills needed for interaction with a computer. The results indicate that children can use a WP, but further studies are clearly needed.

In conversations with Arthur Melmed, Dr. Allen Newell of Carnegie-Mellon University offered the opinion that the most critical factor affecting the success of an AD system for children would be the ease and naturalness of the child-machine interaction. The younger the child, the truer this observation is likely to be, but even with 10-year olds, who can compose interesting essays and type them on a keyboard, the ease and naturalness of interaction will probably be of greater importance than many of the lexical or linguistic aspects of the system's operation.

A well designed system must not only be attractive to children, it must also appeal to teachers and fit the educational system. Machines in the schoolroom have a poor record, and introducing a new one even one designed to be attractive to children may not be easy. Following the conference, Parker Damon provided the following list of reasons why schools have been slow to accept any technology more complicated than the book, the chalkboard, and the ditto machine:

1. Many new devices are promoted before the educational need they are meeting has been established in the mind of the teacher.
2. Many of them are too complicated and too expensive to use on a wide scale. (Remember that as a rule industry spends 9 times more on the equipment inside a plant than on the building housing the equipment. Schools do the reverse. Salaries and fixed operating costs consume 80% of a school's budget. All other supplies, consumable and reusable, contend for the remainder.) As a result of this topsy-turvy allocation of resources, schools are unfamiliar with the task of helping teachers acquire the interest, knowledge, and ability needed to use new equipment and procedures successfully.
3. Some new devices appear at first glance to be similar to something that was tried before without success, a frill to make the teacher's job easier, or a move away from the essentials of sound education.
4. A few devices are not used because of technical difficulties. The typewriter is an example. Typewriters are noisy, heavy, bulky, breakable, difficult to maintain."

Damon sees no reason why an educational WP/D system must fall prey to these problems, but both the system and the manner of introducing it will have to be carefully designed to avoid a deluge of difficulties.

Along with being mechanically and technically easy to use by teachers and students, and being recognized as necessary or helpful to the learning process, the WP/D must also be recognized and responded to for what it is, before it is introduced. The WP/D will upset the status quo of existing school practice. The roles and relationships of teachers and students will be changed. School districts which initially can afford to adopt the WP/D will be set apart from others that do not, or cannot, do so and thereby subject themselves to social and political pressures that should be anticipated.

There are at least two schools of thought regarding how best to introduce computer technology into schools. One point of view advocates thorough inservice preparation for the faculty before the equipment arrives. The topics for such an inservice program might include: philosophical and pedagogical implications; comprehensive training in the use of the equipment; extensive exploration of ways to apply the equipment's use to the curriculum; examination of the everyday problems that may be encountered. The other school of thought advocates that the inservice program accompany the introduction of the equipment into the school. Both views agree that the school's commitment to inservice should be both extensive and enduring.

Problems of design may be easier to solve than problems of implementation. According to Damon, if the first generation of WP/D has the following characteristics, it and the ones to follow may find quicker acceptance by the schools.

1. Inexpensive: $100 or less per unit that one student uses.
2. Easy to operate by a teacher; preferably easy to operate by a third grader.
3. Reliable: durable, relatively maintenance free, repairable.
4. Portable: easy to move about and small enough to fit on a student's desk.
5. Quiet: does not present the classroom with either a visual or aural distraction.
6. Magical and useful: does something that cannot now be done easily and should be done frequently.

It is not likely that the first models that will be rigged up to explore the usefulness of these systems will be optimal — that would be like asking the Wright brothers to begin by inventing the Concorde. Nevertheless, important advances have recently been made in easing — indeed, erasing — the barrier between users and computers (see, e.g., Bolt, 1979), and the possibility of exploiting some of these new techniques should not be overlooked.

Whereas the goal of a commercial WP/D would be to produce better writing, the goal of an educational WP/D would be to produce better writers. The difference can be illustrated by the question of how the system should respond to mistakes. For example, if a user misspells a word, should the system ignore it, signal that a mistake has occurred, refuse to continue until it is corrected, or correct it unobtrusively? If spelling were corrected automatically, proper names, foreign expressions, and unusual words would show up as misspellings. In an educational system, teachers would presumably want the system to give whatever response could be shown to teach children the most about spelling while not inhibiting writing.

Since the number of lexical entries that can be stored in a handheld AD system will be limited, at least initially, educators may wish to develop special word-books for particular academic subjects (e.g., for teaching foreign languages), and to abandon the goal of casting a general dictionary in that form. A single device, perhaps with interchangeable cassettes, could serve a variety of school subject, and might be marketed as automated glossaries accompanying particular textbooks. Even small vocabularies in a handheld AD might be useful to bilingual students. Participants at the conference disclaimed any superior foresight concerning the best use of such devices; their advice was to make something available and hope that it would inspire users to explore its potential applications and to suggest needed improvements.

Some educators may resist this kind of teaching device on philosophical grounds. Language is so uniquely human, so important socially and personally, that to learn it from a machine may seem objectionable, if not abnormal. On the other hand, children are going to be exposed to machines of this type after they leave school; the transition should be easier if they are already familiar with them. We suggest that any children involved in the initial tests should be at least ten years old and free from any indications of autism—their social and conversational skills should already be well established in a human context—and that any teachers involved should be screened for a positive attitude—teachers who objected to such systems could unconsciously communicate that attitude to their students.

Because we have so little experience with language-based computers in the classroom, the development of AD devices would benefit from early and continuing information on user-system interactions. Such information, at the outset, might be largely descriptive. Classrooms would be selected in which teachers were favorable to the project and students were culturally and scholastically diversified. Observations, recorded by eye and hand, as well as videotape, could document spontaneous activities with the device generated by individual children or groups of children; the conditions under which they generate these activities; the nature of children's participation in teacher-initiated and self-initiated encounters, and their reactions to these learning episodes. Similarly, teacher uses of the device could be recorded with special attention paid to the development of new (unanticipated) uses and to the integration of computer games and lessons into the curriculum. Initially, the research aim would be to produce natural histories of AD try-outs that could serve two purposes: help designers make more informed choices from among alternative hardware and software features; begin to identify characteristics of classroom situations, students and teachers which seem to affect utilization.

As improved devices become available and knowledge accumulates about how these work in the classroom, NIE might plan a larger-scale systematic study of experiences, with one or more systems. These descriptive studies would, in turn, lay the basis for experimental interventions and evaluative research designed to test the impact of the devices in various classrooms and their learning potential for individual children. The model of educational research we have in mind is that of the "agricultural station" which participates in the development of new programs and techniques as well as in their experimental test.

SCIENTIFIC BASE

Although a scientific base adequate for an initial implementation of AD or WP/D systems already exists, further research will be required in order to realize the full potential of such systems. The following discussion covers some of the more obvious lines of research that seem relevant, but it is far from exhaustive.

Lexicographic Research

The systems considered here can be viewed as special applications of the results of a new field of linguistic research that some have called computational lexicography. A typical project in this field might call for the automatic disambiguation of English text, i.e., for the identification of the intended sense of each word in connected text. For example, given the sentence, 'The pig is in the pen,' the system should automatically assign to 'pen' the meaning enclosure, not writing implement. Amsler and White (1979) would track this relation through lexical entries, e.g., one sense of 'pen' is 'enclosure for animals;' one of the senses of 'pig' is 'young swine,' and one sense of 'swine' is 'mammal,' and 'mammal' is 'animal;' thus an intersection is found and the appropriate sense of 'pen' in this sentence is identified. In order to conduct this search, Amsler and White's system deals with individual senses, and for a given dictionary (the nouns in the *Merriam-Webster Pocket Dictionary* in work completed to date) constructs a tangled forest of the hierarchical trees indicating which senses occur in the definitions of which other senses. Their result resembles a thesaurus. Sedelow (1976) has proposed a similar program, except that she has taken *Roget's Thesaurus* as the data base.

In a letter to Arthur Melmed, Robert F. Simmons (with W. P. Lehmann, a principal investigator in the Dictionary Project reported by Amsler and White) commented that "The taxonomic data base will immediately support a system that can teach the hierarchical structure of English with each definition. "I suspect that many variations of this teaching game can be discovered with some thought, and it will surely generalize to the teaching of foreign languages." Collaborative research between these lexicographers and cognitive psychologists could quickly establish educational uses for this hierarchical data base.

Linguistic Research

In recent years generative transformational linguists have become increasingly interested in the role of the lexicon. For example, an interesting convergence of syntactic and lexical theory appears in the recent work of Joan Bresnan at MIT. Bresnan (1978) has proposed that a "functional" structure be included in the lexical entry for each verb, and that this functional structure provide a direct mapping from the logical argument structure of a verb into its various syntactic contexts. The functional structure for the verb 'hit,' for example, would be something like NP_1 HIT NP_2, which maps the logical structure *x* HIT *y*, into such syntactic contexts as 'The arrow hit the target' or 'The target was hit.' A parsing system capable of identifying noun phrases and prepositional phrases, and recognizing which phrases serve as logical arguments of which verbs, should go a long way toward extracting the underlying logical argument structure of a sentence.

For longer stretches of text, some linguists (e.g., Longacre, 1968; Grimes, 1972) have proposed techniques of discourse analysis. As Sedelow points out, almost all of the categories that have been proposed for discourse analysis depend on an understanding of word meanings. It might be possible to automate discourse analysis, but only in conjuncion with some version of an automated dictionary.

Psychological Research

It might be of some practical value for psychologists to establish age norms for performance on the various word games and puzzles that will be used. Work on estimation of vocabulary size, more or less abandoned for twenty years, might be revived, with educational goals explicitly defined.

Such research is of less intellectual interest, however, than the attempts to characterize the general structure of peoples' long-term memory for words, and especially the studies of how such structures grow in children. For example, is a named category of objects learned by learning a list of criterial and characteristic properties, or is it learned by identifying a prototypical instance and then judging other instances on the basis of their similarity to the prototype? To what extent does learning the meaning of a word depend on acquiring a repertoire of appropriate motor responses to it? Are the alternative senses of a word conceptually distinct, like a list, or is there some abstract core meaning that is particularized differently in different contexts of use? How do we recognize metaphorical uses of a word? And so on and on. These and many other similar questions are currently being studies by cognitive psychologists. Although the answers could help to shape a theory of what children should be learning from their interactions with AD or WP/D systems, the design of such systems in the near future will have to proceed without the benefit of such a theory. It is to be hoped that these new systems, and the extensive data they will enable us to collect, will contribute to the development of better cognitive theories of semantic knowledge and ultimately to better education.

A different line of psychological research might view these systems as merely one more instance of the growing list of machines that our society expects children to interact with. How are children interacting with the electronic and mechanical devices we already have? What are the psychological consequences of inequitable access to devices considered to be desirable? Do children who have access to logically complex devices solve problems differently from children who do not have such access? Members of the conference had their own convictions about these matters, but they were supported by anecdotes, not by scientific evidence.

Educational Research

The proposed evaluation is, of course, an example of educational research. Of special importance to educators will be information on whether teachers of reading and writing are willing to rely on such systems; on the attitides of parents and taxpayers toward such developments; and on the (possibly inequitable) effects of home systems for the education of children from wealthier families.

Little concern has been expressed in this report for the social, moral, or ethical implications of this educational methodology, but those are questions of central concern to many educators. It may be desirable or necessary to impanel an advisory group chartered to review those aspects of the work before any substantial federal funding is proposed.

Comment on Research

Improved hardware for these lexical systems is not the first priority. Existing technology is already adequate for an initial evaluation of AD and WP/D systems; NIE can play an important role by simply coordinating resources already available and activities already under way. The principal need is the development and testing of new software, broadly defined. If large federal expenditures are needed, it is to be hoped that, in addition to NIE, the Advanced Research Projects Agency of the Department of Defense or the National Science Foundation might also become interested.

It is possible that a project of this kind could have worthwhile consequences for research, even if new funding is not forthcoming. Recognizing what information is needed in order to construct and improve such practical systems can sometimes focus and clarify research goals.

RECOMMENDATIONS

The following recommendations to the National Institute of Education grow out of the discussion summarized in the preceding pages:

1. The NIE should coordinate a continuing program of exploration of the educational use of AD and WP/D systems in appropriate social and educational contexts.

2. In order to promote the development of educational systems worth exploring, the NIE should seek private or federal funding for:

   a. The design and construction of handheld AD with as large a vocabulary as possible at the present time; special attention should be paid to the attractiveness of the device to young children.

   b. The design and construction of an experimental WP/D system for test with older children, with several terminals and a large vocabulary stored in the central processor's RAM circuits; facilities should be provided for experimenting with the form and organization of lexical information.

3. In order to coordinate the above activities, to formulate various subgoals more specifically, and to draw other interested persons into the planning of this work, the NIE should sponsor additional meetings and conferences as they deem necessary and appropriate.

REFERENCES

Amsler, R. A., and White J. S. Development of a Computational Methodology for Deriving Natural Language Semantic Structures via Analysis of Machine-Readable Dictionaries. Final report, NSF Project MSC77-01315, U. of Texas at Austin, 1979.

Bolt, R. A. Spatial Data-Management. Cambridge, Mass.: Architecture Machine Group, MIT, 1979.

Bresnan, J. A realistic transformational grammar. In M. Halle, J. Bresnan, and G. A. Miller (Eds.), *Linguistic Theory and Psychological Reality*. Cambridge, Mass.: MIT Press, 1978.

Brown, R., and McNeil, D. The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 325-337.

Bryan, R. Abstract thesauri and graph theory applications to thesaurus research. In S. Y. Sedelow (Ed.), *Automated Language Analysis*. Lawrence, Kansas: University of Kansas, 1973.

Carey, S. The child as word learner. In M. Halle, J. Bresnan, and G. A. Miller (Eds.), *Linguistic Theory and psychological Reality*. Cambridge, Mass.: MIT Press, 1978.

Fox, M. S., Bebel, D., and Parker, A. The Automated Dictionary: A Preliminary Analysis. Carnegie-Mellon University, 5 June 1979, revised 19 July 1979.

Grimes, J. E. The Thread of Discourse. Technical Report No. 1, NSF Grant GS-3180, Cornell University, 1972.

Longacre, R.E. Philippine Languages: Discourse, Paragraph, and Sentence Structure, Summer Institute of Linguistics, 1968.

Miller, G. A. Automated dictionaries, (Appendix)

Miller, G. A. Lexical meaning. In J. F. Kavanagh and W. Strange (Eds.), *Speech and Language in the Laboratory, School, and Clinic*. Cambridge, Mass.: MIT Press, 1978a.

Morris, W. and Morris, M. *The Word Game Book* New York: Penguin.

Oldfield, R. C., and Wingfield, A. Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, 1965, *17*, 273-281.

Papert, S. *Logo Book*. Unpublished manuscript, 1979.

Quillian, M. R. Semantic memory. In M. Minsky (Ed.), *Semantic Information Processing*. Cambridge, Mass.: MIT Press, 1968.

Schneiderman, M. H., Reber, A. S., and Hainline, L. Children's anagram solutions as an index of linguistic and cognitive development. *Child Development*, 1978, *48*, 765-772.

Sedelow, S. Y. Analysis of "natural" language discourse. *AFIPS Conference Proceedings*, 1976, *45*, 1027-1034.

Sherman, D. A computer archive of machine readable dictionaries. Read at Meeting of Dictionary Society of North America, London, Ontario, 28 June 1979.

APPENDIX

*Automated Dictionaries**

by George A. Miller

A dictionary is a treasury of information about a language. It contains information about spelling, pronunciation, etymology, syntactic category, meanings, accepted usages, even pictures, for thousands and thousands of different words. It is valuable for both writers and readers. It is also valuable to teachers and students, because it is a codified representation of a large body of linguistic knowledge that literate adults are expected to have.

Nevertheless, most adults would probably agree that they do not use dictionaries as often as they should. One reason is that referring to a dictionary interrupts the flow of thought. It takes time, first to find the dictionary, then to find the word, then to select the desired information from the lexical entry. Few adults—and even fewer children—have the sustained attention span required to bridge such interruptions.

If there were sufficient value in the project, the automation of a dictionary would be well within the power of existing technology. It would be possible, for example, to type a word into a computer keyboard and have the machine pronounce it. Any information stored with that word could be displayed on a CRT almost instantaneously. The tedious task of searching through an alphabetical listing is one that machines can perform much faster than humans. A central computer could easily serve a large number of classrooms. Let us refer to such a system as Automated Dictionary Model 1, or AD-1 for short.

Let us imagine that AD-1 is available and consider whether any interesting research projects might result. The most obvious question to ask is whether anyone would use it, or whether it would be as neglected as its printed progenitor. User studies might suggest ways to make the system easier to refer to, or more accessible to more people. It is likely, however, that the system would be ignored unless special efforts were made to educate people in its value, or special tasks were devised to use the system in integral ways.

A second line of research, therefore, would be to develop interesting and educational tasks in which AD-1 played a useful role. Some examples:

*Reprinted from *Testing, Teaching and Learning, Report of a Conference on Research on Testing,* National Institute of Education, October, 1979.

- Spelling. A word is pronounced and the student types it into the keyboard. If the spelling is correct, AD-1 confirms it; if not, various responses might occur, ranging from nonrecognition to corrected spelling suggestions (based on tabulations of the most frequent spelling mistakes).
- Word recognition. Letter sequences could be presented and the student asked whether or not they spell an English word. AD-1 could serve as referee in student competitions.
- Anagrams. the student could be challenged to form as many words as possible from a given collection of letters, and given access to AD-1 to settle dubious instances. (Scrabble, word pyramids, etc.)
- Word search. The dictionary definition of some sense of a word could be given and the student challenged to think of the word. Guesses would be typed into AD-1, which would display the definitions of that word, until a matching definition was found. The definitions might be provided in the form of clues for the solution of crossword puzzles.
- Picture naming. A picture from the dictionary might be displayed and the student challenged (as in word search) to think of its name.

These and other dictionary games could be explored for their intrinsic interest to children and for their educational value. From a computerized record of a student's use of AD-1 it should be possible to assess ability in spelling, vocabulary size, or knowledge of multiple meanings and, if desired, reduce those assessments to comparisons with other students or with external criteria.

A third line of research would concentrate on the development of more sophisticated models. Two alternatives suggest themselves a priori. One line would explore programs for using the dictionary data base in a variety of different ways of presenting tasks and keeping records. A second line would explore additions to the data base of various kinds.

As an example of the first kind, it might prove desirable to be able to key the dictionary to particular texts that children were reading. If a new word appeared in the text, the teacher might wish to display only that sense of the word that was appropriate in the context, thus savings children the difficult task of selecting the appropriate sense. Special software would be required to enable the teacher to make such modifications in the standard display. Or it might be desirable to edit the dictionary in other ways—simplifying definition, inserting more pictures, giving more sentences illustrating usage, deleting etymology, or whatever—which would also require special software. Or the various tasks (exemplified above) might be presented by the system, rather than the teacher, which would also complicate the system.

Such revisions, leading to what we might call AD-2, should be undertaken in close collaboration between systems programmers and educational research workers familiar with the potential uses of the system.

It is not impossible to imagine elaborate and expensive software innovations that would have no real value to teachers or students.

Versions that might be called AD-3 should consider additions to the data base itself. Perhaps the most obvious example would be to add a thesaurus to the dictionary, thus providing cross references between lexical entries that are valuable in searching for apposite words. Alphabetic versions of Roget's thesaurus are probably adequate to support initial explorations in this direction at the present time.

A request for information from a thesaurus elicts a set of words. Computers can do interesting things with sets that would be quite laborious to do by hand. They can, for example, find the intersection of two sets. Thus, it should be possible to type two words into AD-3 and have displayed only those words that are listed under both entries—thus sharply narrowing the range of search.

Various teaching-testing tasks might be set on AD-3. For example, two words might be selected with nonintersectiong sets of associates, and the student challenged to find the shortest path between them. A trace of the student's exploration should provide an interesting picture of his or her sensitivity to alternative meanings of words. Once the path was found, the student might be asked to return to the dictionary and select the particular senses of each word that are involved in the path. (It might also give the student an intuitive feeling for the theory of finite sets.)

Dictionaries do not itemize hyponyms. That is to say, under both 'table' and 'chair' they will give the superordinate term, 'article of furniture,' but if you look up 'furniture,' you will not find 'table' or 'chair.' (The exception is crossword dictionaries, which specialize in hyponyms.) This information could also be added to an automated dictionary (at least, for nouns). In some cases—animal and plant taxonomies, for example—this lexical information is part of the subject matter of some courses.

However, the most valuable additions to an automated dictionary—grammar and general knowledge—would also be the most difficult to make in the present state of the art. There is something artifical about any use of words in isolation. In order to deal with words as they occur in normal communication, the system must be able to recognize sentences and to assign interpretations to them that are reasonable, given the context and our presupposed general knowledge about the world and the events that occur in it. AD-4, incorporating grammar, and AD-5, incorporating grammar and general knowledge, are still beyond the reach of current science and technology. Enough progress has been made in these directions, however, that we can at least speculate about such systems and their educational implications.

It would be possible to add that component of a grammar known as the "phrase structure rules" directly to the entires in an automated dictionary, simply by adding to each entry the grammatical contexts in which that word can occur. Such a representation would be enormously redundant (the rules for forming noun phrases, for example, would be repeated in the entry for every noun), and there are reasons to believe it could not be complete (the transformational component of the grammar cannot be caputred lexically). The only scheme that is considered plausible at the present time is to add grammatical rules as a component separate from the lexical component, in which case the system should probably be referred to as an automated grammar, rather than an automated dictionary.

Relevant evidence might be obtained, however, by a careful analysis of children's errors. If those errors can be characterized in terms of syntactic classes of words, rather than in terms of individual words, it would provide strong evidence for an independent grammatical component and against the information of even phrase structure rules in the lexical component. Evidence available at the present time suggests that children learn grammatical rules abstractly, not redundantly for each particular word, because they often overgeneralize a new rule.

An automated grammar (AG-1) is feasible at the present time, given the various parsing systems that have been developed by computational linguists. These systems have frequently been used in question answering systems: the question is parsed and compiled as a program for retrieving information from a prestored data base, then answered with a simplified declarative syntax. None of these systems is capable presently of dealing with the full range of grammatical English sentences, and most can process only a small vocabulary of terms directly related to the data base and the questions that can be asked of it.

AG-1 could support two lines of research, one leading to more sophisticated systems, the other exploring its educational applications. Along the first line, it would be useful to see whether a dictionary could provide an interesting data base for a question answering system. Along the second line, one would attempt to devise educational tasks using AG-1:

- Parsing. The student tries to parse a simple sentence (one known to be correctly analyzed by AG-1) by typing its labelled bracketing. If the student's parsing is correct, it is accepted; if not, various responses might occur—rejection, acceptance of only the correct parts, suggestions as to what is wrong, etc.

- Syntactic ambiguity. Some sentences can be parsed in more than one way. Students might compete to find all correct parses, with AG-1 serving as referee.

- Sentence anagrams. The student could be challenged to form as many sentences as possible from a given set of words, and AG-1 used to setting whether any disputed ordering was grammatical.

- Privileges of occurrence. The student could be given a sentence with one word missing and asked what words could fill the blank: AG-1 would accept or reject, relative to that sentence, any word that the student typed. The task should teach a distinction between sentences that are syntactically ill-formed and sentences that are grammatical, but semantically anomalous.

• Editing. AG-1 should be able to do most of the relatively mechanical corrections that teachers make in student compositions. The advantage, in addition to freeing teacher time for instruction that AG-1 could not provide, would be immediate feedback to the student while his or her communicative intentions are still operative.

• Question answering. Questions about words could be phrased in standard English, so that the various tasks developed for the automated dictionary would not have to be posed in some special code.

Future developments of AG-1 might lead to a linguistic equivalent of the arithmetic BUGGY system, but we are presently far from understanding children's mistakes well enough to simulate their alternative grammars in a computer.

We are even further from AD-5, a system that could interpret sentences in the light of general knowledge about the world. As a metaphor, it is easy to speak of adding an encyclopedia to the dictionary in the system's data base, in which case the system should probably be called an Automated Encyclopedia (AE-1). Not until the representation of large bodies of information in a computer memory is better understood, however, will it be feasible to implement such a system. Preliminary research might begin, however, with particular bodies of information relevant to specific academic subjects. Question answering systems with a data base on, say, the American Revolution, might be experimentally feasible at the present time.

A full scale Automated Encyclopedia could be a powerful teacher, and might be taken as a distant goal to work toward. Meanwhile, systems at the level of AD-1 to AD-3 could not only support interesting educational research, but could provide valuable experience to guide the development of more sophisticated systems.

☆ U.S.G.P.O. 625-308/1302-1762