A PROPOSAL FOR COMPUTER-ASSISTED INSTRUCTION IN VOCABULARY

ROBBINS BURLING

Department of Anthropology, University of Michigan, Ann Arbor, U.S.A.

THE PROBLEM OF VOCABULARY

Vocabulary is the step-child of modern foreign language pedagogy. We have lavished great effort in developing sophisiticated instruction in grammatical and phonological subtleties, but we have done far less with words. Several factors have conspired to encourage this bias.

First, the biases of language pedagogy reflect the biases of much of modern linguistics. Phonology and syntax form patterns that linguists find appealing, and it has been tempting to apply our expanding theoretical knowledge of these topics to practical language teaching. The lexicon, by comparison, seems to form a relatively disorganized heap that is resistant to orderly patterning. Linguistic analysis of vocabulary has yielded relatively little that might inspire derivative pedagogical applications.

Second, the greatest effort in recent language pedagogy has been directed toward production—teaching students to talk. Speaking demands a firmer control over phonology and syntax than does listening but, reciprocally, a listener has a more urgent need for vocabulary than for phonology or syntax. An imperfect speaker can resort to synonyms and circumlocutions but an imperfect listener has no such option, since the person to whom she is listening may use any word in the language. A speaker may be able to manage quite well with 'little' and 'very little'. A listener or reader had better know 'small' and 'tiny' as well. As language pedagogy is redirected toward a greater emphasis on comprehension, vocabulary is certain to receive more attention than it has in the past.

The third reason for the relative neglect of vocabulary is its sheer bulk, for this discourages a direct assault. It is all too easy to be dismayed by the subtle semantic complexities of even a single word. A speaker ought to know whether a word is formal or colloquial, and she even ought to know its precise degree of formality. What are its connotations? Is it used more by women than by men? Is it childish? Technical? The complexities are so vast that it may seem safer to limit students to a relatively small vocabulary, and hope that these few words can be mastered, before the entire mass is piled upon them.

Finally, a fourth reason for the neglect of vocabulary is that all fields of modern pedagogy, language teaching among them, have reacted rather strongly against 'mere' memory. We do not ask our students to memorize dates, or state capitals, or presidents, or poems, or the twelves table, as often as our grandparents asked their students to do

so. We have sought, instead, to give our students understanding. We want the facts we teach to fit into larger patterns and we want them to have relevance for larger issues. If facts are to be learned at all, we suppose that they should be learned as part of some larger and more meaningful pattern, or theory, or context. Rote memory has been denigrated. This probably helps us to feel that the patterns of phonology and syntax, in which linguists have taken such delight, are more important, or more deserving of study, than the many thousands of words in a lexicon, for these seem to call for simple old fashioned rote memory. We seem almost to have forgotten how to memorize, and we hesitate to demand memorization of our students. We have hoped, instead, that, if the central patterns of the language could be learned first, the lexicon would then follow along relatively painlessly, and the student would not be forced to resort to too much 'mere' memory.

For all these reasons, we have done little to encourage our students to work deliberately on vocabulary and we have done little to help them to memorize. The students have been left to discover for themselves just what a terrible burden the vocabulary is. This is likely to hit them with discouraging force when they first advance in their reading beyond exercises that have been carefully limited to a restricted vocabulary. The first reasonably 'natural' texts will have countless words that the students do not know and that they cannot successfully guess. Unless they are willing to read with only a partial and imperfect understanding, they will find themselves spending long and dreary hours with the dictionary. Then, later, when students finally escape their classroom and actually attempt to communicate in the language, they are likely to find that the natives use a dauntingly diverse vocabulary, a range of words with which they cannot cope. Comprehension, and with it dialogue, is then in danger of collapsing on the shoals of limited vocabulary.

I believe that an excessive attention to production in the early stages of language learning and a reluctance to encourage rote memory have both contributed to the halting progress that foreign language students have typically made with vocabulary, but even when they begin to face the bulky reality of the lexicon, and even when they begin to resort to measures of self-help, they face formidable problems. When a learner hears a word spoken, it usually comes mixed in among other words of the discourse. Unless the surrounding words and phrases are clearly understood it may be very difficult to extract the unknown word from its context. It may be difficult even to ask a sensible question about it. If too many words are misunderstood, the learner hardly knows where to start. For the first few months of practical use with the language, it may even be difficult to know where one word ends and the next one begins. When reading, one can always resort to a dictionary, but the aural equivalent of a dictionary does not exist. Learners find it exceedingly difficult to fight their way into rapid conversation where words so easily blur into one another.

Reading presents its own set of problems. Here to be sure, words come nicely isolated between white spaces. These clear word boundaries, along with the ability to go slowly and to think over various possible meanings that a word might have, give students a chance to try to guess the meaning of words from the context, or even to decide that some words can be safely ignored. With sufficient patience, crucial words can even be tracked down in the dictionary. But then, after a word is found, what should the student

do next? If she makes no attempt to memorize the words that she has so laboriously looked up (or, for that matter, so laboriously guessed) she will surely face the debilitating experience of having to track down some of the same words a dozen times. For most of us, one trip to the dictionary is simply not enough to fix a word permanently in our minds. Constant reference to the dictionary, however, or even constant interruptions to guess again at the same old word, makes reading so agonizingly slow that it seriously limits the amount of text through which a student can work in a reasonable amount of time. Simplified texts with restricted vocabulary can reduce the struggle, but they are usually dull reading and they only postpone the evil day. In many ways, the sooner we can get students to real texts, and to material that is suitable to their age and interests, the better.

It is so discouraging to look up the same word over and over again that some students, in desperation (and sometimes in defiance of their teachers) resort to rote memory. They write out long lists of words that they think they ought to know, or they construct hundreds of flash cards, but, here, students face still another dilemma: they have no way of knowing which of the unknown words that they encounter are most likely to reappear. All too often, in a desperate attempt to overcome the terrible hurdles posed by unknown vocabulary, students try heroically to memorize words that are not destined to appear again in their reading for several hundred pages. If only students could focus on the most common words, and if only they could do so without the long debilitating hours with the dictionary, they would have a far better chance of expanding their vocabulary with some degree of efficiency.

Readers, then, are pulled between three unhappy choices. If they resist looking up words they may miss a good deal of the meaning, and their vocabulary is likely to expand only very slowly. If, on the other hand, they look up most unknown words but make little or no deliberate attempt to memorize their meaning, they may understand what they read, but they will read quite slowly and their vocabulary will still expand at a discouraging pace. If, finally, they try to memorize a large number of the words they look up, they will be able to read even less—since they must now not only spend dreary time with the dictionary but even drearier time at attempts at memorization. Soon, they will begin to feel it would be more useful, and surely more pleasant, to read more and memorize less. Somewhere between these three extremes (reading without looking up, looking up without memorization, and earnest memorization) there must be an optimal point where one looks up a limited number of words and tries to memorize the meaning of an even more limited number, but it is very hard for a student to find that optimum.

In a kind, but possibly misguided, attempt to spare our students the pain of rote memory, we have sometimes advised them to minimize the use of the dictionary and to let the meanings of the words crystalize slowly in their minds as they are encountered in many contexts. Of course it is true that a full understanding of a word can only be achieved by encountering it repeatedly in a wide variety of contexts, but precisely for this reason, learners are simply forced to begin with something partial and something simpler. In point of pragmatic fact, a one or two word translation will often give a student an excellent first approximation of the meaning of a word, and this may be all that she needs and all that she can cope with at first. We acknowledge the value of a simple translation when we give students glossaries and dictionaries, but we sometimes act as if these are unfor-

tunate crutches, resorted to by those who lack the imagination to abstract the meaning of a word from its original context.

Context is undeniably crucial for refining the meaning of words, but we seriously exaggerate, and we make life terribly difficult for our students, if we do not recognize that the first and most difficult step of all can be vastly eased by a quick translation. If the student can remember this translation the next time the word appears, she will avoid one distracting struggle with the dictionary. This will make reading a bit more rapid, and this means, in turn, that the word will appear again more quickly and more often. This, in its turn, will allow context more quickly and effectively to provide the subtle refinements of meaning that full understanding of the word requires.

SIMULTANEOUS READING AND LISTENING

I wish to propose a technique for helping students understand both aural and written language. It is a technique that recognizes the central importance of vocabulary in gaining an adequate ability to comprehend a new language, but the technique should minimize the time spent leafing through the dictionary at the same time that it maximizes the utility of the time devoted to memorization. My proposal has three parts. The first two are relatively minor and can be disposed of quickly, but only after they have been described can I go on to the third and more complex part of my proposal that involves computer assistance.

The first part of my proposal is so simple that any teacher could put it into effect in the classroom tomorrow. I suggest that every text that the students are asked to read ought to be accompanied by a taped version, so that they can hear the text as well as read it. They could even hear it repeatedly if that proves helpful. Having an aural version of written texts has several advantages. It helps to fix pronunciation in one's mind and it helps to avoid spelling pronunciations. It also forces students to read, and to try and understand, at something like normal speed. A spoken version, moreover, comes with the normal intonation and emphasis of speech and these provide important clues to the structure of the language that the written text, by itself, omits. A fluent reader spontaneously groups words in a way that help to make the message understandable to the listener. Finally, if new words are heard, as well as seen, they might become lodged in memory a bit more quickly.

The second part of my proposal is really the reciprocal of the first part, although it would be a bit more difficult to put into effect. It ought to help students to gain a fluent ability to comprehend the spoken language, if they could have extensive practice with recordings of reasonably natural conversations, complete with colloquial vocabulary, and delivered with the full use of fast speech rules. Rapid natural conversation is notoriously difficult for new learners to understand, and yet, if they are to participate in conversations, this kind of language is crucial. A recorded version of a written text cannot serve as an adequate substitute for colloquial speech. The spoken and written forms of most languages, probably all languages, differ in so many ways, that not even an extensive exposure to read-aloud versions of written texts can prepare learners for the fast phonology,

deletions, false starts, interruptions, and colloquial vocabulary (to say nothing of the overlaps and background noise) of natural conversations.

Recordings of entirely spontaneous conversations may pose insuperable problems as pedagogical tools, however. They are too cluttered with overlaps, with references to the visible context that cannot be understood from a sound recording, and with the shuffling of papers and the clinking of glasses. It is not difficult, however, to stage minimally rehearsed conversations in which two or three people act out a situation as naturally as they can, but in conditions where a good quality recording can be made, and where references to the visible situation can be avoided. Even the highest quality recordings are very difficult to use alone, however. Fast speech rules reduce some words to mere blips that a learner cannot catch and cannot possibly find in her dictionary. Some very common and very colloquial words (or noises such as *umhm, hm*? and *ohoh*) are not found in our dictionaries at all. Yet it is precisely these blips and noises that a learner must understand easily if she is to participate in anything approaching natural conversation.

Learners could be helped immeasurably with this kind of language if their recordings were accompanied by a written transcript in which the words are spelled out in full. A transcript would show students how to interpret bits of phonology they hear, and it would help them to find the connections between the fast colloquial language and the fuller, more formal, language of writing. When all else fails, a transcript would also allow unknown words to be chased down in the dictionary.

The first two parts of my proposal are alike in suggesting the simultaneous use of spoken and written versions of the same texts, and this may make them seem to be almost identical suggestions. The spoken and written languages are so different, however, that practical student experience with the two kinds of lessons, as well as the practical task of preparing them, would be very different. The third and final part of my proposal is far more radical and ambitious than either of these first two, however, and it is this third part that addresses most closely the dilemmas of vocabulary that I discussed earlier.

THE COMPUTER

I envisage a student sitting at a computer terminal or, even better, at a personal computer, with both a keyboard and a screen. By manipulating the keyboard, the page of text that she wants to study can be called to the screen. (For reasons to be explained shortly, the text must be double spaced.) Ideally, the keyboard would also directly control a tape recorder so that the aural version of the text would automatically accompany the written version that the student sees on the screen, but we may have to settle, at first, for a separate tape recorder. In any case, I want the student to be able to hear everything she reads and to be able to see everything that she listens to.

I would want the student to start by simultaneously listening and reading to a passage straight through at least once, possibly several times, trying to understand as much as possible without looking for translations. She should try to guess at as many meanings as possible and get a feeling for both the sound and the look of the passage. Sometimes,

particularly with language samples that originate as spoken conversations, it would be best to begin by trying to understand the aural version by itself, and to turn to the transcript for help only with the more murky parts.

Next, the student should work through the passage more carefully. This time, when she comes to a word whose meaning she needs but that she cannot guess, she should be able to ask the computer for a short translation. The easiest way for the learner to do this would be to build technology that would allow her to touch the right spot on the screen directly with a light pen. At present, however, it would be simpler and less expensive to rely on a standard computer cursor. This is a mark, such as an underline, that can be moved to the desired location on the screen by means of keyboard controls. Whether by a light pen or a cursor, the student should be able to indicate the word for which a translation is needed, and then, by the magic of modern computer programming, a brief translation of the unknown foreign term, usually a single word, could appear on the screen in the empty line directly above the unknown word. At the same time, the unknown foreign term might change to bold face or italics so that it will stand out from its context. Following the translation, I would like a number to appear that would offer a rough indication of the word's frequency. Thus '(14)' might indicate that the word lies between the 1300 and 1400 most common words in the text.

Most often this short translation would be all that the student would need. It should, after all, give the particular meaning of the particular word in that particular context. She might, however, occasionally want a fuller dictionary treatment of the word. She could, of course, turn to an ordinary dictionary, but it might also be helpful if she could call for a fuller definition through her computer. In any case, the student would then need to decide whether or not this is a word that is important enough to try to memorize. If she knows most of the 2000 most common words, of her new language, she may feel that a word that is among the 1400 most common should certainly be given some attention. On the other hand, she may already have been quite sure of the word's meaning. Perhaps she only wanted to check it to make dead sure. In such a case she might be willing to leave it alone. Many other factors might encourage her either to make a special effort to memorize a word or to regard it as too rare (or too easy) to warrant a struggle at just that point. If she does want to work on memorizing it so that she can avoid looking it up again in the future, she should be able to press an appropriate control key that would enter the word into a computer file to which she can refer later.

'Looking up' an unknown word in this way, noting its meaning, deciding whether or not to memorize it, and then asking the computer to copy it into a list, could be done quite quickly, far more quickly than turning the pages of a dictionary and more quickly even than checking a note at the bottom of the page. The ideal would be to interrupt the flow of the sentence as little as possible. The reader would then read on until she reaches another unknown word, at which point she can repeat the process. Whenever a new translation is requested, the previous translation might disappear from the screen so that it does not become cluttered with translations, but it might be useful for the bold face type or italics of the unknown foreign word to remain. In this way, once a passage has been read through, the learner would be able to look back at the text and review the words that she had looked up, calling once again for their translations when she must. Often, a modest but quick review of this sort will considerably increase the rate of vocabulary retention.

In this way, the student would be able to read along relatively rapidly and relatively undistracted by fumbling for the meaning of unknown words. At the same time she will construct a list of the most necessary vocabulary that she wants to bear down upon. This list should certainly include far less than the total number of words whose meaning she checks, but some listing of the most important unknown words and some concerted effort at memorizing them ought to make further reading more rapid and more satisfying. The student will certainly want to cut down gradually on the number of words that she has to look up.

The learner will gradually accumulate a computer file of words that she wants to work on. When she is ready, she should be able to call for this file to appear on her screen in the form of a column of foreign words. She can inspect the list and try to remember the meanings of the words. When she is doubtful, she can use her light pen or cursor to call for the translation. If she wants to have a list available for study after she leaves her terminal, and if a printer is available, she can have the list, or part of it, printed, or of course she can copy it onto paper by hand. As long as she confines her attempts at memorization to the more common words, she can have a reasonable hope that they will soon appear again in her reading, and this will give her another chance to face them in context and to consolidate them in her memory. In addition to cutting down drastically on the time spent shuffling through a dictionary, she will also cut down the time spent in memorizing because she will avoid any struggle with the less common and less useful words. The fact that she will be concentrating on the words that are most likely to appear again, makes them the ripest for rapid consolidation. The method will allow the student to move through a text relatively quickly and this should add to her satisfaction.

FLEXIBILITY

I have described what can be regarded as the 'default' form of the computer program. This is the form in which the program might operate unless the learner gives it explicit instructions to the contrary. If the learner prefers, however, the flexibility of the computer would allow a number of alternative procedures. The learner might, for instance, like to have the translations of all the words that she requests should stay on the screen, instead of having each one fade at the time she calls for the next one. She might like to have the translation of every word appear on the screen simultaneously. She might like to have all the translations of words below some given frequency appear automatically and without having to request each one individually, so that she would know immediately that these are words with which she should not struggle just now. This would let her avoid almost all struggle with the rarer words. She might wish to cancel the bold-face component of the program. When looking at the list of words that she constructs, she might prefer to have all of the translations visible for simultaneous inspection instead of having to ask for each one individually. It would be possible to introduce a counter into the program so that the computer would remember how many times a learner asks for the translation of a word. She could then decide whether or not to try to memorize it by how many times she has had to ask for its meaning, rather than by how often it occurs in some word count.

The words that the learner selects for special attention might ordinarily be listed in the

order in which she has added them to her list. Most often this will be in the same order as they are found in the text. Recalling the context in which the word appeared often helps to fix it in memory, and listing the words in the order of their appearance may be the most helpful way to remind the learner of the context. However, words that are very similar are easily confused and so it may be helpful to bring similar words together. The learner can be given the means to reshuffle and reorder the words on her list in any way she chooses, but one obvious way to bring some similar words together is to ask the computer to provide an alphabetical list. Some students might also like to have the computer provide a randomized listing. Freedom to construct sub-lists, to add new words to old lists, or to remove old words once they seem to have been learned sufficiently well, should not be difficult to provide. Clearly there are many forms of flexibility that could be built into the program in order to provide alternatives that would be adaptable to varied learning styles.

PREPARATION OF THE PROGRAM

A program of the sort I propose would demand painstaking preparation, but, once prepared, it would allow self-instruction for any number of students. The following steps would be required to prepare materials.

(1) First, the texts must be selected and prepared in both aural and written versions. For texts that start in a written form, this is hardly a problem. In order to provide an aural version, it is only necessary to have a native speaker record the text on tape. I feel that a well-rounded program would need colloquial spoken texts as well, however, and since these are less as readily available than written texts, they might have to be specially recorded with the help of two or three native speakers of the target language. Transcripts of these recordings, written in full spelling, would then have to be prepared. This is a time consuming, but straightforward task. (An appealing alternative, that someone should try, would be to obtain the sound track, or even better the full video recording, of highly colloquial target language movies or TV programs. If the written scripts of these shows could be obtained as well, a good deal of the lesson designer's work would already have been done.)

(2) Next comes the mechanical task of typing the written form of the text into the computer. This, again, is a straightforward, if tedious, job.

(3) A rather elaborate computer program would have to be designed and tested. It ought to offer a maximally efficient yet flexible style of lesson. Once the program is designed, of course, it could be used freely for any number of texts.

(4) The most demanding and time consuming task of all would be to assign a short translation of each word in the text, or to each idiom of several words. This job could be simplified in two ways. Depending upon the level of students for whom the materials are designed, it would be legitimate *not* to give translations for the 100, or 500, or 1000 most common words in the text. Since even the 100 most common words constitute a substantial fraction of a running text, omitting them will cut down the labor considerably. Second, since the text will already have been entered into the computer by this time, it should be possible to assign a translation the first time a word appears and then ask the computer to copy this translation wherever the same word appears in the text thereafter. This, unfortunately, cannot be done blindly, since, depending on the context, a single word in one language often requires several alternative translations in another language.

Thus, each entry would have to be individually checked. Someone with a good command of both languages would have to work through the text checking each word to make sure that it has an adequate translation for its particular context. This is, undeniably, a lengthy task, but it is by no means as great a job as giving a complete translation.

(5) The computer can then be asked to provide an alphabetical print-out of all the words that occur in the text, and when a single word has received more than one alternative translation, the print-out can subclassify the words according to the varied translations. Thus homographs will be separated. At the same time, the print-out can show how often each word occurs in each of its translations. This raw list will include many collections of closely related inflectional forms (e.g. crash, crashed, crashing, crashes), along with separate sub-entries of the same word (or of closely related words) that have been assigned varying translations in different occurrences. This list would then have to be worked over entry by entry. Words that are minor alternatives, in particular, inflectional forms, would be grouped together, and when the alternative meanings that have been assigned to a single foreign term are considered sufficiently alike, they too, can be grouped together.

The information about which forms and which translations are to be grouped together must then be fed into the computer and the word frequencies recalculated. The frequencies of 'crash', 'crashing' etc. will be combined as will, for instance, examples of a word that has sometimes been translated by 'little' and at other times by 'small'. The computer can then be instructed to assign a number to each word that shows its frequency and which will appear on the screen when a translation is called for. These numbers will show the learner how frequently the word appears in the text she is working on. For some purposes it might be helpful to supply instead, or in addition, an indication of the word's frequency in a larger or more general sample of the language, but for the learner's immediate purposes, the frequency in the text at hand may be more important. That, after all, is what she must struggle with just now, and if a word occurs only once in that particular text, she may be well advised not to worry about it, even if it has a considerably higher frequency in other texts.

(6) Finally, the alphabetical list could also be used to construct a fuller dictionary, so that each word could be given a more general definition, one that will indicate the range of meanings that this word has in this particular text. It might be desirable to supplement this information with a more general definition of the word as used in a wider variety of contexts. It would even be possible to give the student computer access to various sentences in which the word appears so as to let her study the word in a variety of contexts. In one way or another, each word could be treated in a way that would be helpful for those occasions when the student feels the need for a fuller description of its use than is possible in a literal interlineal translation.

I see no serious problem in designing the computer program to accomplish a course of study such as I have described. Nor do I see any problem, other than that of time and effort, in developing the language materials. Once developed, this system should ease some of the most frustrating aspects of language learning. Of course, we may worry about finding ourselves tied to elaborate computer equipment, but this equipment is spreading

with such avalanche speed that accessibility is unlikely to be a problem for long. We may be even more worried about tying our student's eyes to the flickering green screen, but perhaps even this worry is less acute for the children of the television generation than it would have been to the generation of their parents.