# The Power of CALL

**Edited by**

**Martha C. Pennington**

Dr. Martha C. Pennington
English Department
City University of Hong Kong
83 Tat Chee Avenue
Kowloon Tong, Kowloon
Hong Kong
enmcpenn@cityu.edu.hk

# CONTENTS

# 7. A Principled Consideration of Computers and Reading in a Second Language

Tom Cobb and Vance Stevens

## 7.1. INTRODUCTION

It has often been noted that CALL lacks a solid research base (Dunkel, 1987, 1991; Roblyer, 1988). The problem lies mainly in two areas: inadequate reference to theories of language acquisition (Hubbard, 1992, this volume), and inadequate description of what students actually do, if anything, with specific CALL programs (Chapelle, 1990; Chapelle, Jamieson, and Park, this volume). The arguments made in this chapter in favor of using text manipulation activities to develop reading skills in a second language will attempt to address these problem areas. Evidence from research on student use of text manipulation will be presented.

## 7.2. WHAT IS READING COURSEWARE?

It is not at all clear what language teachers expect reading skills development courseware to do. While making insightful predictions concerning the impact of CD-ROM and laser printers, two devices neither widely used nor understood when his article was written, Wyatt (1989) placed use of computers in reading on a continuum constrained by development of orthographic recognition skills at the low end and mechanical/meaningful tasks such as comprehension exercises at the other. "Revolutionary" applications extended only to annotation (i.e., hypertext), modeling of productive reading strategies, and interactions with branching plot and adventure/simulation programs. While extolling the "raw potential" of the medium, Wyatt (1989) noted that "almost none of the existing courseware for second and foreign language reading skills has moved beyond the stage of directly paralleling the activities found in printed textbooks" (p. 64).

Teachers often assume that reading courseware might do something similar to what they do as a reading class activity. Indeed, much reading courseware does attempt to emulate what might be done in a classroom; hence the "reading comprehension" approach, where a passage is presented on screen followed by questions. In such courseware, computers can make existing techniques more efficient for the learner, in that feedback is immediate and interactive, possibly highlighting areas of the

text where attention could be most productively directed. The main drawback regarding the "efficiency" of this approach is the inordinate amount of time needed by developers to prepare each CALL lesson. For example, software which highlights context clues assumes that help has been set up for "*every* relevant word in every reading passage" [italics in original] (Wyatt, 1989, p. 73). Adding to the frustration is the work wasted if content which the software is tied to (e.g., in textbooks) is later deemed inappropriate and replaced in the curriculum. For these reasons, tools for producing this type of courseware are prone to lie fallow on developers' shelves after only one harvest.

More recently, computers have been used in reading in ways which do not emulate traditional methods of teaching and learning reading. Development along these lines has been directed not so much at the creation of new courseware, but at devising ways of making connections between an emerging battery of software tools and the proliferation of machine-readable text. One focus of this chapter, then, is to examine such connections in light of recent thinking on how reading skills are developed in a second or foreign language.

Hypertext is one means of making such connections. In its simplest form, hypertext allows annotations to on-screen text to be displayed on request. In more sophisticated implementations, hyperlinks can be developed to almost anything imaginable: e.g., video or sound segments, pathways into reference databases, annotations made by other readers, etc. These links might give students access to background and reference information, e.g., on-line access to tools such as dictionaries and encyclopedias (see Ashworth, this volume, for examples).

An example of the evolution of such courseware can be seen in the development of *Where in the World is Carmen Santiago?* and its offshoots (*Where in Space...*, *Where in the USA...*). In these programs, users try to solve a crime by discerning clues that enable them to track down a criminal moving freely throughout the virtual world (or space, or the USA). Solution of the mystery depends on "world knowledge" which, if lacking, may be augmented from a database of information supplied on a disk that contains appealing sound and animated graphics. More recently, CD-ROM versions of the program have come out, greatly increasing the amount of information and imagery that can be made available to crime-stoppers, as well as enhancing the sophistication with which this information can be accessed. In the CD-ROM version, the screen becomes a mouse-driven console providing a video telephone, a computer sub-screen for database access, a notepad, and a video window where pictures are displayed. The program produces a plethora of spoken discourse via the sound card, and whatever is spoken is generally printed out on the computer sub-screen (giving students who read it the benefit of vocalization). Many other CD-ROM-based multimedia packages offer similar rich mixes of reading and sound. *The Animals!*, for example, offers hyper-linked video and still-image tours and explorations of The San Diego Zoo. Authentic, native-level instructional text is spoken to users and also printed on the screen for those who prefer to

read it or who may have difficulty in following the spoken discourse (the category into which many second language learners would fall). Similarly, hyperlinked resource packages such as Microsoft's *Encarta* can immerse students into media-enriched target language environments, in which the comprehension of authentic written discourse is both encouraged and facilitated by sound and image.

Thinking along these lines, we might envision students solving similar language-enriched learning tasks by accessing authentic real-world databases over local or global networks, exploring the databases via hyperlinks, and annotating the materials or reading the annotations of others to achieve some result or resolution. The potential of these media in providing both a text-rich substrate for second language learning and the means and motivation for these materials to be used is becoming more apparent to those engaged in teaching and learning languages, as these powerful tools become more readily available and commonplace on networked microcomputers and stand-alone PC's.

When readers have widespread access to such tools, the concept of reading itself may change. Tuman (1992) argues that an "on-line literacy" is emerging which, while empowering some readers by allowing them to interact in compelling ways with text and with each other, will also lead to the demise of the author as the qualified and ever-present guide to a reader's private, sustained, and critical reading experience. Reading could soon be characterized by zapping one's way aimlessly around the "docuverse" of available materials. Thus, as with any application of technology to pedagogy, researchers will need to characterize the nature of the reading that takes place when learners are granted access to corpora and databases[1] and assess what affect this might have on second language reading in particular. Our own experience suggests that there is no guarantee that making large and varied amounts of on-line text available automatically promotes particularly deep processing, even when the task is in a motivating, pleasurable game format and other types of information are on offer. So, before we turn our students loose to cruise the information highway, we need to decide what they can use there and roughly to what effect.

Having speculated about on-line reading in the not-so-distant future, we would like to step back to a point where we are more certain of our position. The remainder of this chapter will suggest how students can be presented with copious amounts of text, along with exercises which we believe train strategies in comprehension of that text for language learners. In developing a theory supporting such implementation, we expand somewhat Wyatt's notion of courseware for reading, taking the concept beyond what is typically done in classes where reading is "taught". In particular, we support the text manipulation concept as a second language reading activity, as it is readily implementable on most present-day computer-based learning configurations, and as it is of particular value to students learning to read in a second or foreign language. Moreover, it can make use of the large amounts of text now becoming

available without departing totally from a pedagogy that we at least know a little about.

### 7.2.1.    Access to Text, the Computer-Based Reading Advantage

One of the most significant recent developments which impacts computer-based reading is the proliferation of and improved ease of access to machine-readable text. Text comes in over e-mail, is scanned from printed documents, is downloaded from CD-ROM databases in university libraries, is purchased as huge corpora from commercial suppliers, or is captured in endless streams from close-captioned television broadcasts. Consequently, experienced as well as less-skilled readers can anticipate increasingly wider access to text in a format which can be exploited in computer-based programs of reading instruction.

One of the most interesting aspects of computerized text is that almost all of it is authentic discourse. In light of Higgins' (1991a, p. 5) definition of authentic text as "anything not created by a teacher for the purpose of demonstrating language at work", the question then arises whether second and foreign language learners can cope with real-world written discourse. Happily, indications are that they can.

Bacon and Finnemann (1990) examined perceptions of general language learning (as reflected in attitudes, motivation, and choice of strategy), gender, and willingness to deal with authentic input for first-year Spanish students at two U.S. universities. They wanted to know whether these perceptions could be associated with comprehension, satisfaction, and strategy use in situations of authentic input. The results suggest that students perceive the value of authentic text to their learning and that they are not unduly constrained in processing it.

Similarly, Allen, Bernhardt, Berry, and Demel's (1988) study of 1500 high school foreign language students indicates that subjects were able to cope with all authentic texts they were presented with at three levels of difficulty. In an offshoot of that study, Bernhardt and Berkemeyer (1988) found that secondary level learners of German could cope with authentic texts of all types, and "that target language and level of instruction was a more important correlate of comprehension than was text difficulty" (Bacon and Finnemann, 1990, p. 460). Finally, Kienbaum, Russel, and Welty (1986) found from an attitudes survey that elementary level foreign language learners express a high degree of interest in authentic current events materials. These results all suggest that use of authentic text in second language reading can be motivating and not unduly daunting to second language learners.

The foregoing is of particular interest in light of Kleinmann's (1987) suggestion that reading courseware rarely provides adequate amounts of comprehensible input. Kleinmann found no significant differences in learning when a selection of twenty computer-based reading programs was used to teach reading compared to conventional reading materials, and he reasoned that the drill-and-practice nature of the CALL material prevented greater strides in learning by failing to address higher order reading skills, hence the need for more text. In his words:

> If we accept the notion that comprehensible input in the form of text material that is interesting, relevant, and at an appropriate level of complexity is crucial to second language development (Krashen & Terrell) then the nonsignificant findings with respect to the effect of CAI compared to non-CAI in the present study are easily understood. Very little of the available reading skills software meets these criteria of comprehensible input, especially for more advanced learners ... . Moreover, it will be necessary to develop software that stimulates general learning strategies that have been correlated with successful language learning, e.g., guessing, attending to meaning, self-monitoring (Rubin, Stern), as well as more specific strategies relating to particular skill areas. For reading skills development, strategies such as skimming, scanning, and context utilization will be important. (Kleinmann, 1987, p. 272)

So there is a prima facie case for channeling appropriate parts of the text stream through reading courseware designed for language learners. However, if beginning or intermediate learners are to be exposed to large amounts of authentic text, clearly they will need something to do with this text besides attempting to read or use it in their academic courses as if they were native speakers. Intermediate learners may be able to search through various on-line textbases, perhaps seeking answers to questions on a worksheet. However, because scanning for specific information requires only a modest degree of engagement with either high-level themes or low-level details of a text, this is not the type of reading development most beneficial to second language learners.

This chapter argues that TM (text manipulation) templates can engage students at higher cognitive levels while presenting them with virtually limitless amounts of comprehensible input in the form of authentic texts. Although scanning is not a skill that cloze activities encourage (Alderson, 1980; Feldmann and Stemmer, 1987; Nunan, 1985; Windeatt, 1986), work with text manipulation such as on-line cloze exercises may promote awareness of contextual help in restoring degraded messages (Bachman, 1982, 1985; Jonz, 1990) while exposing learners to a considerable amount of comprehensible input, assuming that learners take advantage of the amount of text that can be made available. And it appears from the results of the studies described above that use of authentic, ungraded text, rather than posing insurmountable problems for second language learners, might instead provide opportunities for the exercise of higher order processing skills called for by Wyatt, Kleinmann, and others.

### 7.2.2. Templates for Text Manipulation: Developer's Convenience or Sound Instructional Design?

It is not hard to see the attractions of linking text manipulation technology to the stream of on-line text becoming available. Copious amounts of

machine readable text, on the one hand, coupled with ease of implementation, on the other, makes appealing a template approach, where the courseware incorporates an algorithm which can be applied to any text supplied, realizing quantum savings in implementation time. Indeed, the distinctive feature of TM program design is that the program is able to deal with any text whatever.

TM systems can be quite varied, although they all have in common the algorithmic deconstruction of on-screen text for a learner to put back together. Some common types are the following:

- Cloze and other systematic deletions (suffixes, auxiliaries, transition markers, all function words, etc.);
- Scrambled elements (chunks, sentences, or paragraphs);
- Sentence boundary identification;
- Crosswords;
- Hangman or concentration-type word matching or supplying.

The developer's task is to find machine-readable features of text that correspond to something readers need to pay attention to, as indicated by either observation or theory. For example, if readers are observed to pay little heed to sentence boundaries, then an algorithm can be written to detect the surface features of sentence boundaries and then eliminate them throughout a given text so that the reader focuses on them by a process of re-insertion. Because such features are common to all text, one great advantage of a template approach is that texts of almost any genre can be shared among a set of driver TM programs.

On-line help can also be designed to take advantage of this commonalty of generic text. The only limitation is that the help must come from the text itself (or from the larger textbase from which the text derives) and be computable by an algorithm rather than coded ad hoc or "canned" (see Pennington, 1992a, for a discussion of the problems associated with "canned" CALL). Within this constraint, help can be any kind of information the text can provide which is relevant to the task at hand, from letting learners take a peek at the target reconstruction, to granting access only to that part of the context which will enable them to make inferences. One option made possible by the potentially large amount of text available is to provide help in the form of a concordance on the word the learner is trying to discover, with that word masked in the concordance output, giving learners richer context, but not the answer. The authors' present experiments are looking at user responses to on-line concordance as a help system for various word-level TM activities.

Text manipulation ideally uses any text, "raw" from its authentic source. However, the TM concept extends to cases where text is altered or annotated slightly to adhere to the particular requirements of the template, but in such a way that alterations do not render the text unusable by other text manipulation programs. For example, *Hopalong*, an implementation of the "speed read" approach to reading instruction developed by John Higgins, highlights text to guide the eye from chunk to chunk at a measured speed. All that the developer (e.g., the teacher or the

curriculum specialist) must do, after selecting the text, is to denote the chunks with carriage returns (in the case of *Hopalong*, the comprehension questions must be written in as well, but as these are in a separate file, the integrity of the original text is maintained). The already chunked text can be used directly in another of Higgins' programs, *Sequitur*, which displays the first chunk of text and has the student rebuild the entire passage by discerning the next chunk in sequence from among several proposed (i.e., the correctly sequenced chunk plus two distractors taken at random from the pool of not-yet-used chunks found elsewhere in the text file). The chunked text can in turn be used in a variety of other text manipulation programs which format the text according to sentence and paragraph boundaries (sentence-ending strings and blank lines, respectively), so that the integrity of sentences and paragraphs is essentially unaffected by the chunking required by *Sequitur* and *Hopalong*. Furthermore, the text can be part of a larger corpus used in concordancing or other forms of text analysis, from which still other text-based activities may be drawn (such as the concordance help feature noted in the preceding paragraph).

Thus, a wide variety of reading activities can be performed on any text considered relevant to the learners, who might prefer to restore or unscramble components from an article in a recent issue of *The Wall Street Journal* rather than perform equivalent operations in their graded reading workbooks. Or if the students prefer the graded readers, then these can form the text matrix. Whatever motivates the students is suitable text.

From a developer's point of view, the advantages of this approach to CALL implementation are obvious. However, the history of technology in education should alert us to the potential dangers of too-easy marriages of technology and instruction, which sometimes hides the fact that one partner has been made to adapt to the other (instruction to technology in this case). Dick (1991) has noted, with regard to the development of interactive multimedia instructional systems generally, that as the technology becomes more sophisticated the pedagogy tends to become more simplistic, often becoming detached entirely from any basis in instructional research.

The question to be addressed in the rest of this chapter is whether the TM approach yields corresponding benefits to second language learners, and particularly to their skills in reading. In arguing that it does, the authors will show how the activities students perform in text manipulation exercises are commensurate with current theories regarding productive reading strategies and environments favoring the development of second language reading ability.

## 7.3. THEORETICAL BACKGROUND: TM AND READING THEORY

For most people, reading is more agreeable and efficient on paper than on screen (Heppner, Anderson, Farstrup, and Weiderman, 1985). However, on-screen reading has the potential for overt interactivity. A reader can

send a message via the screen to a text, and then the text, properly coded, can send a message back to the reader. A paper text, by comparison, suggests a reader responding to a text whose fixed and independent meaning he/she must discover. Of course, for a skilled reader the process is interactive, whether via paper or screen text, except that with a paper text the interaction is mainly invisible, occurring in the mind of the reader.

Some notion of interactivity between reader and text characterizes virtually all post-behaviorist models of the reading process (various applications of the term "interactive" to the study of reading are discussed in Lesgold and Perfetti, 1981). In these models, the skilled reader is far from a passive responder to print stimuli, but rather a questioner, judge, summarizer, comparer, predictor, hypothesizer, and elaborator, depending somewhat on the type of text and a great deal on the reader's prior knowledge and current goals. A text's meaning for a particular reader is gradually constructed through the dynamic flow of information between reader and text, both "top-down" (reader to text) and "bottom-up" (text to reader). Of course, no two readers are likely to construct identical mental models of a given text, inasmuch as they bring to it different knowledge bases, purposes, and information processing strategies.

However, interaction with a text, although characteristic of skilled reading in the native language, is often problematic for second language readers, even those at a relatively advanced level of proficiency. The second language reader characteristically resembles B.F. Skinner's reader, passive before the text in attempting to extract its secret meaning. This characterization often holds true even for second language readers whose reading in their native language is highly interactive. The reasons for the prevalence of a non-interactive style of reading in a second language are many. Second language readers may not have automated one or more of the component processes of reading in the second language, such as word decoding and recognition, resulting in working memory overload and diversion of attention away from the construction of a text model. Or, at a higher processing level, readers may not be familiar with semantic or discourse schemata specific to the culture of the second language, so that they have no preactivated scaffolding to help them summarize and organize the details of the incoming text, and quickly face overload. For these and related reasons, many second language readers experience reading as a one-way flow of information coming from the text to them, and never send messages of the types suggested above back to the text. So one objective for second language reading courseware might be to encourage the automatization of certain controlled processes such as decoding; or it might inform the learner about certain discourse schemata or in some other way attempt to establish the preconditions for eventual interaction. Perfetti (1983) has advocated such a role for courseware with regard to young first language readers, and Frederiksen (1986) has implemented and tested related ideas in a second language context.

Text manipulation courseware attacks the problem in a different, but complementary, way. TM simulates the target activity itself, rather than giving practice in any of its preconditioning or component processes. At any of a number of levels of processing, text manipulation externalizes the otherwise invisible reader-text interaction and gives the reader supported practice in real interaction with the text. Readers faced with a text that has been deconstructed in one of the ways described above must operate on it by questioning it or hypothesizing about what it might mean or how it might fit together. Readers have no choice but to interact with the text if they want to engage in the computer-based reading activity: passive meaning-extraction is not an option. Admittedly, the simulations of interaction provided by a TM system may not be perfect ones. Many of the typical TM operations that must be performed to reconstruct a text involve cognitive processing at a level not far below the surface of the written text, whereas the target interaction is actually deeper, i.e., it is semantic. Nonetheless, we assume that a second language reader who, for example, uses the mouse to drag boxed sentences of a text into their proper place in discourse order, is doing something akin to what skilled native language readers do unconsciously when they read—such as puzzling out the logical connection between two sentences or supplying a bridging inference from memory or from the preceding text. Further, when the boxed sentence has been placed, we assume that the TM system's mechanical feedback then simulates the far more subtle confirmatory or disconfirmatory feedback supplied for the skilled reader by subsequent text itself.

How successfully TM operations simulate the high-level interactions that characterize skilled reading, and with what degree of transfer, are empirical questions. The best-case scenario is that the habit of interaction is transferable to on-paper reading regardless of the exact level of the interactions provided by a TM system. In any event, the alternative is worse—many second language readers get no interaction with text from solo reading and only second-hand and/or delayed interaction from classroom reading.

So far, then, we are arguing that TM is capable of tapping text in ways which we can currently implement and that the interactive model of skilled reading can serve to guide, control, and evaluate. However, alert readers (highly interactive ones, armed with appropriate schemata) will have noticed that this interactive-simulation idea of TM is phrased in a particular conceptual framework, that of information processing or cognitive psychology. Such readers may also be aware that adapting such a framework raises some controversies. In the battle with behaviorism, cognitivism may have seemed unified, but now that "we are all cognitivists", the subdivisions are assuming more importance. For example, even given an interactive view of skilled reading, how do we know that readers who are skilled interactants in their first language need support for a similar target interaction in a second language? It could be that higher level skills involving inferencing and integration with prior knowledge are completely transferable from the native

language. If so, it would be redundant to encourage learners to practice these skills and, worse, a diversion of time and attention from where it is needed—such as at lower levels of cognitive processing involving lexical knowledge and lexical access, where positive transfer is generally low or nil. A good deal of first language research locates the typical source of reading deficit at the lower rather than the higher level of skills (Perfetti, 1983, 1985; Stanovich and Cunningham, 1991), and the case has been extended to reading in a second language (Polson, 1992; Segalowitz, 1986). If true, this would be a serious argument against further development of TM, especially a new generation of it designed to exploit the proliferation of machine-readable text. We believe the argument is false, but must dredge up a little history to frame the issue.

### 7.3.1.    The Background to Interactive Reading: Reading as Writing

The interactive version of reading, with the reader contributing to the construction of text meaning in conjunction with the text itself, is often considered an attractive account of this ultimate human activity. In fact, this account rests on the less attractive fact that human working memory is far too limited for behaviorist theory to have much applicability to reading. The constant theme in cognitive studies from Miller (1956) onward is that the mind uses various tricks, like chunking and prediction, to compensate for processing limitations. Experiments have shown even simple acts of perception to be "knowledge-driven" to varying degrees, and more so complex information processing like reading. For example, on the level of word perception, Tulving and Gold (1963) found that deformed words were better perceived when primed by more context, in other words, by more prior expectation. On the level of discourse, Bransford and Johnson's (1972) "laundry story" showed that not only immediate comprehension but also subsequent memory for a story was determined by prior expectation. The studies are legion; the theme is that expectation, especially well-structured expectation (in the form of models, schemas, scripts, grammars, and other kinds of frameworks) is needed to cope with the otherwise overwhelming flow of incoming information. Such structures are also important in view of how much typically gets left out of texts and yet is required for their comprehension, to be supplied from the reader's store of "default", or schematic, knowledge (Minsky, 1975; Schank and Abelson, 1977).

To those interested in educational applications, the pedagogy of reading implied by this version of human information processing seemed straightforward. The application came mainly from Smith (1971) and Goodman (1967) under the heading "reading as a psycholinguistic guessing game". In their model, reading is barely perception-driven at all—at least, not after the first few sentences to set the scene. Having made predictions at various levels, from various contextual sources, and having activated the relevant schemas, the skilled reader "feedsforward" through the text, merely "sampling" from the words themselves and stopping for a closer look when there are mismatches with predictions. The role of text is thus changed from authoritative to

merely suggestive. In the frameworks of both Smith and Goodman, the reader constructs the text almost as much as the writer, and the beginning reader should be encouraged to be as constructive as possible. The crucial point as concerns pedagogy is that readers should be discouraged from any major effort to pay close attention to the text itself, such as careful word decoding.

"Reading as writing" was very much the original basis of the text manipulation concept. The deformed on-screen text simulates, and at the same time exaggerates, the limited usefulness of any text surface as given. A "storyboard" with every word masked apart from the title is essentially Goodman's idea of what any text "really" looks like to the brain: a set of suggestive symbols encoding a message to be reconstructed through interaction with any prior and contextual information sources available. This notion is opposed to that of a text as a set of fixed signs whose single meaning is to be determined linearly from the combined independent meanings of the words.[2]

The applicability of psycholinguistic reading theory to second language reading seemed obvious (Clarke and Silberstein, 1977; Coady, 1979), and by the 1970's the theory had assumed the status of dogma in EFL/ESL practice (see Grabe, 1991, for more background). Clearly, if native speakers must bring a lot of their own information to the act of reading, then the second language learner brings even more. If reading is a guessing game even when most of the words and discourse conventions are familiar, how much more of a guessing game it must be when a large proportion of the words and discourse conventions are unknown or not well-understood.

This view of reading suggests providing second language readers with a practice environment in which to develop guessing and related strategies, especially one that feeds back to the guesses in shorter loops than are provided naturally in the reading process. Therefore, in the late 1970's the case seemed strong for developing TM, and the theory matched the technology becoming available.

### 7.3.2. Problems with Reading as Reconstruction

Given the enormous influence of the Smith-Goodman view of reading in both first language and second language instruction, its prescriptions and effects have been remarkably little researched. Perhaps this is because the theory, as a processing model, is actually quite short on specifics, as Perfetti (1985) maintains. Perhaps it seemed as if the copious psychological evidence for top-down processing made testing of the "obvious" instructional application unnecessary (an assumption that is almost never justified). Least researched of all, of course, have been the CALL applications of the model. And many involved in TM believe that to undertake such research now would be irrelevant, as the reading theory underpinning this predictive model of reading has already started to unravel.

It was probably inevitable that the Smith-Goodman theory of reading would come in for some criticism during the late 1970's and 1980's,

since the pendulum has been swinging in first language reading all this century between expectation-driven and perception-driven reading theories, with the latter currently ascendant (Adams, 1990, provides good background to this theory). Fashion aside, however, some novel research paradigms and techniques emerged in these years that seemed to produce genuinely new information about the nature of skilled reading, resulting notably in the expert-novice comparison (Lesgold, 1984). Unexpectedly, in several studies seeking to identify the actual characteristics that divide skilled readers from unskilled, guessing and predicting often came in quite low on the list.

Sampling from a very large pool, Mitchell and Greene (1978) argued that Goodman's eye movement data could represent any number of underlying cognitive processes, and that when a less ambiguous measure was used, no evidence at all of the use of prediction in skilled reading would emerge. Their consistent experimental finding was that reading speed is not a function of the degree of predictability of a text. Balota, Pollatsek, and Rayner (1985) examined the visual mechanisms of reading directly and concluded it was simply not true that reading is driven mainly by "expectations and predictions about forthcoming information", with visual information providing only confirmatory evidence of predictions. Perfetti, Goldman, and Hogaboam (1979) discovered that while contextually predictable words are identified a little more quickly than unpredictable ones, even skilled readers' predictions are accurate at a rate of only 20-30% and therefore this cannot be the basis of their success. Graesser, Hoffman, and Clark (1980) found that for good readers, neither speed nor comprehension is significantly affected by the degree of syntactic predictability of additional words in a sentence, although weak readers are significantly aided by higher predictability. Possibly the most persuasive evidence is provided by Stanovich and West (1979, 1981), who uncovered an effect similar to that of Graesser, Hoffman, and Clark (1980), but for semantic predictability: Good readers are aided by semantic predictability, moderately and unconsciously, but weak readers rely on it strategically, to the extent that they are thrown off when their predictions are wrong.

The theme emerging from this research was that poor readers guess and predict a good deal, because they do not know enough words, do not know them well enough, or cannot quickly enough recognize visually those words which they know phonologically so as to beat the rate of information decay in working memory. A coherent sequence of studies on this subject is reviewed in Perfetti (1985). Study after study in the 1980's showed speed of context-free, expectation-free, word decoding to account for the main part of the variance in multiple regression analyses in which numerous reader attributes were pitted against general reading comprehension as the dependent measure. The instructional implication is that practice in rapid word recognition, not practice in guessing, is what can turn weak readers into strong.

The decoding issue was slow to arrive in second language reading theory, possibly because reading-as-predicting had become such a domi-

nant view (as suggested by Grabe, 1991). However, a sign that the tide is turning can be found in a number of the contributions to Huckin, Haynes, and Coady (1993), which qualify severely the nature, role, importance, and conditions of guessing in reading in a second language. Coady, as noted above, was one of the original importers of psycholinguistic notions of reading into second language acquisition theory. The emergence of findings counter to guessing theory suggests that CALL reading software, rather than promoting the development of strategies in predicting and hypothesizing, would be better devoted to helping learners develop the ability to automatically decode the highest frequency words. In fact, some large-scale CALL projects now seem headed in that direction (for example, Coady, Magoto, Hubbard, Graney, and Mokhtari, 1993).

If second language theory and practice were to embrace the latest first language reading theory as quickly and thoroughly as it once did the so-called psycholinguistic theory, then we shall inevitably all be teaching word lists and rapid decoding via our various media. Selinker (1992) characterizes EFL/ESL as a field fond of throwing out the little it achieves in periodic swings to discover ever newer and more exciting theoretical underpinnings. Second language reading research is bound to follow the lead of first language research in significant ways, given the relative size and gravitational pull of the two enterprises. In any case, it is no doubt true that there is a greater role in reading in a second language for more specific vocabulary and word recognition training, particularly at the early stages, as argued by many of the contributors to the volume by Huckin, Haynes, and Coady (1993). However, an argument can be made for encouraging second language reading researchers to be more discriminating about what they borrow from first language research and how they interpret and adapt it (also the view of Grabe, 1991).

### 7.3.3.   Reading in a First and a Second Language: Same or Different?

First language reading research does not map onto second language reading in any simple or obvious way. Even Perfetti (1985), an arch foe of guessing theory, suggests as much, noting that:

> Skilled reading is, by definition, a very fluent process. If a skilled reader fixates three or four words per second, around the normal rate, where is there time to guess? Moreover, if he is skilled at reading, why bother? Reading is much easier than guessing. The case may be different in, for example, reading in a foreign language that is incompletely mastered. There is plenty of time to guess in such cases and perhaps enough payoff for doing so. (p. 26)

Studies looking into subtle differences between first language and second language reading are somewhat sparse. However, a number have attempted to replicate some of the first language reading experiments mentioned above with second language readers and obtained rather different results. For example, the key Stanovich and West experiment mentioned above was replicated in Quebec by Favreau and Segalowitz

(1983) with skilled and less skilled bilingual readers, and patterns of context sensitivity were found that did not confirm the Stanovich and West results. What Stanovich and West characterized as less skilled readers' over-reliance on and yet poor use of contextual information was found precisely to characterize slow but otherwise highly skilled second language readers. In other words, both weak first language readers and skilled second language readers appear to be strategically reliant on context to recognize a large proportion of words, and yet not very successful in using the information context offers. Therefore, skilled, flexible, automated use of context apparently does not automatically transfer from the first to the second language, even when the foundations for such transfer appear to be in place.

Second language readers' apparent context-insensitivity even when otherwise highly skilled in reading is not an extensively documented phenomenon, yet it appears to exist. For example, it appears in a series of mainly unpublished studies described in McLaughlin (1987) and McLeod and McLaughlin (1986). The latter study compared the read-aloud errors of both more and less skilled second language readers against those produced by first language readers in terms of meaningfulness, or contextual goodness-of-fit. One sentence in the text the subjects read was, "She shook the piggy bank, and out came some money" (McLeod and McLaughlin, 1986, p. 115). Predictably, if young first language subjects did not know the word *money*, they might replace it with *dimes*, a semantically reasonable alternative. But if second language students did not know the word *money*, they tended to replace it with something orthographically similar but contextually violating, such as *many*. This tendency was even more interesting with the advanced ESL students in the study. Advanced students made far fewer errors than beginners, as one would expect, but of those that remained, just as large a proportion were context violating or non-meaningful. This phenomenon was confirmed by McLaughlin (1987) in a cloze test given to both advanced and beginning ESL readers as well as to native speakers. The advanced readers scored significantly higher than beginners, but once again the point of interest is in the character of the errors that remained: only 20% of beginners' errors were plausible within the context, and for advanced readers the figure was only 29%; for native speakers, the figure was 79%. In other words, if recognition was not automatic, there was no workable strategy for producing a reasonable guess. A few other experiments confirm the existence of this phenomenon in second language learners, including Arden-Close's (in press) work in Oman; and it has been noted as well in first language studies (e.g., Oakhill, 1993).

Thus, direct instruction or practice in reading-as-interaction or even reading-as-educated-guessing makes some sense in principle in the second language context, whatever other realities may exist in the case of reading in the first language. Therefore, the recent decoding movement has not made the idea of reading as interaction, or its applications such as TM, untenable. Psycholinguistic reading theory has not then been unraveled as much as it has been moderated, supplemented, and specified, in

that it has been shown to have a special relevance in the second language case. In order to delimit the potential for TM to train the reading process in the second language case, a number of empirical questions need to be explored:

> (1) To what extent does work on text manipulation software produce context sensitivity for second language readers of various types and at various proficiency levels?

> (2) Does TM produce a more interactive reader, who habitually integrates text information with his/her own prior knowledge in such a way that non-grammatical sequences become impossible?

> 3) Who needs TM-based training, all or merely some readers, and how do we find out?

> (4) If training is required of both high-and low-level reading skills, what are the optimal proportions and sequencing of these skills?

> (5) What, if anything, do learners actually do with reading courseware of different kinds? What are the variables in their behavior and the outcomes of that behavior? What strategies (if any) seem to emerge in a CALL reading context?

This excursion into theoretical background, we would argue, builds a plausible case for text manipulation in line with what is currently known about the reading process, and suggests a number of hypotheses for empirical research and a rationale for doing that research. Where should one begin an empirical examination? We follow Long's (1980) argument that in second language acquisition research, the research cycle ideally moves from descriptive, to correlational, to experimental studies, and that no phase should be skipped. Chapelle (1990)—see also discussion by Chapelle, Jamieson, and Park, this volume—has proposed the applicability of this cycle to the CALL area, and, as mentioned before, notes that the descriptive phase has hardly begun.

Cloze is one template for reading and language skills development on which a substantial body of research has been carried out toward describing what reading skills it exercises, and it is also a template which lends itself well to text manipulation. Therefore, it seems reasonable when embarking on a course of inquiry into text manipulation to take as an example what has been done with computer-based implementations of cloze.

## 7.4. SECOND LANGUAGE READERS AND CLOZE

### 7.4.1. Some Problems with Cloze

Lee's (1990) survey of the previous decade of research on reading examines several genres of research instrument including cloze. The section on beginners draws heavily on Nunan (1985), who finds that:

> unlike more advanced learners, beginning language learners are less able to perceive (or perhaps utilize) intratextual relationships when carrying out written cloze tests ... . Beginning language learners are not able to take in the text as an integrated expression of ideas, when the text is violated by blanks. This finding may be a by-product of the fact that the text itself, as presented to readers, is not an integrated expression of ideas. (p. 5)

Similarly, Douglas (1981) finds that advanced second language readers, unlike native-speaking readers, are more reliant on local redundancy in a text than they are on longer range redundancy in their completion of cloze exercises.

If, as has already been noted, second language readers are not even able to perceive non-degraded text as an integrated expression of ideas, it is not surprising that a degraded text such as a cloze passage would be even more impenetrable. This partially explains Feldmann and Stemmer's (1987) finding that of twenty subjects in their study of C-tests (which are similar to cloze tests), only two attempted to skim the entire text as instructed, and they gave the task up as impossible because of the gaps. Cohen, Segal, and Weiss (1985) instructed students to skim cloze passages first, and reported a similar breakdown. Alderson (1980) gives further evidence of students not treating cloze passages as integrated readings, and concludes that "the nature of the cloze test, the filling in of gaps in connected text, forces subjects to concentrate on the immediate environment of the gap ... ." (p. 74). He further finds that varying the amount of context has no predictable bearing on the ability of either native speakers or non-native speakers to solve cloze tests: "Neither native nor nonnative speakers were aided in their ability to restore deleted words, or a semantic equivalent, or a grammatically correct word, by the addition even the doubling, of context around the deletion" (Alderson, 1980, p. 72).

If paper-based cloze poses such problems for language learners, then one advantage to computer-based cloze is enhancement of the reader-to-text interaction made possible when the gaps violating the text respond to the students' attempts at recovering them. In the first place, students receive feedback as they go, to whatever degree granted by the program designer. Secondly, as Feldmann and Stemmer note, it is possible that as the text is resolved, the learners have more and more redundancy at their disposal to elucidate unsolved blanks, and students working on computer-based cloze activities have the added advantage of knowing

whether blanks solved have been filled in correctly or not (incorrect words left intact in paper-based cloze passages might further skew meaning). So, whereas a computer-based cloze passage may initially appear indecipherable to students, they are at least handed a set of tools to work with in attempting to tease the message out of the text as they render it gradually less degraded. That second language learners are in fact able to work effectively in interaction with computer-based cloze has been borne out in at least one study (Stevens, forthcoming). Before discussing that study, however, we will discuss a crucial choice to be made by researchers in collecting data in such studies.

### 7.4.2. Examining Text Manipulation Non-Intrusively

Because of the interest of cloze to researchers as a measure of language proficiency, learner strategies when working on cloze passages have been extensively examined, though not necessarily as computer-based exercises. One useful description occurs in the work of Feldman and Stemmer (1987), who found that in solving C-tests, solution of gaps was either "automatic" or "non-automatic", i.e., spontaneous or considered. In the latter case, recall strategies were used, leading to delay, to giving up, or to activation of another recall strategy. Once an item was recovered, evaluation strategies were employed to check appropriateness (also used for automatic recovery), leading to acceptance or rejection of the item for that blank. Since production problems (e.g., spelling) could still occur after recall of the item, application strategies might also have to be employed.

Since they felt that their use of student introspection as one means of generating data was a factor in their study, Feldmann and Stemmer comment on gathering data on cognitive processes intrusively. An "intrusive" protocol is one for which the act of gathering data interferes with the process under study; for example, where the presence of video equipment or the need to "think aloud" causes learners to monitor their behavior more closely than they might if left to their own devices. In solving cloze passages or C-tests, students are constrained in what they can process simultaneously. Signal data limits occur when the quality of the data is eroded, as with phone call interruptions, or in the case at hand, with the blanks in a cloze exercise. Memory data limits occur when language items are encountered which the learner does not know or has forgotten. Furthermore, there are resource limits, where the learner is given too much to process beyond his/her capabilities. Focus on multiple tasks can be maintained until one task starts drawing attention preponderantly from the others. Feldmann and Stemmer (1987) suggest that having to think aloud could interfere with the subjects' ability to focus properly on the task under study.

In order to get a clear picture of actual self-access use of TM, some researchers opt for non-intrusive research techniques. Unobtrusively tracking key presses of second language students performing computer-based cloze activities in unmonitored self-access situations has yielded evidence of engagement in interaction with the text of the type noted

above—e.g., hypothesis formation, testing, and reformation. In a study of 100 cloze paragraphs completed by second language learners at university level, Stevens (forthcoming) found that students successfully used feedback from the program to substantively complete 36 of the passages (with 22 of those paragraphs entirely completed). However, there is also evidence in the same data of students giving up on passages which they had started: 49 of the interactions were essentially nil sessions, where students logged on, checked things out, and logged off again with little or no interaction; and a further 16 quit after working only within the first sentence. Although it is not clear if this minimal time spent on computer is because the students were unable to complete the passages or simply did not want to complete them, the latter possibility seems more likely, as use of the hint and help features built into the program practically guarantee solution of any problem by anyone who persists.[3]

We therefore find that many of the students in the Stevens study were simply "window shopping", just looking for something to do for a few minutes, but not in the mood for cognitive engagement. This appears to be fairly typical student behavior, and indeed, many computer users, not just students, enjoy browsing. There is probably nothing inherent in the medium that would elicit this outcome other than the fact that it was possible, given the circumstances of the investigation, to gather data unobtrusively, without students knowing that they were being monitored, and so these data were collected. Such data might not have emerged in an intrusive study. This is but one way that results from non-intrusive studies might contradict those from intrusive ones. As another example, Windeatt (1986), in a study where screens were videotaped as students thought-aloud about their reading processes while going through the text and were later interviewed about their experiences, found that while reading the students made little use of program help features (see also Hubbard et al., 1986). The unobtrusive studies of Stevens (1991a-c, forthcoming) suggest, however, that students working under self-access conditions tend to abuse help features rather than to apply more self-reliant cognitive strategies in solving the problems they encounter. If whether students know they are being monitored is a factor in their use of computer-based help, then whether a study is intrusive or not is itself an important consideration in assessing the results.

## 7.5. LEARNER CONTROL ISSUES

There is some evidence that students who rely excessively on program-supplied help are not learning as much as those who try to solve problems through their own self-generated trial-and-error feedback. Pederson (1986), for example, demonstrated differences in cognitive processing when comparing students who had access to help in the form of the option to review reading passages while answering comprehension questions as contrasted with those to whom such access was denied. In the author's words: "The results indicate that passage-unavailable treatment always resulted in a comparatively higher comprehension rate

than occurred in counterpart passage-available treatments regardless of the level of question or level of verbal ability" (Pederson, 1986, p. 39). In other words, "greater benefit was derived from the subjects' being aware that they were required to do all of their processing of the text prior to viewing the question" (Pederson, 1986, p. 38). It follows then that in using text manipulation as a means of having students engage in "reading as guessing", help should not be allowed to such an extent that guessing is suppressed.

One strategy frequently noted when students use TM programs is a tendency to proceed linearly rather than holistically, as one might be expected to do if reading a passage and drawing inferences from outside the immediate context. Edmondson, Reck, and Schroder (1988) tracked nine secondary level students doing a combined jumbled sentence/paragraph exercise called *Shuffle* and noted a tendency for students to use "frontal-attack" strategies; that is, to take the first available sentence and try to place it, or to build from the first sentence to the next, and so on. Accordingly, Windeatt (1986) found that his subjects completed computer-based cloze blanks in a predominantly linear fashion, even though the system did not require it—perhaps because they did not like to scroll from screen to screen—and similar findings have consistently obtained in more recent work by the present authors (e.g., Stevens, forthcoming). If, as Windeatt suggests, this tendency to proceed linearly with computer-based exercises occurs at the expense of more holistic strategies, then it may be that a more effective implementation would encourage or even force students to jump around in the text instead.

The possibility (indeed, the likelihood) that students may not of their own free will choose a pathway through CALL materials leading to optimal learning suggests a re-examination of the magister-pedagogue dichotomy introduced by Higgins (1983, 1988) which has strongly influenced CALL software development over the past decade. Rather than the computer acting as a pliant slave which unquestioningly obeys all student commands (the role favored in the dichotomy), it may be that an entity which aids the learner on demand while exercising enlightened authority over the learning process is more conducive to learning. But how much authority can a program exert without depriving students of benefits of autonomous learning (thus tending to be a magister, in terms of the dichotomy)?

One problem with allowing learners control over their own learning is getting them to take advantage of available options. How, for example, can students be encouraged to select and learn to interpret unfamiliar forms of feedback? Bland, Noblitt, Armstrong, and Gray discovered in a SYSTÈME D implementation that although students had access to both dictionary and lexical help, they avoided lexical help for fear of getting lost in it. "We were initially surprised at the very few queries of this nature in the data" (Bland, Noblitt, Armstrong, and Gray, 1990, p. 445). Furthermore, in an attempt to reverse the outcome of the Stevens' (1991a) Hangman study, where it was found that 53% of the students were

touring the material with unacceptably low levels of cognitive engagement, the program was reconfigured to present varying amounts of context surrounding the target word when demanded by the user. The demand feature comes at the cost of points, the idea being for students to request just as much context as they need to solve the problem. On examination of the first set of data after the revised program was implemented, it was found that cognitive engagement remained at about the same level and that the students were not using the context feature, probably because the program failed to make them aware of it. These are just two examples of the caveat that simply providing options to students by no means ensures that they will use them.

One of the present authors is finding much the same thing in his research into learners' use of on-line concordancing with keyword masked as the help in a systematic deletion exercise (*Textpert*). In this study, learners' use of concordance help in self-access was virtually non-existent, in spite of their previously having tried it in a practice session, and also (in the practice session) having doubled the success rate of either a no help or dictionary help option. In order for the experiment to continue, the system had to be reconfigured three times to make the concordance window unavoidable (*Petwords*). Admittedly, spontaneous use of the concordance increased with familiarity, but not entirely in proportion to the increasing advantages it produced, both on-line and later on in classroom paper-and-pencil cloze tests for the same vocabulary items.

## 7.6.  CONCLUDING REMARKS TO CHAPTER 7

In this chapter, we have attempted to broaden the notion of reading courseware beyond a replication of what instructors might try to teach in a reading class to courseware that emulates the reading process. We submit that text manipulation, besides being easy to program and to implement in an educational context, is capable of promoting interactive reading. Most importantly, TM programs are able to take advantage of the increasingly widespread access to machine-readable text and are thus potentially able to supply learners with substantial amounts of input, which teachers—or the learners themselves—can filter to ensure that it is comprehensible input.

Designers of TM programs are often accused of succumbing to expediency at the expense of pedagogical merit in churning out text manipulation templates. In this chapter, an attempt has been made to explain how text manipulation programs enhance the reading process by promoting interactions with the text. In particular, TM can provide feedback that enables second language readers to perceive meaning when work with the same text might be too difficult for them if attempted via less interactive means.

It now seems that whatever processes are instigated by TM are beneath the learner's level of conscious perception or attention. This should come as no surprise, as the same applies to much of language learning. Although the "reading as guessing" model from which TM is

derived has been challenged, it is shown in this chapter that reading in the native language differs enough from second language reading that much of this criticism applies only obliquely to the second language case. Hence, there remains a plausible scenario for pursuing the development of TM materials, particularly in second language reading, and in conjunction with other types of materials aimed at a lower or less holistic level.

However, this plausibility must be supported by more definite evidence that TM actually produces differences in skill acquisition over alternatives, on-line or off. In this chapter, the importance of making empirical inquiry into positions taken with regard to TM is stressed, and work on cloze is taken as an example of one such line of inquiry. Notes of caution are sounded in interpreting results without taking into account the degree of intrusion in the process afforded by the protocol, and also in assuming that features built into a program will as a matter of course be used as expected by students.

Developers at this point should take advantage of the descriptive data available and feed it back into the design process, particularly that part of the process relating to learner control. As pointed out by Chapelle and Mizuno (1989), the issue of optimal degree of learner control over CALL had not yet been investigated in the latter part of the 1980's. With investigation now tentatively under way, it is fair to say that the issue of learner control is still far from being resolved in CALL or TM, as is also true in the wider world of computer-assisted instruction (see Steinberg, 1989). We are finding that we may have to make our TM programs somewhat more magisterial if we wish to obtain significant research results from their use. The questions we must now address concern what we need to do so that our learners will use them most effectively as well.

NOTES TO CHAPTER 7

1. Ease of access would be an important variable determining the nature of the reading process based on corpora and databases.

2. Although a TM routine insists in the end on a single exact surface reconstruction, the focus on a set text can be de-emphasized to some extent by imaginative programming.

3. In this study no attempt was made to identify individual students; thus, there were no violations of privacy, and also no compunction on students to concentrate on the task unless self-motivated to do so.