First and Second Language Reading Strategies: Evidence from Cloze

Patricia L. Carrell and Joan G. Carson with Dong Zhe

Georgia State University

In a variety of different studies, cloze tests have been used to examine the similarities and differences between native and nonnative speakers' performance on the same task and across languages. In the present study, in addition to using cloze tests to examine native and nonnative speakers' performance both on the same task and across languages, cloze performance is examined as a window on native and nonnative readers' strategies. The performances of native speakers of Chinese and of English on a random deletion cloze in their native language, and, in the case of the Chinese, also in English as a second language were closely examined for what they reveal about native and nonnative readers' strategies and about reading in the two different languages. Results indicate that native and nonnative speakers perform similarly on cloze in a given language (English) — that native and nonnative speakers seem to use the same strategies on cloze in English — and that readers perform differently in different languages (Chinese and English). Chinese and English readers' responses tended to be similar in English, with differences due to whether the language was the reader's native or foreign language. However, Chinese readers' responses were not similar on the Chinese and English clozes.

INTRODUCTION

Cloze tests, originally introduced as a method for testing the readability of text, are often used as measures of language proficiency and reading comprehension. The traditional cloze procedure is a technique in which every nth word of a text is deleted, and the reader's task is to reconstruct the text by filling in the blanks. Because a cloze task requires the reader to supply words that are appropriate morphologically, syntactically and semantically, cloze tests are assumed to be a good measure of language proficiency. And, because a cloze task forces the reader to use context to predict meaning (i.e., the appropriate word), cloze tests are also assumed to mirror and measure reading comprehension.

Much discussion of cloze tests has focused on the question of whether cloze is sensitive to intersentential phenomena or whether it can only provide information about lower order skills, such as clause-level syntactic phenomena. If the cloze is to be accepted as a valid test of reading comprehension, then it must also measure

Patricia L. CARRELL is Dean of the College of Public and Urban Affairs and Professor of Applied Linguistics/English as a Second Language at Georgia State University. Her research interests include text and reader factors, including language, which affect reading in English as a foreign/second language. Her most recent publications have appeared in the TESOL Quarterly, Applied Linguistics and the Modern Language Journal.

Joan CARSON teaches graduate courses in the M.S TESL program at Georgia State University. Her research interests are in academic language and biliteracy. She has published in TESOL Quarterly, Language Learning and the Journal of Second Language Writing.

DONG Zhe received his Ph.D. in Educational Administration from Georgia State University.

The authors may be contacted via Patricia L. Carrell, Office of the Dean, College of Public and Urban Affairs, Georgia State University, P.O. Box 4018, Atlanta, Georgia 30302-4018.
discourse constraints across sentences. Although cloze testing is not uncontroversial, much evidence suggests that a cloze does measure reading comprehension beyond the immediate context of the deleted item (see Jonz, 1990, for a review of this debate). As such, the cloze test has become an important tool for second language researchers investigating similarities and differences between first and second language reading performance.

Studies show that native and nonnative speakers often perform similarly on cloze tests. Carroll, Carton, and Wilds (1959) studied English and French or German adult bilinguals and found no difference in mean scores between the readers’ first and second languages. Oller and Conrad (1971) found no differences between advanced nonnative speakers and college freshmen native speakers of English, although there were differences between beginner and intermediate nonnative speakers and college graduate student native speakers. Alderson (1980) found that while native and nonnative speakers’ performance on cloze tests were significantly different, there was considerable overlap. Interestingly, an exact word scoring procedure yielded the smallest difference between the two, while semantically acceptable scoring produced the greatest difference. Alderson concluded that “in all essential respects the cloze task is the same for native and nonnative speakers” (p. 73).

Jonz’s (1987) study of native and nonnative speakers showed that nonnative speakers, like native speakers, attended to both bottom-up and top-down cues in reading, suggesting that similar strategies may underlie similar cloze performance. Do native and nonnative speakers use the same reading strategies?

Several early studies presented evidence of relationships between various reading strategies and successful or unsuccessful second or foreign language reading (Hosenfeld, 1977; Block, 1986; Hauptsman, 1979; Knight, Padron, & Waxman, 1985; Sarig, 1987). Other researchers found evidence that native and nonnative speakers used similar reading strategies on cloze tasks (Cziko, 1980; Hauptsman, 1979; Markham, 1987). In fact, cloze is often suggested as a pedagogic technique for teaching successful reading strategies to second language learners (Borodkin, 1991; Meyer & Tetrault, 1986; Raymond, 1988).

Yet the picture is more complex than suggested by these studies and proposed pedagogies. A study by Pritchard (1990) shows that comprehension is affected by the interaction of reading processing strategies and cultural content schemata. Furthermore, research reported by Anderson (1991) shows that there are no simple correlations or one-to-one relationships between particular strategies and successful second or foreign language reading comprehension. Rather than a single set of processing strategies which significantly contributed to successful reading comprehension, Anderson concludes from his data that successful second language reading comprehension is a matter of the reader controlling a wide and flexible repertoire of strategies and knowing how to use that repertoire strategically.

In a 1986 study, Barnett related performance on certain types of cloze items to lexical ability and grammatical ability in foreign language reading. She used a 50-item multiple-choice rational deletion cloze test which she described as follows: “The correct choice for half the blanks depended upon a knowledge of syntax; the other half depended upon a knowledge of vocabulary or upon attention to semantics with the text” (p. 345). Thus, subjects’ performance on the one type of item reflects their use of syntactic strategies, their performance on the other type of item reflects their use of lexical or semantic strategies. The study reported in this article similarly viewed subjects’ responses to cloze items as indicative of their use of reading strategies.

To move beyond simply investigating second or foreign language reading and to attempt to compare that reading to first language reading introduces additional complexities, as the recent AILA Review Number 8 (1991) on Reading in Two Languages attests. We may be interested in the behavior of the same readers in both their L1 and their L2, or we may equally be interested in comparing the reading of L2 learners to that of native speakers of a particular language.

The picture of cloze performance as similar for native and nonnative speakers on the same task suggests that readers tend to use similar strategies when performing the same cloze task regardless of their different language backgrounds. However, this picture becomes more complex when we look at cloze performance across languages. Tzeng (1983) summarized research that indicates that different orthographic systems may be processed differently. “...reading skills acquired in one orthography may not be the same as those acquired in another orthography if these two orthographies have different script–speech mapping rules” (p. 92).

However, it is not just orthography that causes differences. Grundin, Courtney, Langer, Pehrsson, Robinson, and Sakamoto (1978) administered comparable cloze tests in English, Japanese, and Swedish to 150 fifth graders in three countries. Japanese readers scored higher (52%) than either English (45%) or Swedish (44%). The researchers concluded that “a given cloze percentage need not have the same meaning or significance in different language areas. In fact, it is even possible that the cloze procedure does not measure exactly the same aspect of reading ability in different languages” (p. 55-56).

Finally, Turner (1989) tested Francophone ESL learners on their cloze performance in both languages. The best model to explain the results was one with three orthogonal factors: L1, L2, and a general cloze-taking ability (CTA) factor. This general CTA factor was independent of the first two, and was relatively greater than either the L1 or the L2 factor in both L1 and L2 cloze performance, although
smaller in the L1 than in the L2. Turner hypothesizes that the language factors (L1 and L2) do not, in fact, measure the same abilities in the respective languages. Rather, she suggests that the L1 factor measures higher-order skills, such as the ability to use textual constraints, whereas the L2 factor measures the ability to use lexical and syntactic constraints.

Thus, despite the fact that L2 speakers may perform similarly to native speakers on the same cloze tests, there is evidence that readers may, in fact, perform differently when doing cloze tasks in different languages. If this is true, then it calls into question the notion that good L1 readers may be good L2 readers because they use the same reading strategies in L2 that make them successful readers in their L1. It may not be the case that good L1/L2 readers use the same strategies in both languages. Or good readers in both L1 and L2 may use different strategies in each language; i.e., they may be good readers in each language, but for differing reasons.

If readers read differently in different languages, what accounts for the similarity between native and nonnative speakers reading in the same language? Do reading strategies vary with the language? Or do L1 strategies allow readers to be successful in the L2?

In this paper, we present and discuss the results of a study designed to explore reading strategies used by readers in both their L1 and their L2, as well as comparing L2 reading to that of native speakers of the L2. Since no language processes are directly open for observation (and even when they are, as in the case of eye movement measurements taken during reading, the raw data requires interpretation), like other researchers, we make a cognitive leap between our observed data and what we believe that data represents. Similar to the longer traditions of interpreting reading miscue data as windows on reading processes, and of interpreting learner errors as windows on the language development process, and similar to the practice (Barnett, 1986) of examining certain types of cloze items as drawing upon certain types of abilities (or strategies), we reasoned that readers' responses to cloze items would provide a "window" on reader strategies, since all responses which were not exact word replications would provide evidence of what readers were trying to do, of the strategies they were utilizing as they tried to make sense of the text.

METHOD

Data were gathered as part of a much larger study of reading and writing relationships in first and second language (Carson, Carrall, Silberstein, Kroll & Kuehn, 1990).

Subjects

One group of subjects consisted of 60 native speakers of Chinese from five different countries (Taiwan, N=29; Malaysia, N=12; People's Republic of China, N=10; Singapore, N=5; Hong Kong, N=4), with almost one-half from Taiwan. They were studying at one of nine different universities from all across the United States. Forty two (42) were enrolled in basic freshman English composition courses as fully matriculated students, the others were enrolled in pre-academic intensive English programs (Advanced, N = 7; Intermediate, N = 11).

The second group consisted of 28 native speakers of English, traditional freshman composition students from one of the participating universities, who provided baseline data.

Materials

The reading test consisted of two cloze passages, one in Chinese, one in English. Reading passages to be clozed were selected using the following criteria: (a) the topic of the passage was of general interest; (b) the passage was an authentic text intended for native readers with high school level reading skills; (c) the passages exhibited the same rhetorical organization (comparison and contrast organization was arbitrarily selected; and (d) the length of the passage had to be between 325-400 words. A native Chinese speaking consultant assisted with the Chinese cloze, selecting and clozing the text from a Chinese newspaper (People's Daily, June/July 1987) on the differences in family environments. The English text was on the differences in the effects of nature vs. nurture in personal development and was taken from an advanced ESL text (Adams & Dwyer, 1982).

To keep the reading tasks as comparable as possible across languages, we chose to use a fixed-ratio random deletion cloze, deleting every 7th word. The first sentence of each passage was retained intact. Clozing the Chinese text proved somewhat difficult due to the fact that there was no clear definition of a word as an orthographic unit in the Chinese writing system. The Chinese consultant's decision was that a word had been identified if a character or characters could stand alone with the same meaning. For example, if the word for "sun" consisted of the characters for "bright" and "ball," then those two characters were said to constitute one word - "sun." The English passage contained 53 blanks, the Chinese 44.

Instructions, which included a sample sentence with words written in the blanks, were written in English and translated into Chinese. The Chinese instructions allowed subjects to write in either complex or simplified characters, although the Chinese text was written in complex characters, the more universally readable by the Chinese. All passages were then typed and the space allotted for each cloze item was standardized across items and across languages.

Procedures

The data were collected at the beginning of academic terms during the fall of 1989. Subjects were given between 30 and 45 minutes to complete each task. L1 and L2
reading tasks were counterbalanced and administered within a two week period. No dictionaries were allowed, and subjects were instructed to read through the entire passage before going back and attempting to fill in the blanks.

**Scoring and Reliabilities**

Cloze passages were scored by two native-speakers for each language, using an acceptable scoring procedure. Figure 1 shows the scoring categories and criteria for each category. Each response to each item was categorized into one of the eight categories. Figure 2 illustrates the scoring categories by providing the beginning of the English cloze text with examples of responses to the third blank item and their respective categorization.

<table>
<thead>
<tr>
<th>ACCEPTABLE</th>
<th>exact, consistent, or acceptable; semantically and syntactically acceptable; see below.</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXACT</td>
<td>exact; same word as that in original text.</td>
</tr>
<tr>
<td>CONST</td>
<td>consistent; a lexical synonym as might be found in a dictionary or thesaurus; grammatically and semantically perfect.</td>
</tr>
<tr>
<td>ACCEPT</td>
<td>acceptable; although the meaning may vary from the original word, it makes sense in the context of the entire text and does not change the gist of the text meaning even though it is not a synonym; both syntactically and semantically acceptable.</td>
</tr>
<tr>
<td>SEMA</td>
<td>semantically acceptable; semantically acceptable according to preceding criteria, but not syntactically acceptable.</td>
</tr>
<tr>
<td>SYNA</td>
<td>syntactically acceptable; syntactically acceptable in terms of sentence syntax; could stand as an acceptable sentence, although meaning differs from meaning of original word or from text level meaning.</td>
</tr>
<tr>
<td>UNACC</td>
<td>unacceptable; unacceptable semantically and syntactically.</td>
</tr>
<tr>
<td>BLANK</td>
<td>blank.</td>
</tr>
<tr>
<td>UNREAD</td>
<td>unreadable; not able to decipher what is written.</td>
</tr>
</tbody>
</table>

Figure 1: Scoring Categories.

**TEXT:**

People have wondered for a long time how their personalities and behaviors are formed. It's not easy to explain why one person is intelligent and another is not, or why one is cooperative and another is competitive.

Social scientists are, of course, extremely interested in these types of questions. They want to explain why we possess certain characteristics and exhibit certain behaviors.

<table>
<thead>
<tr>
<th>EXACT</th>
<th>characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONST</td>
<td>qualities</td>
</tr>
<tr>
<td>ACCEPT</td>
<td>personalities</td>
</tr>
<tr>
<td>SEMA</td>
<td>characteristic, manner, characters</td>
</tr>
<tr>
<td>SYNA</td>
<td>behaviors, thoughts, ideas</td>
</tr>
<tr>
<td>UNACC</td>
<td>behavior, reaction, belief, acts</td>
</tr>
</tbody>
</table>

Figure 2: Example of Scoring Categories.

Thus, the scoring categories: **Exact, Consistent, and Acceptable** signify that the reader was using both syntactic and semantic reading strategies, integrating formal, linguistic, or "language-based" strategies with meaning-based, or "reading" strategies. The category: **Semantically Acceptable** signifies that the reader was using essentially a semantic but not a syntactic reading strategy, ignoring formal, linguistic, or "language-based" strategies and going with meaning-based or "reading" strategies. The category: **Syntactically Acceptable** signifies that the reader was using essentially a syntactic but not a semantic reading strategy, relying on formal, linguistic, or "language-based" strategies and ignoring meaning-based, or "reading" strategies.

Interrater reliabilities for the English raters varied between $r = .94$ for the categories **Semantically Acceptable** and **Syntactically Acceptable** to $r = 1.00$ for **Exact, Unreadable, and Blank**. Interrater reliabilities for the Chinese raters varied between $r = 1.00$ for **Exact** and **Blank**, to $r = .89$ **Consistent**, $r = .88$ **Unacceptable**, $r = .72$ **Acceptable** and **Semantically Acceptable**, $r = .47$ **Syntactically Acceptable** and $r = .34$ **Unreadable**. The relatively low correlation for the **Unreadable** category probably reflects the low number of unreadable items. The relatively low reliabilities
for Syntactically Acceptable may indicate that this category has rather marginal status in Chinese, an isolating language in which "syntactic but not semantic acceptability" has little relevance. That is, the fact that Chinese words are not inflected means that the distinction between what is semantic and what is syntactic is not as clear morphologically as it is in English.

Data Analyses

Because the number of responses (blanks) differed slightly between the Chinese and English clozes, the data were standardized as percentages of responses falling into each of the eight categories. Data were analyzed using non-parametric χ² tests for contingency tables to determine whether the distributions of the responses across the categories differed or not.

RESULTS

Distributions of responses by percentages in each category, for each group of subjects on the Chinese and English clozes are reported in Table 1. To analyze these cloze data for what they might reveal about similarities and differences in, and relationships between, first and second language reading strategies, the data were subjected to two different sets of inferential statistical analyses. First, χ² analyses were performed on Chinese subjects' performance on the Chinese (L1) and English (L2) clozes. In other words, we compared the results shown in the first two columns in Table 1. In these analyses, which we will discuss first, we were examining the strategic behavior of the same subjects in their L1 (Chinese) and in English as their L2. Second, we performed χ² analyses on the English clozes to see what was revealed about the strategies used by L1 (native English speakers) and L2 (native Chinese speakers) in their performance on the same English cloze. In other words, we were comparing the results shown in the second and third columns in Table 1. Finally, although we did not (and cannot) perform any direct statistical analyses comparing the L1 strategic performance by the Chinese native speakers on their native cloze and by the English native speakers on their native cloze, because neither the subjects nor the items are directly comparable, we will make some informal observations about these differences. In these informal observations, we will be comparing the results shown in columns 1 and 3 in Table 1. These informal observations will suggest that performance differences may be related to differences in the two languages.

<table>
<thead>
<tr>
<th>Chinese Subjects (N = 60)</th>
<th>English Subjects (N = 28)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td>English</td>
</tr>
<tr>
<td>Cloze</td>
<td>Cloze</td>
</tr>
<tr>
<td>C-C (L1)</td>
<td>C-E (L2)</td>
</tr>
<tr>
<td>Acceptable</td>
<td>68.9</td>
</tr>
<tr>
<td>Exact</td>
<td>21.4</td>
</tr>
<tr>
<td>Const</td>
<td>23.8</td>
</tr>
<tr>
<td>Accept</td>
<td>23.7</td>
</tr>
<tr>
<td>Sem A</td>
<td>9.1</td>
</tr>
<tr>
<td>Syn A</td>
<td>4.8</td>
</tr>
<tr>
<td>Unacc</td>
<td>12.9</td>
</tr>
<tr>
<td>Blank</td>
<td>3.5</td>
</tr>
<tr>
<td>Unreadable</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table 1: Distributions of Responses by Percentages; Chinese and English Subjects on Chinese and English Clozes.

The first χ² test of the distributions of responses across the eight categories, comparing the percentages from columns 1 and 2 in Table 1, showed these distributions to be statistically significantly different, χ² = 33.93, df = 7, p < .001. A closer examination of Table 1 shows a number of interesting things about these results. First, as would be expected from subjects performing in their first language as compared to a second language in which they do not have native-like competency, subjects performed differently in the proportions of their responses which were Consistent, Acceptable and Unacceptable. Furthermore, Table 1 shows that these differences are in the expected directions, namely higher percentages of Consistent and Acceptable scores on the L1 cloze and higher percentages of Unacceptable scores on the L2 cloze. However, on the Exact responses, while there are differences in the percentages of responses, these differences are not in the direction we might expect. Table 1 shows that subjects achieved higher percentages of Exact responses on the L2 cloze than on the L1 cloze. More will be said about this in connection with the other comparisons we will make. However, for now, it should be noted that from the other Exact mean score in Table 1, the English cloze seems to have lent itself to higher percentages of exact word responses, overall, than the Chinese, whether by native or non-native speakers. Finally, there were sample differences in the percentages of Chinese subjects' Semantically Acceptable and Syntactically Acceptable responses on the English and Chinese clozes. Table 1 shows that, not surprisingly, Semantically Acceptable responses were proportionally higher for
the L1 cloze than the L2 cloze. However, proportionally more Syntactically
Acceptable responses were produced for the L2 English cloze than the L1 Chinese
cloze. This latter result may be surprising, but again, as can be seen in Table 1,
the English cloze seems to have lent itself to production of syntactically acceptable
responses to a greater extent than the Chinese, whether by native or non-native
speakers. Also, as we previously mentioned, the Syntactically Acceptable category
may not be a particularly meaningful one for the Chinese language, and hence for
the Chinese cloze.

In the second set of analyses, the distributions of responses on the English cloze
were examined for the similarities and differences exhibited by the Chinese (L2)
and English (L1) subjects. In other words, the last two columns in Table 1 are being
compared. The resulting $\chi^2 = 27.34, df = 7, p < .001$, indicates that these distributions
were statistically significantly different. Re-examination of Table 1 shows that the
biggest sample differences were in subjects' performance on all categories of
responses except Consistent and Syntactically Acceptable. Table 1 shows, that,
not surprisingly, native English speakers produced greater percentages of Exact
and Acceptable responses, while the Chinese speakers produced greater percentages
of Unacceptable responses. The Chinese subjects produced higher percentages of
Syntactically Acceptable responses than did the native English speakers. Again,
this may be evidence of the Chinese using a "language" or a "syntactic" strategy in
their English reading. Also, interestingly, but not surprisingly, we note that the
differences in the Semantically Acceptable category are in the direction of higher
performance by the Chinese. The Chinese are often able to come up with a
semantically but not syntactically acceptable choice. This may be evidence of the
use by the Chinese of a "reading" strategy as opposed to a "language" strategy in
English as their L2.

DISCUSSION

The central points which emerged from our review of the literature were that: (1)
native and nonnative speakers perform similarly on cloze in a given language, (2)
native and nonnative speakers seem to use the same strategies on cloze in a given
language, (3) readers perform differently in different languages. Our results tend to
support all three of these points. First, the Chinese readers apparently did not use
the same strategies on the Chinese and on the English clozes, and second, the Chinese
and English readers apparently tended to use similar strategies in English, with
some differences due to native versus foreign language.

Our assumption is that the various categories of responses to cloze tests provide a
window on reader strategies, analogous to the way in which miscue analysis of oral
reading provides an on-line indication of oral reading processing. If so, then our
results can be interpreted as showing that the Chinese subjects reading in their L1

and L2 were using different general strategies in performing the task in their L1
and in their L2, or that the task is different in the L1 and in the L2. The higher
proportion of Syntactically Acceptable responses produced by the Chinese subjects
for the L2 cloze than for the L1 cloze suggests that the task in English lent itself to
the use of that formal, linguistic, or "language-based" strategy. This result was
confirmed by the similar performance of native English speakers in the production
of Syntactically Acceptable responses.

When we compared L1 (English) and L2 (Chinese) readers on the English cloze
we noted that the L2 readers tended to produce higher proportions of Semantically
and Syntactically Acceptable responses than the L1 subjects, who, of course, tended
to be able to produce responses which were completely acceptable, both semantically
and syntactically. In the L2 data we believe that we are seeing the use of both
"reading" strategies and "language-based" strategies. We think that Semantically
Acceptable responses may reflect the use of "reading" strategies, while
Syntactically Acceptable responses may reflect the use of "language" or
grammatical strategies. Completely Acceptable responses reflect the efficient
coordination of both reading and language strategies.

Our research also supports that of Anderson (1991), showing that the picture of
reading strategy use is not only a complex one, but also further showing that what
seems to be important is that being a good reader means the ability to use strategies
flexibly, in particular, as required by the linguistic features of the given language in
which the reading occurs.

REFERENCES

Cliffs, NJ: Prentice Hall.

Language Learning, 30, 59-76.


Block, E. (1986). The comprehension strategies of second language readers. TESOL
Quarterly, 20, 463-494.

Publishers.


