Sources of difficulties in cross-cultural communication and ELT: The case of the long-distance “but” in Chinese discourse

Hao Sun
Indiana-Purdue University, Fort Wayne

ABSTRACT

The use of discourse markers has been reported as one of the problems in cross-cultural communication as well as one of the learning difficulties encountered by native speakers of Chinese and other Asian learners. However, sources of such difficulties or “ambiguity” have not been fully investigated or examined. Using elicited performance data produced by native speakers of Chinese and English respectively, this study reports findings on a distinctive long-distance feature of the marker “but” in Chinese discourse and its possible impact on comprehension when such a pattern is used in English discourse. The revelation identifies sources of difficulties in cross-cultural communication as well as informs English language instruction.1

Introduction

Successful communication relies on many factors, one of which involves cohesion—the way in which meanings within the text relate to each other. Such linguistic devices are crucial for signaling relations and interpretation of meaning. Halliday and Hasan (1976) identify four different types of conjunctions in English, among which adversatives constitute an important subcategory. Words such as “but” and “however” belong to adversatives; they both represent a textual meaning of changing direction, or of being contrary to expectation. Examining discourse markers including “because,” “well” and “but,” Schiffrin (1987, p. 9) states that these cohesive devices “are clues used by speakers and hearers to find the meanings which underline surface utterances”.

Research has shown repeatedly that different languages and cultures may have different conventions governing the norms of speech and interaction styles (e.g. Bailey, 1997; Gumperz, 1982; Hymes, 1972; Saville-Troike, 2003; Scollon & Scollon, 1990; Spencer-Oatey, 2004; Sun, 2002, 2004, 2005; Tannen, 1981; Young, 1994). Proposing that the interpretive process is shaped by interpretive conventions, Gumperz (1982) discusses “contextualizing cues,” which include features such as prosodic styles, dialect, code and style switching, and nonverbal features of language. In addition, lexical and syntactic options also serve “contextualizing functions.” These contextualizing cues, as Gumperz (p. 131) points out, “are habitually used and perceived but rarely consciously noted and
almost never talked about directly” for the most part. In order to ensure successful cross-cultural communication, it is necessary to examine discourse conventions in different languages and to uncover differences in our contextualizing cues, including grammatical systems, if there are indeed important differences.

Studies in cross-cultural communication have reported difficulties in communication and comprehension in interactions involving learners of English with regard to the use of conjunctions or discourse signals. Problems in communication between native speakers of English and Chinese learners of English were discussed in, for example, Scollon (1993), Tyler, Jeferries and Davies (1988) and Young (1982). One of the features of the Korean and Chinese graduate students’ discourse that was confusing to their American audience found by Tyler, Jeferries and Davies (1988) is the use of discourse markers “however”, “but”, and “and”. Scollon (1993) examined conversational data between a native speaker of Chinese and a native speaker of American English. Scollon illustrated that the Chinese learner’s English discourse seemed to be confusing to the native speaker of English, as the marker “but” was used by the English language learner in functions that differed markedly from the ways it would be used by native speakers, namely, for additive, adversative, causal, and repetitive functions instead of only the adversative function, thus the term “accumulative ambiguity” (Scollon, 1993).

In addition, learning difficulties have been observed and reported. Examining students’ writing in English, Field and Yip (1992) reported that second language learners of English, particularly native speakers of Cantonese, used conjunctions more frequently in writing than native speakers, and these conjunctions were used more frequently in sentence-initial positions. An important issue that needs to be examined is the cause(s) of such phenomena, as Scollon (1993) rightly pointed out, so that relevant issues can be addressed effectively pedagogically. However, sources of such difficulties with the use of the marker “but” have not been fully examined, which underlines the need for the current study.

Examining the use of “but” as a discourse marker in Chinese to investigate whether there is any major difference from the way the marker is used in English2, Sun (2005) reported findings that one particular feature of “but” in Chinese discourse constitutes a pattern saliently different from the ways in which “but” functions in English. Specifically, the discourse marker “but” in Chinese may occur in a syntactic position which is separated from its target clause3, a phenomenon described in the current article as a long-distance “but”. The

---

1 An earlier version of this article was presented at the 2005 World Congress of Applied Linguistics in Madison, WI. This study was supported by a Summer Faculty Grant from Indiana-Purdue University Fort Wayne. I thank the anonymous reviewers of this journal for their helpful comments and suggestions. I am most grateful to Yili Li, who kindly provided very valuable insights and useful feedback on earlier drafts of this article as well as important assistance with the statistical analysis of this study. I also greatly appreciate the support of Avon Crismore, Baofang Fan, Lidan Lin, Ying Sun and Sidury Christiansen, who graciously assisted me in my data collection. Most importantly, this project would not have been possible without the help and efforts of the volunteers who participated in this study. Any errors in this article, however, remain my sole responsibility.

2 The term discourse marker will be used throughout this article to refer to the word “but” instead of the word conjunction or other possible terms.

3 I am using the term “target clause” to refer to the clause that is supposed to be syntactically introduced and semantically modified by “but,” following the English discourse convention.
following is an example of a long-distance “but” (hereafter LDB for short). We first present the original segment of the Chinese text, followed by a translation of the Chinese text with the structural position of the “but” intact. The LDB under discussion is in bold.

我开始工作的时候，我老板给我一个项目，就是做一个买家。...但是大家知道，98年的时候没有很多的供应商在用互联网，所以我没有办法用互联网搜索。所以我就到了很多的贸易公司，看他们的目录。但是因为他们五点就关门了，或者他们的目录没有他们最新的产品，或者他们的样品师找不到一个合适的产品。所以我就总是找不到合适的产品。4(阿里巴巴.中国)

When I started my job, my boss asked me to look for certain products for our company. As you all know, there were not many suppliers using internet in 1998 so I had to visit many companies and looked at their catalogues. But because these companies closed at 5 o’clock, or that theircatalogues didn’t include their newest products, or that their sales personnel couldn’t find the appropriate products for me, I was never able to find the appropriate products.

As we can see, the translated English version maintaining the original Chinese syntactic structure of the “but” renders a long-distance “but” (in bold), which is set apart from its target clause “I was never able to find the appropriate products”, with three adverbial clauses of reason inserted between the LDB and the target clause. This is due to the grammatical structure of the Chinese language, which requires that adverbial clauses precede the main clause syntactically5 (e.g. Kirkpatrick, 1993; Wang, 1999). In such cases, the discourse marker “but” serves its adversative function in a “remote” manner, connecting meanings crossing sentential or clausal boundaries. In fact, the target clause may not occur until after all the relevant adverbial clauses have been introduced, which may render comprehension difficult for speakers unfamiliar with such a discourse and syntactic structure, including native speakers of English. Such a distributional position is saliently different from how the marker “but” is used in English.6

Due to the syntactic features of Chinese mentioned above, an LDB may also involve the co-occurrence of multiple “conjunctions” such as “but because…,” “but although…,” or “but if…,” if we analyze such structures in English syntactic terms. Not all these words, however, are considered conjunctions in Chinese grammar (e.g. Chao, 1968; Li & Thompson, 1981). There is also an “enveloping” feature in Chinese (Kirkpatrick, 1993), which is described as a “pregnant unit.”7 Native speakers of English find such a phenomenon confusing and difficult to

4 A more appropriate way to translate this paragraph, one might argue, is to use “however” instead of “but.”
5 This is not possible in Chinese, nevertheless, since Chinese does not differentiate between “but” and “however.” There is, however, some difference between the spoken and written language in this regard. For example, the word 然而 is only used for the written language.
6 It needs to be made clear that this is only one of the ways in which the discourse marker “but” is used in Chinese and it is certainly not the most frequently encountered pattern.
7 Kirkpatrick (1993) discusses a “because-therefore” unit in which lower level units are contained.
process. Although the phenomenon of LDB is common in Chinese discourse, it has not received adequate attention in English language teaching, nor has the issue been examined fully in Chinese linguistics to uncover patterns of use.

Observed patterns in Chinese as discussed in Sun (2005) exhibit some similarities with previously reported findings about learners’ interlanguage data (e.g. Scollon 1993; Tyler, Jeferries and Davies, 1988), indicating a possible effect of native language transfer at a level beyond sentences. Learners’ difficulties do not seem to all derive from pedagogical influences, as there is similarity between the data from the native language and interlanguage, one of the criteria for measuring transfer discussed in Jarvis (2000).

Pedagogically, it is also important to emphasize the need to approach language skills development from a discourse perspective. In fact, many researchers have repeatedly stressed the need to incorporate a discourse perspective for both grammatical analysis and language teaching (e.g. Celce-Murcia & Larsen-Freeman, 1999; Celce-Murcia & Olshtain, 2000; Lock, 1996; McCarthy, 1991) as well as the inadequacy of sentence grammar in dealing with conjunctions (e.g. Schiffrin, 1987; Stubbs 1983). Particularly in the case of adversatives, the traditional way of grammar instruction cannot uncover learners’ difficulties completely, as some issues may not exhibit themselves at the sentence level. The hidden dimension of adversatives beyond sentence boundaries in the Chinese language may have blurred learners’ perceptions and awareness of the differences between Chinese and English adversative use, which in turn has increased difficulties in learners’ use of adversatives in English. It is only when language use and learner language use are examined at a discourse level in addition to the sentence level will issues of discourse markers be uncovered fully and issues of learning difficulties be addressed effectively.

The present study

Purpose

Pedagogically motivated in an attempt to uncover sources of learning difficulties in English language learning and teaching as well as to inform cross-cultural communication, this article intends to address the following research questions:

1. If the discourse marker “but” is used in English discourse in a long-distance manner, does it constitute a cause of difficulty for comprehension? If so, how does it affect different native language groups?
2. Does the distance “traveled” by an LDB add to comprehension difficulties?
3. What is the effect of a “focused change” on comprehension performed by different native language groups? The term “focused change” refers to a replacement of an LDB in the original version of the text with an adjacent “but” (in a reconstructed version), which is syntactically positioned next to the target clause it is supposed to introduce.
**Participants**

Two different groups of volunteers participated in this study: thirteen native speakers of Chinese and seventeen native speakers of American English, a combined total of thirty individual participants. The Chinese participants, who were all residing in Shanghai, China at the time of this study, are college graduates with either a master’s or bachelor’s degree. As several participants obtained their master’s degrees from one of the most prestigious universities in China, which requires high scores for admission, these participants’ Chinese language skills were assumed to be fairly high. In addition, all the participants learned English as a foreign language for at least ten years. While most of them work in the areas of finance, accounting, or banking, some of them in fact read and write English on a daily basis, as is required by their jobs. No other detailed information about the participants was available as the participants were volunteers contacted through the help of other volunteers for this study.

The American participants were all college undergraduate students at a university in the Midwest of the United States at the time of data collection for this study. A minimum of composite score of 950 on SAT I was required for admission at this university, based on information from published sources (e.g. Barron, 2005). They were enrolled in either a basic writing or an introductory literature course. Some of the participants were attending college as traditional students while others were working adults, who were taking college courses and working part time or full time at the same time.

**Data Collection**

The data of the present study consist of elicited performance of listening comprehension after the participants completed listening tasks for six short passages recorded (See Appendix I). All the six paragraphs were in Chinese originally, obtained from the Internet and then translated into English for this study. The most important criterion of selection of the texts was the presence of an LDB. Other considerations include topic and vocabulary comprehensibility. Data collected include news reports, interviews, commentaries, and a workshop talk. The topics of the selected paragraphs range from grocery shopping, jobs, and natural disaster to economics and politics. The total time of the recording of the six paragraphs is 3.73 minutes (224 seconds). The average number of words is approximately 68 per paragraph. While the shorter paragraphs are less than half a minute in length, the longest paragraph is 52 seconds, with the average length of each paragraph being 37 seconds. The recording was made by the researcher, who was a faculty member at a university in the US, teaching a variety of graduate and undergraduate college courses.

Participants were instructed to listen to each paragraph once, and then complete three multiple choice questions for each paragraph, unless instructions specified otherwise. In the case of the American participants, the recording was played while they were physically present in a classroom and the participants’

---

8 We were keenly aware of the relatively small number of participants in this study. Therefore, any observations and generalizations based on our statistical analysis are tentative.
responses were recorded in writing. For the Chinese group, as the participants were not students on campus or residents living in the same locale, the recording was made available online in the form of sound files so that the Chinese participants were able to access the recording online. No difference in the quality of the sound was observed between the two types of input. In the case of the use of sound files, participants’ performance and responses were recorded in written form as well and sent to the researcher electronically by a volunteer assistant.

Research design

The six paragraphs were divided into two subcategories: Original Structure (OS) and Reconstructed Structure (RS). The original structure refers to paragraphs that contain no focused change from their original versions except for minor modifications where necessary (such as changing the word “famine” to “lack of food” to ensure understanding by all the participants), while the reconstructed structure involves a focused change by removing the discourse marker “but” from a long-distance position. Specifically, whereas the original version may contain the word “but” in a position that is one, two, or three clauses away from the target clause it modifies, the reconstructed version places the word “but” in an adjacent position next to the target clause, which is similar to the syntactic behavior of “but” in English (An example of both the original and reconstructed structures are included in Appendix II). While P1 (P1 stands for Paragraph 1 hereafter, and it is the same for other paragraphs), P2, and P3 belong to the OS category, P4, P5, and P6 belong to the RS category (See Appendix III for all the English texts used for the listening tasks as well as the original Chinese texts).

In addition to examining the effect of the absence versus presence of an LDB on comprehension, the distance between an LDB and its target clause was another issue set out to be examined, namely, whether all LDBs exert the same effect, as measured by both comprehension accuracy and participants’ ranking of difficulty level for each paragraph in the OS category.

To address the question of the distance issue, the three paragraphs selected in the OS category also involve some differences, although all of them contain at least one LDB respectively. The differences lie in both the length of the texts and the structures. In the OS category, P1 is much shorter with 44 words, and P2 has 60 words. P3 is the longest and has 84 words, almost twice as many words as P1 does. Structurally, in the case of P1, the LDB is separated from its target clause by two clauses with no other syntactic complications. On the other hand, in both P2 and P3, conjunctions other than “and” (such as “therefore” and “so”) are also present in between the LDB and its target clause, and the separation thus involves a much longer section as well as a more complex structure in each case. Of the three paragraphs in the OS category, P3 is the most complex, with three different clauses separating the LDB and the target clause. Paragraphs with these differences were selected so that possible effect of LDBS in different positions could be examined.

In the RS category, three original Chinese paragraphs with similar lengths compared with those in the OS category were selected. As a result, P5 is similar
to P1 with only 47 words, while P6 is close to the length of P2, with 71 words. P4 is the longest in the RS category, with a total of 83 words, and is similar to P3, which has 84 words in length.

Comparisons were made in terms of comprehension accuracy, perceived levels of difficulty, and identified area(s) of difficulty. For each of the six paragraphs, participants answered three multiple choice questions addressing the three issues respectively. Participants needed to answer all three questions for each paragraph unless their answer for question two (which asked the participants to rank the difficulty level of the paragraphs for comprehension) was choice “A.” If participants selected A for question number two, it indicated that they considered the given paragraph “easy,” which rendered the completion of question three (which asked the participants to identify any source of difficulty) unnecessary. In such cases, participants ended up answering only two questions instead of three for a given paragraph.

Of the three questions for each paragraph, question one is focused on the content of the paragraph, which is intended to assess participants’ general comprehension of the message. If a participant’s comprehension is poor across all or most of the six paragraphs, his/her English language proficiency will be reflected in question number one across all the paragraphs (which has been the case with a few Chinese participants). Such information will be useful in examining, analyzing, and interpreting comprehension accuracy for a given group as well as for individual participants’ evaluation.

Question two addresses each participant’s assessment of the ease with which he/she processed and interpreted the meaning of the paragraph. The question reads “Is the speaker’s meaning easy to follow?” A participant thus needed to select from among the three options available: Choice A stands for the fact that the paragraph is easy, Choice B states that the paragraph is “Somewhat easy,” and Choice C indicates that the paragraph is “Not easy.”

The third multiple choice question asks the participants to identify a particular source (or sources) which has led to perceived difficulties. There are four choices provided for question three: (1) unfamiliarity with the subject, (2) difficult vocabulary choices, (3) problems in sentence structures, and (4) the order in which ideas are presented. Choice 4 is intended to identify the long-distance “but” as a cause of comprehension difficulty. Participants can identify either one or more than one area as source(s) of difficulty for comprehension. Appendix 1 provides an example of the multiple choice questions.

Results

Findings are presented below in the form of answers to the research questions set out at the beginning of this article.

Q. 1. Does the use of an LDB in English discourse constitute a cause of difficulty for comprehension? If so, how does it affect comprehension for different language groups?
In this section, we address the issue whether an LDB constitutes a source of difficulty for comprehension by using two major measurements: comprehension accuracy, a measurement of performance, and identification of areas of difficulty in comprehension, a measurement of perception. We first examine the result of comprehension accuracy, to be followed by an examination of participants’ identification of “order” (of ideas) as a/the problematic area. Table 1 below shows comprehension accuracy of both groups across the six paragraphs. Figure 1, on the other hand, compares the accuracy percentage of the two groups’ performance.

### Table 1
Comprehensive Accuracy (percentage)

<table>
<thead>
<tr>
<th>Group</th>
<th>Original Order (OS)</th>
<th>Reconstructed Order (RS)</th>
<th>Group Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Paragraph 1</td>
<td>Paragraph 2</td>
<td>Paragraph 3</td>
</tr>
<tr>
<td>No.</td>
<td>%</td>
<td>No.</td>
<td>%</td>
</tr>
<tr>
<td>Chinese (13)***</td>
<td>9</td>
<td>69</td>
<td>9</td>
</tr>
<tr>
<td>American (17)</td>
<td>16</td>
<td>94</td>
<td>16</td>
</tr>
<tr>
<td>Paragraph Average</td>
<td>81.5</td>
<td>81.5</td>
<td>70.0</td>
</tr>
</tbody>
</table>

*The number refers to the total number of participants whose answers are correct.
**The percentage refers to the accuracy percentage of the group for a given paragraph.
***There are two individuals in this group who, out of six paragraphs, only had one answer correct for content comprehension, indicating very low level of listening comprehension skills.

### Figure 1
Comparison of group accuracy for comprehension
(a) Comprehension accuracy

1. Comprehension accuracy for different paragraphs

Table 1 shows each group’s average accuracy percentage for each paragraph. As explained earlier, P1, P2, and P3 all belong to the OS category while P4, P5, and P6 all fall under the RS category. A comparison of the average accuracy percentages of each paragraph by the two groups shows that P3 (OS) has the lowest average of 70%, indicating it to be the most difficult text to comprehend, followed by 79.5% in P6 (RS), leaving P4 with 89.5%, the highest percentage point for accuracy, indicating P4 to be the easiest for comprehension.

To determine whether there was any statistically significant difference in comprehension accuracy between different categories of text, a Mann-Whitney one-tailed test (Hatch & Lazaraton, 1991) was run to compare the accuracy percentages between the OS and the RS paragraphs. The results showed significant differences between the OS and the RS categories within both the Chinese and the American group (Chinese group: $z = -2.443$, $p < .01$; American group: $z = -1.597$, $p < .05$).

2. Comprehension accuracy by different groups

As shown in Table 1, the American group shows a higher percentage for accuracy, with an average of 84% across all the six paragraphs. The Chinese group, comparatively speaking, does not seem to be too far behind, with its 76.8% for comprehension, given that the participants’ comprehension was based on their performance in their second (or 3rd) language.

To determine whether there was any statistically significant difference in comprehension accuracy between the two native language groups in terms of OS and RS categories, a Mann-Whitney one-tailed test (Hatch & Lazaraton, 1991) was also run to compare the accuracy percentages between the Chinese and the American groups for both the OS and the RS paragraphs. The results did not indicate statistically significant differences in terms of either the OS or RS categories. (OS: $z = -1.049$; RS: $z = - .544$).

3. A marked case of difficulty

Although the American group has a fairly high percentage of comprehension for P1, P2, and P4, it shows significant decrease in comprehension accuracy for P3, a paragraph which features the use of an LDB with three clauses separating the marker and its target clause. Interestingly, the Chinese group does not obtain a high percentage of comprehension for P3 either, although it is the same case with P1 and P2. However, in comparison, its performance pattern shows a very different picture compared with the American group with regard to the possible impact of the LDB across the different paragraphs. Specifically, the Chinese group features a steady 69% distribution across P1, P2, and P3, as shown in Table 1, which indicates that there is no change in accuracy performance across the three paragraphs.
(b) Identification of area(s) of difficulty

The second type of evidence with regard to the issue whether an LDB constitutes any difficulties for comprehension relies on participants’ identification of perceived areas of difficulties in comprehension, as shown in Table 2 below.

Table 2
Identification of “order” (of ideas) as a/the problematic area

<table>
<thead>
<tr>
<th>Group</th>
<th>OS (Original Structure)</th>
<th></th>
<th>RS (Reconstructed Structure)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P1</td>
<td>P2</td>
<td>P3</td>
<td>No.*</td>
</tr>
<tr>
<td>Chinese</td>
<td>2</td>
<td>15</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>American</td>
<td>1</td>
<td>5.9</td>
<td>4</td>
<td>24</td>
</tr>
</tbody>
</table>

Number*: The number of participants who identified “order of ideas” as a/the problem area.

Figure 2
Identification of “order” as a/the problematic area (percentage)

Of the two groups, the American group clearly indicated that the order of ideas was a source of difficulty for comprehension, as is reflected in the increased percentages across the three paragraphs in the OS category. Specifically, the American group shows a steady rise from 5.9% in P1 to 24% in P2, and then to 65% in P3, exhibiting a clear pattern of increase in terms of participants’ identification of the “order” (of ideas) as a/the problematic area for comprehension. In fact, while approximately one fourth (24%) of the American participants in this group considered the order of ideas a major contributing factor for comprehension difficulty in P2, the percentage grew to a majority of almost two thirds (65%) in P3.
The Chinese group showed a very different picture with regard to participants' perceptions about the "order" issue. As is shown in Table 2, although a few participants identified "order" as a problem for comprehension, the group held the same 15% percent across the three OS paragraphs, showing a consistent pattern with its performance on accuracy (with a steady 69% across all three paragraphs).

Q. 2. Does the distance “traveled” by an LDB add to comprehension difficulties?

As is shown in Table 1 and Table 2, the three paragraphs in the OS category did not result in similar percentages for comprehension accuracy for the American group, and P3 is the paragraph with the lowest comprehension accuracy percentage. In addition, it is also clear that native speakers of American English did not consider the three paragraphs in the OS category as exhibiting the same levels of difficulty, as can be seen from Figure 3 below.

Specifically, the American group obtained a score of 1.06 for P1, perceiving P1 to be the easiest. In contrast, its ranking of P3 scored 1.94, the highest of all the scores, with 1.76 for P2 in between. The Chinese group, however, did not exhibit the same level of ranking although there was a slight increase across the three paragraphs, with 1.46 in P1, 1.62 in P2, and 1.85 in P3. In addition, the Chinese group showed no change in comprehension accuracy in the OS category, as discussed earlier.
Q. 3. What is the effect of a “focused change,” which involves the replacement of an LDB, on comprehension performed by different native language groups?

In particular, P3 and P4 constitute a useful comparison since both P3 and the original P4 involve a three-clause separation between an LDB and the target clause. As a result of the focused change in P4 (a reconstructed version), the presence of an LDB in P3 and the absence of it in P4 constitute a major difference between the two paragraphs, which renders the comparison between the two meaningful, given that the number of clauses separated by the LDBs in both P3 and the original P4 Chinese version are the same.

Compared with P3, comprehension accuracy for P4 increased for both groups, as can be seen in both Table 1 and Figure 1. The Chinese group obtained a smaller increase of sixteen percent, while the American group showed an increase of twenty-three percent. If we examine the average accuracy percentages of both groups, the percentage increased from 70% in P3 (OS) to 89.5% in P4 (RS), a marked difference and contrast. In brief, P4 clearly indicates the effect of the absence of an LDB. However, the picture is not as clear in the case of P5 and P6.

In the case of P5, there is only one clause between the LDB and its target clause; P5 is therefore more similar to P1 than P3 in terms of the number of clauses between the LDB and its target clause. It is also more similar to P2 with regard to text length. It was therefore expected that P5, similar to P4 and P6, would be fairly easy for comprehension since the LDB was removed in the reconstructed version.

As the result indicates, however, P5 did not turn out to be as easy as expected for the American group, with its 71% for comprehension accuracy. In fact, the accuracy percentage for P5 is the same as the percentage for P3 for the American group. In comparison, the Chinese group performed fairly well on P5 with 92% for accuracy, the highest percentage for the group across different paragraphs, as can be seen in Table 1 and Figure 1. It seems that other issues must have contributed to the American group’s lower performance on accuracy in P5, which are addressed in the Discussion section below.

In the case of P6, the performance on accuracy by the Americans and the Chinese almost seems to have reversed. While the American group improved in P6 for comprehension and obtained 82% (compared with their performance of comprehension in P5), the Chinese group decreased its accuracy in P6 from their percentage for P5, obtaining only 77% for comprehension. Possible factors that have contributed to the performance of both groups are explained in the Discussion sections as well.

Discussion

Difficulties for comprehension

Analysis based on examination of both comprehension accuracy and identification of difficult areas tentatively suggests that the presence of an LDB
has added difficulties in comprehension for participants in some cases, as is mostly clearly indicated in the case of P3 for the American group, as can be seen in the group’s lower comprehension percentage and its high score in identification of the text as involving problems in the order of ideas. Specifically, the American group has an accuracy rate of 94% for both P1 & P2 in the OS category, showing a relatively high rate of comprehension first. However, its comprehension accuracy for P3 drops 23% in P3, achieving an accuracy percentage at only 71%. What is more important is the American group’s significant increase to 65% for identification of the “order” (of ideas) as a difficult area in P3, lending support to the observation above with regard to the difficulties encountered as well as perceived in P3 by the American participants. The fact that P3 involves a three-clause LDB separation is likely to have constituted a main cause of such difficulties.

The American group’s increased percentages across three paragraphs in identification of the order of ideas as a difficult area for comprehension seem to be consistent with the group’s demonstrated decrease in comprehension from P1 (94%) to P3 (74%), suggesting that the perceived difficulties in comprehension can partly account for decreased percentage in accuracy of comprehension in P3. Such an observation is further supported by the American group’s scores for difficult level ranking: this group shows a low score of 1.06 for P1 while its ranking of P3 reaches the highest (1.94) score, as shown in Figure 3.

The combined objective and subjective measurements provide evidence that the presence of an LDB in P3 constituted a source of difficulty for native speakers of American English in comprehension of the passage. However, if we merely rely on comprehension accuracy percentages between the OS and RS categories without considering either the participants’ identification of problematic areas or their difficulty level ranking, findings seem to suggest otherwise. Specifically, the American participants seemed to have performed better for comprehension in the OS category than they did in the RS category, as the percentages shown in Table 1; the result of the Mann-Whitney test also indicates that the difference is significant (P<0.05), which is inconsistent with the findings obtained from combined measurements for P3 rendered above.

There are several variable issues that might have contributed to the results of the inconsistent findings: first and foremost, the number of participants in this study constitutes one of the major issues. Given that there was a total of thirty participants between the two groups, the small number of subject pool did make statistical analysis and interpretation difficult. Other major issues include the differences between the texts in the OS and RS categories, which will be addressed in the sections to follow. Given that the comparison between P3 and P4 is supported by both objective and subjective measurements, and that the results of the comparison between the OS and RS categories do not seem to yield a consistent and meaningful pattern, the rest of our discussion in this section on difficulties of comprehension will be focused on the results obtained from P3.

Why would an LDB in P3 cause more difficulties for the American participants? Following English discourse and syntactic conventions, an interlocutor typically encounters and processes the target clause following an adversative marker “but.” In the case of an LDB, however, Americans might feel the “adversative” meaning is somewhat “floating,” and its cohesive target may
be difficult to identify. One might wonder: To what is “but” in reference? The meaning of the “but” in such cases might not be easy to follow for an audience who may not be familiar with such use of the discourse marker.

Such difficulties therefore may derive from differences in grammatical convention and discourse expectations or “contextualization cues” (Gumperz, 1982), real time processing constraints, and short term memory capacity (Scollon, 1993). Specifically, in cases involving an LDB, while processing all the incoming information received, an interlocutor also needs to connect the adversative meaning perceived to relevant parts of the ensuing discourse, which might only occur until after several seemingly remotely related clauses have been presented. The supposedly familiar discourse signal may seem to have been misplaced or misused in such cases, and may result in communication that is difficult to process or comprehend.

The data from the Chinese group exhibits a different pattern. For comprehension accuracy in the OS category, the Chinese group maintained 69% across all three paragraphs. In addition, this group’s perception percentage remained 15% across all the three paragraphs with no variation. Such a pattern seems to indicate that regardless of the differences among P1, P2, and P3, the presence of an LDB didn’t constitute as much a challenge for the Chinese participants as it did for the American group. The fact that comprehension accuracy for the Chinese group in the OS category is much lower than the American group is not unexpected, however, given that English is a second (or third) language for the participants, not to mention the fact that a few individuals in this group consistently produced very low accuracy percentages for comprehension.

For the Chinese group, however, there is some improvement in the RS category over the OS category in accuracy percentages, with 85%, 92%, and 77% for P4, P5, and P6 respectively, compared with the 69% performance across all the paragraphs in the OS category. As a result, the overall accuracy percentage for the RS category is higher than that in the OS category. One possible interpretation for such a phenomenon is that comprehension tasks in the RS categories were also easier for the Chinese participants. In addition, other factors (to be addressed later) may also have contributed to the results.

Examining the perception data presented in Table 2, we noted that only a few of the participants in the Chinese group identified “order” as a problematic area, and only two participants out of the thirteen identified “order of ideas” as a problematic area more than once. Each of these two participants, in addition, only correctly answered 17% of the content questions across the six paragraphs respectively, which indicates a very low level of their English skills in listening. Given this is the case, their identification of the “order of ideas” needs to be interpreted with caution. It is also interesting to observe that the group’s overall ranking of “order of ideas” as a problematic area remains unchanged across three paragraphs for the OS category, which is certainly not the case with the American group. In short, it seems that the marked difficulty perceived by the American group in P3 was not perceived the same way by the Chinese group.
The distance issue

The results of this study also indicate that not all texts involving an LDB are necessarily more difficult. It seems that the farther the discourse marker "but" is separated from its target clause, the more difficulties such a structure may add to processing and comprehension. This observation is supported by, firstly, the American group’s sharp decrease in comprehension accuracy in P3, as compared with their higher accuracy percentages in P1 and P2. Secondly, this group’s increased percentages of recognition of difficulties across the three paragraphs in the OS category, as can be seen in Figure 3, provide further evidence of participants’ own perceptions about difficulties in comprehension. While P1 involves the use of an LDB, the distance between the marker and its target clause is not as extreme as P3, nor is its structure as complex. While P2 involves two clauses between the LDB and its target clause, P3 involves three clauses inserted separating the LDB and its target clause, rendering it the most difficult of the three paragraphs in the OS category. The lower percentage of comprehension in P3 can thus be accounted for partly by the increased distance between the marker and its target clause.

Another factor that may have added to comprehension difficulties is the co-occurrence of additional “conjunctions” or “conjunctive adverbs” such as “therefore” and “so” in addition to the LDB; the clearest case is manifested in P3, with the use of “because,” “and,” “so,” and “therefore” occurring in the three clauses in addition to the use of the marker “but.”

Comparisons also indicate that the two different language groups did indicate different difficulty levels with regard to the occurrence of an LDB. Native speakers of American English have shown a much higher percentage in their identification of “order” of ideas as a/the problem area in two out of the three original paragraphs (See Figure 2). On the other hand, the Chinese group did not seem to experience any more difficulties in comprehension in P3 than in P1 or P2, nor did they indicate such differences in their perception as reflected in their judgment, which is in sharp contrast with the American group’s perception and performance. Given such findings, we can tentatively state that native speakers of Chinese did not perceive the presence of an LDB to have caused as much difficulty in comprehension as the American group, based on very limited evidence. Likewise, their comprehension performance within the OS category did not seem to show variation due to the presence of LDBs occurring in different positions.

The effect of focused change

Focused change involves the reconstruction of paragraphs with the presence of an LDB originally, which are the cases with all the paragraphs in the RS category. P3 and P4 offer a most compelling case due to their similar paragraph length and the number of clauses in between the LDB and the target clause. A comparison of comprehension accuracy shows that both groups gained in P4 compared with their percentages in P3. The American group showed a 23% jump from 71% in P3 to 94% in P4 for comprehension. In addition, the perceived level of difficulties for the American group decreased dramatically to 24% for P4 instead of 65% in
the case of P3. The Chinese group, on the other hand, gained 16% for comprehension, moving from 69% in P3 to 85% in P4. Chinese participants’ decreased perception of 7.7% (as an area of difficulty in P4), compared with their percentage of 15% in P3, seems to be congruent with the increase in comprehension in P4, although the difference in either case is far from being similar to the case with the American group.

Focused change involves P5 and P6 as well, but the results present blurred pictures and challenges for analysis and interpretation. Although it is assumed that P4, P5 and P6 will be less difficult than they would be in their original version, P5 and P6 did not lead to higher percentages for comprehension or lower levels of perceived difficulties by both groups. Based on the data and reexamination of the two paragraphs, we reasoned that other factors might have been involved in P5 and P6 respectively. What is clear is that in both cases, the mere absence of an LDB didn’t make the two paragraphs necessarily easy pieces to comprehend for all participants, as explained below.

In the case of P5, the Chinese group obtained high percentage points in comprehension accuracy. In contrast, P5 was not an easy paragraph for the American group. Part of the reason for the lower comprehension percentage of the American group might be the participants’ unfamiliarity with the subject and context: The topic of this paragraph is about the Chinese money market, and the context involves the financial policies after China joined the World Trade Organization. While the participants’ identification of difficult area did not seem to suggest a strong bias in favor of the Chinese participants, the Chinese group performed relatively well for comprehension. The difference in the participants’ background knowledge and their schemata may have been an important factor for such observed differences, which was not anticipated in the text selection process, mostly because structurally this text was not deemed difficult or complex for comprehension. Furthermore, the omission of some words in the original text in an attempt to remove any identifiable sources of the text may have rendered the text more difficult to process as a whole for the American group, while the issue of the “money market” may have been more familiar to the Chinese participants. In any case, comprehension difficulties in P5 (if they did occur for some participants) couldn’t have resulted from the presence of an LDB since it doesn’t contain an LDB. Issues such as context familiarity, on the other hand, may have largely contributed to difficulties encountered or perceived in comprehension.

P6 is another interesting case for both the current study and for future studies of research on translation. Its reconstruction may have resulted in some new problems for comprehension for the Chinese group, it seems. As a result of the change, the syntactic relationship between the two clauses may have been less clear to the participants in spite of the absence of an LDB, and consequently the interpretation of the meaning may be more difficult. As a matter of fact, with hindsight, syntactically, P6 is a hybrid of some sort in the sense that it is not following structural patterns of spoken English, nor is it actually following Chinese syntactic patterns completely, since the removal of an LDB has effected necessarily other syntactic modifications as well. This is perhaps why P6 ended up with
fairly low comprehension percentages for both groups, with 77% for the Chinese group and 82% for the Americans, rendering it the second most difficult text for comprehension in spite of its being a reconstructed text. The percentages suggest that the removal of an LDB may have actually created other unexpected problems for comprehension, an extremely interesting area for further investigation.

Other issues

It was also found that the presence or absence of an LDB was not the only factor affecting comprehension of the listening tasks. Other factors may have affected the performance of participants as well.

As mentioned earlier, the English language proficiency of the participants in the Chinese group is an important factor in this study. In fact, there were two participants who produced only one answer correctly out of the six questions on content comprehension respectively, and there was one participant who had only two correct answers out of the six content questions. Given the mixed level of proficiency skills, our findings and statistical analysis need to be interpreted with caution, whether it is the objective measurement of participants’ comprehension or their subjective assessment of their perceptions. Although the recruiting of the volunteers did include the consideration of the proficiency level, it was difficult to control the participants’ proficiency. That was primarily because it was volunteer based recruiting, and the volunteers were contacted through the help of other volunteers.

Other variables not focused on in this study may have played a part in the performance as well such as the issue of translation effect (as mentioned in the case of P6). It is also possible that some participants were more familiar with some topics (as explained in the case of P5) and less so with other topics. Familiarity, knowledge, and discourse performance are all related, especially for second language learners, an issue discussed in Zuengler (1989). The wide range of genres from which the data were drawn also constitute a variable: there was discourse data from interviews and a workshop talk as well as data from commentaries and news reports, which span across both the spoken and written language. Although no participants were assumed to be particularly more familiar with a certain type of genre, the different types of text used may have added one more variable in the blurred picture of the findings.

Pedagogical implications

Discourse markers are important as navigation signals for communication and comprehension. Pedagogically, it is necessary to raise instructors’ awareness of the significance of discourse markers in language instruction and their own understanding of such uses based on examination of discourse data. In fact, the role of discourse data cannot be overstated; the observed patterns mentioned in this article, for example, are unlikely to surface or reveal themselves, if we only examine sentence-based grammar or read about traditional formal analysis of a given language. Issues with discourse markers are more likely to be uncovered
and addressed when we utilize discourse or corpus data to examine and discuss language use at a level beyond sentences. It is also necessary for instructors to examine pedagogical materials to ensure that the coverage and explanations of the target language are accurate and adequate, given that there are often differences and seemingly similar aspects between the native language and the target language such as the case discussed in this article.

We also need to help our learners develop keen awareness of patterns and conventions of discourse markers in English as well as the difference between their native language and English in the distribution patterns of discourse markers and their use in discourse. Such awareness, reinforced by focused, interactive, and meaningful practice, will enable learners to express themselves most effectively, accurately, and successfully in cross-cultural communication. This is particularly important for students aiming to engage in professional, business, and cross-cultural communication or for advanced learners who need highly proficient skills in verbal interaction.

Specifically, it would be beneficial to encourage students, if their language proficiency levels and learning skills are appropriate, to handle tasks such as discovery learning or even collect corpus data to discover patterns with regard to discourse markers and conjunctions in English, either in groups or individually. Such discovery learning will be particularly useful for more advanced learners. Learners can also be provided with discourse data or some observed patterns to enhance their awareness. Many current exercises for conjunctions only provide practice at the sentence level (e.g. multiple choice questions, error correction, blank filling, paraphrase using cues provided), which may not uncover some patterns of use or learner difficulties at the discourse level. Communicative and interactive activities or tasks beyond the sentence level, which require learners to be engaged in language use for real purposes, are more useful and meaningful, provided that the lower level language needs are addressed as well.

**Conclusion**

This study demonstrates that the use of discourse markers such as “but” has considerable impact on comprehension. In particular, a long-distance “but,” one of the features of the Chinese language, was examined; its possible impact on comprehension in English discourse, if transferred, was measured. The use of such a discourse marker has not hitherto been given due attention in either Chinese linguistics or English language teaching for learners whose native language is Chinese.

The LDB feature is shown, based on limited evidence, to have resulted in difficulties in comprehension in some text for the American participants in particular, who were not familiar with such use of a discourse marker. The observation was corroborated by the American participants’ identification of the text as involving problems in the order of ideas. In addition, the findings suggest that the distance between the LDB and its target clause did seem to make some difference in comprehension, but the observation was again based on limited evidence. Focused change was shown to be effective in one case, while results
from other cases were inconclusive.

Results from the Chinese group also presented a mixed picture: on the one hand, there is evidence suggesting that the presence of the LDB did not constitute as much difficulty for the Chinese group; on the other hand, the Chinese group also produced higher accuracy percentages in the reconstructed paragraphs than the paragraphs in the OS categories with the presence of an LDB, which may suggest that texts involving LDBs might also be more difficult to comprehend for native speakers of Chinese. As other factors might have been involved in this study beyond the LDB issue, it is difficult to draw a definitive conclusion on the issue of the impact of LDBs on comprehension by Chinese participants.

In brief, the findings of this study have presented limited yet revealing evidence of the impact of an LDB for comprehension for the American participants although the Chinese group did not present better performances for texts involving LDBs. However, such a phenomenon may be largely due to the language proficiency issue of the Chinese participants.

The results of this study have also provided revelations for research in the future. First, a larger pool of participants is essential. It is also necessary to ensure Chinese participants’ English proficiency levels are well measured. In addition, it will enhance the comparability of the OS and RS texts if both OS and RS texts are based on the same original piece of text, and if these texts are used as comprehension tasks for two different groups. This way, effect of focused change will be easier to identify and assess.

This study advocates the need to incorporate a discourse-based approach to language teaching and learning. The observed impact on comprehension, though limited, also lends support to the significance of teaching and learning of discourse markers.

Pedagogically, this study underlines the importance of developing English language learners’ communicative and discourse competence beyond the sentence level, for only when we examine language use at a discourse level can we uncover and realize the significance of issues such as discourse markers or uncover learner difficulties.

References


Appendix I: Listening Task Questions

"Looking for products" (Paragraph 4)

1. Why wasn’t the speaker able to find the appropriate products in 1998? One of the reasons is that ________.
   (A) these companies were not open on weekends
   (B) these companies’ catalogues didn’t carry information about their newest products
   (C) the prices of the products were too high
   (D) I don’t know

2. Is the speaker’s meaning easy to follow?
   (A) Yes          (B) Somewhat easy          (C) No

3. Please circle all relevant reasons if you have selected either answer (“somewhat easy”) or C (“No”) in question 2.
   (A) You are unfamiliar with the subject.
   (B) Difficult vocabulary choices used
   (C) Problems in sentence structures
   (D) The Order in which ideas/parts of the sentence are presented

Appendix II: Example of Original Structure vs. Reconstructed Structure

Paragraph 4: “Looking for Products”

Chinese Text

Translations of English

Original Structure (A: with an LDB)
When I started my job, my boss gave me an assignment to look for certain products for our company. As you all know, there were not many suppliers using internet in 1998, so I wasn’t able to do research via the internet and had to visit many companies and looked at their catalogues. But because these companies closed at 5 o’clock, or that their catalogues didn’t include their newest products, or that their sales personnel couldn’t find the appropriate products, I was never able to find the appropriate products.

Reconstructed Structure (B: without an LDB)
When I started my job, my boss gave me an assignment to look for certain products for our company. As you all know, there were not many suppliers using internet in 1998, so I wasn’t able to do research via the internet and had to visit many companies and looked at their catalogues instead. But I was never able to find
the appropriate products because these companies closed at 5 o’clock, or that their catalogues didn’t include their newest products, or that their sales personnel couldn’t find the appropriate products for me.

*The marker “but” under discussion is in bold while the target clause is underlined.

Appendix III: Texts for listening Comprehension

1. * "Neighborhood"

Chinese Text
搬了家, 小区里没有农贸市场了, 买菜都要到很远的地方, 听起来好像不是什么大事, 但是因为菜每天都要买, 每天都要走远路, 的确很发愁。

English translation for listening
We moved recently and there are no vegetable markets in our new neighborhood. It may not sound like a big deal, but because we need vegetables everyday, and we have to walk a long way for the vegetables, it is indeed very frustrating.

2. "Africa"

Chinese Text
而非洲连年战乱饥荒, 每天都有大量平民死去, 但是因为太过平常, 不具备新闻价值, 所以虽然死的人更多, 但是很少见过有谁主动给非洲孩子捐款。

English translation for listening
On the other hand, war and shortage of food occur very often in Africa, and people there die every day. But because such events are common and not considered "sensational" to be included for news coverage, so although more people have died in Africa than in the recent earthquake, voluntary donation for African children has been rare.

3. "EU (European Union) Constitution"

Chinese Text
宪法的表决，目前有三种可能性：第一种就是直接宣布这部宪法已经死亡，但是因为在其它的一些国家已经得到通过，而且也很有可能在其它国家得到通过，所以卢森堡首相，现任欧盟主席容克主张继续表决，所以直接宣布宪法死亡的可能性不大。

English translation for listening
The first possibility with regard to the EU Constitution is that it might be declared dead. But because the constitution has already been approved in some European countries, and it is also very likely to be passed in other countries, so the current European Chairman Mr. Ronk is in support of the continuation of the use of a general vote to let the people decide the future of the constitution. Therefore, it is unlikely that the EU Constitution will be declared dead.

4. "Looking for products"

Chinese Text
我开始工作的时候，我老板给我一个项目，就是做一个买家。...但是大家知道，98年的时候没有很多的供应商在用互联网，所以我没有办法用互联网搜索。
Sources of difficulties in cross-cultural communication and ELT

When I started my job, my boss asked me to look for certain products for our company. As you all know, there were not many suppliers using internet in 1998 so I had to visit many companies and looked at their catalogues. But I was never able to find the appropriate products because these companies closed at 5 o’clock, or that their catalogues didn’t include their newest products, or that their sales personnel couldn’t find the appropriate products for me.

5. "Finance"

Chinese Text
我们知道在过去几年之内中国的金融业取得的改革成就非常巨大，在这方面改革的进程也非常之快，但是由于中国加入世贸组织，就要更多地向世界开放中国的金融市场。

English translation for listening
We are fully aware of the tremendous accomplishments of our financial industry in the past few years as well as the speed of our progress, but we still need to make our money market more open to the world since we have now joined World Trade Organization.

6. "Technicians"

Chinese Text
他说，四川是我国"三线"建设的重点地区，一些国有企业、研究所、军工企业有很多人才，尤其是技术人才，但是虽然有些国有企业已经到了破产的地步，许多技术人员在这个企业整天无所事事，拿的工资也很低，可就是不愿意离开。

English translation for listening
Mr. Chen pointed out that the central region of our country is a key area, and there are many talented people such as technical experts in various enterprises and research institutions. But many technicians are unwilling to leave their jobs although some of these state owned enterprises, institutions, and military enterprises are already on the brinks of bankruptcy; technicians working in these enterprises receive very low pay and have nothing to do at work.

*For the sake of convenience for reference and discussion, the first three paragraphs belong to the OS category while the last three belong to the RS category. However, the actual order in which the participants heard the texts was slightly different so that the two categories were mixed intentionally rather than being completely separated.

**Some changes were made in translation to either make the texts more comprehensible or to remove any obvious indications of some specific terms, which might have provided advantages or disadvantages for comprehension to either one of the groups.