An empirical demonstration of contrastive rhetoric: Preference for rhetorical structure depends on one’s first language

MARCUS TAFT, DENISSE KACANAS, WINNIE HUEN, and RAMONY CHAN

Abstract

English written texts were produced by a group of monolingual speakers, as well as Chinese-English and Spanish-English bilinguals. These were randomly presented to another set of participants from the same three language groups for rating. The raters were unable to identify the language background of the authors of the transcripts, yet they were found to prefer the way the arguments were presented in the transcripts of their own language group. In contrast, there was no preference for the content of the arguments of the three language groups. A discourse analysis identified several aspects of the texts that might have led to the own-language preferences for rhetorical structure. The study provides empirical support for the notion of contrastive rhetoric.

1. Introduction

When evaluating rhetorical information presented in a written text (such as marking an essay), one is influenced by the organization of the arguments and their relevance to the topic, and this is an evaluation based upon what one expects to find in a well-constructed discourse (Farrell 1996). In the case of English, a native speaker expects the logic to be developed in a linear fashion (e.g., Kaplan 1966), typically following a deductive organization with a topic statement followed by specific evidence used to support the argument. However, such a linear structure is not necessarily the universal expectation because different cultures appear to encourage different types of rhetorical structure (e.g., Connor 1996; Kaplan 1966, 1972; McCool 2009; Purves 1988). This means that the evaluation of a piece of written work by a person of a different language background to the writer may well be adversely affected by the mismatch between the preferred rhetorical structure of the writer and reader. So, even if a non-native English speaker has largely mastered the syntax and vocabulary of English, they may still be at a disadvantage when it comes to...
communicating their ideas. Such a situation might be considered unfair to that person because, while the accuracy of grammar and word use can be objectively determined, rhetorical structure may merely be a matter of preference. The study presented here directly examines whether the evaluation of non-native argument structures is indeed subjectively biased by language background.

2. Contrastive rhetoric

Examination of the rhetorical style used in different languages has focused on contrastive approaches to the way in which ideas are presented. For example, it has been claimed that Japanese text follows a more inductive approach to exposition than does English text (e.g., Hinds 1983; Kubota 1998; Maynard 1996), and that Chinese writers are more indirect in their text structure than English writers (e.g., Matalene 1985; Wu and Rubin 2000; Yang and Cahill 2008). Spanish text often comprises elaborate sentences and a tendency to be less concrete than English, along with fewer explicit indications as to how the sentences connect with each other (e.g., Monroy-Casas 2008; Montaño-Harmon 1991). It follows, then, that the rhetorical structures used when writing in a second language (L2) might be carried over from the native language (L1), and claims of such transfer from L1 to L2 discourse have been made.

Kaplan (1966) was the first to put forward such a notion of “contrastive rhetoric” and supported it by examining English texts produced by native speakers of different languages. Kaplan (1966) claimed that L1 influences could be observed in these texts, including indirect structure in the discourse of speakers of Oriental languages and digressions from the central idea in the discourse of speakers of Romance languages (such as Spanish). Since that time, studies have supported the idea of contrastive rhetoric (see e.g., Hinkel 2005; Silva 1993). Such research has largely compared the rhetorical structure of L2 texts with the rhetorical structure produced by the same individuals when writing in their L1 (e.g., Hirose 2003; Kubota 1998; Uysal 2008) or with the rhetorical structure produced in the same language by native speakers (e.g., Hyland and Milton 1997; Kobayashi 1984; Reid 1992). There is now considerable evidence against the simplistic characterization that Kaplan (1966) proposed for the rhetorical structure of particular languages (e.g., Hirose 2003; Hyland and Milton 1997; Kubota 1998; Mohan and Lo 1985; Monroy-Casas 2008; Yang and Cahill 2008). Nevertheless, studies of inter-language discourse organization typically reveal either similarities between L1 and L2 texts written by the same individual or differences between native and non-native writers, even if the nature and degree of those differences is influenced by such
factors as the genre of the discourse (e.g., Connor 2002; Hyland 2004; Taylor and Chen 1991) and L2 proficiency (e.g., Kubota 1988; Scarcella 1984; Uysal 2008). Such findings can be taken to mean that there is some impact of L1 rhetorical organization on L2.

There is an alternative possibility, however, and that is the idea that L2 discourse structure is not so much the result of transfer from L1 to L2, but a demonstration of poor organizational skills when writing in L2 (e.g., Kubota 1998; Mohan and Lo 1985). When Hong Kong Chinese show greater indirectness when writing English than do native English speakers, Mohan and Lo (1985) argue that this is simply a reflection of their lack of organizational skills in producing English discourse, presumably because such skills are not nurtured within the Hong Kong educational system.

One potentially fruitful way in which to explicitly examine transfer from L1 to L2 is to compare the L2 writing of individuals from different L1 backgrounds. If individuals of two different languages show systematic differences in their L2 rhetorical structure despite similar exposure to that L2, it would imply that the specific nature of the L1 was having a differential impact on L2 performance, and that it is not simply a matter of poor skills in L2. This was the original focus of Kaplan (1966), yet there has been surprisingly little subsequent research that directly compares non-native speakers of different backgrounds on the structure of their L2 writing. The studies that have looked at this (e.g., Hinkel 2003; Reid 1992; Scarcella 1984) indicate that L2 rhetorical patterns do have different characteristics depending on the L1 of the writer, hence providing support for the idea that rhetorical patterns are transferred from L1 to L2.

3. Rhetorical preference and the present study

Rather than attempting to analyze the structure of L2 texts written by different L1 groups, the present study focuses on inter-language preferences for writing style in what appears to be the first experimental examination of contrastive rhetoric. Native English speakers (i.e., monolingual Australians), Chinese-English bilinguals, and Spanish-English bilinguals are asked to rate a set of English argumentative texts written by other members of the same three language groups, where the language background of the writers is anonymous. By having both native and non-native English speakers evaluate the texts written by other native and non-native speakers, it can be established whether non-native texts are devalued by native speakers. Importantly, it can be further established whether such devaluation, if it occurs, is due to the argument simply being of poorer quality (e.g., Mohan and Lo 1985) or due to a mismatch between the expectations of the reader and writer based on their L1. The former
would be evidenced by lower ratings of the quality of non-native texts, not just by native speakers, but by non-native speakers as well as. In contrast, if the quality of rhetorical structure is subjectively biased by language background, then Australians should prefer the structure of texts written by Australians, Chinese should prefer the structure of texts written by Chinese, and Spanish should prefer the structure of texts written by Spanish.

Finally, in order to ensure that the preferences being examined are based on the rhetorical structure of the argument rather than its conceptual content, raters are asked to evaluate the structure and content separately. It can then be seen whether differential ratings of one language group over another are observed specifically when rhetorical organization is being considered rather than the ideas being conveyed.

4. Method

4.1. Participants

All participants were recruited from the student population at the University of New South Wales, primarily from the pool of first-year psychology students (apart from two Spanish raters who were personal contacts of one of the authors). The non-native speakers were recruited from this pool by asking specifically for individuals who were either Cantonese speakers from Hong Kong or Spanish speakers from South America. These two groups were selected both because they were readily available amongst the university population, and because Chinese and Spanish putatively make use of quite different rhetorical structures (e.g., Kaplan 1966; Reid 1992). All non-native participants were asked to estimate how long they had been fluent in English, with the definition of “fluency” being left up to the individual.

There were 10 “authors” from each of three language groups: Australian monolingual English speakers (three male, seven female; mean age 20.80, sd 2.53), native Cantonese speakers (two male, eight female; mean age 20.20, sd 3.05), and native Spanish speakers (two male, eight female; mean age 19.90, sd 2.42). The mean estimated number of years of English fluency was 6.20 (sd 2.47) for the Chinese and 6.40 (sd 1.96) for the Spanish, as based on each participant’s own estimation.

Another set of 16 participants from the same three population groups served as “raters”: English (five male, 11 female; mean age 19.31, sd 1.58), Chinese (five male, 11 female; mean age 20.00, sd 1.71), and Spanish (seven male, nine female; mean age 21.00, sd 4.07). The mean estimated number of years of English fluency was 7.09 (sd 1.99) for the Chinese and 6.22 (sd 2.65) for the Spanish.
4.2. Materials and procedure

The topic that was chosen for the authors to write about was selected on the grounds that it was not highly emotional and that the content was culturally neutral: “Do you think that holding the Olympic Games has a positive or negative impact on the host country? Explain why.” The texts were handwritten in English by the participants, and then subsequently typed up for presentation to the raters. Consistent with the fact that they had met the requirements for university entry, the grammatical ability of the non-native English writers was high, though almost all had minor grammatical imperfections (such as lack of subject/verb agreement, or misuse of articles). These superficial errors were corrected by a native English speaker to avoid low-level clues to non-nativeness, but word order and vocabulary selection was left intact.

Each rater was presented a different set of five transcripts from each language group with each transcript being ultimately evaluated by eight raters from each group. The transcripts were randomly presented to each rater, and the language background of the authors was not identified. The raters were asked to read each transcript and to respond on a seven-point scale: (a) “To what extent do you agree with the argument?” and (b) “Regardless of the opinion of the writer, how well structured is the argument?” The definition of “structure” was left open for the rater to interpret in whichever way they wanted, but it should have been clear from the wording and ordering of the questions that structure was meant to be considered independent of content.

After completing this task, the raters were told that the transcripts were actually composed by monolinguals, Chinese bilinguals, and Spanish bilinguals. They were then asked to re-read each transcript and attempt to identify the language group to which each author belonged.

5. Results

Analyses were carried out separately for the ratings of agreement and the ratings of structure (i.e., rhetorical preference). The main effects were not of interest, with the focus being on the interaction between the language background of the author and the language background of the rater. The three interactions (English and Chinese, English and Spanish, Chinese and Spanish) were tested using MANOVA, with rater background being a between-groups factor and author background a within-group factor.

5.1. Ratings of agreement

The mean ratings of agreement with the content of the discourse are depicted in Figure 1. Neither rater nor author background had any impact, with $F < 0.47$ for all interactions.
5.2. Rhetorical preference

In contrast, the ratings of rhetorical structure showed clear interactions with language background, such that all groups of raters preferred the rhetorical structure of the authors from their own language group relative to either of the other two groups: English × Chinese, $F(1,45) = 12.97, p < 0.001$; English × Spanish, $F(1,45) = 12.56, p < 0.001$; and Chinese × Spanish, $F(1,45) = 29.56, p < 0.001$. In order to see these interactions most clearly, they are separately depicted in Figures 2a, 2b, and 2c respectively.

5.3. Identification of language background

The identification data were analyzed in terms of the mean number of times a transcript was correctly classified in terms of the author’s language background by the eight raters from each language group. These data are presented in Table 1. There were no significant main effects or interactions, all $F$’s < 0.92. In fact, all means were very close to chance (i.e., 1 correct out of 3, or 2.667 out of 8).

![Figure 1. Mean ratings of agreement (out of seven) for English, Chinese, and Spanish authors and raters. Standard deviations in parentheses.](image-url)
Figures 2a–c. Mean ratings of argument structure (out of seven). Standard deviations in parentheses.
6. Discussion

The highly significant statistical interactions reveal a clear preference for the rhetorical structure of one’s own native language in texts written in a second language. This is the case even when the rater is ignorant of the language background of the authors of the texts. Not only was knowledge of language background withheld from the raters, but they were also unable to consciously identify the language background when explicitly asked to do so. Thus, the rhetorical preferences that were demonstrated were not biased by the recognition of and identification with an in-group (e.g., Tajfel 1982). That is, higher ratings were not a result of raters simply knowing that they belonged to the same language group as the author. Nor could the higher ratings have been based on any systematic differences in discourse content between language groups, because none were observed in relation to ratings of agreement. Rather, the rhetorical preference must have arisen from a preference for a particular type of discourse structure.

It is possible that the monolingual English speakers rated the bilingual texts lower than the monolingual texts simply because the grammar of the former was awkward (even though minor grammatical corrections had been made). However, even if this were the case, the most dramatic feature of the results is the fact that the Chinese bilinguals preferred the Chinese texts and the Spanish preferred the Spanish texts, even relative to the native English texts. Such an affinity with one’s own language group, and only for ratings of rhetorical structure, implies a sensitivity to the discourse organization of one’s native language, resulting in a subjective bias when evaluating the written work of others.

It needs to be ensured, however, that the proficiency of the Chinese and Spanish authors was similar, given that rhetorical organization in L2 is likely to be influenced by L2 proficiency (e.g., Kubota 1998; Mohan and Lo 1985; Uysal 2008). In relation to this, the two groups were matched on the number

<table>
<thead>
<tr>
<th>Background of rater</th>
<th>Identified language background</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“English”</td>
</tr>
<tr>
<td>English</td>
<td>2.75 (0.70)</td>
</tr>
<tr>
<td>Chinese</td>
<td>2.50 (0.93)</td>
</tr>
<tr>
<td>Spanish</td>
<td>2.50 (0.54)</td>
</tr>
</tbody>
</table>

Table 1. Mean correct identifications of authorship (out of 8) by the English, Chinese, and Spanish raters. Standard deviations in parentheses.
of years they believed they had been fluent in English. So, unless the Chinese and Spanish authors differed systematically on their estimation of “fluency,” there is little basis to think they differed in proficiency. Moreover, since all members of the two bilingual groups of authors were undergraduate university students matched on average age, there is little basis for arguing that there might have been a systematic difference in their exposure to English. Importantly, the fact that the monolingual English raters gave equal ratings to the arguments of the Chinese and Spanish groups (4.34 vs. 4.40, respectively) further attests the likelihood that they were of equivalent proficiency. Even if they differed, however, it would not explain why Chinese raters preferred the Chinese texts and Spanish raters preferred the Spanish texts. If a discourse was structured in a particular way because the writer lacked organizational skills, it is likely to have been devalued by all raters regardless of language background.

Although the aim of this study was simply to demonstrate the impact of language-specific rhetorical structures on the evaluation process, it is necessary to give some consideration to which aspects of the texts might have led to the pattern of preferences. It was left up to the raters themselves to decide how to interpret the notion of a “well-structured” argument, so it is possible that there were cultural differences in relation to which aspect of the discourse to focus on (e.g., the cohesiveness of the clauses, the use of paragraphing, the signposting of anaphoric reference, etc.). However, even if this were the case, it would not explain why the different language groups preferred the texts of their own language group over the others. It seems that the explanation for that would have to lie in what was produced by the authors rather than what the raters focused on.

In order to provide at least some information about the possible basis for the differential language preferences, the discourse was analyzed in terms of several characteristics that might be expected to differ between the different language backgrounds.

6.1. Discourse analysis

It has been suggested that discourse written by Chinese speakers is shorter and less elaborated than that of English speakers (e.g., Taylor and Chen 1991), whereas that written by Spanish speakers tends to include digressions and repetition as well as flowery sentence structures (e.g., Farrell 1996; Monaño-Harmon 1991). As such, it might be expected that the sentences found in the native English transcripts were longer and more numerous than in the Chinese ones, with the reverse being true for the Spanish transcripts. Table 2 presents various measures of such structural complexity for each language group. For all post-hoc analyses, variability was reduced by removing the single highest
and lowest score from each group, and Fischer’s LSD (Least Significant Difference) test for multiple comparisons was used to establish any statistical differences.

Although not differing at all in terms of the overall number of words used, the Chinese group actually used more sentences than the other two groups, with the average length of each sentence being shorter. There was no difference between the English and Spanish texts. The larger number of sentences used by the Chinese could mean that they were more circuitous in their argumentation, repeating points and listing a variety of ideas (cf. Kaplan 1966), or it could mean that they were more detailed in the presentation of their arguments. To examine this, the rhetorical structure of the transcripts was analyzed by classifying the relationships between their component clauses and then quantifying them. The rhetorical relationships identified by Mann and Thompson (1988) were used as a guide for this classification, but only a few of their fine-grained categories were used, sometimes collapsing across more than one. The following are the most informative classifications that were made.

(a) **Justification**: Where a clause aims to increase the belief in and acceptance of an adjacent clause (e.g., “*Holding the Olympic Games can be positive*” + “*because it raises the international profile of the host country*”).

(b) **Elaboration**: Where a clause adds more detail to the information contained in the previous one (e.g., “*It will bring overseas countries’ attention to the host country’s organization skills*” + “*and their efficiency in planning and conducting the Games*”).

(c) **Outcome**: Where a clause describes what follows from the previous one (e.g., “*If they do a good job in controlling the budget*” + “*the economic benefits will be positive*”).

(d) **Contrast**: Where a clause provides an alternative view to the previous one (e.g., “*While the thousands of people arriving in a city may inconvenience those living there*” + “*most may find it an opportunity for nationalistic pride*”).

Table 2. *Measures of structural complexity for the English (E), Chinese (C), and Spanish (S) transcripts, along with t values (d.f. = 14 in each case).*

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Chinese</th>
<th>Spanish</th>
<th>E vs C</th>
<th>E vs S</th>
<th>C vs S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Words</td>
<td>246</td>
<td>253</td>
<td>233</td>
<td>&lt; 1</td>
<td>&lt; 1</td>
<td>1.48</td>
</tr>
<tr>
<td>Number of Sentences</td>
<td>10.75</td>
<td>13.63</td>
<td>9.75</td>
<td>2.22</td>
<td>&lt; 1</td>
<td>2.99</td>
</tr>
<tr>
<td>Words per Sentence</td>
<td>23.92</td>
<td>18.90</td>
<td>24.28</td>
<td>2.66</td>
<td>&lt; 1</td>
<td>2.85</td>
</tr>
</tbody>
</table>
Table 3. Percentage of clausal relationships belonging to the different categories for the English (E), Chinese (C), and Spanish (S) transcripts, along with t values (d.f. = 14 in each case).

<table>
<thead>
<tr>
<th>Category</th>
<th>English</th>
<th>Chinese</th>
<th>Spanish</th>
<th>E vs C</th>
<th>E vs S</th>
<th>C vs S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Justification</td>
<td>15.7%</td>
<td>13.3%</td>
<td>19.6%</td>
<td>t = 1.05</td>
<td>t = 1.16</td>
<td>t = 2.21</td>
</tr>
<tr>
<td>Elaboration</td>
<td>25.4%</td>
<td>14.6%</td>
<td>19.2%</td>
<td>t = 3.16</td>
<td>t = 2.21</td>
<td>t &lt; 1</td>
</tr>
<tr>
<td>Outcome</td>
<td>6.6%</td>
<td>13.6%</td>
<td>13.1%</td>
<td>t = 3.27</td>
<td>t = 2.91</td>
<td>t &lt; 1</td>
</tr>
<tr>
<td>Contrast</td>
<td>10.0%</td>
<td>5.6%</td>
<td>6.2%</td>
<td>t = 2.20</td>
<td>t = 1.81</td>
<td>t &lt; 1</td>
</tr>
<tr>
<td>Repetition</td>
<td>5.9%</td>
<td>8.4%</td>
<td>9.9%</td>
<td>t = 1.31</td>
<td>t = 1.85</td>
<td>t &lt; 1</td>
</tr>
</tbody>
</table>

(e) Repetition: Where a clause has the same content as a previous one (e.g., “there can be seen three benefits” + “and thus positive impacts”).

Table 3 reports the percentage of relationships between adjacent clauses that correspond to each of these categories for the three language groups.

These data suggest that the fact that the Chinese used more sentences than the monolinguals did not arise from the former elaborating upon and justifying their arguments more strongly. Instead of adding to the basis for their argument, they tended to describe the consequences of the point they were making (i.e., “outcome”). The amount of repetition was not statistically significant, though it should be noted that seven of the Chinese transcripts provided a final statement summarizing the arguments, whereas only three of the English transcripts did this. The reason why the Chinese used more sentences than the Spanish is not obvious from the rhetorical analysis because the two groups did not differ on most measures. However, the Spanish provided a greater justification for their arguments. Like the Chinese authors, the Spanish authors included more final summaries (six cases) than did the monolinguals. The use of contrasting arguments was most characteristic of the monolingual group.

The above analysis gives no indication whether the Spanish texts were characterized by particularly flowery language (e.g., Monaño-Harmon 1991). A count of the types of words used, however, gives some suggestion that they might have been. There are at least two potential indicators of an elaborate writing style that were most prevalent in the Spanish transcripts. First, the Spanish used nominalizations of verbs and adjectives more often than did the monolinguals (0.85% of words vs. 0.45% respectively, t(14) = 3.05, p < 0.01), though not more often than the Chinese (0.66%, t(14) = 1.41, p < 0.1). Nominalization is often a form of indirectness (e.g., Hinkel 1997) and increases complexity by generating an abstract noun (e.g., by saying “we can see the
bonding between the different countries” rather than “we can see the different countries bond”). Second, adjectives are a form of embellishment and the Spanish used them more often (5.1% of words) than both the monolinguals (3.7%), t(14) = 2.41, p < 0.05, and the Chinese (3.8%), t(14) = 2.23, p < 0.05.

Obviously, own-language preferences can only emerge if there are systematic differences between the texts written by individuals of different language backgrounds. The purpose of the discourse analysis was simply to demonstrate that at least some of these differences can be identified, even if it cannot pinpoint all possibilities.

Clearly, the differences between languages is quite subtle because it was apparent that the raters were unable to differentiate the three groups consciously. Instead, the rhetorical differences emerged in the form of a preference for texts produced by the language group of the person doing the evaluation. This is a very striking finding, particularly the difference observed between the two bilingual groups, whose experience with English was comparable. Such differential language preferences cannot be explained in terms of the L2 text simply being poorly organized as a result of the lower proficiency of the L2 writer (e.g., Mohan and Lo 1985). Instead, they provide clear support for the notion of contrastive rhetoric, where the structure of discourse in a second language is influenced by characteristics of the native language of the writer.

The results of this study emphasize the subjectivity of rhetorical structure. It is important, therefore, that bilingual speakers of a language be aware of the preferred rhetorical organization of that language if they are to avoid being disadvantaged when having their written work evaluated by a native speaker. While this might be well-known to teachers of English as a second language (see e.g., Hinkel 2005, 2006), it may not be so well-known to the bilingual speakers themselves. The next empirical step, then, is to try to identify those specific aspects of text that affect differential rhetorical preferences, perhaps by examining the impact of manipulations of different rhetorical structures on preference ratings. In the light of the knowledge gained from such research, bilingual speakers could potentially adapt their writing style to the expectations of the native reader.

References

An empirical demonstration of contrastive rhetoric


_Correspondence address:_ m.taft@unsw.edu.au