Divergent news representations of Lien Chan’s visit to China
A corpus-based lexical comparison between the *China Post* and the *China Daily*

Honglei Wang
Northwestern University

By combining the Critical Discourse Analysis (CDA) approach with the corpus linguistic method, this study compares the lexical choices employed by two English newspapers, i.e. the *China Daily* (CD) and the *China Post* (CP), in their coverage of Taiwan’s KMT (Kuomintang) Chairman Lien Chan’s visit to China in April 2005. The comparative analysis of the two corpora of news reports is conducted at three levels: word frequencies, part-of-speech (POS) frequencies and semantic category frequencies. The paper argues that the differences in the news representation of Lien’s visit are attributed to the different socio-political functions performed by these two newspapers: while the CD depicts the event as a momentous one that has received nation-wide plaudits in order to echo faithfully China’s policy of peaceful reunification, the CP presents a more comprehensive picture of the event in that it also reports in detail some incidents, including violent protests, triggered by Lien’s visit that has great impact on the domestic social life of Taiwan.

**Introduction**

In the last two decades, beginning with van Dijk (1988) which uncovers the racist prejudices implied in the various linguistic features of the Western news reports, a large body of studies has been devoted to the critical research in the impact of power and ideology on the language employed in the media discourse. Most of these studies are based on empirical analysis of news texts and theoretically, they draw on critical linguistics and critical discourse analysis (CDA) developed by such scholars as Fairclough (1989, 1992, 1995), Fowler (1991, 1996) and van Dijk (1997a, 1997b, 1998, 2006). These two closely related theories share several
essential underlying assumptions, for example, that various linguistic features, whether lexical choices or syntactical structures, serve to reproduce the existing social reality and power relations, and that ideological and political implications are embedded in this reproduction process. The ultimate aim of the critical approach to the study in discourse and language is at unmasking the ideological motivations underlying the linguistic choices and discourse features.

Inspired by those studies based on the media language in the Western world, there has emerged a profusion of studies devoted to the critical analysis of the language utilized by the media institutions in China and Taiwan, and most of these studies point to the conclusion that these media institutions manipulate language to represent those politically significant events in the ways that are compatible with their political or ideological standpoints. The linguistic features that have been examined by these studies include headlines (Shih and Soong, 1998), quotations (Kuo, 2003; Li, 1996), syntactical structures and lexical choices (Fang, 1994, 2001; Kuo, 2001; Kuo and Nakamura, 2005) and other features (S. Wang, 1993; Yu, 1995).

Despite the extensive studies in how lexical choices are employed by the news agencies in either China or Taiwan to represent the social and political reality, they are mostly qualitative in nature and focus only on a few selected lexical items. Given the fact that meanings are usually expressed through collocations of a group of lexical items (Leech, 1974), a quantitative analysis of a wide range of lexical items in a large corpus of news reports will reveal a more comprehensive picture about how words in the news texts work ideologically.

Stubbs (1997) is one of the first scholars suggesting that a corpus-based quantitative analysis should be combined with a qualitatively oriented critical discourse analysis to strengthen the conclusion arrived at. The advantage of a corpus-based study is also recommended by other researchers, for example, Biber, Conrad and Reppen (1994: 169) observe that “text corpora provide large databases of naturally occurring discourse, enabling empirical analyses of the actual patterns of use in a language, and, when coupled with (semi-)automatic computational tools, the corpus-based approach enables analyses of a scope not otherwise feasible”. Partington (2003: 6) also remarks that “a corpus-based analysis adds an empirical dimension to introspection and marries quantitative and qualitative research methodologies”. Following these observations, some CDA-informed studies (e.g. Bayley, 1999; Hardt-Mautner, 1995; Krishnamurthy, 1996) rely on a quantitative analysis of large corpora in order to corroborate their conclusions concerning the relationship between language and ideology. Another characteristic of the recent corpus based CDA studies (e.g. Bayley, Bevitori and Zoni, 2004; Bayley and Vicente, 2004; Orpin, 2005) is that they often involve a comparison between at least two corpora since “by and large, we are not methodologically justified in interpreting
the significance of a particular linguistic event unless we can compare it with other similar events” (Partington, 1998: 146).

Moreover, most of these studies tend to focus on lexical items in the corpora, and this is understandable for several reasons. The most obvious one is that compared with other linguistic features, lexical choices are much easier to be examined in the corpus and the other more important one, according to Bayley and Vicente (2004: 238), is that “the nuances of particular words may reflect political orientations- or discourse formations- towards a phenomenon that has become socially problematic.” Therefore, lexical items may provide the most direct cue to how the media institutions interpret the social reality and how the institutions intend the reader to understand the social reality as well.

By combining a CDA approach with a corpus linguistic method, this paper compares the lexical choices employed by two English newspapers, i.e. the *China Daily* (CD) and the *China Post* (CP), in their coverage of Taiwan’s KMT (Kuomintang) Chairman Lien Chan’s visit to China in April 2005. In the following, Section 2 describes the database and methodology of my study. Section 3 presents a detailed comparative analysis of the two corpora of news texts taken from the two newspapers. Section 4 concludes this study by proposing a possible reason to account for these discrepancies in the reporting of Lien’s visit.

**Database and methodology**

Since this study is based on the news reports on Lien Chan’s visit to China, some background information about this event needs to be mentioned briefly here. This visit was initiated by Lien Chan, the Chairman of the Kuomintang (KMT) Party or the Nationalist Party, which has remained the largest opposition party in Taiwan since Chen Shui-bian, from the Democratic Progressive Party (DPP), was elected to the presidency in 2000. Although this visit lasted only several days, from 26th April to 4th May in 2005, it has attracted much attention from both China and Taiwan, because this was the first time that the KMT set foot on China since it retreated into Taiwan after the defeat by the Communist Party of China (CPC) in the civil war in 1949. Since then, the political status of Taiwan has remained a controversial issue contended by both sides: while China insists that Taiwan is an inseparable part of China that will be reunited sooner or later, the ruling regime in Taiwan regards Taiwan as an independent country. The relations between the two sides worsened considerably when Chen became the president of Taiwan, implementing his pro-independence policy in an attempt to separate Taiwan from China permanently. In order to restrain Chen’s intentions, the National People’s Congress of China passed the Anti-Secession Law on 14th March, 2005, which
authorizes Beijing to resort to the military force if Taiwan declares its independence. Soon after, Lien declared at a news conference that his visit to China was aimed at seeking a peaceful means of easing the tense relations between China and Taiwan. In the ensuing months, other opposition parties in Taiwan (the People First Party and the New Party) visited China successively.

Given the political sensitivity of Lien’s visit, this event was reported extensively by the media in both China and Taiwan. I select the news reports about this event from two newspapers: the CD, the leading English newspaper in China and its counterpart in Taiwan, the CP. The corpus on which this study is based consists of 74 news reports from the CP, dating from 14th April to 6th May with a total of 42256 running words, and 97 news reports from the CD, dating from 9th April to 9th May with a total of 49332 running words. Since the common practice in the corpus analysis is to use the smaller corpus to compare the larger one, in this research I choose to compare the CP with the CD. The two computational tools that I use in this study are WordSmith Tools 4.0 (Scott, 1999) and Wmatrix (Rayson, 2003). While the most obvious advantage of the former is its ability to compare the frequencies of the same word between the two corpora, the beauty of the latter is that it can get each word in the corpus parsed and tagged with the appropriate part-of-speech (POS) and semantic category, so that the frequencies of those POSs and semantic categories between the two corpora can be compared.

Data analysis

In this section, with the help of WordSmith Tools 4.0 and Wmatrix, I compare the lexical choices between the two corpora in terms of word frequency, POS frequency and semantic category frequency.

Some notes on the notations used in this section are in order here. Since in this study, I compare the corpus of the CP with that of the CD, in Tables 1 and 2 below, O1 indicates the observed frequency of a particular item in the CP and O2, the observed frequency of the item in the CD. And %1 and %2 values show the relative frequencies of an item in the CP and the CD respectively. Furthermore, in this paper, the log-likelihood value (LL value) is calculated for the sake of measuring the extent of difference in the relative frequency of a particular item between the two corpora. The larger the absolute value of the LL parameter is, the more different the two corpora are in terms of the relative frequency of the target item. Besides, the LL value calculator also shows a plus or minus symbol before the LL parameter to indicate that a particular item is more or less frequent in the CP than in the CD, hence, “+” indicates the more occurrence of the target item in the CP relative to the CD, and “−” the less occurrence of the item in the CP relative to the CD.
Comparison at the word level

I use WordSmith Tool 4.0 to produce a list of keywords which is based on a comparison of the observed frequencies of all the words in the two corpora. This list consists of two sections: the first is composed of those words which occur more frequently in the CP than in the CD (Table 1) and the second composed of those words which are used more frequently in the CD than in the CP (Table 2). For the sake of comparison, in the following two tables, I list only the top nine words with the highest LL values in each section, since these significantly overused or underused words are most likely to exhibit the difference between these two corpora.

Table 1. The overused words in the CP

<table>
<thead>
<tr>
<th>Item</th>
<th>O1</th>
<th>%1</th>
<th>O2</th>
<th>%2</th>
<th>LL</th>
</tr>
</thead>
<tbody>
<tr>
<td>president</td>
<td>293</td>
<td>0.67</td>
<td>30</td>
<td>0.04</td>
<td>+298.62</td>
</tr>
<tr>
<td>China</td>
<td>647</td>
<td>0.94</td>
<td>235</td>
<td>0.47</td>
<td>+283.49</td>
</tr>
<tr>
<td>Chen</td>
<td>305</td>
<td>0.70</td>
<td>67</td>
<td>0.16</td>
<td>+211.21</td>
</tr>
<tr>
<td>government</td>
<td>226</td>
<td>0.52</td>
<td>35</td>
<td>0.06</td>
<td>+192.86</td>
</tr>
<tr>
<td>police</td>
<td>89</td>
<td>0.20</td>
<td>0</td>
<td>0</td>
<td>+140.12</td>
</tr>
<tr>
<td>party</td>
<td>713</td>
<td>0.97</td>
<td>500</td>
<td>0.84</td>
<td>+111.73</td>
</tr>
<tr>
<td>if</td>
<td>158</td>
<td>0.38</td>
<td>36</td>
<td>0.06</td>
<td>+106.40</td>
</tr>
<tr>
<td>said</td>
<td>67</td>
<td>0.16</td>
<td>2</td>
<td>0</td>
<td>+ 89.79</td>
</tr>
<tr>
<td>time</td>
<td>77</td>
<td>0.18</td>
<td>88</td>
<td>0.17</td>
<td>+ 86.52</td>
</tr>
</tbody>
</table>

Table 2. The overused words in the CD

<table>
<thead>
<tr>
<th>Item</th>
<th>O1</th>
<th>%1</th>
<th>O2</th>
<th>%2</th>
<th>LL</th>
</tr>
</thead>
<tbody>
<tr>
<td>historic</td>
<td>37</td>
<td>0.07</td>
<td>109</td>
<td>0.23</td>
<td>-43.78</td>
</tr>
<tr>
<td>of</td>
<td>16</td>
<td>0.04</td>
<td>88</td>
<td>0.17</td>
<td>-42.86</td>
</tr>
<tr>
<td>years</td>
<td>119</td>
<td>0.32</td>
<td>300</td>
<td>0.51</td>
<td>-42.21</td>
</tr>
<tr>
<td>win-win</td>
<td>19</td>
<td>0.04</td>
<td>92</td>
<td>0.18</td>
<td>-40.12</td>
</tr>
<tr>
<td>relations</td>
<td>24</td>
<td>0.05</td>
<td>59</td>
<td>0.14</td>
<td>-39.06</td>
</tr>
<tr>
<td>exchange</td>
<td>54</td>
<td>0.12</td>
<td>161</td>
<td>0.39</td>
<td>-38.35</td>
</tr>
<tr>
<td>in</td>
<td>638</td>
<td>0.92</td>
<td>10333</td>
<td>0.98</td>
<td>-37.57</td>
</tr>
<tr>
<td>sides</td>
<td>32</td>
<td>0.06</td>
<td>109</td>
<td>0.23</td>
<td>-31.83</td>
</tr>
<tr>
<td>both</td>
<td>61</td>
<td>0.14</td>
<td>156</td>
<td>0.37</td>
<td>-27.82</td>
</tr>
</tbody>
</table>

My analysis in this section begins with the word president, which has a marked contrast in the frequencies in the CP and the CD in that it occurs as many as 293 times in the CP but only 30 in the CD. A comparison of the concordances of this word in
the two corpora reveals the fact that the different use of this word may reflect the different interpretations of Taiwan’s political status by the two newspapers.

A closer analysis of the concordances of all the tokens of this word in the CP shows that an overwhelming majority of tokens refer to Chen Shui-bian, the president of Taiwan and head of the DPP, whether it is used alone or co-occurs with the name of Chen Shui-bian. And most of these concordances, as illustrated in the examples below, are employed to describe Chen’s reaction to Lien Chan’s visit and the specific activities undertaken by him before or during Lien’s visit to China.

The concordances of *president* in the CP
1. …a series of decisions reached by President Chen Shui-bian in a high level meeting a week ago…
2. …talks with Beijing officials in China. President Chen said no political party had the power with…
3. …peace pact with Beijing leaders. President Chen Shui-bian has said he would support Lien’s v…
4. …discuss his planned China trip with President Chen Shui-bian. Lien’s visit is tentatively late…

Besides, there are about forty tokens of *president* referring to Hu Jintao, the president of China and General Secretary of the CPC. Subsequently, I search for the concordances of *president* in the CD and find several interesting facts worth considering. First, rather than being used to refer to Hu Jintao or Chen Shui-bian, a majority of the tokens of this word refers to the head of an organization, as illustrated below:

The concordances of *president* in the CD
1. …May 2, and may even meet Wang Daohan, president of the Association for Relations Across …
2. …the first-ever KMT delegation headed by Vice President Chiang Pin-kung from Taiwan toured …
3. …”good news for all Chinese,” Hunter Huang, president of the Washington-based National Asso…
4. …he added. Jing Gong, honorary president of the Coordination Council of Chinese Americans …

Second, I have also found two cases in the CD where the name of Chen Shui-bian is prefaced with the title of president, with the word *president* cited in a quotation mark. According to Li (1996), the quotation mark in the Chinese journalistic language may imply an ironic rejection, hence, in the case of the CD, the invalidation of the legitimate presidency of Chen Shi-bian.

The concordances of *president* used in quotation marks in the CD
1. …improve relations with the mainland. Taiwan “*president*” Chen Shui-bian of the pro-independe …
Third, I also search in the corpus of the CD for the concordances containing “Chen Shui-bian”, finding that the CD mentions this name in those contexts where the political background of Lien's visit is introduced and where Chen is described as the most arduous supporter of separation of Taiwan from China.

Finally, in contrast with the CP in which Hu Jintao is addressed frequently as the president of China, in those reports in the CD concerning the meeting between Lien and Hu Jintao, the latter is often addressed as the General Secretary of the CPC, rather than as the president of China. According to Kuo (2003) and Jaworski and Galasinski (2000), address forms and referring forms in political contexts help to reconstruct the political reality and justify ideological stances. Therefore, the sharp contrast in the use of the referring forms for Chen Shui-bian and Hu Jintao is deeply rooted in the different ways that the CD and the CP interpret the present relations between China and Taiwan.

In the case of the CD, since the fundamental and consistent position of the CPC of China is that Taiwan is part of China and will be reunited with China, the CD never acknowledges the legitimacy of the presidency of Chen Shui-bian in its news reports, and consequently the only title preceding the name of Chen Shui-bian is taiwan lingdaoren, which means the leader of Taiwan. Currently, even in some programs broadcast by the pro-China Phoenix TV station in Hong Kong, whenever those Taiwanese interviewees refer to Chen as zongtong “president”, the Chinese subtitles appearing at the bottom of the screen will be taiwan lingdaoren (Kuo, personal communication). This position taken by the CD also bears on in what capacity Hu Jintao is supposed to hold talks with Lien Chan. In order to produce among the audience an impression that the meeting between Hu and Lien is one between the two parties within the same country, Hu is reported by the CD as the General Secretary of the CPC, in lieu of the president of China.

In contrast, the CP, as a local newspaper circulated in Taiwan, does not intend to impugn Chen's presidency of Taiwan, and thus still addresses him by using the title of president.

Two other lexical items that exhibit significant over-frequencies in the CP compared with the CD are government and police, both of which reveal some aspects of Lien's visit selectively reported in the CP. In the case of government, most of the concordances underscore the various means by which the DPP, as the ruling party of the central government in Taiwan, voiced their opposition to Lien's visit, for example, by warning that Lien's visit still needs the government authorization, or by accusing Chiang Pin-Kung, vice chairman of the KMT, of violating the constitution of Taiwan for his visit to China as a prelude to Lien's visit. The second word police, which tends to occur in the reporting on the violent
actions against Lien’s visit, will be analyzed with reference to two other related words in Section 3.3.

The concordances of government in the CP
1 …political party had the power to deal with China without previous government authorization…
2 …the government has initiated an independent legal investigation into the delegation’s trip to …
3 …with Beijing. He repeated the need for Lien to receive government authorization for his upc…
4 …agreement with the ruling party in China should be a government-to-government matter, Wu..

Table 2, which contains those significantly overused words in the CD, however, presents an alternative picture of Lien’s visit, and a close analysis of the concordances of some words suggests that it is the strategy of the CD to accentuate some aspects of Lien’s visit, such as the wide acclamations he received from the people in China, the potential consequence of his visit, particularly the possible improvements in the cross-Strait relations resulting from his visit, etc. The word on the top of Table 2, historic, is one of the most frequently used terms to characterize Lien’s visit to China.

The concordances of historic in the CD
1 …opposition Kuomintang (KMT) party, is expected to make an historic visit to the mainland in…
2 …Chiang Pin-kung and other KMT members made the historic mainland visit between March 2
3 …cultural exchanges in the wake of the historic KMT visit. "Setting up hurdles to bilateral exch…
4 …at Lukou International Airport. "I am very happy we finally took a first historic step," Lien…

Functionally similar to the word historic, two other words that are frequently used in the CD to underline the possible contribution made by Lien Chan to the improvement of the cross-Strait relations are win-win and exchange, with their concordances listed below:

The concordances of exchange in the CD
1 …achieving common economic development. The inter-party exchange conforms to the public’s…
2 …understanding about the mainland and exchange views with mainland leaders on common c…
3 …the city of Beijing will do the utmost to improve exchange and cooperation, enhance mutual…
4 …The two forums are intended to help facilitate exchange of ideas across the Straits and reflec
The concordances of *win-win* in the CD

1. …help lay some groundwork for exploring a *win-win* solution so that both sides can enjoy a
2. …mutually beneficial, peaceful and *win-win* future for both sides is of concern to all people,” he..
3. …Chinese mainland and Taiwan to join hands for common prosperity. The *win-win* situation...
4. …"mutually beneficial and helpful, coexistent and *win-win*" relations across the Straits, he said.

**Comparison at the level of part of speech (POS)**

After each word in the two corpora is parsed and assigned by Wmatrix the appropriate part-of-speech according to the UCREL CLAWS7 Tagset system, I obtain a list that compares the parts-of-speech in terms of the extent of difference in their frequencies in the two corpora. This list shows that among the top twenty parts-of-speech with the highest LL values, except for the nineteenth one, JJ, which denotes what are called general adjectives in this tagging system, all the other parts-of-speech fall within the range of functional words, denoting no semantic meaning at all. It is noted that the class of general adjectives is listed in the UCREL CLAWS7 Tagset system on a par with another three classes of adjectives: general comparative adjectives (e.g. older, better, stronger), general superlative adjectives (e.g. oldest, best, strongest) and cantenative adjectives (e.g. “able” in “be able to”, “willing” in “be willing to”). General adjectives, which are supposed to include the most commonly used adjectives functioning as prenominal modifiers, are most likely to reveal the way that Lien's visit is described by each newspaper, and the analysis in this section will focus on this kind of adjectives. Furthermore, I scrutinize with caution the concordances in which these general adjectives are used in order to ensure that these adjectives do appear in the contexts where they are used to modify Lien's visit.

The result produced by Wmatrix indicates that there are altogether 3621 general adjectives in the CD but 2704 in the CP, and that the LL value between the two corpora in this case is 33.44. Tables 3 and 4 show the general adjectives with more than 30 occurrences in the CD and the CP, respectively.

As indicated in the Tables 3 and 4, among the adjectives with more than 30 tokens, those highlight the positive aspects of Lien's visit occupy a much larger proportion in the CD than in the CP, and the CD also contains a much richer variety of word types used to describe the significance of this event than the CP does. These include: *historic, win-win, common, peaceful, mutual, new* and *welcome*, with some of the concordances of *common* and *peaceful* listed below.
Table 3. The most frequently used general adjectives in the CD

<table>
<thead>
<tr>
<th>Item</th>
<th>POS</th>
<th>Raw Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td>JJ</td>
<td>220</td>
<td>0.46</td>
</tr>
<tr>
<td>historic</td>
<td>JJ</td>
<td>109</td>
<td>0.23</td>
</tr>
<tr>
<td>win-win</td>
<td>JJ</td>
<td>92</td>
<td>0.18</td>
</tr>
<tr>
<td>economic</td>
<td>JJ</td>
<td>79</td>
<td>0.17</td>
</tr>
<tr>
<td>common</td>
<td>JJ</td>
<td>72</td>
<td>0.17</td>
</tr>
<tr>
<td>peaceful</td>
<td>JJ</td>
<td>53</td>
<td>0.12</td>
</tr>
<tr>
<td>mutual</td>
<td>JJ</td>
<td>43</td>
<td>0.09</td>
</tr>
<tr>
<td>general</td>
<td>JJ</td>
<td>41</td>
<td>0.08</td>
</tr>
<tr>
<td>New</td>
<td>JJ</td>
<td>38</td>
<td>0.07</td>
</tr>
<tr>
<td>welcome</td>
<td>JJ</td>
<td>38</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Table 4. The most frequently used general adjectives in the CP

<table>
<thead>
<tr>
<th>Item</th>
<th>POS</th>
<th>Raw Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td>JJ</td>
<td>157</td>
<td>0.38</td>
</tr>
<tr>
<td>Ruling</td>
<td>JJ</td>
<td>65</td>
<td>0.15</td>
</tr>
<tr>
<td>political</td>
<td>JJ</td>
<td>59</td>
<td>0.14</td>
</tr>
<tr>
<td>democratic</td>
<td>JJ</td>
<td>57</td>
<td>0.13</td>
</tr>
<tr>
<td>public</td>
<td>JJ</td>
<td>51</td>
<td>0.12</td>
</tr>
<tr>
<td>progressive</td>
<td>JJ</td>
<td>43</td>
<td>0.11</td>
</tr>
<tr>
<td>violent</td>
<td>JJ</td>
<td>43</td>
<td>0.11</td>
</tr>
<tr>
<td>international</td>
<td>JJ</td>
<td>42</td>
<td>0.10</td>
</tr>
<tr>
<td>economic</td>
<td>JJ</td>
<td>39</td>
<td>0.09</td>
</tr>
<tr>
<td>national</td>
<td>JJ</td>
<td>39</td>
<td>0.09</td>
</tr>
<tr>
<td>historic</td>
<td>JJ</td>
<td>37</td>
<td>0.07</td>
</tr>
</tbody>
</table>

The concordances of common in the CD

1. …striving for a peaceful reunification. Meanwhile, the bill also shows the common will and stro…
2. …interests of the Taiwanese people and goes against the common aspiration of people across the…
3. …both sides can enjoy a favourable environment for achieving common economic development.
4. …work hard to improve cross-Straits relations and promote common development of both sides…

All rights reserved
The concordances of peaceful in the CD

1. “Finding a reciprocal, mutually beneficial, peaceful and win-win future for both sides is of c…
2. “…warm feelings”, and enabled the general public to see the hope of peaceful cross-strait reun…
3. “…trend of public opinions to build a peaceful and win-win future through reconciliation and dia…
4. “…who oppose ”Taiwan independence” to strive for a peaceful, stable, co-operative future across..

These selected concordances are in line with the observation made earlier that what is reported in the CD makes the reader believe that Lien's visit has political and historical implications. Since his visit is considered by the CPC as a positive response to its policy of peaceful reunification, the CD, politically affiliated with the national policy of the state led by the CPC, describes in great detail his visit, especially such facts as how he was hailed by the people in China, and the rapturous applause with which his visit was greeted by China. A further piece of evidence is that during Lien’s stay in China, CCTV (China Central Television, the highest official TV station of China) never ceased the live coverage of Lien’s activities; therefore, the audience kept updated with the detailed information about his visit, for example, his tour of the Mausoleum of Dr. Sun Yat-sen (the founding father of the KMT), his visit to the tomb of his grandmother, and even his return to his former primary school located at Xi'an.

What is reported on Lien's visit in the CP, however, demonstrates several differences from what is reported by the CD. First, the positive significance of his visit receives much less coverage by the CP given the fact that among all the most frequently used adjectives in Table 4, there is only one word, historic (with 37 occurrences), which appraises the historic importance of this event with reference to the relations between China and Taiwan. Another striking difference is that the CP releases much less information about the specific activities undertaken by Lien during his stay in China, sketching only several cities that Lien visited and his two other important activities: his delivery of speech at Peking University and his meeting with Hu Jintao, the General Secretary of the CPC.

Second, as evidenced by their concordances, the majority of the word types in Table 4 appear in the contexts where the reactions induced by Lien's visit among the several political forces in Taiwan are reported. And among all these reactions, those from the ruling DPP draw primary attention by the CP, given the fact that the DPP acts as the strongest force supporting the independence of Taiwan. Furthermore, as indicated in Table 4, three words have very high frequencies: ruling, democratic and progressive, with the latter two often co-occurring with each other. The concordances of these three words convey such information that the CP devotes much space to describing the attitude towards Lien’s visit held by the
DPP, particularly by Chen Shui-bian. No such information, however, is included in the CD.

The concordances of *ruling* in the CP

1. …in contact with government officials. *Ruling* party and government officials claim Lien shun…
2. …leader, is also planning to visit China. A *ruling* party spokesman urged Lien and Soong to reach…
3. …Lien Chan by government officials and *ruling* party politicians, Chen said in the same ann…
4. …trip, which concludes May 3, with some *ruling* party lawmakers suggesting the trip amounts to…

The concordances of *democratic* and *progressive* in the CP

1. …Wen-chung of the ruling *Democratic Progressive* Party said that China will invade Taiwan only…
2. …party *Democratic Progressive* Party officials had threatened to take legal steps against Lien…
3. …*Democratic Progressive* Party legislative whip William Lai also warned Lien against breaking…
4. …after the ruling *Democratic Progressive* Party administration put weeks of pressure on Lien to…

**Comparison at the level of semantic category**

In this section, the CP and the CD are to be compared with the help of Wmatrix in terms of the semantic categories employed in the two corpora. After the words in the two corpora are tagged with the semantic categories according to the UCREL

<table>
<thead>
<tr>
<th>Item</th>
<th>O1</th>
<th>%1</th>
<th>O2</th>
<th>%2</th>
<th>LL</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>W3</td>
<td>261</td>
<td>0.63</td>
<td>1007</td>
<td>0.93</td>
<td>−320.37</td>
<td>geographical terms</td>
</tr>
<tr>
<td>G2.1</td>
<td>351</td>
<td>0.84</td>
<td>57</td>
<td>0.12</td>
<td>+280.53</td>
<td>crime, law and order</td>
</tr>
<tr>
<td>G1.1</td>
<td>649</td>
<td>0.94</td>
<td>269</td>
<td>0.47</td>
<td>+222.27</td>
<td>government etc.</td>
</tr>
<tr>
<td>Z6</td>
<td>337</td>
<td>0.81</td>
<td>146</td>
<td>0.32</td>
<td>+107.98</td>
<td>negative</td>
</tr>
<tr>
<td>N1</td>
<td>345</td>
<td>0.82</td>
<td>971</td>
<td>0.90</td>
<td>−93.82</td>
<td>numbers</td>
</tr>
<tr>
<td>E3-</td>
<td>159</td>
<td>0.39</td>
<td>41</td>
<td>0.08</td>
<td>+92.58</td>
<td>calm, violent and angry</td>
</tr>
<tr>
<td>S8+</td>
<td>217</td>
<td>0.51</td>
<td>406</td>
<td>0.80</td>
<td>−70.49</td>
<td>helping or hindering</td>
</tr>
<tr>
<td>A7+</td>
<td>464</td>
<td>0.85</td>
<td>289</td>
<td>0.48</td>
<td>+67.37</td>
<td>definite (+ modals)</td>
</tr>
<tr>
<td>Q2.1</td>
<td>879</td>
<td>0.99</td>
<td>622</td>
<td>0.89</td>
<td>+34.05</td>
<td>Speech and communicative</td>
</tr>
</tbody>
</table>
Semantic Tagset system, I obtain a list which compares the frequencies of the semantic categories employed in the CP and the CD respectively. Table 5 below shows only the top 8 semantic categories, which exhibit significant differences in frequency between the two corpora, and which are ranked in terms of the absolute value of the LL parameter.

The first category, W3, whose LL value between the two corpora is as high as 320.37, designates the geographical terms. The conspicuous overuse of this category of words in the CD than in the CP is most likely to be attributed to the fact that the former devotes much space to describing in great detail the activities of Lien Chan, who visited Nanjing, Xi’an, Beijing and Shanghai successively during his stay in China. The CP, however, does not mention much detail of what he did during his stay in China, thus, the geographical terms in its news reports are less frequent than those in the CD.

Following this category are four categories, which occur much more frequently in the CP than in the CD, and which are to be discussed subsequently.

The first category with the LL value of 280.53 is G2.1, which denotes crime, law and order, and semantically, this category is closely related to that of E3- (with a rather high LL value of 92.58), indicating calm, violent and angry. The two tables below exemplify some words that belong to each of these two categories respectively.

**Table 6.** The selected words of G2.1 in the CP

<table>
<thead>
<tr>
<th>Item</th>
<th>Category</th>
<th>Raw Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>police</td>
<td>G2.1</td>
<td>89</td>
<td>0.20</td>
</tr>
<tr>
<td>law</td>
<td>G2.1</td>
<td>75</td>
<td>0.18</td>
</tr>
<tr>
<td>legislative</td>
<td>G2.1</td>
<td>26</td>
<td>0.06</td>
</tr>
<tr>
<td>policemen</td>
<td>G2.1</td>
<td>24</td>
<td>0.05</td>
</tr>
<tr>
<td>security</td>
<td>G2.1</td>
<td>22</td>
<td>0.05</td>
</tr>
<tr>
<td>legal</td>
<td>G2.1</td>
<td>19</td>
<td>0.04</td>
</tr>
<tr>
<td>rule</td>
<td>G2.1</td>
<td>11</td>
<td>0.03</td>
</tr>
</tbody>
</table>

**Table 7.** The selected words of E3- in the CP

<table>
<thead>
<tr>
<th>Item</th>
<th>Category</th>
<th>Raw Frequency</th>
<th>Relative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>violence</td>
<td>E3-</td>
<td>27</td>
<td>0.06</td>
</tr>
<tr>
<td>whip</td>
<td>E3-</td>
<td>22</td>
<td>0.05</td>
</tr>
<tr>
<td>force</td>
<td>E3-</td>
<td>16</td>
<td>0.04</td>
</tr>
<tr>
<td>clashes</td>
<td>E3-</td>
<td>10</td>
<td>0.02</td>
</tr>
<tr>
<td>riotous</td>
<td>E3-</td>
<td>9</td>
<td>0.02</td>
</tr>
<tr>
<td>attack</td>
<td>E3-</td>
<td>5</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Among all the words in the above two tables, my analysis focuses on police which occurs 89 times with the relative frequency of 0.20% and violence which occurs 27 times with the relative frequency of 0.06%. Collectively, the concordances of these two words suggest how some alternative facets of Lien’s visit are represented in the CP’s news reports.

The concordances of police in the CP
1 …umbrellas and knives and let off fire crackers. Police first arrested two taxi drivers, Chen Ming-…
2 …charged with violence after police painstakingly reviewed video footage and photographs of the…
3 …they would be arrested immediately. Meanwhile police arrested two more suspects, a sling-shot…
4 …investigate. Police believe the same people were stirring up violence in the opposition protests…

The concordances of violence in the CP
1 …pro-independence lawmakers and a radio host in inciting the violence. Hundreds of supporters…
2 …secret triad organizations were stirring up violence, following the clash which saw hundreds…
3 …exercise his public authority to respond properly when the violence broke out,” Lin said, refer…
4 …pro-reunification activist Lin Ta-ching were also charged with violence after police painstaking…

A close analysis of the concordances of these two words points to the fact that most of them are used to describe the riots in Taipei triggered by Lien’s visit to China because some pro-independence Taiwanese protested against his visit before his departure and the police were mobilized to prevent any riot from happening. Given the complex political situation in Taiwan before Lien’s visit, for example, the long standing hostility between China and Taiwan, the division among the Taiwanese between the pro-independence and the pro-reunification (Kuo and Nakamura, 2005), the consistent policy of the KMT striving for the eventual reunification of Taiwan with China and more importantly, the recent ratification by China of the Anti-Secession Law which was labeled by Chen Shui-bian as an “aggression” law, Lien’s visit received diametrically opposite treatments from two groups of Taiwanese: while the supporter of reunification consider his visit as a sensible move to better the relations between China and Taiwan, the supporters of independence consider it as an act of betrayal. Since these violent incidents impinge significantly on the domestic social life in Taiwan, it is natural that the CP present a detailed representation of them, which in turn, gives rise to a very high frequency of such words as police, violence and so on.

In contrast, these violent incidents are not reported by the CD, nor by any other media in China. The most plausible reason that the CD ignores these details
may be that these facts will blemish the successfulness of the policy of the peaceful reunification promoted by the CPC, which also plans to undermine the pro-independent DPP by inviting Lien Chan to visit China. The absence of these facts in the CD lends convincing support to the observation by Galasinski and Marley (1998), which point out that the discursive construction of the political reality by the media entails the decision as to which aspects to include. Usually the news agencies report on only those aspects of the reality that they consider significant in the way that is compatible with their political and ideological stances.

Next to the category of G2.1, comes the third category of G1.1, which represents those words whose meanings relate to the category of government in the UCREL Semantic Tagset system and among all 649 tokens of words under this category, president has 293 tokens, and government 226 tokens. As discussed in Section 3.1, the concordances of these two words are associated with the reactions of the ruling DPP towards Lien’s visit.

The fourth category which displays a significant overuse in the CP relative to the CD is that of Z6, which indicates negative words. Wmatrix shows that among altogether 337 tokens of words under this category in the CP, there are 220 occurrences of not, and its concordances show that most often, it is used on the occasions where the ruling DPP refuses to accept the negotiation terms provided by Beijing.

The concordances of not in the CP
1 …political sweeteners, saying it could not accept Beijing’s attack on President Chen Shui-bian’s
2 …conspiring to diplomatically isolate Taiwan for a long time and he could not see any good will
3 …He urged Beijing leaders not to miscalculate the cross-strait situation, especially mainstream
4 …any substantive concession if Lien does not win authorization from the government for his talks

Bush (2005) observes that the most thorny issue which still prevents the CPC and the DPP from making any substantive progress towards the peaceful negotiation is that both sides have different interpretations of the political status of Taiwan: while the CPC insists that any negotiation between China and Taiwan must be based on the “One China Principle” which claims that Taiwan is an inseparable part of China, the DPP contends that Taiwan and China are two different countries on an equal footing. This dispute is also evidenced in some concordances (examples 1, 2 and 3) listed above, in which the DPP expresses the position that they will not accept China’s offer to negotiate on the terms of the one China principle. Although there are only 92 occurrences of not in the corpus of the CD, its concordances show that this word is used in diverse senses with no specific implication related to the political standpoint of the DPP.
Discussions and conclusion

In Section 3, by comparing the two corpora of the news reports concerning the same event (i.e. Lien Chan’s visit to China) in terms of word frequency, part-of-speech frequency and semantic category frequency, I come to the following conclusion: the lexical choices are manipulated by both the *China Daily* and the *China Post* in such a way that Lien’s visit is represented quite differently, with the result that some facets of this event are omitted and some others highlighted by each newspaper. For example, the CD depicts Lien’s visit as a successful and historic moment that will make tremendous contribution towards better relations between China and Taiwan; therefore, the news reports in the CD give a full account of his specific activities in China, the warm treatment by the Chinese government and its people, and the positive assessment of his visit from various sources. In the case of the CP, however, the news reports present a more comprehensive but less optimistic picture of this event. According to the corpus data, apart from only a small number of tokens of *historic* that emphasize the importance of his visit, much more space in the CP is devoted to describing such details as how the ruling DPP negatively responded to Lien’s visit, the seriousness of the riots occasioned by his visit and the mixed reactions towards this event among the local Taiwanese. Another sharp contrast with the CD is that the CP makes no mention of the details of the activities of Lien Chan during his stay in China.

With this conclusion, the question worth considering is why there are such obvious discrepancies in the reporting of the same event. Despite the fact that these two newspapers are published in China and Taiwan respectively, they are not much different with respect to the political status of Taiwan and its relations with China. For instance, both uphold the eventual reunification of Taiwan with China, although the CP also hopes to maintain the status quo between China and Taiwan. I suggest that what results in the discrepancies in the reporting of Lien’s visit may be related to other socio-political factors which influence the news reporting, among which, the most likely is the difference between China and Taiwan in the political and institutional environment in which the media operate. In fact, the vital importance of this factor in the media discourse was noticed long time ago by Thompson (1990), who remarks that an analysis of texts would be a superficial one without reference to the political, historical and social conditions within which the object of analysis is produced, constructed and received. In the earliest version of the critical discourse analysis expounded by Fairclough (1989), discourse is considered as a social practice which involves the institutional conditions in which language is produced. Later, Fairclough (1998: 145) observes that “the analysis of discourse practice is probably best thought of as actually a complex of different sort of analysis, including more discursive aspects of institutional process (e.g.
practices of producing TV news programs), as well as sociocognitive aspects of discourse processing”.

With the significant socio-institutional role in determining the discourse as articulated as above, it is argued that the differences between China and Taiwan in the functions and working mechanisms of the media in general give rise to the different representations of Lien's visit by the CD and the CP. As observed by several previous studies (Fang, 1994, 2001; Kluver, 1996; Kuo, 2001), the fundamental function of the media (including newspapers) in China is to disseminate the ideas of the government and to report those facts that the CPC decides to have the people know, hence, “it is assumed that the content of the PRC newspapers will reflect the dominant ideology of the state” (Fang, 1994). Following the recent commercialization of the press in China, this vital role of “throat and tongue” assigned to the media remains unchanged although the contents appear more appealing and realistic (Liu, 1998). As one of the leading news institutions in China, and given the sensitivity of Lien's visit in the cross-Strait relations, it is natural that the CD shoulders the responsibility of depicting this event as a momentous one that receives the nation-wide plaudits, in order to highlight the successfulness of the policy of peaceful reunification carried out by the CPC. Meanwhile, the CD intentionally omits the details of those opposing voices from the supporters of pro-independence and the ruling DPP in Taiwan, which are considered by the CPC as challenging its “One China” policy and hence not worth releasing to the audience in China.

As to the institutional environment in which the media work in Taiwan, however, it is markedly different and more complex. Beginning with the retreat of the KMT to Taiwan in 1949, the following nearly four decades witnessed a stringent control of the media by the KMT under the martial law to legitimate its ideology of anti-communism (Chen, 1998). In the 1980s, Taiwan saw some democratic reforms introduced, together with the ending of the authoritarian regime in 2000 when the DPP took over presidency. These political changes brought forth a series of significant transformations in the press industry in Taiwan, for example, the legitimating of the promotion of political independence from China, more freedom enjoyed by the press to release information or comment on the performance of any government official, the diversification of the contents of newspapers, more attention accorded to the public opinion, and a more balanced approach to reporting political events (Wang and Lo, 2000). These tendencies are also manifested in the CP’s reporting on Lien Chan’s visit. In general, the CP presents a more comprehensive picture of this news event than the CD does. With its pro-unification ideology, the newspaper, on the one hand, expresses its support to and approbation of Lien's visit both in its editorials and some news reports; on the other hand, it also devotes a considerable portion of space to reporting the domestically important
incidents resulting from Lien's visit, for example, the attitude and reactions of the ruling DPP, the reactions among the local Taiwanese, particularly those violent protests before Lien's departure to China, thus presenting a more balanced and comprehensive account of the various aspects of Lien's visit.

Notes

* The original paper was presented at the 5th International Conference on Chinese Sociolinguistics held at Peking University from 8th to 10th December, 2006. I am grateful to Peter K. W. Tan for giving comments on my earlier drafts. Especially, I am grateful to Saihua Kuo for her spending much time discussing with me about my paper, mailing reference materials to me, and providing many valuable suggestions on how to improve my previous versions. Thanks should also be extended to Vincent B. Y. Ooi for providing me the access to the WMatrix software.

1. Although Lien Chan's visit began on 26th April and ended on 4th May in 2005, the continuous reporting of this event by both the CP and the CD lasted much longer than the duration of his visit to China. Therefore, the time span covered by the news reports of the two corpora is longer than his actual stay in China.

2. Due to the limit of space, among all the concordances of the node word produced by the computer software, I select only four in order to illustrate the contexts in which it occurs, except when the total number of the concordances is less than four. Moreover, it is also the usual practice in the corpus linguistic literature to present truncated concordances of the node word.

3. The introduction of this system is available at http://www.comp.lancs.ac.uk/ucrel/claws7tags.html.

4. The introduction of this system is available at http://www.comp.lancs.ac.uk/ucrel/usas/.

References


