How Chinese Native Speakers Handle Written Style Material in Reading and its Application in Second Language Acquisition

Yongfang Zhang
The University of Akron

In this paper, a new approach was employed to investigate the relationship between written Chinese and spoken Chinese. In a pilot study, how Chinese native speakers handle written style materials was investigated and compared with Chinese L2 learners. Preliminary findings reveal that both groups are aware of the variation of written levels in different genres. However, the written elements selected by Chinese native speakers include low frequency content words, content words related to the text theme, as well as idioms and poetic sentences. In addition, native speakers can interpret the written elements through their strong vocabulary knowledge, lexical network, and ability to apply their schema in reading. In contrast, L2 learners select more function words, classical words and idioms as written elements, and they lack the ability to interpret. The findings bear direct application on linguistic research and second language acquisition, and call for further research.

0. Introduction

There are many researches on the relationship between spoken language and written language in English. However, such research is rare in Chinese (Zhang and Yang 1994). As Zhang and Yang (1994) and Zhang (2000) pointed out, many researches in Chinese only analyzed a single or a few features, and a systematic study of spoken Chinese and written Chinese is still missing. In addition, many analyses were relied on the researchers' personal judgment and thus the reliability of their findings is questionable. Moreover, many findings were obtained from small scale corpus, and are hard to be generalized.

This paper aims to propose a new reception approach for investigating the relationship between spoken Chinese and written Chinese from the analyses of how Chinese native speakers handle written materials. In a pilot study, I investigated how Chinese native speakers rated different genres in terms of written levels, what elements in texts they picked as written elements, why they categorized the elements into written, and how they interpreted the written elements. In addition, this research can also benefit Chinese second language (L2) acquisition. When we can understand how Chinese native speakers handle the written materials, we can guide Chinese L2 learners towards a

conscious and effective learning of both spoken and written Chinese. Therefore, this pilot study also investigated how advanced-level Chinese L2 learners handle written Chinese.

In this article, I will first review the findings in spoken language and written language in general, and summarize the limited research on Chinese. Then I will describe my pilot study on how Chinese native speakers handle written materials in reading, followed by the description of my investigation on how Chinese L2 learners handle written materials in reading and the comparison between the two groups of subjects. In the final section, I will draw my conclusions and point out the applications of my preliminary study in Chinese linguistic research and second language acquisition.

1. Research in General

There are many researches on the relationship between spoken language and written language in general. Spoken language and written language share common core such as similar sentence patterns (Blankenship 1962) and are related closely from historical perspective. In addition, many genres can have both spoken feature and written feature (Tannen 1982). Therefore, most researchers treat them as a continuum instead of a dichotomy (Tannen 1982). Scholes (1999) treated spoken language and written language as variant ways of coding a single underlying language, though they have some differences. Georgakopoulou and Goutsos (1997) pointed out that there are varying degrees of variations and some are technology-mediated.

Researchers have also analyzed the differences between spoken language and written language. According to Tannen (1982: xi), the difference between spoken language and written language is "broadly interdisciplinary". With the linguistics as the core, it encompasses anthropology, psychology and literature. The spoken language and written language are also different in terms of interaction to linguistic context, information density (Tannen 1982), formality and purpose of language (Biber 1988). Basically, the previous studies on spoken language and written language can be categorized into two groups. The first group, which is symbolized as the Producing Group, refers to the researches focusing on the process of generating the two variations of the language (spoken and written). The second group, Product Group, focuses on the product and the reception of spoken and written language.

Producing Group. Researchers have found that spoken language uses phonics, and thus is done through paralinguistic channels with emotional or interpersonal dynamics, while written language uses graphs and thus uses both content and verbal channels (Gumperz et al. 1984; Kay 1977; Tannen 1982). Related to this aspect, spoken language is temporal and evanescent while written language is spatial and stable (Chafe 1982; Li and Thompson 1982). From the receiver's respect, listener uses ears as sensatory organ and relies on memory to retain information, while reader uses eyes as sensatory organ and can halt reading and turn back to earlier passages (Jahandarie 1999). In most cases, spoken language is produced impromptu while written language is planned

(Georgakopoulou and Goutsos 1997; Stubbs 1980). Chafe (1982) has found that the information flow in spoken language is from one idea to the next at a rate of one every two seconds while in written language there is plenty time to move and integrate a succession of ideas into single linguistic whole. Therefore, researchers (Chafe 1982; Georgakopoulou and Goutsos 1997; Stubbs 1980) claimed that written language is decontextulized and thus detached, while spoken language is context-bound and thus involved. The speaker interacts with listener and has more interpersonal involvement, while during writing there is only one-way information flow from the writer to the reader (no turn taking) and the reader is detached from the process (Chafe 1982; Gumperz and Gumperz 1981; Kay 1977; Tannen 1982). Vygotsky (1962) described spoken language as dialogue while written language as monologue.

Product Group. As a result of the differences in the producing process, spoken language and written language also show differences in terms of product. Spoken language uses simple structures and is condensed like telegraph (Li and Thompson 1982), while written language uses more complicated structures and thus is more complex, elaborate and explicit (Georgakopoulou and Goutsos 1997; Stubbs 1980). In addition, spoken language has more redundancy and is fragmented with loose structure, while written language is more integrated with complete structure (Chafe 1982; Christensen 1994; Zhang and Yang 1994; Zhang 2000). Moreover, spoken language and written language show difference in dimensions such as phonology, lexicon, syntax and discourse. While spoken language uses prosodic devices which is evanescent and requires more short term memory, written language uses punctuation which is permanent and thus requires diverse types of memory (Halliday 1989; Jahadarie 1999). Compared to spoken language, written language uses more attributive adjectives (Blankenship 1962) and has varied vocabulary, and fewer words referred to speaker (Devito 1966). Written language uses more subordinative clauses and devices (Chafe 1982). In addition, verbs are tended to be used more in oral discourse while nouns to be used more in written discourse (Blankenship 1962; Halliday 1989). However, Halliday (1989) proposed to use the degree of density as the criteria: written language is lexically dense with simple grammar while spoken language is grammatically dense with simple vocabulary.

2. Research in Chinese

Most Chinese scholars treat spoken Chinese and written Chinese as *yuti* 语体 'two register variations' (Liu 2005). Register is defined as the language used by communicators in specific contexts to complete different communicative purposes. Therefore, it is a functional variation of the general language and the variation reflects on linguistics (Wang 1987). Influencing factors in register include communicational motivation, purpose, settings, producer agents, and linguistic expression. The differences in these factors define subtypes of spoken Chinese and written Chinese. A most cited categorization was proposed by Wang (1987). According to him, spoken Chinese is

divided into *suiyi tanhua* 随意谈话 'causal talk' and *renzhen tanhua* 认真谈话 'serious talk', while written Chinese is divided into *kexue yuti* 科学语体 'scientific register', *yishu yuti* 艺术语体 'literate register', *zhenglun yuti* 政论语体 'political argument register' and *gongwen shiwu yuti* 公文事物语体 'official document register'. Another categorization is to divide spoken and written Chinese into *pingshi yuti* 平实语体 'vernacular' and *yishu yuti* 艺术语体 'elegant style'.

Most Chinese discourses are stretched in the continuum between the spoken polar and the written polar. Spoken Chinese discourse can have written language features while written Chinese discourse can also have spoken features. In *A grammar of spoken Chinese,* Chao (1968) listed different types of spoken discourses from artificial to authentic, in regarding to genre, context (formal vs informal) and preparation (planned vs unplanned). These types covered spoken Chinese with different degree of formality and written features, and thus implied the integration between spoken Chinese and written Chinese. According to Chao (1968: 18), the most artificial spoken Chinese is "monologues read stiffly from manuscript" while the most authentic one is "things said in response to some situation or to something in the stream of consciousness". More artificial spoken language discourses have more written features. Chao (1968) also indicated that the difference between spoken and written Chinese is also related to formality of the setting.

Different register variations show both similarities and differences in linguistics and pragmatics with linguistics as the core. As variations of the same language, spoken Chinese and written Chinese share common core which is legible to and used by the majority people using Chinese. According to Gu and Zhou (1989), common core includes fundamental grammar and vocabularies with high frequency. The most often examined differences in linguistic features are in the dimension of lexicon. Chao (1968: 18) pointed out that "different types of discourse make more difference in sentence structure and in the expressive elements than in other aspects of grammar". Generally, the majority vocabulary in Chinese can be used in all kinds of variations. However, some vocabulary can only be used in certain register. Spoken Chinese has small range of lexicon. It often uses high frequency words, dialect words, slang, proverbs and *xiehouyu* 歇后语 'a two-part allegorical saying'. It seldom uses classical Chinese expressions. In contrast, written Chinese has a varied vocabulary. It uses classical Chinese (Li and Thompson 1982), idioms, more attributive adjectives, and fewer words referred to speaker and listeners.

3. Investigation on Chinese Native Speakers

In this section, I will describe my investigation on how Chinese native speakers handle written materials in reading. Specifically, I investigate their ratings of different genres in terms of written levels, their selections of written elements in texts, their reasons to categorize items into written, and their interpretation of the written elements. Due to the preliminary characteristics of the pilot study and the small number of subjects

in some part of research, the findings are very tentative and the generalization is not suggested beyond this study. Further studies must be done for the generalization.

Pilot Study. The pilot study was conducted as a grounded survey of 160 Chinese native speakers following the interview of 13 different Chinese native speakers during the summer of 2007. All subjects were college students. First, 13 subjects were asked to read 5 short news articles and rate the written levels of each article. Then, they were asked to read a long prose and rate the written level, underline the written style elements and interpret the written elements in spoken language. They were also asked to provide reasons why they categorized the underlined elements to written Chinese.

The following extensive study surveyed 160 new subjects. First, the 160 subjects were asked to rate the written levels of different genres. The genres were divided into three categories: newspaper articles, novels and internet articles in blogs. Newspaper articles included articles on current affairs, sports, entertainments, living and editorial sections. The online versions of all these genres were also included. Novels included classic novels, contemporary novels, modern novels and foreign novels translated into Chinese. Internet articles in blogs included celebrity blogs, vernacular blogs and blogs written by classmates. Then, they were asked to read the highest rated news article elected by the previous 13 subjects and underline the written style linguistic elements.

All texts were selected from the Chinese L2 textbooks for intermediate to advanced level. The five short news articles have 100, 135, 118, 105, and 104 characters, with an average of 112.4 characters per article. Throughout the study, the subjects were asked to rate using a 7-point scale where 1 means "the most spoken" and 7 means "the most written". 7-point scale is used because it has a clear middle point. In addition, between the middle and end points there are just two choices, which capture variations in opinions without presenting too many choices or too few.

Rating of Genres. The ratings have shown that these Chinese native speakers are aware of the variations of registers in different genres. The 160 Chinese native speakers rated current affair news the genre with the highest written style (6.04). The next highest was classical novel (s=5.94). The newspaper articles on sports, entertainment, living and editorial sections were rated to be 3.97, 3.23, 2.87 and 4.68, respectively; the contemporary novels, modern novel and translated foreign novels were rated to be 4.62, 3.97 and 4.63, respectively; the articles in celebrity blogs, vernacular blogs and classmate blogs were rated to be 4.08, 3.11, and 2.42 respectively; the online articles on current affairs, sports, entertainment, living and editorial were rated to be 5.52, 3.61, 2.97, 2.78 and 3.92 respectively.

My observations agree with Biber (1988) and Georgakopoulou and Goutsos (1997). Biber (1988) has proposed that variations between spoken and written languages are related to purpose of the language. Among the three types of blogs, the celebrity blog has highest level of written style. My explanation is that the celebrity's blogs are oriented

to attract more attentions and they intentionally use more written style to be elegant. Georgakopoulou and Goutsos (1997) have asserted that some variations are technology-mediated. The finding that the ratings of written style of online newspaper articles follow the same order of paper based newspaper articles, yet to lower level of written style, implies the differences in written level varied by media.

Among the novels, classical novel is rated most written, followed by translated foreign novel and contemporary novel, and then modern novel. This agrees with the observations of Li and Thompson (1982) that classical Chinese is used in written Chinese only. In addition, this also implies that the subjects believe that only the best novels in the history can survive and still be widely welcomed today, and tend to consider them more elegant and thus written.

Short News Articles. The ratings by 13 native speakers have demonstrated that different texts belonging to the same genre can have variations of registers. Five short news articles were rated to be 5.5, 3.92, 4.69, 4.19, and 4.11, respectively. In addition, these speakers rated news more written since most ratings were above the mid-point 4 in the 7-point scale. Moreover, the written level not related to the length of the article since the article with the most characters was rated the least written while the article with the fewest characters was rated the most written. However, the real factors were not investigated in this study.

The low occurrence frequency content words endow the article more written style than function words. As shown in the grounded survey of 160 new subjects, more content words and phrases were considered as written elements. The content words which were picked by more than 1/3 subjects include 竣工 *jungong* '(project) be completed' (f=136, p=85%)，交付 *jiaofu* 'submit' (f=130, p=81.25%)，兼有 *jianyou* 'also have' (f=115, p=71.88%)，库容 *kurong* 'storage capacity' (f=113, p=70.63%)，效益 *xiaoyi* 'profit' (f=105, p=65.63%)，综合 *zonghe* 'comprehensive, synthetical' (f=102, p=63.75%)，典礼 *dianli* 'ceremony' (f=92, p=57.5%)， 使用 *shiyong* 'use' (f=87, p=54.38%), 日前 *riqian* 'the other day, a few days ago' (f=85, p=53.13%)，调节 *tiaojie* 'adjust' (f=80, p=50%), 位于 *weiyu* 'be located, be situated' (f=65, p=40.63%)，正式 *zhengshi* 'formally' (f=60, p=37.5%). Here *f* is defined as the number of subjects considering the word as a written element, while *p* is the ratio of the *f* value over the number of subjects (160). In contrast, only two functions words were picked by more than 15% subjects (p>15%). One is the conjunction word 并 *bing* 'and' (f=46, p=28.75%), and the other is an auxiliary verb 可 *ke* 'can' (f=26, p=16.25%).

In addition, all of the selected content words are low frequency Chinese words as none of them are included in *Xiandai hanyu pinlü cidian* 现代汉语频率词典 *'A Frequency Dictionary of Modern Chinese'*. The observations agree with previous findings that written Chinese has bigger range of lexicon while the spoken Chinese is often limited to the high frequency words (Chao 1968). A small corpus study on written

words in two dictionaries (Li 1998; Lü 1994) reveals that some bi-character words tend to become one character in written Chinese, such as 可以 *keyi* 'can', 应该 *yinggai* 'should', 比较 *bijiao* 'comparatively', 本来 *benlai* 'originally', 因为 *yinwei* 'because' are used in spoken Chinese while 可, 应, 较, 本 and 因 are used in formal written style. However, in my study, 并 *bing* 'and' and 可 *ke* 'can', the abbreviated forms of bi-character words 并且 *bingqie* and 可以 *keyi*, were picked by small number of subjects. I believe the reason is the high frequency of these two abbreviated words, as both are included in *Xiandai hanyu pinlü cidian 现代汉语频率词典*

Long Prose. 13 Chinese native speakers rated the long prose, g*uiyu fanxiang* 鲑鱼返乡 'Guiyu's home return' to be more written with an average rating of 4.46. Overall, they selected 132 words, idioms and poetic sentences as written style elements, among which 54.5% was selected 1 to 3 times while the rest 45.5% was selected 4 to 11 times. Among the 407 total frequencies of written style elements, 179 (43.98% of the total frequency) is related to the theme of the prose such as terms for hometown, homesick and return to hometown across the words, idioms and poetic sentences. Poetic sentences were picked 50 frequency (12.3%) while idioms were picked 34 (8.35%). None of the words is included in *xiandai hanyu pinlü cidian.*

The approaches to interpret written words reflect native speakers' strong vocabulary knowledge, lexical network and application of schema in reading. Chinese native speakers provided the following approaches to explain the written style words. (1) correspondence between classical words and today's words (e.g., explaining 度 *du* 'spend' as 过 *guo*,非 *fei* 'no' as 不是 *bushi*, 即 *ji* 'that is' as 就 *jiu*), (2) changing one character in the words (e.g., explaining 驱车 *quche* 'drive the car' as 开车 *kaiche* 'drive the car', 确然 *queran* 'indeed so' as 的确 *dique* 'indeed', 糅合 *rouhe* 'mix' as 结合 *jiehe* "combine, link'), (3) extending the abbreviated one character word into bi-character words (e.g., explaining 因 *yin* 'because' as 因为 *yinwei,* 已非 *yifei* 'already not' as 已经不是 *yijing bushi*), (4) deleting a character in the word (e.g., explaining 苦涩 *kuse* 'bitter and astringent; agonized' as 苦 'bitter, suffering',浓溢 *nongyi* 'dense and overflow, excessively dense' as 浓 'dense', 圆周 *yuangzhou* 'circumference' as 圆 'circle'), (5) deleting rhetoric or cultural portion (e.g., explaining 三丈之内 *sanzhang zhinei* 'inside three zhang's distance' as 离得很近 *lide henjin* 'very close',游子回归 *youzi huigui* 'return home by man travelling far away from home' as回国了 *huiguo le* 'go back to one's country' or 回来了 *huilai le* 'come back'), (6) using words with similar meaning in the theme (e.g., *xiangtu qingjie* 乡土情结 'feeling knot of one's hometown' is explained as *lianjia* 恋家 'homesick, long for home').

Native speakers' categorizing written elements and the expressions used in their interpretation reflect these native speakers' understanding of the discrepancy between

written Chinese and spoken Chinese. In general, there are three types of reasons. The first reason is that some elements are seldom used in spoken Chinese (e.g., "口语中不常用" 'not often used in spoken Chinese', "很少这么用" 'very rarely used in this way', "不大用" 'not frequently used', "不这么用" 'not used in this way', "不会这么说" '(people) do not say this way', "不用说" 'no need to say'). When explaining the words, the subjects used phrases such as "平常都说" 'normally (people) all say', "一般说" 'in general (people) say', "大家都说" 'everybody says'. Some subjects also mentioned that there is no equivalence in spoken Chinese. Examples include "口语中找不到对等的词语", 'Could not locate a corresponding word in spoken Chinese', "口语里没有这个词或者口语里借用了这个词" 'there is not such as word in spoken Chinese or spoken Chinese borrowed this word (from written Chinese)'. The subjects used such phrases to interpret: "直接说明, 就这么简单" 'explain directly, and that is simple', "简单地" 'simply'.

The second explanation reflects the difficulty in understanding written Chinese. Answers include: "不明白" 'do not understand', "不理解" 'do not understand', "好像很绕口" 'seems very unsmooth'. On the contrary, subjects felt that spoken language is not hard to understand, as see in one subject's explanation "我觉得读起来感觉比较顺，然后他说的话吧，你读完以后基本上就能理解，比较可以体会作者写这篇文章的意思。口语化还比书面语的成分还要多一些。"[1]

The third type of explanation reflects the concept that written Chinese is more elegant and specialized, use more classical words, and often used by scholars. Examples include "古文里就有" 'existed in classical essays', "古文的表达方法" 'expressions from classical essays', "文绉绉的" 'genteel', the specialized (e.g., "挺专业" 'very specialized'), and elegant ("有文采"). 12 subjects categorized poetic sentence into written Chinese. Some subjects elaborated their reasons such as "很少在口语中突然引经据典，然后引一段诗，除非那些教授" 'in spoken Chinese people rarely suddenly use classical sentences and stories, and then quote a paragraph of poem. The only exception is those professors。), "如果平常说话根本不可能出现这样的句子" 'such sentences will not appear in normal speech at all', "他（指作者）是知识分子，他会用一些非常有中国韵味的诗句来表达内心的心情。我觉得这样用非常有文采" 'he is a scholar. He will use some poetic sentences with deep Chinese aroma to express his feeling. I feel such a use is very elegant'. In their explanation, the subjects used phrases such as "现在意思是" 'the meaning now is', "说得比较罗嗦一点" 'to say a little bit wordy', "说得没有诗意一点" 'to say with less poetic flavor'.

---

[1] The following is a rough translation of the explanation: I feel that my reading of this article comparatively smooth, after reading you can generally understand the author's speech, you can understand the meaning the author wrote this article. There are more spoken style elements than written style elements in this article."

In addition to the discrepancy, subjects also noticed the possibility of changing the use of linguistic elements from one variation into another. 10 subjects categorized idioms into written Chinese. Some subjects noticed the use of idioms in spoken language. They felt that idioms originate from written Chinese but if after being used in spoken Chinese often or for a long time, people will feel them more spoken. Their explanations include: "成语还是经常会应用的吧 (在口语中)" 'idioms are often used (in spoken)', "遍体鳞伤是口语，还挺常见的" '*bianti linshang* is spoken, and can be seen quite often (in spoken)', "斩钉截铁，遍体鳞伤，都是书面语吧，但是现在口语都用了" '*zanding jietie*, *bianti linshang*, are written language, however, now they are used in spoken language as well', "我觉得像成语之类的，有些词用得多了，就会觉得是口语，用得少了，就会觉得是书面语。" 'I feel that words like idioms, if being used more, people will categorize them into spoken, if being used less, people will categorize them into written'.

To sum up, in this preliminary study, I investigated how Chinese native speakers rated different genres in terms of written levels, what elements in texts they picked as written style elements, why they categorized the elements into written, and how they interpreted the written elements. Findings include: (1) the Chinese native speakers in my pilot study are aware of the variation of register in different genres. News and classical novels are rated very written, (2) in addition to idioms and poetic sentences, low frequency content words and content words related to theme are often considered as written elements, (3) the subjects have strong vocabulary knowledge, lexical network and schema of the text theme, and are able to interpret the written elements.

4. Investigation on Chinese L2 Learners

In this pilot study, 16 advanced-level Chinese L2 learners were surveyed. All these learners have finished at least three years of formal college instruction of Chinese in the United States. These learners were first asked to rate the written levels of genres. Then they were asked to read three of the five short news articles and rate the written levels of each. Finally, they were asked to read g*uiyu fanxiang* 鲑鱼返乡 'Guiyu's home return', rate the written level, underline the written style elements and interpret them. This aims to detect L2 learners' understanding of and competence in written Chinese.

Rating of Genres. The ratings demonstrated that Chinese L2 learners are also aware of the variations of registers in different genres in the similar manners with native speakers, despite some differences. As illustrated in Figure 1, the ratings by native speaker and L2 learners has very close pattern. The L2 learners rated classical novel, which is the second most written genre by native speakers, to be most written (7.00). The most written genre by native speaker, current affair news, is rated as the third most written by L2 learner. The biggest different rating between the native speaker and L2 learner is about the celebrity blogs. When the L2 learner rated it as the second most

written genre, the native speaker rated the fifth most written genre in Figure 1 with a rating of 4.08, which is slimly over the middle point of the scale.

The newspaper articles on current affairs, sports, entertainment, financial, and editorial sections were rated to be 4.95, 3.89, 3.00, 4.30, and 4.34 respectively; the contemporary novels, modern novel and translated foreign novels were rated to be 4.75, 3.59, and 5.00, respectively; the articles in celebrity blogs, vernacular blogs and classmate blogs were rated to be 5.50, 3.00, and 3.50 respectively; the online articles on current affairs, sports, entertainment, financial and editorial were rated to be 4.84, 3.83, 2.88, 4.67, and 4.00 respectively.

In each categories of writings (news, novels and blogs), the L2 learner showed same opinion as native speakers. Chinese L2 learners also rated the genres in the category of novels in the same order as native speakers: classical novels > translated foreign novels > modern novels > contemporary novels. In the category of news articles, L2 learners also rated the current affair news to be most written, followed by editorial, sports and entertainment news. In the category of blogs, both native speaker and L2 learner rated the celebrity blog is most written. Although the L2 learner rated classmates blogs to be more written than vernacular blog, which is different from native speakers, both of them consider the genres are fairly spoken style as the rates are below 4, the middle points of the scale.
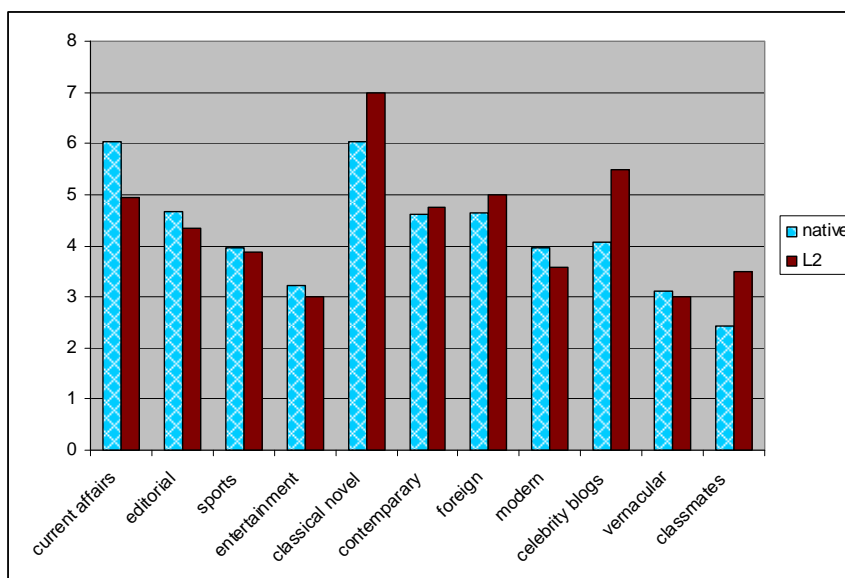

Figure 1. Comparison of written level of different genres as rated by native speaker (the first column) and L2 learners.

Short News Articles. The Chinese L2 learners showed very different opinions on the written elements from the native speakers. When reading the three short news articles, L2 learners picked more single-character abbreviated words, and classical character 为 *wei* 'be', as written elements. As showed in Figure 2, the most picked words were 为 (f=8, p=50%), 可 (f=7, p=43.75%),目前 (f=7, p=43.75%), 位于 (f=7, p=43.75%),并 (f=6, p=37.5%), 典礼(f=3, p=18.75%), 效益(f=3, p=18.65%), 综合 (f=2, p=17.5%), 竣工 (f=2, p=17.5%), 交付 (f=2, p=17.5%) and 兼有 (f=1, p=6.25%). Moreover, some Chinese L2 learners showed difficulties in identifying an element boundary, such as treating "工典礼后" as a phrase in "竣工典礼后" 'after the celebration of the completion'.
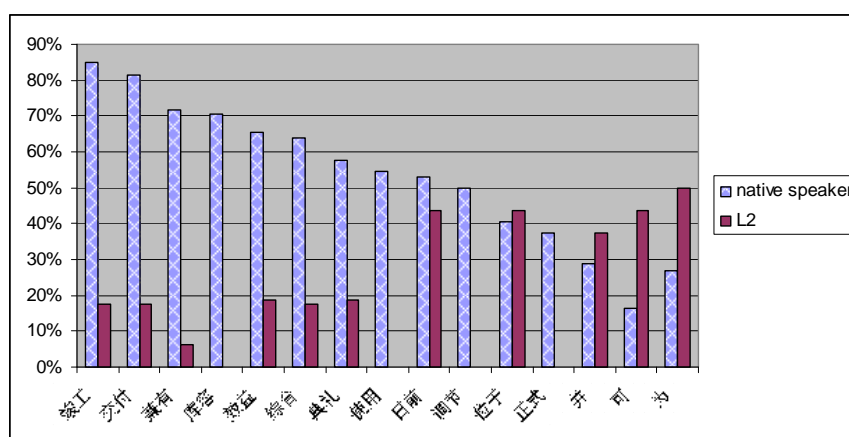

Figure 2. Comparison of the percentage of native speaker (the first column) and L2 learners (the second column) that considered the studied Chinese words as written elements.

Long Prose. Chinese L2 learners rated the long prose *guiyu fanxiang* 鲑鱼返乡 to higher written level than Chinese native speakers. While the native speaker rated the article to be 4.46 in average, the L2 learners rated it to be 5.5 in average. The density of content and vocabulary, and more use of idioms may explain the discrepancy between L2 learners' higher rating of the long prose than native speakers and their lower rating of the short news articles than native speakers.

Compared to native speakers, L2 learners picked more single-character function words as written elements. Among the total 122 frequency, they picked single-character words 45 times (p=36.89%) including 之 *zhi* 'of' (f=5), 而 *er* 'and' (f=3), 与 *yu* 'and' (f=2), 何 *he* 'what' (f=2), 却 *que* 'however' (f=2), 即 *ji* 'that is' (f=2), 已 *yi* 'already' (f=2), 无 *wu* 'no' (f=2), and 以 *yi* 'can' (f=1). The next highest one was idiom (f=15, p=12.3%), followed by words related to the theme (f=14, p=11.48%). Poetic sentences were only selected twice, which counted 1.64% of the total frequency.

In addition, Chinese L2 speakers showed great difficulties in interpreting the written elements in spoken Chinese. In most cases L2 learners did not provide their

interpretation of the written style elements. Moreover, Chinese L2 learners made some mistakes when identifying the boundary of words or phrases, such as identifying 糅 in 糅合 *rouhe* 'mix', 缠缱绻 in 纠缠缱绻 *jiuchan qianquan* 'fast tangle'，涩 in 苦涩 *kuse*，国度 in 回国度长假 *huiguo du changjia* 'return home country for a long vacation'. This type of mistakes indicates the lack of vocabulary knowledge, and more bottom-up reading, while native speakers focus on top-down reading. As native speaker subjects wrote: "阅读时以大义为主", "一般看文章的时候不会特别关注其中的一些句子，不会细抠一些句子。" ('during reading the reader should focus on the gist' and 'in general when reading articles, the reader will not pay special attention to some sentences, and will not analyze some sentences in details').

To sum up, Chinese L2 learners has very similar awareness of the variations of register in different genres. However, they show very different opinion from native speaker about written elements. L2 learns more often considered single-character function words and classical words (including idioms) as written elements. During the survey, the L2 learner also showed some difficulties in interpreting the written elements.

5. Conclusions and Applications

This preliminary research first revealed that different genres have different written style degrees, and both native speakers and advanced level L2 learners are awareness of this variation. News articles on current affairs and classical novels are considered to be very written, while contemporary novels and news articles on sports and entertainment are not very written.

However, native speakers and L2 learners showed very different understanding of written elements and different competence in handling the written elements. Native speakers considered low frequency content words and content words related to text theme, as well as idioms and poetic sentences as written elements. Moreover, the subjects have strong vocabulary knowledge, lexical network and schema of the text theme, and thus are able to interpret the written elements. In contrast, L2 learners considered more single-character function words and classical words as written elements, in addition to idioms. But their ability to interpret the written elements was missing.

This study proposed a new approach for linguistics research. This approach investigates the process of reading, instead of comparison between written style text and spoken style discourse, which is the products of reading. Though the means of survey, interview, and think aloud protocol, the subjects are asked to demonstrate their reading process, such as rating the written level of a text and linguistic elements, and the readers' interpretation of the written elements. This approach has shown to be effective as it revealed the distinctive opinions about written elements by L2 learner and native speakers. In addition, since more subjects are involved, this approach can also provide more objective and reliable findings than those based on the researchers' personal judgment.

The findings can also provide some guidance on Chinese second language acquisition. Since the L2 learner has already built the correct awareness about the variation of register in genres, fewer resources should be spent in this aspect. On the other hand, the L2 learner lacks the awareness of content word as the written elements so that more efforts should be focused on learning of the content words. The L2 learner also showed difficulties in understanding the abbreviated form of bi-character words. This means learner need to build up their familiarity of Chinese abbreviations.

The findings can also implement current L2 instruction theory, especially the sequential approach. On the basis of the primacy of the spoken language, some scholars (Walker, 1984; Everson, 1994, 1998, 2002; Everson & Ke, 1997; Ke, 1996, 1998a, 1998b; Ke & Shen, 2003; etc) have proposed a sequential approach in the development of four language skills. Progressing from speaking/listening to reading/writing, rather than dealing with all four skills equally from the beginning. In the sequential approach, reading is introduced after the learners have built up some foundation of spoken Chinese. Moreover, what the students read first is what they have learned in their spoken classes. The strength of this approach is that it can help learners become familiar with Chinese orthography and avoid learners decoding in beginning reading. However, the approach is silent with respect to instructions on how to transition from spoken-style materials to authentic written-style materials. The findings in this preliminary study suggest that genres, together with the lengths of texts, density of content and familiarity of the learners with the text themes should be considered when designing the reading material for the transition. As a result of this study, modern novels will be the most suitable genre for this transition.

This study is preliminary and the findings are tentative and should not be generalized. More research is needed to obtain more convincing findings. However, this study makes the efforts to construct new approach from the language users, and fill the gap in the sequential reading approach in current Chinese L2 reading instruction.

REFERENCES

Biber, Douglas. 1988. *Variation across speech and writing.* Cambridge: Cambridge University Press.

Biber, Douglas. 1995. *Dimensions of registers variation: A cross-linguistic comparison.* Cambridge: Cambridge University Press.

Blankenship, Jane. 1962. A linguistic analysis of oral and written style. *Quarterly Journal of Speech* 48. 419-22.

Chafe, Wallace L. 1982. Integration and involvement in speaking, writing and oral literature. In Tannen (ed.), 35-53.

Chao, Yuanren. 1968. *A grammar of spoken Chinese.* Berkeley: University of California Press.

Christensen, Matthew B. 1994. *Variation in spoken and written mandarin narrative discourse.* Unpublished Ph.D. dissertation, Ohio State University.

Gumperz, Cook J., John J. Gumperz. 1981. From oral to written: The transition to literacy. *Variation in writing,* ed. by Whiteman M. F., 89-109. Hillsdale, NJ: Erlbaum.

DeVito, J. A. 1966. Psychogrammatical factors in oral and written discourse by skilled communicators. *Speech Monographs* 33. 73-76.

Everson, Michael E. 1994. Toward a process view of teaching reading in the second language Chinese curriculum. *Theory into Practice* 33(1). 4-9.

Everson, Michael E. 1998. Word recognition among learners of Chinese as a foreign language: Investigating the relationship between naming and knowing. *The Modern Language Journal* 82. 194-204.

Everson, Michael E. 2002. Theoretical developments in reading Chinese and Japanese as foreign language. *Literacy and the second language learners,* ed. by Sullivan J. H., 1-16. Greenwich, CT: Information Age Publishing.

Everson, Michael E., Chuanren Ke. 1997. An inquiry into the reading strategies of intermediate and advanced learners of Chinese as a foreign language. *Journal of the Chinese Language Teachers Association* 32(1). 1-20.

Georgakopoulou, A. & Goutsos, D. 1997. *Discourse analysis: An introduction.* Edinburgh, EH: Edinburgh University Press.

Gu Huangchu 顾黄初, ZHOU MEIZHEN 周梅珍. 1989. Shilun yutixue yu yuwen jiaoxue 试论语体学与语文教学. *Gu Huangchu yuwen jiaoyu wenji* 《顾黄初语文教育文集》(2002), ed. by GU HUANGCHU 顾黄初. Beijing: Renmin Jiaoyu Chubanshe. http://219.239.238.40/200410/ca605135.htm (March 2006).

Gumperz, John J., Hannah Kaltman, O'Connor, M. C. 1984. Cohesion in spoken and written discourse: Ethnic Style and the Transition to Literacy. *Coherence in spoken and written discourse*, ed. by Deborah Tannen, 3-19. Norwood, NJ: Ablex Publishing Corporation.

Halliday, Michael Alexander Kirkwood. 1989. *Spoken and written language.* Oxford: Oxford University Press.

Kay, Paul. 1977. Language evolution and speech style. *Sociocultural dimensions of language change*, ed. by Ben G. Blount, Mary Sanches, 21-33. New York: Academic Press.

Ke, Chuanren. 1996. An empirical study on the relationship between Chinese character recognition and production. *Modern Language Journal* 80(3). 340-50.

Ke, Chuanren. 1998a. Effects of strategies on the learning of Chinese characters among foreign language students. *Journal of the Chinese Language Teachers Association* 33(2). 93-112.

Ke, Chuanren. 1998b. Effects of language background on the learning of Chinese characters among foreign language students. *Foreign Language Annals* 31(1). 91-100.

Ke, Chuanren, Michael E. Everson. 1999. Recent Research in CFL Reading and Its Pedagogical Implications. *Mapping the Course of the Chinese Language Field* (Chinese Language Teachers Association Monograph Series, Volume III), ed. by Madeline Chu, 189-203. Kalamazoo, MI: Chinese Language Teachers Association.

Jahandarie, Khosrow. 1999. Structural differences between speech and writing: the linguistic evidence. *Spoken and written discourse: a multi-disciplinary perspective,* ed. by Khosrow Jahandarie, 131-50. Stamford: CT, Ablex Publishing Corporation.

Li, Charles N., Sandra A. Thompson. 1982. The gulf between spoken and written language: a case study in Chinese. In Tannen (ed.), 77-88.

Li Xingjian 李行健 (ed.). 1998. *Xiandai hanyu guifan zidian* 《现代汉语规范字典》. Beijing: Yuwen Chubanshe.

Lu Bingfu 陆丙甫. 2002. Duiwai hanyu jiaoxue zhong de kouyu bijiao 对外汉语教学中的口语比较。*Duiwai hanyu luncong (di er ji)* 《对外汉语论丛，第二集》, JING CHENG 竟成(ed.), 352-59. Shanghai: Shanghai waiyu jiaoyu chubanshe.

Liu Zhuang 刘壮. 2005. Teaching written Chinese and systematized research 重视书面语教学，进行系统研究. *Chinese Language Learning 汉语学习*, 2005(4). 58-62.

Lü Shuxiang 吕叔湘 (ed.). 1994. *Xiandai hanyu babaici* 《现代汉语八百词》. Beijing: Shangwu Yinshuguan.

Scholes, Robert J. 1999. *A linguistics approach to reading and writing*. Lewiston/Queenston/Lampeter: The Edwin Mellen Press.

Stubbs, Michael. 1980. *Language and literacy: The sociolinguistics of reading and writing*. London: Routedge.

Tannen, Deborah. 1982. The oral/literate continuum in discourse. In Tannen (ed.), 1-16.

Tannen, Deborah (ed.). 1982. *Spoken and written language: Exploring orality and literacy.* Norwood, NJ: ABLEX Publishing Corporation.

Walker, Galal. 1984. "Literacy" and "reading" in a Chinese language program. *Journal of the Chinese Language Teachers Association* 19(1). 67-84.

Wang Dechun 王德春. 1987. *Yuti lüelun* 《语体略论》. Fuzhou: Fujian Educational Press.

Vygotsky, Lev Semenovich. (1962). *Thought and Language.* Edited and translated by Eugenia Hanfmann and Gertrude Vakar. Cambridge: The M.I.T. Press.

Yuyan Jiaoxue Yanjiusuo 语言教学研究所. 1986. *Xiandai hanyu pinlü cidian* 《现代汉语频率词典》. Beijing: Beijing Language Institute Press.

Zhang Xueqian (Tiun Hak-Khiam) 张学谦. 2000. A multi-dimensional analysis of spoken and written Taiwanese register 台语口语及书面语体的多面向分析. *Language and Linguistics* 1. 89-117.

Zhang Xueqian (Tiun Hak-Khiam) 张学谦, Yang Yunyan (Iu U) 杨允言. (1994). *Taiyu yuti yanbian fenxi* 台语语体演变分析. http://203.62.42.21/giankiu/ Keoe/KKH/Register/kiatanpoko/kiatanpoko.asp (April 2006).