Data-Driven Adapting for Fine-Tuning Chinese Teaching Materials: Using Corpora as Benchmarks

Wei Bo, Jing Chen, Kai Guo and Tan Jin

Abstract While there have been a considerable number of corpus-based studies informing the content of teaching materials, direct explorations of corpora by teachers to adapt source texts (i.e., data-driven adapting) for classroom teaching remain a largely unexplored area. This chapter examines how teachers adapt new texts in a more comprehensible manner for L2 Chinese learners using an online system, *Chi-Editor*. *Chi-Editor* was developed to automatically assess text complexity and tag Chinese words and sentences for text simplification purposes. The evaluation of a text in terms of its level of difficulty and annotation of difficult words and long sentences in the text are produced based on the data mining of linguistic features from a corpus of roughly 350 widely-used textbooks, selected from an anthology of L2 Chinese teaching materials and packages produced by over 1000 publishers around the world. To investigate both the process and outcome of data-driven adapting using *Chi-Editor*, a case study was conducted, involving a team of teachers working on the adaptation of texts. Results are discussed in terms of the effectiveness of the data-driven adapting practices by teachers in a classroom setting. Overall, the results contribute strong evidence that teachers can learn and benefit from data-driven adapting and support the notion that corpus data, including linguistic features, can be employed to facilitate the text simplification process. Implications are also given for integrating the data-driven adapting process into regular teacher-prepared L2 Chinese materials for classroom teaching.

W. Bo
Dali University, Dali, China
e-mail: bowei_2010@163.com

J. Chen · T. Jin (✉)
Sun Yat-sen University, Guangzhou, China
e-mail: jintan6@mail.sysu.edu.cn

J. Chen
e-mail: chenjing@mail.sysu.edu.cn

K. Guo
Shanghai Jiao Tong University, Shanghai, China
e-mail: dearmrk@163.com

X. Lu and B. Chen (eds.), *Computational and Corpus Approaches to Chinese Language Learning*, Chinese Language Learning Sciences,
https://doi.org/10.1007/978-981-13-3570-9_6

1 Introduction

In second language teaching and learning, text adaptation is often employed to ensure comprehensibility of the reading texts by L2 learners (Yano et al. 1994). In the L2 classroom, it is especially common for adapted texts (Young 1999) to be used. Previous studies on the process of text adaptation and its outcome have examined the use of both teacher intuition and automatic tools in text adaptation, with a common goal to find objective, consistent criteria that could be used to efficiently evaluate and improve the readability of texts by L2 learners. Research on the intuitive approach has examined how texts are perceived and adapted by teachers (e.g., Green and Hawkey 2011). In the absence of corpus data, teachers usually rely on their own intuitive judgments, sometimes in combination with guidance provided in a textbook, to adapt texts. Compared to the intuitive approach, the use of computational tools has the unique potential of being both more efficient and consistent. Recent research has also shown that data-driven computational tools have the added benefit of enhancing the accuracy of text adaption, by highlighting and annotating candidate words and sentences that contribute to the complexity of a text (Jin and Lu 2018).

The purpose of this chapter is to show how the use of a data-driven approach for adapting texts for L2 Chinese classrooms is more effective than solely relying on teacher intuition. To this end, we investigate the process and outcome when teachers adapt texts using the online tool, *Chi-Editor* (see Sects. 3.1–3.3), designed with corpus data for benchmarking. We focus in particular on how texts adapted using this tool differ from those adapted based on teacher intuition, and whether L2 Chinese teachers perceive it as a useful facilitative tool for text adaptation. In doing so, the authors hope to showcase how data-driven text adaption may be effectively integrated into regular teacher-prepared L2 Chinese materials for classroom teaching.

2 Literature Review

Since the late 1990s, there has been a tendency to advocate the use of either authentic or simplified texts as input to L2 learning, especially for L2 learners at the beginning and intermediate levels (e.g., Johnson 1981; Tomlinson et al. 2001). However, little empirical evidence has been provided to validate the effects of using either simplified or authentic texts on L2 development (e.g., Cummins 1981; Goodman 1986; Krashen 1981, 1985), and a more realistic question is whether the use of simplified or authentic texts leads to different learning outcomes. In an effort to investigate the extent of such differences and their implications for L2 learning, Crossley et al. (2007) measured linguistic features—including the lexical, syntactical, and discourse differences—that characterize authentic and simplified texts, respectively. To compare the linguistic features of authentic and simplified texts, over 250 linguistic and cohesion features were employed to configure the computational tool *Coh-Metrix*, developed to assess the coherence and cohesion of reading texts (Graesser et al.

2004). Crossley et al. analyzed a corpus of 81 simplified texts with 21,117 words and one of 24 authentic texts with 15,640 words using this tool. Subsequent statistical analysis revealed significant differences between simplified and authentic texts in syntactic complexity, word information and cohesion. These differences have useful implications for selecting texts as input to L2 learning. For instance, the simplified texts showed greater cohesion than did the authentic texts, indicating that adapted texts may be more appropriate for beginner and intermediate level L2 learners.

Research has shown that text adaptation is necessary in preparing teaching materials for L2 learners of certain levels in reading, but it remains opaque as to how text adaptation should be conducted in practice. Adaptation is regarded as a creative art (Wesman 1971), often improvised or carried out in intuitive ways in practice. In other words, it needs to be revealed how the process of text adaptation proceeds and whether there exist identifiable patterns of adaptation. To this end, Green and Hawkey (2011) conducted a case study of item writer practices based on qualitative analyses to standardize the rules for text adaptation. Four trained item writers working on the International English Language Testing System (IELTS) were selected as the subjects to be observed. Stimulated recall interviews and the subjects' writing reflections revealed common strategies such as text deletion, consolidation, expansion, substitution, and insertion. Results showed that during the item writers' adaptation process, they increased the proportion of frequent word types and decreased that of less frequent words. Item writers also reflected on their practices of reducing redundancy and technical language, changing styles, deciding on potentially sensitive issues and relationships between texts and test items when they adapted texts in order to make them appropriate for the proficiency levels of the L2 test takers.

It seems that in Green and Hawkey's study, item writers used a defined set of strategies for writing tests and that these strategies were deployed exclusively in isolation from feedback by the reader, that is, there is no direct feedback mechanism from the L2 learners on the effects of adaptation on the linguistic features of texts. To reveal the impact of different levels of text adaptation on reading comprehension, Crossley et al. (2012) used *Coh-Metrix* to quantitatively establish a link between the comprehension of adapted texts and proficiency levels. Three hundred news texts were simplified into three different levels, i.e., beginner, intermediate, and advanced. Fourteen indices were employed to measure linguistic features related to cohesion, linguistic sophistication, and surface-level variables, such as word frequency, lexical diversity, spatial cohesion, temporal cohesion, and syntactic complexity. The results showed that beginner-level texts are generally less lexically and syntactically sophisticated than the advanced-level ones and that the former contains more cohesive features than the latter. This quantitative study indicates that lexical, syntactic, and cohesive features are generally the best indices for classifying different levels of L2 texts.

Chinese Mandarin, a language in the Sino-Tibetan family, has linguistic features that differ significantly from English. The English language is based on clearly identifiable word units (which are of one or more syllables in length), whereas Chinese characters can be either a word or part of a word. For example, in English the word "computer" is one string of letters, whereas in Chinese, 电脑 (*dian nao*) is two sep-

arate characters, each of which has its own meaning: "电 (*dian* means electric)" and "脑 (*nao* means brain)". Thus, when processing a text, in English, "computer" would be a single word; whereas "电脑 (*dian nao*)" could be analyzed as either one or two words (i.e., either two characters combined or two separate characters). We, therefore, have two distinct ways to process text in Chinese: at the level of the character (i.e., character-based) or at the level of the 'word' (i.e., word-based). At the level of the word, processing text in English and in Chinese is operationally very similar. For example, in both languages, word-count, part-of-speech diversity, and frequency can be treated with very similar processes. However, at the level of the character, there is no equivalent in English. Therefore, in contrast to English, such features as the complexity of character strokes and usage-based frequency of characters are commonly applied to measure character complexity. Such criteria have been widely employed for developing teaching materials for both elementary native speakers and L2 learners. As a national standard, *the Graded Chinese Syllables, Characters and Words for the Application of Teaching Chinese to the Speakers of Other Languages* (Ministry of Education and State Language Commission, the People's Republic of China 2010, *GCSCW* hereafter) has been developed for the purpose of proficiency testing and has become integral to the L2 Chinese evaluation system (Liu and Ma 2010). Corpus techniques are central in *GCSCW* research in order to rank and design the characters and words by frequency. In addition, other natural language processing technologies are also applied in *GCSCW* research, such as automatic processing of word-segmentation and word-frequency statistics. The most notable feature of *GCSCW* research identifies three levels (beginner, intermediate, and advanced) by combining both character-based and word-based approaches and this combination approach has now become established as the standard to develop L2 Chinese teaching materials.

These multilevel linguistic features were also applied to level the readability of L2 Chinese texts by Sung et al. (2015), who used thirty linguistic features of L2 Chinese as well as 1578 classified texts to evaluate the accuracy of text leveling for instructional purposes—specifically, for teaching materials. Sung et al. produced a text that defines the levels comparable to the Common European Framework of Reference (CEFR) according to L2 Chinese experts; a readability assessment system was later created using the 30 linguistic features developed from previous studies. The F-score selection method was used to evaluate the relative importance of linguistic features. In the final system for L2 Chinese leveling, words and characters receive equal weight, and multiple features are used for both, such as the average of vocabulary levels, high-level words, mean square of vocabulary levels, two-character words, and intermediate stroke-count characters. Leveling, i.e., assigning a "level" to a given text, helps teachers, and learners select proper texts to enhance learning at an appropriate proficiency level. Another application suggested by this study is that the authors or editors also benefited from the leveling system when examining the linguistic features of L2 teaching materials they are editing.

In sum, previous studies have revealed two particularly important insights regarding text adaptation for L2 teaching materials: on the one hand, there is qualitative evidence showing that teachers employ adaptation strategies (e.g., text deletion, con-

solidation, expansion, substitution, and insertion) with certain patterns rather than treating adaptation as a purely "improvised art" (e.g., Green and Hawkey 2011); on the other hand, quantitative research has revealed a link between text adaptation and reading comprehension, and the key linguistic features (such as word frequency and spatial cohesion) at play in this link (e.g., Crossley et al. 2012). Such strategies used in text adaptation in previous studies are termed "teacher intuition" in this chapter. In the evaluation of L2 Chinese proficiency levels, a national standard for Chinese characters and words is being actively promoted by the *GCSCW* (Ministry of Education and State Language Commission, the People's Republic of China 2010). A more detailed empirical examination of this *GCSCW* standard was undertaken by Sung and his colleagues (Sung et al. 2015) using both the character-based and word-based approach to analyze a set of widely-used L2 Chinese teaching materials. The *GCSCW* and Sung et al.'s findings together demonstrate the following two aspects: first, the standard promoted by the *GCSCW* facilitates sound guidelines for editing L2 Chinese teaching materials and adapting learning texts; second, the relative importance of words and characters varied according to proficiency level (Sung et al. 2015). Therefore, the term "leveling" both in Sung et al. and in this chapter refers to evaluating the difficulty of linguistic features consistent with the L2 Chinese teaching syllabus. However, the relationship between the linguistic features and the result of text adaptation still has not been addressed by any empirical research. In other words, under the guidance of linguistic features in *GCSCW*, how L2 Chinese texts are adapted for pedagogical purposes (i.e., leveling of text difficulty, finding of linguistic features, etc.) remains unknown. The current study, therefore, attempts to fill this gap.

3 Introduction to the Online Tool *Chi-Editor*

To provide data-driven support for teachers in adapting new texts in a more comprehensible manner for L2 Chinese students, the online system, *Chi-Editor*, has been developed to automatically level texts in terms of linguistic complexity, to tag Chinese words and sentences as well as to report profiles of Chinese characters and words (Jin and Li 2016). The function of leveling and tagging in *Chi-Editor* are based on the data mining of linguistic features from a corpus compiled from roughly 350 widely-used textbooks, which were selected from an anthology of L2 Chinese teaching materials and packages produced by over one thousand publishers around the world (Base for International Chinese Teaching Materials Developing and Teacher Training 2017); the reporting function generates the lists of Chinese characters and words provided by *GCSCW*, which established a standard combining both character-based and word-based approaches.

We take a text as an example to illustrate the three main functions—leveling, tagging, and reporting—provided by *Chi-Editor* as follows. The text is extracted from the Level 3 book *I Want to Be a Lawyer* in a graded Chinese reader series titled *Friends* (Confucius Institute Headquarters 2014). The Preface of the textbook series

Fig. 1 Interface of text typing/pasting

states that Level 3 in all six levels uses 600 words in accordance with the new HSK (an abbreviation of *Hanyu Shuiping Kaoshi*, the Chinese Proficiency Test). Level 3 of HSK requires that students can read basic Chinese materials related to daily life and find specific information from paragraphs of familiar content according to the *International Curriculum for Chinese Language Education* (Confucius Institute Headquarters 2015). After typing or pasting the text into *Chi-Editor*, the resulting analysis can be accessed online (see Fig. 1, retrieved from http://www.languagedata.net/editor/) but the text length must fall within the limits of 100–5000 characters.

3.1 Leveling in Chi-Editor

After inputting a verification code and pressing “Start Analysis” (see Fig. 1), the user will see the results in a coordinate graph indicating the level of the text (see Fig. 2). The upper section of Fig. 2 summarizes the leveling result for the original text using several indices, namely, the overall level of difficulty (hereafter LD), CEFR level, L2 Chinese syllabus level, mean length of sentence, length of the longest sentence and length of the text. The lower section provides guidance on how the LD value is to be interpreted. The horizontal axis represents the LD, which corresponds to six levels in CEFR with real values ranging from 1.0 to 4.0; the accuracy of these values in judging text difficulty reaches over 90% as supported by empirical research (Lin 2016). The vertical axis represents L2 Chinese syllabus levels according to the *International Curriculum for Chinese Language Education* (Confucius Institute Headquarters 2015).

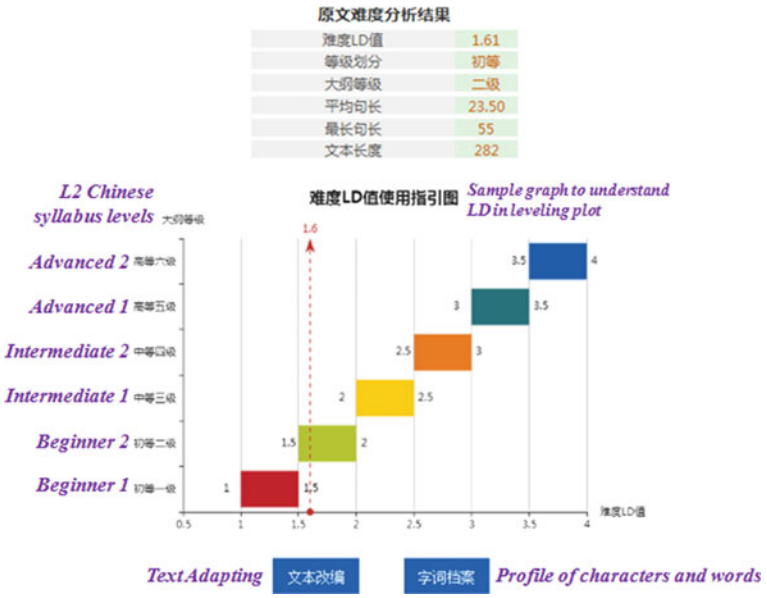


Fig. 2 Interface of leveling plot [English translation given in purple italics]

3.2 Tagging in Chi-Editor

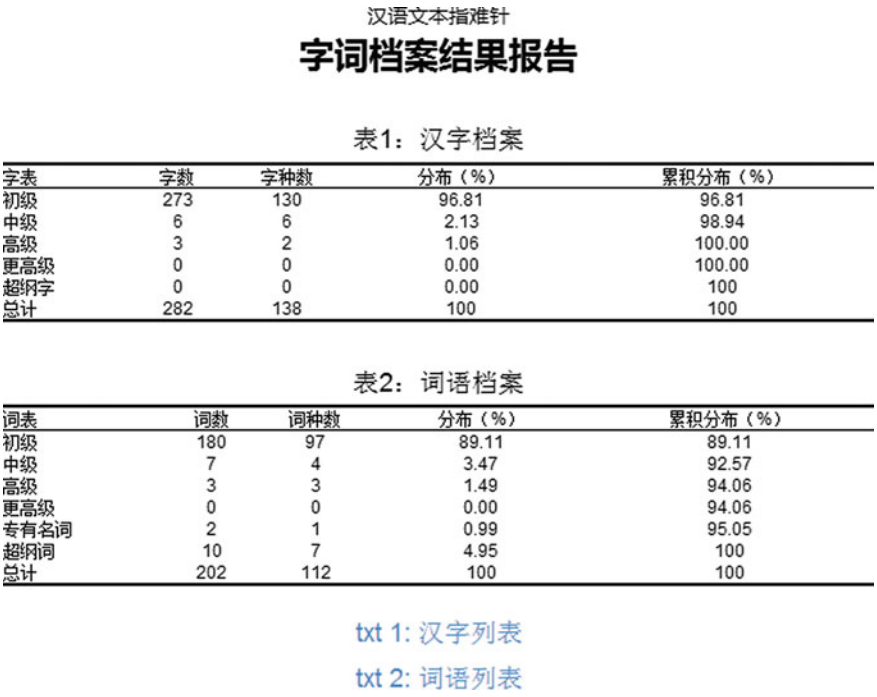
More detailed information on the functions of the adapting interface can be found when users click the blue button *Text Adapting* as shown at the bottom of Fig. 2. Figure 3 shows the tagged elements that may help teachers or learners adapt. To operate this tagging interface, it is important to be aware of the following procedures. First, the interface provides tagged words and sentences throughout the text adaptation process. Words can be tagged based on (1) word level and (2) word frequency or usage-based examples. Different colors are used to distinguish different word levels: beginner, intermediate, advanced, higher advanced, words beyond vocabulary (i.e., words that are beyond the vocabulary of the specific level; hereafter WBV), and proper nouns (PN). In the example, in Fig. 3, red words such as "摩托车 (motorcycle)" and "拜佛 (worship the Buddha)" are in the WBV level, green words like "胖 (fat)" and 辛苦 (laborious)" belong to the intermediate level. As for word frequency and usage-based examples, links are provided to allow users to explore examples of how a word is used in the corpus of. Further, all words at the WBV level are tagged with word frequency information to help teachers or learners judge their level. For example, the word "摩托车" (motorcycle) is marked as "22" to its upper right to indicate its word frequency, while the word "拜佛" (worship the Buddha) is marked as "2". Thus, users can easily identify the vocabulary levels of these words by comparing those numbers. Sentence tagging focuses on the longest sentence in the text and underlines it to draw user attention to it since the length of the longest

Fig. 3 Tagging for adapting

sentence is an important contributor to text complexity. Second, for the adapting function, *Chi-Editor* provides a Microsoft Word-like window where one can adapt the colored words through addition, deletion, or substitution. Finally, *Chi-Editor* can reanalyze results and reassess subsequent attempts. Using this text, as an example again, users could first separate the longest sentence into three and four short sentences before the reassessment automatically changes the LD value from 1.61 to 1.60 in three sentences and from 1.61 to 1.59 in four sentences. Moreover, the LD value changes from 1.61 to 1.49 and the syllabus level from grade 2 to grade 1 when researchers substitute such words "辛苦,饭菜,总是,想念" with "累,菜,常常,想".

3.3 Reporting in Chi-Editor

Referring back to Fig. 2, users can click the blue button on the bottom right of the coordinate graph called *Lexical Profiling* during or after reassessment. The third main function—reporting, which is based on lists of Chinese characters and words by *GCSCW*—is illustrated in Fig. 4. The upper table presents information about characters, while the lower table presents information about words in the same text. Both reports include the number of types and tokens of characters or words at different levels, as well as coverage of those characters or words. These statistics are useful to be aware of as users try to generalize the level of the text. One more detailed report can be displayed by clicking the links known as *txt 1* and *txt 2* as shown in Fig. 5, in which the left portion represents the profile of characters, and right portion represents the profile of the words. In short, a *Chi-Editor* report allows teachers and learners to generalize the difficulty of texts by presenting information graphically.


汉语文本指难针

字词档案结果报告

表1：汉字档案

字表	字数	字种数	分布（%）	累积分布（%）
初级	273	130	96.81	96.81
中级	6	6	2.13	98.94
高级	3	2	1.06	100.00
更高级	0	0	0.00	100.00
超纲字	0	0	0.00	100
总计	282	138	100	100

表2：词语档案

词表	词数	词种数	分布（%）	累积分布（%）
初级	180	97	89.11	89.11
中级	7	4	3.47	92.57
高级	3	3	1.49	94.06
更高级	0	0	0.00	94.06
专有名词	2	1	0.99	95.05
超纲词	10	7	4.95	100
总计	202	112	100	100

txt 1: 汉字列表

txt 2: 词语列表

Fig. 4 Reporting for adapting

序号	字	字频	等级	比例	累计比例
1	爷	20	初级	7.09%	7.09%
2	我	17	初级	6.03%	13.12%
3	他	13	初级	4.61%	17.73%
4	很	10	初级	3.55%	21.28%
5	一	8	初级	2.84%	24.11%
6	是	7	初级	2.48%	26.60%
7	国	7	初级	2.48%	29.08%
8	的	6	初级	2.13%	31.21%
9	中	5	初级	1.77%	32.98%
10	做	5	初级	1.77%	34.75%
11	人	4	初级	1.42%	36.17%
12	也	4	初级	1.42%	37.59%
13	天	4	初级	1.42%	39.01%
14	着	4	初级	1.42%	40.43%
15	但	3	初级	1.06%	41.49%
16	每	3	初级	1.06%	42.55%
17	都	3	初级	1.06%	43.62%
18	喜	3	初级	1.06%	44.68%
19	欢	3	初级	1.06%	45.74%
20	好	3	初级	1.06%	46.81%

txt 1

序号	词	词频	等级	比例	累计比例
1	我	15	初级	7.43%	7.43%
2	他	13	初级	6.44%	13.86%
3	爷爷	10	初级	4.95%	18.81%
4	很	9	初级	4.46%	23.27%
5	的	6	初级	2.97%	26.24%
6	中国	5	初级	2.48%	28.71%
7	做	5	初级	2.48%	31.19%
8	是	4	初级	1.98%	33.17%
9	也	4	初级	1.98%	35.15%
10	着	4	中级	1.98%	37.13%
11	人	3	初级	1.49%	38.61%
12	每天	3	超纲词	1.49%	40.10%
13	都	3	初级	1.49%	41.58%
14	喜欢	3	初级	1.49%	43.07%
15	看	3	初级	1.49%	44.55%
16	了	3	初级	1.49%	46.04%
17	泰国	2	专有名词	0.99%	47.03%
18	我们	2	初级	0.99%	48.02%
19	在	2	初级	0.99%	49.01%
20	但	2	初级	0.99%	50.00%

txt 2

Fig. 5 Profiles in reporting

4 Research Questions

As discussed above, text adapting in L2 Chinese teaching is pragmatic, but the effect of the use of online tools on the text adaptation process and its outcome remains vague. It is unknown whether there are patterns in using the online tool for text adaption and whether such adapting contributes to improvements in teaching practice. The current study aims to compare teachers' perceptions of adapting by teacher intuition and adapting by using *Chi-Editor*. The specific research questions addressed are: (1) How does *Chi-Editor* affect teachers' adapting process, and (2) compared to adapting by teacher intuition, does adapting with *Chi-Editor* bring about different outcomes with its leveling, tagging and reporting functions?

Table 1 Teacher profiles

Teacher	Gender	Educational background	Relevant experience (years)	Pedagogical adapting
A	Female	Ph.D.	3.5	Experienced
B	Male	Ph.D.	>6	Experienced
C	Female	M.A.	>6	Experienced
D	Male	Ph.D.	>6	Experienced

5 Methodology

5.1 Interview Subjects

Four experienced L2 Chinese teachers participated voluntarily in this study. These four teachers had been selected based on three criteria: (1) educational background with postgraduate degrees and/or doctorate degrees in applied linguistics, (2) relevant practical experience including teaching and researching L2 Chinese, and (3) pedagogical attempts on text adapting for the relevant teaching level (see Table 1 for a profile of the four teachers).

5.2 Procedures

The four teachers followed four adapting procedures: (1) Training: the researchers briefed these teachers about adapting methods; all four teachers used two text samples to get familiar with *Chi-Editor* and reported to the researchers on how they used *Chi-Editor* to adapt the sample texts. (2) Familiarization: the four teachers became familiar with the graded L2 texts. Teachers then practiced using *Chi-Editor* with three graded L2 Chinese texts (beginner, intermediate, advanced). At this point, they made judgments independently without receiving any feedback or help from the researchers. They were allowed to judge text levels intuitively and focused on familiarizing themselves with different text features. (3) Pilot Adapting: the four teachers adapted these three graded texts again for pilot adapting, and subsequently received feedback from the researchers. This would help the four teachers see clearly what linguistic features had been adapted and how a text's level had been modified by *Chi-Editor.* (4) Adapting: the four teachers were required to accomplish independently the following two adapting tasks: (a) adapting one upper-intermediate text to a lower-intermediate level first using teacher intuition (followed by an interview), then using *Chi-Editor* (also followed by an interview); and (b) adapting a different text, of lower-intermediate level, to cater to L2 Chinese learners at beginner level. Tasks (a) and (b) are illustrated in Fig. 6.

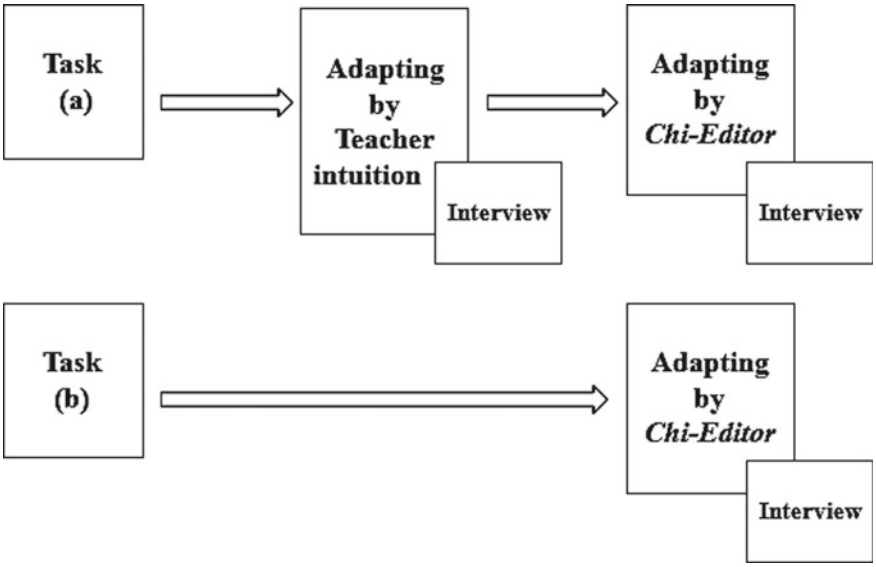


Fig. 6 Task procedure of adapting

5.3 Texts

For the abovementioned two assignments, we selected two passages, one of upper-intermediate level and the other of intermediate level, for two reasons. On the one hand, significantly different linguistic features exist between the beginning and the advanced levels (see Crossley et al. 2012), so we focused on adapting from upper-intermediate level to basic-intermediate level in task (a). On the other hand, pedagogical adapting requires more simplified texts for beginner level (or for basic-intermediate at most), so we asked interviewees to adapt to the beginner and the basic-intermediate levels in tasks (b) and (a), respectively. Therefore, we selected two passages in the series book of graded Chinese readers published by the Confucius Institute Headquarters (2014): the text in task (a) comes from Level 3 entitled *My Grandfather*, and the text in task (b) comes from Level 5 and is entitled *How the Internet Changed Our Lives*.

After the four teachers completed the whole set of adapting activities, qualitative interviews were carried out by the researchers to provide insights into the adapting process and outcomes of using *Chi-Editor*. The interviews were semi-structured, and three topics were designed according to the functions of *Chi-Editor* to investigate both the process and outcomes of data-driven adapting: (1) leveling, (2) tagging, and (3) reporting. An interview frame encompassing the three topics was first developed for guiding the interview, and it was then sent out for expert review before the formal interviews were conducted (see Fig. 7).

INTERVIEW FRAME

Topic 1 Leveling
Questions:
(1) When using traditional teaching methods, how did you assign a level to the teaching materials extract?
(2) When using *Chi-Editor*, how did you assign a level to the teaching materials extract?
(3) Did you notice any differences between the two results in assigning a level to the teaching materials extract?
(4) What did you learn, if anything, when using the *Chi-Editor* leveling function?

Topic 2 Tagging
Questions:
(1) When using traditional teaching methods, what text features did you identify in your adapting process?
(2) When using *Chi-Editor*, did you use the tagging function and if so what text features did you use and how?
(3) Did you notice any differences between text features identified when not using *Chi-Editor* compared to those identified by the *Chi-Editor* tagging function?
(4) What did you learn, if anything, when using the *Chi-Editor* tagging function?

Topic 3 Reporting
Questions:
(1) When using your teacher intuition and experience, how would you describe the process you used to adapt the teaching materials extract provided?
(2) When using *Chi-Editor* how did you use the report it generated to adapt the teaching materials extract provided?
(3) Did you notice any differences in your adapting process between using *Chi-Editor*, with its reporting function, and not using it?
(4) What did you learn, if anything, when using the report generated by *Chi-Editor*?

Fig. 7 Interview frame

The interviews were conducted in Chinese, which is the native language of all four teachers. All four interviews were recorded, resulting in a total of four hours of recorded audio, excluding another two hours for adapting by teacher intuition in task (a), which included reading, thinking, and analyzing by each interviewer. The audio files were transcribed into a total of 9023 Chinese characters by the researchers, including quotation marks.

During the qualitative analysis phase of this research, analytical categories emerged when the researchers were listening to the audio files and transcribing the interview. Based on the analytical categories, a coding scheme was made by two of the researchers, which was later sent for expert review and finalized.

6 Findings

After the interviews were transcribed, interviewees' comments were categorized to describe their usage of *Chi-Editor*. Among all dimensions provided by *Chi-Editor,* tagging (with 56% comment percentage) was the most frequent topic mentioned by the interviewees, followed by leveling (32%) and reporting (12%). The following presents the four interviewees' comments on the topics of leveling, tagging, and reporting.

6.1 Interview: Leveling

Leveling was the first fundamental step interviewee were asked to review in their adapting assignments, and information about language proficiency of the interviewee was noted. Procedurally, interviewees first examined text adapting in the traditional teaching setting; second, they provided reflections of their experiences using *Chi-Editor*. As for the indices used during the leveling process, the four interviewees observed similarities between the traditional teaching setting and using *Chi-Editor*, both of which focused on both words and sentences. The statements given below, by Interviewee B, are commonly found in the interviewees' description of such similarities.

Extract 1 (Interviewee B, #T1-B2)
I judged the text level mainly by the level of words, especially the words on intermediate and advanced levels of HSK word lists.

Commonly, the four interviewees focused on the word levels to discussing the similarities between the two different ways of leveling. In addition, language points—such as collocations, chunks, and complements in Chinese Mandarin—also attracted their attention. For example, Extract 2 was again excerpted from Interviewee B's comments on the traditional teaching setting.

Extract 2 (Interviewee B, #T1-B3)
…they (referring to "多得多" and "多多了," both of which mean "much more than") are complements in the sentences "网上的资料比学校图书馆的多得多 (There is more information on the Internet than in school libraries)" and "网上的电视剧比电视上多多了 (There are more TV shows on the Internet than on television)"; although their meanings are similar, their linguistic structures are completely different, especially after combining comparable structures (it refers to the "bi" structure in Mandarin Chinese).

While *Chi-Editor* made use of the character list in text leveling, the four interviewees did not seem to utilize either level or frequency information of characters. This indicates that many L2 Chinese teachers default to word-based teaching approaches instead of character-based teaching approaches, as illustrated in Extract 3 by Interviewee C:

Extract 3 (Interviewee C, #T1-C2)
It is hard to say attention toward characters is useless when a kind of approach conducts leveling for the purpose of text adapting. But as a L2 Chinese teacher, I would usually pay more attention to characters only if the learners are early beginners. According to your requirements (referring to requirements from the interviewer/researcher), I assumed it to be the level of basic-intermediate rather than a true beginner level. So it does not work if I pay any attention to the characters. In other words, I feel it is tough to determine which character affects the text level and which does not.

In this claim, Interviewee C expresses a presumption that learners' levels result from two different teaching approaches. Namely, the beginner level requires a more character-based teaching approach, and after that, a word-based teaching approach should serve as the backbone under the communication approach. Interviewee A voiced an opinion on a different aspect regarding the character-based approach: "*In my opinion, character-based teaching materials are really boring, and learners can benefit nothing from them when communicating in spoken Chinese*" (Interviewee A, #T1-A4).

It is apparent that the L2 Chinese teachers regard the word-based teaching approach as the more effective alternative; the four interviewees reported similar sentiments regarding Topic 1. Specifically, we find that teacher intuition, in identifying word levels and linguistic features, exactly matches the criteria used by *Chi-Editor*. Nevertheless, *Chi-Editor* as an online system can provide more detailed quantitative data for text leveling in terms of linguistic features, such as mean length of utterance (MLU; Ellis 1999), the LD value, the total number of words and characters, etc. It seems possible that interviewees are often also referring to external frameworks, such as the HSK test or other established teaching syllabi, to refine their teacher intuition. Moreover, the preference for using a word-based approach is also based on teacher intuition rather than any particular theoretical basis.

6.2 Interview: Tagging

Tagging is typically implicit when one is learning to adapt texts, but it is explicit in *Chi-Editor* processing due to its data-driven design. In commenting on the tagging function of *Chi-Editor*, interviewees showed great interest in the explicit tagging interface, which, as introduced in Sect. 3.2, includes such functions as underlining the longest sentence, marking word frequency, and citing examples linked to teaching materials. Interviewee A first noticed how she leveled the text using the implicit tags in her mind.

Extract 4 (Interviewee A, #T2-A1)
(By teacher intuition) I basically leveled the text with marks derived from difficult words and grammar already in mind, and these marks made me more confident in

text leveling. However, I would not write them down or list them in a quantitative way. I memorized these marks, and they permeated a part of my judging instinct.

As for the helpfulness of explicit tags offered by *Chi-Editor*, all interviewees commented on the positive effect on teaching practice because the tagging results could help interviewees rethink a specific word or grammar level. Interviewee B described *Chi-Editor* as a "flexible friend" in that teachers could find more information with data sources rather than merely concludes the word frequency and word difficulty by color and number on the interface.

Extract 5 (Interviewee B, #T2-B6)
During the process of using *Chi-Editor*, I found it to be a helpful friend who could tell you the places where we should pay attention to in the text. For example, when I was processing the text with *Chi-Editor*, it marked "摩托车 (the motorcycle)" as a WBV; Meanwhile, it was also marked with a relatively high word frequency and a link to data sources. At first glance, you might assume this word falls under WBV, but because of its high frequency (the frequency rank was 22), you would need to link it to teaching materials as a data source. After checking the data source, it showed this word's frequency was 9 in beginning materials, 9 in intermediate materials, and 4 in advanced materials. Your judgment was clear that this word should not be categorized as WBV in the teaching syllabus. On the contrary, the data source showed a frequency of 9 in intermediate materials, so this word instead became a new lexical point in your teaching syllabus for the beginner level. In another example, in the same text processing, the program marked "泰国 (Thailand)" as a PN, without any hint about word frequency aside from the blue color indication. However, after you linked it to the data source, you would find the word frequency was 114, a relatively high frequency. According to my intuition from teaching Chinese, you could keep and teach this word if you met a Thai student, or you could delete or change this word if you do not want to teach it. The *Chi-Editor* just marked this word in blue in case you might deal with this scenario. Another word, "拜佛 (worship the Buddha)," which was a WBV word for non-Thai students but an all-level-fitted word for Thai students, was only 2 according to the data source despite the word frequency.

While teachers found the tagging of the longest sentences provided *Chi-Editor* to be useful, they indicate that they paid attention to other long sentences or complex components of certain sentences as well. Interviewee D, for example, stated:

Extract 6 (Interviewee D, #T2-D3)
The online available *Chi-Editor* provides the tag of the longest sentence, and it truly reminds me to pay more attention to this sentence. However, according to your requirement of adapting the text from upper-intermediate level to lower-intermediate level, I might look carefully for something other than the longest sentence tagged by *Chi-Editor*. For example, the attribute "住在中国的 (who lives in China)" is difficult because the attribute is similar to the relative clause in English, and it probably lifts the difficulty of the full sentence "我可以跟我住在中国的朋友发电子邮件 (I can email my friend who lives in China)". In that case, although this sentence is not the

longest one according to the online system, this sentence should be noticed due to the attribute part in it. As for an adapting solution in this example, I think it is better to separate the sentence into two, or to substitute the attribute with a simple one, such as "中国的 (Chinese)".

Tagging not only demonstrates the analytical result of *Chi-Editor* but also immediately computes text difficulty levels. Interviewee C provides the operational details and discusses the longest sentence as follows:

Extract 7 (Interviewee C, #T2-C11)
I have noticed the longer sentences when I am asked to level the text, and I just kept thinking about how the situation would change if I shorten the sentences. *Chi-Editor* underlined the longest sentence in the text, which confirmed my original judgment (the sentence was "吃完饭后, 爷爷喜欢一边唱歌一边洗碗, 我就在旁边看着他洗, 听着他唱小时候的歌, 每天早上他都骑摩托车送我去学校, 再给我一些零花钱。"). The difference is, I could easily shorten or separate this sentence, and I tried calculating this again and again using *Chi-Editor* to make sure the difficulty could be reduced. In the longest sentence, for example, the LD value before my adaptation was 1.61, and the length was 57 characters. The LD value then changed to 1.60 (length = 32 characters) and 1.59 (length = 32 characters) when I tried to separate the sentence into three single sentences and four single sentences, respectively. In that case, you could select the proper LD value after adapting.

6.3 Interview: Reporting

Reporting in *Chi-Editor* summarizes information about the characters and words used in the text in a list format. Most comments on the reporting function touched upon teachers' perceptions of how the resulting word/character lists could be properly used. Interviewee B expressed this viewpoint as follows:

Extract 8 (Interviewee B, #T3-B1)
I typically focused on the WBV and PN after adapting a text, even if the text was not a typical long passage. I assumed that the WBV and PN would be just as effective in this small passage. So, I focused on their word frequency and word percentage. However, I wasn't guessing the level when I encountered a WBV or PN unless the word/character list was reported in *Chi-Editor*. For example, the program shows the word "总是 (always)" as a WBV, but you do not need to worry about it because the word frequency was 2 and the word percentage was 0.98%. Most importantly, *Chi-Editor* will tell you the result, and you don't need to do the calculations by yourself. You can imagine the effect numbers have on word frequency and word percentage in a long text.

Interviewee D agreed that using the reporting function was helpful in adapting a long text. She went a step further, stating that "…if I need to edit or revise my textbook on passages, I'm certain that numbers and calculations in the reporting could help where I need to change the word or the character in a lesson". The following provides more details:

Extract 9 (Interviewee D, #T3-D2)
The word list told me that some word frequencies were really high, but it might make the whole text seem redundant, at least to a native speaker. For example, the word "爷爷 (grandfather)" has been used in this text 11 times according to the word frequency in reporting. Because the word is a noun, I feel that the number is too high due to Chinese being a topic-dominant language. Subsequently, I read the text again more closely. The adapting of this word is supposed to allow discourse cohesion, and this word could usually be replaced or substituted by a pronoun "他 (he/him)".

In relation to the reporting function, interviewees agreed that character/word lists were less effective in a short text, but agreed that these lists would be very useful in highlighting the distribution of words and characters by word frequency, word percentage and word percentage within a certain range in a long text.

7 Discussion and Conclusion

This study generated qualitative data to assess teachers' perceptions of the differences between the intuitive approach and data-driven approach (using *Chi-Editor*) to text adaptation. Our findings indicate that teachers can be trained to use *Chi-Editor* and to take advantage of its benefits.

First, despite the different syllabi referred to by Chinese teachers and *Chi-Editor*, the interviewees noted similarities in leveling by intuition or with *Chi-Editor*, particularly with reference to the consideration of word difficulty and sentence length. The major difference between the two lied in the use of character information by *Chi-Editor* and the disregard of such information by teachers when leveling by intuition. However, studies on Chinese information processing show that Chinese characters contribute useful information to text leveling. These differences bring to light the question of how much weight character level should carry in text leveling by teacher intuition.

Second, interviewees benefited from the analysis and tagging provided by *Chi-Editor* during text adaptation process. It is interesting to note that a high percentage of comments made by the interviewees were on tagging. In general, they agreed that the tagged elements offered directly useful information for their consideration that is not necessarily available to them intuitively. For example, Interviewee C in Extract 7 mentioned said that she could "shorten or separate a sentence without hesitation" when using *Chi-Editor's* longest sentence tagging function, and that she no longer needed to keep "thinking about the situation if [she] shortened the wrong

sentence" as she would have done when using teacher intuition. Interviewee A in Extract 4 indicated that when not using *Chi-Editor*, she relied on her own instinct and "impressions," whereas she became more "confident" when using the tagging function in *Chi-Editor*.

Third, regarding the reporting function of *Chi-Editor*, the interviewees reported that the word list provided by *Chi-Editor* could help them better understand the levels of the words. They did not, however, seem to attach much value to the character list reported, mostly due to their preferences for the word-based approach to teaching. Additional comments indicated that the reporting function may be more practically useful with longer texts and that the reported word frequency and coverage information could be useful when editing or revising a text.

Overall, these findings show that *Chi-Editor* facilitates a text adaptation process that is in some ways similar to one that relies on teacher intuition, such as the use of word difficulty information for leveling and the focus on reducing difficult words and long sentences in adaptation. Meanwhile, it is critical to note that the data-driven approach and the intuitive approach should best be integrated in complementary ways. *Chi-Editor* is developed on the basis of the *GCSCW*, which was rooted in pedagogical practices in the first place. Utilizing data-driven technology, it provides rich, detailed information that can inform teachers' text adaptation process. At the same time, many other dimensions of text complexity are not yet captured by *Chi-Editor*, and teachers' expertise would certainly be necessary to complement *Chi-Editor* in considering those dimensions in text adaptation.

Acknowledgements This research was supported by the Fundamental Research Funds for the Central Universities (16wkpy09). We are indebted to Dr. Xiaofei Lu, Dr. Baichuan Li, and Dr. Le Wang for their suggestions and feedback on earlier drafts of the chapter.

References

Base for International Chinese Teaching Materials Developing and Teacher Training. (2017). https://www.cntexts.com/. Accessed July 9, 2017.

Confucius Institute Headquarters (Hanban). (2014). *I want to be a lawyer* (Level 3). Beijing: Beijing Language and Culture University Press.

Confucius Institute Headquarters (Hanban). (2015). *International curriculum for Chinese language education*. Beijing: Beijing Language and Culture University Press.

Crossley, S. A., Louwerse, M. M., McCarthy, P. M., & McNamara, D. S. (2007). A linguistic analysis of simplified and authentic texts. *The Modern Language Journal, 91*(1), 15–30.

Crossley, S. A., Allen, D., & McNamara, D. S. (2012). Text simplification and comprehensible input: A case for an intuitive approach. *Language Teaching Research, 16*(1), 89–108.

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. *California State Department of Education Office of Bilingual Education, Schooling and Language Minority Students: A theoretical framework*. Sacramento, CA: Author.

Ellis, R. (1999). *Understanding second language acquisition*. Shanghai: Shanghai Foreign Language Education Press.

Goodman, K. (1986). *What's whole in whole language*. Portsmouth, NH: Heinemann Educational Books.

Graesser, A., McNamara, D., Louwerse, M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments, and Computers, 36*(2), 193–202.

Green, A., & Hawkey, R. (2011). Re-fitting for a different purpose: A case study of item writer practices in adapting source texts for a test of academic reading. *Language Testing, 29*(1), 109–129.

Jin, T., & Li, B. (2016). *Chi-Editor: A data-driven tool for Chinese text adaptation*. Guangzhou: LanguageData. Retrieved from www.languagedata.net/editor.

Jin, T., & Lu, X. (2018). A data-driven approach to text adaptation in teaching material preparation: Design, implementation and teacher professional development. *TESOL Quarterly, 52*(2), 457–467.

Johnson, P. (1981). Effects on reading comprehension of language complexity and cultural background of a text. *TESOL Quarterly, 15*(2), 169–181.

Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Oxford University Press.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. London: Longman.

Lin, X. (2016). *The design and realization of an online tool for assessing text complexity of L2 Chinese reading materials*. (Unpublished undergraduate dissertation). Sun Yat-sen University, Guangzhou, China.

Liu, Y. L., & Ma, J. F. (2010). The development of the graded Chinese syllables, characters and words: Exploring the new perspective of global Chinese education. *Chinese Teaching in the World, 24,* 82–92.

Ministry of Education & State Language Commission, the People's Republic of China. (2010). *The graded Chinese syllables, characters and words for the application of teaching Chinese to the speakers of other languages*. Beijing: Language and Literature Press.

Sung, Y. T., Dyson, S. B., Chen, Y. C., Lin, W. C., & Chang, K. E. (2015). Leveling L2 texts through readability: Combining multilevel linguistic features with the CEFR. *The Modern Language Journal, 99*(2), 371–391.

Tomlinson, B., Dat, B., Masuhara, H., & Rubdy, R. (2001). EFL courses for adults. *ELT Journal, 55*(1), 80–101.

Wesman, A. G. (1971). Writing the test item. In R. L. Thorndike (Ed.), *Educational measurement* (pp. 99–111). Washington, DC: American Council on Education.

Yano, Y., Long, M., & Ross, S. (1994). Effects of simplified and elaborated texts on foreign language reading comprehension. *Language Learning, 44*(2), 189–219.

Young, D. J. (1999). Linguistic simplification of second language reading material: Effective instructional practice? *The Modern Language Journal, 83*(3), 350–366.