

December 2019

## A comparative study of the Chinese characters in the Graded list and the EBCL list

Qi Zhang  
*Dublin City University*

George Xinsheng Zhang  
*Centre for Modern Languages, Richmond University, London, UK*

Yuanhui Gao  
*The Education University of Hong Kong*

Follow this and additional works at: <https://engagedscholarship.csuohio.edu/cltmt>



Part of the [Chinese Studies Commons](#), and the [Modern Languages Commons](#)

[How does access to this work benefit you? Let us know!](#)

---

### Recommended Citation

Zhang, Qi; Zhang, George Xinsheng; and Gao, Yuanhui (2019) "A comparative study of the Chinese characters in the Graded list and the EBCL list," *Chinese Language Teaching Methodology and Technology*. Vol. 2: Iss. 2, Article 4.

Available at: <https://engagedscholarship.csuohio.edu/cltmt/vol2/iss2/4>

This Article is brought to you for free and open access by the Chinese American Faculty and Staff Association at EngagedScholarship@CSU. It has been accepted for inclusion in Chinese Language Teaching Methodology and Technology by an authorized editor of EngagedScholarship@CSU. For more information, please contact [library.es@csuohio.edu](mailto:library.es@csuohio.edu).

# A Comparative Study of the Chinese Characters in the Graded List and the EBCL List

Qi Zhang (Corresponding Author)

Dublin City University, Republic of Ireland

George Xinsheng Zhang

Richmond University, UK

Yuhui Gao

The Education University of Hong Kong, Hong Kong

## ABSTRACT

This study scrutinises the Chinese characters in two reference works compiled as part of two projects: the Graded list and the EBCL list. Situated in the character-based theory, the characters compiled in both are examined in depth from the perspectives of character complexity, character structure, semantic and phonetic radical transparency, frequency of character, and frequency of word formation. The results show that the two lists share similarities in terms of distribution of characters of different structures, level of semantic radical transparency, and frequencies of character and word formation. The character complexity of the Graded list is higher than that of the EBCL, while the level of phonetic radical transparency of the Graded list is lower than that of the EBCL list. With the Graded list, the high demands of character learning need to be borne in mind; this is due to the visual load of character complexity and the pronunciation information provided in phonetic radicals. Against a backdrop of increasing demand for CFL pedagogical materials, the differences and similarities between the two lists analysed and discussed in this study contribute to their pedagogical applications in teaching and learning Chinese characters.

*Keywords:* Graded list, EBCL, character, complexity, structure, transparency, frequency

## 1. Introduction

The learning and teaching of Chinese as a foreign language (henceforth CFL) has been rapidly growing in scale with the development of China's economy. Looking around the world, for example, there was a 41 percent increase in the number of students learning Chinese at The General Certificate of Secondary Education (GCSE) level in the UK between 2010 and 2017 (Tinsley & Board, 2017). Importantly, Chinese accounted for the largest number of A-level language entries in 2016 in the "other languages" category (ibid.). Similarly, Chinese is one of the five languages most often taught in foreign language programmes in the United States (Wiley & Garcia, 2016), and consequently funding has been allocated to secondary schools in the US to support the increasing demand for studying CFL (Zhou, 2011). Moreover, *Australian Curriculum: Languages Chinese* takes different learner groups into account and offers three different pathways for second language learners, heritage learners, and native speaking learners (Mollering, 2016). In Ireland, it has recently been announced that Chinese will be introduced as an exam subject on the Leaving Certificate curriculum within five years, as part of a 10-year strategy to improve the foreign language skills of Irish students (Department of Education and Skills, 2017; Osborne, Zhang, & Zhang, 2018). It is estimated that there are more than 1.1 million people that registered in Confucius Institutes and Classrooms<sup>1</sup> CFL classes at different levels worldwide (Luo & Guo, 2012).

Since this figure does not include students enrolled in non-Confucius courses, the actual number of CFL learners is higher.

The growth in the number of CFL learners has led to a need for CFL tests and teaching materials. Hanyu Shuiping Kaoshi (the Chinese Language Proficiency Test, henceforth HSK) is the only official examination offered by mainland China to evaluate CFL proficiency. By 2016, there were 1,066 HSK exam sites in 125 countries and regions, including 695 outside China. Through 2013, about 560,000 people had taken Chinese proficiency exams including the HSK around the world (Confucius Institute Annual Report, 2016). In addition to the HSK, there are also a variety of tests to examine CFL learners in China and around the world, such as the Test of Chinese as a Foreign Language (TOCFL) in Taiwan and the Assessment of Performance toward Proficiency in Chinese provided by the ACTFL (American Council on the Teaching of Foreign Languages, 2019). Under these circumstances, a large number of character lists have been published as supporting material for people taking those tests.

Set within the character-based theory, the current study aims to identify the similarities and differences of two-character lists – the Graded list and the European Benchmarking Chinese Language Proposed List (EBCL list) – and their potential pedagogical implications. The paper first reviews previous literature of the Graded list and the EBCL list. Using statistical analyses, the current study scrutinises the two-character lists from five aspects: character complexity, structure type, radical transparency, frequency of characters, and frequency of word formation, in order to in turn offer insights into future CFL material development and guidance on the character acquisition by CFL learners.

## 2. Literature Review of the Graded List and the EBCL List

One of the earliest character lists was the *List of Frequently Used Characters in Modern Chinese*, published in 1988 by the China National Language Committee to provide a reference for Chinese language education and dictionary design in China. The list of 3,500 characters was based on the principle of their being the most commonly used (Zhou, 2002).

The China National Language Committee consulted with nine of 27 modern Chinese dictionaries and ten common word lists from different resources; it also took into consideration the frequency of characters in a corpus between 1928 and 1986, as well as the frequency of word formation of these characters (Fu, 1988). According to Fu (1988), the list provides clear guidance for language practitioners and learners, since the mastery of the 2,500 most commonly used characters on the list is the threshold for managing daily communication in Chinese. This character list laid the foundation for the development of other character lists later (Zhu, 2013).

The earliest version of a character list for CFL pedagogy was published in 1992 by Hanban, entitled *Syllabus of Graded Word and Characters for Chinese Proficiency* (hereafter the Syllabus) (Sun, 2013). It offered references and guidelines to overall CFL pedagogy, teaching material design, in-class practice and language evaluation (Zhou & Li, 2008). However, even after revisions and a new edition were published in 2001, the Syllabus no longer reflects language use in real life (ibid.) and cannot cope with the changing environment of international Chinese education (Sun, 2013). Under these circumstances, *Graded Chinese Syllables, Characters, and Words for the Application of Teaching Chinese to the Speakers of Other Languages* was compiled and published in 2010. It was the first list to take character syllables into consideration when being compiled (Hanban, 2010). Therefore, the Graded project is a “three-dimensional standard system” consisting of syllable, character, and vocabulary (Hanban, 2010). In addition to offering guidance for CFL character acquisition, it has also contributed to the development of syllable corpus and computer-

assisted CFL tests (Li, 2011). The Graded project is a key NOCFL research project that can be extensively adapted and widely used for CFL teaching (Liu & Ma, 2010). For example, it offers guidelines for designing textbooks, classes, and tests for both CFL practice internationally and the MHK in China, as well as being a reference for Chinese dictionaries (ibid.).

The Graded project is mainly drawn from two sources. The first is a large-scale dynamic discourse corpus consisting of 3.7 billion Chinese characters (Li, 2011), including TV and radio conversational discourse, auditory media discourse, and print and online media discourse (Hanban, 2010). The second source is various character and vocabulary lists, such as those published in 1988 and 1992, *The Key to Chinese Speech and Writing* by Joel Bellassen and Pengpeng Zhang (1997), and dictionaries such as *The Essential Chinese Dictionary* by Xu and Yao (2007). Experts were also invited to help make sure that the characters were commonly used in real communication (Hanban, 2010).

Characters compiled in the Graded project are classified into three levels (Hanban, 2010). A list of 900 characters is categorised as Level 1/basic level, so the Graded list<sup>ii</sup> in this study refers to these 900 characters for CFL beginner learners. Another 900 characters are categorised as Level 2/intermediate level, and an additional 900 characters are categorised as Level 3/advanced level.

The Graded list (the 900 characters for CFL beginners) investigated in this study is built primarily on eight sub-lists of characters. Five of the lists come from a large dynamic discourse corpus. The other three are (i) the most frequent 900 characters from a conversational discourse corpus, (ii) 900 characters from *The Key to Chinese Speech and Writing*, and (iii) a preliminary list of 900 characters for further investigation (Liu & Ma, 2010). Characters were compared across these eight sub-lists and eventually 900 were selected, with less than 5% entailing consultation with experts (ibid.). That is to say, a very low level of subjectivity was involved in deciding characters for the Graded list.

The EBCL List is an EU-funded project involving partners such as Rennes II in France, Freie University Berlin in Germany, the University of Rome in Italy and SOAS University of London in the UK (EBCL, 2019). The project started in 2010 against the backdrop of increasing demand for Chinese language courses and a need for consistency and standardisation in Chinese language pedagogy in Europe (Wang, Song, & Suen, 2012; Zhang, 2011). Two Chinese language proficiency levels – A1 and A2 – have been released on the EBCL official website (EBCL, 2019). Therefore, the EBCL list discussed in this study refers to the A1 and A2 characters published in 2012 as part of the EBCL project's 'Can-do Statements'. There are 320 characters in A1 and 630 characters in A2. A2 includes the characters in the A1 list, as well as new ones; that is to say, there are 320 characters in A1 and an additional 310 in A2. This study examines these 630 characters.

The EBCL list was prepared through a comparison of four existing character lists for CFL beginners (EBCL Character Methodology; Allanic, 2012):

- (1) Threshold of 255 “active characters” (or “productive characters”)<sup>iii</sup> for Chinese Teaching Programme LV3 (Chinese as third foreign language) published by the Official Bulletin developed under the direction of Joel Bellassen.
- (2) Threshold of 202 “active characters” (or “productive characters”) and 98 “passive characters” (or “receptive characters”) for Chinese Teaching Programme Middle Schools Level 1 developed under the direction of Isabelle Pillet and Joel Bellassen.
- (3) 487 characters constituting the 581 EBCL core lexical items (EBCL Vocabulary Methodology; Allanic & Shu, 2012).
- (4) 300 characters selected from the Graded list by a panel of international experts.

As both the Graded list and the EBCL list are key references for Chinese character study, it is understandable that overlapping materials were used to compile these lists. For

example, resources developed by Joel Bellassen – *The Key to Chinese Speech and Writing* and *Chinese Teaching Programme* – were used in both lists. Even though the Graded list was one of the key resources when the EBCL list was prepared, the latter takes into consideration other character lists and pedagogical materials used as references in Europe, in order to develop a list appropriate to and applicable in the European context. This paper aims to investigate the characters in the Graded and EBCL lists in order to identify their similarities and differences, to provide valuable information for the future application of each list.

The Common European Framework of Reference for Language (CEFR) was published by the Council of Europe in 2001. It is used as the guideline for language teaching and examination setting and has been influential in Europe (Hulstijn, 2007). The CEFR outlines three levels of language proficiency: A for Basic user, B for Independent user, and C for Proficient user. Each level can be further divided into two grades, meaning there are six grades in total (Council of Europe, 2018). The CEFR is now widely used to indicate proficiency levels in the development of curricula, assessment instruments and exams (Little, 2013). As mentioned earlier, the EBCL project was built upon CEFR and developed the A1 and A2 levels (EBCL, 2019; see Table 1).<sup>iv</sup>

The New HSK introduced by Hanban in 2009 also has six grades (Zhang, Xie, Wang, Li, & Zhang, 2010). It is suggested that these six grades are equivalent to those outlined in the CEFR (Hanban Test Centre, 2019). The following table exhibits the corresponding levels between the New HSK and the CEFR, as well as between the EBCL and the CEFR (Hanban Test Centre, 2019; see also Table 1).

Table 1. *The grades of new HSK and EBCL, with CEFR equivalent*

Graded <sup>a</sup>	New HSK	CEFR	EBCL
Level 3 (Advanced)	HSK (6)	C2	–
Level 2 (Intermediate)	HSK (5)	C1	
	HSK (4)	B2	
Level 1 (Basic) <sup>b</sup>	HSK (3)	B1	
Level 1 (Basic) <sup>b</sup>	HSK (2)	A2	A2
	HSK (1)	A1	A1
Note:	- No further information released from the official website a. As shown later, further research is needed to examine if this is the best way to match the levels in this column with others, and it is therefore shaded grey. b. The characters in this level will be compared with those in A1 and A2 in EBCL.		

However, the levels benchmarked against the CEFR in the above table are indeed self-claimed by each of the projects and therefore can be disputed. In addition, it is difficult to see the corresponding levels between the Graded project and the other three (the New HSK, CEFR and the EBCL project). Although only words rather than characters are stated at each level of the New HSK, Lü (2010) summarises the number of characters involved in the new HSK 1, 2, and 3 based on the words provided (see Table 2). If the number of characters is

used as a reference, even the basic level in the Graded project is equivalent to the levels above HSK 4 (see Table 2).

Table 2. *The grades from the Graded Project, the New HSK and the EBCL Project, based on character numbers*

Graded Project (characters)		New HSK (characters)		EBCL Project (characters)	
Advanced	(2700)	HSK 6	-	-	
Intermediate	(1800)	HSK 5	-	-	
Basic	(900)	HSK 4	-	-	
		HSK 3	(623)	A2	(630)
		HSK 2	(344)	A1	(320)
		HSK 1	(172)		
Note:	- No information provided in the official document or in previous research.				

According to Hanban (2010, p. III), Level 1 in Graded is called “Basic level, which can also be [interpreted as] a level of popularisation [the Chinese language]” (一级 (初级, 也称为普及化等级) [yiji, ye chengwei pujihua dengji]).<sup>v</sup> In contrast, ‘basic’ and ‘simple’ are mentioned in the descriptions of HSK 2 and HSK 3 (Hanban Test Centre, 2019).

HSK 2: Test takers who are able to pass the HSK (Level II) have an excellent grasp of basic Chinese and can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters.

HSK 3: Test takers who are able to pass the HSK (Level III) can communicate in Chinese at a basic level in their daily, academic and professional lives. They can manage most communication in Chinese when travelling in China.

(Hanban Test Centre)

In other words, when the descriptions of different levels are compared, the Basic level in the Graded project seems to be equivalent to HSK 2 or HSK 3, which can be matched to the A1 and A2 levels in the EBCL project (see Table 1). EBCL supporting documents indicate that CFL learners should have acquired 320 characters to reach level A1 and a total of 630 characters to pass level A2. The 630 characters from the EBCL list and the 900 from the Graded list seem to offer a guideline for CFL learners when studying Chinese characters in order to become a basic user of the language. On the other hand, it is worth comparing the two lists of characters in order to gain an in-depth understanding of how the Graded list is related to the EBCL list, so as to compare the guidance they offer to CFL beginners in terms of the study of Chinese characters.

### 3. Character-based Theory

The character-based theory is built upon the notion that the foundation of the Chinese language is the character (Wang, 2000). While a word is the smallest meaningful unit in alphabetic language systems, a character is the smallest meaningful unit in Chinese (Zhang, 1992). A character can constitute a word in itself, though a combination of various characters also makes up a Chinese word (Osborne, Zhang, & Zhang, 2018; Sun, 2006). The character-based theory therefore emphasises this difference between Chinese, as a Sino-Tibetan language, and Indo-European languages (Lincoln, 2015; Lü, 2010; Wang, 2006; Wang, 2000; Yang, 2011; Zhang, 1992).

The orthographic structure of Chinese characters consists of three tiers: several strokes construct a radical, and one or more radicals form a character (Shen & Ke, 2007). Approximately 80% of characters in modern Chinese are semantic-phonetic compound characters (Kuo et al., 2015), each of which involves a semantic radical providing cues to the meaning of the character and a phonetic radical indicating the pronunciation. For example, the character 姑 (gū – aunt) is a compound character with 女 (nǚ – woman) as the semantic radical on the left and 古 (gǔ – old) as the phonetic radical on the right. The majority of semantic-phonetic compound characters have a left-right and top-down structure (Shu, Chen, Anderson, Wu, & Xuan, 2003).

However, the correspondence between a phonetic radical and the actual pronunciation of the character that the radical is within is unreliable and unsystematic (Xu, Chang, Zhang, & Perfetti, 2013). As a result, it is recommended that characters be the primary focus in CFL learning (Wang, 2000). One of the applications of the character-based approach in CFL pedagogy entails learning Chinese characters by acquiring meaning and pronunciation through the writing of characters (Lincoln, 2015). The importance of character writing has also been pointed out in experimental psychology research, especially its impact on reading competence (Cao et al., 2013; Zhang & Reilly, 2015). The repeated writing practice helps learners to develop a motor memory trace, which may ultimately assist CFL learners in connecting the meaning of a character with its semantic radical, as well as linking the character's pronunciation with the phonetic radical. In other words, the logographic nature of Chinese language leads us to emphasise the study of characters, which is a way to respect the characteristics of the language (Lü, 2015).

Due to the pictographic origin of the Chinese writing system, the traditional approach to studying Chinese language acknowledged slower development in character writing compared to character recognition, and consequently suggested a separation of the teaching of reading and that of writing. In pre-PRC China, Chinese children learned Chinese characters through reading 'Three, Hundred, Thousand': *Three Character Scripture* (三字经), *Hundred Family Names* (百家姓) and *A Thousand Characters* (千字文). As the names suggest, Chinese children learned to read approximately two thousand characters through these textbooks. On the other hand, they learned to write much simpler characters (e.g., with fewer strokes) (Lam, 2011).

The development of CFL pedagogical material also takes into account the characteristics of the Chinese writing system, and a few Chinese language textbooks have adopted the character-based approach. For example, almost all Chinese language teaching materials in France follow the example set by *Méthode d'initiation à la langue et à l'écriture chinoises* (Bellassen, 1989), practising character-based theory, such as *Snowballing Chinese* (Lü, 2015; Wang, 2000). The Graded project also moved away from the conventional way of "prioritising words over characters" used in the *Syllabus* in 1992 (Hanban, 2010, p. XI). Instead, the sifting of characters was put first and was considered the most important step when the graded syllables, characters, and words were compiled in the Graded project (Liu & Ma, 2010).

The current study is situated in the character-based theory, with the view that characters, as well as components of characters (radical, structure, etc.), are an integral part of learning the Chinese language. The main goal of the research is to scrutinise the differences and similarities of two lists of characters from the perspectives of character complexity, character structure, radical transparency, frequency of character, and frequency of word formation. In this case, the study will provide valuable details for the future applications of characters from the two lists, such as for the development of pedagogical materials built upon the character-based approach. As shown in Section 2, the Graded list and the EBCL list were compiled generally based on three methods: overlapping characters from different existing character lists, frequently used characters computed from large-size corpora,<sup>vi</sup> and characters chosen by consultants based on their expertise and experience. It remains unknown whether and to what extent the two lists differ from each other in terms of the nature of the characters. This study therefore focuses on the characters themselves and investigates the two lists based on the nature of the characters shown below, particularly their similarities and differences.

## 4. Methodology

Of the 900 characters of the Graded list and the 630 of the EBCL list, 586 overlap (see Figure 1). Due to the difference in the total number of characters on the two lists, this study employed IBM SPSS (Version 24) to conduct various statistical analyses, rather than simply examining the descriptive data.

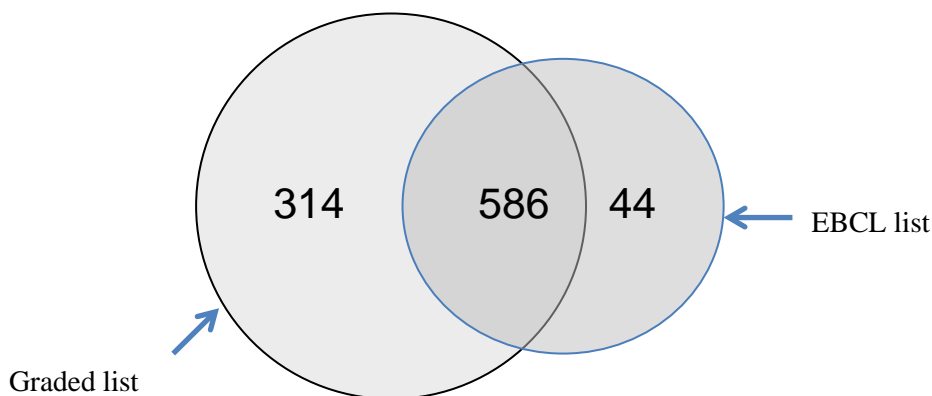


Figure 1. Overview of the number of characters in the Graded and EBCL lists

The nature of each character on the Graded and EBCL lists is analysed based on five aspects: (1) character complexity, (2) structure type, (3) radical transparency, (4) frequency of characters, (5) frequency of word formation. The former three aspects can indicate the level of difficulty of the characters in a list that are expected to be acquired by CFL learners, whereas the latter two aspects suggest the level of usefulness of each list.

(1) Character complexity refers to the number of strokes (Changizi & Shimojo, 2005; Shu et al., 2003). As the human visual system tends to prefer character lengths of approximately three strokes which can be stored in the visual short-term memory (Changizi & Shimojo, 2005), complexity can affect the learning of a character.

(2) Chinese characters can be categorised into two types – integral and compound – based on structure. Integral characters cannot be further divided into radicals, whereas compound characters are usually composed of two or more radicals (Wang, Wu, Zhao, Ni, & Zhang, 2016; Zhang & Reilly, 2015).



Based on the position of the radicals in the character, compound characters can be further classified into four structures: top-bottom (including top-middle-bottom), left-right (including left-middle-right), half-surround and surround. All characters from the Graded and EBCL lists can be categorised and coded into five structures: integral, top-bottom, left-right, half-surround, and surround. In addition to character complexity, there has been research showing the potential effect of character structure on the acquisition of Chinese characters (Liu, 2008; Zhang, 2008). Therefore, the types of character structures were also considered in the current study.

(3) As explained earlier, a phonetic radical usually indicates the pronunciation of a semantic-phonetic compound character, whereas a semantic radical contributes to the meaning of this character. Previous studies (e.g., Wang, Wu, Zhao, Ni, & Zhang, 2016; Zhang, Wang, & Yin, 2014) have found that semantic and phonetic radicals can have different effects on character recognition. Adapted from the classification in Shu et al. (2003), radicals can be coded into four levels: transparent, semi-transparent, opaque, and unclassified (see Table 3), depending on the amount of information that a person can rely on to detect the meaning or pronunciation of a character. The more transparent a radical is, the more reliable the clue it provides to learners in terms of the meaning or pronunciation of a character.

Table 3. *Character radical transparency classification (adapted from Shu et al. 2003)*

	Semantic Transparency	Phonetic Transparency
Transparent	The character has the same meaning as its radical; or the character belongs to the category that its radical represents; or the meaning of the character is directly related to the meaning of its radical.	The character has the same pronunciation as its phonetic, including the tone; or the character has the same syllable as its phonetic, but a different tone.
Semi-transparent	The meaning of the character is indirectly related to the meaning of its radical; or the extended meaning of the character is directly or indirectly related to the meaning of its radical.	The character has the same final as its phonetic, but a different onset; or the character has the same onset as its phonetic, but a different final.
Opaque	The meaning of the character is unrelated to the meaning of its radical.	Either the character or the phonetic radical has more than one pronunciation; or the character is pronounced with a totally different syllable from its phonetic, but the tone can be the same.
Unclassified	It is difficult to define the radical of a character due to simplification or other reasons.	The character lost its original phonetic at some stage in the past.

In other words, character complexity, structure, and radical transparency can influence the acquisition of Chinese characters to some degree and so can serve as indicators of the level of character difficulty from different aspects. Therefore, this study analyses characters

from these three aspects in order to see whether and to what extent the two lists of characters have a similar level of character difficulty. Such an investigation of the EBCL and the Graded lists can also demonstrate the different expectations of CFL learners set by the two projects in terms of character acquisition.

(4) The frequency of characters refers to the frequency of occurrence of each character in the modern Chinese language. The more frequently a character is used, the more likely it is that a learner needs to use it for daily communication, and so the more useful it is. As character frequency is a common factor that CFL pedagogical materials take into consideration (Jiang, 2006; Li, 2014), both the EBCL and Graded lists are built upon previous character lists that have calculated the character frequency from large-size corpora. However, it remains unknown if the two lists have the same level of or different levels of frequency.

(5) Frequency of word formation refers to the number of times a character appears with one or more other characters to form a word. It can show the productivity of a character (i.e., word-building ability) (Jiang, 2006). Therefore, the study of characters with high frequency of word formation is very likely to benefit learners, in that they acquire more words through their reserve of characters. Please note that the calculation of frequency of word formation does not include single-syllable vocabulary or single-character words. Take the character 的 (de), for example. It can be a single-character word, 的 (de), used as an auxiliary word. It can also combine with another character to form the two-character word 的士 (dīshì – taxi). The former is not included in the calculation of frequency of word formation.

Frequency of character and frequency of word formation are calculated through programming in Perl script in a corpus consisting of 1.25 million sentences (2.79 million words) extracted from LDC corpora.<sup>vii</sup> As shown above, both frequencies are helpful in indicating how useful characters are for real-life communication or for assisting in the study of new words.

## 5. Data Analysis

### 5.1 Character Complexity

Following the definition from previous research (Changizi & Shimojo, 2005; Shu, Chen, Anderson, Wu, & Xuan, 2003), the number of strokes of a character is used here to indicate the complexity of the character. In terms of complexity, Table 4 shows the descriptive data from the Graded and the EBCL lists. A one-way ANOVA was conducted to examine if the difference in character complexity reached statistical significance.

Table 4. *Descriptive results of character complexity in the Graded and EBCL lists*

	N	Mean	95% Confidence Interval for Mean		Minimum	Maximum
			Lower Bound	Upper Bound		
Graded	900	7.95	7.75	8.14	1	19
EBCL	630	7.59	7.37	7.82	1	16
Total	1530	7.80	7.65	7.95	1	19

Levene's test of the homogeneity of variances shows that the variances in two groups are equal ( $p = .617$ ). The one-way ANOVA test result indicates that the character complexity of the Graded list (mean = 7.95) is significantly different from that of the EBCL list (mean = 7.59):  $F(1,1528) = 5.47, p = .019$ . That is to say, the character complexity of the Graded list is significantly higher than that of the EBCL list.

## 5.2 Character Structure

All characters in both lists are categorised into five structure types. Since the lists are intended for CFL beginners, there is also a sizable number of integral characters (17.7% and 21.4, respectively) which are less complex and may be easier to learn (see Table 5). In addition, there are 741 compound characters in the Graded list and 495 in the EBCL list. The analysis shows that the most common type of compound character structure in both the Graded and the EBCL list is the left-right structure (44.9% and 43.3%), followed by the top-bottom structure (26.7% and 24.9%).

Table 5. *Character structure in the Graded and EBCL lists*

Type of character	Graded	%	EBCL	%
Left-right	404	44.9	273	43.3
Top-bottom	240	26.7	157	24.9
Half-surround	90	10.0	60	9.5
Surround	7	0.8	5	0.8
Sub-total	741		495	
Integral	159	17.7	135	21.4

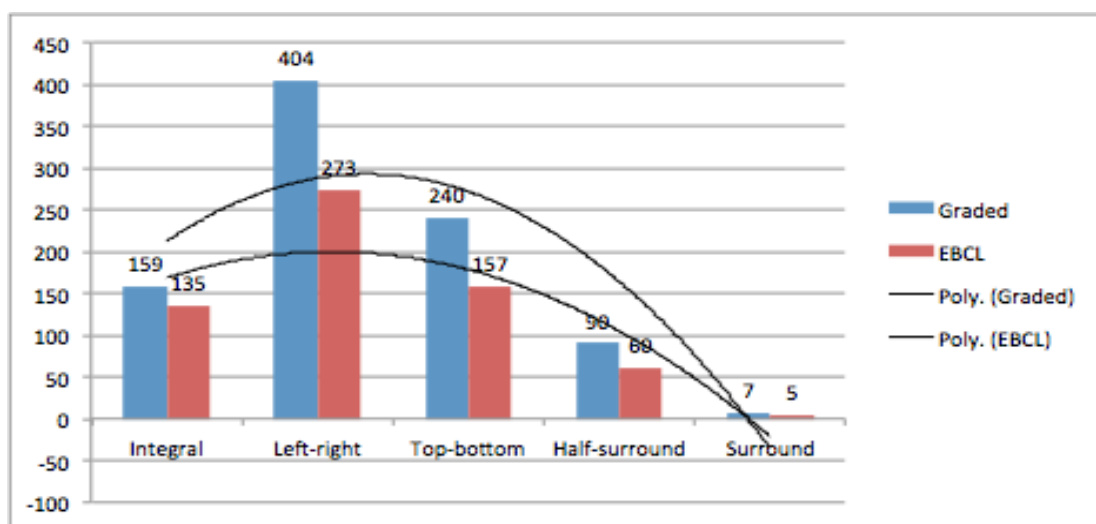


Figure 2. Character structure in the Graded and EBCL lists

As shown in Figure 2, the Graded list and the EBCL list demonstrate a similar tendency in the distribution of character structures. A non-parametric test – the Mann-

Whitney test – was carried out to examine if the differences between two lists are statistically significant. The result shows that the Graded list does not seem to differ from the EBCL list in the distribution of character structure:  $U = 271879.00$ ,  $z = -1.45$ ,  $p = .147$ ,  $r = -.037$ .

### 5.3 Radical Transparency

All the compound characters in the two lists were analysed according to radical transparency. Please note that integral characters are also included for the analysis. As a result, around half of the characters on both lists cannot be classified in terms of the semantic transparency of characters (see Table 6). 21.8% of characters on the Graded list and 21.7% of characters on the EBCL list have opaque semantic radicals, followed in frequency by semi-transparent radicals (18.1% and 18.7% respectively). The fewest characters have transparent semantic radicals (4.8% and 4.1%).

Table 6. *Transparency of semantic radicals in the Graded and EBCL lists*

Semantic transparency	Graded	%	EBCL	%
3 Transparent	43	4.8	26	4.1
2 Semi-transparent	163	18.1	118	18.7
1 Opaque	196	21.8	137	21.7
Sub-total	402		281	
UNC	498	55.3	349	55.4

Also due to the fact that the integral characters were considered, the majority of characters have unclassified phonetic transparency in the Graded and EBCL lists (95.2% and 64.9% respectively). As shown in Table 7, the largest number of characters have transparent phonetic radicals (2.2% and 17.5%), followed by semi-transparent radicals (1.4% and 11.4%). The fewest characters have opaque phonetic radicals (1.1% and 6.2%).

Table 7. *Transparency of phonetic radicals in the Graded and EBCL lists*

Phonetic transparency	Graded	%	EBCL	%
3 Transparent	20	2.2	110	17.5
2 Semi-transparent	13	1.4	72	11.4
1 Opaque	10	1.1	39	6.2
Sub-total	43		221	
Unclassified	857	95.2	409	64.9

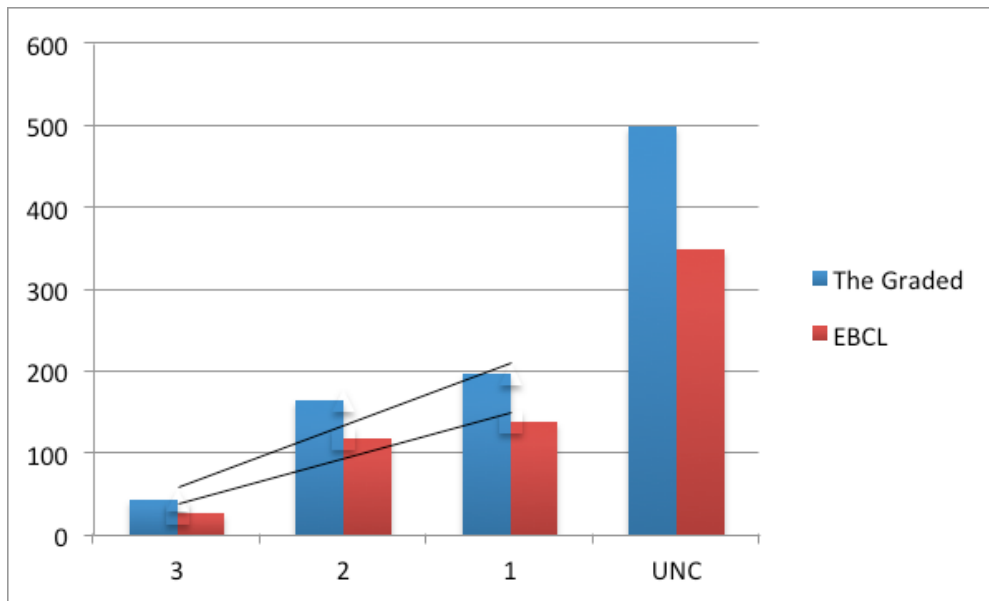


Figure 3. Transparency of semantic radicals in the Graded and EBCL lists

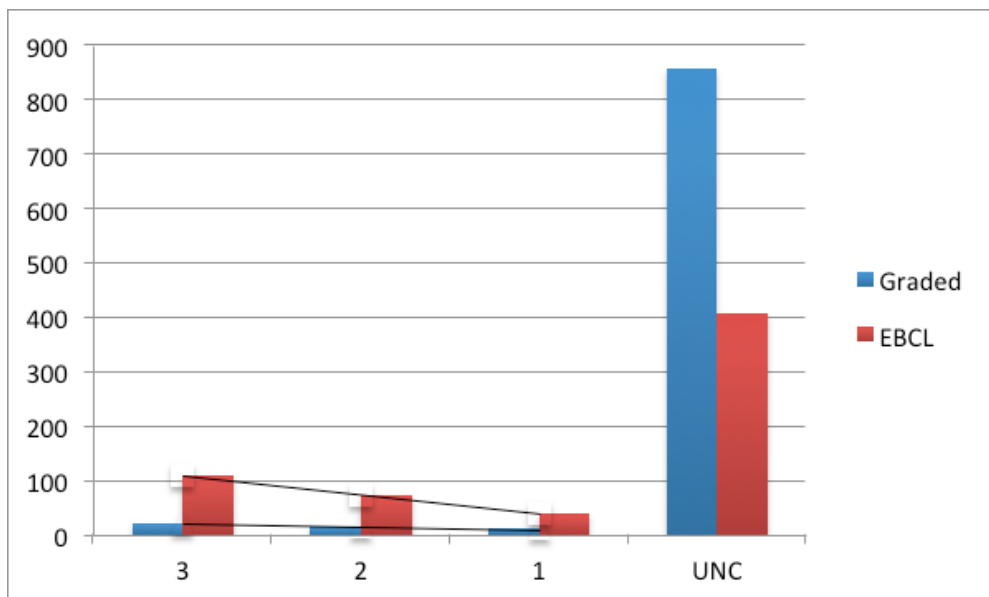


Figure 4. Transparency of phonetic radicals in the Graded and EBCL lists

As shown in Figure 3, the two lists have a similar distribution of characters in terms of semantic radical transparency. However, they seem to show different patterns of character distribution in terms of phonetic radical transparency (see the trend lines in Figure 4): while the character distribution for the EBCL list demonstrates a decreasing trend, that for the Graded list was likely to be level across 1-3 transparency. A non-parametric test – the Mann-Whitney test – was carried out to examine if the differences between two lists are statistically significant in terms of radical transparency. The semantic radical transparency of the Graded list was not significantly different to that of the EBCL:  $U = 282903.00$ ,  $z = -.078$ ,  $p = .938$ ,  $r = -.02$ . In contrast, the phonetic radical transparency of the EBCL list was significantly higher than that of the Graded list:  $U = 197333.50$ ,  $z = -15.402$ ,  $p = .00$ ,  $r = -.394$ .

### 5.4 Frequency of Character and Frequency of Word Formation

The average frequency of character is 39022.11 for the Graded list and 43989.65 for the EBCL list (see Table 8). There is unlikely to be a vast difference in both lists in terms of character frequency, which is confirmed by an independent t-test:  $t(1225.85) = -1.249$ ,  $p = .21$ . As mentioned earlier, one of the key methods employed by the Graded project was to calculate the frequency of characters in large-size corpora in order to develop the character list. The EBCL list was also compiled with suggestions based on the frequency of character occurrence in Chinese language. Therefore, it is unsurprising to find that both lists demonstrate a similar level of character frequency.

Table 8. *Descriptive results of character frequency in the Graded and EBCL lists*

	N	Mean	95% Confidence Interval for Mean		Minimum	Maximum
			Lower Bound	Upper Bound		
Graded	900	39022.11	34443.20	43601.03	19	1379119
EBCL	630	43989.65	37663.15	50316.15	19	1379119
Total	1530	41067.57	37322.84	44812.29	19	1379119

The average frequency of word formation is 30608.02 for the Graded list and 33359.32 for the EBCL list (see Table 9). This is not a vast difference, and is supported by an independent t-test:  $t(1254.2) = -1.178$ ,  $p = .239$ . Both lists aim to include characters that are frequently used in modern Chinese language. High frequency of characters indicates that these characters frequently occur in many words. It is therefore unsurprising that the two lists share a similar level of frequency of word formation, based on a similar level of character frequency.

Table 9. *Descriptive results of frequency of word formation in the Graded and EBCL lists*

	N	Mean	95% Confidence Interval for Mean		Minimum	Maximum
			Lower Bound	Upper Bound		
Graded	900	30608.02	27864.70	33351.34	19	486552
EBCL	630	33359.32	29686.78	37031.86	19	486552
Total	1530	31740.91	29530.90	33950.92	19	486552

## 6. Discussion

Five aspects of all characters – complexity, structure, transparency, frequency of character and frequency of word formation – were analysed. Drawing on the findings

outlined above, this section focuses on discussing main similarities and differences identified between two lists and their pedagogical implications. The highest percentage of characters on both lists are left-right structures, and the least common type of characters are surround structures (see Table 5). As shown in Table 5, there are 741 compound characters on the Graded list, in comparison to 495 on the EBCL list. Of these 741 and 495 compound characters, the majority have the semantic radical on the left (272 on the Graded list and 199 on the EBCL list) and the phonetic radical on the right (603 and 193).<sup>viii</sup> According to Shu et al. (2003), over 70% of compound characters have a left-right structure with the semantic radical on the left and phonetic radical on the right. That is to say, although the Graded and EBCL lists contain a mere fraction of all Chinese characters, they are similar in a way of reflecting the nature of Chinese characters in general.

Furthermore, previous research indicates that exposure to a large number of semantic-phonetic structures is useful for beginners, in that it helps them to develop a good sense of the position constraints of Chinese characters (Anderson et al., 2013; Tong & McBride, 2014), which contributes to character learning. Therefore, a reasonable number of semantic-phonetic compound characters on the Graded and EBCL lists can help CFL beginners in their study of the Chinese language.

The second similarity between the Graded and EBCL lists is that they have similar character occurrence frequency and word formation frequency. The frequency effect usually suggests to what extent character acquisition has taken place naturally and implicitly through exposure to the Chinese language. The advancement of technology makes it possible to calculate the frequency of characters in large-size corpora. Since both lists are built upon such big data analysis, they consist of characters that are frequently used in modern Chinese society, and are consequently fairly useful in offering guidance for CFL beginners. The study of these characters contributes to learning Chinese in order to cope with basic communication in daily life.

Although both lists have a similar distribution of characters in terms of character structure (see Figure 2 in Section 4.2) and semantic radical transparency (see Figure 3 in Section 4.3), there are differences between them in two aspects: character complexity and phonetic radical transparency. The character complexity of the Graded list is significantly higher than that of the EBCL list (see Table 4 in Section 4.1). The more complex a character is, the higher visual load there is for a CFL beginner to process. Previous studies show that character complexity can affect the learning of Chinese characters, especially in writing (Feng, 2002; Liu, 2008; You, 2003). It is therefore more appropriate for CFL beginners to first learn integrals and characters of less complexity, and to then learn more complex characters. In this case, the EBCL list seems to be more thoughtful from this perspective.

In addition, while the difference in semantic transparency between them does not reach statistical significance, the phonetic transparency of the EBCL list is higher than that of the Graded list (see Figure 4 in Section 4.3). The more transparent a phonetic radical is, the more information it provides to learners regarding pronunciation. Previous research suggests that, in comparison with semantic radicals, phonetic radicals may make a more significant contribution to the recognition of Chinese characters (Wang, Wu, Zhao, Ni, & Zhang, 2016; Zhang, Wang, & Yin, 2014). The relatively low phonetic radical transparency of the Graded list may to some extent impose difficulties in learning compound characters. In comparison, the EBCL list may better help CFL learners to acquire the pronunciation of semantic-phonetic compound characters.

Both lists include more compound characters with different levels of semantic radical transparency (44.7% for the Graded list and 44.5% for the EBCL; see Table 6 in Section 4.3) than those with a certain level of phonetic radical transparency (4.7% and 35.1% respectively; see Table 7 in Section 4.3). Although it may be convenient to suggest that the

potential positive effect of phonetic radical transparency on Chinese language learning should have been considered when compiling these lists, empirical research is needed in order to decide how many characters with some sort of phonetic radical transparency a list for CFL beginners should include.

Further to the examination of the characters themselves, it is worth discussing the pedagogical implications of these findings. Character lists are usually employed as a reference for curriculum planning, textbook design, and assessment development. The application of these two lists may not vary significantly in relation to character structures and frequency of occurrence. As discussed earlier, a substantial number of semantic-phonetic structures and frequently encountered characters from learners' social milieus contribute to the study of Chinese language. However, specific to textbook design, this also depends on the repeated occurrence of a character in a textbook in order to enhance its memorisation. CFL teachers need to take this into account in language pedagogy, for example, by encouraging repeated use of new vocabulary through activities and tasks inside and outside the classroom.

Since the character complexity of the Graded list is significantly higher than that of the EBCL list, any curriculum or assessment using the Graded as a reference implies higher demands on students in terms of character composition, in particular for writing. While both lists show a similar level of semantic radical transparency, the phonetic radical transparency of the Graded list is significantly lower than that of the EBCL list. The advantage of phonetic radical transparency when it comes to character recognition is less exploited in the Graded than in the EBCL list. Interestingly, the Graded project is the first to offer a three-dimension system of syllable, character, and vocabulary. In other words, as well as the character list examined in the current study which mainly considers the frequency of character occurrence, the Graded project also provides a syllable list which consists of characters with frequently used syllables. Further research is recommended to investigate whether phonetic radical transparency – which, even unsystematically, is related to a character's syllable – is better presented in the Graded syllable list.

## 7. Conclusion

This study entailed conducting an in-depth examination of the characters on the Graded list and the EBCL list. Similarities between the two lists include the distribution of characters in terms of five structures, the level of phonetic radical transparency, and the frequency of character and word formation. Although the exact equivalence between the Graded list and the EBCL list remains unclear (see Table 1 in Section 2), these similar traits may suggest that the two lists do indeed provide guidance to CFL beginners. Although the character complexity of the Graded list is higher than that of the EBCL, further research is needed to investigate what level of character complexity is appropriate for CFL beginner learners. Both lists contain more characters with a certain level of semantic radical transparency than characters with a certain level of phonetic radical transparency. The EBCL list, which is primarily for European learners, demonstrates better phonetic radical transparency than the Graded list, which is intended for teaching Chinese to speakers of all languages. The Graded list is clearly designed to be suitable for CFL learners worldwide and to be compatible with computer-assisted Chinese language testing (Hanban, 2010).

This study is one of the first to be based on the character-based theory through its emphasis on the character itself, and therefore focuses on characters only. Other important factors, such as the ratio of characters to words – an indicator of the use of characters as morphemes to construct words – are not explored. In addition, the two lists have been compiled with different intentions and principles. The Graded is designed for all CFL learners worldwide, and built upon character frequency. In contrast, the EBCL is designed



specifically for European learners, and emphasises characters used in words, with its character selection procedure based on character occurrence in vocabulary items. Therefore, it is too soon to conclude that one list is better than the other and future research may further examine character lists from other perspectives, such as considering other factors (e.g., the ratio of characters to words) or compiling principles.

Nevertheless, the current study is one of the first to compare two lists of characters against the backdrop of increasing demand for CFL pedagogical materials. The significance of such research is that it can enhance the understanding of both lists, for the better application of both. It is hoped that the differences and similarities between the two lists analysed and discussed here will shed some light on the pedagogical application of the two lists in the learning and teaching of Chinese characters.

### References

- Allanic, B. (2012). *EBCL character methodology: Method of preparing the character list for Level A1*. Retrieved from <http://ebcl.eu.com/wp-content/uploads/2012/04/Methodology-A1-character-list.pdf>
- Allanic, B., & Shu, C. (2012). *EBCL vocabulary methodology: Methodology of EBCL vocabulary work (A1-A2)*. Retrieved from <http://ebcl.eu.com/wp-content/uploads/2012/05/EBCL-Vocabulary-Methodology.pdf>
- American Council on the Teaching of Foreign Languages. (2019). *ACTFL assessment of performance toward proficiency in languages*. Retrieved from <https://www.actfl.org/assessment-professional-development/assessments-the-actfl-testing-office/aappl>
- Anderson, R. C., Ku, Y.-M., Li, W., Chen, X., Wu, X., & Shu, H. (2013). Learning to see the patterns in Chinese characters. *Scientific Studies of Reading*, 17(1), 41–56. <https://doi.org/10.1080/10888438.2012.689789>
- Bellassen, J. (1989). *Méthode d'initiation à la langue et à l'écriture chinoises (Initiation to Chinese language and writing)*. Paris: La Compagine.
- Bellassen, J., & Zhang, P. (1997). *A key to Chinese speech and writing*. Beijing: Chinese Language Teaching Press.
- Cao, F., Vu, M., Chan, D. H. L., Lawrence, J. M., Harris, L. N., Guan, Q., Xu, Y., Perfetti, C. A. (2013). Writing affects the brain network of reading in Chinese: A functional magnetic resonance imaging study. *Human Brain Mapping*, 34(7), 1670–1684. <https://doi.org/10.1002/hbm.22017>
- Changizi, M. A., & Shimojo, S. (2005). Character complexity and redundancy in writing systems over human history. *Proceedings of the Royal Society B: Biological Sciences*, 272(1560), 267–275. <https://doi.org/10.1098/rspb.2004.2942>
- Confucius Institute Annual Report. (2016). *Confucius Institute annual development report 2016*. Retrieved from Confucius Institute website: <http://www.hanban.org/report/2016.pdf>
- Council of Europe. (2018). *Common European framework of reference for languages: Learning, teaching, assessment* (p. 235). Retrieved from Council of Europe website: <https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989>
- Department of Education and Skills. (2017). *Languages connect: Ireland's strategy for foreign languages in education 2017-2026*. Retrieved 24 September 2018, from [https://www.education.ie/en/Schools-Colleges/Information/Curriculum-and-Syllabus/Foreign-Languages-Strategy/fls\\_languages\\_connect\\_strategy.pdf](https://www.education.ie/en/Schools-Colleges/Information/Curriculum-and-Syllabus/Foreign-Languages-Strategy/fls_languages_connect_strategy.pdf)
- EBCL. (2019). EBCL Website. Retrieved from <http://ebcl.eu.com/>

- Feng, L. (2002). Influence factors of Chinese character shape recognition of non-Chinese students [Fei hanzi beijing liuxuesheng hanzi xingyin shibie de yingxiang]. *Chinese Character Culture [Hanzi Wenhua]*, 3, 47–49.
- Fu, Y. H. (1988). The development of modern Chinese characters commonly used in appendix: Common words and common words language [Xiandai hanyu changyong zibiao de yanzhi]. *Application and Research*, 2, 22–31.
- Hanban. (2010). *The graded Chinese syllables, characters and words for the application of teaching Chinese to the speakers of other languages*. Beijing: Beijing Language and Culture University Press.
- Hanban Test Centre. (2019). *An introduction to the new HSK*. Retrieved from <http://www.chinesetest.cn/userfiles/file/dagang/HSK-koushi.pdf>
- Hoosain, R. (1992). Psychological reality of the word in Chinese. *Psychological Reality of the Word in Chinese*, 90, 111–130.
- Hulstijn, J. H. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal*, 91(4), 663–667. <https://doi.org/10.1111/j.1540-4781.2007.00627.5.x>
- Jiang, X. (2006). The effects of frequency and productivity on character learning by elementary L2 learners from alphabetic language backgrounds. *Acta Psychologica Sinica*, 38(4), 489–496.
- Kuo, L.-J., Kim, T.-J., Yang, X., Li, H., Liu, Y., Wang, H., Li, Y. (2015). Acquisition of Chinese characters: The effects of character properties and individual differences among second language learners. *Frontiers in Psychology*, 6, 986. <https://doi.org/10.3389/fpsyg.2015.00986>
- Lam, H. C. (2011). A critical analysis of the various ways of teaching Chinese characters. *Electronic Journal of Foreign Language Teaching*, 8, 57–70.
- Li, C. N., & Thompson, S. A. (1989). *Mandarin Chinese: A functional reference grammar*. Berkeley and Los Angeles, California: University of California Press.
- Li, F. (2011). The publication of the graded Chinese syllables, characters and words for the application of teaching Chinese to the speakers of other languages [‘Hanyu Guoji Jiaoyuyong Yinjie Hanzi Cihui Dengji Huafen’ chuban]. *Chinese Teaching in the World [Shijie Hanyu Jiaoxue]*, 1, 119.
- Li, R. (2014). Characteristics of the Chinese language and teaching Chinese as a second language [Hanyu de tedian yu duiwai hanyu jiaoxue]. *Language Teaching and Linguistic Studies*, 3, 1–10.
- Lincoln, P. Z. (2015). Character-based, word-based, or two-way approach. *NUCB Journal of Language, Culture and Communication*, 16(2), 51–58.
- Little, D. (2013). The common European framework of reference for languages and the European language portfolio: Some history, a view of language learner autonomy, and some implications for language learning in higher education. *Language Learning in Higher Education*, 2(1), 1–16. <https://doi.org/10.1515/cercles-2012-0001>
- Liu, L. (2008). The effects of character complexity and structure on the study of Chinese characters among international students in China [Bihua shu yu jieou fangshi dui liuxuesheng hanzi xuexi de yingxiang]. *Language Teaching and Linguistic Studies [Yuyan Jiaoxue Yu Yanjiu]*, 1, 89–96.
- Liu, Y., & Ma, J. (2010). The graded Chinese syllables, characters and words for the application of teaching Chinese to the speakers of other languages. *Chinese Teaching in the World [Shijie Hanyu Jiaoxue]*, 1, 82–92.
- Lü, D. (2015). *A comparative Study of Two ‘Character-based Approach’ Textbooks [Liangben Faguo zibenwei jiaocai de bijiao yanjiu]* (MA Dissertation). Shanghai International Studies University, Shanghai, China.

- Lü, H. (2010). A study of the Chinese characters used in the vocabulary lists for Level 1, 2 and 3 of the New HSK [Xin HSK yi, er, san ji cihui dagang yongzi qingkuang yanjiu]. *Heilongjiang Social Science [Heilongjiang Shehui Kexue]*, 122(5), 106–108.
- Luo, F., & Guo, X. (2012). The situation of the international promotion of Chinese [Hanyu guoji chuanbo zhuangkuang]. In Language Situation in China Team (Ed.), *Language situation in China: 2011 [Zhongguo yuyan shenghuo zhuangkuang Baogao 2011]* (pp. 91–97). Beijing: Commercial Press.
- Mollering, M. (2016). Australian language policy and the learning and teaching of Chinese. In R. Moloney & H. L. Xu (Eds.), *Exploring innovative pedagogy in the teaching and learning of Chinese as a foreign language* (pp. 19–38). London: Springer Singapore.
- Osborne, C., Zhang, Q., & Zhang, G. X. (2018). Which is more effective in introducing Chinese characters? An investigative study of four methods used to teach CFL beginners. *The Language Learning Journal*, 1–17.  
<https://doi.org/10.1080/09571736.2017.1393838>
- Shen, H. H., & Ke, C. (2007). Radical awareness and word acquisition among nonnative learners of Chinese. *The Modern Language Journal*, 91(1), 97–111.  
<https://doi.org/10.1111/j.1540-4781.2007.00511.x>
- Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Development*, 74(1), 27–47.  
<https://doi.org/10.1111/1467-8624.00519>
- Starr, D. (2009). Chinese language education in Europe: The Confucius Institutes. *European Journal of Education*, 44(1), 65–82.
- Sun, C. (2006). *Chinese: A Linguistic Introduction*. Cambridge: Cambridge University Press.
- Sun, J. (2013). The analysis of the non-lexical-items in the 'Graded Chinese Syllables, Characters and Words for the Application of Teaching Chinese to the Speakers of Other Languages' ['Hanyu Guoji Jiaoyu yong yinjie hanzi cihui dengji huafen' zhong feici chengfen shoulu pingxi]. *TCSOL Studies*, 4, 59–64.
- Tinsley, T., & Board, K. (2017). *Language trends 2016/2017: Language teaching in primary and secondary schools in English survey report* (p. 105). Retrieved from British Council website:  
[https://www.britishcouncil.org/sites/default/files/language\\_trends\\_survey\\_2017\\_0.pdf](https://www.britishcouncil.org/sites/default/files/language_trends_survey_2017_0.pdf)
- Tong, X., & McBride, C. (2014). Chinese children's statistical learning of orthographic regularities: Positional constraints and character structure. *Scientific Studies of Reading*, 18(4), 291–308. <https://doi.org/10.1080/10888438.2014.884098>
- Wang, J. (2006). *Zi-centred theory and the cognitive method in TCFL [Zi benwei yu renzhifa de duiwai hanyu jiaoxue]* (MA Dissertation). China East Normal University, Shanghai, China.
- Wang, L., Song, L., & Suen, L. (2012, September 6). *Applying CEFR to teaching and assessing Chinese as a foreign language: A proposal from the EBCL project*. Presented at The 12th International CercleS Conference, London School of Economics and Political Science, University of London, UK.
- Wang, R. (2000). Reflections on the French character-based Chinese textbook [You Faguo zibenwei hanyu jiaocai yinfa de sikao]. *Chinese Teaching in the World*, 53(3), 89–98.
- Wang, X., Wu, Y., Zhao, S., Ni, C., & Zhang, M. (2016). The effects of semantic radicals and phonetic radicals in Chinese phonogram recognition. *Acta Psychologica Sinica*, 48(2), 130–140.
- Wiley, T. G., & Garcia, O. (2016). Language policy and planning in language education: Legacies, consequences, and possibilities. *The Modern Language Journal*, 100(S1), 48–63. <https://doi.org/10.1111/modl.12303>

- Xu, L., & Yao, X. (2007). *Essential Chinese dictionary*. Beijing: Foreign Language Teaching and Research Press.
- Xu, Y., Chang, L.-Y., Zhang, J., & Perfetti, C. A. (2013). Reading, writing, and animation in character learning in Chinese as a foreign language. *Foreign Language Annals*, 46(3), 423–444. <https://doi.org/10.1111/flan.12040>
- Yang, X. (2011). Achievements and problems of Chinese as a foreign language textbooks in our country [Woguo duiwai hanyu jiaocai bianxie de chengjiu yu wenti]. *Journal of Hubei University*, 4, 31–34.
- You, H. (2003). The impact of stroke number, radical number and typological structure on the mastery of Chinese characters by learners of non-Chinese character culture circle [Bihua shu, bujian shu he tapu jieyou leixing dui feihanzi wenhuaquan xuexizhe hanzi zhangwo de yingxiang]. *Chinese Teaching in the World [Shijie Hanyu Jiaoxue]*, 64(2), 72–82.
- Zhang, G. X. (2011, July 1). *European benchmarking Chinese project*. Presented at the BTLCS Annual Conference, Edinburgh, UK.
- Zhang, J, Xie, N., Wang, S., Li, Y., & Zhang, T. (2010). A report on researching and producing the new HSK. *China Examinations*, 9, 38–43.
- Zhang, Jijia, Wang, J., & Yin, C. (2014). The role of phonetic radicals and semantic radicals in phonetics and semantics extraction of phonogram characters: An eye movement study on components perception. *Acta Psychologica Sinica*, 46(7), 885–990. <https://doi.org/10.3724/SP.J.1041.2014.00885>
- Zhang, Jinqiao. (2008). On the stroke effect, word frequency effect and morpheme frequency effect in the recognition of Chinese one-character words by foreign students [Liuxuesheng hanyu danzici shibie de bihuashu xiaoying, cipin xiaoying he cisu pinglü xiaoying]. *Journal of College of Chinese Language and Culture of Jinan University*, 1, 22–29.
- Zhang, P. (1992). The enlightenment of ‘A Key to Chinese Speech and Writing’ in France [Hanyu Yuyan Wenzi Qimeng yishu zai Faguo huode chengong de qishi]. *Language Teaching and Linguistic Studies*, 1, 119–125.
- Zhang, Q., & Reilly, R. G. (2015). Writing to read: The case of Chinese. *29th Pacific Asia Conference on Language, Information and Computation*, 345–354.
- Zhou, G. (2002). Two points regarding List of Frequently Used Characters in Modern Chinese [Guanyu ‘Xiandai Hanyu Changyong Zibiao’ de liangdian yijian]. *Chinese Language [Zhongguo Yuwen]*, 3, 266–269.
- Zhou, J., & Li, H. (2008). Revision opinions about A-grade word of the Syllabus of Graded Word and Characters for Chinese Proficiencies [Dui ‘Hanyu Shuiping Cihui yu Hanzi Dengji Dagang’ jiajici de xiuding yijian]. *Working Papers for the College of Chinese Language and Culture in Jinan University [Jinan Daxue Huawen Xueyuan Xuebao]*, 3, 35–42.
- Zhou, M. (2011). Globalization and language order: Teaching Chinese as a foreign language in the United States. In L. Tsung & K. Cruickshank (Eds.), *Teaching and learning Chinese in global contexts: CFL worldwide* (pp. 131–150). London and New York: Continuum International Publishing Group.
- Zhu, L. (2013). A review of the Graded Chinese Syllables, Characters and Words for the application of teaching Chinese to Chinese native speakers: The application of single-character syllables in the teaching of Putonghua [Hanyu muyu jiaoyu yinjie hanzi cihuibiao kaocha: Jiantan Putonghua jiaoxue dui dutizi yinjie de yingyong]. *Applied Linguistics [Yuyan Wenzi Yingyong]*, 2, 109–114.

- 
- <sup>i</sup> Confucius Institutes are initiated by the National Office of Chinese Language Council International to promote and support the teaching of Chinese language and culture worldwide (Starr, 2009).
- <sup>ii</sup> The Graded list refers solely to the 900 characters in this study. In contrast, the Graded project refers to the whole document, including lists of characters and vocabulary items at different levels.
- <sup>iii</sup> Active characters refer to those required to be reproduced by learners, whereas passive characters are those expected to be recognised.
- <sup>iv</sup> Both the EBCL ‘Can-do Statements’ and a talk by Paternico (2012) indicate that there are five sub-levels (A1.1, A1, A1+, A2, A2+) for the Basic User level, in comparison with two sub-levels (A1 and A2) for the equivalent user in CEFR. However, only A1 and A2 characters are available on the EBCL website.
- <sup>v</sup> Various translations can be given to ‘pujihua’, as it generally means disseminating something to reach as many people as possible.
- <sup>vi</sup> This is indirectly mentioned in the EBCL Character Methodology. One of the sources for the selection of characters was the EBCL vocabulary list. A study of the EBCL Vocabulary Methodology (Allanic & Shu, 2012) shows that the compilation of vocabulary items was also based on a comparison of the selected words with the most frequently used words in the Chinese language. Therefore, the EBCL list also takes into consideration the frequency of character occurrence.
- <sup>vii</sup> The corpora consisted of LDC2002E18, LDC2003E07, LDC2003E14, Hansards portion of LDC2004T07, LDC2004T08 and LDC2005T06. A range of sources, such as news from various newspapers, news agencies and online blogs, comprise this corpus, which is usually used for machine translation (Prof. Junhui Li, personal communication).
- <sup>viii</sup> Please note that a semantic or phonetic radical can be placed in other positions (left, right, top, bottom, inside, outside). Here, we are only concerned with the positioning in the majority of compounds.

## About the Authors:

### 1. Qi Zhang (corresponding author)

Dr Qi Zhang is an assistant professor in Dublin City University. Her recent research interests are primarily in the field of Chinese language education, including teaching and learning Chinese as a foreign language, language attitudes and pedagogy, and Chinese language study among ethnic minorities in China. Email: qi.zhang@dcu.ie

### 2. George Xinsheng Zhang

Prof. George Zhang is a professor of Chinese and Director of the Centre for Modern Languages in Richmond University, London, and a former Director of SOAS Language Centre, University of London. He has participated and managed a few projects on Chinese language learning and teaching, including EBCL which is the first EU funded non-European language CEFR benchmark project. Email: george.zhang@richmond.ac.uk

### 3. Miss Yuhui Gao

Miss Gao is a postgraduate student of Chinese language studies at the Education University of Hong Kong, Hong Kong.