THE GRAPHABET AND BUJIAN APPROACH AT ACQUIRING HANZI (CHINESE CHARACTER) WRITING SKILL

CHEN-YA HUANG

University of Hong Kong

Abstract
Learning how to write orthographically correct Hanzi (otherwise known as the Chinese character) is a major hurdle facing students studying Chinese. The difficulty arises from the visual complexity of Hanzi, the opaqueness, i.e. diminished correspondence of sound to orthography, and the traditional method of learning Hanzi, which is monopolized by rote repetitive copying, excessive demand on memory, and lacking of any means of creating an auditory memory of the structural organization of individual Hanzi. As a result, the novice student has to invest a great deal of time and effort trying to master Hanzi, and is often deterred from continuing study. Yet, despite the seriousness of the problem, very little research has been carried out on its solution. This paper proposes a new approach at improving the ability to write Hanzi, through understanding Hanzi as strings of subunits stacked in two-dimensional space, and composed from 21 high frequency recurring shapes herein called graphabets. The combined use of graphabets and bujian can provide a means of creating an auditory memory of the structural organization and significantly decrease the memory load through chunking, as well as facilitating the use of computer feedback for learning purpose.

Keywords: Chinese character, character writing, orthography, teaching method, second language learning.

1. WRITING HANZI, A MAJOR HURDLE IN LEARNING CHINESE

The Chinese language is considered one of the hardest languages to learn by people who are not native to the language. As a result, there is a very high drop-out rate for Chinese as foreign language (CFL) students. In Australia, even with a target of only having to learn 500 Hanzi, the drop- out rate for CFL students in secondary schools was 94% (Australian government, 2010).

Huang, C.-Y. (2017). The graphabet and bujian approach at acquiring Hanzi (Chinese character) writing skill. Contribution to a special issue on Teaching and Learning of Reading and Writing in the Multilingual and Multicultural Context, edited by Shek Kam Tse and Elizabeth Loh. L1-Educational Studies in Language and Literature, 17, 1-34.
h*ttps://doi.org/10.17239/L1ESLL- 10.17239/L1ESLL-2017.17.03.07*
Corresponding author: Chen-Ya Huang, Department of Medicine, Hong Kong University, Pokfulam Road, Hong Kong, email: cyhuang@hku.hk

Apart from not sharing vocabulary with with European languages, two major hurdles exist for CFL students: Learning how to master the tonal changes and learning how to read and write Hanzi (otherwise known as Chinese character). Learning how to write Hanzi in the sense of producing not just a shape but an appropriate Hanzi of the correct orthographic form, regardless of aesthetic or calligraphy value, is especially difficulty for CFL students. As a result, some teachers have even championed the practice of delaying introduction of Hanzi to CFL students.

It should be recognized from the onset however that this difficulty with learning how to write Hanzi is not exclusive to CFL students. Even Chinese as first language (L1) students encounter great difficulties. Since the problem is universal to all students of Hanzi, although this essay will focus on the problem of how to teach the CFL student to write the orthographic structure of Hanzi correctly, experience with Chinese as first language students will be referred to as well.

2. THE IMPORTANCE OF AND DIFFICULTY WITH WRITING HANZI

Handwriting involves an action perceptual coupling, which has been shown to have positive effects on reading in the alphabetical language (Graham & Hebert, 2010 Kiefer, Schuler, Mayer, Trumpp, Hille & Sachse, 2015; von Koss Torkildsen, Morken, Helland & Helland, 2016) Not surprisingly then, handwriting has also been reported to have an enhancing effect on the ability to read Chinese (Tan, Spinks, Eden, Perfetti & Siok, 2005; Cao, Vu, Chan, Lawrence, Harris, Guan, Xu & Perfetti, 2013; Zhang & Reilly, 2015). Thus, Tan et al performed experiments which showed that copy writing skill contributed to variance in reading by beginner and intermittent readers from a Beijing primary school, whereas whereas phonological awareness did not. In addition, after nonverbal IQ was controlled, pseudo character writing ability explained 18.3% of variance in reading by beginner reading whereas picture drawing only explained 3%. Cao et al found that lexical decision accuracy and decision time were enhanced more by character writing than by pinyin training. Zhang and Reilly showed that character recognition was enhanced more by having to draw the character rather writing its pin yin form.

When handwriting essays even native speakers will tend to use only words which they are capable of writing, failing to use the full repertoire of words they are in possession of when speaking, if they cannot recall how to write those words. This means that the amount Hanzi they are able to write will also limit their ability to express their thoughts and ideas fluently. Alternately, they will interrupt their thought stream in an attempt to capture the word. Either way, the failure to know how to write a Hanzi has detrimental effect on the composition of the essay, affecting coherence and communication of idea (C. J. Tseng, 2008). Thus, improvement in the acquisition of Hanzi hand writing skill will likely improve Hanzi recognition and understanding as suggested by the previous researchers and also enhance fluency in literary writing by hand.

Whilst only a coarse assessment of salient features is needed for visual recognition and reading of script, writing the correct word requires possession of fine detailed knowledge of its internal configuration. It is easier to recognize and read a Hanzi text than to write it spontaneously or to dictation. This disparity between the ability to read aloud and understand, and the ability to write affects all users of Hanzi, whether L1 or CFL students. It was implicitly recognized by the Chinese Ministry of Education which directed in 2002 that the target ability to write is set at only a third to a half of the reading vocabulary for primary one and two L1 students in China, and rising to about 83% of the vocabulary at more senior level (Ministry of Education PRC, 2002).

Since even first language users have problems learning how to write Hanzi, it is not surprisingly then that the acquisition of writing Hanzi becomes a major hurdle in learning Hanzi. In France, of the 805 Hanzi in Langue Vivante I, students are only expected to write 505. Jen and Xu found in a study of 55 first year CFL students in the United States, with proficiency in 550 Hanzi, that the students could only write out correctly 39% of what they could speak or read. They reported that 91% of the students who gave up after one year of study blamed difficulty with writing Hanzi as the cause (Jen & Xu, 2000).

The difficulty with writing Hanzi arises from its visual complexity; its opaqueness i.e. diminished correspondence of pronunciation sound to orthography; and the traditional method of learning Hanzi, which is monopolised by rote repetitive copying, excessive demand on memory, and reliance on visual and grapho-motor memory without access to an auditory memory of the internal configuration of individual Hanzi.

2.1 Visual configuration as hurdle

There are two versions of written Hanzi in common use. Mainland China uses mainly a so called simplified version of Hanzi, and Taiwan uses the so called traditional or regular version of Hanzi. Both versions are used in Hong Kong, Macao, and among Chinese diaspora worldwide. CFL students these days tend to learn the simplified version of Hanzi. But it should be emphasized that about half of Hanzi in common usage by CFL students are in fact common to both versions. For example, amongst the 800 Hanzi which constitute the class A Hanzi of Hànyǔ Shuǐpíng Kǎoshì (HSK) or Chinese Proficiency test, only 356 (44.5%) are structurally and therefore visually different when the traditional and simplified forms are compared. Amongst the 804 Hanzi in HSK class B only 407 (50.6 %) were different structurally. Of these, the majority consist of a change in the structure without change in semantic or phonological transparency (the presence or absence of a reliable semantic or phonological identifier), nor changes in the likelihood of structural ambiguity. Since the structural change mostly in the form of delating components, or substituting for simpler components, without violating traditional rules for orthographic construction, there is no

dramatic change in the way a Hanzi is written, only simply substituting different components: e.g. writing 车 instead of 車, 专 instead of 專, and writing 转 for 轉; or writing in sequence 宀玉 resulting in 宝 instead of the sequence of 宀王缶貝 and 寶. Hanzi has a long tradition of allographs, different visual and structural configuration of an identical meaning and pronunciation. Simplified and traditional Hanzi versions can be considered simply as the recognition of different allographs as the official or regular version of the same Hanzi. This essay will therefore deal with Hanzi as a whole, pointing out only exceptions where relevant.

For CFL students who are used to the concept of words being a string of identifiable alphabetical letters, the visual appearance of a two-dimensional script requires a major shift in perception and a need to identify the salient features of this new script. A CFL student dealing with a two or three alphabet length word such as “TO”, “MY”, “OX”, “BUS”, “CAN”, need only discern two or three salient components, a changed alphabet or change in the letter order. But the same size Hanzi can be much more complex.

A simple comparison will show that using the same font size, a Hanzi is only about 2 alphabet length long, with a height of 1.5 alphabets. But within this small area, a large number of geometrical shapes may reside, and a change of 0.5 alphabet size may signify marked change of meaning and sound. The beginner CFL student has therefore to develop much finer visual discriminatory skill than required in their own native language. It is not surprising then that 天 may be mistaken for 夫, 白 for 臼, or 申 for 甲, and the left upper corner of 登 confused for the left upper corner of 祭.

Unlike alphabetical words which can vary in length, all Hanzi are fixed in size and square. The internal configuration is however more complex and various typology classifications have been proposed[1]. Basically however, a Hanzi may be one of four types, singletons, tiered, horizontally partitioned, or encased. True singletons are those which are not encased and cannot be further separated into subcomponents by horizontal or vertical lines. A tiered Hanzi may appear along the vertical axis as those which can be separated into two tiers of Hanzi by a horizontal line e.g. 需; those which can be separated into three tiers by two horizontal lines e.g. 崽, 篦, and those with even more tiers e.g. 鱻. A Hanzi on any one tier, may be a singleton, e.g. 我, a bipartite where a vertical line can separate the Hanzi into two component Hanzi e.g. 明, a tripartite which can be separated by two vertical lines into three parts e.g. 衍. or those which are either encased within a square box e.g. 因, or partially encased within an open receptacle like component e.g. 同, 安, 函, 反, 症, 匠, 司, 句, 可, 迫, 起.

[1] *See e.g. Su, P. C., 苏培成 (2001).* 二十世纪的现代汉字研究 *[Research on modern Hanzi of the 20th century]. 书海出版社; Ki, W. W., Lam, H. C., Chung, A. L. S., Tse, S. K. et al. (2003). Structural awareness, variation in theory and ICT support. L1 Educational Studies in Language and Literature, 3, 53-78; Wang N., 王寧 (2013).* 漢字構形學講座. *[Lectures on the structural studies on Hanzi]. Sanmin publishers Taipei.*

The encased component again can be either single tiered, or several tiered. A few singletons have a single central stem with symmetrical underarm structures, e.g. 夾, 奭, 乖. The majority of bipartite, tripartite, and two-tiered Hanzi are phono-sematic compounds. In addition to providing clues to obtaining semantic and phonological information, the typology is also the prime determinant of stroke order used in writing Hanzi.

Thus, compared to an alphabetical word of equal physical size, a Hanzi is much more structural dense, visually crowded and complex in configuration. Studies have shown not only that structural density as reflected in stroke numbers has an adverse effect on accuracy of writing, but bipartite structured Hanzi are more prone to miswriting (You, 2003; L. P. Liu, 2008). It is obvious that without effective explicit instruction, the novice will find great difficulties, requiring much time and effort to overcome.

2.2 Phonological opacity as hurdle

In the widely accepted dual route theory of reading and writing, when attempting to write a word, a writer can use either a direct lexical route, and produce the word from a mental lexicon of known words, or to use a sublexical route assembling a word configuration by means of phonology to grapheme conversion. In the latter, the writer has to convert the sound of the target word (phonology) into a visual symbol (grapheme). Most alphabetical script enjoys a correlation between the alphabetical combination in a word, and its pronunciation. This transparency means that it is easier to known its pronunciation, and easier to reassemble from memory when writing.

Hanzi on the other hand is a fairly opaque script, with much less correspondence between the configuration and pronunciation. Although modern Hanzi are mostly phono-semantic compounds which has phonological component (bujian) providing phonological information and a semantic bujian acting as a semantic signifier, the constituent phonological bujian is not always a reliable guide to the pronunciation. Even if tonal differences are ignored, only 38% of phonological bujian extracted from a corpus of 27380 Hanzi were found to represent a single sound value. Another 29% represent two possible pronunciation sounds (Huang unpublished). Thus, more than 40% of phono-semantic compounds are not phonologically transparent, their phonological signifier having less certain sound representation. Furthermore, a sizable minority of Hanzi lack phonological signifier, or are associative compounds such as 室, 涉, 就, 位, 規 in which none of the components contribute phonologically to the sound of the whole Hanzi. In addition, there are a large number of homophones, i.e. Hanzi which have the same sound but different visual configuration and meaning. The writer therefore will encounter considerable difficulty assembling the Hanzi configuration according to sound using the sublexical route and is reliant on the mental lexicon of known Hanzi which is limited at the beginner stage. The mental lexicon is

also liable to shrink with disuse. Thus, the person trying to write in Hanzi will have greater difficulty than the person writing with an alphabetical script. In the aforementioned study by Jen and Xu, their CFL students could not produce any response in 17% of the target Hanzi, and made mistakes in 44%, 74% of which were homophones, and 26% constructive in nature (Jen & Xu, 2000).

In contrast to the opaque relationship between pronunciation and configuration, the presence of a semantic bujian in a Hanzi means however that there is information about the meaning of the word. Hanzi has therefore better semantic transparency than alphabetical script. This is useful for reading and accessing the meaning of a Hanzi, and in the process of writing, can help to narrow down the choice after the initial successful phonology to grapheme transformation to a set of homophones.

3. LIMITATION OF CURRENT METHODS FOR ACQUIRING HANZI WRITING ABILITY

Despite the problem of learning how to write facing CFL students, publications on Hanzi learning strategies have hitherto been overwhelmingly concentrated on how to recognise, rather than how to write (Jiang & Cohen, 2012). The monograph length paper on teaching of Chinese as a foreign language in the United States by Ke and Li (2011) reviewed extensively research on learning to speak, grammar, and reading and many other aspects of Chinese studies yet made no mention of any study on the teaching of how to write. There has apparently not been any serious attempt to develop a teaching method for learning how to be able to write Hanzi without misconstruction of the orthographic configuration. Students are simply expected to do rote copying. Analysis of writing errors committed by beginner students has led usually to a mere call for improvement in the ability to differentiate various strokes and bujian (see for example Shi, 2000; Yang, 2007). No writing skill specific teaching method has resulted.

The assumption seems to be either that having the ability to recognize a Hanzi, output production of the correct configuration will naturally follow, or the problem can be circumvented by having the CFL student type with a keyboard instead of learning how to write Hanzi.

Methods which teach recognition of Hanzi may however vary in the degree they affect writing ability. Methods such as learning Hanzi through by rationale 字理識字, or by wild association 奇特聯想 have not made claims of improving the ability to master writing skill.

Some methods for teaching recognition have claimed both improvement in understanding and writing ability in Chinese children. The intensive learning of character 集中識字 method teaches analysing each Hanzi for its semantic or phonetic radicals, and has been reported to facilitate the understanding of Hanzi as well as result in a 90% rate of accuracy in writing (Liu, Ma & Tong, 1979). Si claimed that with Dispersive learning of character method 分散識字, in which characters are taught when encountered, 88.5% of 2049 Hanzi taught in a Chinese primary school was written

correctly to dictation (Si, 1978, p.12). Using a family of character 字族文 approach where characters which share common features are taught together, is said to have achieved 94% accuracy in writing (Sichuan, 1994). However, these results do not seem to have been replicated and these teaching methods also do not seem to have been used much amongst CFL students.

3.1 Bujian and writing acquisition

Bujian is defined as any component which can be used to construct a Hanzi orthography. Defined in such broad sense, they are made up of several categories: Hanzi; abbreviated versions of Hanzi with semantic properties e.g. 氵, 阝, 艹; and components of Hanzi which are not recognized Hanzi e.g. the simple dot and line, as well as more complex forms like 羊, 夫, and 丬. Since Hanzi is often formed from combination of two or more Hanzi, the number of bujian which are also Hanzi in their own rights would already be quite large. Indeed, our own analysis suggest that there are close to 2000 phonological bujian alone in a corpus of 30,000 (Huang unpublished). Even fundamental bujian i.e. those which cannot be subdivided further into any other bujian, are found to number in the hundreds, 439 being found in a corpus of 6097 traditional Hanzi (Chen, Chiang, Chiou & Sung, 2011), and 514 amongst the most commonly used 3500 simplified versions of Hanzi (Ministry of Education PRC, 2009).

Obviously, teaching bujian indiscriminately would not be cost effective as it would simply be like teaching nearly all Hanzi encountered. How then could bujian be used for the purpose of enhancing the learning of writing skill?

The use of bujian 部件識字 in learning Hanzi by first language learners have not been reported to affect writing ability. In teaching bujian as a means of learning Hanzi to CFL students, ability to read and understand rather than writing improvement seem to be the goal (Cai, 1997; Chu, 2009; Wang & Koda, 2013). A search through the literature found however that of the only two previous published papers describing attempts at testing a method for learning how to remember Hanzi for writing purpose amongst CSL students, both involved bujian. In one study, Liu & Jian (2003) found that a recall writing method was superior to repetitive copying. The former method involves cognitive analysis of the components of an unfamiliar Hanzi, before trying to write. Thus, it shows that explicit awareness of the composition of Hanzi can lead to more accurate and improved writing ability. There was no attempt at creating however a concept about the serial order that the components occupy in the configuration. In the other study, Xu and Padilla tested the use of chunking through teaching students how to recognize bujian and radicals. Although they tested writing ability, the results were presented as a composite of writing, and association with pinyin. The effect on writing ability is therefore uncertainty even though there was beneficial effect on the composite score in the short-term period (Xu & Padilla, 2013). In any case, they found a stronger effect on understanding the

meaning rather than the composite of writing and pinyin. Neither of these two studies employed auditory enunciation of the bujian for memory consolidation purpose. Nor was semantic or phonological transparency explicitly targeted.

3.2 Rote copying for writing acquisition

Rote copying seems to be the most commonly employed method of learning how to write Hanzi for either first or second language learners. Its cost effectiveness has never been studied however. Rote copying is basically initiating a grapho-motor act to depict what is seen visually. Repeated writing the same Hanzi creates a grapho-motor program memory in the brain, which can be recalled when needed for writing output. Rote copying has however several problems: it is time consuming and boring. Furthermore, for rote copying to be successful in building up a reliable memory of the Hanzi configuration, the copying has to correct, without mistakes being made each time. If an error creeps in at any stage, it may become ingrained, resulting in an incorrect configuration being constructed repeatedly, and resulting in the consolidation of an incorrect memory. Thus, success in rote copying depends on continued attention by the writer, and strict adherence to the stroke order.

In an alphabetical language, the letter order is the sequence of the letters in a word. The letter order is essential to the pronunciation of the word as well as its visual appearance. Each word requires a fixed letter order. It is doubtful that a word's meaning or sound will be remembered if the letter order is remembered only haphazardly in a random manner. In Hanzi, the sequence of its subunits has traditionally called stroke order.

If a Hanzi is treated as a pictogram, with the student starting and finishing each time at at different point, without having a concept of stroke order in the writing of the component units, then it will be difficulty if not impossible to form or consolidate any permanent re-accessible perceptual-motor memory trace. Thus, following the stroke order is an essential condition if rote copying is to be successful.

Although the stroke order for Hanzi has been standardized (State language commission, 2000; Ministry of Education ROC, 2008), with only some discrepancies between mainland China, Taiwan and Hong Kong (J. Y. Huang, 2005), and is usually taught to students learning Hanzi in China, its importance is not always understood or respected amongst CFL students. Numerous reports of writing errors containing reversed positions of radicals in phonological-semantic compounds in CFL students studying in China points to a lack of attention to stroke order (see for example: Xiao, 2002). Whilst rote writing appeared to be the most commonly strategy used by CFL students in America or Germany, and knowledge of stroke order exist to some extent, explicit instruction on the use of strokes and radicals do not seen to be prevalent amongst university students. (J. J. Tseng, 2000; Yin, 2003; Shen, 2005). The question whether teaching of correct stroke order is necessary or useful also remains a research question for English schools (Carruthers, 2012). Hanzi also seem to be treated mainly as pictures and little attempt is really made to teach stroke order in

Australian primary and secondary schools (Orton, Tee, Gong, McCulloch, Zhao & McRae, 2012). The situation may however be changing, as stroke order is now being taught in newer German primary school text (Liu & Su, 2014).

3.3 Chunk load and auditory memory in rote copying

Rote copying means repeated hand movements to duplicate what is visually seen. Stroke order allows the hand movements to follow a repeated reproducible pattern. Each discrete movement constitutes a motor chunk in the working memory. If a Hanzi contains a large number of strokes the chunk number involved will also be large, unless a sequence of individual chunks is consolidated into a larger size chunk such as a specific identifiable spatial configuration. Thus, rote copying of high stroke number or visually complex configuration Hanzi requires the ability to identify discrete component chunks. If this cannot be done, the motor movements chunk number may overwhelm working memory capacity, with difficulty in establishing a useful reliable memory resulting.

Rote copying as a means of learning how to create a memory of the organizational configuration of a Hanzi is often accompanied by pronunciation of the name of the Hanzi (Shen, 2005). This can create an association between the auditory sound and the grapho-motor program being formed in the brain. However, during traditional rote copying there is no attempt to create an auditory map of the configuration, similar to for example saying E, L, E, P, H, A, N, T for elephant. Thus, in an alphabetical language, there is not only a grapho-motor and visual memory of a word, there is also the possibility of an auditory memory of the letter order, which is also the actual configuration of the word. However, establishing an auditory memory of the internal configuration is rarely considered important and students are seldom instructed to orally describe the configuration of a Hanzi. At most mnemonics are used for a few Hanzi, (See e.g. Chiu, 2003; J. H. Liu, 2008; Shao, 2008; Long, 2008; Hsu, 2011; S.Y. Wu, 2015) and used deliberately to a limited extent in the component teaching method (Su, 1991)

Traditionally therefore, unlike alphabetical language users who can rely on a triad of visual, auditory, and grapho-motor memory to remember a word, Hanzi learners can only rely on visual and grapho-motor memory. This is despite the fact that having more sensory information is conducive towards better memory.

A better recall is found when bimodal presentation using both audiovisual stimuli is used instead of unimodal stimulus employing either auditory or visual stimulation (Mastroberardino, Santangelo, Botta, Marucci & Olivetti Belardinelli, 2008). Wu demonstrated that when students saw Hanzi at the same time as hearing the pronunciation, they were better at recognising low density (6-7 strokes) Hanzi in the short term, and the high density (13-14 strokes) Hanzi in the long term. The bimodal presentation also helped with better ability to write correctly low density but not high density Hanzi in the long term (Y. Wu, 2014).

What Wu demonstrated is that the auditory input can indeed help to promote retention of Hanzi for reading and writing. The failure for the bimodal presentation to help writing of higher density Hanzi however shows that the configuration of denser Hanzi cannot be accurately remembered by using the auditory information of the whole Hanzi.

As it stands, rote copying as a means of Hanzi acquisition is not only time and effort intensive, but is also prevented from achieving its desired result by failure to adhere to stroke order, chunk overload, and lack of simultaneous auditory information.

4. IS LEARNING HANDWRITING HANZI REDUNDANT?

In line with the trend in electronic and mobile means of messaging and faced with this difficulty in writing Hanzi, it is not surprising that in some institutions, CFL students are expected only to be able to type Chinese language essays rather than writing by hand. The typing is almost always by means of pinyin, and therefore is essentially just accessing a list of candidate Hanzi homophones for visual selection, thus bypassing the need for having detailed knowledge of the structure of the target Hanzi. Whilst this method allows the CFL students to express their thoughts using their oral vocabulary, the effect on actual linguistic acquisition remains uncertain. Since homophones are common, it can be expected that using this method will increase the chance of producing homophonic mistakes if the candidate list includes several visually approximate choices. However, study on CFL students' performance in producing the correct Hanzi using typing as a means of writing Hanzi has been notable for its absence. The only published paper showed an appalling 33% rate of error (Xie, 2011). This is similar to the 32.6% homophonic errors when hand writing reported by Jen and Xu (2000). Whilst newer computer technology may help to improve the resultant accuracy rate, the fundamental problem of differentiating between homophones remains hidden but unresolved. There has been no study comparing the merit of using typing in place of, or in parallel with versus more traditional method of handwriting Hanzi in terms of learning. The effect on reading skill and accuracy has also not been reported. The real world need for handwriting when an electronic appliance is not available for help or output is also not addressed. Whilst CFL students who study for interest may afford to be unable to handwrite, CFL students in Hong Kong and elsewhere who need day to day Chinese literacy and handwriting ability for occupational purpose will not find typing by pinyin an adequate substitute. Until, such studies validate the educational value of typing in place of handwriting, the need for learning to handwrite and to recognize the features of individual Hanzi remains.

5. A NEW PARADIGM FOR HANZI WRITING ACQUISITION

Despite the fact that the essay liùshū 六書, first mention around second century BCE, and whose contents were known by the first century CE, had clearly stated that pictographs are only one of six types of Hanzi, popular myths still propagate the view that Hanzi are pictographs representing stylized depiction of actual scenes. This view is shared by many second language students on their first contact with Hanzi.

The misunderstanding comes basically from assuming that scripts are universally structured as horizontally aligned strings of subunits. The visually complex Hanzi as well as the Koran alphabet based script Hangul can however be understood better as a two-dimensional script with such strings stacked in both the horizontal and vertical axis. The alphabetical word "down" is written by a temporal sequence of the letter string d, o, w, n, giving a temporal sequential appearance of d, do, dow, down. The temporal sequence of writing the string 一丨丨一乚 results in a temporal sequential appearance of 一十廾廿世. Thus, the Hanzi 世 is the two-dimensional equivalent of the the horizontal string 一丨丨一乚. 采 is another case in point, writing / first, then moving vertically down to write horizontal a sequence of 丶丶/ creating 丷 beneath the /, before moving downwards again to write 一十, then placing the oblique lines horizontally to the left and the right of the 十 creating 才 then 木, in order to complete the configuration of 采. Thus, instead of writing in one dimensional space horizontally, Hanzi writing utilizes two dimensions. This arrangement results in a visual image as well as an economic use of space, shortening the horizontal length at the cost of increasing density and visual crowding.

5.1 What are the fundamental units underlying Hanzi?

If Hanzi is understood as a string or a number of strings stacked along two dimensions, then what are the subunits which form this string?

A visual inspection of a bipartite Hanzi, may show it being made up from two components or bujian. For example, 粉 is a string of two bujian 米 and 分, the former providing semantic information and the latter phonological information. But then, these two bujian are clearly themselves strings of subunits and are therefore not the final subunits in question.

Hanzi is usually thought to be built up from single strokes such as a dot, or a unidirectional line (vertical, horizontal, oblique, angulated or wavy). This concept arises from the way Hanzi is written using a brush and the influence of brush writing on Hanzi orthography (Wang, 2014); as well as the use of stroke number to index Hanzi in dictionaries, has reinforced the idea that Hanzi is thus formed. But the concept of stroke for the purpose of learning Hanzi, in the sense of describing its configuration rather than just the motor act of penmanship is a relatively modern idea. Unlike with alphabetical language where the number of alphabet in a language is long settled, since as many as 35 types of strokes exist, there is still argument about how many

types of strokes should be used for pedagogic purpose for teaching Hanzi writing acquisition (see for example: Wang & Su, 2015; J. C. Huang, 2016). Currently, angulated lines are rather neglected. All angulated lines are lumped under the term zhé, despite that fact that amongst the 593 core Hanzi shared in common by class A Hanzi of Hànyǔ Shuǐpíng Kǎosh (HSK), the Test of Proficiency-Huayu (TOP) and the 805 Hanzi in Langue Vivante I curriculum (Chang, 2008), only 53 (8.9%) of the traditional form and 79 (13.3%) of the simplified form are formed from just single stroke graphabets, and only 13 (2%) did not involve any angulated line.

There are several problems with the concept of stroke as the underlying subunit in Hanzi orthography.

The concept of stroke being the underlying subunit in Hanzi rests on the assumption that the act of writing is equivalent to the conceptual product being produced. This is obviously a different assumption from alphabetical language, where the alphabetical letter is the conceptual product and underlying subunit of the language. Although an upper-case alphabet may need to be written with several strokes, and even lower case cursive letters require a dot, horizontal and oblique crossing lines to write, these strokes obviously are not considered the subunits of alphabetical language.

When strokes are treated as the fundamental subunits of Hanzi, Hanzi is prone to be continually viewed as a picture to be drawn with strokes, much like painting with brushstrokes. Frequently this leads to confusing the objectives of calligraphy and aesthetics with meaning that configurations convey, and overemphasis on producing shapes which are purely artistic embellishment without ideational representation. This can be problematic at times, further substantiating the view that Hanzi is difficult to learn. For example, students are often penalized for failing to produce a hook at the lower end of a Hanzi, when with the exception of 干 and 于 and their derivatives, the hook conveys no semantic message and is only a late calligraphic addition to Hanzi font appearance. Likewise, the lowest horizontal bar in the left bujian is written tilted upwards toward the right in bipartite Hanzi. As a result, semantic bujian such as 土, 血, 王, 止 no longer resemble their original selves. So, should students be penalized if they ignore such artistic embellishment if they wrote a bar instead of a tilt? And are they to remember these bujian as comprising of a tilt or a bar? And does the semantic bujian 孑 in 孩 (child) represent 子(son) or 孑 (alone, or right arm amputee)? Clearly, artistic license and stroke shape as seen in fonts can clash with semantic identity

Strokes composed of just single dots or lines alone may also be difficult to use for oral communication of the two-dimensional structure of Hanzi. 坐 would be a sequence of / \ / \ 一丨一; 交 would be • 一 / \ / \. The initial sequence of 一/ 丶 can lead to either of 灰, 戍, or 达, when \ is treated as a dot (State Language Commission, 2000). Hearing the sequence 一一丨一 a listener would not know whether to expect the final Hanzi to be 王 or 式. A string of the strokes 丨冂一一 may constitute different configurations seen in 日, 門, 段, 闫. 目 is different from 旦, and 且, not only because

etymologically it was an entirely independent conceptual symbol but also because in writing terms it is (⺕ followed by a 一 whose ends terminate on touching the vertical lines on both sides) whereas 旦 is (⺕ followed by a 一 whose ends terminate on touching the vertical lines on both sides, then by another 一 which does not connect to the above) and 且 is (⺕ followed by a 一 which touches both vertical lines but extends beyond these points of contact). This extra information in italics are what determines the visuospatial position and differentiates the three Hanzi; information missing when only stroke numbers or sequence (丨𠃌一一一) is considered.

Furthermore, the use of a few, usually five to eight units, to describe the anatomy of Hanzi leads inevitably to memory overload. For 3479 high frequency simplified Chinese Hanzi, the average stroke number was 7.24; the corresponding traditional Hanzi stroke number was 9.4. (Guo, 2009) This means that the number of strokes in many Hanzi will exceed the human working memory capacity of 4±1 "chunks" of compressed information, or 7±2 uncompressed item (Mathy & Feldman, 2012). This is also a major reason why there has been little attempt to make use of the auditory memory of Hanzi configuration.

5.2 Graphabets as the fundamental units

An alternate view developed over the past decade consider however that Hanzi can be understood better as being formed from not just single lines and dot, but also from more complex shapes like angles, crosses, and boxes (Table 1).

Analysis of uniquely shaped geometrical units within Hanzi has isolated 21 bujian which are highest in frequency of occurrence (Table 2). Only 木 and 土 occur with higher frequency than the least frequent appearing of these bujian, 𠂇, 凵, 目 and 丰. Both 木 and 土 can however be formed be constructed using the other 21 units and are therefore excluded from the set.

In addition to being high frequency definable geometrical shapes, multi-stroke shapes included in this set are also distinct conceptual units which will lose their meaning if further broken down. Furthermore, no further geometrical shapes should be required to be used in the shaping of modern Hanzi. The resultant 21 shapes have been named graphabets (字母部件) to distinguish them from other bujian.

The graphabet set has been tested against both traditional and simplified modern Hanzi and found to be sufficient for describing all Hanzi as strings of graphabets (C. Y. Huang, 2006; Huang & Chan, 2012a). The only exceptions are unusual shaped Hanzi which are still extant in modern Chinese script such as 𠂆 and some other allographs of rare usage.

Table 1: Graphabets, name, variant, definition, and examples

	Variant	Definition	Name	Hanzi containing graphabet
•		dot	diǎn	汁 凡 卜
一		Horizontal line	héng	二 上 工
丨	亅	Vertical line	shù	下 巾 小
丿		Slash down or up	Piě (tí)	彡 才 扣
丶		backslash	nà	又 半 久
十		Horizontal Cross	shí	汁 土 斗
𠂇	乂	Oblique line crossing bar or another oblique	chā	右 大 爻 老
𠂆	3 乙 𠃌	Convoluted shape with 2 changes in direction	wān	夕 久 之 讠 阝
乚	𠃊	L shape	jiǎo	刂 爿 以 儿
𠂉	八	Inverted V	jiān	全 木 谷
𠂎	𠂉 く	Left pointing angle	niǔ	午 幺 女 水 衣
𠂏	>	Right pointing angle	zhuǎn	了 又 子 水
冂	冖 几	Receptacle open at bottom	mào	巾 凡 宀 内 贝
凵		Receptacle open at top	dōu	山 出 廿 心
厂		Receptacle open to right	chǎng	后 反 長
𠃌		Receptacle open to left	zhé	羽 已 力 马 永
丰	≠	Double bar line cross	chuàn	未 牛 青 夫
口		Empty Box	kǒu	中 只 吊 豆
日		Box with bar	rì	早 東 車 白
目		Box with two bars	mù	貝 自 見 鼎
口		Box with content, written in the sequence of 冂, content, and then closure line.	wéi	田 四 西 因

TABLE 2: Frequency of graphabets in 593 core Hanzi

	Traditional	simplified		traditional	simplified
●	295	268	𠂎	49	56
一	424	379	冂	110	110
丨	256	230	凵	32	21
丿	332	309	厂	63	41
¥	132	143	𠃌	140	122
十	208	194	丰	55	40
𠂇	74	75	口	148	110
𠂆	66	102	日	75	55
乚	119	111	目	34	14
𠂉	146	128	口	50	31
𠂎	113	105			

Of the graphabets, 9 are single dot or line, 4 of which being angulated. The other 12 would be considered to be multi-stroke in nature, several of which also form angles. Previous descriptions of Hanzi using strokes tend to lump all angulated lines as one. In contrast, the three categories of angulated lines, sharp angles, ninety degree angles, and convoluted, are identified separately and represented here. These angulated graphabets form distinct geometrical shapes which help to differentiate Hanzi configurations as well as conveying often distinct semantic signals, e.g. 纟, 巛, 灬, 又, 予, 子, 马, 爿. ^ defines a relationship of the two oblique lines, differentiating its shape from the other possible spatial relationship between the two oblique lines seen in 乂, 刃 or 匆, 豕. 厂, 冂, and 匚on the other hand are important in recognising a Hanzi as being partly encased.

Since the great majority of Hanzi require multi-stroke graphabets in the configuration, identifying them is essential. All the multi-stroke graphabets existed as semantically distinct units in the oracle bone period, and are therefore distinct concepts which should not be further broken down to smaller units. The only except is 丰, which existed as a frequent occurring structural component in modern font, but is derived from a variety of semantic source. Recognition of the multi-stroke receptacle like shapes 冂, 凵 permit recognition of vertical embedding Hanzi, ensuing again certain stroke rules for writing as well as heightened awareness of difference from visual mimics like 齐, and 门. Common Hanzi like 互, 穿, 英, 又, 友, 支 and 女 form eye catching pseudo 口 like structures, so that explicit instruction on recognizing the square shaped 口 is likely to promote early differentiation of these separate configurations. Likewise, the teaching of the graphabet 𠂉 will promote recognition of the triangular 又 more clearly, separating it from 女; while explicit instruction on 目 will bestow early appreciation that approximate Hanzi like 耳, 身, and 且 are different.

Although Graphabets can be considered therefore as a subset of bujian, they are best considered independently because they enjoy greater importance than other bujian, and can be used both for keyboard input as well as indexing and computation purposes (Huang & Chan, 2012b). In the following discussion, therefore, the term bujian will refer only to the other bujian which are not graphabets.

The left to right descending oblique line 丶 is not listed as a stroke or bujian in mainland China, but is in fact the sixth most used stroke in China. (Tseng, Chang & Chen, 1965) and is considered both as a stroke and a bujian in Taiwan, Hong Kong and Macao. (Government of Hong Kong, 2002; Ministry of Education PRC, 2009; Institute of information sciences, Academia Sinica, n.d.). It forms part of Hanzi components with a V shape, such as 羊. We consider this separate from a dot because when the etymology of these Hanzi are considered, the lines that descend to the right and left are symmetrical, For example, 羊, 米, 曾, 尚, 立 all show right and left lines of equal length. The left oblique line is in our opinion different from a dot as depicted in some modern fonts.

A number of convoluted lines are all grouped together in the graphabet wān. This is because individually each convoluted shape does not appear frequently, but as a group they are numerous in appearance. 𠂇 which is multi-stroke is included in this group because of its derivation from a Ↄ or › like shape. Convoluted lines tend to form combinations with other graphabets without much overlapping. Thus, wān shù (convoluted shape followed by a vertical line) lead to 𠂆, wān héng (convoluted shape followed by a bar) are seen in 马鸟龍与, whilst, wān mào (convoluted shape followed by an upper receptacle) usually lead to Hanzi like 角, 奐, 奂.

Some minor variations may exist in the shape of a graphabet. Included within the graphabet 𠆢 for example are the ^ in 木, 共, 具 and the Hanzi 人, 入, 八. Variation in the shape of the ^ in traditional and simplified versions of the same Hanzi 貝, 贝, or different font representations of 全 and 俞 shows clearly that the visual variants within a type may sometimes interchange. 人 and 入 are also now consolidated into the same index signifier (bushou) category in mainland Chinese dictionaries. The inverted V in Hanzi such as 兵, 共, 兴, 興, 真, has always been included in the bushou 八 section for visual reasons. All box like structure with a single horizontal bar in the middle are considered the same graphabet, thus including 日, 曰, and the 曰 in Hanzi such as 白, 香, 東, 易, 龟 or 甲 which are semantically unrelated to the sun 日 or speech 曰. Although the two Hanzi 日 and 曰 are different in semantic meaning, 日 is in fact frequently depicted as a 曰. For example, the 曰 in 是, 早, 景, 昜 are all in fact variants of 日.

For a linguistic representation, such as Hanzi, to be recognised as such requires distinction from a jumble of nonspecific lines and shapes. To do so, the viewer has to match stored information in the brain about salient features in Hanzi, with what is seen. The graphabets provide distinct geometrical patterns which are easy to recognise and can provide the first anchor to visual recognising a Hanzi. A dot is easy to miss in a visually crowded Hanzi, and the straight lines tend to also lose much of their distinctive features in a visually crowded situation. Angles and squares however tend to stand out even in these circumstances.

It is easy to see how the graphabets increases the transparency in Hanzi configuration. Stroke arrays of 一乚丨 could mean either 午 or 廿, but the two Hanzi are separated easily when using graphabet because 廿 would be identified as 一凵 and午 as 一乚丨. Under conventional terms, a string of the strokes 丨冂一一 may constitute different configurations seen in 日, 門, 段, 闫, 伊. But using the graphabet concept would define 日 a specific spatial configuration and differentiate it from the other spatial configurations. By means of defining such spatial configuration, graphabet more accurately describes visuospatial configurations than the bare information available from using strokes alone. 5 of the graphabets are of the encasing type and therefore enhance awareness of a Hanzi possessing encasing configuration which dictates specific rules of sequence order.

Other graphabets also provide a sharper focus on the geometrical shapes and spatial configuration which underlines Hanzi. For example, 十 defines the spatial configuration of a straight vertical line crossing a horizontal line. A Hanzi containing 十 will therefore look differently from a Hanzi with a horizontal and vertical line which meets but does not intercept, such as that in 工, 正, 耳, 亞, 上, and 非. Thus, 十 and 一 are the constituent components of 土, whereas 一丨一 are the constituent components of 工. The awareness of 十 being a distinct unit is likely to bring earlier recognition of important bujian like 艹, 土, and 木. Lattice like structures such as 用, 典, and 隹 are composed from either 十 or 丰, so that early introduction of these graphabet will made the student confident of being able to deal with these visually complex looking Hanzi. Familiarity with 十 will also help the student to be aware of difference from mimics like 长.

Similarly, 𠂇 and 乂 defines a spatial configuration where the oblique line intercepts a horizontal or oblique line at the midpoint, as seen in 左, 孝, 在, different from the non-intercepting horizontal and oblique lines in 石, 百, 衣, and 元. 右 is therefore a string of two graphabets, 𠂇 and 口 whereas 一/ 口 are the string equivalent of 石. The sharpened recognition of 乂 will also help to differentiate Hanzi with visually near intercepting lines like 米 and 釆.

Conceptually, graphabet are culture and custom neutral. It therefore avoids the problem of stroke order for 十, which is usually written horizontal preceding vertical in the usual Chinese stroke sequence, but vertical preceding horizontal in Japanese Kanji writing when part of another Kanji. Similarly, 厂 in 馬 is written with the vertical preceding the horizontal bar in Taiwan and Japan, but in a reverse order in Hong Kong.

5.3 How to use graphabet for Hanzi acquisition?

Many Hanzi can be seen to be formed by combination of just a few graphabets (table 3). Writing out the string in one dimension and speaking out the underlying graphabet will provide simultaneous input of visual, grapho-motor and auditory information about the configuration of the Hanzi.

On introducing Hanzi to the student, demonstrating that many Hanzi are formed from just two or three graphabet shapes and sound can rapidly decrease fears about learning to write or understand the Hanzi configuration. The sequence for graphabet is the same as the sequence according to the stroke order, only that in this case, a graphabet is the unit instead of a stroke. Moving on, tiered Hanzi with few graphabet per tier like 香 could be shown as an array with the top tier preceding the lower tier (/十ㅅ日 Piě shí jiān rì). Bipartite and tripartite Hanzi similarly could be written or read orally as arrays of graphabet starting from the left to the right.

TABLE 3: Examples of Hanzi as graphabet strings

Hanzi	Meaning	Graphabet strings	Auditory
大小	Size (big and small)	ナ ¥ 丨𠂉	chā nà shù jiān
自己	self	ノ 目 ㄱ一乚	Piě mù zhé héng jiǎo
中文	Chinese language	口丨 丶一乂	kǒu shù diǎn héng chā
早上	morning	日十 丨一 一	rì shí shù héng héng
下午	afternoon	一丨· 𠂉十	héng shù diǎn niǔ shí
冬天	winter	夂 ¥ · · 一ナ ¥	wān nà diǎn diǎn héng chā nà
火車	train	¥ ノ𠂉 一日十	nà Piě jiān héng rì shí
火车	train	¥ ノ𠂉 一𠂆十	nà Piě jiān héng rì shí
父母	Father and mother	𠂉乂 乚ㄱ· 一·	jiān chā jiǎo zhé diǎn héng diǎn
女子	girl	𠂇/ 一 𠂆十	Niǔ Piě héng zhuǎn shí

In the case of encasing graphabet from the top such as 冖, 厂, 冂, the encasing graphabet is always written before the encased graphabets. In the case of the lower encasement 凵, and other multi-graphabet lower encasement 辶, 廴, the encased contents are written first. Hanzi containing encasing graphabet which requires more complex graphabet order can be taught at a later stage.

Students are encouraged to orally enunciate the graphabet array and write in the air, to consolidate their memory of the configuration. Rote copying is usually not necessary, the Hanzi need only to be written a few times for familiarisation and practice. If the graphabet sequence can be orally recited, there is little difficulty handwriting on paper. Young children can be taught a hand movement depiction of the graphabet to improve their interest. A sequence of graphabet movements would then resemble a dance sequence to them.

The importance of displaying the graphabet composition of a Hanzi in a horizontal array and naming them should be stressed. Beginner students tend to miss dots or miscount the number of horizontal bars. Requiring them to recite or write out the array encourages paying close attention to the configuration. Such visual mistakes as confusing 氵 for 讠, 衣 for 农, 反 for 斤, 廴 for 辶, or mistaking the number of horizontal bars in 且 and 直 can be easily detected on a horizontal array. 女 is surprisingly difficult to write correctly for beginner CFL students probably because of its stroke order and pseudo square like shape, as well as resembling 𠂇. Displaying the horizontal array as 𠂇/一, it no longer poses any problem.

Using such an approach, the student will understand that visually complex Hanzi singleton which CFL students find difficult like 我, are just formed from 5 graphabets

(/十/乂•). The configuration can be remembered as “Piě shí tí chā diǎn”. By means of remembering the array visually and through sound, the writer will be able to distinguish 我 from the visually approximate Hanzi 找.

Graphabets underlie Hanzi, whether traditional or simplified, only that the frequency of usage may slightly differ, with more convoluted and right pointing angles amongst simplified Hanzi, and more of 目 in traditional Hanzi. Whether a Hanzi is traditional or simplified, the process of unzipping the two dimensional Hanzi into a string is the same, the only difference being the graphabets used for the different configurations. Thus, 車 and its simplified allograph 车 are in graphabet terms 一日十 and 一厶十, whilst the Hanzi for ghost 鬼, is written with a 田 in traditional Hanzi form, but with a 日 in mainland China with the diagonal penetrating the 日. The graphabet sequence therefore is for traditional Hanzi 鬼, / 口十 / 乚厶·, and for simplified Hanzi 鬼[2], / 日 / 乚厶·.

Because of its two-dimensional nature, some Hanzi may have similar graphabets but differ in minor details of length or the relative positioning of the graphabets. In such cases, oral enunciation will help to differentiate. For example, 土 can be differentiated from 土 because the former is a long armed 十. 力 and 刀, and 刁 are different because 力 is 乛丿 zhé with a long intercepting Piě, whilst 刁 is 乛/ (upward oblique zhé tí instead of zhé Piě). Similarly, both 太 and 犬 are written with 𠂇乀•, but the former has the dot below, and the latter has dot on the right shoulder of 大. Saying it out will help to clarify the differentiation.

As a rule, there is no difficulty recognizing the graphabet in Hanzi. Some fonts however may take artistic liberty and confuse dots and lines. 標楷體 or its equivalent should be used in case of doubt. The problem with font artistic embellishment has been discussed in the previous section 5.1. As for hooks, which are usually artistic embellishments and not cognitively important except in differentiating 干 and 于 or their derivatives, the 十 in 于 can be remembered orally as a hooked shí (shí gōu). Similarly, 牛 and 手 are differentiated orally with the latter being a hooked 扌.

Kanji, the Japanese equivalent of Hanzi, and Hanja, the Korean equivalent of Hanzi, can also be transformed into strings using graphabets, whilst noting differences in the configurations. For example, the Kanji equivalent for the Hanzi 真 is 眞, with 目 as the third graphabet; whilst the Kanji equivalent of 帶 is 帯, using different graphabet strings for the top tier. Similarly, Hanja uses 敎 instead of 教, thus starting the Hanja with a 乂 instead of a 十.

Pronouncing Chinese is often considered difficult by CFL students. Whilst we have allowed students to use if they wish their own names for the graphabets, with few

[2] *Because the unicode is the same for both traditional and simplified allographs, to appreciate this difference, the appropriate font should be used: the traditional 鬼 read with 新細明 font and the simplified with sim sun font.*

exceptions, CFL students seem to be comfortable with using the Chinese terms as they feel that they are also learning to speak Chinese at the same time. Students can also test each other as to the identity of the Hanzi heard as an array of graphabets.

With older children and adults, typing and exploring on a graphabet keyboard and use of a graphabet-based electronic dictionary is encouraged as it will reinforce the remembrance of the graphabet string as well as creating interest in the language.

The graphabet method is sometimes criticised for not emphasising the hook or the tilt in the horizontal stroke of a semantic radical. However, students are observed to become more aware of the aesthetic value and style after a while and their Hanzi gradually become the same as that of native writers. Not burdening them with aesthetic requirements in the beginning appears therefore a sensible approach without undesirable long term consequence.

5.4 Graphabet and the use of technology

Since nearly all Hanzi can be transcribed into horizontal array of graphabets, graphabets can also be used for keyboard input of Hanzi. Graphabets can be used to type in Hanzi using apps (Easy Chinese input pad) on mobiles or soft wares on computers. The advantage of a graphabet input system is that whilst pinyin input system may consolidate the ability to speak, the graphabet system consolidate the ability to write, providing feedback if there is faulty remembrance of the array and thereby the configuration of the Hanzi.

Because typing on a graphabet keyboard uses the same sequence as in writing, there is also near seamless transfer to pen writing ability. As the graphabet sequence in Hanzi moves from left to right, or from top to bottom downwards, and semantic radicals in Hanzi usually are positioned similarly, the user will be exposed to semantic bujian grouping, and therefore acquire increased awareness of phono-semantic compounds. In addition, whilst pinyin typing will make the user aware of the many homophones that exist, using a graphabet typing system educates the user to recurring patterns such as seen in 京哀享豪亮高, 尚掌堂常裳賞, 卷券眷養 and the realization that the many thousands of Hanzi are really made up from just a limited number of recurring units.

Using such a graphabet based keyboard, students can hit keys randomly and discover by themselves how Hanzi is formed from limited recurrent units. The simple combination of two graphabets like 十 and 口 can form 古, but combined with other graphabets become 舌, 固, 咕, 胡, 克, 故 etcetera.

5.5 Advantage of the graphabet approach

Using graphabets instead of strokes as the subunits of Hanzi strings provides several advantages in learning the structure of Hanzi and facilitating its writing. Firstly, it promotes clarity in configuration. Graphabets highlight salient frequently recurring

forms for visual recognition and for chunking during motor learning of the grapho-motor sequence. A sharpened awareness of discrete units means less chance of mistaken identification. Mistakes which CFL students often make, confusing 𠂉 and 𠂊, 𠂇 and 𠂈 are obviously visual in nature, and vocally articulating the graphabet composition should avoid such errors. Emphasising the features of angulated lines will decrease the risk of miswriting 八 instead of 儿 in 元 and 見, or miswriting 儿 for 八 in 頁. Mistaking 日 for 目 is common in novice students, with substitution of 日 for 目 in 眼, and substituting 目 for 日 in 春 or 明. Awareness of the discrete graphabets 日 and 目 also can be expected to decrease the risk of miswriting these Hanzi as well as other miswriting the 月 in 有 as 目, the 且 in 姐 as 目, or the 曰 in 書 as 目. CFL students are prone to be affected by their original language. Novice CFL writers are prone to write 灬as KK, and the lower component of 官 as a B, as well as treating 多 as AA. This is unlikely to occur when they learn the first two as 𠂉 • and as 丨冂一冂一, and 夕 as •, each time enunciating the graphabet name as visualizing and writing down the graphabets.

CFL students who are Japanese or other nationals with previous knowledge of Kanji, as well as Chinese students learning Japanese are prone to confuse Hanzi and Kanji forms which are visually approximate. Graphabets are helpful by highlighting the differences which are easily missed by the eye, e.g. 勇 for Hanzi and 勇 for Kanji, 直 for Hanzi and written as 十目𠃌 in Kanji, and 瘦 in Hanzi, and 痩 in Kanji. The Hanzi 傲 in traditional font is written with a 土方 in the middle, whereas in simplified Hanzi it is a sequence of 丰一/ 冫. By using the concept of graphabet, there is also no need to burden the novice Japanese CFL student with having to relearn how to write 右, 若, 青 just because the stroke order is different between Hanzi and Kanji.

Intercepting lines are very common in Hanzi, being able to identify 十ナメ and 丰 enables the student to distinguish other types of interceptions in Hanzi, such as those seen in 巾, 力, 书, 急 which involve encasing graphabets. Singleton configurations such as 我, 承, or bujian which are not Hanzi or are rare Hanzi, e.g. 𠦂, 𫩏, 寅, and consisting of intercepting lines or visually approximate components which are difficult to differentiate also can be described using graphabets, making their apparently complex structure much easier to understand. Similarly, dense looking Hanzi like 垂 or 華 are easily discernible and made easy by converting them into the graphabet strings of / 十十丨一一 and 十十一十十丰 respectively. Since a horizontal string of symbols is familiar to those from an alphabetical background, it can reduce the intimidation and anxiety brought about by an unfamiliar orthographic form. Furthermore, displaying the graphabet array according to "graphabet order" rules would clarify the writing sequence, ensuring that the configuration of the Hanzi would not be "drawn" as a picture with random sequences and therefore difficult to recall.

A further advantage with the use of graphabets is a decrease of memory chunks, thus reducing memory load and facilitating retention. For example, in the core 593

Hanzi the average number of strokes in traditional Hanzi is 9.82; whilst the average number for the same corpus written with simplified Hanzi is 7.88. However, the corresponding average number of graphabets for the corpus is 6.91 and 5.59 respectively. Both of these are comfortably within the working memory capacity of 7±2 uncompressed item. Thus, a Hanzi like 賽 has 17 strokes, and its simplified version 赛 14 strokes. It is however 8 graphabets for both the traditional simplified forms. A Hanzi like 卖 which has 8 strokes requires only 6 graphabet. Similarly, the ten stroke Hanzi 高 requires only 5 graphabets to remember. 香 has its stroke chunk size of nine also reduced to just four.

Most importantly, the use of graphabets makes possible oral enunciation of the internal configuration of Hanzi, and thus the creation of a corresponding auditory memory. Since graphabets can be given names, they can provide a means of enunciating the internal configuration of Hanzi and therefore provide auditory memory of the structural pattern of each Hanzi (Huang & Chan, 2012a). For example, the simple Hanzi 以 is surprisingly often miswritten by CFL students, writing either the dot as a slash or on top of the 𠂆. Speaking out “jiǎo diǎn jiān” and writing at the same time a sequence of 𠂆丶人 will markedly reduce the chance of such miswriting occurring, as writing differently would create a discrepancy between the auditory and visual memory, with the grapho-motor output. Mistakes seen in CFL students such as writing 规 both as a mirror image and with the wrong component (贝夫) is unlikely to happen if the graphabet sequence ≠ \冂/乚 has been spoken out and written in sequence. Nor is one likely to write 廾 in place of 丷 for 前, after having said its graphabet sequence out loud. The component part above 示 in 祭, is often wrongly written as the visually approximate 𥫱 by students. But when this component as written with graphabets, “wān diǎn tí, zhuǎn nà”, is read out, the auditory difference from 𥫱 “zhuǎn diǎn pié pié nà” is distinctly different and unmistakable. A common mistake by native or CFL students such as confusing 礻 for 衤 is also easily avoidable if the point that the former has a dot (diǎn) and the latter a left pointing angle (niǔ) is highlighted. Problems with remembering how many horizontal lines are in 真, 直, 且, 具, and 貝 is also easily overcome by orally enunciating the number of lines.

Not only can the use of graphabets confer auditory input and create an auditory memory engram to describe the structural configuration of each Hanzi, it also makes rote copying redundant. Repeating the string sequence of a Hanzi orally and writing in the air is just as good for consolidating the memory and can be done without a pen and pencil or any other writing instruments, and in any physical environment.

A further advantage of graphabets is the possibility of typing on a graphabet keyboard. This can provide feedback as to whether the internal configuration of a Hanzi is remembered or not, and therefore reduce the need for rote writing further. It provides an alternate means of typing and communicating to pinyin, whilst enhancing handwriting ability.

Novice CFL students sometimes would write each component of a multipartite Hanzi separately from the other. For example, 鹏 is written as 月月鸟. This problem rises from the failure to understand the physical size of a single Hanzi. With repeated use of the graphabet typing system, the student becomes familiar however with the spatial distribution and relative proportion of each part of the Hanzi. Whilst CFL students who write with the pen without sufficient preparation often produce strangely disproportionally shaped Hanzi, our experience has been that those who type several times before attempting to write tend to produce more normal looking Hanzi.

Another advantage in being able to type Hanzi whose pronunciation is unknown is an enhanced ability to learn newly encountered words. Thus empowered, students will be encouraged to explore independent reading of text containing new Hanzi words. When they encounter a new Hanzi, they can easily copy out the Hanzi using their knowledge of deconstruction, then use a graphabet sequence based dictionary (easy Chinese dictionary), or type it on a graphabet keyboard to seek its meaning on the internet. This overcomes the problem of pinyin based dictionaries, which require students to know the phonetic value of a Hanzi first, or traditional dictionaries which require knowing the bushou as well stroke numbers.

6. USE OF BUJIAN FOR WRITING SKILL ACQUISITION

After a Hanzi has been learnt, and its configuration string understood, it is obviously redundant to use the graphabets to describe it when it appears as a component of another Hanzi. Thus, larger and fewer memory chunks can result from the use of bujian.

As discussed previously, a bujian can have several properties. All bujian are by definition structural components of Hanzi. Bujian which are also recognized Hanzi may serve as phonological bujian to provide phonological information. A subset of these Hanzi also may in addition have semantic signifier ability acting as semantic bujian. A bujian which is not a Hanzi will usually have no semantic or phonological property and is purely a structural component. Some Hanzi in the role of a component bujian may also lose its phonological or semantic signifier property, and only behave as a structural bujian. For example, 六 has neither semantic nor phonological signifier ability when such a configuration forms a structural part of the Hanzi 交, 兗, 冥. Similarly, the 立 in 音 also does not carry any meaning or phonological information related to the Hanzi 立. It is just a section of the string of graphabets which form 音. The 米 in 粪 and 奥 likewise have nothing to do with rice. Conceptually therefore, an embedded bujian does not necessarily carry the same meaning as the free standing Hanzi it may visually resemble, or when it occupies the position of a semantic or phonological signifier.

Theoretically, therefore a bujian can be used to improve writing accuracy by improving semantic, phonological, and structural information as well as reducing chunk numbers and providing auditory memory.

6.1 Limitation to the use of bujian for Hanzi acquisition

However, certain constraints exist in real world experience.

As discussed previously, many phonological signifiers are insufficiently reliable. Although a significant number of reliable phonological bujian still exist, because of pessimism about their reliability, no published list exists for pedagogic purpose. Furthermore, although phono-semantic compound Hanzi is dominant in modern Hanzi, its dominance is less prominent on early contact with learning Hanzi. Of the 593 core Hanzi common to CFL students in mainland China, Taiwan, and France, only 384 (65%) are phono-semantic compounds. A survey of the common textbooks used for college level CFL students in America found also that only about half of the Hanzi in them were phono-semantic compounds (Fan, 2010). Furthermore, using bujian as phonological signifier is more difficult at the beginner stage. In the European Union and United Kingdom, the target for secondary school Chinese lessons is 600-1,000 Hanzi over 3-6 years. Most college level education in the United States aims at 500 Hanzi over 4 years (Tomizawa, Matsumoto & Endo, 2013). The beginner student will therefore encounter only a few Hanzi using the same phonological bujian at any one time. In the core Hanzi list, most phonological bujian appear only once or twice. Even in the 1604 Hanzi belonging to HSK class A and B list there are only a few reliable phonological bujian like 主, 采 which occurs three times or more. CFL students in American colleges find also only 14.5%-27.1% phono-semantic compounds containing reliable phonological signifiers in the textbook they use (Fan 2010). When a student is only in possession of a limited number of Hanzi, the phonological signifier is not likely to be functionally useful.

Semantic bujian are useful in signifying the semantic nature of the Hanzi. Because semantic bujian are usually positioned on the left side of a phono-semantic compound, they are also the beginning of a string. The text context will therefore help to trigger off the likely semantic bujian which can then further trigger off the grapho-motor program of the appropriate Hanzi with the right sound. Amongst competing homophones, semantic bujian can also narrow down the choice in the selection of writing output. They also tend to occur with much higher frequency even within the small set of Hanzi which the beginner student encounters. Thus, semantic bujian can become useful to the beginner student. However not all bujian provide reliable semantic information (P. R. Huang, 2009; C. Y. Huang, 2015). Only those which are reliable could impart useful semantic information. Of 115 semantic bujian whose semantic reliability exceeds 80% of Hanzi relying on them as semantic signifiers, only 66 were deemed worthy of teaching (C. Y. Huang, 2015). 馬 马 for example is not a reliable semantic bujian because of Hanzi in common modern usage such as 驕骄, 騙骗, 驗验, 駐驻. Similarly, the very common semantic bujian 艹 and 忄 are also less than reliable semantic signifiers for Hanzi in common usage. Thirdly, adult CFL students probably require more detailed explanation of the semantic bujian, as they are liable

to use these mistakenly if insufficiently informed. For example, substituting 米 for 饣 in 饼, and using 饣 for 米 in 糕.

6.2 How to use bujian effectively for Hanzi acquisition

At the early stage of encountering Hanzi, bujian is therefore best used in its role as a structural component for reducing chunk numbers and providing auditory memory for the purpose of remembering the configuration of the Hanzi. Bujian like 馬, 马, 灬, 艹, whilst unreliable as semantic signifier is nevertheless of frequent occurrence and therefore would be useful as structural bujian. A bujian like 隹 which will appear 15 times in the traditional Hanzi version of the core 593 Hanzi set, and 4 times in the simplified version, is useful for the same reason even if it rarely appears in isolation and is neither a useful semantic nor a reliable phonological signifier. Non Hanzi configurational bujian such as 丷 are also frequent occurrence and are therefore useful for the same reason. In using bujian to remember the configuration of Hanzi all bujian once learnt can be used, irrespective of whether they are Hanzi or non-Hanzi bujian. In addition, the bujian should be orally enunciated. Thus, 课 can be said to be a string of 讠 and 果 (yán guǒ); 端, 立山而 (lì shān ér). As discussed earlier, bipartite Hanzi are more liable to be miswritten by CFL students. Deconstructing into a string will convert the bipartite into two or more distinct units and make it easier.

The great majority of dense Hanzi of high stroke numbers is bipartite in nature, and therefore are composed of smaller discrete parts, i.e. bujian. Of the 79 commonly used simplified Chinese Hanzi possessing 21-22 strokes each, 28 are singletons. But of these, even complex singletons are often just strings of a few chunks. For example, 夔 is just a string of five bujian chunks 首止巳八夂, 鬻 of 4 chunks, 龢 of two chunks, and therefore all become quite manageable. If a student enunciates the name of these parts, an auditory memory of the Hanzi is created, enhancing the chance of remembering how to write out the string correctly and construct the Hanzi.

The need to enunciate presupposes that the bujian has a name. For bujian which are also Hanzi, a sound equivalent obviously would exist. Most non Hanzi bujian, such as 氵, 羊, would also have a name, although not all such names are universally accepted. For bujian composed of just two or three graphabets however, just describing it in terms of strings of graphabets may be equally as useful. Similarly, bujian which are rarely used Hanzi, are perhaps also better remembered in terms of its graphabet string composition.

Bujian alone are obviously useful for describing Hanzi once they have been encountered and learnt. But graphabets may still be necessary to form a bujian when it has not been previously encountered and learnt, or if it is a Hanzi such as 监, 画, which simply cannot be formed using bujian alone, or is of such uncommon use that teaching it would be a diversion with little beneficial outcome.

Thus, saying that 炼 is made up from 火 and and the right sided bujian of 拣 would be quite appropriate once 拣 has been learnt, but on initial contact with the Hanzi 拣, saying that it is made up from 扌 and the right sided bujian of 拣 would make no sense. This bujian thus will need to be described at that stage as the graphabet string 一𠂉冂八.

The usefulness of bujian is indeed markedly limited by the inconvenient fact that the appearance of individual bujian frequently follows rather than precede Hanzi which contain it. Thus, 多 is likely to be learnt before its component bujian 夕, 家 ahead of the bujian 豕, 睡 ahead of 垂, 藍蓝 ahead of 監监, and 跳 ahead of 兆. Furthermore, if a bujian has low recurrence rate in the textbooks, the value of teaching it on encounter becomes also doubtful.

In other case, a Hanzi may consist of parts which are non- recurring or low recurring configurations bereft of meaning or sound. A combination of bujian and graphabets are also especially useful for remembering the internal configuration of uncommonly used bujian. The bujian 黍 is a case in point. This is a rarely used Hanzi, which appears in modern usage only as a component of two frequently said and read but less commonly written and phonologically different Hanzi, 漆 and 膝. 鬳 is another bujian seen in few Hanzi, and nowadays only in the commonly said but seldom written Hanzi 噫 in the bound morpheme 噴噫 (sneeze). Teaching the configuration of these unusual bujian with graphabets and component bujian but not teaching the whole bujian would seem more useful.

7. SUMMARY OF METHOD

In summary, the graphabet and bujian method for improving the ability to produce an appropriate and correct Hanzi configuration consist of the following steps:

1) Learn the basic 21 graphabets, their geometrical shapes and names. If the students are taught these before they master Chinese intonation, they may consider using names for the graphabets in their own national language initially as a bridge over, although in our experience students tend to prefer plunging into the Chinese names directly.
2) Encourage students to detect the graphabets in common Hanzi, so that they know how to recognize them. Teach Hanzi and key structural non Hanzi bujian that can be formed from a combination of just a few graphabets such as 上, 下, 工, 木, 早, 山, deconstruct these Hanzi so that they are strings of graphabets and enunciate the string so as to create an auditory memory of the Hanzi.
3) Learn the basic visual structure of Hanzi, and the basic rules which govern the sequencing of graphabets and bujian.
4) Enunciate each Hanzi as a string of graphabet to establish an auditory memory of the configuration.

5) Write out each Hanzi as a string of graphabet and as its final form.
6) Once a Hanzi or its abbreviated form has been learnt, then it too can be used in the string. Learn to use combinations of graphabets and bujian to form strings from Hanzi encountered in the text, and read out the strings, committing it to auditory memory.
7) Learn reliable semantic and phonological bujian as they present, giving explicit instruction on how to identify at such times.
8) Practice using the graphabet based keyboard to type out the Hanzi so as to provide feedback on the knowledge of the configuration.
9) Children can practice remembering Hanzi as a string of graphabet based hand movements.
10) Handwrite Hanzi only after having acquired the ability to correctly write out the string or typing out the string on a graphabet based keyboard.
11) Use a graphabet based dictionary to search for Hanzi.

8. IS THE METHOD USEFUL IN PRACTICE?

The graphabet and bujian method of learning to write Hanzi has evolved over the last decade. Unfortunately, the majority of CFL students in Hong Kong are underprivileged migrant children from South East Asia. Schools that have to teach such students would not allow experiments to be carried out on them, so that permission to do controlled trials on a large scale and over a long period has so far been unobtainable. There has therefore been only four preliminary proof of concept studies. A brief summary of these studies is hereby provided.

8.1 Attitude towards learning Chinese

CFL students are known to suffer frequently from anxiety. Anxiety has been blamed as one of the reason for drop out from further studies. Motivation is another factor which may affect continuation in study.

Research question: Can the use of the graphabet and bujian method reduce anxiety and affect motivation amongst CFL students over time?

Research design: A self-evaluation survey of attitude towards Hanzi and Chinese learning before and after a period of learning using the graphabet based method.
Instrument: The instrument used in the survey was a simple self-administered ten points analogue scale (1 minimum and 10 maximum) evaluation of (a) interest in learning Chinese (b) degree of difficulty with Hanzi.

Participants: 14 adult woman CFL students attending a community centre interest class where both oral and written Chinese was taught at the same time. 10 of the participants were South East Asians and 4 from Europe or America. None could speak fluent Chinese, and only two had attended previous Chinese classes. With the exception of the 4 from Europe or America who were deliberately seeking to learn Chinese, the others were learning Chinese for everyday usage in the community.

Learning sessions: The students were taught spoken Chinese and written traditional form of Hanzi using Cantonese, the local language used in Hong Kong for pronunciation of the language. Medium of teaching was English. The texts used were both tailored teaching lessons and textbooks for native speakers in primary schools. Exposure to Hanzi follows immediately after learning of the oral language. Pin yin supplement was not used but students were allowed to write down using their language any phonetic notation they wish.

Procedure: The survey was carried out before commencement of studies, and after 50 hours of tuition. At this stage, usually about one hundred Hanzi would have been taught using the graphabet and bujian method. The participants were asked to evaluate their interest in learning Chinese and perceived degree of difficulty with Hanzi by giving a score on the analogue scale on each of the two issues. The results were analysed using a simple Chi square statistics, using a scale of 7 and above reflecting good interest, and a difficulty level of 5 or more being unacceptable.

Findings: The results showed marked improvement in interest (poor interest in 13 participants prior to the study and only 1 participant after, $\chi^2(1) = 7.143$, $p < 0.01$) and drop in perceived unacceptable difficulties (all 14 participants before, and only 1 participant after, $\chi^2(1) = 10.286$, $p = 0.001$) (Huang & Chan, 2011).

8.2 Attitude towards learning Chinese II

A similar study using the same methodology was carried out on 26 male Hong Kong secondary school CFL students from South East Asia in a Hong Kong secondary school. They have already been exposed for variable periods to Chinese school lessons. The lesson sessions used school text and in addition tailored teaching material using the graphabet and bujian method. Teaching of the method was supplementary to their usual school lessons and the rate of teaching of Hanzi was determined by their usual school teacher. The interest in Chinese was low in all 26 before introduction of the method, but remained low only in 3 after 50 hours of study, with the teaching of about a hundred Hanzi by means of the graphabet and bujian method ($\chi^2(1) = 15.385$, $p < 0.001$). Perceived unacceptable difficulty in Hanzi was felt by all 26 beforehand and still present in 12 after ($\chi^2(1) = 0.154$, $p = 0.695$). Despite the improvement in interest, the secondary school students unlike the adult students failed as a group to achieve statistical improvement in their perception of difficulty about learning Hanzi. Contradiction between the way Hanzi is taught by their regular school teacher and this method which is taught as a supplementary method, as well as the small sample size may be some of the reasons for the discrepancy.

8.3 Performance at IGCSE Chinese language examination

The IGCSE Chinese language examination is the exit examination that Hong Kong CFL school children can sit for in their secondary school after 10 years of formal education.

The research question was whether the graphabet and bujian method could facilitate faster and effective learning when used by adults in a community centre.

Participants: A group of adult CFL housewife attendees at a community centre were encouraged to participate with the goal of passing the examination. 9 students, 7 from South East Asia and 2 from Europe entered the project, out of a total of 15 students who originally entered the CFL course in the centre. Only the ones from Europe had some earlier experience with learning Chinese.

Procedure: The students were given spoken and written Chinese lessons twice a week, each time lasting two hours. The medium of instruction was English, the Hanzi pronunciation and oral language taught were all in Cantonese. Traditional form of Hanzi was taught. Pin yin was not used but students were allowed to use notations if desired. All students were taught to write Hanzi using the graphabet and bujian based method described in this paper, and to to use the graphabet keyboard. They were all told to handwrite only after becoming familiar with writing Hanzi as a string and frequent practice writing with the graphabet computer software system.

Texts used included tailored teaching material as well as formal textbooks.

The endpoint was passing the examination.

Result: After 260 hours of study, they took the examination resulting in in 8 out of 9 of the students passing, 3 at grade A level and 5 at B or C level.

Discussion: Participants were adults and motivated. Some of them were hoping that passing the examination would secure job opportunities. These variables might have bestowed a positive impact on the outcome. Another factor which might have affected the students is the fact that several of them would copy down new Hanzi which they encounter in the streets or everyday encounters, type it out and find out its meaning and usage on the internet. It is possible that this behaviour which broadened their Hanzi reading exposure benefitted their performance in the reading and writing part of the examination. It is also possible that graphabets enhanced their ability to read because of heightened awareness of distinguishing features in the Hanzi. Lastly, whilst these participants are mostly housewives with limited time for study, not having to do rote copying and having feedback from the use of graphabet based keyboard might have meant that they have more efficient use of study time. Whilst the above explanations are all conjectures, the result does demonstrate that the graphabet and bujian method is worthy of a larger study with participants of different degrees of motivation.

No published information is available regarding the passing rate of ethnic minority students from Hong Kong schools who took this test. Many Ethnic minority students in Hong Kong stop learning CFL and never took the examination. A study published in 2005 found that only 71 of 200 secondary school ethnic minority students

interviewed had undertaken CFL studies, and only 9 were still doing CFL. (Ku, Chan & Sandhu, 2005).

8.4 Effect on kindergarten students

The authors of the only short term controlled study were interested in the question whether the graphabets could be useful if taught to learn Hanzi amongst kindergarten students, who are either first language or CFL.

Participants were from two K3 classes in a kindergarten school. Two classes, each of 15 students, were compared; one class was taught using the traditional method and one by using the graphabet method.

Teaching Methodology: The graphabets were taught as hand movements as well as sounds and visual symbols. The school textbook provided for 23 new Hanzi in the form of six sentences to be taught to the students during 6 separate classes during the month.

Test instrument: The researchers designed a scoring card and a writing sample sheet. The 23 Hanzi were presented to the children one by one on individual cards.

Procedure: Reading, writing and stroke order ability were tested before commencement of the study and after a study period of one month. All 23 Hanzi in the school curriculum were presented to the subjects to read, one Hanzi at a time on a card. The researchers then chose randomly from the list of Hanzi which the child could read out correctly, and asked the child to write it out. Video recording of their writing attempts were analysed for their use of the stroke order. Because the reading ability was unbalanced at the start the researchers did not report on the reading outcome at the end of the experiment. One of the experimental group students missed the post teaching test; as a result, the researchers excluded one from the control group from the final analysis. The results were analysed using a SSPS package for Ancova analysis.

Results showed that prior to training, the writing and stroke order ability were similar in the two groups but reading ability was better in the control group as some of the participants had been taught to read the words before. After one month, the experimental group using the graphabet method was however superior to the the group using traditional methods in both writing ability ($F(1, 28) = 12.718$, $p = 0.001$) and stroke order awareness ($F(1, 28) = 8.918$, $p = 0.006$) (Chan-Wong & Tang, 2011).

9. SUGGESTION FOR FURTHER STUDIES

There have been few experimental studies on methods of improving writing output ability in Hanzi. What studies there are, whether published or in thesis work have been short term. Students who are highly self-motivated or of high aptitude may be able to achieve good competence in a language, regardless of which method of learning being used. A large investment in time and effort may also achieve such results for a less talented student. A useful method of teaching writing ability should

be applicable to the average student, and having high cost effectiveness ratio, with the student not having to sacrifice other studies just to achieve competence in Chinese.

The proof of concept studies reported above have looked at several different issues which should be assessed at a future study on the usefulness of the graphabet and bujian method or other methods of teaching how to write better. These proof of concept studies are limited either by lack of a control group, or limitation in observation time. A future study should compare the effect of different methods on student anxiety and perceived difficulty which may impact on willingness to carry on study, the student dropout rate, the amount of time and effort a student need to invest in acquiring Hanzi writing proficiency, and finally the result in proficiency in writing over a substantial period of time, as well as overall impact on the academic achievement reached. Obviously, if a student devotes a great deal of time and effort, achievement in Hanzi or Chinese studies may be at the expense of other academic subjects. As for the effect on Hanzi and Chinese learning, errors in writing, such as failure to produce a Hanzi, configuration errors as well as homophone substitution should all be noted, as also the quality of essay production. Such studies should be controlled studies, comparing this method against the de-facto method of rote copying, or other methods yet to be revised. Such controlled studies should be stand alone, and the students are not exposed to different methods of analysing Hanzi at the same time period.

10. CONCLUSION

Current methods of acquiring competence in writing Hanzi have high time and effort cost, and is probably a major cause for the poor uptake and high drop-out rate amongst CFL students.

This paper discusses a new method of teaching how to acquire proficiency in Hanzi writing. Graphabets, a subset of 21 bujian are suggested to be the fundamental units of Hanzi organization instead of strokes. Use of graphabets is likely to decrease memory chunk load, create a clear clarification of the structural architecture of Hanzi, and contribute to the creation of an auditory memory of the architecture. In addition, technological teaching facility providing a feedback and self-learning through the use of a graphabet based keyboard will cut down the need for rote copying for learning purpose. The graphabet method is suitable for the beginner stage of Hanzi learning and, together with the use of bujian, as the vocabulary grows will likely reduce the difficulty of Hanzi learning and decrease the attrition rate seen amongst CFL students.

Although this method has not been able to undergo rigorous empirical studies, proofs of concept studies have shown that in the short term at least, the method does decrease learner anxiety and sense of perceived difficulty with Hanzi. It also has shown efficacy in actual performance both in children and adults. Further empirical studies are warranted.

There is no known conflict between this method and methods of teaching recognition or understanding of Hanzi.

REFERENCES

Australian government department of Education, employment and workplace relationship (2010). *Current state of Chinese, Indonesian, Japanese, and Korean language education in Australian schools.* Education Services Australia.

Cai, Y. H., 崔永華 (1997). 漢字部件和對外漢字教學 [Components of Hanzi and teaching of Hanzi to foreigners]. *語言文字應用* [Applied Linguistics], *23*, 49-54.

Cao, F., Vu, M., Chan, D. H., 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

Carruthers K. (2012). *The teaching and learning of Chinese in schools: Developing a research agenda to support growth.* London, UK*:* Institute of Education, University of London.

Chen, H. C., Chang, L. Y., Chiou, Y. S. & Sung, Y. T., 陳學志, 張璨勻, 邱郁秀, 宋曜廷, 張國恩 (2011). 中文部件組字與形構資料庫之建立及其在識字教學的應用 [Chinese orthography database and its application in teaching Chinese characters]. *國立臺灣師範大學教育心理與輔導學系教育心理學報 (閱讀專刊)* [Bulletin of Educational Psychology, National Taiwan Normal University (Special Issue on Reading)], *43*, 269-290.

Chan-Wong, K. C. & Tang, H. K., 黃潔貞、鄧浩堅 (2011 Dec). 在香港幼稚園應用中文字母教識字的初探 [Preliminary investigation on the use of Chinese graphabets in Hong Kong kindergartens]. 第十屆世界華語文教學研討會 [Paper presented at 10th World Chinese language teaching conference Taipei, Taiwan ROC].

Chang, L. P., 張莉萍 (2008). 對外漢語字集 [Corpus of Hanzi for foreign teaching]. 台灣華語文教學年會暨研討會 [Paper presented at Taiwan Chinese language teaching annual conference 2008 Taiwan].

Chiu, C. L., 邱慶玲 (2003). *避免小學生寫錯別字之教學策略* [Teaching strategies for preventing elementary students from incorrectly writing and using characters]. MA thesis, Hsinchu College of Education, Taiwan. Retrieved from http://handle.ncl.edu.tw/11296/ndltd/71276830229261456858

Chu, C. (2009). The component-oriented net-weaving approach to character teaching: Rationale and application. Paper presented in the 2nd International Conference on Chinese Language Pedagogy, UC Berkley.

Fan, H. M. (2010). *Developing orthographic awareness among beginning Chinese language learners: investigating the influence of beginning level textbooks*. Ph.D. thesis, University of Iowa. Retrieved from http://ir.uiowa.edu/etd/496/

Government of Hong Kong SAR, Office of the government Chief information officer, (2002). 香電腦漢字楷體字形基礎部件表 [Table of basic components for Kai style Chinese font in Hong Kong].

Graham, S. & Hebert, M. A. (2010). *Writing to read: Evidence for how writing can improve reading*. *A Carnegie Corporation Time to Act Report.* Washington, DC: Alliance for Excellent Education.

Guo, S. L. (2009). Dynamic counting and contrast between the stroke numbers of simplified and traditional Chinese characters, *Journal of Beihua University (Social Sciences)*, *10*(2), 50-56.

Hsu, S. H., 许淑华 (2011). 对外汉字教学策略探讨 [Discourse on strategy to teach Hanzi as foreign language]. *海外华文教育* [Overseas Chinese Education], *58*, 33-57.

Huang, C. Y., 黃震遐 (2006). 人腦、電腦，和識字寫字 [The human brain, computer, reading and writing]. In 《*识字教育科學化論文集粹*》 [Selection of scientific educational papers on Hanzi learning] (pp. 431-436). 中國輕工業出版社 [Chinese light industry publishing company].

Huang, C. Y., 黃震遐 (2015). 部首意符的教育價值 [The educational value of semantic radicals amongst index signifiers (Bushou)]. *漢字漢文教育* [Hanzi and Han language education], *37,* 7-23.

Huang, C. Y. & Chan, Y. L. R, 黃震遐、陳躍良 (2011). 漢字字母教學法： 理論和實踐。[Graphabet teaching method; theory and practice]. *漢字漢文教育* [Hanzi and Han language education], *26*, 383-397.

Huang, C. Y. & Chan, Y. L. R., 黃震遐、陳耀良 (2012a). *中文字母寫字口訣* [Mnemonics using graphabets for writing characters] 青田教育中心、中文環境系統(香港)有限公司出版 [Greenfied Education centre and Sinovista Systems (Hong Kong) Ltd].

Huang, C. Y. & Chan, Y. L. R. 黃震遐、陳耀良 (2012b). 必先利其器：漢字教育的字典和電腦發展 [First sharpen the tools: dictionary and progress in computer use in Hanzi education]. *漢字漢文教育* [Hanzi and Han language education], *28*, 215-234.

Huang, J. C., 黃金城 (2016). 論漢字筆畫 [On the strokes of Chinese characters]. *云南師范大學學報(對外漢語教學與研究版)* [Journal of Yunnan Normal University (Teaching and Research on Chinese as a Foreign Language)], *14*, 48-61.

Huang, J. Y., 黃靜吟 (2005). *漢字筆順研究* [Study of stroke order in Hanzi]. 駱駝出版社 [Camel publishing Ltd Taipei].

Huang, P. R., 黃沛榮 (2009). *漢字教學的理論與實踐* [Theory and Practice of Chinese character teaching and learning]. Tapei Lexus shuju.

Institute of information sciences, Academia Sinica (n.d.). 中央研究院資訊所. *漢字部件及組字規則* [Hanzi components and rules for forming an ideograph].

Jen, T. & Xu, P. (2000). Penless Chinese character reproduction, *Sino-Platonic Papers*, *102*, 1-15.

Jiang, X. & Cohen, A. D. A. (2012). Critical review of research on strategies in learning Chineseas both a second and foreign language. *Studies in Second Language Learning and Teaching*, *2*(1), 9-43. https://doi.org/10.14746/ssllt.2012.2.1.2

Ke, C. R. & Li, Y. H. A. (2011). Chinese as a foreign language in the US. *Journal of Chinese Linguistics*, *39*(1), 177-238.

Kiefer, M., Schuler, S., Mayer, C., Trumpp, N. M., Hille, K. & Sachse, S. (2015). Handwriting or typewriting? The influence of pen- or keyboard-based writing training on reading and writing performance in preschool children. *Advances in Cognitive Psychology,* 31, *11*(4), 136-46. https://doi.org/10.5709/acp-0178-7

Ku, H. B., Chan, K. M. & Sandhu, K. K. (2005). A Research report on the education of South Asian ethnic minority groups in Hong Kong. *Research Report Series No. 11, Centre for Social Policy Studies Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong.*

Liu, J. H., 刘居红 (2008). 对外国学生汉字书写偏误的分析兼谈汉字教学 [Analysis of Hanzi miswriting by foreign students and a discussion of the teaching of Hanzi]. *喀什师范学院学报* [Journal of Kashgar Teachers College], *29*, 94-97.

Liu, L. P., 刘丽萍 (2008). 筆畫數與結構方式對留學生漢字學習的影響 [The impacts of the stroke number and structural way of Chinese characters on foreign students in learning Chinese characters]. *語言教學與研究,* [Language Teaching and Linguistic Studies], *1*, 89-96.

Liu, Y. M. & Jiang, X., 柳燕梅, 江新 (2003). 欧美学生汉字学习方法的实验研究'回忆默写法与重复抄写法的比较 [An experimental study on European and American students' Chinese character learning strategies: A comparison between recall-based writing and repetitive copying]. *世界汉语教学* [World Chinese Teaching and Learning], *1*, 59-67.

Liu, M. H., Ma, S. Z. & Tong, Z. Y., 劉曼華, 馬淑珍, 童正玉 (1979). 集中識字試驗二十年 [20 years of experience with the concentrated learning of character experiment]. *人民教育* [People's Education], 11, 24-27.

Liu, W. M. & Su Y. Q., 劉惠敏 蘇育全 (2014). 德國華語文小學教材之漢字教學研究 [A teaching research on the characters of Chinese language textbooks for the elementary schools in Germany]. *中原華語文學報* [Chung Yuan Journal of Teaching Chinese as a Second Language], *14*, 83-114.

Long, W.H. 龙韦宏 (2008). 小学生错别字泛滥的原因与对策 [Reason and countermeasures against the tide of miswritten and incorrect Hanzi amongst primary school students]. *湖南教育(语文教师)*[Hunan Education (language teacher)], *6*, 7-8.

Mastroberardino, S., Santangelo, V., Botta, F., Marucci,F. S. & Olivetti Belardinelli M. (2008). How the bimodal format of presentation affects working memory: An overview. *Cognitive Processing* (1), 69-76. https://doi.org/10.1007/s10339-007-0195-6

Mathy, F. & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition*., *122*(3), 346-62. https://doi.org/10.1016/j.cognition.2011.11.003

Ministry of Education PRC, 中華人民共和國教育部 (2002). *全日制義務教育課程標準* [Ministry of Education, People's Republic of China. Standard for full day Obligatory education curriculum].

Ministry of Education PRC, 中华人民共和国教育部 (2009). *现代常用字部件及部件名称规范*[Specification of common modern Chinese character component and component names].

Ministry of Education ROC, 中華民國教育部 (2008). *常用標準字體筆順手冊* [Handbook of stroke order for commonly used standard glyphs]. Tapei Taiwan.

Orton, J., Tee, J., Gong, J., McCulloch, J., Zhao, Y. L. & McRae, D., (2012). *Profiles of Chinese language Programs in Victorian schools.* Melbourne, Australia: Chinese Teacher Training Centre, The University of Melbourne.

Shao Y. Q., 邵迎强 (2008). 预防和纠正错别字的方法 [Method of preventing and correcting miswritten and incorrect Hanzi]. *甘肃教育* [Gansu education], *7*, 39.

Shen, H. H. (2005). An investigation of Chinese-character learning strategies among nonnative speakers of Chinese. *System*, *33*(1), 49-68. https://doi.org/10.1016/j.system.2004.11.001

Shi, Z. Y., 施正宇 (2000). 外国留學生字形书写偏误分析 [Error analysis of Chinese characters written by foreign learners]. *汉语学习*[Chinese Language Learning], *2*, 38-41.

Si, X., 斯霞 (1978). *斯霞教育經驗選編*。[Selections of the teaching experience of Si Xia]. Beijiing, China: People's Education Press.

Sichuan Jingyan county education department curriculum unit, 四川省井研县教育局课题组 (1994). 字族文识字" 教学实验研究 [A Study on the teaching experiment of "Learning Chinese characters by the character families"] *教育研究* [Educational Research], *5*, 49-53.

State Language Commission, 国家语言文字工作委员会 (2000). *现代汉语通用字笔顺规范* [Stroke order rules for common words in modern Han language], [Beijing, China: Language and Literature Press].

Su, J. B., 蘇靜白 (1991). 循著漢字構字規跡建立識字教學體系 [Constructing a character learning system following the tracks of Hanzi organization]. *汉字文化* [Chinese Character culture], *2*, 22-24.

Tan, L. H., Spinks, J. A., Eden, G. F., Perfetti, C. A. & Siok, W. T. (2005). Reading depends on writing, in Chinese. *Proceedings of the National Academy of Sciences* USA, *102*, 8781-8785. https://doi.org/10.1073/pnas.0503523102

Tomizawa, S., Matsumoto, K. & Endo, M. (2013). Effective Chinese and Japanese Character instruction using an etymological CALL approach. In S. Dhonau (Ed.), MultiTasks, MultiSkills, MultiConnections (pp. 71-88). Milwaukee, WI: Central States Conference on the Teaching of Foreign Languages.

Tseng, C. J., 曾啟瑞 (2008). *國小學童寫作能力發展之研究* [The research of pupils' development on writing ability in elementary schools]. M.A. thesis, National Tai Tung University. Retrieved from http://ir.lib.nttu.edu.tw/bitstream/987654321/341/1/200884165812.pdf

Tseng H. C., Chang, L. H. & Chen C.K., 曾性初, 陈履祥, 陈昭宽 (1965). 汉字的各种笔画的使用频率的估针 [The relative frequencies of the various stroke types of the Chinese ideogram]. *Acta Psychologica Sinica 3*, 212-214.

Tseng, J. J., 曾金金 (2000). 德國學生學習漢字的情況及其學習策略 [Character learning strategies among German students]. *語言研究特刊* [Language Research supplement], 323-327.

von Koss Torkildsen, J., Morken F, Helland W. A. & Helland T. (2016). The dynamics of narrative writing in primary grade children: writing process factors predict story quality. *Reading and Writing: An Interdsiciplinary Journal, 29*(3), 529-554. https://doi.org/10.1007/s11145-015-9618-4

Wang, G. Y., 王貴元 (2014). 漢字筆畫系統形成的過程與機制 [Process and mechanism of Chinese character stroke system]. *語言科學* [linguistic sciences], *72*, 549-560.

Wang, H. W. & Su, Y. X., 王汉卫, 苏印霞 (2015). 現代漢字筆畫系統的簡化、排序及命名 [The simplification, sorting, and naming of the stroke system of modern Chinese characters]. *語言文字應用* [Applied Linguistics], *1*, 19-26.

Wang, J. & Koda, K. (2013). Does partial radical information help in the learning of Chinese characters? In Erik Voss et al. (eds.), *Selected proceedings of the 2011 Second Language Research Forum* (pp. 162-172). Somerville, MA: Cascadilla Proceedings Project.

Wu, S. Y., 伍順英 (2015). 關于語文教學中低年級學生易錯字的有效教學研究 [Research into effective teaching method in language teaching for easily miswritten Hanzi by middle or low grade students]. *教育教學論壇* [Education Teaching Forum], *42*, 251-252.

Wu, Y. (2014). The effect of mixed-sensory mode presentation on retaining graphic features of Chinese characters. In Z. H. Han (Ed.), *Studies in second language acquisition of Chinese* (pp. 126-41). Bristol U.K.: Multilingual Matters.

Xiao X. Q., 肖奚強 (2002). 外國學生漢字偏誤分析 [An analysis of character errors of foreign learners of Chinese]. *世界漢語教學* [Chinese Teaching in the World], *2*(12), 79-86.

Xie, T. W. (2011). “手写”还是“电写”——电脑输入中文引起的讨论. [“Handwriting” vs “electronic writing” which way we should go in teaching Chinese?]. *Journal of Chinese Language Teaching and Research in the U.S.*, *6*(1), 98-102.

Xu, X. & Padilla, A. M. (2013). Using meaningful interpretation and chunking to enhance memory: The case of Chinese character learning. *Foreign Language Annals*, *46*(3), 402-422. https://doi.org/10.1111/flan.12039

Yang K. G.,楊開貴(2007). 外籍學生中文錯字之研究 [Chinese character errors from learners of Chinese as a second language]. M.A. thesis retrieved from http://theses.dyu.edu.tw/ Dayeh University Taiwan.

Yin J. J. H. (2003). Improve the efficiency of Chinese character teaching at the elementary level in American universities: issues and tactics. *Journal of Yunnan Normal University*, *1*, 6-11.

You, H. J. (2003). 筆畫數、部件數和拓撲結構類型對非漢字文化圈學習者漢字掌握的影響 [How is the acquisition of Chinese characters by learners from the non Chinese character cultural sphere affected by stroke numbers, component numbers, and topological structural patterns?]. *世界漢語教學* [Shijie Hanyue Jiaoxue , *64*, 82-81

Zhang, Q. & Reilly R. (2015 Nov). Writing to read: the case of Chinese. Paper presented at 29th Pacic Asia Conference on Language, Information and Computation (pp. 341-350). Shanghai, China.