Focusing on a Subset of Scripts Enhances the Learning Efficiency of Second Language Writing System

Ching-Pong Au
Community College of City University
6/F, AC2, City University of Hong Kong,
Tat Chee Avenue, Kowloon Tong, Hong Kong
chingpau@cityu.edu.hk

Yuk-Man Cheung
Hong Kong Baptist University
SCE, 2/F, Franki Centre, 320 Junction Road
Kowloon Tong, Hong Kong
cheungym@hkbu.edu.hk

Charles Chen, Jr.
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
ctchen@polyu.edu.hk

Abstract

Memorizing the whole set of graphemes is generally accepted as the first step of learning a phonogramic language. However, it is demanding for L2 learners to familiarize the whole inventory of graphemes in advance if the language has a relatively large inventory size. We propose that learning a subset of graphemes would largely enhance the learning efficiency by reducing the memory burden. With homophony minimized, effort of acquiring vocabulary in elementary stage can be greatly reduced. In this paper, the writing system of Thai is used to illustrate the main idea. Besides, the method may also be extendable to Japanese and Korean, which grapheme inventory sizes are smaller.

1   Introduction

There is a general assumption in many language textbooks that L2 learners are able to comprehend and produce scripts of any second language just because they have acquired the reading and writing ability in their first language. However, different processes and strategies are actually involved in L1 and L2 writing systems (L1WS and L2WS) and should not be ignored by the learning material designers or the language teachers (Bassetti, 2006). English is an international language learnt globally by most people and textbooks of English are perhaps the most convenient model being imitated by textbooks of other languages for L2 learners. In the first lesson of learning a L2WS, the first thing that comes to our mind is the 26 letters. Therefore, learning how to read and write kana and hangul are expected in the first chapter of Japanese or Korean textbooks, because they function like the 26 letters of English as the most basic building blocks in the writing system (WS, hereafter).

Although it may take some time for learners to "swallow" graphemes like kana or hangul, it is not a daunting task for most people to master (at least to recognize) them well enough for general usage. However, there are also WSs which are less "learner-friendly" – Thai WS is one of them. In Hong Kong, we observed through classroom teaching that many L2 learners of Japanese or Korean can memorize kana or hangul reasonably well within weeks or even days; while L2 learners of Thai tend to give up learning to read and write for good, and confine their learning to verbal language.

In the present thesis, we propose that (1) WSs that have relatively large inventory size could be learnt more efficiently through reordering the learning sequence; (2) the usage frequencies of the basic writing units are important criteria for deciding on the learning sequence; (3) learning materials designed for L2WS adult learners should be different from those for L1WS children learners, as two groups have had different experience before they learn the target WS.

In our discussions in this paper, we would like to take learning Thai scripts as an illustration for

the concept. In section 2, some properties of the Thai WS are compared with those of other languages. In section 3, the Thai WS (symbols for consonants, vowels and tones) will be briefly introduced and the difficulties faced by L2 learners will also be highlighted. Functional approach seems to be a more effective way for learners to acquire Thai WS. Based on the estimated usage frequencies of each consonant letter from an online dictionary, an optimized learning sequence is proposed in section 4. In the earliest steps of the sequence, beginners are recommended to restrict themselves to only a subset of consonant letters. They should not start learning the other remaining letters before they are highly familiarized with the subset in the context of basic daily life vocabularies. Each Thai consonant has a name to disambiguate the homophonic consonant letters. In section 5, we will explain why the names are burdens to foreign learners although it may be helpful to native Thai children to remember the consonant letters. In section 6, the method proposed is applied to Japanese and Korean. Finally, an overall conclusion will be drawn in section 7.

2   Comparing the Writing Systems

To measure a WS, there are 10 useful properties: inventory size, complexity, frequency, ornamentality, distinctivity, variability, phonemic load, grapheme size, grapheme load and letter utility, according to Altmann (2008). In some of the above properties, "L(etter)", "G(rapheme)" and "P(honeme)" are carefully distinguished in the definition of the properties. Letters refer to the basic writing units of a language, such as the 26 letters in English; Graphemes are the units that can minimally distinguish the meanings in writing. For example, "ph" in "phone" is a grapheme as it makes contrast with "z" in the word "zone". Phonemes are the phonological units that can minimally distinguish the meanings in speech. For example, in English, there are 24 consonant phonemes (see Altmann and Fan, 2008: 151-154 for details).

Some of the above properties are useful guidelines for us to optimize the learning path for second language learners. In this section, some properties of Thai WS are compared with those of other languages in order to see why Thai scripts are more difficult to learn than other phonogramic languages.

In terms of graphemes, Thai has a larger G-inventory size than Korean hangul and Japanese kana, than English letters. However, its G-inventory size is certainly much smaller than those of logograms such as Chinese characters or Japanese kanji.

Among the properties, grapheme load and phonemic load are the measures of how many graphemes a letter can represent and of how many phonemes a letter/grapheme can represent. However, to estimate the burden of learning L2WS, grapheme load and phonemic load are not sufficient. Phonological transparency is another parameter to examine the difficulty of a WS. The English WS is much less transparent than the Japanese kana. For example, "a" in English can represent /aː/ in "father", /æ/ in "bat", /ɔː/ in "ball", etc. and not all of them can be predicted, but Japanese kana is almost totally transparent because it is almost a one-to-one mapping system, with only a small number of exceptions (Cook and Bassetti, 2005: 7-10). With this concept, the grapheme-phoneme transparency in Thai should be quite similar to Japanese kana and Korean hangul, but much more predicable than English and the logogramic systems including Chinese character and Japanese kanji. Although Thai graphemes are as predictable as hangul and kana, the rules predicting the phonemes from the letters are far more complicated than the two. As you will see in section 3.3, not all diphthongs are direct combinations of the corresponding monophthongs and glides. Moreover, to predict tones accurately, the tone marks, open or closed syllables, types of initial consonants ("high class", "mid class" or "low class"), types of final consonants (sonorants or obstruents) and vowel lengths (long or short) all play roles as the phonic rule conditions.

3   The Writing System of Thai Language

To understand the difficulties learners face in learning Thai WS, a brief introduction will be given in this section. Thai WS consists of 44 letters for consonants, 18 symbols for vowels and 4 tone marks (For more comprehensive descriptions, see Diller, 1996).

Obviously, Thai WS have three main properties that make learning it a more difficult task than learning kana or hangul:

(1) The G-inventory size is large and many homophones are in the consonant system;
(2) Complicated phonic rules are required to predict the actual tone value of each syllable.
(3) The forms of vowels vary under different phonological conditions and not all diphthongs are direct combinations smaller units.

3.1 Inventory Size and Homophony

Due to phonological changes in history, some consonants which were distinctive in the past became homophones in the modern Thai (Smalley, 1994: 194-195; Diller, 1996). In many cases, the original written forms are still in use nowadays despite the sound changes.

	1	2	3	4	5	6	7	8	9	10
	M	M	H	H	L	L	L	L	L	L
Vel	ก	ข	ฃ	ค	ฅ	ฆ	ง			
	k	kʰ	kʰ	kʰ	kʰ	kʰ	ŋ			
Pal	จ	ฉ	ศ	ช	ซ	ฌ	ญ	ย		
	tɕ	tɕʰ	s	tɕʰ	s	tɕʰ	j	j		
Ret	ฏ	ฎ	ฐ	ษ	ฑ		ฒ	ณ	ร	ฬ
	d	t	tʰ	s	tʰ		tʰ	n	r	l
Den	ด	ต	ถ	ส	ท		ธ	น		ล
	d	t	tʰ	s	tʰ		tʰ	n		l
Lab	บ	ป	ผ	ฝ	พ	ฟ	ภ	ม	ว	
	b	p	pʰ	f	pʰ	f	pʰ	m	w	
Glo		อ		ห		ฮ				
		ʔ		h		h				

Figure 1: Consonant Letters in Thai[1]

In the consonant chart (figure 1), the IPA below each script is its actual phoneme at syllable initial position in modern Thai, but the columns represent manners of articulation and the features of aspiration and voicing in historical sense:

- Columns 1&2: Unaspirated, unvoiced obstruents
- Columns 3&4: Aspirated, unvoiced obstruents
- Columns 5&6: Unaspirated, voiced obstruents
- Column 7: Aspirated, voiced obstruents
- Column 8: Nasals
- Column 9: Approximants (and Trills)
- Column 10: Lateral approximants

It is important to highlight that, in modern Thai, consonants in Columns 3, 5 and 7 all changed to aspirated, unvoiced stops or affricates as shown in IPA, although the columns had aspiration and/or voicing contrasts before. Consonants (fricatives in the past) in Column 6 were also devoiced.

On the other hand, in the consonant chart, different places of articulation: vel(ar), pal(atal), ret(roflex), den(tal), lab(ial) and glo(ttal) are classified into six rows. It can be easily observed that the rows 'ret' and 'den' had merged completely. All in all, a lot of homophones arise from the mergers in the dimensions of manners, places, aspiration and/or voicing. Homophony is one of the sources that make Thai WS complicated.

3.2 Complicated Phonic Rules for Tones

In figure 1, the consonants are divided into three groups. They are traditionally named "Mid class", "High class" or "Low class", based on the voicing and aspiration properties of the consonants:

- Mid class (M, hereafter) consonants include those which were historically unaspirated, unvoiced obstruents (i.e. Column 1 & 2). Consonants in Column 1 are unaspirated, <u>voiced</u> in modern Thai, but it was sprung from Column 2;
- High class (H, hereafter) consonants include aspirated, unvoiced obstruents (i.e. Columns 3&4);
- Low class (L, hereafter) consonants include all voiced consonants (Columns 5-10) in the past inventory, although Columns 5-7 are currently unvoiced in modern Thai.

One of the major functions of the M-H-L classification is to be part of the conditions of the phonics rules for tones. It is not surprising for voicing/aspiration correlating with tones as many similar phenomena can be found in sound changes of various Chinese dialects and other languages.

Thai language has five tones. They are called "mid-level tone", "low tone (tone 1)", "falling tone (tone 2)", "high tone (tone 3)" and "rising tone (tone 4)". The tone values are 33, 21, 41, 45, 14

---

[1] It is controversial whether อ should be transcribed as a glottal stop, as glottal stop seems not to be contrastive with zero (See Noss 1964; Harris, 2001 for details).

respectively. (Remark: The "high", "mid" and "low" in tones and H, M, L consonants are totally two different concepts.)

Thai WS has four symbols to mark tones. They are put above the initial consonant (if a vowel occupies the place, the mark is put above the vowel). The phonic rules for tones are listed in table 1. Among the three types of consonants, the rules for M consonants are the easiest to remember: ่, ้, ๊ and ๋ are tones 1, 2, 3 and 4 respectively; "live" syllables without tone mark is tone M; dead syllables without tone mark is tone 1. The rules for H and L consonants are more complicated. (see table 1).

	"Live" Syl. w/o tone mark	"Dead" Syl. w/o tone mark	่	้	๊	๋
Mid	M	1	1	2	3	4
Low	M	2/3	2	3		
High	4	1	1	2		

Table 1: Phonic Rule for Thai Tones[2]

In Open Syllable | In Closed Syllable

Figure 2: Combinations Representing Vowels

---

[2] A "live" syllable is one (1) with only a long vowel as rime or (2) with a nasal consonant as final, while a "dead" syllable is one (1) with only a short vowel as rime or (2) with stop consonant as final.
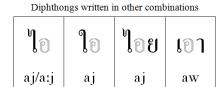
3.3 Inconsistency in Written Forms of Vowels and Diphthongs

In figure 2, vowels (in black) can be represented by a combination of one to three components (อ in grey is an initial consonant; น in grey is a final consonant). For long vowels (upper matrixes of figure 2), patterns in open and closed syllables are almost identical except /ɯː/, which requires an extra อ in an open syllable. However, cases of short vowels (lower matrixes) are rather inconsistent. Non-high short vowels are completely different between open (lower, left) and closed (lower, right) syllables.

Besides, some diphthongs are irregular (figure 3), although the majority of larger units are combined with the smaller units:

(1) diphthong = monophthong + glide;
(2) triphthong = /-a/ diphthong + glide.

Figure 3: Irregular Combinations

4 A Usage-based Approach in Design

As Thai has a relatively large inventory size, a good learning sequence would help improving the learning efficiency. In this section, we intend to look for the optimized learning sequence by considering the usage frequencies of the consonant letters.

4.1 Measuring Usage Frequencies

In our research, the entries of the 44 Thai consonant letters in a Thai-English online dictionary (http://www.thai-language.com) were counted. Based on the numbers of entries, the consonant letters were reordered from most to least entries. Figure 4 shows the numbers of entries (in logarithmic scale, base 10), and a significant drop can be detected after the 31st letter. It is obvious

that the numbers of entries in the dictionary count only the words with the letters as word initial. However, the data are still sufficient to estimate the relative usage frequencies, in order to reveal the "usefulness" of the letters for the learning material design.

4.2 Grouping and Reordering Consonants

In order to look for a desirable learning sequence, the phonological nature of the letters was examined. Learning high frequency letters in the early stage will let learners utilize the WS as early as possible. However, simply adopting the most-to-least sequence as the learning sequence would make the learning process unordered and not systematic. It would not be easy for learners to remember the letters. Therefore, besides the usage frequencies, forming phonologically-correlated subsets are also important criteria for a good learning sequence.

According to the finding in figure 4, the consonant letters can be primarily divided into two subsets based to their usage frequencies: high on the left and low on the right in figure 5. In order to seek for phonologically-correlated subsets, the information about the phonological origin of the letters in figure 1 is also extracted and listed in two tables. The first observation from the left table is that all phonemes (including the zero initial, อ) can all be found. Besides the fricatives (/f/, /s/ and /h/) and the aspirated obstruents (/k^h/, /p^h/, /t^h/ and /ts^h/), all other letters are not repeated in the left table. The repeated ones are only either H or L consonants. Comparing the repeated letters as shown in table 2, it is obvious that letters in high usage frequencies all come from Columns 3-6 in figure 1 (H3, H4, L5 and L6). Therefore, it is quite reasonable to put the three letters: ก(24th), ศ(28th) and ธ(29th) to later stages of the learning sequence.

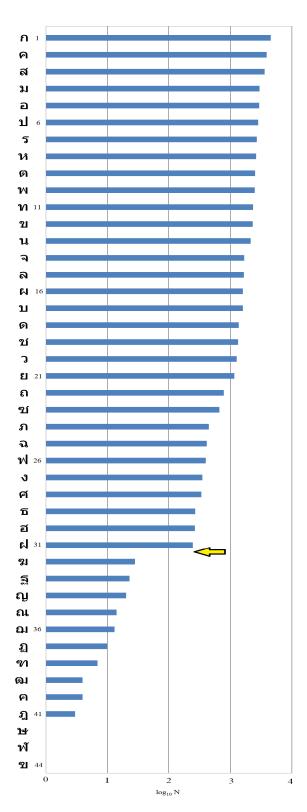| Consonant	Rank / MOA		
k^h	2/L5 > 12/H3		
p^h	10/L5 > 16/H3	> 24/L7	
t^h	11/L5 > 22/H3	> 29/L7	
ts^h	14/L5 > 25/H3		
f	26/L6 > 31/H4		
s	3/H4 > 23/L6	> 28/H4	
h	8/H4 > 30/L6		

Table 2: Ranking of Corresponding Consonants

Figure 4: No. of Entries for Each Thai Consonant in a Thai-English Online Dictionary

Cons.	IPA	POA	MOA	Rank
ก	k	1_VEL	M2	1
ค	kʰ	1_VEL	L5	2
ส	s	4_DEN	H4	3
ม	m	5_LAB	L8	4
อ	ʔ	6_GLO	M2	5
ป	p	5_LAB	M2	6
ร	r	3_RET	L9	7
ห	h	6_GLO	H4	8
ต	t	4_DEN	M2	9
พ	pʰ	5_LAB	L5	10
ท	tʰ	4_DEN	L5	11
ข	kʰ	1_VEL	H3	12
น	n	4_DEN	L8	13
จ	ts	2_PAL	M2	14
ล	l	4_DEN	L10	15
ผ	pʰ	5_LAB	H3	16
บ	b	5_LAB	M1	17
ด	d	4_DEN	M1	18
ช	tsʰ	2_PAL	L5	19
ว	w	1_VEL	L9	20
ย	j	2_PAL	L9	21
ถ	tʰ	4_DEN	H3	22
ซ	s	2_PAL	L6	23
ภ	pʰ	5_LAB	L7	24
ฉ	tsʰ	2_PAL	H3	25
ฟ	f	5_LAB	L6	26
ง	ŋ	1_VEL	L8	27
ศ	s	2_PAL	H4	28
ธ	tʰ	4_DEN	L7	29
ฮ	h	6_GLO	L	30
ฝ	f	5_LAB	H4	31

Cons.	IPA	POA	MOA	Rank
ฆ	kʰ	1_VEL	L7	32
ฐ	tʰ	3_RET	H3	33
ญ	j	2_PAL	L8	34
ณ	n	3_RET	L8	35
ฌ	tsʰ	2_PAL	L7	36
ฎ	t	3_RET	M2	37
ฑ	tʰ	3_RET	L5	38
ฒ	tʰ	3_RET	L7	39
ฅ	kʰ	1_VEL	L6	40
ฏ	d	3_RET	M1	41
ษ	s	3_RET	H4	42
ฬ	l	3_RET	L 0	43
ฃ	kʰ	1_VEL	H4	44

Cons. : Consonant letter
IPA: IPA transcription of letters in modern Thai
POA: Place of Articulation that the letter was in the past
MOA: Manner of Articulation that the letter was in the past
Rank: Ranking of letters from highest to lowest frequency

Figure 5: Usage Frequencies of Thai Consonants

After removing ก, ศ and ธ, we still have to choose between the two sets of homophones in the first two columns of table 2, in order to form a complete set of 21 letters, containing all contrastive phonemes. There are three reasonable choices:

(1) The set with highest frequencies regardless of their phonological properties:
2/L5, 10/L5, 11/L5, 14/L5, 26/L6, 3/H4, 8/H4;
(2) The set with all L consonants:
2/L5, 10/L5, 11/L5, 14/L5, 26/L6, 23/L6, 30/L6;
(3) The set with all H consonants:
12/H3, 16/H3, 22/H3, 25/H3, 31/H4, 3/H4, 8/H4.

As mentioned before, choice (1) may not be proper because mixing H and L consonants may cause confusion, as syllables have different tones. Between two choices with consistent consonant types, choice (2) sounds more preferable than choice (3), not only because it has a higher average ranking, but also other non-repeated letters in the left table of figure 5 belong to either M or L consonants. If we opt for choice (2), learners are only required to deal with the phonic rules for M and L consonants at the beginning stage and leave H consonants to later stages. Besides, one more advantage for choice (2) is that loanwords from English are mainly written in L and/or M consonants, although some H consonants are also occasionally used in loanwords. It is advantageous for most foreign learners to recognize the letters and gasp the first bunch of vocabularies rapidly through these déjà vu.

After deciding the earlier stages, what remains are ก, ศ and ธ and the 13 letters in the right table of figure 5. The two /kʰ/s ranked 40th and 44th are officially replaced by /kʰ/s ranked 2nd and 12th respectively and are no longer in use in modern Thai. They are certainly the last step, if the learners insist to learn. A general observation from the remaining 14 is that, except /s/ ranked 28th and /j/ ranked 34th, all others belong to either "L7" or "retroflex". Although we now know the fact that rare letters are aspirated, voiced obstruents and/or retroflex consonants, it is not necessary for the learners to know them unless they are interested in the historical linguistics of Thai. As the usage frequencies are very low compared to the 28 consonant letters in the left table of figure 5, learners even do not need to border whether they

are M, H or L. It seems more reasonable to memorize the entire word or phrase with the pronunciation as a whole, while learners may or may not occasionally come across a couple of them in the whole life. This process resembles people learning Chinese characters or other logogramic languages.

However, among the 14 low frequency letters, some of them are actually more useful than others, regardless of their phonological properties. Some letters may not be used in many words but they do appear in a few high frequency ones. For instance, they can be found in the following words:

(1) /k^h/ 32nd in "kill", "cloud", etc.
(2) /t^h/ 29nd in "she", "flag", etc.
(3) /p^h/ 24th in "language", "greedy", etc.
(4) /n/ 35th in "you", "Mr/Ms", "father", "mother", etc.
(5) /j/ 34th can be found in "big", "Japan", etc.
(6) /s/ 28th "country", etc.
(7) /s/ 42nd in "sorry", "language", etc.

Learners can learn these letters after they are familiarized with the 28 letters in the left table of figure 5. At this stage, taking the /n/ 35th as an example, learners can be told, "ณ is another form of น. It is only used in some specific words and the /n/ in /k^{h}un/ "you" is one of them, using ณ."

To summarize, table 3 lists our proposed learning sequence. Learning subsets 1 and 2 is similar to learning a WS of a phonogramic language with a small inventory size, while learning subsets 3, 4 and 5 is similar to learning a WS of a logogramic language.

Subset 1 (21 letters)	No homophones. All are M or L consonants	กจดตบปอ คชทพ/ซฟฮ /งนม/วยรล
Subset 2 (7 letters)	All are H consonants	ขฉถผ/สฝห
Subset 3 (7 letters)	Other consonants in high frequency words	ฆธภณญศษ
Subset 4 (7 letters)	Other consonants in low frequency words	ฎฏฐฑฒฌฑ
Subset 5 (2 letters)	Not used in modern Thai	ฅฃ

Table 3: Proposed Learning Sequence

4.3 Learning Sequence of Vowels

To enhance the learning efficiency, the learning sequence of vowels should also be reordered.

Although it can be based on the relative usage frequencies in a similar way as the consonants, simply doing the same thing for vowels can be problematic in the considerations of pedagogy. There are two more important aspects we have to consider further.

First, besides the usage frequencies of the symbols, the simplicity of words' meanings is other important concerns. It is much easier for beginners to pick up basic words with simpler meanings at the beginning. Words with high usage frequency consonants can be easily chosen to form a list of simpler basic words. The same can be done for words with high usage frequency vowels. However, if both criteria of consonants and vowels are applied at the same time, many simple words are eliminated from the list, because words with both high frequency consonant(s) and vowel(s) are not necessarily simple in meaning. One way to deal with the problem is to take usage frequencies of consonants as the primary selection criterion for the learning sequence and those of vowels as a criterion in lower priority.

Second, the complexity of the symbols is another concern. Thai consonants and vowels are both complicated for learners, but their complications are different in nature. Contrast to vowels, the phoneme of a consonant can be represented by different homophonic graphemes and the grapheme-phoneme relationship is unpredictable phonologically. On the other hand, the vowel system is complicated because one single vowel can be represented by a combination of up to three components (e.g. /ɤ/ and /ɔ/ in open syllables are combinations of 3 components, see figure 2). Besides, a component can be used in different vowels (e.g. "เ" can be found in /e/, /eː/, /ɤ/ and /ɤː/, see figure 2). However, the grapheme-phoneme relationship of vowels is rather transparent and predictable.

With the considerations of the complexity of vowel formation, most traditional textbooks tend to introduce, first, monophthongs (long and short), then, diphthongs and finally, triphthongs. We have no objection to taking complexity as a criterion in deciding learning sequence, but we do believe that the relative usage frequency of vowels and the semantic simplicity of the words taught in beginner level should also be important criteria for ordering learning sequence. We now propose a nine-step

learning sequence as follows in order to enhance the learning efficiency:

(1) All long monophthongs – Long vowels are almost same in the open and closed syllables, so the learners can acquire them without having a prior knowledge of open and closed syllables.

(2) ไ and ใ (/aj/ or /a:j/): - They are used in many high frequency grammatical words such as wh-question words, negation, modal verbs; They can only be used in open syllable.

(3) Short monophthongs /i/ and /a/ in both open and closed syllables – They are commonly found in many high frequency words; There is only one component in each of these monophthongs; They are basic building blocks of several other short vowels or diphthongs.

(4) Short monophthongs /u/ and /ɯ/ in both open and closed syllables – They are found in a few high frequency words. Although there is only symbol in each of these monophthongs, they are not used as part of other monophthongs.

(5) Other short monophthongs in closed syllables – They are quite common.

(6) Diphthongs in irregular combinations other than the two symbols in (2) – Some words are quite common but the combinations may be harder to remember due to the irregularity.

(7) Diphthongs in regular combinations – They are combinations of (1)/(3)/(4)/(5) and a glide consonant.

(8) Triphthongs in regular combinations – They are combinations of (6) and a glide consonant.

(9) Other short monophthongs in open syllables – They are rarely used in high usage frequency words.

The above list shows a reasonable learning sequence of vowels based on the three parameters:
- relative usage frequency of vowels;
- simplicity in formation of vowels; and
- simplicity in words' meanings.

Vowels with higher usage frequency, fewer components and more simple meaning words are arranged in earlier stages. As subjective judgments have been made on whether the words are appropriate for beginners, some textbook writers or course developers may find a slightly different sequence more suitable when the course is taught in different cultures or to people in various ages.

Learning basic daily life vocabularies together with the letters are very important for reinforcing the memory of orthography. Memorizing the letters in the above order alone without learning the vocabularies are not as effective. The consonant and vowel letters should be learnt in parallel so that vocabularies written in Thai scripts can also be learnt at the same time. For example, at the beginning, some consonants in subset 1 can form simple nouns with long vowels. Then, some simple nouns and verbs formed by more consonants in subset 1 and irregular /aj/s in the following lesson. Learners could "bootstrap", or initiate the acquiring process, more easily, with subsets 1 and 2 plus different vowels or diphthongs until they have learnt a couple hundreds of words. Then, they could proceed to subset 3. They could simply ignore subsets 4 and 5 until they become intermediate or advanced learners.

4.4 Tone Marks and their Phonic Rules

After settling down the consonants and vowels, the next question will be the tone marks and the phonic rules for tones. The L2 learners are recommended to learn the values of the five tones at the very beginning, even earlier than the consonant subset 1.

Instead of learning the phonic rules for tones consciously, the learners are recommended to learn only the relationship between the consonant-vowel spelling and the pronunciation of the whole word. However, the learners should be told that they must remember the tone mark as part of the spelling by heart without linking to its actual pronunciation of tone: "Just treat the tone mark of each syllable as a kind of decoration on one hand, and remember the tone of each syllable on the other hand." The process is similar to the learning of Chinese characters or Japanese kanji. When the learners are building up their lexicon of Thai gradually, the tone mark-toneme correspondence in their minds will emerge in a subconscious way without paying much effort on "calculating" tones syllable by syllable. Since each syllable lasts for several hundred milliseconds in a normal speech, it does not make sense to spend several seconds to calculate the tone of each syllable before reading the text aloud. Memorizing tones and the tone marks separately is also what native speakers do normally. They acquire the tones in their speech by heart from their L1 environment. Later on, when

they learn to write, they remember the spellings by heart, although some native speakers may learn a reverse version of table 1 to predict the spellings from the sound of tone.

Some speakers tend to remember written and spoken forms separately when they learn a second language. In a recent study about Cantonese speakers learning Korean as a second language, learners could pronounce hangul spellings more accurately for words than non-words. This indicates that remembering the spoken forms by heart instead of predicting pronunciations from spellings is the main strategy the speakers use (Au and Cheung, 2014). This subconscious strategy is also applicable to people who are learning other more complicated WSs such as Thai.

5 Consonant names are burden to non-native adults

Another advantage of learning only a subset without homophones is to get rid of the names of the Thai consonant letters. Similar to English's "A for apple", "B for boy", "C for cat", Thai children learn the name of each Thai consonant letter, which composes of one consonant-vowel syllable (e.g. /t^h/ + /ɔ:/) and one noun that contains this particular consonant letter:

- ฑ /t^{h}ɔ:^{33} mon^{33} t^{h}o:^{33}/ (Ramayana character)
- ฒ /t^{h}ɔ:^{33} p^{h}u:^{51} t^{h}aw^{51}/ (old man)
- ท /t^{h}ɔ:^{33} t^{h}a^{45} ha:n^{14}/(soldier)
- ธ /t^{h}ɔ:^{33} t^{h}oŋ^{33}/ (flag)

The nouns help in disambiguating homophonic consonant letters. In the above example, all four letters called /t^{h}ɔ:^{33}/ become distinguishable orally.

As their native language, Thai children have already learnt a certain amount of spoken forms of Thai words at home before learning the written forms in school. They can remember all letters more easily with the letters' names because they have probably acquired some, if not all, sound-meaning relationship of the words such as "old man", "soldier" and "flag". On the contrary, as the letters and the nouns are both new to the foreign learners, the names are in fact extra burdens to the L2 learners. In our proposal, when only one of the four is learnt in the subset 1, learners can simply call ท as /t^{h}ɔ:^{33}/. In subset 2, although there is another /t^h/, as the tone is different, ฒ can be

called /t^{h}ɔ:^{14}/ without mentioning its name. In this case, the learning of names can be postponed to later stages of the study after learners gain enough vocabularies and knowledge of Thai culture in order to know the disambiguating names of the letters.

6 Applying to learning of other languages

The concept we propose is also applicable to other languages, although their WSs may not be as complicated as Thai scripts.

Similar to table 3 for Thai WS, Japanese WS can be divided into three subsets:

- subset 1 is hiragara (e.g. あ、い), the basic set of graphemes;
- subset 2 is katakana（e.g.ア、イ), another set of graphemes used mainly for loanwords; and
- subset 3 is kanji, adopted Chinese characters.

To enhance the learning efficiency, instead of teaching both sets of kana (hiragana and katakana) at the start of a beginner course, some teachers may teach hiragana first and postpone the teaching of katakana until learners become more familiar with the hiragana through acquiring large amount of vocabularies. This is a way to reduce the interference between two sets of kana.

In the case of the WS of Korean language, although each phoneme is represented by only one consonant letter in hangul, letters with similar pronunciations and shapes can also be learnt in different stages:

- subset 1 includes ㅂ/p/, ㄷ/t/, ㄱ/k/, ㅈ/tɕ/, ㅅ/s/, ㅎ/h/, ㅁ/m/, ㄴ/n/, ㅇ/ŋ/ and ㄹ/l/;
- subset 2 is the tense consonants ㅃ/p̈/, ㄸ/ẗ/, ㄲ/k̈/, ㅉ/tɕ̈/ and ㅆ/s̈/, which are formed by doubling the first five consonant letters in subset 1; and
- subset 3 is the aspirated consonants ㅍ/p^h/, ㅌ/t^h/, ㅋ/k^h/ and ㅊ/tɕ^h/, which have similar shapes as the first four consonant letters in subset 1.
(Remark: The Korean letters are only transcribed phonemically. Phonetic and allophonic details are not included here)

Similarly, learning efficiency can be improved by focusing on only a subset of graphemes at the beginning. Consonants in simpler shapes (subset 1) can be learnt first, with considerable amount of simple vocabularies that are strictly formed by the first ten consonants. After having enough exposure to subset 1, learners can start learning the five tense consonants in subset 2. The phonetic contrasts of the obstruent consonants in the two subsets should also be introduced in this stage. After the learners are familiarized with the 15 consonants under the context of simple words and sentences formed by them, they can proceed to subset 3 and learn the concept of aspiration.

Although all 19 consonant graphemes are phonemically contrastive in Korean language, the three groups of obstruent consonants may not be perceptually or consciously distinctive to many non-native speakers (Au and Cheung, 2014). Therefore, allocating these "pseudo"- homophonic obstruent consonants (e.g. ㅂ, ㅃ and ㅍ) into three different subsets resembles the case of Thai shown in table 3.

In terms of usage frequencies and relatively complexity of grapheme shapes, the tense and aspirated consonants are lower, compared to the other ten consonants. Thus, it should not be difficult for course developers to gather sufficient basic sentence structures and everyday vocabularies for compiling the first few lessons without using subsets 2 and 3.

This arrangement will not only reduce the memory burden of the second language learners at the beginning stages, but also highlight the phonetic contrast among the three groups of obstruent consonants.

7 Conclusion

The present paper tries to demonstrate that the efficiency of a second language WS with a relatively large letter/grapheme inventory size (such as Thai) could be enhanced by separating the large inventory into (at least) two subsets, based on the usage frequencies. The high frequency subset(s) can be learnt early through knowing the grapheme-phoneme correspondence, while the low frequency subset(s) can be learnt in later stages in the way similar to learning a logogramic language (such as Chinese characters), by remembering the spelling and the sound individually.

Theoretically, the learning sequences proposed would improve efficiency and effectiveness of acquiring a new second language writing system, although the actual improvement needs to be substantiated by future researches on students' acceptability and performance.

References

Gabriel Altmann and Fengxiang Fan. 2008. *Analyses of Script: Properties of Characters and Writing Systems*. Mouton de Gruyter.

Ching-Pong Au and Yuk-Man Cheung. 2014. Phonological Tendencies in Korean Spoken by Hong Kong Cantonese Learners. *Proceedings of 5th ICPM*, Gwangju, South Korea, 263-266.

Benedetta Bassetti. 2006. Learning Second Language Writing systems. Centre for Languages, Linguistics and Area Studies, Guide to Good Practice. 29 Nov., 2006. https://www.llas.ac.uk/resources/gpg/2662

Vivian. J. Cook and Benedetta Bassetti. 2005. *Second Language Writing Systems*. Clevedon, UK: Multilingual Matters.

Anthony Diller 1996. Thai and Lao Writing. In P. T. Daniels and W. Bright (eds.). *The World's Writing Systems*, 457-466.

Jimmy G. Harris. 2001. States of the Glottis of Thai Voiceless Stops and Affricates. In M. R. K. Tingsabadh and A. S. Abramson (eds.), Essays in Tai Linguistics. Bangkok: Chulalongkorn University Press.

Richard B. Noss. 1964. *Thai Reference Grammar*. Washington: Foreign Service Institute.

William A. Smalley. 1994. *Linguistic Diversity and National Unity: Language Ecology in Thailand*. Chicago and London: The University of Chicago Press.

Dictionary in http://www.thai-language.com