



# Linguistic Variation in Chinese Characters: Knowledge Essential for Teachers

**Ho-Cheong Lam**

(hclam@eduhk.hk)

The Education University of Hong Kong, SAR Government, China

---

## Abstract

“Hong Kong and Taiwan adopt the same standard of traditional Chinese characters.” “Characters are pronounced only in those ways specified in the dictionary.” These are some simplistic notions that teachers commonly hold about the standard of correctness of characters. But if we closely examine the actual usage of characters in everyday practice, there exists subtle variants of the characters. This paper reports on the results of our project to identify linguistic variation in characters for raising teachers’ awareness in a teacher education course. Our project focused on those characters in the curriculum and, by gathering multiple sources of data, endeavored to find out what categories of variation existed in these characters. Two such categories of variation, namely, variation in written form and variation in stroke order, are explained in full detail. These results have practical implications to teaching since problems often arise when children adopt a variant of characters different from that of the teachers. Knowledge about possible variants of characters is thus necessary for teachers to appropriately respond to the children. Multiple standards, which teachers should interpret as helpful references rather than dogmatic rules, are recommended. We conclude the paper with suggestions for future research on practical linguistic knowledge for teachers.

---

## 1 Introduction and background

Teachers have the responsibility to be sure about the correctness of the content they teach. Because of this, teachers who teach the Chinese language in Hong Kong often hold to the belief that there exists an absolute standard of the language, on which they have to rely. For instance, young learners are taught to adopt solely the dictionary’s prescribed pronunciation of a given character in Cantonese (i.e. the variety of Chinese that is widely spoken by about 90% of the Hong Kong population), and write characters according to the standard forms provided by the Education Bureau (EDB) of the Hong Kong government. And so on. These reflect a simplistic understanding of the existence of a single standard of characters, which is commonly held in the minds of teachers. But, if we closely examine how characters are used in daily lives, we will find that despite uniformity on the whole, subtle linguistic variation actually exists in the characters.

Lack of attention of teachers to the existence of linguistic variation has led to the occurrence of a number of problems. For instance, when parents teach children to write characters at home, they follow what they were taught in their own school days. Later on, when they receive feedbacks from the teachers, the written forms of their children are marked as wrong, which have to be corrected. This is because what the children have written does not fully conform to the current standard of EDB. As another example, when teachers use computers to prepare teaching materials, the printed forms of the characters are found not to align with the standard written forms in Hong Kong, but rather with that in Taiwan. The computer fonts of traditional characters that the teachers use are

product of Taiwan. Furthermore, there are occasions in which teachers, despite being native Cantonese speakers, are asked by the principals to pronounce a character in a way they have never heard of before. This is due to the difference in the pronunciations specified in the dictionary and those actually used in everyday lives. Against the background of these problems, we believe a project to identify linguistic variation in characters for raising the awareness of teachers is timely and of crucial importance.

The rest of this paper is organized as follows: Next section will discuss the notion of variation in Linguistics. Then we will discuss the background, objectives, and method used in our project, followed by a thorough discussion of two categories of variation in characters we identified, which is the main thrust of this paper. After that, recommendation of the adoption of multiple standards for teachers will be made. Finally, this paper will be concluded with future research direction on practical linguistic knowledge.

## 2 Linguistic variation

Before going into our project, we would like to clarify the notion of variation used in this paper, which we have borrowed from the discipline of Linguistics. This is different from that of variation theory in the phenomenography research tradition, which we, the authors, often use as the theoretical and methodological framework in our research. The former is a concept in Linguistics (concerned with what kind of variation exists in a language); while the latter is a theory of learning (concerned with what makes learning happen).

Most existing works in the literature on linguistic variation (or language variation) have focused on analyzing the sound patterns of languages. For example, Radford, Atkinson, Britain, Clahsen, and Spencer (2009, p. 47) defined phonological variation as ‘it is the existence within the speech of a single community of more than one possible realization (or variant) of a particular sound.’ In other words, to express a certain meaning, speakers of a speech community may use two or more variant sounds without a change in meaning. Actually, ‘everyday speech contains a great deal of phonetic variation that speakers [simply] pay little or no attention to. (Czaykowska-Higgins & Dobrovolsky, 2010, p. 65)’ Despite its common existence, variation in languages has limits. ‘There is considerable variation in the speech of any one individual, but there are also definite bounds to that variation: no individual is free to do just exactly what he or she pleases so far as language is concerned. (Wardhaugh & Fuller, 2014, p. 6)’ This means the alternation in sounds is not drastic and occurs only in limited ways.

Methodologically, to investigate variation in languages, linguists generally adopt a descriptive rather than a prescriptive approach. They try to describe the variation as it is in the actual use of the language in a speech community rather than to lay down rules to define the ‘correct’ use of the language (i.e. the case of language standardization; See Garvin, 1993; Milroy & Milroy, 1999). ‘First, and most important, linguistics is *descriptive*, not prescriptive. A linguist is interested in what *is* said, not what he thinks *ought* to be said. He describes language in all its aspects, but does not prescribe rules of “correctness.” [italics in the original] (Aitchison, 1999, p. 4)’

Since existing research has mainly been developed from the analysis of phonetic languages (e.g. English), little has been discussed about variation in Chinese, which is a non-phonetic language. Chinese consists of thousands of characters. Variation can occur in the written forms of these characters themselves as well as in the ways people produce the written forms such as the orders in which they write each of the strokes in the written forms (c.f. allograph instead of allophone). Although there is the possibility to examine variation in these novel aspects of the language, existing efforts have mainly been devoted merely to the standardization of characters, where a prescriptive, not a descriptive, approach was adopted.

For example, 陳越 (1965) proposed several principles that prescribed what should be the ‘correct’ written forms of characters (e.g., the printed forms, whenever possible, should align with the handwritten ones). 費錦昌 (1997) and 傅永和 (2000) discussed what aspects of characters had to be standardized (e.g. the names, types, and orders of the strokes in the written forms). Other literature

concerned the standard stipulated by the government. For example, 王寧 (2013) explained the principles the Ministry of Education used to define the standard written forms of characters in mainland China. 高更生 (2002) and 費錦昌 (2000) took a historical perspective and reviewed the past development of the standard written forms of characters of the mainland China government. Besides, while the above research examined the government standard of one place; other research adopted a comparative approach. 胡雙寶 (1993) compared the differences in the standard written forms between Taiwan and mainland China. 黃靜吟 (2005) made a comparison of the standard stroke orders between the same two places.

Apart from the context of standardization, variation in Chinese has scantily been explored. The paucity of existing research, with the adoption of a descriptive approach, was at most piecemeal and addressed variation in only limited aspects of characters such as Cantonese pronunciation (Bauer & Benedict, 1997; 李新魁、黃家教、施其生、麥耘、陳定方, 1995). To the best of our knowledge, a comprehensive and coherent picture of the discussion of a wide range of categories of variation in characters is not available. Let alone the availability of such a treatment specifically for the professional development of teachers.

### 3 The present project

By way of background, we are teacher educators with a passion to facilitate kindergarten teachers to improve their teaching of Chinese. In this paper, a case is made for the importance of enhancing kindergarten teachers' awareness of linguistic variation in characters. To do this, we designed and implemented a teacher education course. As discussed earlier, a coherent picture of variation in characters was lacking in the literature. We thus had to identify for ourselves the various categories of existing variation, which were for use as part of the content in our course. This project has eventually developed into a book titled *Discerning the Differences in Chinese Characters 大同小異: 辨識漢字不同處* (Lam, in press). The ideas presented in this paper are brief sketches of more thorough explanations included in the book.

Although this paper draws on the case that kindergarten teachers in Hong Kong teach children Chinese as the first language, the same issue on linguistic variation and standard of characters should also apply to the case of teaching Chinese as a second and foreign language.

#### 3.1 Research questions

We began with the premise that teachers know all of the characters they teach in the curriculum. Since there is no such recommended list of characters to be taught at the kindergarten level, there is a great deal of variation in the numbers of characters taught across kindergartens. Moreover, the characters that are taught come mostly from the list of 3,171 characters that is meant for the primary school curriculum. Similarly, it is this recommended list of characters for the primary school curriculum that was focused on in this project. Based on this set of characters, we aimed to find out what linguistic variation existed in these characters. In contrast, our aim was not to exemplify linguistic concepts in typical textbooks with characters rarely used in daily lives. Our general question was: What type of linguistic knowledge about these characters do teachers need in order for them to teach children characters better? It was in this context that we identified the categories of variation and sought to answer the following specific questions: What are the different categories of linguistic variation in characters? In what ways do characters vary in each of these categories? What are the different situations in which such variation occurs?

#### 3.2 Data collection and analysis

To answer the above research questions, we gathered multiple sources of data for analysis. For instance, tests were conducted with people who needed to teach children to write characters. They were asked to write the same list of characters. Difference in the written forms they produced as

well as difference in the stroke orders they used were analyzed. In addition, we also looked up dictionaries and contrasted government documents that prescribed the standards in Hong Kong (課程發展處中國語文教育組, 2007), Taiwan (教育部國語推行委員會, 2008), and mainland China (國家語言文字工作委員會, 1997). Furthermore, we also observed how characters were actually used in our daily lives (e.g. environmental prints).

To analyze the data, we did not fit the variation we identified into some pre-determined categories in mind (e.g. which variants are correct or incorrect according to the dictionary). Rather, we put aside our presuppositions of the ‘correct’ standard and let the different ways characters varied emerged. This practice of analyzing the data was motivated by our phenomenography research tradition, which we believe is consistent with a descriptive approach.

#### 4 Categories of variation in characters

In what follows, we will report on two categories of linguistic variation that we identified. In this paper, variation refers to the different realizations of the same character. In other words, the character denotes the same meaning but its written forms or stroke orders vary in different situations. Since this sort of variation is rarely mentioned in textbooks, even native speakers often do not consciously notice its existence. In their minds, only the ‘correct’ forms or stroke orders exist. Due to space limitations, in the following sections, we will explain two categories of variation, namely, variation in written form and variation in stroke order. Full report of all categories is available in our book.

##### 4.1 Variation in written form

The first category of linguistic variation we identified had to do with the written forms of characters. The same character was written slightly differently in different situations, yielding different variant forms. Normally the variation consisted of only a limited number of variant forms. Otherwise, if the written forms of the same characters were all different, it would have become impossible for people to use the written forms to communicate their ideas.

##### 4.1.1 Across individuals

The first situation in which variation occurred was as follows: When different people wrote the same character, the written forms they produced varied. To fully reveal the possible range of variants in this situation, we conducted a test with a group of 115 participants, all of whom needed to teach children to write characters in their everyday work. The same list of characters was read aloud to them character by character in the form of a word (e.g. ‘The 快 of 快樂 “happy” /faai3lok6/’). They then wrote down the corresponding characters (i.e. 快) on a piece of worksheet. The worksheet consisted of a number of boxes 5 cm x 5 cm each in size. Each box was for one character and had a footnote such as ‘faai3 樂’ to serve as a reminder to the participants. All of the variant forms the participants produced were recorded and then analyzed. We aimed to find out what the subtle differences were in these variant forms. The results are shown in Table 1.

Table 1. Variant forms between individuals

Character	Variant forms and the corresponding number of occurrences				Other forms
快 ‘fast’	快 <sup>HK</sup> 76	快 28	快 10		1
骨 ‘bone’	骨 <sup>HK</sup> 74	骨 25			16
荔 ‘lychee’	荔 <sup>HK</sup> 44	荔 39	荔 11	荔 10	11

Character	Variant forms and the corresponding number of occurrences		Other forms
告 'to tell'	告 90	告 <sup>HK</sup> 25	0
慈 'kind'	慈 113	慈 <sup>HK</sup> 1	1







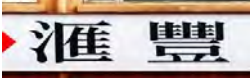
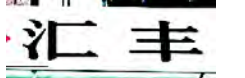
For example, as shown in the table, there were three ways in which the participants wrote the character 快 'fast'. The three variant forms differed only on the left-hand side of the character (i.e. the 'heart' component). In the table, the numbers under each of the variant forms indicate the number of participants who had produced the corresponding variant forms, that is, showing how often each of the variant forms occurred. For instance, 76 (66.1%), 28 (24.3%), and 10 (8.7%) participants produced the three variant forms of the character 快 respectively. Besides, there is a symbol 'HK' added to one of the three variant forms, indicating the standard written form of EDB. For the character 快, EDB had chosen to use the first variant form as the standard. The analysis of characters other than 快 is also tabulated in Table 1.

For most of the characters, EDB adopted the use of the same variant forms as those the majority of our participants produced. In other words, the EDB standard generally agreed with the actual practice of the participants. However, worthy of noting was the case of the character 慈 'kind', for which there were two different variant forms that differed in the component at the top. The majority of the participants (113 out of 115, 98.3%) produced the first variant form; while only 1 participant (0.9%) produced the second. However, it was the second variant form that EDB had adopted as the standard. In this specific case, the EDB standard did not align with the practice of our participants.

#### 4.1.2 Across environmental prints

In addition to examining the different variant forms people produced, we also observed how the written forms of characters varied in environmental prints (i.e. the second situation). Table 2 shows pictures of the different variant forms of the same characters in our observation. We took these pictures in different public places mostly in Hong Kong (only a few in Shenzhen). Children might encounter these environmental prints in their daily lives. Their exposure to the characters was not limited only to the environments in the kindergartens. A total of 1,250 pictures of 81 characters were taken and analyzed.

Table 2. Variant forms in environmental prints

Character	Photos		Variant forms
教 'to teach'		vs. 	教 <sup>HK</sup> vs. 教
角 'corner'		vs. 	角 <sup>HK</sup> vs. 角
滑 'to slip'		vs. 	滑 <sup>HK</sup> vs. 滑
匯 'to gather together'		vs. 	滙 vs. 滙 <sup>HK</sup> vs. 汇 <sup>CN</sup>









One interesting case was the character 匯 'to gather together' of 滙豐 'The Hongkong and Shanghai Banking Corporation Limited'. The bank had chosen to use as their sign the variant form with the three-dot component 'water' on the left (i.e. 滙); while the standard written form of EDB,

which also aligned with the standard in Taiwan, had the three-dot component placed inside the 匚 (i.e. 匯). Historically, the character was formed by combining the two characters 淮 and 匚 together, where the 淮 /waai4/ served to signify the sound of the character /wui6/. Perhaps, that was the reason why the three-dot component was not separated from the 隹 in the standards in Hong Kong and Taiwan. However, the standard written form in mainland China (indicated with a ‘CN’ in the table) had the three-dot component placed on the left, which interestingly aligned once again with the sign of the bank in simplified characters.

#### 4.1.3 Across government standards









Apart from the observation of environmental prints, we also inspected the government documents that specified the standards in Hong Kong, Taiwan, and mainland China (i.e. the third situation). Both Hong Kong and Taiwan use the traditional characters. Consistently, the majority of the standard written forms in the two places were found to be identical. However, there were a number of cases in which the two places had adopted different standard written forms. The character 育 ‘rear’ was a case at issue. For the standard written form in Hong Kong, the writing of the character started with the dot at the top; while, in Taiwan, it started with the folded stroke. As the result of this, the total number of strokes of the character was 8 in Hong Kong and only 7 in Taiwan. See Table 3.

Table 3. Variant forms across standards in Hong Kong and Taiwan

Character	Standard in Hong Kong	Standard in Taiwan
育 ‘rear’	 8 strokes in total. The first stroke was the dot at the top.	 7 strokes in total. The first stroke was the folded stroke at the top.
告 ‘to tell’	 The upper component was 牛 with the vertical stroke passing through the lowest horizontal stroke.	 The vertical stroke of the upper component did not pass through the lowest horizontal stroke.
慈 ‘kind’	 14 strokes in total. The upper component was 卅.	 13 strokes in total. The upper component consisted of two dots and one horizontal stroke.
骨 ‘bone’	 In the lower component, the vertical stroke on the left was straight and the two horizontal strokes were parallel to each other.	 In the lower component, the vertical stroke on the left was slanting and the two horizontal strokes were converging.







Although Hong Kong has adopted the use of traditional characters; while mainland China has adopted simplified characters, there are a substantial portion of the characters, of which the written forms are the same in the two places. Only in a few exceptions, differences were found to exist in the standard written forms of the two places. For example, as shown in Table 4, in Hong Kong, there were two folded strokes on the right of the character 及 ‘to reach’; while in mainland China, the two strokes merged together into one double-folded stroke. The total number of strokes was thus 4 in the former but 3 in the latter.

Table 4. Variant forms across standards in Hong Kong and Mainland China

Character	Standard in Hong Kong	Standard in mainland China
及 'to reach'	 4 strokes in total. There were two folded strokes on the right.	 3 strokes in total. There was one double-folded stroke on the right.
角 'corner'	 The vertical stroke in the middle did not pass through the lowest horizontal stroke.	 The vertical stroke in the middle passed through the lowest horizontal stroke.
花 'flower'	 8 strokes in total. The upper component consisted of two 十 s. The slanting stroke at the bottom right did not go through the vertical bend hook.	 7 strokes in total. The upper component consisted of one horizontal and two vertical strokes. The slanting stroke at the bottom right went through the vertical bend hook.
骨 'bone'	 10 strokes in total. The two strokes in the middle of the upper component pointed towards the right.	 9 strokes in total. The folded stroke in the middle of the upper component pointed towards the left.

In other cases, the written forms of the same characters differed not in the strokes or components but in the spatial organization of the components in the characters. This means the variant forms contained exactly the same components but the overall configurations were different. For example, in Hong Kong, the character 羣 'flock' had its two components 尹 and 羊 arranged in a top-bottom configuration, while in Taiwan and mainland China, the two components were arranged side by side in a left-right configuration. See Table 5, where HK, CN, and TW indicate the standard written forms in Hong Kong, mainland China, and Taiwan respectively.

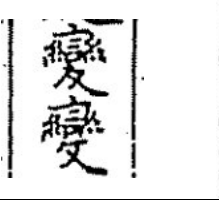
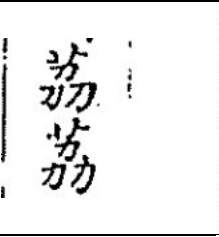
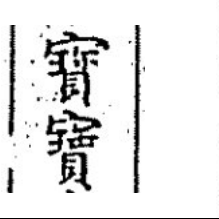
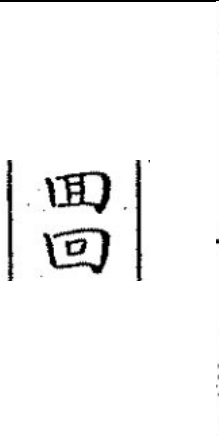
Table 5. Variant forms in configuration across standards in the three places

Character	Left-right configuration 	Top-bottom configuration 
羣 'flock'	 TW, CN	 HK
峯 'mountain peaks'	 TW, CN	 HK

#### 4.1.4 Across time in history

Thus far we have discussed the differences in the standards of different places. Standards across different points in time in history may also differ (i.e. the fourth situation). Table 6 shows Ganlu Zishu 干祿字書 (顏元孫, 1999), which was an official guide for the use of those people who took the imperial examination in the Tang dynasty (618 - 907). In the table, '下正 "bottom correct"' was found to indicate that the variant form at the bottom was 'correct'. As can be seen, these 'correct' written forms in the Tang dynasty occasionally differed from the standard written forms of EDB today. For example, as shown in the table, the highest horizontal stroke in the 言 of the character 變 'change' was much longer than that in the current EDB standard (i.e. 變).

Table 6. Variant forms across time in history

Character	Standard of Ganlu Zishu	
變 'change'		並上俗 下正
荔 'lychee'		並上俗 下正
寶 'treasure'		並上通 下正
回 'to return'		並上俗 下有從回者並準此 下正諸字

#### 4.1.5 Across roles as character or component

In the above cases, the characters served as a character on their own. The next situation we had explored, through the observation of actual usage of characters, occurred when a character acted as one of the components in another character (i.e. the fifth situation). In this case, the written form of a character on its own differed from that of the same character as a component in another character. For example, the three-dot component 'water' historically came from the character 水 'water'. Although they were of the same character, their written forms (as a component and as a character) were clearly different, having 3 strokes in the former but 4 strokes in the latter. See Table 7.



Table 7. Variant forms arises as character or component

Character	Written form as a character	Variant forms as a component
水 'water'	水	泉 海
人 'human being'	人	坐 休
心 'heart'	心	思 恭 快

In addition, worthy of mentioning was the case of the character 釜 'pot', which was made up of the two components 父 'father' and 金 'gold'. When the two components were historically combined together, two of the strokes of the two components merged together. Thus, the total number of strokes of the character 釜 was 10, which was smaller than the sum of the numbers of strokes of the two components 父 and 金 individually (i.e. 4+8 = 12). See Table 8.

Table 8. Variant forms with strokes merged together across roles as character or component

Character	Written form as a character	Written form as a component	
釜 'pot'	釜 10 strokes in total	父 4 strokes in total	金 8 strokes in total

#### 4.1.6 Discussion

In summary, this section has discussed five situations in which the written forms of characters varied in terms of strokes, components, and configurations. One point worth discussing concerns the historical origin of a character, which educators often use for arguing for the correct written form of the character. For example, historically, the upper component of the character 告 'to tell' originated from the character 牛 'cow'. The character 告 originally meant 'to pray' as in 禱告 'to pray' and a prayer was said usually together with the use of an animal such as a cow. Because of this, educators often argue that the written form of the upper component should have its vertical stroke in the middle go through the lowest horizontal stroke in order to align with the written form of the character 牛. Indeed, this is the standard written form adopted by EDB.

But, if we admit this argument, we will run into problems with other characters such as 養 'to raise', the upper component of which is 羊 'sheep' *joeng4*, which serves to signify the sound of the character /*joeng5*/. In this case, should we also have the vertical stroke of the 羊 in 養 pass through the lowest horizontal stroke like that of the 牛 in 告? If not, why do we follow the historical origin of 牛 but not that of 羊? Table 9 shows the hypothetical written forms of a list of characters with their components strictly in alignment with their historical origins. Are these hypothetical forms acceptable? If not, why should we accept the EDB standard written form of the character 告?

Table 9. Hypothetical written forms with components strictly aligning with historical origins

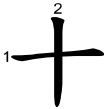




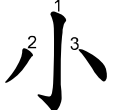
Character	養 'to raise'	休 'to rest'	飯 'rice'	釜 'pot'
Hypothetical written form	養	休	飯	釜

## 4.2 Variation in stroke order

The second category of linguistic variation in characters has to do with the stroke orders of characters, that is, the orders in which the strokes of characters are written. Generally speaking, there are principles that govern the orders of how we should write each of the strokes in a character. For example, for the character 川 ‘river’, we should go from left to right (i.e. the principle) and write the stroke on the leftmost first, and then the one in the middle, followed by the rightmost one.

Such general principles of stroke order are commonly taught in schools. Teachers ask children to write the strokes of characters according to these general principles. However, after a character has been written down on a piece of paper, it will become very difficult, if not impossible, for the teachers to tell the actual order in which the children have used to produce the strokes. In other words, the teachers in reality do not have a clear idea of whether children indeed follow the principles. Little has basically been explored in the orders of how people write each of the strokes of characters.

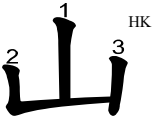

Table 10. General principles of stroke order

General principles	先橫後豎 Horizontal then vertical	先撇後捺 Left slanting then right slanting	從上到下 From top to bottom
Character			
General principles	從左到右 From left to right	先外後內再封口 Outside, inside, then closing	先中間後兩邊 Middle, then both sides
Character			

### 4.2.1 Across individuals

To explore variation in the stroke orders of characters, we had attempted to identify the different orders in which people wrote each of the strokes of the same list of characters (i.e. the first situation). A test was conducted with 95 participants whose daily work involved teaching children to write characters. We read aloud the same list of characters to them and asked them to write the characters in the boxes provided on a worksheet. The boxes were 11 cm x 11 cm large with a tiny reminder of the characters (e.g. 出). After the participants had written each of the characters, they were asked to number the strokes in the orders they had produced the character. The numbers were placed at the beginning end of the strokes where the participants started to write the strokes (e.g. the leftmost end of a horizontal stroke that went from left to right). Upon completion, the worksheets of all participants were collected and analyzed. Table 11 shows the results.

Table 11. Variant Stroke Orders Across People

Character	Variant stroke orders	Other variations	
山 ‘hill’	 74	 14	7

Character	Variant stroke orders			Other variations
火 'fire'				
飛 'to fly'				5
出 'to go out'				23
母 'mother'				4
	70	12	2	11

There were a limited number of variant stroke orders that the participants used to write the strokes. For example, as shown in the table, the participants used only two stroke orders to write the three strokes of the character 山 'hill'. 74 (77.9%) of them began with the vertical stroke in the middle; while 14 (14.7%) of them started with the folded stroke on the left. The former was the standard stroke order of EDB.


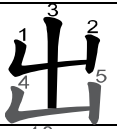


For most of the characters, the standard stroke orders of EDB were consistent with the practice of the majority of our participants. There were yet a few exceptional cases such as the character 母 'mother'. When the participants wrote the last three strokes of the character, 70 (73.7%) of them produced the horizontal stroke first, followed by the upper dot and the lower dot. 12 (12.6%) of them went from top to bottom, producing the upper dot, the horizontal stroke, and then the lower dot. Only 2 (2.1%) of them produced the upper and lower dot first, followed by the horizontal stroke. The last stroke order, which were used by only 2.1% of the participants, was however the standard stroke order of EDB.

#### 4.2.2 Across government standards

In the second situation, we had inspected the government documents on the standard stroke orders of different places. Table 12 shows the standard stroke orders of those characters that the standards of Hong Kong and Taiwan were the same (in the second column), while that of mainland China was different (in the third column). As an example, for the character 乃 'be really', both the standards in Hong Kong and Taiwan were to produce the left slanting stroke first, while that of mainland China started with the double folded stroke on the right.

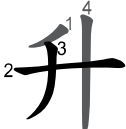
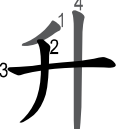




**Table 12. Variant stroke orders across standards in Hong Kong and Mainland China**

Character	Standards in Hong Kong and Taiwan	Standard in mainland China
乃 'be really'		

Character	Standards in Hong Kong and Taiwan	Standard in mainland China
出 'to go out'		
母 'mother'		

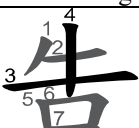
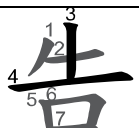
In Table 13, the standard stroke orders of Hong Kong and mainland China were the same but that of Taiwan differed.

**Table 13. Variant stroke orders across standards in Hong Kong and Taiwan**

Character	Standards in Hong Kong and mainland China	Standard in Taiwan
升 'to go up'		
皮 'skin'		
我 'I'		

It was noteworthy that the different stroke orders sometimes resulted from the use of different standard written forms in the three places. For example, both Taiwan and mainland China adopted the variant form of the character 告 'to tell' with the vertical stroke of the upper component not passing through the lowest horizontal stroke. As such, the standard stroke order of the last three strokes of the upper component were consistent with that of the character 土 'soil', that is, the upper horizontal stroke, the vertical stroke, and then the lower horizontal stroke. As discussed earlier, Hong Kong adopted the variant form of the upper component as that of the character 牛 'cow'. The standard stroke order of the last three strokes was accordingly consistent with that of 牛, that is, the upper horizontal stroke, the lower horizontal stroke, and lastly the vertical stroke. Thus the adoption of the variant written form of a character occasionally had a bearing on the adoption of the stroke order for use to write the variant form. See Table 14.



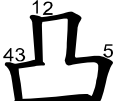

**Table 14. Variant stroke orders resulting from the adoption of different variant forms**

Character	Standard in Hong Kong	Standard in mainland China and Taiwan
告 'to tell'		

Besides the above, the stroke order of a character also sometimes was found to depend on how the delimitation of the strokes of the character was determined. Examples were the characters 凹

‘concave’ and 凹 ‘convex’. Hong Kong, mainland China, and Taiwan all regarded the two characters as having 5 strokes. But the 5 strokes were delimited differently. For example, for the character 凹 in Hong Kong, the stroke on the left and the stroke at the bottom were regarded as one folded stroke (the fourth stroke), while in Taiwan and mainland China, they were regarded as two separate strokes (the first vertical and the fifth horizontal stroke). See Table 15.

**Table 15. Variant Stroke Orders Resulting from Different Delimitations of Strokes**



Character	Standard in Hong Kong	Standards in mainland China and Taiwan
凹 ‘concave’		
凸 ‘convex’		

### 4.2.3 Discussion

The above has explored two situations, in which the stroke orders for writing each of the strokes of the same characters differed. One may wonder that it is now more often for people to use computers than to write characters by hands on a sheet of paper. Does the existence of variation in stroke order further point to the conclusion that stroke order actually does not matter? Should we accept children to use any stroke orders to write characters as long as the correct written forms are eventually produced? Furthermore, the principle to write from left to right was developed in favor for people using right hands to write. Should left-handed children be allowed to write from right to left, neglecting the from-left-to-right principle?

If one believes that stroke order is now of no value, that is, children should be allowed to write the strokes freely, it will be interesting to see whether the following stroke order of the character 萌 ‘to bud’ is acceptable. One of our students (Ng, 2017) once observed a child, who was learning Chinese as a second language, wrote a character in the way similar to producing the character 萌 in the order of the top left ‘十’, the bottom left ‘日’, the top right ‘十’, and finally the bottom right ‘月’. This means the child fanatically followed the principle to write from left to right and broke up the ‘grass’ component at the top into two separate ‘十’s. If we accept children to use any stroke orders, is this stroke order of the character 萌, which reflects the breakup of a component, acceptable? See Table 16.



**Table 16. Observed Stroke Order that Might Not be Acceptable**

Character	Observed stroke order	Standard in Hong Kong
萌 ‘to bud’		
	From left to right	From top to bottom

If the above observed stroke order is unacceptable, it will be interesting to observe in the following case that we indeed accept the use of the stroke order for a character that actually breaks up one of the components of the character. When we write the character 園 ‘garden’, we normally write the 阝 first, then the 袁, and finally the 一 at the bottom (i.e. ‘closing the door’), which is also the standard stroke order of EDB. However, historically, the character 園 was made up of the two components 囗 and 袁, where 囗 ‘to surround’ signified its meaning. Thus, the stroke order we use to

write 園 reflects a breakup of the component 口 into 凵 and 一. In other words, we begin with the component 凵 but, before its completion, we start to write the other component 袁. Do we accept a hypothetical stroke order in which the component 口 is kept intact and is completed before the start of the component 袁? If no, why do we in this case accept the breaking up of the component 口, but not for the case of the grass component of the character 萌?

**Table 17. Stroke Order Reflecting the Breakup of a Component**

Character	Standard in Hong Kong	Hypothetical stroke order
園 'garden'		
	Breaking up the component 口	Keeping the component 口 intact

## 5 Conclusion

Thus far we have discussed two categories of linguistic variation, namely, variation in written form and variation in stroke order. Although we have pointed out a few specific cases in which the standard of EDB was found to differ from the practice of our participants (e.g. the written form of the character 慈 'kind' and the stroke order of the last three strokes of the character 母 'mother'), in light of the majority of characters, the EDB standard was basically found to be in line with the practice of the participants. Thus, on the whole, we recognize the value of establishing the EDB standard, which can serve as a useful common ground for teachers, parents, and other stakeholders to work collaboratively to facilitate the learning of children.

We believe the existing problems lie in how teachers interpret the EDB standard. Should teachers interpret it as a dogmatic rule such that it is the only correct answer? Or is the EDB standard only a helpful reference and there exists multiple correct standards of characters? We, the authors, tend to side with the latter position of the adoption of multiple standards.

To elaborate this position, we would suggest Hong Kong teachers to adopt the EDB standard in teaching situations where only one standard is allowed, for example, in designing teaching materials, in the teachers' own demonstration of characters in classrooms, etc. But teachers should regard variants of characters produced by children that conform to any one of the standards in Hong Kong, Taiwan, mainland China or other widely accepted dictionaries as also correct, that is, adopting all these multiple standards. For example, the writing of the character 告 'to tell' of a child should be regarded as correct no matter whether the child has the vertical stroke in the upper component passing through the lowest horizontal stroke or not. This is because the written form of the child correspondingly conforms to the EDB standard or the standard in Taiwan. Thus variants of characters are regarded as correct as long as they agree with any one of the standards.

However, it must be pointed out that the adoption of multiple standards does not imply the absence of right or wrong answers. A variant of a character that agrees with none of the standards should be considered as incorrect. For example, if a child has the upper end of the vertical stroke in the character 告 'to tell' only touch the highest horizontal stroke without passing through it at the top (i.e. 午 rather than 牛), the written form of the child does not agree with any of the standards and thus should be regarded as incorrect. Similarly, writing the character 萌 'to bud' in the order of '十', '日', '十', and '月' as mentioned earlier, which is unacceptable by any of the standards, should be judged as incorrect.

To make judgments as above, teachers must know more than their own particular variants of characters, which may differ from the also-correct variants of the children. As such, we believe

teachers have to be aware of as many variants of characters as possible. In this light, we argue for the necessity of the knowledge of linguistic variation in characters for teachers to teach characters to children. Without such knowledge, teachers will not be able to respond to the children appropriately during their actual usage of the characters with the adoption of a variant different from that of the teachers.

However, this does not imply that children must also know the same thing as the teachers. In our opinion, it is questionable if children have to be aware of all the possible variants of characters. For instance, it may be more valuable to foster children's interest in literacy than to ask children to spend time on figuring out each of the variants. Let alone the drilling of the variants in a mechanical way. Our belief is that knowledge of linguistic variation is necessary only for teachers.

## 6 Future directions

To reflect upon our project, we would like to rethink what knowledge is essential for teachers to improve their teaching of characters. In preparation courses for pre-service teachers, we often teach them simple and general knowledge about characters, for example, the written forms, sounds, and meanings. But when children learn to use characters in their real lives, unavoidably they will encounter the more complicated and messier reality about the characters (i.e. the linguistic variation). Our concern is thus, there is a gap between what is taught in teacher preparation courses and what teachers actually need to know in their teaching practice in kindergartens. Teachers generally know the characters well but are their practical linguistic knowledge about the real characters enough for them to suitably respond to their children in teaching? What other forms of practical linguistic knowledge than those mentioned in this paper do teachers need? How does such knowledge of teachers relate to their practice of teaching in schools? Practical linguistic knowledge for teachers is certainly an area worthwhile for more attention of future research.

## References

- Aitchison, J. (1999). *Linguistics: An introduction (2<sup>nd</sup> ed.)*. London, UK: Hodder & Stoughton.
- Bauer, R.S. and Benedict, P.K. (1997). *Modern Cantonese phonology*. Berlin: Mouton de Gruyter.
- Czaykowska-Higgins, E. & Dobrovolsky, M. (2010). Phonology: The function and patterning of sounds. In W. O'Grady, J. Archibald, M. Aronoff, & J. Rees-Miller (Eds.). *Contemporary Linguistics: An Introduction (6<sup>th</sup> ed.)* (pp. 59–113). Boston, USA: Bedford / St. Martin's.
- Garvin, P.L. (1993). A conceptual framework for the study of language standardization. *International Journal of the Sociology of Language*, 100–101(1), 37–54.
- Lam, H.C. (in press). *Discerning the differences in Chinese characters* [大同小異：辨識漢字不同處]. Hong Kong: Oxford University Press.
- Milroy, J. and Milroy, L. (1999). *Authority in language: Investigating standard English (3rd ed.)*. London, UK: Routledge.
- Ng, P.K.E. (2017). *A phenomenographic study of the learning and teaching of Chinese characters to non-Chinese speaking (NCS) preschool children*. Unpublished Ed.D. dissertation, the Education University of Hong Kong, Hong Kong.
- Radford, A., Atkinson, M., Britain, D., Clahsen, H., & Spencer, A. (2009). *Linguistics: An introduction (2<sup>nd</sup> ed.)*. Cambridge, UK: Cambridge University Press.
- Wardhaugh, R. & Fuller, J.M. (2014). *An introduction to sociolinguistics (7<sup>th</sup> ed.)*. West Sussex, UK: John Wiley & Sons.
- 傅永和 (2000)。《字形的規範》。北京：語文出版社。
- 國家語言文字工作委員會 (主編) (1997)，《現代漢語通用字筆順規範》。北京：語文出版社。
- 教育部國語推行委員會 (主編) (2008)，《國字標準字體教師手冊》。台北：教育部。
- 常用國字標準字體筆順學習網：<http://stroke-order.learningweb.moe.edu.tw>

- 李新魁、黃家教、施其生、麥耘、陳定方（1995）。《廣州方言研究》。廣東：廣東人民出版社。
- 王寧（2013）（主編），《通用規範漢字表解讀》。北京：商務。
- 胡雙寶（1993）。〈海峽兩岸用字異同議〉，《漢字文化》，3期，16–19頁。
- 課程發展處中國語文教育組（主編）（2007），《香港小學學習字詞表》。香港：香港教育局。香港小學學習字詞表網站：[http://www.edbchinese.hk/lexlist\\_ch/](http://www.edbchinese.hk/lexlist_ch/)
- 費錦昌（1997）。〈現代漢字筆畫規範芻議〉，《世界漢語教學》，2期，10–16頁。
- 費錦昌（2000）。《漢字整理的昨天和今天》。北京：語文出版社。
- 陳越（1965）。〈談字形規範化問題（2）〉。《文字改革》，4期，10–11頁。
- 顏元孫（1999）。〈干祿字書〉，載於魯仁（主編），《中國古代工具書叢編（5）》（827–834頁）。天津：古籍出版社。
- 高更生（2002）。《現行漢字規範問題》。北京：商務印書館。
- 黃靜吟（2005）。《漢字筆順研究》。台北：駱駝出版社。