Theory on Chinese Character Derivation[*]

ZHU Sheng-yu

Tianjin University, Tianjin, China

Character derivation means that if a graph represents several meanings, a new graph based on the original one will be created to bear one or two of the meanings, which is a natural law in Chinese writing system. The old graph is called original character, and the new generated one is called derived character. Two kinds of phenomena—derivation of cognate words (同詞孳乳) and differentiation of unidentical words (異詞別異)—promote the derivation of Chinese characters. In return, derived characters not only bear meaning of the original one, but also serve as a symbol of an independent word and consolidate the graph-meaning relationship. It deserves much attention to the law of the process of character derivation.

Keywords: character derivation, derivation of cognate words, differentiation of unidentical words, interactivity, consolidation of the graph-meaning relationship

Introduction

Graphs are employed to represent words. As visual symbols, the development of words will induce the change of graphs representing those words. Many scholars have done some research on this phenomenon, such as Qiu (1988), F. Y. Wang (1989), N. Wang (2002), and so forth[1]. In general, the production of derived graphs is due to kinds of motivations: derivation of cognate words[2] (同詞孳乳) and differentiation of different words (異詞別異). The latter includes two situations. The first is that a graph is borrowed to write another word that has no literal semantic-graphic relation; the second refers to homographs (同形字), which means that only one graph writes two or more different words—the graph is graphically corresponded to each word it represents. Derivation of cognate words is easy to understand, which means that when different words are derived from the same root, distinctive external forms are also required as symbols. The essence of differentiation of unidentical

---

ZHU Sheng-yu, Ph.D., Lecturer of Research Centre for Linguistic Sciences, Tianjin University.

[1] All these scholars gave definitions about Character Derivation; Also, they talked about the methods of Character Derivation. Six methods are listed in Qiu Xi-gui's *Wénzìxué gàiyào* (Qiu, 1988; 2007, pp. 223-254). Wang Feng-yang's *Hàn zìxué* (Wang, 1989) made a detailed and comprehensive research on the subject. Wang Ning's *Introduction to Configuration of Chinese Characters* (Hànzì Gòuxíngxué Dǎolùn) (Wang, 2002; 2015) focuses on Character Derivation induced by the derivation of cognate words. In addition, Cheng Rong (1992), Zhang Xi-feng (1995), and Hao Shi-hong (2008) also did some research on Character Derivation.

[2] Here there may be a puzzle about what a word is in Chinese language, which refers to the "identity of a word" (詞的同一性). Admittedly, there are some differences on "identity of a word" between English and Chinese. If a phonetic form writes one or more but semantic related meanings, we think that the bundle of related meanings belong to an identical word. However, Chinese words (a unit of phonetic-meaning combination) are visually displayed with graphs that are closely related to the meaning. So when exploring the "identity of a word" in Chinese, we should also consider the visual graphs. In a graph form graphically corresponds with the meaning of a word unit, we regard it as an identical word. With more details see Zhang Lian-rong (2000, pp. 130-185).

(different) words is when a same graph form is used to represent unidentical word units there is also the need for distinctive external forms as symbols, in order to express ideas accurately. In return, if both situations' words can be distinguished with visual graphs, those graphs not only take some meanings or functions from the original words, but also consolidate the derived words. The interactivity between Chinese words and characters can be generalized as two points: words as motivated force of graphic change and the solidification of words due to graphs.

Words as Motivated Force of Graphical Change

We can observe this phenomenon on the basis of relation between a graph and its meanings. Every graph is created based on literal meaning (本義)[3], so there is an intimate connection between graph and the meaning of words. After the graph is created, when it is used in language, it will develop out some extended meanings (引申义) through mechanism of metaphor and metonymy. Not only that, when a graph is bonded with its word, it gets a sound from the word in the language[4], and the graph-sound relation is gradually solidified in our cognition. So when a word is too abstract to create a graph for it or when people do not know how to write the graph representing a particular word, they are inclined to borrow a existed graph to write that word[5]. As to the word, the graph is called borrowed character; and as to the graph, the meaning is also borrowed from other words. In any case, the meanings a graph represents can be divided into three categories: literal meaning, extended meanings, and borrowed meanings[6].

A graph has its own literal meaning, from which extended meanings are developed; besides, the graph can be possibly loaned to write other words. Thus, it seems that many senses or functions are represented by just one graph, which can be called meaning-expansion of graphs. Meaning-expansion is the motivated power of character derivation. Character derivation can be defined as following: meaning-expansion of graph form X motivated the creation of graph form X' to bear one or two meanings of X, namely $X \rightarrow X'$. X is called the original graph and the X' is called derived graph.

Strictly speaking, we should distinguish two different kinds of derivations: words' derivation and characters' derivation, which are on different levels. As to a single Chinese character, literal meaning, extended meanings, and borrowed meanings can be classified into two categories: literal meaning and extended meanings are written by one graph, so they belong to an identical lexical item—polysemy. Borrowed graph and homograph can be analyzed to represent unidentical lexical items—homonymy[7]. It is not very convenient and efficient to communicate if a lexical form has too many meanings. Both polysemy and homonymy can lead to communicative conflicts and thus trigger the change of lexical form. The derivation of polysemy into different

---

[3] We assume that a lexical item has many senses, but not every sense can be treated as the base or reason to create graphs. Literal sense refers to one of the senses that we can get through the analysis of the graph. The literal sense may not be the original one that the graph represents.

[4] In principle, a graph itself, just as a symbol, has no sounds, and it only acquires the sound by the connection with the words it represents.

[5] In this case, there is no graphic-semantic relation between the graph and the word, so we usually call this kind of graph as borrowed characters (or loan characters).

[6] Borrowed meanings can also be called loan usages. Sometimes, a graph is borrowed to write several meanings.

[7] Here we use homonymy to refer a word that is spelled the same but is different in meaning or origin, the sound either the same or different. E.g., hé (何) represents two unidentical word units: One is the literal meaning "負荷" (load), and the other is borrowed meaning interrogative pronoun. The two word units are phonetically related. The Oracle-Bone Inscriptions (abbr. OBI) graph ⟡ is used to represent two unidentical word units: One is "月" (moon), and the other is "夕" (eve). The two word units are not phonetically related.

lexical items is called derivation of cognate words. And the differentiation of homonymy into various lexical forms is called differentiation of unidentical words. The former happens in the level of language, and the later exists in the level of symbol of the language. Although they are in the different levels, but if we focus on the view of graphs writing those words, they will both lead to the process: X→X'. X' effectively deprives one or two meanings of X, so X loses that meanings at some point of time, thus concepts or senses can be expressed more precisely. For both X and X', they have less meanings or functions. In this sense, X' is called derived characters of X.

On the surface, it seems that the derived graph X' is created out of the meaning-expansion of the original graph X. However, we have to remember the fact that there are two kinds of situations behind. One is that the derivation of cognate words leads to the derivation of characters[8], the other is that the derivation of characters is appealed as the purpose of differentiating homonyms. There are essential differences between the two above-mentioned situations, but they can be settled with the help of changing lexical forms.

Both derivation of cognate words and differentiation of different words should resort to some methods. Theoretically, both of the situations can use phonetic change or graphic change. Nevertheless, there are different manifestations with the two methods applied to the situations.

(1) For derivation of cognate words. It is acknowledged that phonetic change can be used to distinguish cognate words. For instance, distinguishing meaning by means of the four tones (四聲別義) is very common in Chinese. E.g., shù (數: number) is a noun, while shǔ (數: to count) is a verb. Yī (衣: clothes) is a noun, while Yì (get dressed) is a verb (Sun, 2007)[9]. In addition, change of graph forms also works well. Examples are as follows:

(a) 正-征-政: In the OBI (Oracle-Bone Inscriptions), the graph zhèng/zhēng (正) has at least three kinds of meanings: 祭名 (sacrifice), 適當 (prope), and 征伐 (go and fight; war). It seems a little burden for the graph 正 to bear those meanings synchronically, at least for language users at that time, so a new graph is extremely needed to reduce the burden of the graph 正. As a result, graph zhēng (征) was created during the late Shang dynasty, which was specially employed to represent the meaning 征伐 (go and fight; war). However, with zhēng (征) largely used, it develops some new meanings, such as 征月 (lunar January)[10], 征收 (levy), 貢稅 (tribute and taxes), 政事 (official affairs), 人名(name of people), etc. Like before, zhèng (政) was created to specially represent the meaning 政事 (official affairs). In a word, zhēng (征) is called derived character of zhèng/zhēng (正), and zhèng "政" is the derived character of zhēng (征).

(b) 工-功: In the Bronze Inscriptions, gōng (工) can represent lots of meanings, e.g., 工匠 (artisan), 官吏 (officer), 功績 (achievement; merit), 工室 (a factory for casting weapons), and so on. In order to bear the meaning 功績 (achievements; feats) of the graph 工, a new graph gōng (功) was created purposely (Schuessler, 2007), so gōng (功) is the derived character of gōng (工).

Examples like mentioned above are plentiful in archaic Chinese. As we can see, both phonetic change and graphic change can take effect in distinguishing lexical items, independent on each other. Sometimes, not only sound changes, but also the graph. Usually it is difficult to determine whether sound changes earlier than graph

---

[8] Monosyllable and disyllable obey to different rules in derivation respectively. Here we just talk about the monosyllables.
[9] For more details, please see "Study on Chinese Word Formation by Tone-Shifting" (漢語變調構詞研究) by Sun Yu-wen (孙玉文, 2007).
[10] In the Bronze Inscriptions, both "正月" and "征月" are existing with the same meaning, but "征月" is not as common as "正月". E.g., *YīnZhōu JīnWénJíChén* (殷周金文集成) NO.5.2695 YuanFāngDíng (員方鼎).

or the opposite.

(2) For differentiation of unidentical words. Phenomena of loan characters are very common in archaic Chinese. But theoretically each character has its own literal sense and extended meanings. Therefore, if a graph has loan usages, it will be a great burden for the graph to represent so many meanings or words. Thus there is the impetus for differentiation. For example:

(a) 辟-避-劈-壁: According to *Shuōwén jiězì*, the literal sense of bì (辟) refers to 法 (law). Besides, it can be used to express 君王 (king), 官吏 (officer), 罪行 (crime), 懲罰 (penalize), 治理 (govern; administrate), and so forth. Meanwhile, bì (辟) was borrowed to write the following words: 躲避 (avoid; keep away from), 劈开 (split), 墙壁 (wall), etc. In order to specially bear the loan usages of bì (辟), the graphs bì (避), pī (劈), and bì (壁) were created.

(b) 莫-暮: The literal sense of mò (莫) is 日暮 (sunset), but most of the time it was loaned as a negative pronoun, so a semantic determinative rì (日) was added to mò (莫) to form the new graph mù (暮), which specifically represents the meaning of 日暮 (sunset) while mò (莫) functions as the negative pronoun.

Examples like that are numerous as well. According to our observation, it seems that pure phonetic change cannot work well when used to differentiate loan usages, but we know little about the reason. There are some possible conclusions based on theoretic reasoning. Firstly, the condition of borrowing characters is that the pronunciations of the two words are the same or close enough. If the pronunciations are close instead of the same, it is not necessary to change sound, because they are originally different. Secondly, the number of phonetic units in a language is always limited, so if phonetic change happened, it would be the same with another pronunciation, which makes no practical difference in oral communication. Finally, phonetic changes, sometimes, are locally established by the people through long social practice, but this kind of establishment may not be accepted by others. Therefore, loan usages are usually graphically differentiated, because the production of graphs can be countless.

To sum up, derived characters refer to those new graphs generated in the process of differentiating cognate words and homonyms, which can be illustrated as following:
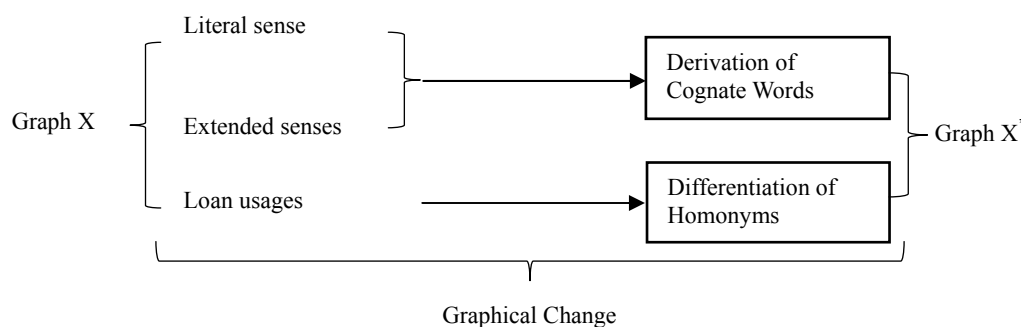

Figure 1. Theory of Graphical change.

Graphical change is one of the means to differentiate lexical items. As we know, there are many ways to realize graphical change, and many specialists have done research on this (Zhang, 1995; Hao, 2008)[11], so we will not waste space here.

---

[11] The most commonly used way for graphical change is to add a determinative to the original graphs. With more details, see Zhang Xi-feng (1995) and Hao Shi-hong (2008).

Graphs' Function in Solidification of Words

Graphs are the external forms of words, and words are the contents of that forms. Derived character not only carries off one or two meanings from the original character, but also plays a great role in the solidification of words it writes, which can be observed through the following facts:

Firstly, derived character is the symbol of the independence of a lexical item. As we mentioned above, both phonetic change and graphical change are the means of differentiating lexical items. Although graphical change does not necessarily lead to the independence of a lexical item, it can be an efficient symbol of the independence cognitively and psychologically. For instance, tián (田) has already appeared in the OBI, which represents two closely related meanings: 田地 (farmland) and 田猎 (hunt). In the late Shang period, tián (畋) was created to specially represent the meaning 田猎 (hunt). Before tián (畋) appeared, we say that 田地 (farmland) and 田猎 (hunt) belong to the same word—a word has two meanings. With the advent of tián (畋), we deem that tián (田) and tián (畋) represent two independent but cognate words. The same situation can be seen from fǎn (反) and fǎn (返). In the Bronze Inscriptions, fǎn (反) bears so many meanings, such as 歸來 (come back), 歸還 (give back), 反叛 (rebel), etc. Then a derived graph fǎn (返) was produced to represent the meaning 歸來 (come back), so we regard fǎn (反) and fǎn (返) as two characters writing cognate words. In brief, the new generated graph is of significance in judging the identity of word.

In addition, derived characters can make the semantic domain of the words written limited and explicit. most of the derived characters are generated by adding determinatives (意符) to the original characters. Not any determinative can be employed every time when creating derived characters. In fact, the determinative must be chosen according to the meaning the derived character is about to represent. The choice is limited by two factors. On the one hand, determinative should be selected from the pool of extant graphs. Theoretically, every graph can serve as a determinative or a phonetic element, but there is lack of balance among graphs which can work well enough as determinatives. On the other hand, suitability and uniformity between determinatives and the meaning the derived character is about to represent should be considered. Whether used to represent the literal meaning, extended meanings, or loan usages, the derived characters, generated by adding determinatives, ought to graphically relate to the meaning they write, namely, conforming to the Theory of Graph-Meaning Corresponding (形義統一原理). In return, the added determinatives make the semantic domain of the words that the derived characters write more limited and explicit, which can be seen from series of derived characters with the same added determinatives. For instance[12]:

A. 巳-祀、石-祏、御-禦、舌褐、冊-𣬉
B. 多-奲、井-妍、良-娘、喜-嬉、帚-婦
C. 田-畋、學-斆、正-政、吾-敔
D. 般-盤、齊-齍、它-宧

All the derived characters of Groups A and B exist in OBI. All the derived characters in Group A were created by adding shì (示) to the original ones, indicating that the semantic field is related to sacrifice. In Group B, with nǚ (女) added, all the derived characters were used as female's names in OBI. In Group C, tián (畋) and xué (斆) appeared in OBI, while zhèng (政) and yǔ (敔) emerged in the Bronze Inscriptions. The same point is that all the derived characters are formed by adding pū (攴) whose semantic field is closely related to action behavior. Derived characters in Group D, appearing in the Bronze Inscriptions, are formed by adding mǐn (皿),

---

[12] For all the group of characters below, the former are the original characters, and the later are derived characters.

which tells us that the words these derived characters represent belong to the vessel semantic field.

Finally, most of the derived characters produced during the earlier times are still applied now instead of being weeded out, which demonstrates their indomitable vitality in Chinese. As some scholars pointed out, the number of phonograms (形声字) approximately occupy 87% of the total amount of Chinese characters. Moreover, character derivation is the most productive way for creating phonograms. Therefore, it is not difficult to imagine derived characters' power of consolidation of words. Graphic change plays a better role than phonetic change in consolidation.

In conclusion, the interactivity between Chinese characters and words can be vividly illustrated with the following Figure 2:
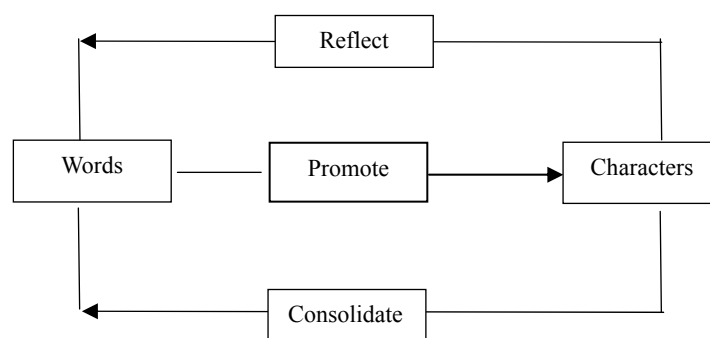

Figure 2. Interactivity between words and characters.

The development of words or meanings in Chinese language promotes the graphic change in Chinese characters; in return, the derived characters work as a symbol of independent lexical items and consolidate the graph-meaning relationship, making the perfect co-adaptation between Chinese writing system and the language.

References

Cheng, R. (1992). 古文字分化問題探討 (Discussion on the differentiation of ancient Chinese characters). *語言文字應用 (Application of Linguistics), 4,* 64-75.

Hao, S. H. (2008). 古漢字同源分化研究 (The study on the cognate derivation of Ancient Chinese Characters). Hefei: Anhui University Press.

Qiu, X. G. (2007). *文字學概要* (*Chinese writing*). Beijing: The Commercial Press.

Schuessler, A. (2007). *ABC etymological dictionary of old Chinese*. Honolulu: University of Hawai'I Press.

Sun, Y. W. (2007). *漢語變調構詞研究* (*Study on the Chinese Word Formation by Tone Change*). Beijing: The Commercial Press.

Wang, F. Y. (1989). *漢字學* (*The science of Chinese characters*). Changchun: Jilin wenshi chubanshe.

Wang, N. (2015). *漢字構形學導論* (*Introduction to construction of Chinese characters*). Beijing: The Commercial Press.

Zhang, X. F. (1995). 分化字的類型研究 (Study on the types of differentiated characters). *語言教學與研究 (Language Teaching and Linguistic Studies), 1,* 96-107.