

## A SEMANTIC FRAME ANALYSIS OF CHINESE CHARACTER: RELEVANCE AND LIMITATIONS

Olabinjo Yewande, M (Ph.D.)  
Department of Linguistics, African and  
Asian Studies, University of Lagos

### **Abstract**

*Chinese characters are one of the earliest forms of writing known to man. They represent one of the four independent inventions of writing in human history to be universally accepted by scholars (Jianyu Li 2012). Chinese characters are carriers of Chinese history, culture and visualisation of Chinese perception. The Chinese language, from the discovery of oracle bone inscription to the contemporary form, has gone through several stages of development that emanates features generally studied by scholars. Using the frame semantic framework, this work looks at the principles of character formation in the Chinese language. The phono-semantic principle is one of the most prominent features of the characters. In the collation of 7,697 characters in 100 CE, 82.3% of the characters are phono-semantic characters--Xu Shen's <Shuo Wen Jiezi >, which account for character pronunciation and/or relevant semantics for the character formed, while also on several occasions give pointers to the historical etymology of the characters. This research analyses the semantic and phonetic frames present in phono-semantic characters, their relevance, and limitations.*

**Keywords:** *semantic frame; phonetic frame; Chinese character*

### **Introduction**

The exact time the Chinese character was invented has been a constant source of debate by scholars. Some schools of thought believe Chinese writing has been in existence since the late Neolithic period. They argued that varieties of marks on pottery pieces found to be as old as 4800 BC are Chinese characters (Boltz 1994)<sup>1</sup>. With this assertion, it means

---

<sup>1</sup> Boltz, W. G. (1994). The origin and early development of the Chinese writing system. New Haven, CT: American Oriental Society

writing was invented in China about a thousand years earlier than anywhere else. This assertion has since been discredited as the inscriptions are considered not coherent enough to be termed a language. However, the earliest inscription undoubtedly recognized as a Chinese character was found in the late Shang dynasty. The late Shang dynasty inscription, commonly known as the Oracle Bone script, existed between 1200-1046 BCE. They were written on oracle bones and turtle shells.

After this form of inscription was discovered, a bronze inscription (1046-221 BCE) was found. The bronze inscription was used during the Zhou dynasty. The characters during this period were often inscribed on bronze artefacts. After the bronze script, there was the Seal script, the Clerical script, and the regular script called the Kaishu. However, during the mid-twentieth century, the government enacted a unification reform. The reform aimed at simplifying the complicated Chinese character into a more simplified form in order to promote literacy.

This work is divided into introduction, theoretical framework, evolution and features of Chinese characters, Phono-Semantic compounds, phonetic frame, semantic frame, relevance of Phono-semantic frames, Limitations of phono-semantic frames, and conclusion.

### **Theoretic framework**

This paper employs the frame semantic framework. Frame semantics is a linguistic theory developed by Charles Fillmore, which explains how people understand the meanings of words and concepts based on contextual "frames." According to this theory, a "frame" is a mental structure of related concepts and experiences that helps people interpret language. "Frame Semantics" as enacted in 1976, introduces the concept of frames in an article presented at the annual meeting of the Linguistic

Society of America. Fillmore began exploring how meaning is understood within structured sets of related concepts and experiences, which he called frames.

Frame Semantics has practical applications in linguistics, cognitive science, artificial intelligence (AI), Machine Translation, Educational Tools, Language Teaching, and natural language processing (NLP). By structuring and connecting concepts around frames, it aids the interpretation of language based on context. Frames are conceptual structures that provide context for elements of interpretation; their primary role in an account of text understanding is to explain how our text interpretations can (validly) leap far beyond what the text literally says.

This work applies Frame Semantics to the phono-semantic features of Chinese characters by analyzing how meaning is constructed through the interaction between phonetic and semantic elements in the characters. The framework analyzes the manner at which meanings are framed by the combination of phonetics and semantics.

### **Evolution and features of Chinese Characters**

The ancient Chinese characters, before the simplified form, were mostly pictograms; they were representations of physical objects. They are the fundamental inscription from which all the other abstract inscriptions emanate. The characters received modifications over different historical periods.



fig.1

A pictogram closer to the image of the actual object “the sun” was found in the Oracle Bone Script as well as the Bronze Inscriptions, as illustrated in *fig1*. The initial square shape was modified and added in the Clerical Script, while the regular script perfected the shape to what it is in the present day. Therefore, Chinese characters are carriers of Chinese history, culture, and the visualisation of the Chinese way of perception. For instance, any characters relating to the hand intuitively carry a character frame for hand “手 扌 - shou”, e.g., 打 - *hit*、拉 *pull*、拿 *take*、接 *receive*、指 *-point*、抓 *grab*. These characters display the Chinese ideology on the formation and construction of the characters representing the various meanings attached.

The final character in the simplified form presents features that have been studied by scholars. Scholars in the Han dynasty stated that there are six essential features to character formation in the Chinese language, which they called “Liu shu”. The term “Liu shu” was first used in a book called <Zhou Li>, while describing the methods used in educating the Chinese elites on Chinese character. The method was referred to as <Liu Yi – Six Arts>. Xu Shen in <Shuo Wen Jiezi > published in the Han dynasty (c. 100 CE), stated the six Arts as the basic principles on which the structure and meanings of characters are formed. The principles are: pictograms (象形 *Xiàng xíng*), Ideograms (指事 *Zhǐ shì*), compound ideograms (会意, *Huì yì*), phono-semantic compounds (形声, *Xíng shēng*), rebus (phonetic loan characters) (假借, *Jiǎ jiè*), derivative characters (转注, *Zhuǎn zhù*).

Pictograms are the earliest form of Chinese character development; they originated from the image of the object they represent in the natural world. Pictograms were used mostly to represent tangible objects and not abstract objects. They have subsequently served as a building block for *a lot of* characters in the other features. While pictograms are visual

representations, ideograms are representations of abstract ideas or concepts using simple symbolic elements. The composition is usually the concept it represents. They are used to represent ideas that pictograms cannot capture. Oftentimes, the abstract idea is added to the pictogram image already in existence to create a new word meaning. For instance, Běn 本 formed from the pictogram character for tree 木 with a horizontal stroke added to the pictogram to form the word for tree root.

Pictograms and ideograms were combined to create the compound ideogram, most likely as a result of the urge to create more relatable terms. Combinations of two or more pictograms or ideograms are known as compound ideograms. For instance, a combination of three trees 木 *mù* formed the character for forest 森 *sēn*. The characters combined can occasionally be the same as in 森 *sēn* or different characters as in 明 *míng*. 明 *míng* is a combination of two individual characters for the *sun* and *moon*, respectively, to mean *clear*. Derivative Characters are features of characters with similar etymological roots that have since evolved to represent different meanings but maintain a common structural, historical or semantic link. A rebus is a situation in whereby a character is used to represent a different concept. While not creating a new character, a word with a similar phonetic sound is used to represent another. The new concept sometimes renders the old one obsolete. It is, however, noteworthy that derivative characters and rebus are not listed in some Chinese texts as word formation features; works like <*Xian dai Hanyu*> (2011) did not list derivative and rebus.

A phono-semantic compound is the combination of a semantic component and a phonetic component to form a new character. This work studies the phono-semantic compounds in Chinese characters, their features, capabilities, and constraints.

### Phono-semantic Compound

Phono-semantic characters are the most commonly used character formation process. In Xu Shen's <Shuo Wen Jiezi >, in the collation of 7,697 characters, 82.3% of the characters are phono-semantic characters. A phono-semantic compound, as earlier mentioned, makes use of the combination of a semantic component and a phonetic component to form a new character.

This form of character records pronunciation and semantics embedded in the character. They are compound characters with two parts merged to form one. However, the merging is not similar to the compound ideogram merging, as compound ideograms are mainly composed of pictograms and ideograms. Phono-semantic Compounds can combine with all the other features to make up their character. Pictograms, ideograms, and compound ideograms, as seen below, can all be combined to form phono-semantic compounds.

- i. *Pictograms:* 木 (tree) 、 手 (hand) 、 心 (heart)
- ii. *Ideogram:* 本 (root) 、 上 (up) 、 下 (down)
- iii. *Compound ideogram:* 姑 (Aunt) 、 从 (from) 、 明 (bright)

Examples of such Phono-Semantic characters are 菇 gū、 志 tǎn、 淋 lín、 体 tǐ. 菇 gū -combinations of a pictogram and a compound ideogram; 志 tǎn、 恣 zì, and 体 tǐ combine ideograms and pictograms.

The distribution of the phonetic and semantic features in a character varies; their positions are not fixed. Their distribution is inherently dependent on the structure of the character. The most frequently observed distribution includes:

- i. Phonetic left - semantic right as in (攻-gōng、 战-zhàn、 切-qiè)
- ii. Semantic left - phonetic right as in (河-hé、 挑-tiāo、 肝-gān)

- iii. Semantic top - phonetic bottom as in (芳-fāng、宇-yǔ、竿-gān)
- iv. Semantic bottom - phonetic top as in (袋-dài、貸-dài、型-xíng)
- v. Semantic outer - phonetic inner as in (阁-gé、园-yuán、固-gù)
- vi. Semantic inner-phonetic outer as in (闻-wén、问-wèn)

Semantic left-phonetic right is the most common of the structural forms. An important feature of phono-semantic characters is that they create a semantic frame for the characters associated. They either create a semantic frame or a phonetic frame, which in turn creates an association among characters with similar frames. The frames enable easy comprehension and assimilation of characters with similar frames.

### **Phonetic Frame**

A Phono-semantic compound, as earlier mentioned, has a combination of two parts, namely the phonetic radical and the semantic radical. Characters that share phonetic components often belong to the same “phonetic frame” and share similar pronunciations. The phonetic radical gives a pointer to the pronunciation of the final character compound. According to <Xian dai Hanyu> 2011 (*it is always better to direct all references to the author(s) and not the book title*), one fourth of the total phono-semantic characters have the same pronunciation. For instance: 蝗、螳、磺、璜, they are all pronounced as Huáng, with the same tone.

Phonetic radicals are the oldest way to learn several characters at once. Therefore, in the mind of a language learner, there is a phonetic frame creating an association of phonetic sounds framing all such sounds together.

There are no radicals dedicated to be phonetic radicals; a radical can be a phonetic radical in one character and a semantic radical in another. For example, the radical Mù 木 in the compound 根 gēn is a semantic radical, while the same 木 mù is a phonetic radical in 沐 mù. This implies that a radical creating a phonetic frame in one character can create a semantic frame in another.

Generally, phonetic features in phono-semantic characters appear in different forms (Jianyu et al, 2012<sup>2</sup>), which are characters with the same pronunciation, tone variation, sounds with vowel or consonant variation, and characters where the pronunciations differ both in consonant and vowel.

Characters with the same pronunciation as mentioned above involve phono-semantic characters with similar tones and phonetics. The only way to differentiate them is through their semantic radical; otherwise, they can be mistaken for the same characters. The second form involves some phonetic frame presenting words with variation in just their tones. These classes of characters have the same pronunciation but different tones. For example: 轮、论、抡 – they are all pronounced as *lun*, but they all carry different tones, *lún*/*lùn*/*lūn*, respectively. The third form is phonetic frames with just a vowel or consonant variation, they present a case whereby only the initials or a vowel of the final differ. For example: 汀、灯、厅 are pronounced *dīng*, *dēng*, and *tīng* respectively, where *dīng* differs from *dēng* by only a vowel *e*, while in the case of *tīng*, the initial *t* differs, but a speaker can still deduce they all belong to a single phonetic frame.

The final form is one whereby there are no similarities in pronunciation,

---

<sup>2</sup> Jianyu et al (2012) indicated four forms of the phonetic features of phono-semantics. But we were able to categorize the forms to five in number.

both in consonants and vowels, in the phonetic feature. 罪<sup>zhuì</sup> and 排<sup>pái</sup> were cited as examples for this form, stating that they both contain the same phonetic radical but have different pronunciations. The above-listed are the only four forms listed by Jianyu et al (2012).

However, this work observed a fifth form whereby there are sounds similar by association. Sounds similar by association refers to the form in which characters with a particular phonetic sound can all end or contain similar finals. For instance, characters with the radical 仓 will usually end with ‘ang’ as their phonetic frame, as in 疮<sup>chuāng</sup>, 抢<sup>qiǎng</sup>, 苍<sup>cāng</sup> pronounced *chuāng\qiǎng\cāng* respectively. This form can commonly be observed in lots of characters. The difference in the associated character goes beyond just a vowel or consonant, as observed in *chuāng\qiǎng\cāng* above, more than one vowel or consonant varies in the character, but they all have partial phonetic frame similarity.

Generally, the phonetic frames in Chinese characters can be divided into two types. The first type can be equally utilised as characters. For example: 仓<sup>cāng</sup>, 非<sup>fēi</sup>, 木<sup>mù</sup>, 丁<sup>dīng</sup>, etc. These characters have their individual meaning; therefore, they can be used as words on their own. The second type are characters that are no longer in use as words or cannot stand on their own, and cannot be used as characters. Examples of such characters are 畷<sup>shù</sup>, 琴<sup>qín</sup>, etc.

### Semantic Frame

The semantic frame is the most important feature in getting the meaning of the character, while the phonetic frame only gives pronunciations associated with the character; as in language, the semantic frame helps make sense of the whole character. The phonetic frame can be random, but the semantic frame is precise. They give a specific category for the phono-semantic compound. For example, the character 扛<sup>kāng</sup> also gang and the character 扛<sup>kāng</sup>. Although they both share the same phonetic

frame, however, their semantic frame distinguishes them from similar characters. 扌*kang* carried the radical for *hand*, indicating its meaning has something to do with the *hand*; the semantic frame will link it with characters with similar frames. The same will apply to 土*gang*, which carries the *earth* radical, indicating its meaning has to do with earth.

Most semantic frames are usually in the form of pictograms that directly depict the item they represent. This makes it easy to recognise the semantic category of the character by merely observing the semantic frames. For example: 流*liú*、樓*lóu*、摸*mō*、物*wù*; their semantic frames 氵、木、扌、犮 are all pictograms giving a hint of each character's category. 氵 has *water* as its frame, 木 *tree*, 扌 with *hand* or an action done using the hand, and 犮 with *animal* as frame. These frames are all indicators of the semantic categories of the words, respectively. This helps in the understanding of the structure of the characters.

### Relevance of Phono-semantic frames

Phono-semantic frames are no doubt important to Chinese character learning. They basically help in the understanding of the ancient Chinese ways of thinking, and the ideas behind characters' formation process. For example, the characters: 拉*la* (*pull*), 抓*zua* (*grab*), 挂*gua* (*hang*), are all characters whereby *hands* are used to perform them, hence the idea of carrying a radical indicating hand.

They are useful in the differentiation of characters with similar tones and sounds. There are characters which have similarities both in phonetics and tone; such characters can only be distinguished through their semantic frame.

An etymological trace to character is made possible by Phono-semantic frames. For example, the character for *gun* 枪 has the semantic feature 木, indicating the use of wood. This is a direct indication of the object during

ancient times, when *guns* were made from wood. Although *guns* are no longer made using wood, however, this gives learners relevant information about the history of the character formation.

The phonetic frame of a phono-semantic character is responsible for the provision of the sounds in the pronunciation of characters with similar frames. This implies that words with different semantic frames can have the same or similar pronunciations. For example, 柑 *gān*, 甘 *gān* – both characters have exactly the same pronunciation as a result of their phonetic frame, likewise 柑 *gān* and 甘 *gān*.

Phonetic frames can aid language users in the learning of multiple characters using just the sounds with similarity. 换 唤 煊 all have *huan* as their pronunciation; a learner can utilise the phonetic frame to assimilate all the different characters. Also, differentiation of characters with similar outlook is made possible with phonetic frames. For example, the characters: 沧 *cāng* and 沦 *lún*, at first glance, might look the same, but the phonetic feature distinctively differentiates them.

### **Limitations of phono-semantic frames**

With the need to unify and simplify the complex characters of the old times, the government enacted a unification reform in the mid-twentieth century. The simplification of these characters initiated a lot of changes in the general structure of the Chinese characters. These changes incited a modification in the understanding of the phono-semantic feature of characters. During the unification, some characters were eradicated, while some were merged into one. For example: 谢 *xiè*, the word was used not just to say *thank you*, but also to apologise, but during the simplification, its meaning was shrunk to just one. Also, the character 谢 *xiè* was originally “*xié*” with a different meaning (Cao Wei 2001).

While phono-semantic frames give pointers to the etymology of characters, one of the main challenges to their characters is the change in object type. As earlier mentioned, the character for *gun* 枪 *qiāng* has the semantic feature 木 indicating the use of wood, but since the invention of guns to date, changes have been made to guns. With the current state of guns, the most appropriate semantic frame will definitely be 金 *īn* –metal and not *wood*.

Changes in character meaning also bring about inadequacies in the accuracy of phono-semantic frames. For instance, *táng* 塘 meaning pool, in traditional Chinese, *táng* 塘 means a pool bank (Cao Wei 2001)<sup>3</sup>. The semantic frame for the character *Tǔ* 土 ‘Earth’ will not be appropriate to describe a pool in contemporary times, largely because pools come in different forms that have no similarity with earth.

There are times when the semantic frame can bring about confusion in the grasping of character meaning. This can happen when a character’s semantic frame clashes with another that previously existed. Hence, making it impossible to rely on the frame to navigate through characters. A typical example will be the character 月 *yue*, it is a pictogram of the moon which occurs as a semantic radical in characters like *zhāo* 朝 –Morning, *qī* 期 *period*, *lǎng* 朗 *bright*, where their meaning has to do with light and brightness. However, in the same light, the same radical is used as the semantic feature in characters that have to do with parts or organs of the body. Characters like 肌 *jī* –muscle, 肝 *gān* –liver, 胃 *wèi* –stomach, 脏 *zàng* internal organs. In this kind of situation, reliance on the phono-semantic frame to understand or grasp the structure of characters cannot be relied on.

Similar to the semantic frames, the relevance and reliance of the phonetic

---

<sup>3</sup> Cao Wei. 2001. Hanyu Jingjiang. Beijing university press.

frame in character have also been challenged in the modern Chinese language. The percentage of characters categorised as having phonetic frames is low in comparison to their semantic counterpart. In the early 50's, out of the 2000 characters published by the Chinese government, only 388 characters, that is 19.4%, present an accurate phonetic frame.

A major component that has affected the phonetic frame over the course of time is the change in character form. Cao (2001), while discussing this phenomenon, cited examples of characters which have gone through changes such that their phonetic frame is no longer apparent by merely looking at the characters, hence, bringing about uncertainty in the accuracy of their feature. For example, the phonetic frame for 寺*si* is 之*hi*, which is not even traceable from its character; the same also applies to other characters like 夜*è*, where its phonetic frame *yì* 亦; and 布*bù*'s phonetic frame is 土*tǔ*, which also does not reflect or is no longer traceable in the character.

Another problem is that some characters do not produce the same sounds as their phonetic frame; they have an entirely different sound, which makes it difficult to predict sound from its frame. An example of such a character is 寿*shòu* characters with a phonetic frame that includes 禱 *dǎo*、涛 *tāo*、筹 *chóu*。 The character 勺 *sháo* could be even more complex, as they create characters with completely different pronunciations and unrelated semantics that are hard to trace. The only pointer to their frame will be the radical 勺 *sháo*. Examples of such characters are: 灼 *zhuó*、约 *yuē*、炮 *pào*、的 *de*。

### Conclusion

Phono-semantic has a combination of two parts, namely the phonetic radical and the semantic radical. Characters that share phonetic components often belong to the same “phonetic frame” and share similar pronunciations, while the ones that share the same semantic category

have the same “semantic frame”. Semantic frame gives an inclination of the category of word meaning. Phonetic and semantic frames are major connecting blocks connecting several characters, hence making them an important tool in Chinese language learning. The significance of the frames cannot be overemphasised; however, there are obvious shortcomings associated with the frames in connection with the change in characters over time. These frames connect several characters, which therefore creates a frame for language learners to rely upon. This work analyses the different frames in Chinese character, their relevance, and limitations.

### **Reference**

- Boltz, W. G. (1994). *The origin and early development of the Chinese writing system*. New Haven, CT: American Oriental Society.
- Cao, W. (2001). *Hanyu Jingjiang*. Beijing University Press.
- Feng, L. (2023). Research on the formation of Chinese characters. *J Neuro Spine*, 1(1), 1–15.
- Huang, B., & Liao, X. (2011). *Xian dai Hanyu*. Higher Education Press.
- Jianyu, L., Zhou, J., & Mao, X. (2012). Analysis of the structure and function of phono-semantic compounds based on complex networks. *International Conference on Artificial Intelligence, ICAI*.