

Introduction to the Japanese Writing System

文字入門

by

Rafael E. Beermann

copyright 2006

Introduction to the Japanese Writing System

Historical background of Kanji

In the 5th century Japanese buddhist monks brought Chinese texts written in Chinese language to Japan. Since the Japanese language itself had no written form at that time, the Chinese characters (so-called **Kanji** or 漢字) were adapted. At that time the original texts would have been read according to the Chinese language. Even documents written by Japanese scholars were, so to speak, imitations of Chinese texts in grammatical, morphological, and syntactical respect, although Chinese and Japanese are completely different languages. Parallel to these efforts there were attempts made to detach kanji from all Chinese peculiarities and to use them as a tool of genuine rendition of Japanese (as a phonetic alphabet for a transcription).

Development of Hiragana

Later other efforts were made to re-adjust Kanji. A limited set of characters were used as purely phonetic representation of Japanese words, their meaning was ignored. Kanji written in curvilinear style is the origin of the syllabary **Hiragana**, which is still used as one of the three main Japanese writing system. In the following two examples of the development are shown: The Hiragana か which is pronounced “ka“ has been derived from the Kanji 加. The left part 力 was rounded, while the right part 口 was simplified to 丷. The origin of the Hiragana き “ki“ is the complex Kanji 幾. The whole round shape of き looks simple but also abstract. These examples make clear that the origin of some Hiragana can easily be assumed, whereas the source of other can not be so easily detected. Below you see other examples of the development of the syllabary Hiragana according to the contemporary system of ordering. Each of the 46 basic Hiragana has developed from equal number of Kanji by simplification of the whole form:

original Kanji	sound	development	syllabary Hiragana	sound
安	[a]	→	あ	[a]
以	[i]	→	い	[i]
宇	[u]	→	う	[u]
衣	[e]	→	え	[e]
於	[o]	→	お	[o]
加	[ka]	→	か	[ka]
幾	[ki]	→	き	[ki]
久	[ku]	→	く	[ku] etc.

Quiz: Which character could be a Kanji and which one could be a Hiragana ?

1 案 2 と 3 の 4 新 5 朝 6 下 7 す 8 し 9 時 10 せ



2,3,7,8,10 are Hiragana

Development of Katakana

Nearly at the same time, but completely independent of this movement another attempt was made to suit the Kanji to Japanese interests and conditions. Monastery students simplified some frequently used Kanji to single constituent elements. So another limited set of characters were used as a phonetic alphabet for a transcription of Japanese words. This system is the origin of the syllabary **Katakana**, which is also still in use. Each Katakana has an equivalent Kanji as its origin. Only *parts* of a Kanji and not the whole structure was taken as drafts to build this simple script and the rest was omitted. The development of some of the 46 basic Katakana is shown:

original Kanji	sound	development	syllabary Katakana	sound
伊	[i]	→	イ	[i]
江	[e]	→	エ	[e]
加	[ka]	→	カ	[ka]
己	[ko]	→	コ	[ko]
仁	[ni]	→	ニ	[ni]
比	[hi]	→	ヒ	[hi]
利	[ri]	→	リ	[ri]
呂	[ro]	→	ロ	[ro] etc.

Quiz: Which character could be a Kanji and which of them are Katakana ?

1ヌ 2水 3ト 4侍 5ノ 6人 7明 8シ 9森 10ヤ



1,3,5,8,10 are Katakana (No. 6 looks like a Katakana but it is a genuine Kanji !)

Difference between Kanji, Hiragana, and Katakana

It depends on the point of view to define Hiragana and Katakana as original Japanese written forms or as a taking apart of a handful of Kanji to a simple tool of phonetic transcription. Compared to the usually complex character of Kanji with its numerous strokes it is quite obvious that Hiragana and Katakana are characterized by their proportionally **simple** layout. While Hiragana are **round** in shape, Katakana have an **angular** form:

幾

き

キ

Kanji

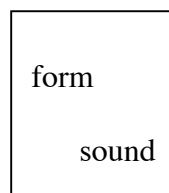
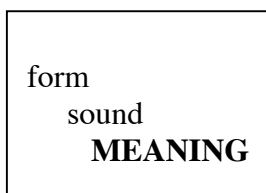
Hiragana

Katakana

But the decisive difference between Kanji and the two syllabaries can be revealed the following :

Kanji

Hiragana & Katakana



While Hiragana and Katakana are just abstract letters like the letters of the alphabet with their specific **form** (e.g. き or キ) and a specific **pronunciation** (in both cases [ki]) *without* a meaning, each Kanji has besides these two properties also a **meaning**.

The Kanji 幾, the prototype of the above mentioned Hiragana き has also to be read [ki], but it has additionally the meaning of “a certain number“. Correspondingly, all other Kanji have an ascribed meaning:

母mother 父father 心heart 侍Samurai 女woman 男man 魚fish 感feeling etc.

Quiz: Which are Kanji, which are Hiragana, and which are Katakana ?

1 神 2 ネ 3 ね 4 家 5 ヌ 6 馬 7 ぬ 8 龜 9 ま 10 マ 11 ウ 12 す 13 手 14 顔 15 ン



1,4,6,8,13,14 are Kanji; 3,7,12 are Hiragana; the rest is Katakana

Range of use

Nowadays all written Japanese texts are composed to **50% of Hiragana**, although its function is limited: e.g. for words for which no Kanji exist (particles, suffixes, and adverbs), for words where the Kanji is too difficult to read or remember, inflectional endings of verbs and adjectives, to give the pronunciation of Kanji in a small reading aid called *furigana*. (see more: Hiragana Handbook)

The quota of **Katakana** in contemporary average text is quite small (less than 6%).

It is most often used for transcription of words from foreign languages, for country names and foreign place and personal names, for onomatopoeia, letters used to represent sounds, technical and scientific terms, such as the names of animal and plant species and minerals. (see more: Katakana Handbook)

The **Kanji** finally are used for nouns, stems of adjectives and verbs, and Japanese names.

In some texts you also find **Roman letters** as a quotation and **Arabic numerals**.

A contemporary Japanese text usually consists of a mixture of Kanji, Hiragana and Katakana:

日本では、長さの単位はメートル (m) を、重さの単位はグラム (gr) を使っています。

K K h h K h h K K h k k k k h K h h K K h k k k h K h h h h h

The unit of length in use in Japan is the meter (m) and the unit of weight is the gram (gr).

Nihon de wa, nagasa no tan'i wa meetoru (m) o, omosa no tan'i wa guramu (gr) o tsukatte imasu

This sentence consists of

9 Kanji (K) :

日本長単位重単位使

15 Hiragana (h) :

ではさのはをさのはをっています

7 Katakana (k):

メートルグラム

3 Roman letters:

m g r

2 Japanese signs of comma: 、 、

1 Japanese sign of period: 。

0 Arabic numerals

Number of Kanji

The total number of Kanji is disputed. The largest Kanji dictionary of Japan the *dai kanwa jiten* 大漢和辞典 contains about 50,000 characters while recent mainland Chinese dictionaries contain 80,000 and more characters. But most of these are *not* in common use in Japan.

Japanese children learn **1,006** so-called *kyooiku kanji* 教育漢字 in six years of elementary school. Further the *jooyoo kanji* 常用漢字 which are **1,945** characters consisting of all the *kyooiku kanji*, plus an additional **939** Kanji are taught in junior high (3 years) and high school (3 years). Then, there are **983** more Kanji found in people's names, the so-called *jinmeiyoo kanji* 人名用漢字. But the Ministry of Justice adds over the years more Kanji to this special group. We can roughly guess, that there are approximately **3,000** Kanji in frequent use. In publishing much more Kanji are in use, but the characters outside this lists are usually given small reading aid in Hiragana. In the late 1970s about **6,300** Kanji were determined as a Japan Industrial Standard (JIS) for word-processors.

Why learning Kanji

Kanji are used to write the **core** of the Japanese words. In the above mentioned sentence the words *Japan, unit, length, unit, weight, and use* are written in Kanji, i.e. the essence of this sentence.

Japanese has a relatively simple sound structure, i.e. it has a very large number of **homophones** (words that sound the same like *I* and *eye*). If we did not have the chance to *write* these words differently we were not able to distinguish them. A very popular tongue twister *kisha ga kisha de kisha shita.* (*the reporter came back to his firm by train.*) demonstrates us clearly this problem.

These three *kisha* written in Kanji 記者 汽車 婦社 can make the sentence unequivocal while a notation in Hiragana きしゃ きしゃ きしゃ or Katakana キシャ キシャ キシャ would make the sentence virtually ununderstandable.

In case of a word with a *single* sound like [ki] the problem of homophones is much more disastrous:

木 wood, 鬼 devil, 季 season, 黄 yellow, 機 machine, 着 to wear, 気 soul, 記 chronicle, 期 period, 騎 horse rider, 旗 flag, 生 life, 城 castle, 姫 princess, 亀 turtle, 祈 prayer, 危 danger, and approximately **50** more words are pronounced all as [ki] !

Although there are many other practical reasons for the existence of Kanji in Japanese, after this striking encounter we have to say, already unrestrainedly:

Kanji are absolutely necessary !

Essence of Kanji

Each Kanji consists of a certain number of **strokes** starting with one (the Kanji for the number 1 一) up to sixtyfour. The **stroke order** (so-called *hitsujun* 筆順), the order and direction in which Kanji are written is strictly prescribed. The rules for stroke order were developed to facilitate writing *and* reading, but also to give a Kanji a balanced and esthetical appearance. A falsely written Kanji creates a completely different visual representation, and in some cases it can even be unreadable, i.e. you cannot decode it unequivocally.

Strictly speaking the strokes of Kanji are *not* straight geometrical lines drawn with a straightedge and such geometrically appearing characters are in any case quite rare:

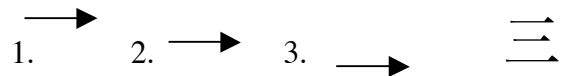
三 田 十 日 目 口 品 古

The majority of Kanji is composed of more or less curved diagonal lines, dots, and angular or winding strokes like:

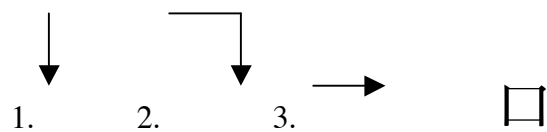
人 母 女 風 必 馬 龜 兔

Stroke Order

As a general rule, Kanji are written **from left to the right** and **from the top to the bottom**. The Kanji 三 mentioned above which means the number *three* is written in following manner:



The Kanji 口, which means the *mouth* (the sixth Kanji), is written in following way:



There are some more detailed rules you will learn during the writing lessons on the base of concrete examples like:

- horizontal lines are written from the left to the right and vertical lines are written from the top to the bottom
- horizontal before vertical
- when the character ends in a horizontal stroke at the bottom, vertical stroke is written before a horizontal
- vertical strokes that cut through a character are written after the horizontal strokes they cut through
- horizontal strokes that cut through a character are written last
- right to left diagonals are written before left to right diagonals
- vertical central strokes are written before vertical or diagonal outside strokes
- left outside strokes are written before right outside strokes
- outside enclosing strokes are written before inside strokes
- bottom enclosing strokes are written last
- left vertical strokes are written before enclosing strokes
- minor strokes (dots) are usually written last

In www.kanjistep.com you will find a wonderful visual demonstration of the stroke order.

If you are interested in far more detailed information, it will be very helpful to read the introductions of any Japanese-English character dictionaries. The one of the two poles among many others is the dictionary written in a traditional way by **Andrew N. Nelson** (1962) and its counterpart work might be the differently ordered dictionary by **Jack Halpern** (1990).

Writing Characters

The syllables Hiragana and Katakana as well as the Kanji should be written in the **same size**; i.e. the proportion of all letters must be unified although the three character types have different shapes and they even have different number of strokes:

三田十日目口品古人母女風必馬龜兎

but *not*: 三田十日 目口品 古人母女風必馬 龜兎

This is another example for the uniformity of writing of characters in a running text:

日本人の名前はほとんど漢字で書かれています。

Almost all Japanese personal names are written in Kanji.

but *not*:

日本人の名前はほとんど漢字で書かれています。

Classification of Kanji

The Kanji are classified into six categories since two thousand years. This traditional way of classification is problematic, since some of them are not clearly defined, some refer to the structure and some to usage. However we can adapt this classification partly to get an idea of the essence of Kanji:

pictograph 象形 (shookei)

These Kanji are **sketches** of the object they represent. The current form of the character is sometimes very different from the original, and it is now uneasy to discover the origin.

Pictographs constitute only an *inconsiderable* fraction of modern Kanji:

口 mouth 目 eye 人 human 女 woman 馬 horse 龜 turtle etc.

indicative 師事 (shiji)

These Kanji are also called *logograms*. They represent abstract matters such as numbers or directions. Indicatives make up also a *tiny* fraction of modern Kanji:

一 one 二 two 三 three 上 up 下 down 中 middle 半 half etc.

ideograph 会意 (kaii)

These Kanji are also called *compound ideographs*. These are a combination of pictographs that combined to produce an overall meaning. Ideographs constitute a *tiny* fraction of modern Kanji:

日sun + 月moon = 明 bright 口mouth + 鳥bird = 鳴 to twitter 口mouth + 犬dog = 吠to bark

semasio-phonetic 形声 (keisei)

These Kanji are by far the *largest* category (**90% !**). Typically (or more precisely ideally) they are made up of *two* components, one of which indicates the *meaning* (or the semantic category), and the other the *pronunciation* (the sound). The most ideal type of a *semasio-phonetic* Kanji consists of exactly *two* components, the *left* half indicates to which *category* or group this Kanji belongs to and the *right* half indicates how to *pronounce* it:

example of an ideal semasio-phonetic Kanji

<i>left half:</i>	<i>right half:</i>
semantic category	sound
金	同

Kanji: 銅
 meaning: copper
 sound: doo

The official name of the left part which indicates the meaning is the **radical** (a word that means the *root* or the *origin*). There are **214** historical radicals. A radical is comparable with a name of a group like a family name. For instance the radical or the group of metal 金 has a lot of members like:

銅 copper, 釘 spike, 銀 silver, 錆 rust, 鍵 key, 銭 coin etc.

With help of this radical we are generally able to guess the group or the semantic category of a Kanji and with the right part of those Kanji 同 丁 良 青 建 羹 (they are again genuine Kanji with a meaning and sound but they are used just as suppliers of the sound) you can identify the sound. Approximately 50% of the semasio-phonetic Kanji are created according to this easily noticeable system of the separation in left (radical) and right (sound) parts. But at the same time this means that the other 50% are aberrations! Many other parts can be a radical and nearly every Kanji contains from two to eight radicals. So the process of finding a Kanji in a dictionary, i.e. to find its meaning or its sound is often a time-wasting und discouraging trial-and-error process.

On- and Kun- reading

As we already know Kanji was imported first from China around the 5th century. During this process the Japanese incorporated both the original Chinese reading of the notions and their purely native Japanese reading for all things for which they had no written form.□

On-reading (*on-yomi* 音読み) is the Sino-Japanese reading, a rough approximation of the Chinese pronunciation of the character at the time it was introduced. Some Kanji were introduced from different parts of China at different times. The process of import lasted hundreds of years and so we have multiple *on-yomi*, and often multiple meanings.

On-yomi primarily occur in multi-Kanji **compound words**, many of which are the result of the adoption of Chinese words for concepts that either did not exist in Japanese or could not be articulated as elegantly using native words. This borrowing process is comparable to the English borrowings from Latin, Norman French, and Greek since Chinese-borrowed terms are often more specialized, or considered to sound more erudite or formal, than their native counterparts. The Kun-reading (*kun-yomi* 訓読み) is the native Japanese reading associated with the meaning of a Kanji. As with *on-yomi*, there can be multiple kun-readings for the same Kanji, and some Kanji have no *kun-yomi* at all.

Examples of *on-yomi* and *kun-yomi*:

Kanji	Meaning	on-yomi	kun-yomi
北	north	hoku	kita
南	south	nan	minami
西	west	sei/zai	nishi
東	east	too	higashi
京	capital	kyoo/kei	miyako

When to use which reading

Unfortunately there is no reliable rule when to use *on-yomi* or *kun-yomi*. You need to memorize the pronunciation on an individual basis, one word at a time. However, here are two facts that are worth remembering.□

On-yomi is usually used when the Kanji is a part of a compound (two or more Kanji are placed side by side).

Examples: The points of the compass *south-east* and *south-west* are always read as compounds *nantoo* 南東 and *nansei* 南西 but never *minamihigashi* or *minaminishi*. In the same way you would call the capital of Japan *tookyoo* 東京, but never *higahimiyako*.

Kun-yomi is used when one Kanji is used isolated on its own, either as a complete noun or as adjective stems and verb stems.

Example : If you want to say *east and west* you use *higashi to nishi* 東と西, but never *too to sei*.

These rules of thumb have many exceptions. *Kun-yomi* are quite capable of forming compound words, and on the other hand some *on-yomi* Kanji can also be used as words in isolation. And sometimes on- and kun-yomi may be even combined as on-kun or kun-on.

Examples: In the list of vocabulary of our first lesson you find the word *namae* 名前 which means *name*. It is a compound read in *kun-yomi*. (the *on-yomi* of 名 and 前 are *mei* and *zen*)

The single Kanji 本 (*book*) is read *hon*, which is indeed a *on-yomi*. (kun-yomi of 本 is *moto*)

The compound 仕事 (work) is a typical case of an *on-kun*. It is read as *shigoto* (shi=on, goto=kun) Fortunately, almost all other words written in Kanji follow the above mentioned rules of thumb.

Final Remarks

This rudimentary introduction to the Japanese writing system makes us after all clear, that the syllabaries Hiragana and Katakana are quite easy to learn in a few weeks. On the other hand learning of Kanji as a whole system is a lifework. The first 500 is the hardest hurdle. Although Kanji is not a pictograph it will be very helpful to use its strong *symbolic* quality by studying the etymology of each Kanji (this is the sophisticated and academic method), or you try to find personal eccentric cribs as crazy as possible in them. My suggestion for the Kanji 古 : it means *old, ancient*, “a transverse section of a coffin with a cross on the top“....