3

Ancient China

FRANÇOISE BOTTÉRO

As in many other cultures, the beginning of Chinese lexicography is rooted in the philological heritage. Interest in classical texts whose meaning had become difficult to understand engendered a significant number of explanations. Following or inserted into the texts, these explanations were subsequently collected to form the first glossaries and collections of synonyms. Later scholars would draw deep from within this rich exegetic tradition to create the first dictionaries.

Before Dictionaries: The First Wordlists

The practice of writing requires methods and then perhaps even manuals to learn how to write. In the case of ancient China, there was also the need to manipulate different styles of writing, depending on the type of documents to be produced. Nothing is known about how the earliest Chinese scribes were educated or what kind of documents they may have used to learn and practise.

Shǐzhòu Piān 史籀篇

The first primer for children of which there are records is the *Shǐzhòu piān* 史籀篇, traditionally attributed to the scribe of King Xuān 宣 of Zhōu, who reigned in the late ninth and early eighth centuries BC. The original text, which comprised fifteen chapters written in large seal script (*dàzhuàn* 大篆 or *zhòuwén* 籀文), has not survived.[1] By the Hàn dynasty (206 BC–AD 220), nine chapters were already lost, but it appears that some 225 graphs were chosen from the chapters which were still extant and included in *Shuōwén jiězì* (for which, see below). Based on the analysis of these characters and their graphic constituents, it is thought that the *Shǐzhòu piān* was probably

[1] Bān Gù, *Hànshū*, 30.1719.

https://doi.org/10.1017/9781316827437.004 Published online by Cambridge University Press

composed under royal command around the fifth century BC, in an effort to introduce an orthographic norm.[2]

Cāngjié Piān 倉頡篇

A series of other wordlists was produced when the first emperor, Qín Shǐ huángdì 秦始皇帝, conquered the Chinese territory in 221 BC. As part of the unification of the Chinese empire, he tried to impose unification of writing on all the ancient states that had developed their own graphic variants. His prime minister Lǐ Sī 李斯 compiled a manual of characters called *Cāngjié piān* 倉頡篇. The original text, in seven sections, consisted of rhyming sentences of four characters, and its study was strongly recommended for those who wanted to get a position in government. Two other texts, one in six sections and one in seven, soon followed, forming a larger work, which was still known as *Cāngjié piān*. With the Hàn dynasty the *Cāngjié piān* was rearranged and augmented for pedagogical needs. Different versions, some including explanations, were also produced. Most of them were lost, but fragments have been found in different places in China, suggesting the importance and the wide diffusion of the text at that time.

Jíjiù Piān 急就篇

Among the ten or so (mnemonic) wordlists produced during the Western Hàn and recorded in the bibliography of the great historical work *Hànshū*, only the *Jíjiù piān* 急就篇 is still available today.[3] This manual, conceived to help its readers learn a text of 2,016 characters rapidly, was written by Shǐ Yóu 史游, under the reign of King Yuán 元, in the second half of the first century BC. It consisted of thirty-two sections of sixty-three characters each, with different lengths of rhyming phrases. The words, very often disyllabic, were enumerated without being embedded in sentences. The vocabulary was organized thematically: family names, fabrics, colours, buying and selling, cereals, vegetables, metallic objects, manufactured objects, aquatic animals, women, servants, sleeping room objects, musical instruments, kitchen, food, the human body, weapons, charts, buildings, terms related to work in the fields, trees, animals, diseases, pharmacopeia, and terms from the religious and ritual domain. In the end, Shǐ Yóu gave a general presentation of government organization with officers' names and titles, a list of the texts officers were supposed to study, and the laws and regulations they should know. The *Jíjiù piān* ended with a paragraph to the glory of the Hàn.

[2] See Pān Yùkūn, '"Shǐzhòu piān" niándài kǎo'. [3] Bān Gù, *Hànshū*, 30.1719–20.

From the study of the *Jíjiù piān* as well as that of the various fragments of the *Cāngjié piān*, we can understand how these manuals were composed, and how they could have been influential for the lexicographic tradition that would arise a century or so later. Not only did they collect words, with close or opposed meanings, but they also tended to put together characters written with the same semantic constituent (a graphic element which enters into the formation of multiple characters, and is supposed to give them its meaning). In the bamboo slips C33–C34 of the *Cāngjié piān,* discovered in Fùyáng 阜陽 in 1977, or in chapter 11 of the *Jíjiù piān,* for example, we can see no fewer than nine characters with the semantic constituent *hēi* 黑 'black', and nineteen characters with the semantic constituent *jīn* 金 'metal', in a row. This no doubt gave Xǔ Shèn (see below) the idea of radical classification, which would later play a role in Chinese dictionaries equivalent to that of the alphabet.

Before Dictionaries: The First Collections of Glosses and Synonyms

Next to wordlists, collections of glosses or synonyms were also produced before and after the development of dictionaries in the second century. Three of them are still extant: the *Ěryǎ*, the *Fāngyán*, and the *Shìmíng*. The *Ěryǎ* was by far the most important lexicographic work of its time.

The Ěryǎ 爾雅 *'Approaching Perfection'*

The *Ěryǎ* is traditionally presented as the oldest Chinese dictionary. The author is unknown but, from the Eastern Hàn dynasty onwards, different dates have been proposed for its composition. It has been said, for instance, to have been written by disciples of Confucius (who died in the fifth century BC), or by the younger brother of the first duke of Zhōu (twelfth century BC). The content and the heterogeneous structure of the text suggest a much later date. The *Ěryǎ* gathers many expressions from the classics and pre-Qín (late third century BC) authors, but it also includes a certain number of terms and toponyms clearly linked to the Hàn. Thus, it should best be considered as a compilation regrouping different glosses, scholia, or texts written between the fifth and the first centuries BC, and compiled around the time of Emperor Hàn Wǔdì 武帝, whose reign ended in 87 BC.[4]

[4] Zhōu Zǔmó, *Wén xué jí*, 675.

https://doi.org/10.1017/9781316827437.004 Published online by Cambridge University Press

The text of the *Ěryǎ* which has been transmitted to the present day is 13,000 characters long and comprises 19 chapters. A preface is said to have existed, and a major early source counts twenty chapters, one of which was presumably the lost preface. One should distinguish two (or even three) parts in the *Ěryǎ*. The first three chapters present lists of words (nouns, verbs, adjectives, grammatical words, adverbs, and so on) with more or less the same meaning, whereas the last sixteen chapters resemble small encyclopedic treatises organized according to different themes and providing information concerning names and expressions related to each theme. In other words, the *Ěryǎ* combines three collections of synonyms (chapters 1 to 3) with sixteen topical glossaries (chapters 4 to 19).

The first chapter, *Shì gǔ* 釋詁 ('Explaining old words'), gathers 173 lists of words that can be found in ancient texts. Each of these lists is defined by a single and more common word that comes at the end of the list (that is, the bottom: Chinese texts were written from top to bottom): for example, '*Qí* 鬐 [鬐 is a rare character] is like *qí* 汽 "vapour".' The lists can include as many as thirty-nine 'synonyms' (with words used in different contexts to refer to the same notion), the longest being the third list of words corresponding to the meaning 'big'. The text starts with words referring to the meaning 'beginning' and ends with those referring to the meaning 'death':

> First, sprout [of a tree], head, basis, initiate, ancestor, primary, foetus, start, set, bud are [like] 'beginning' [*chū* 初, *zāi* 哉, *shǒu* 首, *jī* 基, *zhào* 肇, *zǔ* 祖, *yuán* 元, *tāi* 胎, *chù* 俶, *luò* 落, *quányú* 權輿, *shǐ yě* 始也] . . .
>
> Collapse, death of a prince, death, die, pass away, die are [like] 'die/death' [*bēng* 崩, *hōng* 薨, *wúlù* 無祿, *zú* 卒, *cúluò* 徂落, *yì* 殪, *sǐ yě* 死也].[5]

Next to formal, poetic, or ancient words, we find common as well as dialectal words. An example is a list of first-person pronouns, '*áng* 卬, *wú* 吾, *yí* 台, *yú* 予, *zhèn* 朕, *shēn* 身, *fǔ* 甫, *yú* 余, *yán* 言, *wǒ yě* 我也'.[6] Here, *áng* 卬 is a dialectal word corresponding to *wú* 吾 'we', as *yí* 台 corresponds to *yú* 予.[7] *Zhèn* 朕 was an ordinary southern dialect first-person pronoun in the *Songs of Chǔ* (*Chǔ cí* 楚辭, fourth–third century BC), before it became an honorific first-person pronoun in the third century BC. *Shēn* 身 'body' is a pseudo-pronoun used by a speaker to refer to himself. *Fǔ* 甫 is a courtesy name that is explained as referring to the first person in the *Books of Rites* (*Lǐjì* 禮記, fourth–second century BC). The pronoun *yú* 余, like *yú* 予 'we

[5] *Ěryǎ*, 1.1, 1.173. For matter of convenience I refer to the numbers Xú Cháohuá provides in his *Ěryǎ jīnshù* for each list.
[6] *Ěryǎ*, 1.43. [7] Xú Cháohuá, *Ěryǎ jīnshù*, 27.

[personally]', is often used by poets. *Yán* 言 'to speak' can be used to refer to the speaker(s) of the speech. The more standard *wǒ* 我 'we [the speaker's group]' is used as the *definiendum*.

Words have different meanings, and therefore they may appear more than once. In the next list, another meaning of *zhèn* 朕 and *yú* 余, which we have just seen as pronouns, is made clear as they are brought together with *gōng* 躬 'body, person' in a list of words equivalent to *shēn* 身 'person, ego, self'.[8] Moreover, words in the same list do not necessarily have the same meaning. In the next list, the six words *yí* 台, *zhèn* 朕, *lài* 賚, *bì* 畀, *bǔ* 卜, and *yáng* 陽 are all glossed with the character 予.[9] But this character records in fact two different words: *yǔ* 'give' and *yú* 'we'. Here, *yí* 台 and *zhèn* 朕, refer, as they did in the list discussed above, to the first-person pronoun. *Bǔ* 卜 is said to be used for 'we' in a verse of the *Book of Songs* (tenth–sixth century BC).[10] As for *yáng* 陽, it stands for the homophonous dialectal pronoun *áng* 卬 'we', which was discussed above. In between these four words equated with *yú* 'we', however, *lài* 賚 and *bì* 畀 both mean 'to give', and are being equated with *yǔ* 'give'. So, in this particular case, the *definiendum* has two meanings, and the list combines words with one meaning or the other. This potentially confusing list shows that the *Ěryǎ* was aimed at scholars who already had a good knowledge of ancient Chinese.

The second chapter, *Shì yán* 釋言 ('Explaining words'), gathers 280 smaller lists of synonyms of more ordinary use, such as '*yǒng* 泳 is like *yóu* 游 "to swim"' and '*móu* 謀 "to plan" is like *xīn* 心 "to think"'.[11] It is interesting to note that the second chapter ends with two words meaning 'end': '*mí* 彌 is like *zhōng* 終 "end"'. One is tempted to think that this second section could have been compiled as a development of the first one to form an independent set, and thus count as the third part of the *Ěryǎ*.

With its own independent structure, the third chapter, *Shì xùn* 釋訓 ('Explaining meanings'), gathers 116 lists. The first seventy-five lists present disyllabic expressions, taken from the classics (for instance, the *Book of Songs* and *Book of Documents*, eleventh–third century BC) or other pre-Qín and Hàn texts, which are defined by a monosyllabic word or a sentence. For example: '*Màomào* 懋懋, *mùmù* 慔慔, is *miǎn* 勉 "to make earnest effort".'[12] The expression *màomào* 懋懋 appears in the *Book of Documents*.[13] It had been

[8] *Ěryǎ*, 1.44. [9] *Ěryǎ*, 1.45.
[10] The verse is 卜爾萬壽無疆 'We predict for you a myriad years of life without limit' ('Tiān bǎo' 天保, in *Shísānjīng zhùshū*, 412b); in fact, the subject 'we' has been dropped in this verse, and *bǔ* 卜 is simply a verb meaning 'to divine, predict'.
[11] *Ěryǎ*, 2.97, 2.137. [12] *Ěryǎ*, 3.28; Xú Cháohuá, *Ěryǎ jīnshù*, 136.
[13] *Shísānjīng zhùshū*, 161b.

explained by the earlier Hàn commentator Kǒng Ānguó 孔安國 as *miǎn* 勉 'to make earnest effort' (*miǎnlì* 勉力 'to make effort', *quǎnmiǎn* 勸勉 'to encourage'), just as it is in the *Ěryǎ*. In another passage from the *Book of Documents*, Kǒng Ānguó also explains *mào* 懋 as *miǎn* 勉 'to make earnest effort'.[14] This example shows how glosses given by commentators on the *Book of Documents* or other ancient texts were culled to form the *Ěryǎ*. The end of the third chapter also gathers forty-seven expressions excerpted from ancient texts and explained: for example, '*shéi xī* 誰昔 is equivalent to *xī* 昔 "before, in the past"'.[15]

The other sixteen chapters of the *Ěryǎ* follow a different pattern since they are organized thematically, with the subjects of kinship, architecture, utensils, music, the heavens, the earth, hills, mountains, rivers, bushes and grasses, trees, insects, aquatic animals, birds, beasts, and domestic animals. Chapters 13 to 19 all gather essentially plant and animal names, formal as well as colloquial; as will be discussed in Chapter 6, some later works in the *Ěryǎ* tradition focused exclusively on this semantic domain.

In the sixteen topical chapters of the *Ěryǎ*, most of the formulae used to gloss words and expressions take the shape 'X corresponds to Y' or 'X is called Y' or 'X is like Y.' Thus, for example, '*A zōnggǔ* ["fishing net"] is called *jiǔyù*, a *jiǔyù* is a net for fishing' and '*Zǎi* is like *suì* "year". The Xià called it *suì* ["harvest"], the Shāng called it *sì* ["sacrifice"], the Zhōu called it *nián* ["harvest"], [the legendary emperors] Yáo 堯 and Shùn 舜 called it *zǎi* ["year"].'[16] Subthematic ordering is used to classifiy words and expressions in these chapters.

A close look at chapter 8 of the *Ěryǎ* shows how the topic of the heavens (more precisely, 'Explaining celestial [terms]') is organized under twelve subdivisions. This chapter first provides, under the heading 'Four seasons', the different names for the sky according to the four seasons; then secondly 'The auspicious signs' according to the four seasons; and thirdly 'The disasters', with the different terms for hunger engendered by different kinds of bad harvest. Fourthly to seventhly, under 'Names of Jupiter', it lists the names of the ten and the twelve years of ancient Chinese chronology according to the position of Jupiter in the sky; then the five 'Names of the year' according to the different dynasties (this passage was quoted above); the different names for the moon according to its ten positions in the sky; and the names for the months. Eighthly, the text presents under

[14] *Shísānjīng zhùshū*, 130b. [15] *Ěryǎ*, 3.113.
[16] *Ěryǎ*, 6.4, '縠罟謂之九罭；九罭，魚罔也' (Xú Cháohuá, *Ěryǎ jīnshù*, 178); *Ěryǎ*, 8.9, '載, 歲也. 夏曰歲, 商曰祀, 周曰年, 唐虞曰載' (Xú Cháohuá, *Ěryǎ jīnshù*, 201).

'Wind and rain' the different names of the wind according to the quarter from which it blows or to some of its particularities, and the names for frost, rainbow, hailstone, and thunderbolt, as well as different types of rain. Ninthly, under 'Constellation names', it provides names of different stars and constellations, grouped according to cardinal directions. The names for the different sacrifices are listed according to the four seasons and their beneficiaries (such as heaven, earth, mountains, rivers, constellations, and winds) in tenth place. Then we find the different names for hunting according to the four seasons, followed by two excerpts taken from the classical *Shījīng* (詩經 'Book of poems'), explained and commented on at length under the eleventh heading, 'Military training'. The first excerpt discusses the importance of offering sacrifices before letting armies march, and the second the order of seniority when going to or coming back from war. The twelfth and last part describes the different kinds of banners and flags.

In the *Ěryǎ*, the lexicographic unit is the word, and words are often disyllabic. Some of them have different meanings, others refer to different words, and characters very often stand for homophonous words. Thus, polysemy is partially taken into account as well as homonymy. The *Ěryǎ* was no doubt compiled to help the reading and understanding of the classics and the ancient texts. But the non-homogeneous structure of the text and the lack of systematic ordering makes it difficult to count the *Ěryǎ* as the first dictionary. Yet, as the first attempt to collect semantic glosses and discuss words out of their context (excerpts set apart), the *Ěryǎ* played a very important role for the begining of Chinese dictionaries.

The Fāngyán 方言 *'Regional Words'*

Another glossary of a different kind was produced at the beginning of the first century AD. It is called the *Fāngyán* 方言 ('Regional words'), which is an abbreviation of the full title, *Yóuxuān shǐzhě juédài yǔshì biéguó fāngyán* 輶軒使者絕代語釋別國方言 ('The imperial light carriage emissary explaining discursively the regional words in different states through the ages').[17] Compiled by Yáng Xióng 楊雄, the *Fāngyán* collected synonyms taken from different dialects and languages, gathered by court messengers who had been sent to various regions of China. It is the oldest known Chinese documentation on languages other than Chinese. The text, divided into 13 thematic chapters, contained more than 9,000 characters. Like the beginning of the *Ěryǎ,* the *Fāngyán* first presents lists of synonyms, before naming the area of use for

[17] *Fāngyán jiào jiān* 方言校箋 (1956) is an edition.

each word. It also mentions ancient versus contemporary words, and words of common usage all over the Hàn territory or among a speech community, as well as close or slightly modified pronunciations between dialects.[18] However, Yáng Xióng was in fact more concerned with the different ways to write words rather than with words per se.

The Shìmíng 釋名 *'Explaining Names'*

The *Shìmíng* 釋名 ('Explaining names') was apparently composed by Liú Xī 劉熙, in AD 200.[19] The particularity of this topical glossary lies in the general use of paronomastic glosses (also called puns or phonetic glosses, *shēng xùn* 聲訓) in order to clarify the supposed etymology of some 1,500 words. This method consisted in giving a more or less homophonous word with the entry to show a semantic link between them. These supposed motivations at the basis of the creation of words usually corresponded to folk etymology, as in the following examples.

> *Yuè* 'moon' is associated with *quē* 'lacking': it wanes after being full.
>
> When a man begins to discontinue breathing, we talk about *sǐ* 'death'. *Sǐ* 'death' is associated with *sī* 澌 'disappear': it is to move towards disintegration, *xiāosī*.
>
> *Zēngzǔ* 'great-grandfather': the ones below [i.e. the ones who died later] push the ones above [i.e. the ones who died earlier], so the position of the ancestors is removed and goes further up ['augments' [*zēngyì*] in the ancestral hierarchy].[20]

Liú Xī arranged the words into twenty-seven thematic sections: heaven, earth, mountains, rivers, hills, roads, regions and kingdoms, shapes and bodies, appearances, seniority, kinship, expressions, food, silks, ornaments, clothes, palaces, beds and curtains, written documents, classics and arts, tools, and musical instruments.

Compared to the preceding lexicographic works, if Liú Xī also gathered many glosses from ancient texts and commentators, his interest in the everyday use of words was quite original. He introduced colloquial words that one would hardly find in any other extant glossary, with senses such as 'caress', 'laugh', and 'latrines'. He also discussed the polysemy and the homonymy of some words. *Wàng* 望, for example, is not only explained as 'to look into the distance' in chapter 9, its deviant use referring to a kind of beam is also

[18] Harbsmeier, *Language and Logic*, 77.
[19] *Shìmíng huì jiào* 釋名汇校 (2006) is the standard edition.
[20] *Shiming*, 1.3, '月,缺也 ;滿則缺也'; 27.1, '人始氣絕曰死,死,澌也,就消澌也'; 11.6, '曾祖,從下推上, 祖位轉增益也'.

explained in chapter 17, and the meaning 'full moon' is discussed in the first chapter.[21] When using the formula 'A is associated with A' (A A 也), Liú Xī could record examples of homonymy, for instance '*Bù* "cloth" is associated with *bù* "spread"'(布布也).[22] He showed, moreover, that different words could refer to the same thing not only between dialects, or different periods, but also in what may be called the standard language. Last but not least, he expressed a special interest in the pronunciation of spoken words and went as far as referring to the way sounds should be articulated.

But the *Shìmíng* was too limited by the small number of entries Liú Xī could explain with his ad hoc phonetic glosses. It had little impact on later generations of dictionaries, which essentially focused on the meaning of characters in the classics.

The Invention of Dictionaries in China: *Shuōwén Jiězì*, and *Yùpiān*

The period of the *Ěryǎ*, the *Fāngyán*, and the *Shìmíng*, from around 100 BC to around AD 200, overlaps with that of the first Chinese texts which can uncontroversially be called dictionaries, namely *Shuōwén jiězì* (around AD 100) and *Yùpiān* (AD 543).

Shuōwén Jiězì 說文解字 *'Explain the Graphs to Unravel the Written Words'*

Shuōwén jiězì 說文解字 (abbreviated as *Shuōwén*), which preceded the *Shìmíng* by a hundred years, is the first dictionary of Chinese characters.[23] It was composed by Xǔ Shèn 許慎 around AD 100. For the first time, all the characters included in the work were presented according to a new system of classification invented by the author. This is a tremendous achievement in the history of this non-alphabetical writing system. Xǔ Shèn gathered as many as 9,353 characters and organized them according to 540 'classifiers' or semantic constituents (also called 'radicals' *bùshǒu* 部首) such as 'one', 'woman', 'jade', and 'aquatic animal'. It counted 14 chapters followed by a postface, and also included some 1,163 allographs. Each of the 540 radicals gathered between 1 (the radical only) and more than 400 headgraphs or characters. It appears that when

[21] *Shìmíng*, 9.29 and 12.112 ('look into the distance'), 17.25 ('kind of beam'), 1.73 ('full moon').
[22] *Shìmíng*, 14.14.
[23] Xǔ Shèn, *Shuōwén jiězì* (2006), is an edition – not of the lost original, but of the tenth-century recension by Xú Xuàn (see Chapter 6) which is used in its place.

many entries were collected under a single radical, a certain number of semantic series structured the text.[24] The headgraphs were presented in the small seal-style *xiǎo zhuàn* 小篆, an older style of writing on which Xǔ Shèn based his graphic analysis. For each entry, Xǔ Shèn provided a gloss as well as a graphic analysis into constituents, semantic, phonetic, or both. But he sometimes also gave supplementary information concerning meanings, pronunciations, allographs, or illustrative quotations, as well as encyclopedic material.

The originality of *Shuōwén* is that the lexicographic unit is the character, and no longer the word. Xǔ Shèn wanted to show the necessity of retrieving the proper way to write characters from older-style graphic forms. He was not interested in the basic meaning of words. When he glossed *wǔ* 五 'five' as '[for example] the Five Elements', he did not define it as a number, but instead provided an illustrative quotation, which from our modern perspective corresponds to the mistake of using the *definiendum* in the *definiens*. And then when he said in his graphic analysis that '[The graph 五] has "two" as a semantic constituent, with the Yin and the Yang intersecting between Heaven and Earth', we can see that he wanted to make a connection between the Five Elements and the Yin and Yang theory.[25] In his dictionary, Xǔ Shèn tried his best to define the meaning of a graph according to its semantic constituent. In other words, characters with 'heart' as a semantic constituent were usually defined in terms of psychology, and those with, for example, the semantic constituents 'woman', 'jade', or 'aquatic animal', in terms of those constituents. The use of the older small seal-style graphs allowed Xǔ Shèn to recover graphic constituents before they were modified in modern script and to provide what he saw as a correct analysis of graphs, which could clarify their supposed original meaning in the classics. Thus, *Shuōwén* is not a dictionary of the meaning of words. It is a graphic etymological dictionary in which Xǔ Shèn tried to provide the meaning that best suited the graphic structure of the graph (written word) and its immediate constituents. Xǔ Shèn came close to composing a dictionary of the orthography of words.

This being so, we can understand why polysemy is scarcely touched upon in *Shuōwén*: the supposed original meaning of a graph was unitary. So, Xǔ Shèn usually only gave one gloss for each headgraph. He drew many of his glosses from the rich exegetic tradition, sometimes mentioning his sources (including the *Ěryǎ* and commentators on classical texts), but often keeping silent about them. Thus, we find all sorts of glosses in *Shuōwén*: near-

[24] Bottéro and Harbsmeier, *Chinese Lexicography on Matters of the Heart*.
[25] Xǔ Shèn, *Shuōwén jiězì*, 14B 7b, '五 五行也。从二陰陽在天地閒交午也'.

synonymic glosses ('*Zhōng* "loyal" is [a way of] showing respect'); analytical glosses ('*Lì* "minor official" is someone who keeps order among others'); dialectal glosses ('In Yǎnzhou, they refer to cheating by saying *tuó*'); paronomastic glosses, usually followed by an explanation ('*Jiǔ* "fermented wine" is [evocative of] *jiù* "approach". It is the means by which one approaches the good or evil nature of man'); technical glosses ('A *zǔ* is an auspicious relief on a cóng-type jade'); geographic glosses ('Yǐng is the ancient capital of the Kingdom of Chǔ; it is located 10 li north of the south province of Jiānglíng'); and others.[26] Xǔ Shèn not only dealt with lexical items, but also included proper names in his work. For example, under the radical 'water', the 464 entries are divided into 2 sections. The first section deals essentially with the names of rivers, and the second with lexical items. The rivers are neatly arranged according to the four cardinal directions and the centre, with all the provinces of his time included. Thus, this section dedicated to rivers offers a kind of encyclopedic treatise on Chinese rivers. In other words, *Shuōwén*, which incorporates features of an organized thematic encyclopedia, is much more than a dictionary of orthography and graphic etymology.

Following the *Ěryǎ,* but in a more systematic way, Xǔ Shèn explained words out of their context (although in an orally based culture, the context was probably more present in the reader's mind and recognizable than it may seem to us). Xǔ Shèn also provided a method to organize the entries equivalent to the alphabet in European dictionaries. Yet looking for a character in *Shuōwén* was not an easy task, for the radical classification was so much intertwined with the Hàn philosophical context (for instance, the Five Elements, and Yin and Yang) and Xǔ Shèn's own worldview. For example, for the first character and radical, *yī* 一 'one', Xǔ Shèn referred to traditional Chinese cosmogony, which placed One at the very beginning of the creation of the Universe, and thus did not gloss it as a number but as the metaphysical beginning of all things (characters included): 'Initially, at the great beginning, the Way established through the One. By separation it created Heaven and Earth, and [thus] transformed so as to bring to completion the Myriad Creatures.'[27]

In the postface, Xǔ Shèn provided the first account of the history of the Chinese script, explaining what preceded it and how it came to be invented by the scribe of the legendary Yellow Emperor. He also touched upon the 'Six ways of writing down (words)' *liùshū* 六書: a pedagogical method,

[26] Xǔ Shèn, *Shuōwén jiězì*, 忠 敬也; 吏 治人者也; 詎 沇州謂欺曰詎; 酒 就也, 所以就人性之善惡; 玭 琮玉之璋; 郢 故(古)楚都。在南郡江陵北十里.

[27] Xǔ Shèn, *Shuōwén jiězì*, 1A 1a, '惟初太始, 道立於一, 造分天地, 化成萬物'.

traditionally used to introduce writing to eight-year-old children. Four among these 'six ways' – 'symbolize physical shape' *xiàng xíng* 象形, 'refer [pictorially] to something' *zhǐ shì* 指事, 'indicate shape and sound' *xíng shēng* 形聲, and 'associate ideas' *huì yì* 會意 – referred to character creation, whereas the last two – 'borrow [one graph for another]' *jiǎ jiè* 假借 and *zhuǎn zhǔ* 轉注 – referred to the re-utilization of the others. But it is not clear what exactly *zhuǎn zhǔ* 轉注 or 'refer (pictorially) to something' *zhǐ shì* 指事 mean.[28] If Xǔ Shèn occasionally mentioned that a character 'symbolizes physical shape' *xiàng xíng* 象形 in his work, he never characterized characters in terms of *liùshū*. The *liùshū* were in no way imposed in the body of his dictionary, as some scholars would imagine a few centuries later (see Chapter 6). The importance of the *liùshū* lies in the fact that phonetic constituents in the characters were recognized for the first time. And this, no doubt, made Xǔ Shèn's graphic analysis into semantic and phonetic constituents possible.

Shuōwén was not a book to consult; it was a text one would read from the beginning to the end or memorize, until indexes were added much later to facilitate its consultation. It had a tremendous influence on the dictionaries of later generations. We know that, among the lexicographic works compiled between the second century and the fourth that have not survived, some copied its organization. The *Zìlín* 字林 ('Forest of characters') written by Lǔ Chén 呂忱, for example, gathered 12,824 entries and used the 540 radicals to classify the characters. It was composed to complete *Shuōwén*. The author sometimes added the pronunciation of the entries using methods such as *fǎnqiè* 反切 (for which, see Chapter 6), or *zhíyīn* 直音 (which consisted in providing a homophonous character). Yet none of these character dictionaries reached the level of the *Yùpiān* 玉篇, to which we now turn, in terms of lexicographic improvement.

The *Yùpiān* 玉篇 'The Jade Chapters'

One of the first real Chinese meanings dictionaries is the *Yùpiān* 玉篇, composed in 543 by the very learned and talented Gù Yěwáng 顧野王. The original text, which counted 16,917 headgraphs for 30 chapters, plus a preface, was already lost by the Sòng dynasty (960–1279), and was replaced by the augmented version of Chén Péngnián 陳彭年, *Dàguǎng yìhuì Yùpiān* 大廣益會玉篇 (1013), which included 22,561 entries.[29] Many versions of the *Yùpiān*

[28] See Qiú Xīguī, *Wénzìxué gàiyào*, 100–1.

[29] Chén Péngnián, *Sòng běn Yùpiān* (1983), is an edition of this recension; see also Gù Yěwáng, *Yuánběn Yùpiān cánjuàn* (1985).

were produced as early as the seventh century, with more or fewer modifications. Among them Sūn Qiáng's 孫強 *Yùpiān* (674) and the *Dàguǎng yìhuì Yùpiān* are the most famous. But there were also a Buddhist and a Taoist as well as a Korean and a Japanese version of this work.

Luckily, more than a tenth of the original *Yùpiān* was found at the end of the nineteenth century in Japan, where it had been imported around the end of the ninth century. The study of these fragments (which include sixty-three radicals from seven chapters) shows little change in the organization of the dictionary between the original and the extant version, but, as we shall see, drastic cuts in the presentation.[30] Gù Yěwáng took over Xǔ Shèn's system of classification, but modified it more than it seems. Among the 540 radicals used by Xǔ Shèn, he suppressed 10, and added 12 new ones, which made a total of 542 radicals. As we can tell from the eleventh-century augmented version, the first and the last entries (or radicals) were faithfully reproduced, but the organization of the radicals in the chapters was completely modified to provide a thematic organization. It started with radicals related to the subject 'heaven', followed by 'earth', 'man', 'kinship terms', 'parts of the body', 'buildings', 'plants', and 'musical instruments', and proceeding through other categories including 'animals', to end with 'numbers'. Gù Yěwáng was not interested in the structure of the characters, as Xǔ Shèn had been, but in their meaning. Therefore, he only used the current *kǎishū* 楷書 style to write all the entries (including ancient allographs in the *zhòuwén* and *gǔwén* scripts taken from Xǔ Shèn). One can incidentally notice the importance of *Shuōwén* in the eyes of Gù Yěwáng, who often referred to it. The author of the *Yùpiān* also referred to other lexica or glossaries, many of which have disappeared – for instance, the *Ěryǎ, Fāngyán*, *Cāngjié piān*, *Pí Cāng* 埤蒼, *Guǎngyǎ* 廣雅, *Zìshū* 字書, and *Shēnglèi* 聲類 – for the way they defined words or wrote them differently.

The comparison between the Sòng version and the original version of the *Yùpiān* is very instructive if we are to understand the level of development of Gù Yěwáng's original work. In the case of the entry *diǎn* 典 'canon, law, decree, document, classic', for example, the Sòng version gives only the pronunciation with the *fǎnqiè* 反切 spelling method, and a very simple semantic definition. By contrast, the original *Yùpiān* offered a much more complete and elaborated definition of this term in the classics, focusing on polysemy. Gù Yěwáng first gave the pronunciation with the *fǎnqiè* method. Then, using various examples, he presented the different meanings of *diǎn* 典

[30] Bottéro, *Sémantisme et classification*, 96–105.

in the classics followed by commentators' glosses. The first example is from *The Book of Documents*: 'He had *diǎn* 典, he had patterns.' It is followed by a gloss by Kǒng Ānguó which says '(*Diǎn* 典) refers to the canons.' The second and third examples are from the *Rites of Zhōu* (third century BC), 'He is in charge of the six *diǎn* 典' and 'Two middle-rank servants are attributed to the rulers' wives 典婦', with another Hàn commentator's glosses explaining *diǎn* 典 in these passages as equivalent to *cháng* 常 'law' and *zhǔ* 主 'ruler, direct' respectively. Gù Yěwáng then introduced his own reading of a classical text: 'I, Yěwáng, observe that when "King Shùn commanded Bóyí 伯夷 to *diǎn* 典 his three rites" and "Xià to *diǎn* 典", *diǎn* 典 corresponded to this meaning', namely, 'direct'.[31] Then he referred to a commentary on another text which defined *diǎn* 典 as *fǎ* 法 'law, statute'.[32] He quoted the *Ěryǎ* for the gloss '*diǎn* 典 is like *jīng* 經 "canon, classic"', and *Shuōwén jiězì* for the statement that '*Diǎn* 典 is like the documents of the five Emperors. [The graph] has *cè* "bound documents" on a small table, where they are placed respectfully, as semantic constituents. Another explanation says "*diǎn* 典 are great documents".' Returning to the first person, he added 'I, Yěwáng, observe that in the *Book of Documents*, there are the Yáo *diǎn* 堯典 and the Shùn *diǎn* 舜典. According to Kǒng Ānguó's explanations "they can represent the eternal way/conduct of the preceding generations".' Finally, he turned to other ways of writing the character 'canon' and the character representing the word 'direct': 'The ancient graph is written 簨 and is classified under the bamboo radical. As for *diǎn* 敶 meaning "direct" it is classified under the radical *pū* 攴.'

We can see that Gù Yěwáng gathered all the meanings of the character and the word *diǎn* 典 in the classics. He did not limit his presentation to concrete examples taken from the classics: he also referred to their commentators, added his own opinions, and reproduced glosses from older lexicographic works such as the *Ěryǎ* and *Shuōwén jiězì*. But what is probably more original is that he went as far as providing cross-references between entries within his dictionary. At the same time, his new thematic organization of the radicals facilitated looking for a character in the dictionary, and so did the number attached to the radicals as well as the total number of entries they gathered. The author of the *Yùpiān* intended his dictionary to be easy to consult. Thus, compared with previous lexicographic works, we can see in the original *Yùpiān* a clear step in the direction of a real dictionary. But the *Yùpiān* was

[31] The texts cited are from the *Book of Documents*, in *Shísānjīng zhùshū*, 131b, 134b.
[32] The original text is a passage in the *Book of Songs* which refers to 'The statutes of Wen Wang' (*Shísānjīng zhùshū*, 584c, 588b).

not a dictionary of contemporary meanings of words; it too was limited to the meanings or characters in the classics.

We can only regret that such a masterpiece was lost. Gù Yěwáng's dictionary had a strong influence on later works, such as *Lèipiān* 類篇 (1066), *Zìhuì* 字彙 (1615: see Chapter 6), and *Zhèngzìtōng* 正字通 (1680). Even in Japan, the famous Buddhist monk Kūkai 空海 copied it to write the *Tenrei banshō meigi* 篆隸萬象名義 (see Chapter 10), but suppressed the quotations from the classics as well as Gù Yěwáng's observations. For a long time the term *Yùpiān* was used in the general sense 'dictionary'.

A New Type of Dictionary: The Rhyme Dictionaries

With the growing importance of literary composition, different types of books providing the pronunciation of characters were produced. Some of them distinguished a certain number of rhymes or finals under which they classified the characters.[33] Most of the earlier texts have been lost, apart from the *Qièyùn* 切韻 (see Chapter 6), which was copied, recopied, and augmented, and had an immense influence, so as to create a new genre.

Conclusion

All the lexicographic works I have introduced here were compiled to help in the reading of classical texts or the composition of poetry. Thus, except for the *Shìmíng*, none of them included a wide range of everyday words. The discovery of lexica dating from the Táng dynasty (618–907), at the beginning of the twentieth century in the desert of Dūnhuáng, shows that lists gathering everyday vocabulary according to thematic categories did in fact exist, and that dictionaries did not draw on these. Until the sixth century, monolingual dictionaries were the rule. The *Fāngyán* included non-Chinese items and thus constituted an interesting exception until the translation of the Buddhist canon from the fourth century onwards led to the compilation of certain bilingual lexica (Sanskrit/Prakrit/Pali–Chinese). However, these multilingual dictionaries dealt with a much smaller part of the total vocabulary than the older monolingual dictionaries.

With the increase in characters included in dictionaries from *Shuōwén* onwards, the need was felt to facilitate the retrieval of given entries. When

[33] Bottéro, 'Le developpement des livres de rimes en dictionnaires'.

Xǔ Shèn invented the radical system of classification, he did not think of it as a way to find characters in his dictionary (although for all we know he may well have come to notice the advantages of the radical system for retrieval of lexical entries). He was primarily looking for a system to organize the characters which represented the structured realities of the world in his eyes. Gù Yèwáng considerably modified the system to make it easier to consult. The new system of classification on the basis of rhymes employed in the *Qièyùn* would offer a much more efficient way of retrieving lexical entries.

With the radical and rhyme systems of classification, lexical entries could not be words but had to be characters. However, since characters constitute the units of the Chinese writing system, and since characters write the roots of compound words, looking up characters was of great help for finding the meanings of even complex words consisting of more than one character.

https://doi.org/10.1017/9781316827437.004 Published online by Cambridge University Press