

# John Benjamins Publishing Company



This is a contribution from *Writing Systems, Reading Processes, and Cross-Linguistic Influences. Reflections from the Chinese, Japanese and Korean Languages*.

Edited by Hye K. Pae.

© 2018. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: [www.copyright.com](http://www.copyright.com)).

Please contact [rights@benjamins.nl](mailto:rights@benjamins.nl) or consult our website: [www.benjamins.com](http://www.benjamins.com)

Tables of Contents, abstracts and guidelines are available at [www.benjamins.com](http://www.benjamins.com)

## CHAPTER 9

# Introduction to the multi-script Japanese writing system and word processing

Terry Joyce and Hisashi Masuda

Tama University, Japan / Hiroshima Shudo University, Japan

The general consensus among writing-systems researchers is that the Japanese writing system (JWS) is remarkably complex (Joyce, 2002a, 2011). This introductory chapter consists of two main parts that, respectively, provide an overview of the multi-script JWS and a selective review of psycholinguistic research on Japanese visual word processing. More specifically, after outlining its historical development, Part 2 focuses on the contemporary JWS and on highlighting the complex conventions that simultaneously underlie how the component scripts are employed together in essentially complementary ways while effectively sanctioning its pervasive levels of orthographic variation. In contrast, the shorter Part 3 reflects on how JWS's complexity both poses certain challenges and affords unique opportunities for investigating the complicated interactions involved in word processing.

**Keywords:** Japanese multi-script writing system, Japanese word processing, orthographic conventions, orthographic variation

### 1. Introduction

Given the common consensus among scholars of writing systems that the Japanese writing system (JWS) is so remarkably complex that it is essentially unrivalled among both historical and modern writing systems (Joyce, 2002a, 2011), it becomes almost inconceivable to start an introduction to the JWS without at least briefly noting just a few of their descriptions. Although varying somewhat in candour, the following highly selective sampling is generally representative. For example, in contrast to the directness of DeFrancis (1989: 138) who claims that the Japanese “ended up with one of the worst overall systems of writing ever created”, Coulmas (1989: 122) comments that the JWS “is often said to be the most intricate and complicated writing system ever used by a sizable population”. Similarly, while Fischer (2001: 167) asserts that the JWS's mixture of scripts that are “written

together following arbitrary rules perhaps embody the most complicated form of writing ever devised”, Sproat (2010: 47) remarks that “Japanese is a complex system, certainly the most complex writing system in use today and a contender for the title of the more complex system ever”. And yet, while it is undeniably true that the JWS is rather complicated in nature, it should also be equally acknowledged that, as Yamada (1967) astutely suggested some time ago, the JWS is, by sheer dint of its complexity, also unquestionably of unique importance for its potential to challenge and extend our understandings of writing systems, of written language and even of language itself.

Presenting an introductory overview of the JWS, this chapter consists of two main parts. Commencing with an historical outline, the more substantial Part 2 focuses on describing the contemporary JWS and its component scripts. In contrast, the more compact Part 3 offers a highly selective review of psycholinguistic research on Japanese visual word processing.

## 2. JWS

### 2.1 Historical development of the JWS

Reflecting the conspicuous rarity of truly independent inventions of writing systems throughout human history (i.e., just Sumerian, Chinese and Mayan), obviously, it is merely inevitable matters of geographical proximity and ancient circumstance that the initial emergence of writing within Japan – earliest extant texts being 古事記 /ko-ji-ki/ (712 CE) and 日本書紀 /ni-hon-sho-ki/ (720 CE) (within glosses, hyphens mark kanji-kanji boundaries; periods mark kanji-hiragana boundaries) – was wholly dependent on the Chinese writing system (albeit as mediated via the Korean peninsula) (Coulmas, 1989; Lurie, 2012; Miller, 1967). However, that historical twist has also had far-reaching ramifications that have reverberated down in shaping the contemporary JWS.

Miller (1967: 92) has remarked that the Chinese writing system was generally ‘admirably’ suited to the Chinese language, but, as Lurie (2012: 163) muses, if one were to create an experiment to investigate script adaption for profoundly different languages, “it would be difficult to find a more vivid case of linguistic contrast than that provided by Japanese as it comes into contact with the Chinese script”. Accordingly, the Japanese had to adapt the Chinese writing system consisting of Chinese characters, referred to as 漢字 /kan-ji/ (literally, ‘Chinese characters’) in Japanese, to function in orthographically representing their language. And, as Lurie (2012) convincingly argues, the core adaptive technique utilized was 訓読 /kun-doku/ ‘reading by gloss’. At some risk of oversimplification, kundoku essentially entailed associating kanji with Japanese words and then rearranging

their order to read them according to Japanese syntax (Habein, 1984; Lurie, 2012; Miller, 1967). Moreover, as Lurie (2012) also incisively observes, while there are some parallels with conventional translation, there are also crucial differences in the notions of generation and production. A key insight is that, while primarily a reading practice for rendering a Chinese text interpretable as Japanese, *kundoku* was also, in reverse, a method for implementing written language.

The *kundoku*-mediated account is particularly appealing in explaining how a number of key aspects of the JWS emerged and meshed together within an overarching system. Principal among them is the origin and the continuity of the dual-reading, or dual pronunciation, system of both 訓読み /*kun-yo.mi*/ ‘native-Japanese pronunciations’ and 音読み /*on-yo.mi*/ ‘Sino-Japanese pronunciations’ being associated with kanji. Mirroring the basic morphographic relationship between Chinese characters and Chinese language morphemes, no leap of imagination is required to readily grasp how kanji would have been linked to existing native Japanese morphemes for the same things. For instance, borrowing Coulmas’ (1989) example, it was quite natural that the Chinese character of 人 meaning ‘person’ would have been *glossed*, or became associated, with the Old Japanese morpheme of /*fitō*/, which underlies the modern Japanese *kunyomi* of /*hito*/. However, *kundoku* also had a supplementary counterpart known as 音読 /*on-doku*/ ‘reading by sound’, where characters were read according to Sino-Japanese approximations of their Chinese pronunciations. Although traditional accounts of the adaption process (i.e., Miller, 1967) have assigned a more pivotal role to *ondoku*, current evidence accords greater precedence to *kundoku* (Lurie, 2012). Initially, *ondoku* was only employed rather haphazardly on a lexical basis, such as a reader’s stylistic whim to eschew a particular *kunyomi*, but it steadily became the natural resort for coping with the considerable influx of Chinese loanwords into the Japanese lexicon by the eighth century (Lurie, 2012). Hence, 人 is also associated with a modern *onyomi* of /*jin*/, based on the Old Chinese pronunciation of /*jen*/.

Another particularly compelling piece of the *kundoku*-mediated account relates directly to the emergence of the syllabographic kana scripts during the ninth century. The keen realization is that, although the two kana scripts developed from separate practices, both scripts originated out of a shared core strategy of using Chinese characters as phonographs – ignoring their semantic referents to use just for their phonetic values – that “evolved in order to record *kundoku* readings and only subsequently came to be employed more independently” (Lurie, 2012: 174). Crucially, however, rather than becoming completely autonomous and replacing kanji usage, the kana scripts essentially remained as complementary components of the overall contemporary multi-script JWS. Similar to how the Chinese orthographically represented foreign names, the phonographic strategy was employed in the 万葉集 /*man-yō-shū*/ anthology of Japanese verse (759 CE), and, retrospectively, the term 万葉仮名 /*man-yō-ga-na*/ has often been applied to the set of phonographic

characters used at that time (Lurie, 2012; Miller, 1967; Shibatani, 1990). Over time, kundoku conventions came to include the annotation of Chinese texts with man'yōgana to indicate kunyomi and Japanese grammatical elements, particularly by Buddhist priests engaged in textual studies. However, their scribal practices fostered abbreviations of the man'yōgana, usually by emphasizing a character's distinctive feature, that eventually developed into the contemporary 片仮名 /kata-ka-na/ script. In contrast to abbreviation for katakana, the process for 平仮名 /hira-ga-na/ was one of cursive writing, where, by the early Heian period (794–1185), man'yōgana evolved into a form known as 草仮名 /sō-ga-na/ 'grass style', which, in turn, eventually lead to the even more simplified and cursive forms of contemporary hiragana (Habein, 1984; Lurie, 2012).

Yet another key historical development has been the contemporary policies generally aimed at restricting the number of kanji in daily use (Seeley, 1984, 1991; Twine, 1991). Since the mid-twentieth century, the Japanese government has issued a series of guidelines concerning kanji usage; the first was the 当用漢字表 /tō-yō-kan-ji-hyō/ list of 1,850 kanji issued in 1946, which was followed by the 常用漢字 /jō-yō-kan-ji/ 'characters for general use' list of 1,945 kanji in October 1981, which was revised to include 2,136 kanji in November 2010 (Bunkachō, 2010).

Admittedly, these few fleeting observations are inadequate to fully convey the sense of continuity that characterizes the historical development of the JWS, but its quintessentially multi-script nature was established relatively early during the seventh and eighth centuries and has endured to the present day (Lurie, 2012; Martin, 1972).

## 2.2 Contemporary JWS and its multi-scripts

Generally referred to as 漢字かな交じり文 /kan-ji.kana.ma.jiri.bun/ 'mixed kanji and kana writing', the standard orthographic conventions for the contemporary JWS are to employ its multi-scripts – namely, morphographic kanji (Joyce, 2002a, 2011, 2016), the two syllabographic scripts of hiragana and katakana, the phonemic alphabet of ローマ字 /rōma-ji/ 'Roman alphabet', and 数字 /sū-ji/ 'Arabic numerals' – in essentially separate and complementary ways in representing the Japanese language in writing. These conventions effectively function in visually differentiating between content words – usually represented in kanji but often in katakana and occasionally in rōmaji – and grammatical elements – mainly in hiragana – and also in distinguishing, to a lesser degree, between the main lexical strata of Japanese – usually native-Japanese in kanji and hiragana, Sino-Japanese in kanji, and foreign-Japanese in katakana and increasingly in rōmaji in some domains (Gottlieb, 2008; Igarashi, 2007; Kess & Miyamoto, 1999; Joyce, 2011; Joyce, Hodošček, & Nishina, 2012; Joyce, Masuda, & Ogawa, 2014; Smith, 1996; Taylor & Park, 1995; Taylor & Taylor, 2014; Tranter, 2008).

Naturally, the proportions by which the multiple scripts are mixed in any given text depend on a variety of sociolinguistic factors, including the nature of the content and its context, reflecting intended audiences and publication format, as well as an author's stylistic freedoms and preferences (Joyce & Masuda, 2016, 2017). An initial rough sense of the mixture can be gained from a study by Chikamatsu, Yokoyama, Nozaki, Long and Fukuda (2000) which analysed one year (1993) of Asahi Newspaper's articles to find that, of the approximately 56.6 million character tokens, kanji accounted for 41.38%, hiragana for 36.62%, katakana for 6.38%, punctuation and symbols for 13.09%, Arabic numerals for 2.07% and Latin alphabet for 0.46%. However, as character-token data alone is not so informative, a complementary perspective on typical script proportions and their latitudes can be gained from Igarashi's (2007) analyses of a word-type list extracted from nine magazines, three newspapers and some TV commercials (depending on which of the National Language Research Institute's historic word definitions is employed, Igarashi's list contains either 22,612  $\beta$  unit words (basically morphemes and simplex words) or 7,351  $\alpha$  unit words (complex words)). Across the three genres, the average percentages for the four script-types were 60.72% kanji words, 20.51% hiragana words, 12.69% katakana words and 6.09% alphabetic symbols and numbers, but the percentages also varied noticeably for the different genres. At one extreme, the newspaper sub-list consisted of 72.33% kanji words, 18.24% hiragana words, 5.73% katakana words and 3.81% alphabetic symbols and numbers, while, at the opposite extreme, the TV commercials sub-list consisted of 51.38% kanji words, 20.51% hiragana words, 17.35% katakana words and 10.80% alphabetic symbols and numbers. Moreover, while the average percentages for the magazine sub-list fell between these divergent trends, Igarashi also reported considerable degrees of script-type variation across the range of sampled magazines.

Obviously, no single authentic Japanese sentence can encapsulate all aspects of the JWS, but the sentence in Figure 1 is a fairly illustrative example taken from the Japanese Wikipedia entry for JIS X 0208 (a Japanese Industrial Standard (JIS) character set that will be mentioned again shortly).

---

**JIS X 0208**(ジス X 0208)は、日本語表記、地名、人名などで用いられる6,879図形文字を含む、主として情報交換用の2バイト符号化文字集合を規定する日本工業規格である。

jisu ekusu rei-ni-rei-hachi (...) wa, ni-hon-go-hyōki, chi-mei, jin-mei nado de mochi.irareru rokusen-happyaku-nanajū-kyū zu-kei-mo-ji o fuku.mu, shu.toshite jō-hō-kō-kan-yō no ni.baito fu-gō-ka-mo-ji-shū-gō o ki-tei.suru ni-hon-kō-gyō-ki-kaku dearu.

JIS X 0208 is a Japanese Industrial Standard that stipulates the 2-byte encoded character set primarily utilized for information exchange, which includes 6,879 graphic characters used for Japanese language writing, place names and personal names, etc.

---

**Figure 1.** Example of Japanese sentence, with phonological gloss and translation

As noted already, kanji are generally used to represent native-Japanese and Sino-Japanese content words, including nouns, the stems of verbs and of some adjectives, and some adverbs. For instance, consistent with the morphographic nature of kanji (Joyce, 2002a, 2011), the first string of five kanji in Figure 1 is a polymorphemic word, which consists of 日本 ‘Japan’ + 語 ‘language’ + 表記 ‘representation, writing’, and one of the native-Japanese verbs is 含む /fuku.mu/ ‘include’, where the stem is represented by one kanji and the present-tense inflection is represented by one hiragana character. Hiragana are usually used to represent functional words, including the copula verb and the auxiliary する /suru/ ‘do’ verb, inflectional elements of verbs and some adjectives (when they are referred to as 送り仮名 /oku.ri.ga-na/), grammatical case markers and conjunctions. Figure 1 sentence ends with the copula である /dearu/ in the formal present-tense, has the Sino-Japanese noun 規定 /ki-tei/ ‘stipulations’ combined with する to form the verb ‘stipulate’, and includes the most frequent grammatical marker of の /no/ ‘possessive; nominalization’. Katakana is usually used to represent 外来語 /gai-rai-go/ ‘foreign-Japanese’ (referring to loanwords of foreign origin but excluding those from Chinese), foreign names, animal and plant species names, onomatopoeic expressions, and for emphasis and as glosses. Figure 1 contains one loanword バイト /baito/ ‘byte’ and is also used within the parentheses to indicate the pronunciation of JIS as ジス /jisu/. Rōmaji is usually used to represent foreign words and names, most commonly within advertising and the mass media. The single example in Figure 1 is JIS within JIS X 0208, where it forms an element of the standard reference code as an abbreviation of the established English translation for 日本工業規格. Arabic numerals are also widely used to represent numbers, particularly in scientific and financial domains. There are two Arabic-numeral strings within Figure 1; the first is 0208, as part of the standard’s reference code, which is read digit by digit, and the second is 6,879 for the number of graphic characters, which is read as a number. Clearly, 6,879 is much shorter than the corresponding kanji representation of 六千八百七十九, where digit values must be indicated (which is also obligatory when reading as a number, as the phonological gloss indicates; just as one would read the number in English as ‘six-thousand, eight-hundred and seventy-nine’).

Notably, one earlier kundoku practice to persist is the strategy of margin annotations, now known as ルビ /rubi/. The most common application is – when there is some expectation, such as in children’s reading materials, that a kanji’s pronunciation might not be known by a reader – to indicate it with a kana gloss (also referred to as 振り仮名 /fu.ri.ga-na/) that is usually placed above. However, the basic technique is extendable to any component script for more diverse purposes, such as indicating an alternative pronunciation for effect, offering explanation of a word, and simultaneously evoking additional concepts (Joyce & Masuda, 2016).

While these basic orthographic conventions are generally adhered to, patently, orthographic variation is also an inherent property of the JWS by simple virtue of its multi-scripts. The pervasive nature of orthographic variation, at least for the most common Japanese words, is strikingly evident in the analyses that Joyce et al. (2012) conducted on the corpus word lists (CWLs) that they extracted from the large-scale Balanced Corpus of Contemporary Written Japanese (BCCWJ; Maekawa et al., 2013; see also Joyce, Hodošček, & Masuda, 2017). More specifically, they found that the average number of orthographic variations is 8.44 (min 6.46, max 10.19) for the most frequent 100 short-unit words (essentially simplex words) across the four main word classes of nouns, verbs, i-adjectives and adverbs. For example, five orthographic variants are attested for the noun of 玉葱 /tama-negi/ ‘onions’, nine for the verb of 聞き取る /ki.ki.to.ru/ ‘hear, catch (words)’, 11 for the i-adjective of 面白い /omo-shiro.i/ ‘interesting; amusing’ and 13 for the adverb of 全然 /zen-zen/ ‘not at all; entirely’. As Joyce and Masuda (2016, 2017) illustrate, there are a number of motivating factors for these high degrees of orthographic variation, including aesthetics factors, desires to avoid complex kanji or ones with negative connotations, differentiating meaning nuances, and orthographic balance between words.

### 2.2.1 *Morphographic kanji*

As already alluded to, kanji remain the principal script for the orthographic representation of native-Japanese and Sino-Japanese content words. As also noted, the Japanese government’s official jōyō kanji list specifies 2,136 kanji; it consists of 1,006 教育漢字 /kyō-iku-kan-ji/ ‘education kanji’, that are taught during elementary school, and 1,130 kanji, that are taught at high-school (Bunkachō, 2010; Joyce et al., 2014). It should, however, be quickly stressed that this list does not represent an absolute ceiling on the number of kanji in daily usage. Like so many facets of the JWS, the situation is naturally rather more nuanced and a few comments of further clarification are warranted. In particular, it is helpful to understand that the 2010 revision was not a radical reform of kanji policy; rather, it is more appropriate to regard it as a periodic tweaking of the official list that involved the removal of five characters and the addition of 196 kanji compared to the previous version. It is also beneficial to bear in mind that the list only has guideline status; while generally conformed to in official documents and in newspapers, it is not prescriptive for all written Japanese language. As referred to earlier in Figure 1, JIS X 0208 specifies a set of 6,879 graphic characters for the present era of electronic information exchange, which effectively imposes more practical restrictions on kanji usage in actually specifying 6,355 kanji, including 2,965 level 1 and 3,390 level 2 kanji (Lunde, 1993). Hence, it is also valuable to look at corpus data in order to gain more informed insights concerning actual kanji usage. Also analysing kanji coverage for



their BCCWJ-based CWLs, Joyce et al. (2012) report that, although the revised jōyō kanji only represent 33.03% of kanji types, they account for the majority of kanji tokens at 96.12%. Moreover, while 4,093 of the remaining JIS kanji represent 63.30% of the types, they only account for 3.60% of the tokens, with another 237 kanji making up the final 3.67% of types and only a negligible 0.27% of the tokens. With these few caveats lodged, it becomes clearer that while the jōyō kanji list has considerable status as a de facto standard for functional Japanese literacy, generally, educated Japanese people are still expected to know substantially more kanji that continue to be associated with areas of cultural significance, such as their use in place and family names.

Beyond their numbers, another important aspect of kanji is that they naturally vary a great deal in their visual complexity. For instance, the number of strokes to write the jōyō kanji ranges from one (i.e., 一 /ichi/ 'one') to 29 strokes (i.e., 鬱 /utsu/ 'depression'), with the average count being 10.47 strokes (*SD* 3.79) (Joyce et al., 2012). Obviously, complexity on such a scale is only feasible because the majority of kanji possess internal structure (i.e., basic kanji or variant forms) related to the various formation principles that underlie their creation (Joyce, 2011); the dominant principle being that of 形声文字 /kei-sei-mo-ji/ 'phonetic compound kanji', where a 部首 /bu-shu/ '(semantic) radical' is combined with an 音符 /on-pu/ '(phonological) radical', such as 言 'speak' combined with 吾 /go/ 'I; my' to indicate 語 /go/ 'language'. Analysing the internal structures of jōyō and JIS level 1 (JIS1) kanji in terms of three basic configurations (namely, left-right, top-bottom, and enclosure-enclosed, plus a non-divisible category for the remainder), Joyce et al. (2014) report that 91.3% of jōyō and 92.6% of JIS1 kanji possess these basic configurations. Their analysis also involved the identification of 1,072 and 1,290 component elements for jōyō and JIS1 kanji, respectively.

As acknowledged earlier, one particularly compelling factor of the kundoku-mediated account is in elucidating the origins and persistence of the dual-reading system of both kunyomi and onyomi associated with kanji. Once again, however, the complete picture is rather more nuanced in nature, because there is also variation in the numbers of readings associated with particular kanji. A major factor being that there are actually three different kinds of onyomi, reflecting the fact that Chinese characters were borrowed at different periods and from different regions of China (Lurie, 2012; Miller, 1967; Shibatani, 1990). Although Coulmas (1989: 126) singles out 頭 'head; counter for large animals' as a rather extreme example, associated with four onyomi, /zu/, /tō/, /do/, and /ju/, and six kunyomi, /saki/, /atama/, /kashira/, /kōbe/, /kaburi/ and /tsumuri/, it should also be noted that the Japanese government's guideline policies have generally sought to reduce the number of official pronunciations associated with jōyō kanji. As Joyce et al. (2014) report, although the number of onyomi associated with jōyō kanji

ranges from 0–5 and the number of kunyomi ranges from 0–10, the frequency distribution is skewed towards small sets of associated pronunciations. Accordingly, while 92.0% of jōyō kanji are associated with 0–2 kunyomi and 0–2 onyomi, 34.7% have only one onyomi and no kunyomi and 32.1% have only one onyomi and one kunyomi. However, as explained further within the next sub-section on the kana scripts, reflecting the relatively uncomplicated nature of Japanese phonology, the comparatively high incidences of Japanese homophones are particularly associated with onyomi. For instance, 67 jōyō kanji are associated with the onyomi of /kō/ and 66 are associated with /shō/. As noted earlier, onyomi are derived from Japanese approximations of associated Chinese pronunciations, but, many Chinese phonological distinctions, including tone distinctions, in particular, were effectively lost due to the simpler syllable structure of Japanese. In illustration, Martin (1972: 98f) suggests that the diverse classical Chinese syllables of /ko/, /kau/, /kou/, /kwang/, and /kong/, with either aspirated or unaspirated initial as well as three separate tones, all coalesced as the Japanese onyomi of /kō/.

### 2.2.2 Syllabographic kana

As already stressed, the kundoku-mediated account also provides a persuasive explanation of how from the very beginning the syllabographic kana scripts emerged to fulfil complementary functions that led to the contemporary multi-script JWS. As also noted earlier, although hiragana and katakana developed in separate ways, they essentially overlap in their potential to orthographically represent the syllables of Japanese phonology, or, more precisely, 拍 /haku/ ‘mora’ referring to equal-duration syllables.

Joyce et al. (2017) classify Japanese morae according to three main types of 71 basic, 33 contracted and 64 extended morae, for which there are only a few exceptions to having one-to-one mapping relations to the kana scripts. The basic mora group includes the five Japanese vowels of /a/, /i/, /u/, /e/ and /o/, for which the hiragana are あ, い, う, え, and お and for which the katakana are ア, イ, ウ, エ, and オ, respectively. The contemporary kana sets also consist of 39 consonant-vowel (CV) combinations of nine unvoiced consonants and 20 CV combinations with four voiced consonants and one semi-vowel consonant (although the calculation of 5 times 14 (9+4+1) returns 70, some CV combinations were never or are no longer distinguished from the corresponding vowels). For instance, the k-V combinations of /ka/, /ki/, /ku/, /ke/, and /ko/ are represented by the hiragana か, き, く, け, and こ and by the katakana カ, キ, ク, ケ, and コ, respectively. The basic mora group also includes one moraic nasal, /N/, hiragana ん and katakana ン, and one 促音 /soku-on/ ‘glottal stop’ of consonant gemination, hiragana っ and katakana ッ. One of the rare exceptions to the one-to-one relationship between kana and mora involves the hiragana を and katakana ヲ symbols, which historically corresponded

to /wo/ but are now pronounced as /o/ and are restricted to orthographically representing the ‘object’ grammatical marker. The basic kana sets are often described as consisting of 46 basic characters (5 vowels, 39 unvoiced CV combinations, plus /N/ and /wo/), because the other 20 CV combinations are orthographically derived by adding diacritics to some of the basic characters. Thus, for example, hiragana が *ga* and katakana ガ *ga* are derived by adding ̣, known as 濁音 /daku-on/ ‘voiced’, to か and カ /ka/, respectively, while hiragana ぱ *pa* and katakana ぱ *pa* are derived by adding ̤, known as 半濁音 /han-daku-on/ ‘semi-voiced’, to ば and ぱ /ha/, respectively. The 33 contracted mora group consists of CyV clusters, known as 拗音 /yō-on/, which involve combining 11 of the basic CV mora ending with /i/ vowels with reduced forms of either /ya/, /yu/ and /yo/. For example, /kya/, /kyu/ and /kyo/ are represented by the hiragana きゃ, きゅ, and きょ and by the katakana キヤ, キユ, and キヨ, respectively. The third main type of 64 extended mora relates to the additions to the traditional Japanese mora inventory to cope with the transcription of foreign loanwords and names, for which the katakana script is usually used. The extended morae are somewhat similar to the contracted group in that they all involve either a basic V or CV mora combined with either a V alone or one of the /ya/, /yu/ and /yo/ CV combinations, such as ク /ku/ and a reduced ワ /wa/ combined as クワ /kwa/ and テ /te/ and a reduced ュ /yu/ as テュ /tyu/.

Reflecting the strong societal expectations that Japanese children have generally learnt hiragana before entering elementary school, as Taylor and Taylor (2014) observe, first grade reading materials are initially only in hiragana, but katakana and some basic kanji are taught during the first grade. Even though kanji are graphically more complex, Steinberg and Yamada (1978–1979) have reported that three- and four-year-old Japanese children can find some kanji easier to recognize than kana symbols, because they represent meaningful concepts rather than abstract sounds.

### 2.2.3 *Phonemic rōmaji, Arabic numerals and punctuation*

Japanese contact with phonemic rōmaji can be traced back to Portuguese missionaries in the late 16th and early 17th centuries (Lurie, 2012; Okada, 2016). However, it is only really appropriate to regard it as a component of the JWS since the mid-twentieth century, when first taught at elementary schools (currently introduced during grade 4), and, even now, it remains the most peripheral or niche of the component scripts in terms of usage. In addition to representing foreign names and words, rōmaji is also used to represent Japanese words, especially Japanese-coined English terms, in general media contexts, particularly in advertising and information contexts, such as supplementary glossing of names for stops on public transport systems potentially for the benefit of foreigners. CM /shiemu/

‘TV commercial, ad’ seems to be a model example of rōmaji usage within the JWS; in addition to being almost exclusively represented in rōmaji (99.92% of BCCWJ occurrences; Joyce et al., 2012), it is an abbreviation for ‘commercial message’ which is a Japanese-coined term that is attested only once within the BCCWJ in full in katakana script, and, naturally, is of high frequency within the world of advertising. Yet again, orthographic variety is the norm because alternative transcriptions conventions exist; the government official 訓令式 /kun-rei-shiki/ ‘Cabinet ordinance system’, the へボン式 /hebon.shiki/ ‘Hepburn system’ (proposed by the American missionary James Curtis Hepburn (1815–1911), which is used for the phonological transcriptions provided throughout this chapter), and the oldest but least commonly used 日本式 /ni-hon-shiki/ ‘Nihon system’. Thus, for instance, 富士山 ‘Mount Fuji’ is rendered as either /huzisan/ (kunreishiki), /fujisan/ (hebonshiki), or /hudisan/ (nihonshiki). It also bears mentioning that, while kana-input modes are also available, the most commonly used method of inputting Japanese on computers with the QWERTY keyboard is the rōmaji-input mode (Okada, 2016), but kana-input methods tend to dominate for small screen devices, such as mobile and smart phones.

Even though the jōyō kanji list includes kanji that represent large numbers (up to 京 /kei/ ‘10<sup>16</sup>’) and kanji are commonly used in representing numbers, particularly for vertically-arranged texts, as already noted, Arabic numerals are also frequently used as part of the JWS, especially for scientific and financial texts that tend to be arranged horizontally.

While it is true that Japanese is written without spaces between words, as the sentence in Figure 1 illustrates, the JWS includes a number of punctuation marks; some of which are similar to those of European languages in terms of form and function and some of which are specific to the JWS. As a full description of Japanese punctuation is beyond the scope of this chapter, only a few examples are mentioned. Somewhat differing in form but generally similar in function, JWS has 、 読点 /tō-ten/ as a comma and 。 句点 /ku-ten/ as a full stop. It also has quotation marks that are essentially similar in function, but very different in form; 「...」 鉤括弧 /kagi-kak-ko/ ‘single quotation marks’ and 『...』 二重鉤括弧 /ni-jū-kagi-kak-ko/ ‘double quotation marks’. One punctuation mark that is specific to the JWS is ・ 中黒 /naka-guro/ lit. ‘middle black (dot)’ that usually functions as a separator (of characters or words).

### 3. Psycholinguistic studies of Japanese word processing

#### 3.1 Misleading, but enduring, dichotomies

In their important book reviewing psycholinguistic studies of kanji and kana processing, Kess and Miyamoto (1999) remark that such research has been one of the most active areas of Japanese psycholinguistics (for other reviews and edited collections, see Akita & Hatano, 1999; Chen (Ed.), 1997; Chen & Zhou, 1999; Flores d'Arcais, 1992; Hatta & Saito (Eds.), 1999, 2000; Kaiho & Nomura, 1983; Kess, 2005; Leong & Tamaoka (Eds.), 1998; Paradis, Hagiwara, & Hildebrandt, 1985; Saito, 1997, 2006; Sato, 2015; Tamaoka, 1991, 1994; Taylor & Taylor, 2014; Wydell, 2006; Yamada, 1997). While the JWS's unique multi-script nature undoubtedly opens up many potentially interesting opportunities for investigating the processes of visual word recognition, such as comparing functionally different scripts (i.e., morphographic, syllabographic and phonemic), as Part 2 sought to portray, the JWS's complex orthographic conventions also pose special challenges in terms of not confounding various lexical properties across experimental contrasts. Although highly selective in nature, Part 3 seeks to single out a few studies that substantiate Kess and Miyamoto's core insight of denouncing the simple 'early dichotomies' between phonological routes for kana and semantic routes for kanji, which regrettably continue to endure, as woefully inadequate to account for the complex interactions involved in word processing.

#### 3.2 Studies of single kanji processing

Notwithstanding their morphographic nature (Joyce, 2002b, 2011) – such that, while many kanji do represent *free morphemes* (i.e., simplex words), in the vast majority of cases, kanji represent the constituents of *polymorphemic words* (i.e., verb and adjectives stems and elements of compound words) – a great deal of psycholinguistic research has focused on the single kanji to examine the balance between phonological and semantic activation within lexical retrieval (Flores d'Arcais, 1992; Flores d'Arcais, Saito, & Kawakami, 1995; Mizuno, 1997; Saito, Masuda, & Kawakami, 1998; Sakuma, Sasanuma, Tatsumi, & Masaki, 1998; Shimomura & Yokosawa, 1995; Wydell, 1991; Wydell, Patterson, & Humphreys, 1993). Accordingly, many of these studies have focused on phonetic compound kanji – as already noted, the dominant principle of kanji formation combines a semantic marker with a phonetic marker – and, taken together, they generally demonstrate that the lexical retrieval of single kanji involves rather complex interactions within the activation of orthographic, phonological and semantic information.

For instance, one such early study is by Flores d'Arcais (1992), who argues that, while phonological information appears to be activated before semantic information in the naming task, which particularly emphasizes a kanji's pronunciation, the meanings of component radicals are activated during the recognition of complex kanji, even when not semantically related to the kanji meaning. Consistently, Flores d'Arcais and Saito (1993) have also reported semantic activation for component radicals in a speeded semantic-categorization task, where they observed interference in all critical conditions of graphically similar kanji pairs (仲 'friend' and 伸 'extend'), part-whole related kanji pairs (石 'stone' appears to contain 口 'mouth'), and opaque-component related kanji pairs ('mouth' element of 石 and 目 'eye').

Studies have also reported evidence for the phonologically-mediated activation of kanji words. For example, in a semantic categorization task, Wydell et al. (1993) have reported, in addition to significant effects of visual similarity, significant homophone effects with reactions times longer and greater error rates when responding to homophone distractors compared to correct exemplars. While Sakuma et al. (1998) have also obtained similar results for the same task, they also found that the homophone effect was reduced in a masked condition, but the effects of orthographic similarity remained strong. Moreover, Saito et al. (1998) also provide further evidence for both orthographic and phonological activation using a delayed matching task, where participants judged whether a 'probe' kanji (e.g., 畔) was one of two briefly presented 'source' kanji (e.g., 略 and 伴). Only observing a homophone effect when the probe was orthographically similar to the source kanji, they interpreted their findings as indicating that phonological information is automatically activated for both radicals and whole kanji, even though it is not explicitly required for the task.

### 3.3 Studies of kana processing and kana-kanji comparisons

To the extent that the JWS's orthographic conventions are generally adhered to, one experimental manipulation that can be (cautiously) exploited is orthographic familiarity. In one early naming study of katakana, Besner and Hildebrandt (1987) contrasted orthographically familiar words (i.e., normally represented with katakana) with orthographically unfamiliar words (i.e., normally represented with kanji), as well as non-words. The researchers interpreted their findings of orthographically familiar words being named faster than both orthographically unfamiliar words and non-words as indicating that phonological recoding is not obligatory for familiar kana.

Naturally, a number of studies have also sought to compare the processes of visual word recognition for kana and kanji (Hino, Lupker, Ogawa, & Sears 2003;

Kim, 2012; Shimamura, 1987; Yamada, 1997, 1998). For instance, employing the Stroop task and an interesting paradigm variation, Shimamura (1987) conducted an early comparative study that underscores the important dissociation between word naming and word comprehension. The Stroop effect is where participants are slower to name the ink colour of a printed word when the word itself is the name of a colour that is incongruent with the ink colour (Stroop, 1935). Shimamura reported greater Stroop interference for conflicting colour words represented in kanji compared to katakana, even though katakana representations were named faster. In the paradigm variation, participants were asked to indicate the spatial location of a stimulus, such as a conflicting arrow ( ↓ ), kanji word ( 下 ), or katakana word ( シタ ) indicating ‘down’ in an *up* position. While interference was also observed for the conflicting arrow condition, it was greatest in the kanji word condition, but, again, words represented in kana were named faster than the kanji-orthography words.

More recently, Kim (2012) conducted an eye-tracking study with Japanese adults reading both a conventional multi-script version and a hiragana-only version of a text. Although the participants were equally effective in recounting the passage content in both orthographic conditions, they were much slower to read the hiragana-only version, which required both more and longer fixations compared to the authentic multi-script version. Consistent with the empirical eye-tracking data that indicated that the hiragana-only version was read less fluently, all participants self-reported on experiencing the hiragana-only version as being harder to process and less natural to read.

### 3.4 Studies of compound word processing

As an agglutinative language, Japanese certainly has a substantial degree of affixation, but in contradistinction to Myers’ (2006: 169) pronouncement that “Chinese is the poster child of compounding, the language to cite for an example of morphology without much affixation”, unquestionably, Japanese also offers some of the most intriguing cases for research into the morphology of compound words, given that compounding is a highly productive principle of word-formation involving both Sino-Japanese and native-Japanese morphemes (Joyce & Masuda, 2013; see also Masuda & Joyce, 2018). Particularly germane to the point, Nomura’s (1988) assertion that two-kanji compound words are the most common word structure in the Japanese language is essentially validated by Joyce et al.’s (2014) analyses of the orthographic structures of Japanese words; particularly, their analysis of approximately 215,600 headwords of the 広辞苑 /*kō-ji-en*/ dictionary (Shinmura, 2008) that found that the three most frequent orthographic codes were 2C (two-kanji) at 37.5%, 3C (three-kanji) at 15.1% and 4C (four-kanji) at 8.9%.

As one of the first studies to specifically consider the lexical representation of two-kanji compounds within the mental lexicon, the lexical decision task experiments conducted by Hirose (1992) merit brief attention. Employing a form of constituent-morpheme priming to investigate the pattern of facilitations on lexical decisions for two-kanji compound-word targets following one of three prime conditions (first-constituent, second-constituent, or unrelated kanji), Hirose observed significant priming in both constituent conditions. Moreover, because significantly greater priming was observed for the first-constituent condition compared to the second-constituent, somewhat reminiscent of Forster's (1976) serial search model, Hirose hypothesized that the lexical retrieval of two-kanji compound words might be based on search mechanisms for clustered arrangements of compound words according to their shared first-constituents. This notion would, however, seem to entail some curious repercussions for the representation of related compound words within the mental lexicon. For instance, if clusters are only based on shared first-constituents, related compounds like 学習 /gaku-shū/ 'learning' and 大学 /dai-gaku/ 'university' would not be linked, which would seem to be even more problematic for the synonyms of different lexical stratum, such as Sino-Japanese 登山 /to-zan/ 'mountain-climbing' and native-Japanese 山登り /yama-nobo.ri/ 'mountain-climbing'. Accordingly, Joyce (2002a, 2002b) sought to essentially replicate Hirose's experiments, but by also controlling for and contrasting five word-formation principle conditions underlying the two-kanji compound-word targets. Similar to Hirose's results, compared to the unrelated prime conditions, Joyce observed significant priming effects in the two constituent-morpheme conditions across all word-formation principle conditions. However, quite dissimilar to Hirose's results, Joyce also found that the levels of priming in the two constituent conditions were similar across four out of the five word-formation conditions; a result that is not compatible with search mechanisms that prioritize the first-constituent (see also Masuda & Joyce, 2018).

To account for those findings, Joyce (2002a, 2002b, 2004) proposed the Japanese lemma-unit model (JLUM) as a model of the Japanese mental lexicon, that was largely inspired by the version of the multi-level interactive-activation framework proposed for the Chinese mental lexicon by Taft, Liu, and Zhu (1999). In a significant modification of an earlier model (Taft & Zhu, 1997), Taft et al. (1999) advocated the incorporation of lemma unit representations to mediate the connections between both orthographic and phonological access representations and semantic representations, as a solution to issues of representational redundancy, homographs and semantic transparency. In addition to benefiting from these advantages – which remain problematic areas for Saito's (1997) companion-activation model and Tamaoka and Hatsuzuka's (1998) interactive-activation model, although,



arguably, less so for Ijuin, Fushimi, Patterson, and Tatsumi's (1999) distributed connectionist model of naming – JLUM can also provide a more appealing account of the constituent-morpheme priming results just outlined. Given that the lexical retrieval of two-kanji compound words is assumed to be mediated by mechanisms of spreading activation, it can readily accommodate priming effects for the second-constituent condition. However, the most distinctive feature of the JLUM is that it is the first model of the Japanese mental lexicon that, on the one hand, seeks to capture the nuances of the dual-reading system of both kunyomi and onyomi, and, on the other hand, attempts to unify the processing of the JWS's multi-scripts within a single integrated model (see also Masuda & Joyce (2018) for further explanation).

#### 4. Conclusion

This chapter has attempted to tender a concise yet informative introduction to the multi-script JWS and Japanese word processing; a rather daunting challenge in view of the scholar judgements, acknowledged at the outset, concerning its considerable complexity. After singling out just a few key historical developments, Part 2 presented a succinct outline of the contemporary JWS's standard orthographic conventions that effectively serve both to visually differentiate content words from grammatical elements and to distinguish, to a lesser extent, between the main lexical stratum. It is, however, also paramount to appreciate how both the co-existence of multiple scripts and the highly malleable nature of these standard orthographic conventions combine to produce remarkable levels of orthographic variation (Joyce et al., 2012). Echoing Backhouse's (1984) perceptive remark about the incredible potential for orthographic flexibility, the JWS's multi-scripts undoubtedly foster highly imaginative and innovative ways of thinking about the orthographic representation of language. Writing in Japanese potentially involves making script selections that, arguably, reflect a unique awareness of written language (Joyce & Masuda, 2016, 2017).

Contrastive in terms of both its brevity and focus, Part 3 presented a highly-selective review of psycholinguistic research on Japanese visual word processing. While sounding cautionary notes about methodological challenges for experimental designs, such as the wide prevalence of orthographic variation potentially diminishing the significance of orthographic familiarity contrasts, and about the enduring influences of misleading theoretical dichotomies concerning processing routes, Part 3 sought to highlight with a small selection of examples how the multi-script nature of the JWS unquestionably offers some of the most exciting prospects for

investigating the complex interrelationships between orthography, phonology and semantics, that are foundational issues for more adequately understanding writing systems and word processing.

## References

- Akita, K., & Hatano, G. (1999). Learning to read and write in Japanese. In M. Harris & G. Hatano (Eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 214–234). Cambridge: CUP.
- Backhouse, A. E. (1984). Aspects of the graphological structure of Japanese. *Visible Language*, 18, 219–228.
- Besner, D., & Hildebrandt, N. (1987). Orthographic and phonological codes in the oral reading of Japanese kana. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 335–343.
- Bunkachō [Agency for Cultural Affairs]. (2010). Jōyōkanjihyō [Jōyō kanji list]. Available at <[http://kokugo.bunka.go.jp/kokugo\\_nihongo/joho/kijun/naikaku/pdf/joyokanjihyo\\_20101130.pdf](http://kokugo.bunka.go.jp/kokugo_nihongo/joho/kijun/naikaku/pdf/joyokanjihyo_20101130.pdf)> (13 November, 2016).
- Chen, H-C. (Ed.). (1997). *Cognitive processing of Chinese and related Asian languages*. Hong Kong: The Chinese University Press.
- Chen, H-C., & Zhou, X. (1999). Processing East Asian languages: An introduction. *Language and Cognitive Processes*, 14, 425–428. doi:10.1080/016909699386130
- Chikamatsu, N., Yokoyama, S., Nozaki, H., Long, E., & Fukuda, S. (2000). A Japanese logographic character frequency list for cognitive science research. *Behavior Research Methods, Instruments, & Computers*, 32(3), 482–500. doi:10.3758/BF03200819
- Coulmas, F. (1989). *The writing systems of the world*. Oxford: Blackwell.
- DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems*. Honolulu, HI: University of Hawaii Press.
- Fischer, S. R. (2001). *A history of writing*. London: Reaktion Books.
- Flores d'Arcais, G. B. (1992). Graphemic, phonological and semantic activation processes during the recognition of Chinese characters. In H-C. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 37–66). Amsterdam: North-Holland. doi:10.1016/S0166-4115(08)61886-5
- Flores d'Arcais, G. B., & Saito, H. (1993). Lexical decomposition of complex kanji characters in Japanese readers. *Psychological Research*, 55, 52–63. doi:10.1007/BF00419893
- Flores d'Arcais, G. B., Saito, H., & Kawakami, M. (1995). Phonological and semantic activation in reading kanji characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 34–42.
- Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms: A collection of psycholinguistic studies* (pp. 257–287). Amsterdam: North-Holland.
- Gottlieb, N. (2008). Japan: Language policy and planning in transition. In R. B. Kaplan & R. B. Baldauf Jr. (Eds.), *Language planning and policy in Asia, Vol. 1: Japan, Nepal, Taiwan and Chinese characters* (pp. 102–169). Bristol: Multilingual Matters.
- Habein, Y. S. (1984). *The history of the Japanese written language*. Tokyo: University of Tokyo Press.

- Hatta, T., & Saito, H. (Eds.). (1999). Cognitive processing of the Japanese and Chinese languages I. *Psychologia*, 42(4).
- Hatta, T., & Saito, H. (Eds.). (2000). Cognitive processing of the Japanese and Chinese languages II. *Psychologia*, 43(1).
- Hino, Y., Lupker, S. J., Ogawa, T., & Sears, C. R. (2003). Masked repetition priming and word frequency effects across different types of Japanese scripts: An examination of the lexical activation account. *Journal of Memory and Language*, 48, 33–66. doi:10.1016/S0749-596X(02)00500-4
- Hirose, H. (1992). Jukugo no ninchi katei ni kan suru kenkyū: Puraimingu hō ni yoru kentō [An investigation of the recognition process for jukugo by use of priming paradigms]. *Shinrigaku Kenkyū* [The Japanese Journal of Psychology], 63, 303–309. doi:10.4992/jjpsy.63.303
- Igarashi, Y. (2007). *The changing role of katakana in the Japanese writing system: Processing and pedagogical dimensions for native speakers and foreign learners* (Unpublished doctoral dissertation). University of Victoria, British Columbia, Canada.
- Ijuin, M., Fushimi, T., Patterson, K., & Tatsumi, I. (1999). A connectionist approach to Japanese kanji word naming. *Psychologia*, 42, 267–280.
- Japanese Wikipedia. JIS X 0208 entry. Available at <[https://ja.wikipedia.org/wiki/JIS\\_X\\_0208](https://ja.wikipedia.org/wiki/JIS_X_0208)> (5 November, 2016).
- Joyce, T. (2002a). *The Japanese mental lexicon: The lexical retrieval and representation of two-kanji compound words from a morphological perspective* (Unpublished doctoral dissertation). University of Tsukuba, Japan.
- Joyce, T. (2002b). Constituent-morpheme priming: Implications from the morphology of two-kanji compound words. *Japanese Psychological Research*, 44, 79–90. doi:10.1111/1468-5884.00009
- Joyce, T. (2004). Modeling the Japanese mental lexicon: Morphological, orthographic and phonological considerations. In S. P. Shohov (Ed.), *Advances in psychological research: Volume 31* (pp. 27–61). Hauppauge, NY: Nova Science.
- Joyce, T. (2011). The significance of the morphographic principle for the classification of writing systems. *Written Language & Literacy*, 14(1), 58–81. doi:10.1075/wll.14.1.04joy
- Joyce, T. (2016). Writing systems and scripts. In A. Rocci & L. de Saussure (Eds.), *Verbal communication* (pp. 287–308). Berlin: De Gruyter Mouton.
- Joyce, T., Hodošček, B., & Masuda, H. (2017). Constructing an ontology and database of Japanese lexical properties: Handling the orthographic complexity of the Japanese writing system. *Written Language and Literacy*, 20(1), 27–51. doi:10.1075/wll.20.1.03joy
- Joyce, T., Hodošček, B., & Nishina, K. (2012). Orthographic representation and variation within the Japanese writing system: Some corpus-based observations. *Written Language & Literacy*, 15(2), 254–278. doi:10.1075/wll.15.2.07joy
- Joyce, T. & Masuda, H. (2013). Constituent-morpheme priming study of Sino-Japanese and Native-Japanese compound words. Presentation given at 8th International Morphological Processing Conference, 20–22 June, Cambridge, UK.
- Joyce, T., & Masuda, H. (2016). Just mixed up or a pretty neat idea? Some reflections on the multi-script nature of the Japanese writing system. Presentation given at ‘Understanding writing systems: From core issues to implications for written language acquisition’ – 10th International Workshop on Written Language and Literacy, 12–13 May, Radboud University, Nijmegen, The Netherlands.
- Joyce, T., & Masuda, H. (2017). From conventional to non-conventional forms of Japanese orthographic representation. Presentation given at ‘Writing systems: Past, present (... and future?) – 11th International Workshop on Written Language and Literacy, 29–13 August, Nanzan University, Nagoya, Japan.

- Joyce, T., Masuda, H., & Ogawa, T. (2014). Jōyō kanji as core building blocks of the Japanese writing system: Some observations from database construction. *Written Language and Literacy*, 17(2), 173–194. doi:10.1075/wll.17.2.01joy
- Kaiho, H., & Nomura, Y. (1983). *Kanji jōhō shori no shinrigaku* [The psychology of kanji information processing]. Tokyo: Kyōiku Shuppan.
- Kess, J. F. (2005). On the history, use, and structure of Japanese kanji. In K. Tamaoka (Ed.), *Corpus studies on Japanese kanji* (pp. 1–15). Tokyo: Hituzi Syobo & Lüdenschied: RAM-Verlag.
- Kess, J. F., & Miyamoto, T. (1999). *The Japanese mental lexicon: Psycholinguistics studies of kana and kanji processing*. Amsterdam: John Benjamins.
- Kim, K. (2012). How readers process Japanese orthography in two different texts. In K. Goodman, S. Wang, M. S. Iventosch, & Y. Goodman (Eds.), *Reading in Asian languages: Making sense of written texts in Chinese, Japanese, and Korean* (pp. 144–157). London: Routledge.
- Leong, C. K., & Tamaoka, K. (Eds.). (1998). *Cognitive processing of the Chinese and the Japanese languages*. Dordrecht: Kluwer. doi:10.1007/978-94-015-9161-4
- Lunde, K. R. (1993). *Understanding Japanese information processing*. Sebastopol, CA: O'Reilly & Associates.
- Lurie, D. B. (2012). The development of writing in Japan. In S. D. Houston (Ed.), *The shape of script: How and why writing systems change* (pp. 159–185). Santa Fe, NM: School for Advanced Research Press.
- Maekawa, K., Yamazaki, M., Ogiso, T., Maruyama, T., Ogura, H., Kashino, W., ... Den, Y. (2013). Balanced Corpus of Contemporary Written Japanese. *Language Resources and Evaluation*, 1–27. doi:10.1007/s10579-013-9261-0
- Martin, S. E. (1972). Nonalphabetic writing systems: Some observations. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationships between speech and reading* (pp. 81–102). Cambridge, MA: The MIT Press.
- Masuda, H., & Joyce, T. (2018). Constituent-priming investigations of the morphological activation of Japanese compound words. In H. K. Pae (Ed.), *Writing systems, reading processes, and cross-linguistic influences: Reflections from the Chinese, Japanese and Korean languages*. Amsterdam: John Benjamins.
- Miller, R. A. (1967). *The Japanese language*. Chicago, IL: University of Chicago Press.
- Mizuno, R. (1997). Kanji hyōkigo no onin shori jidōka kasetu no kenshō [A test of a hypothesis of automatic phonological processing of kanji words]. *Shinrigaku Kenkyū* [The Japanese Journal of Psychology], 68, 1–8. doi:10.4992/jjpsy.68.1
- Myers, J. (2006). Processing Chinese compounds: A survey of the literature. In G. Libben & G. Jarema (Eds.), *The representation and processing of compound words* (pp. 169–196). Oxford: OUP.
- Nomura, M. (1988). Niji kango no kōzō [The structure of two-kanji Sino-Japanese words]. *Nihongogaku*, 7(5), 44–55.
- Okada, T. (2016). Japanese uses of the English writing system: A case study. In V. Cook & D. Ryan (Eds.), *The Routledge handbook of the English writing system* (pp. 397–412). Oxon: Routledge.
- Paradis, M., Hagiwara, H., & Hildebrandt, N. (1985). *Neurolinguistic aspects of the Japanese writing system*. Orlando, FL: Academic Press.
- Saito, H. (1997). Shinteki jisho [Mental lexicon]. In Y. Matsumoto, T. Kageyama, M. Nagata, H. Saito, & T. Tokunaga (Eds.), *Iwanami kōza gengo no kagaku 3: Tango to jisho* [Iwanami lectures: Science of language 3. Words and dictionaries] (pp. 93–153). Tokyo: Iwanami Shoten.

- Saito, H. (2006). Orthographic processing. In M. Nakayama, R. Mazuka, & Y. Shirai (Eds.), *The handbook of East Asian psycholinguistics, Volume II: Japanese* (pp. 233–240). Cambridge: CUP. doi:10.1017/CBO9780511758652.034
- Saito, H., Masuda, H., & Kawakami, M. (1998). Form and sound similarity effects in kanji recognition. *Reading and Writing: An Interdisciplinary Journal*, 10(3/5), 323–357. doi:10.1023/A:1008093507932
- Sato, H. (2015). Do different orthographies share the same mechanisms of reading? A review of research on and models for Japanese acquired dyslexia. *Aphasiology*, 29(10), 1–30. doi:10.1080/02687038.2015.1034084
- Sakuma, N., Sasanuma, S., Tatsumi, I. F., & Masaki, S. (1998). Orthography and phonology in reading Japanese kanji words: Evidence from the semantic decision task with homophones. *Memory & Cognition*, 26, 75–87. doi:10.3758/BF03211371
- Seeley, C. (1984). The Japanese script since 1900. *Visible Language*, 18, 267–301.
- Seeley, C. (1991). *A history of writing in Japan*. Leiden: Brill.
- Shibatani, M. (1990). *The languages of Japan*. Cambridge: CUP.
- Shimamura, A. P. (1987). Word comprehension and naming: An analysis of English and Japanese orthographies. *American Journal of Psychology*, 100, 15–40. doi:10.2307/1422640
- Shimomura, M. & Yokosawa, K. (1995). The processing of constituent characters in kanji words in proofreading Japanese sentences. *Psychological Research*, 58, 51–60.
- Shinmura, I. (2008). *Kōjien* [Japanese dictionary] (6th ed.). Tokyo: Iwanami Shoten.
- Smith, J. S. (1996). Japanese writing. In P. T. Daniels & W. Bright (Eds.), *The world's writing systems* (pp. 209–217). Oxford: OUP.
- Sproat, R. (2010). *Language, technology, and society*. Oxford: OUP.
- Steinberg, D. D., & Yamada, J. (1978–1979). Are whole word kanji easier to learn than syllable kana? *Reading Research Quarterly*, 14, 88–99. doi:10.2307/747295
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–622. doi:10.1037/h0054651
- Taft, M., Liu, Y., & Zhu, X. (1999). Morphemic processing in reading Chinese. In J. Wang, A. Inhoff, & H.-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 91–113). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Taft, M., & Zhu, X. (1997). Submorphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23, 761–775.
- Tamaoka, K. (1991). Psycholinguistic nature of Japanese orthography. *Studies in Language and Literature*, 11, 49–82.
- Tamaoka, K. (1994). A review of psychological studies of kana and kanji: A single phonological route to a multiple interactive activation. *Studies in Language and Literature*, 13, 1–18.
- Tamaoka, K., & Hatsuzuka, M. (1998). The effects of morphological semantics on the processing of Japanese two-kanji compound words. *Reading and Writing*, 10, 293–322. doi:10.1023/A:1008012220770
- Taylor, I., & Park, K. (1995). Differential processing of content words and function words: Chinese characters vs. phonetic scripts. In I. Taylor & D. R. Olson (Eds.), *Scripts and literacy: Reading and learning to read alphabets, syllabaries and characters* (pp. 185–195). Dordrecht: Kluwer. doi:10.1007/978-94-011-1162-1\_12
- Taylor, I., & Taylor, M. M. (2014). *Writing and literacy in Chinese, Korean and Japanese* (rev. ed.). Amsterdam: John Benjamins.

- Tranter, N. (2008). Nonconventional script choice in Japan. *International Journal of the Sociology of Language*, 192, 133–151.
- Twine, N. (1991). *Language and the modern state: The reform of written Japanese*. London: Routledge.
- Wydell, T. N. (1991). *Processing in the reading of Japanese: Comparative studies between English and Japanese orthographies* (Unpublished doctoral dissertation). London University, UK.
- Wydell, T. N. (2006). Lexical access. In M. Nakayama, R. Mazuka, & Y. Shirai (Eds.), *The handbook of East Asian psycholinguistics, Volume II: Japanese* (pp. 241–248). Cambridge: CUP. doi:10.1017/CBO9780511758652.035
- Wydell, T. N., Patterson, K. E., & Humphreys, G. W. (1993). Phonologically mediated access to meaning for kanji: Is a rose still a rose in Japanese kanji? *Journal of Experimental Psychology Learning, Memory, and Cognition*, 19, 491–514. doi:10.1037/0278-7393.19.3.491
- Yamada, J. (1997). *Learning and information processing of kanji and kana*. Hiroshima: Keisuisha.
- Yamada, J. (1998). The time course of semantic and phonological access in naming kanji and kana words. *Reading and Writing: An Interdisciplinary Journal*, 10(3/5), 425–437. doi:10.1023/A:1008095920749
- Yamada, T. (1967). Japan – The writing system: Historical research and modern development. In T. A. Sebeok (Ed.), *Current trends in linguistics* (Vol. 2, pp. 693–734). The Hague: Mouton.

