

# *Okurigana*:

## An Analysis from the Viewpoint of Decoding

Keisuke Honda

Keywords: The Japanese writing system; Decoding; *Kanji*; *Hiragana*; *Okurigana*

### 1. Introduction

In the Japanese writing system, *hiragana* graphemes constitute *okurigana* when they are postposed to *kanji* graphemes to spell out a section at the end of words (1a) or elements of compound words (1b).<sup>1</sup>

(1) WRITTEN FORM	a. 食 べ た <sup>2</sup>	b. 食 べ 過 ぎ た
TRANSLITERATION	<i>tabe be ta</i>	<i>tabe be sugi gi ta</i>
GLOSS	tabe-ta eat-PAST	tabe+sugi-ta eat+exceed-PAST

As will be elaborated on below, *okurigana* are basically used for ease of *kanji* decoding (Hoshina 1949, Hayashi 1977, Kondō 2005, Satake 2006). Put differently, one may say that they provide certain external support, without which *kanji* decoding would be difficult or even impossible. What are the difficulties, then, and how can we formulate the way in which *okurigana* reduce them? Besides, do *okurigana* offer sufficient linguistic information required for decoding? If not, what kinds of prior knowledge are required, on the basis of which they can be fully utilised?

To answer these questions, this paper presents a descriptive account of *okurigana* from the perspective of decoding, aiming at providing better insights into the nature of *kanji*, *hiragana* and their interaction in Japanese. The paper is organised as follows. First, section 2 outlines some basic issues, especially what will be referred to as three cases of *kanji* indeterminacy. Next, section 3 presents an analysis of the current

<sup>1</sup> This paper uses *grapheme* to refer to the basic unit of writing (Sproat 2000), without committing to theoretical discussions on the appropriateness of this traditional term (Daniels 1991, 1994, Herrick 1994a, b). For *kanji*, it is used for the so-called *character*, and not for its *component*.

<sup>2</sup> Throughout this paper, examples are phonemically transliterated using the *Kunreisiki* romanisation system (Cabinet 1954), with a few minor modifications. Upper cases are used to indicate Sino-Japanese items. The encoded linguistic items are glossed only when necessary. Where these details are irrelevant, examples are shown in a simpler format (e.g., 食べた *tabe-ta* 'ate').

conventions concerning *okurigana* orthography to formulate *okurigana*'s contribution to the resolution of the *kanji* indeterminacies. Then section 4 discusses some problems caused by orthographically permitted variants of *okurigana*. Finally, section 5 summarises the discussion and draws a conclusion.

## 2. Settings

### 2.1 *Okurigana* as the loci of *kanji-hiragana* interaction

Japanese is substantiated by three sets of graphemes or scripts, namely *kanji*, *hiragana* and *katakana*. Of these, the first two constitute the main components of the grapheme system employed in the current Japanese writing system. Texts are typically composed of both types of graphemes, each carrying out the specific functions sketched below.

*Kanji* graphemes encode linguistic elements which belong to the native (NJ) and the Sino-Japanese (SJ) strata of the lexicon. Characterisation of these elements is an issue of ongoing debate (e.g., Kaiser 1995, 2007); for the present purpose, however, it will suffice to regard them as individual morphemes (e.g., Kōno 1994/1977, Sampson 1985, Rogers 2005).<sup>3</sup> Contrastively, *hiragana* graphemes encode morae or, more precisely, mora-bearing phoneme strings (i.e., (C)V units or segments occupying the syllable-coda position) (e.g., Poser 1992, Ratcliffe 2001, Rogers 2005). Today, they are primarily used for representing non-lexical elements such as affixes, case particles and postpositions. With this kind of division of labour between *kanji* and *hiragana*, the Japanese writing

---

<sup>3</sup> There are two points which seem to justify this postulation. The first point is concerned with the size of the linguistic elements to which individual graphemes correspond. For the SJ set, *kanji* are used for free forms (e.g., 金 for *KIN* 'gold') and combining forms (e.g., 鉱 for *KOO* 'metal mine', only permitted in compounds such as 金鉱 *KIN+KOO* 'gold mine'), which can be generalised as root morphemes. For the NJ set, on the other hand, *kanji* are employed not only for free forms (e.g., 夢 *yume* 'dream'), but also for words which form an inflectional paradigm (e.g., 見 for *mi-ru* 'see-NON.PAST', *mi-ta* 'see-PAST', *mi-ro* 'see-IMP', etc.). The recurrent use of the same graphemes in view, it seems reasonable to say that these graphemes are associated to the recurrent linguistic elements, namely the stem morphemes. These observations afford grounds for positing that *kanji* correspond to morpheme-sized elements.

The second point is concerned with the level of abstraction at which *kanji* operate: it is a palpable fact that the graphemes do not indicate morphophonological alternations in a graphically distinctive fashion. For instance, observe the following case of compounding, which involves vowel alternation in the first element (*e* → *a*) and consonant voicing in the second (*k* → *g*): 酒 *sake* 'rice wine' + 蔵 *kura* 'storehouse' → 酒蔵 *saka+gura* 'rice wine cellar'. Importantly, these graphemes do not change their forms according to the alternations, suggesting that *kanji* relate to linguistic elements at a deeper level than the surface phonemic level. These two points taken together, it seems descriptively adequate to interpret these elements as morphemes.

Note, however, that this generalisation is not without exceptions. In fact, some *kanji* encode polymorphemic items (e.g., 政 *maturi-goto* 'politics') or sub-lexical phoneme strings that denote no meanings on their own (e.g., 葡 *BU* and 萄 *DOO* in 葡萄 *BUDOO* 'grape'). Further research is needed for a more systematic treatment of these issues.

system is characterised by the interaction of different types of graphemes.

Perhaps the most remarkable case of *kanji-hiragana* interaction is instantiated by the use of *okurigana*. As has been already mentioned, formally *okurigana* are *hiragana* graphemes employed to spell out a latter part of written words or elements of compound words. More specifically, they are postposed to *kanji* graphemes in a way that they spell out inflectional endings (2a), a section at the end of free forms or stem morphemes (2b), or both (2c).<sup>4</sup> In this way, *kanji* and *okurigana* amalgamate with each other to co-construct written words.

(2) a.	見 た	b.	且 つ	c.	食 べ た
	<i>mi ta</i>		<i>katu tu</i>		<i>tabe be ta</i>
	mi-ta		katu		tabe-ta
	see-PAST		besides		eat-PAST

## 2.2 The current conventions for *okurigana* orthography

The scope of the present paper is limited to the conventions stipulated by *Jōyō Kanji Hyō* (Cabinet 1981a) and *Okurigana no Tsukekata* (Cabinet 1981b). *Jōyō Kanji Hyō* is an official *kanji* inventory, which contains 1,945 commonly used graphemes (*jōyō kanji*) and 4,087 associated readings. Despite the non-binding nature of this inventory, it is a de facto standard adopted in legal and public documents, approved textbooks, major dailies and various other types of publications.

*Okurigana no Tsukekata*, on the other hand, is a collection of regulations concerning *okurigana* orthography. These regulations are primarily intended as guidelines on applying *okurigana* to *jōyō kanji* graphemes; in fact, however, they are also applicable to other non-official graphemes that are actually in use. As a hybrid of traditional usage and modern attempts of systematisation, *Okurigana no Tsukekata* allows many variants and exceptions besides the basic forms defined by general rules. In general, the permitted variants comprise a smaller number of *okurigana* than the basic forms do. (This point is discussed in some detail in section 4.)

Adopting traditional usage, *Okurigana no Tsukekata* (Cabinet 1981b) constrains the

<sup>4</sup> Adopting the conventional conception of *okurigana*, this paper assumes that even if *hiragana* occupy the post-*kanji* position, they do not constitute *okurigana* if they represent the following:

- i) An element syntactically adjacent to what the given *kanji* represents (including particles and postpositions) (e.g., 夢が *yume ga* ‘dream NOM’); or
- ii) An element of a compound word morphologically adjacent to what the given *kanji* represents (e.g., 見かた *mi-kata* ‘mode of seeing’); or
- iii) The light verb *suru* (e.g., 心する *kokoro+suru* ‘be on the alert’).

applicability of *okurigana* with respect to the types of *kanji* readings. First, they are used only for NJ items, and not for SJ items (as stated in §4-2-4 of the Preface to *Okurigana no Tsukekata*). Second, *okurigana* are not employed for NJ nominals, except for nominal derivatives and a small set of exceptional items (as stated in §2 'Words without conjugation'). Because of these constraints, at present *okurigana* are primarily used for NJ non-nominal items such as verbs, adjectives, adverbs, etc.

### 2.3 *Okurigana* as aides to decoding

Generally speaking, writing can be viewed as a relation, in which “particular (sets of) linguistic elements *license* the occurrence of (sets of) orthographic elements” (Sprout 2000: 9; italics in the original). Once established, this relation can be implemented as mappings in both directions: from linguistic elements onto orthographic elements (i.e., *encoding*), or vice versa (i.e., *decoding*).<sup>5</sup>

In Japanese, *kanji* graphemes are known to cause difficulty in decoding, and previous studies have shown three cases where *okurigana* may be considered to remove or reduce the difficulty.<sup>6</sup> First, Satake (2006) points to the general lack of phonological information on the part of *kanji*. To be more specific, *kanji* graphemes are not analytically structured in a way that smaller graphic constituents indicate sub-lexical phoneme strings (e.g., syllables, morae, phonemes, etc.) regularly and systematically (Kōno 1980/1969).<sup>7</sup> This is illustrated in (3), where the phoneme string *EN* is encoded by a set of *kanji* graphemes with no common graphic constituent. The absence of graphically expressed phonemic analysis constitutes a serious problem for decoding, since it makes it impossible to derive the phonemic form of a *kanji* reading directly from the grapheme’s architecture. Hereafter, this problem will be tentatively referred to as the *indeterminacy of phonemic forms*.

(3) *EN* → {円, 延, 沿, 炎, 宴, ...}

<sup>5</sup> This is not to say that encoding is simply the inverse of decoding, or vice versa. For a discussion on the differences between these mappings in Japanese, see Kawai (1970).

<sup>6</sup> This point deserves explicit emphasis, since much more attention has been given to the encoding side of *okurigana*, often with particular reference to issues surrounding orthographic regularisation (e.g., Kokuritsu Kokugo Kenkyūjo 1952, Ōno 1958, Tsukishima 1970, Uchida 1972/1961, Hayashi 1977, Abe 1989). For a review of these issues, see Abe (1989) and Takebe (1979), inter alia.

<sup>7</sup> In principle, the so-called phonetic components of semantic-phonetic compounds (*keisei kanji*) are designated to indicate the pronunciation of SJ items. Nevertheless, there are discrepancies over the actual effectiveness of the phonetic components in present-day Japanese (Nomura & Itō 1978, Stalph 1989, Saito, Kawakami & Masuda 1995). Moreover, these components provide virtually no information at all about NJ items (except for a small number of *kokuji* graphemes).

Satake (2006) maintains that the function of *okurigana* is to make up for the shortage of phonological information in *kanji* graphemes. Intuitively, this remark captures the empirical fact experienced by the writer-reader of Japanese: it is often the case that one makes reference to the phonological information of *okurigana* to work out uncertain or unknown *kanji* readings. That said, Satake (2006) does not provide a detailed account and, as a matter of fact, very few studies have taken pains to discuss the phenomenon at hand. Accordingly, it remains unclear what kind of contribution *okurigana* make, and how it works for the ease of decoding.

In addition to the indeterminacy of phonemic forms, two more cases of indeterminacy arise from the highly ambiguous nature of *kanji*. As Vance (2002), Kondō (2005) and many others observe, a large number of *kanji* graphemes have several possible readings. This is a situation known as *homography*, where a single orthographic element corresponds to multiple linguistic elements (Vance 2002, Rogers 2005). Homographic *kanji* are inherently ambiguous because these graphemes do not incorporate any constituents marking out the single intended reading in a graphically distinctive manner.

To make the matters more complicated, many homographic *kanji* are potentially ambiguous at two levels (Kondō 2005, Honda 2007). At the morpheme level, on one hand, a single *kanji* may encode several NJ morphemes, several SJ morphemes, or both (4a). At the word-form level, on the other, the same *kanji* may be used for a set of various inflected forms realising a particular stem morpheme (4b). As a result, it is hardly possible to predict the intended reading of a homographic *kanji* by just looking at its structure. The present paper tentatively labels these two kinds of ambiguity as the *indeterminacy of morphemes* and the *indeterminacy of inflected forms*, respectively.

- (4) a. 食 → {*tabe-*, *kuw-*, *kuraw-*, *SYOKU*, *ZIKI*}  
 b. 食 → {*tabe-ru*, *tabe-ta*, *tabe-ro*, ...}

It is widely accepted that *okurigana* bring about an effective solution to these problems, too. At the word-form level, they spell out some or all relevant elements so that the target inflected form is explicitly distinguished from other possible forms (Tsukishima 1970, Sampson 1985, Abe 1989, Vance 2002, and many others). In addition, several researchers have suggested that *okurigana* serve to disambiguate homographic *kanji* also at the morpheme level (Hoshina 1949, Hill 1967, Sampson 1985, Vance 2002, Coulmas 2003, Honda 2005). To sum up their descriptions, *okurigana* specify the single intended morpheme by spelling out a section at the end of its realised form (this point will be further discussed in section 3.3). However, despite the occasional reference to

these points in the literature, still wanting is a systematic and detailed analysis of the mechanism by which *okurigana* serve to reduce ambiguity.

### 3. Indeterminacy resolution

In light of the lack of elaborate investigations, this section attempts to formulate the way *okurigana* can be considered to resolve the three cases of *kanji* indeterminacy. For each case, relevant conventions of *Okurigana no Tsukekata* are analysed to give a formal account of the indeterminacy resolution. The analysis is also aimed at identifying the kinds of prior knowledge that are presupposed for manipulating the information of *okurigana* for the purpose of decoding.

#### 3.1 Resolving the indeterminacy of phonemic forms

As mentioned earlier, the indeterminacy of phonemic forms is ascribed to the lack of graphically expressed analysis of linguistic elements into sub-lexical phonemic strings. *Okurigana* play a crucial role in resolving this problem by spelling out the end of these elements, which in effect adds on partial phonemic analysis and thus specifies the *kanji* readings. Logically, this is applicable in all cases where *okurigana* are present, since the lack of phonemic analysis is a property common to all *kanji* graphemes.

Within the range of *Okurigana no Tsukekata* (Cabinet 1981b), this is most visible in free forms such as (non-derived) adverbs (5a), conjunctions (5b) and, in some exceptional cases, nominals (5c). Here, *okurigana* are applied to *kanji* graphemes in such a way that they repeat the final mora(e) of the intended readings. In this way, they suggest the phonemic form of the entire word and thus resolve the indeterminacy.

(5) a.	再	び	b.	且	つ	c.	辺	り
	<i>futatabi</i>	<i>bi</i>		<i>katu</i>	<i>tu</i>		<i>atari</i>	<i>ri</i>
	futatabi			katu			atari	
	again			besides			surroundings	

Similarly, *okurigana* indicate the realised phonemic forms of stem morphemes by repeating a section at the stem-final position.<sup>8</sup> In Japanese, verb stems can be divided into two classes, namely those ending with a vowel (V-verbs) and those ending with a consonant (C-verbs). Following the traditional usage, *Okurigana no Tsukekata* stipulates

<sup>8</sup> *Okurigana no Tsukekata* permits many exceptions, making the matter far more complicated than is described here. Yet, this does not affect the point being made, as in any case *okurigana* repeat a smaller or larger section at the stem-final position.

that *okurigana* spell out one or more stem-final mora(e) in V-verbs (6a), and the unit comprising the stem-final consonant and the suffix-initial vowel in C-verbs (6b).<sup>9</sup> For adjectives, the basic rule is to spell out the inflectional suffix (i.e., *-i* or *-na*); in reality, however, they also represent stem-final mora(e) in various classes of items (6c).<sup>10</sup>

(6) a.	食	べ	る	b.	飲	む	c.	楽	し	い
	<i>tabe</i>	<i>be</i>	<i>ru</i>		<i>nom</i>	<i>mu</i>		<i>tanosi</i>	<i>si</i>	<i>i</i>
	tabe-ru				nom-u			tanosi-i		
	eat-NON.PAST				drink-NON.PAST			pleasurable-NON.PAST		

At this point, it is important to stress that *okurigana* provide only partial phonological information; that is, they do not repeat the *whole* phonemic form of a *kanji* reading (e.g., \*再ふたたび, \*食たべる, etc.).<sup>11</sup> Accordingly, the phonemic form cannot be simply *derived* from the written form; rather, it must be *inferred* from the partial information provided. Such inference necessarily presupposes two kinds of prior knowledge, namely the knowledge about the phonemic forms of the given linguistic element, and the knowledge about grapheme-to-phoneme correspondence rules. Without them, there is no way of utilising the information of *okurigana* in any meaningful way. To sum up, one may say that *okurigana* resolve the indeterminacy of phonemic forms on the basis that the reader has access to the knowledge of both types.<sup>12</sup>

### 3.2 Resolving the indeterminacy of inflected forms

In section 2.3, it was shown that the indeterminacy of inflected forms arises from the

<sup>9</sup> This description is based on a morphological analysis of Japanese verbs, which differs from the traditional analysis of verb conjugation in several respects. For example, from the standpoint of morphology, the two verbs in (6a) and (6b) can be analysed as *tabe-ru* and *nom-u*, respectively, where *tabe-* and *nom-* constitute the verb stems. In traditional schemes, the same verbs may be analysed into *ta-beru* and *no-mu*, respectively, with *ta-* and *no-* construed as the verb stems.

<sup>10</sup> According to *Okurigana no Tsukekata*, this is applicable to all morphemes whose stems end with *-si* (e.g., *itijirusi-i* ‘remarkable’), *-ka* (e.g., *atataka-na* ‘warm’), *-yaka* (e.g., *odayaka-na* ‘calm’), or *-raka* (*akiraka-na* ‘obvious’). The same is also true for several lexically specified items (e.g., *arata-na* ‘new’).

<sup>11</sup> This stands in marked contrast to *furigana*, or *hiragana* graphemes superposed on *kanji* graphemes to spell out the *kanji* reading in full (e.g., 夢 *yume* ‘dream’; ゆ → *yu*, め → *me*). In this case, the phonemic form can be derived directly from the phonological information *furigana* provide.

<sup>12</sup> This paper assumes that the *reader* may be a human user of the Japanese writing system (native or non-native, learned or learning), or a machine designed to convert writing into speech (such as a text-to-speech synthesis system). The present discussion is concerned with matters internal to the writing system per se, and not with the psycholinguistic process of reading or the system architecture.

problem of homography-induced ambiguity at the word-form level. The recurrent use of the same *kanji* for various inflected forms makes it impossible to determine the intended form by looking at the grapheme itself. Meanwhile, it was also shown that there is a clear consensus among researchers that *okurigana* specify the intended form by indicating the inflectional ending.

The basic principle of *Okurigana no Tsukekata* is that *okurigana* spell out inflectional endings, which consist of suffixes denoting tense, aspect, modality, etc.<sup>13</sup> By extension, *okurigana* are capable of specifying any possible inflectional form, even if it is morphologically highly complex as *mi-sase-rare-tagar-anaka-at-ta* ‘did not want to be made to see’ (7).<sup>14</sup>

- (7) 見 さ せ ら れ た が ら な か っ た  
*mi- sa se ra re ta ga ra na ka ɔ ta*  
 mi-sase-rare-tagar-anak-at-ta  
 see-CAUS-PASS-want-NEG-be-PAST

Interestingly, the spelling out of inflectional endings is not only exhaustive but also surface-oriented in terms of the level of abstraction. In Japanese, there are morphophonological processes collectively known as *onbin*, wherein certain classes of stem-final consonants alternate with other segments when concatenated to some inflectional suffixes. For instance, when the stems *kak-* ‘write’, *yom-* ‘read’ and *tor-* ‘take’ are followed by the past tense suffix *-ta*, they become *kai-*, *yon-* and *toɔ-*, respectively (8a-c).<sup>15</sup> Note that *okurigana* represent the surface phonemic forms rather than the more abstract morphophonemic forms.

- |        |             |          |           |    |             |          |           |    |             |          |           |
|--------|-------------|----------|-----------|----|-------------|----------|-----------|----|-------------|----------|-----------|
| (8) a. | 書           | い        | た         | b. | 読           | ん        | だ         | c. | 取           | っ        | た         |
|        | <i>kak-</i> | <i>i</i> | <i>ta</i> |    | <i>yom-</i> | <i>N</i> | <i>da</i> |    | <i>tor-</i> | <i>ɔ</i> | <i>ta</i> |
|        | ka[i]-ta    |          |           |    | yo[n]-ta    |          |           |    | to[t]-ta    |          |           |
|        | write-PAST  |          |           |    | read-PAST   |          |           |    | take-PAST   |          |           |

With the exhaustive representation of the surface phonemic forms, *okurigana* specify

<sup>13</sup> There are varying views as to what constitute inflectional suffixes in Japanese (e.g., Narrog 1998, Miyaoka 2002).

<sup>14</sup> In the transliteration, the symbol *ɔ* denotes consonant length, which is moraic in Japanese.

<sup>15</sup> In the transliteration, the elements shown in square brackets denote the surface phonemic forms. The symbol *N* denotes the moraic nasal, which may be realised by [n], [m] or [ŋ] as a result of regressive assimilation.





these two kinds of knowledge are available, *okurigana* may be utilised to specify the intended NJ non-nominal morpheme phonographically.<sup>18</sup>

#### 4. Indeterminacy resolution in the permitted variants

*Okurigana no Tsukekata* permits a range of variants, which generally comprise a smaller number of *okurigana* than the basic forms do. According to the regulations, these variants can be used “as long as there are no risks of misreading”. The judgement as to whether they may cause misreading is left to individual writers. Thus, even though priority is given to the basic forms, there is virtually no limit to the use of the variants. In real life, they are often preferred over the basic forms because they consume less space, ink, energy, etc.

Because the permitted variants consist of a smaller amount of *okurigana*, they provide a smaller amount of information that can be exploited for decoding. For one example, the grapheme 祭 encodes two morphemes denoting ‘festival’, namely NJ *maturi* and SJ *SAI*. In its basic form, the former is written as 祭り, with the *okurigana* り *ri* repeating the last mora. In this way, it is differentiated from the latter, which is SJ and thus takes no *okurigana*. However, it is permitted to omit the *okurigana* in question, and this results in a homographic representation that is inherently ambiguous. This variant is actually attested in various festival names across Japan, and whether it should be decoded as *maturi* or *SAI* is lexically specified according to individual festivals (13). From the viewpoint of the present study, this is clearly a setback for decoding.

(13) 葵祭 *aoi+maturi* ‘A Shinto festival celebrated in Kyoto’ Cf. ?*aoi+SAI*

Another example can be found in the NJ compound *uke+tuke* ‘reception’. The basic form for this word is 受け付け, wherein the *okurigana* け *ke* spells out the final mora in both elements (14a). Although the two *kanji* graphemes encode three distinct morphemes respectively (i.e., 受 → {*uke-*, *ukar-*, *ZYU*}; 付 → {*tuke-*, *tukar-*, *FU*}), the *okurigana* indicates the realised phonemic forms and thus specifies the intended morphemes. Nevertheless, this *okurigana* is often dropped in one element (14b) or both (14c). To decode these variants, it is necessary to make reference to one’s lexical

---

<sup>18</sup> To be fair, this conclusion holds true only in cases where *okurigana* do accompany *kanji* graphemes. In reality, the applicability of *okurigana* is heavily restricted by several conditions. To make matters worse, there are some *kanji* that are not sufficiently disambiguated even if *okurigana* are present. For a full discussion, see Honda (2007).

Nonetheless, these facts by no means reduce the importance of *okurigana*’s contribution to the resolution of the indeterminacy of morphemes. Without them, decoding would be entirely dependent on guessing and, at best, inference from the orthographic and syntactic contexts (see footnote 17).

knowledge and check if there are such lexical items as *\*uke+tukar(-i)*, *\*uke+FU*, *\*ZYU+FU* and so forth, before determining that *uke+tuke* is the only possible and existing word. Again, this is a major drawback for decoding.

- (14) a. 受 け 付 け                      b. 受 付 け                      c. 受 付  
           *uke ke tuke ke*                      *uke tuke ke*                      *uke tuke*

It is clear from the above examples that the smaller the amount of *okurigana* employed for encoding, the larger the amount of lexical knowledge required for decoding. A recent survey of 70 Japanese magazines conducted by the National Institute for Japanese Language suggests a general tendency to use the basic forms, at least in non-specialist publications (Kokuritsu Kokugo Kenkyūjo 2006). Still, the sporadic occurrences of the variants are by no means negligible. In view of this fact, it would be safe to say that lexical knowledge is indispensable for utilising *okurigana* for decoding.

## 5. Concluding remarks

This paper has presented a descriptive account of Japanese *okurigana* from the viewpoint of decoding, with particular reference to the current conventions concerning *okurigana* orthography. Previous studies have been reviewed to outline three cases of *kanji* indeterminacy, and an analysis has been conducted to formulate the way *okurigana* can be said to resolve these indeterminacies. Some problems caused by orthographically permitted variants of *okurigana* application have been also discussed.

From the discussion above, it seems reasonable to conclude that the following types of prior knowledge are required to exploit *okurigana* for the purpose of decoding. Further research is needed to investigate the utility of these types of knowledge in other aspects of the Japanese writing system.

### ➤ Linguistic

- Lexical knowledge, especially about phonemic forms and the distinction of NJ and SJ items; and
- Morphological knowledge, especially about inflection, derivation and morphophonemic alternations.

### ➤ Orthographic

- Knowledge about grapheme-to-phoneme correspondences; and
- Knowledge about orthographic constraints, especially those restricting the way *okurigana* can be applied to *kanji* graphemes.

## 【References】

- Abe, Seiya (1989) Jōyō kanji no okurigana. In Kiyoji Satō (ed.), *Kanji to kokugo mondai*, 65-98. Meiji Shoin.
- Cabinet (1954) *Rōmaji no tsuzurikata*. Cabinet Notification No. 1, December 9, 1954. *Kanpō Gōgai*, December 9 edition, 1954. Ōkurashō Insatsukyoku.
- (1981a) *Jōyōkanji-hyō*. Cabinet notification No. 1, October 1, 1981. *Kanpō Gōgai*, October 1 edition, 1981. Ōkurashō Insatsukyoku.
- (1981b) *Okurigana no Tsukekata*. Cabinet notification No. 3, October 1, 1981. *Kanpō Gōgai*, October 1 edition, 1981. Ōkurashō Insatsukyoku.
- Coulmas, Florian (2003) *Writing systems: An introduction to their linguistic analysis*. Cambridge University Press.
- Daniels, Peter T. (1991) Is a structural graphemics possible? *LACUS Forum* 18: 528-537.
- (1994) Reply to Herrick. *LACUS Forum* 21: 425-31.
- Hayashi, Ōki (1977) Gendai kanazukai to okurigana. In Susumu Ōno & Takeshi Shibata (eds.), *Kokugo kokuji mondai*, 135-166. Iwanami Shoten.
- Herrick, Earl M. (1994a) Of course a structural graphemics is possible! *LACUS Forum* 21: 413-424.
- (1994b) Reply to Daniels's reply. *LACUS Forum* 21: 432-440.
- Hill, Archibald A. (1967) The typology of writing systems. In William M. Austin (ed.), *Papers in linguistics in honor of Leon Dostert*, 92-99. Mouton.
- Hoshina, Koichi (1949) *Kokugo mondai gojūnen*. San'yō Shuppan.
- Honda, Keisuke (2007) Homography-induced ambiguity in Japanese *kanji* and the lexical disambiguating function of *okurigana*. *Gengogaku Ronsō* 26: 17-42.
- Kaiser, Stefan (1995) Sekai no moji, Chūgoku no moji, Nihon no moji: Kanji no ichizuke saikō. *Sekai no Nihongo Kyōiku* 5: 155-167.
- (2007) Nihongogaku no shomondai: Moji no ruikeiron to Nihongo no hyōki. *Kokugakuin Zasshi* 108(11): 13-27.
- Kawai, Yoshifumi (1970) Kanji ni okeru yomi to kaki no paradokkusu. *Keiryō Kokugogaku* 55: 23-29.
- Kokuritsu Kokugo Kenkyūjo (1952) *Okuriganahō shiryō*. Kokuritsu Kokugo Kenkyūjo.
- (2006) *Gendai zasshi no hyōki: 1994-nen hakkō 70-shi*. Kokuritsu Kokugo Kenkyūjo.
- Kondō, Takako (2005) Kanji to okurigana. In Tomiyoshi Maeda & Masaaki Nomura (eds.), *Kanji to Nihongo*, 146-163. Asakura Shoten.
- Kōno, Rokurō (1980/1969) The Chinese writing and its influences on the scripts of the neighbouring peoples: With special reference to Korea and Japan. In Rokurō Kōno, *Kōno Rokurō chosaku zenshū* (Vol. 3), 15-102. Heibonsha.

- (1994/1977) Moji no honshitsu. In Rokurō Kōno, *Mojiron*, 2-24. Sanseidō.
- Miyaoka, Osahito (2002) 'Go' towa nani ka: *Eskimō-go kara Nihongo o miru*. Sanseidō.
- Narrog, Heiko (1998) Nihongo dōshi no katsuyō taikai. *Nihongo Kagaku* 4: 7-30.
- Nomura, Masaaki & Kikuko Itō (1978) Kanji no hyōondo. *Keiryō Kokugogaku* 11(7): 306-311.
- Ōno, Yahoko (1958) Okurigana no shomondai. *Gengo Seikatsu* 80: 16-23.
- Poser, William J. (1992) The structural typology of phonological writing. *Linguistic Society of America: Sixty-sixth annual meeting handbook*, 46. Linguistic Society of America.
- Ratcliffe, Robert R. (2001) What do "phonemic writing systems represent? Arabic huruuf, Japanese kana, and the Moraic Principle. *Written Language and Literacy* 4 (1): 1-14.
- Rogers, Henry (2005) *Writing systems: A linguistic approach*. Blackwell Publishing.
- Saito, Hirofumi, Masahiro Kawakami & Hisashi Masuda (1995) Kanji kōsei ni okeru buhin (bushu)-on'in taiō. *Jōhō Bunka Kenkyū* 2: 89-115.
- Sampson, Geoffrey (1985) *Writing systems*. Stanford University Press.
- Satake, Hideo (2006) Kanji to hyōki. In Tomiyoshi Maeda & Masaaki Nomura (eds.), *Kanji no hataraki* 44-64. Asakura Shoten.
- Sproat, Richard (2000) *A computational theory of writing systems*. Cambridge University Press.
- Stalph, Jurgen (1989) *Grundlagen einer Grammatik der Sinojapanischen Schrift*. Harrassowitz.
- Takebe, Yoshiaki (1979) *Nihongo no hyōki*. Kadokawa Shoten.
- Tsukishima, Hiroshi (1970) Gendaigo no seishohō ni tsuite: "Tōyō Kanji Kaitei On-Kun Hyō" oyobi "Kaitei Okurigana no Tsukekata" o chūshin ni. *Gengo Seikatsu* 228: 18-26.
- Uchida, Hyakken (1972/1961) Dōshi no fuhanka gobi: Shin okuriganahō ni kanren shite. In Hyakken Uchida, *Uchida Hyakken zenshū* (Vol. 8), 552-554. Kōdansha.
- Vance, Timothy J. (2002) The exception that proves the rule: Ideography and Japanese kun'yomi. In: Mary S. Erbaugh (ed.), *Difficult Characters: Interdisciplinary Studies of Chinese and Japanese Writing*, 177-193. National East Asian Language Resource Center, Ohio State University.