

## *On the moraic nature of hiragana and katakana*

Keisuke Honda

Keywords: *Hiragana*; *Katakana*; Mora; Syllable; Japanese writing system

### 1. Introduction

This paper discusses the type of phonological unit represented in *hiragana* (平仮名) and *katakana* (片仮名), collectively referred to as *kana* (仮名). More specifically, it aims to address issues surrounding the claim that *kana* are moraic writing systems.<sup>1</sup> This topic has direct relevance to the understanding of *kana* and their relation to the phonological properties of the Japanese language.

*Kana* are two separate sets of graphs (i.e., discrete and distinctive graphic marks) that are used to write Japanese vocabulary items in their phonemic forms. The graphs are used either individually or in fixed combinations to represent sequences of one or more phonemes, and vocabulary items are written in terms of such phoneme sequences. Most (but not all) of the graphs and graph combinations found in *hiragana* have phonetically equivalent counterparts in *katakana*, and vice versa.<sup>2</sup> Some examples are given in (1).<sup>3</sup>

---

<sup>1</sup>Strictly speaking, it is misleading to talk of *hiragana* and *katakana* as independent writing systems. For one thing, normally they are used in tandem with *kanji* (漢字) and other scripts to write Japanese. For another, they are designated for writing certain subsets of the Japanese vocabulary. Roughly, *hiragana* is used for non-lexical elements like particles and affixes, whereas *katakana* is used for various kinds of lexical elements like loan and mimetic words. Because it is unconventional to write Japanese entirely in *kana*, they should not be considered as writing systems in their own rights. However, it is also true that they can be regarded as ‘complete orthographies’ (Faber 1992:118-119) because it is in principle possible to write any utterance in Japanese using either *hiragana* or *katakana* (Smith 1996:210). In this sense, they can be considered as semi-autonomous writing systems, which are different components of the multi-script Japanese writing system.

<sup>2</sup>For example, *katakana* — /R/ and ウイ /wi/ have no equivalents in *hiragana*.

<sup>3</sup>Based on Shibatani (1990:160-173), Itô & Mester (1995) and other related studies, this paper makes the following assumptions about Japanese phonemics (the broad phonetic transcriptions follow the conventions of Saitô 2006:17-96): [1] The basic phoneme inventory of Japanese includes /i e a o u p b t d k g s z m n r h y w R Q N/; [2] /s/ is realised as [s] before /e a o u/ and [ɕ] before /i/; [3] /t/ is realised as [t] before /e a o/, [tɕ] before /i/ and [ts] before /u/; [4] While [ɸ] and [ts] are generally allophones of /h/ and /t/, they are contrastive in some loanwords and therefore constitute separate phonemes /h/ and /c/ in that sub-lexicon; [5] Long vowels are sequences of a given short vowel phoneme followed by /R/ (e.g., [a:] = /aR/); [6] Long or geminate consonants are sequences of a given consonant phoneme followed by /Q/ (e.g., [at]ta = /aQta/); [7] Syllable-final nasals are allophones of /N/ (e.g., [ã̃n] = /aN/, [kĩN] = /kiN/); [8] Palatalised consonants and alveopalatal obstruents are sequences of a non-palatal consonant phoneme followed by /y/ (e.g., [kʲ] = /ky/, [ɕ] = /sy/, [tɕ] = /ty/).

- (1) a. あ /a/, ん /N/, き /ki/, ぎ /gi/, きゃ /kya/, ぎゃ /gya/  
 b. ア /a/, ン /N/, キ /ki/, ギ /gi/, キャ /kya/, ギャ /gya/

*Hiragana* items are given in (1a) and *katakana* items in (1b). The transcriptions indicate that each item represents a particular phoneme sequence, and that every item in the first group has an equivalent in the second. In this light, one can assume that the two *kana* systems take the same type of phonological unit as the basic unit of representation. An important question, then, is how to describe this unit in phonological terms.

Many writing systems take a phonological unit as the basic unit of representation, and are described accordingly as phonemic (e.g., Finnish), moraic (e.g., Cherokee), syllabic (e.g., Modern Yi), and so on.<sup>4</sup> Regarding *kana*, the traditional view is that they are syllabic systems, meaning that the graphs and graph combinations represent individual syllables. This view has been expressed by many studies on the Japanese writing system (e.g., Sansom 1928:41; Miller 1967:93; Koizumi 1978:185; Hattori 1979:207; Satake 1989:1719; Smith 1996:210) and those on the world's writing systems (e.g., Gelb 1963/1952:159; Nakamura 1975:175; Sampson 1985:183ff; DeFrancis 1989:134ff). Hereafter, this view will be referred to as the *syllable-based account*.

An alternative view holds that *kana* are moraic systems based on individual morae. It has been accepted by studies on Japanese and other writing systems (e.g., Martin 1972:93; Hayata 1977:136; Poser 1992; Miller 1994:1, fn. 1; Ratcliffe 2001:3-6; Rogers 2005:61) as well as those on Japanese phonology (e.g., Vance 1987:2-3; Shibatani 1990:158). Also, several studies suggest that *kana* can or even should be viewed as moraic, despite treating them as syllabic in their own descriptions (e.g., Kabashima 1977:37-38; Coulmas 2003:80; Taylor & Taylor 1995:308; Sproat 2000:139-140; Fukumori & Ikeda 2002:42). Henceforth, this type of formulation will be referred to as the *mora-based account*.

Over the past decades, a growing number of studies have adopted the mora-based account. However, a survey of the literature reveals that many studies have taken it for granted without elaborating why it should be preferred over the traditional syllable-based account. It is only recently that studies have begun to discuss the merits and demerits of the mora-based account in relation to the syllable-based account. Given this background, the present paper examines both accounts and attempts to address some related issues that have received relatively little attention in the literature. The

---

<sup>4</sup>This does not entail that such systems represent *only* phonemes, morae, syllables, etc. It is often the case that a single writing system consists largely of graphs representing one type of phonological unit but also includes a smaller or larger number of graphs representing another type (Daniels & Bright 1996; Kōno, Chino & Nishida 2001). Besides, a phonology-based writing system may also contain a certain amount of logography, i.e. writing based on non-phonological properties of a language (Sproat 2000:141-142). It seems that there are few writing systems that can be attributed to a single type of linguistic unit, be it phonological or non-phonological (Gelb 1963/1952:199).

goal of our discussion will be to examine the arguments presented in the literature as well as to outline topics for future research.

This paper is organised as follows. Section 2 presents an overview of the relevant features of the *kana* systems and summarises the main points of the previous accounts. Section 3 provides further arguments for the mora-based account. Section 4 discusses problems arising from fixed graph combinations used to represent long vowel syllables. Section 5 makes some tentative remarks about the possibility of analysing *kana* from an economical perspective. Section 6 summarises the discussion and draws a conclusion.

## 2. Background

### 2.1 *Kana*

When used individually, most *kana* graphs represent either vowel phonemes (hereafter, *V phonemes*) or consonant-vowel sequences (hereafter, *CV sequences*). Sometimes diacritics are added to mark certain qualitative changes. These points are illustrated in (2); from now on, examples will be given in *hiragana* unless *katakana* items require particular attention.

- (2) a. あ /a/, う /u/, お /o/  
b. か /ka/, す /su/, と /to/  
c. が /ga/, ず /zu/, ど /do/

The items in (2a) represent V phonemes and those in (2b) represent CV sequences. In (2c), the diacritic <sup>˘</sup> is added to the graphs in (2b) to indicate the changes in consonant quality. Henceforth, the terms *V graphs* and *CV graphs* will be used to refer to graphs like these. The umbrella term *(C)V graphs* will also be used to cover both.

In addition, *kana* graphs may also be used in fixed combinations. Many such combinations, which may or may not be accompanied by diacritics, represent sequences of a consonant followed by /y/ and a vowel (hereafter, *CyV sequences*). Also, in *katakana* but not in *hiragana*, a number of combinations represent CV sequences that occur only in loanwords and some native mimetic words. The examples in (3) illustrate.

- (3) a. きゃ /kya/, しゅ /syu/, ちょ /tyo/  
b. ウィ /wi/, フェ /fe/, ツォ /co/ (*katakana*)

The items in (3a) represent CyV sequences and those in (3b) represent CV sequences. In both groups, each combination consists of a large graph followed by a small graph. The large graph functions as a (C)V graph when used individually (e.g., き /ki/, ウ /u/). On the other hand, the small graph is the reduced version of a (C)V graph (e.g., ゃ /ya/, い /i/). It is always used in combination with a large one as in the present examples, and never by itself. When put together, the large and small graphs function as a single unit

representing a CyV or CV sequence. Henceforth, such combinations will be referred to as *CyV* and *CV combinations* or, collectively, *C(y)V combinations*.

So far two kinds of graphs (i.e., V and CV) and two kinds of graph combinations (i.e., CyV and CV) have been distinguished. Importantly, they all represent phoneme sequences containing a single short vowel. In addition to these, both *hiragana* and *katakana* employ a small number of graphs representing non-vocalic phonemes that occur only syllable-finally (hereafter, *X phonemes*). All of these are enumerated in (4).

- (4) a. ん /N/  
b. っ /Q/  
c. ー /R/ (*katakana*)

All the graphs in (4a-c) represent *X* phonemes. The third one in (4c) is a *katakana* graph that has no *hiragana* counterpart.<sup>5</sup> These graphs are always used in combination with (C)V graphs or C(y)V combinations to represent sequences that contain the phoneme in question (e.g., とんと /toNto/, とつと /toQto/, トート /toRto/). Henceforth, they will be referred to as *X graphs*.

As just noted, *katakana* uses ー to represent the vowel length /R/. Contrastively, *hiragana* does not have any single graph designated for this phoneme. Instead, vowel length is represented by placing a V graph after a (C)V graph or C(y)V combination. Different V graphs are used according to the vowel quality, as exemplified in (5).

- (5) a. かあ /kaR/ (か /ka/ + あ /a/)  
b. とう /toR/ (と /to/ + う /u/)

Here, /R/ is represented by あ in (5a) and う in (5b). When used individually, these graphs represent the vowels /a/ and /u/, respectively. One can speak of a qualitative alteration in such *hiragana* V graphs, since they represent vowels in some cases and vowel length in others. This raises some issues regarding their treatment, which will be discussed in section 4.2.

## 2.2 Syllable-based account

As already noted in section 1, it has been widely accepted that *kana* are syllabic systems. This notion has been around since at least the late 16<sup>th</sup> or early 17<sup>th</sup> century.<sup>6</sup> Besides, it

<sup>5</sup>In actual texts, this graph is also occasionally used in *hiragana* writing. For example, かあさん and かーさん may be used interchangeably to write the word /kaRsaN/ ‘mommy’.

<sup>6</sup>For example, João Rodriguez (1561-1633), a Portuguese Jesuit missionary to Japan, remarks that the *kana* systems represent syllables (Rodriguez 1995/1620:48). In this regard, one interesting question is whether Rodriguez and his contemporaries ever thought of treating *kana* as moraic. The notion of mora originates in the Latin poetic tradition (Trubetzkoy 1958/1938:169-179). As the Jesuits were highly trained in Latin, it is probable that they were aware of mora as well as syllable

seems to have gone largely unchallenged until relatively recently when the alternative mora-based account began to gain wider attention. Consequently, it is difficult to attribute it to any particular study or studies, or to cite works specifically designed to argue in its defence. For this reason, this subsection summarises those factors that appear to underlie many versions of the syllable-based account found in the literature.

Central to the syllable-based account is the observation that a majority of *kana* graphs represent phoneme sequences containing a single short vowel. This holds for both V graphs (e.g., あ /a/) and CV graphs (e.g., か /ka/). Besides, the same can be said for graph combinations, either CyV (e.g., きゃ /kya/) or CV (e.g., フィ /fi/). Phonologically, V, CV and CyV sequences are monosyllabic in Japanese. Hence, (C)V graphs and C(y)V combinations can be described as representing individual syllables.

This syllabic principle does not apply to X graphs, namely ん /N/, ん /Q/ and *katakana* — /R/. Unlike (C)V graphs and C(y)V combinations, the X graphs represent non-vocalic X phonemes, which cannot form monosyllables by themselves. However, it is also true that these graphs are in a small minority. Presumably for this reason, many studies mention X graphs only briefly, implying that they are exceptions to the syllabic nature of *kana* (e.g., Smith 1996:211). This way, it is possible to characterise *kana* as predominantly syllabic systems that include a small number of non-syllabic graphs.

### 2.3 Mora-based account

In Japanese linguistics, it is common to describe *kana* in terms of mora rather than syllable (e.g., Martin 1972:93; Hayata 1977:136; Vance 1987:2-3; Shibatani 1990:158). Yet, there has been little discussion on why this mora-based account should be preferred over the syllable-based account. Over the past two decades, however, focus has been placed on this issue especially in the context of writing systems typology (e.g., Poser 1992; Ratcliffe 2001:3-6; Rogers 2005:61).

Central to the mora-based account is the notion of syllable weight. In Japanese, various prosodic phenomena, like accent placement in compounds, can be explained by making reference to syllable structures (Kubozono 1994, 1999; Kubozono & Ōta 1998).<sup>7</sup> Following traditional nomenclature, syllables are divided into *light*, *heavy* and *super heavy*.<sup>8</sup> Light syllables contain a single short vowel followed by no other phonemes (e.g., /ka/). Heavy syllables contain a short vowel followed by another phoneme within the syllable boundary (e.g., /kaR/, /kaN/, /kaQ/, /kai/). Super heavy

---

(Prof. Jun Ikeda, personal communication).

<sup>7</sup>Syllable weight also plays an important role in many other languages (Kenstowicz 1994:291-293).

<sup>8</sup>Two notes are in order. Firstly, super heavy syllables are often altered into heavy syllables or resyllabified into two syllables (Kubozono & Ōta 1998:66-73). Secondly, some studies posit /i/ for the syllabic [i] and /J/ for the non-syllabic [i] (hence, [ai] = /ai/ and [ai] = /aJ/) (e.g., Shibata 1962:145), while others do not make this distinction (e.g., Shibatani 1990:160-173). Although these points are important for the phonological analysis of Japanese syllables, they are peripheral to the present discussion of the *kana* systems.

syllables contain a short vowel followed by two phonemes within the syllable boundary (e.g., /kaRN/, /kaRQ/, /kaNQ/). The presence or absence of syllable-initial phoneme(s) does not contribute to syllable weight; thus, both /a/ and /ka/ are regarded as light, /aN/ and /kaN/ as heavy, and /aRN/ and /kaRN/ as super heavy.

One common way to formalise syllable weight is to assume mora, a prosodic unit intermediate between the phoneme and the syllable (Kenstowicz 1994:291-293). In moraic theory, it is assumed that vowels and syllable-final phonemes are associated with separate morae. Hence, a light syllable counts as one mora (monomoraic), a heavy syllable as two morae (bimoraic) and a super heavy syllable as three morae (trimoraic).

The basic idea behind the mora-based account of the *kana* systems is that there is a one-to-one correspondence between graphs and morae. This is illustrated in (6); the hyphen (-) in the transcriptions denotes a mora boundary.

- (6) a. Light syllable = 1 mora = 1 graph  
       あ /a/, か /ka/  
       b. Heavy syllable = 2 morae = 2 graphs  
       かあ /ka-R/, かん /ka-N/, かっ /ka-Q/  
       c. Super heavy syllable = 3 morae = 3 graphs  
       かあん /ka-R-N/, かあっ /ka-R-Q/, かんっ /ka-N-Q/

It is clear from the examples in (6a-c) that the number of graphs matches that of morae. This indicates that all *kana* graphs can be accounted for in terms of a single phonological unit, the mora.

Poser (1992) emphasises this point in his discussion of two putative syllabic systems, namely *kana* and Eskimo.<sup>9</sup> Poser points out that in both systems, a light syllable is represented by one graph and a heavy syllable by two graphs. This observation is taken to indicate that “such systems are surely mischaracterized as syllabic” (p. 46), and that they should be regarded as moraic instead. This point is echoed in several studies on writing systems typology (e.g., Miller 1994:1, fn. 1; Sproat 2000:139; Rogers 2005:61).

Ratcliffe (2001:3-6) presents a more extensive discussion of the moraic nature of the *kana* systems in his comparative analysis of *kana* and Arabic *huruuf*. As a basic assumption, Ratcliffe holds that true syllabic systems must contain distinct graphs for all types of syllables found in the language in question. *Kana* do not meet this requirement because they only contain graphs for light syllables and X phonemes but no separate graphs for heavy syllables. On this basis, Ratcliffe argues that *kana* are “organized on the principle of representing phonological time iconically”, where “each graph corresponds with a single unit of time – one mora” (p. 5). It is further noted that

<sup>9</sup>Poser (1992) is the abstract of a paper presented at the meeting of Linguistic Society of America, 1992. Although a full paper remains unpublished to date, the discussion itself has received wide attention (e.g., Miller 1994:1, fn. 1; Sproat 2000:139-140; Rogers 2005:276-277).

“the number of graphs represents the number of morae, while the qualitative differences in the shape of the graphs represent qualitative differences in the stream of speech” (ibid.).

Ratcliffe (2001:6) admits that this one-to-one correspondence between graphs and morae does not hold for C(y)V combinations like きゃ /kya/ and ファ /fa/. As these combinations consist of two distinct graphs and yet represent single morae, the formers’ relation to the latter is apparently two-to-one rather than one-to-one. Ratcliffe sees this as only a partial exception, saying that “each letter [i.e., graph] in its basic form represents only one mora, although not every possible mora can be represented with a single letter [i.e., graph]” (p. 6). In contrast, Buckley (2006:11-21) attaches more importance to this problem and raises question about the validity of treating *kana* as moraic systems.<sup>10</sup> In sections 3 and 4 below, it will be argued that C(y)V combinations do not pose a threat to the mora-based account if *kana* are analysed in terms of orthographic units rather than individual graphs.

### 3. Arguments for the mora-based account

The review of the literature presented in the previous section reveals that the syllable-based account is faced with two problems. The first one is that it forces the conclusion that *kana* take only one subset of syllables, namely the light syllable, as their basic unit of representation. On this basis, one might argue that *kana* belong to a subclass of syllabic systems that might be called something like ‘light syllabic systems’ (see footnote 10). However, this immediately leads to the second problem, namely that X graphs must be treated as exceptions. On the face of it, this second problem appears trivial because writing systems are apt to include a smaller or larger number of exceptions (see footnote 4 above). However, it would be more preferable to have an analysis that covers both exceptions and non-exceptions in a uniform and concise way.

Such an analysis is provided by the mora-based account. Its strength is that it takes a single phonological unit – the mora – to deal with all phoneme sequences represented in *kana*. For (C)V and X graphs, there is a one-to-one correspondence between graphs and morae. Admittedly, as Ratcliffe (2001) and Buckley (2006) observe, this does not hold for C(y)V combinations because they relate to morae on a two-to-one basis. However, it is also true that they, too, represent monomoraic phoneme sequences: both CyV and CV sequences are light syllables that contain a single short vowel followed by no other phonemes. Again, see section 4 for a further discussion.

---

<sup>10</sup>Buckley (2006) is the handout of the author’s presentation delivered at a Linguistic Colloquium meeting in 2006. Buckley points out several problems with the moraic reanalysis of so-called syllabic systems initiated by Poser (1992) and others. It is claimed that these writing systems should be treated as core syllabic systems based on light syllables. Sproat (2000:139-140) presents a similar analysis, which proposes to include putative syllabic systems under the category ‘core syllabary’. However, it should be noted that Sproat also suggests that *kana* might be described as moraic. In any case, these arguments require a careful examination in future research.

The validity of the mora-based account is further supported by the fact that mora plays an important role in Japanese phonology. As already mentioned in section 2.3, various prosodic phenomena in Japanese can be explained in terms of syllable weight, and this can be formalised in terms of mora. It is also important to note that heavy syllables account for about 25% of all Japanese syllables (Kubozono 1994:10). Given the obvious relevance of heavy syllables in Japanese, it is reasonable to take mora into consideration in the analysis of the *kana* systems.

While the mora-based account finds support in the literature on writing systems typology, a possible counterargument also comes from the same area of research. In section 1, it was mentioned that several studies treat *kana* as syllabic despite acknowledging the possibility (and, sometimes, plausibility) of viewing them as moraic systems. To give one example, Fukumori & Ikeda (2002:42) takes this approach in their discussion of so-called syllabic writing systems. Pointing to the non-syllabic nature of X graphs in *kana*, the authors remark that “it is also possible to consider *kana* as moraic systems” (p. 42).<sup>11</sup> Nonetheless, after presenting a comparative analysis of *kana* and other writing systems, they state their intention to treat all of these as subclasses of syllabic systems. According to one of the authors, this decision was due to the uncertain status of moraic systems in writing systems typology (Takahiro Fukumori, personal communication). As a matter of fact, traditional typologies generally do without positing the category ‘moraic’ (e.g., Gelb 1963/1952; Nakamura 1975; Sampson 1985; DeFrancis 1989), even though the notion of mora itself predates such classification frameworks (see footnote 6). Does this mean that moraic systems should be ostracised from writing systems typology for some reason?

It seems that the answer is negative. Ratcliffe (2001:1) points out that phonology-based writing systems have traditionally been classified based on “a phonological theory which recognizes only the syllable and the segment as potential units of representation”. However, with the development of moraic phonology, mora has come to be recognised as a unit of representation in many writing systems. Moreover, recently several studies have proposed that many of the writing systems traditionally labelled as syllabic should be reclassified as moraic.<sup>12</sup> As noted in section 2.3, Poser (1992) suggests that *kana* and Eskimo should be treated as moraic rather than syllabic. The same suggestion is also made for other putative syllabic systems such as Kpelle, Vai and Linear B. Based on a discussion of these and other writing systems, Poser argues that “the predominance of syllabaries has been vastly overestimated” (p. 46).<sup>13</sup> This argument is echoed by several later studies like Miller (1994:1, fn. 1) and Sproat

---

<sup>11</sup>“かなはモーラ文字であるとも考えられる” (Fukumori & Ikeda, *ibid.*).

<sup>12</sup>Modern Yi (a.k.a. Lolo) is often cited as one of the few examples of true syllabic systems (Poser 1992; Sproat 2000:140; Rogers 2005:277).

<sup>13</sup>Dissatisfaction with the classification label ‘syllabic’ has also been expressed earlier by Hill (1967:97-98), which proposes ‘polyphonemic’ as an alternative.



(2000:139-140). Rogers (2005:276-277) goes so far as to maintain that almost all of the so-called syllabic systems should be reanalysed as moraic (e.g., *kana*, Cherokee, Cree-Inuktitut, Linear B). Given these recent developments, there seems to be no longer strong reason for withholding moraic systems from writing systems typology. This, in turn, provides grounds for treating *kana* as moraic rather than syllabic systems.

#### 4. Dealing with graph combinations

As noted repeatedly, C(y)V combinations do not conform to the principle of one-to-one correspondence between graphs and morae. In section 2.3, it was mentioned that Ratcliffe (2001) treats them as a partial exception, while Buckley (2006) considers them problematic for the validity of the mora-based account. This section argues that these combinations can be handled within the framework of the mora-based account by using the notion of orthographic unit.

##### 4.1 Orthographic unit

It has already been mentioned that C(y)V combinations represent monomoraic phoneme sequences like (C)V and X graphs. This entails that C(y)V combinations differ from (C)V and X graphs only in graphic structure and not in function. This difference can be formalised by introducing Kabashima's (1977:27-29ff) notion of *hyōki yōso* (表記要素), translated here as *orthographic unit*.<sup>14</sup>

Kabashima (1977:28) defines orthographic unit as “the smallest string of graphs that would be dissociated from its corresponding sequence of sounds if segmented into smaller parts”.<sup>15</sup> It is used as an analytical tool for describing writing systems in general, regardless of their type or genealogical background. Kabashima presents an example in *kanji*, which is shown in (7).

(7) 五月雨 /samidare/ ‘early summer rain’ (Kabashima 1977:28)

In the above example, the written form consists of three distinct *kanji* graphs, namely 五, 月 and 雨. Used in this particular combination, the entire graph string represents the word /samidare/ ‘early summer rain’. Contrastively, when they are used individually, they represent different morphs or morphemes unrelated to the word /samidare/. This is shown in (8); each graph can represent more than one morph(eme)s, and this is indicated by the tilde (~) inserted between separate items.

(8) a. 五 /go/~/itu(tu)/ ‘five’  
 b. 月 /getu/~/gatu/~/tuki/ ‘moon’

<sup>14</sup>A literal translation of this term would be ‘writing element’ or ‘orthographic element’. However, this paper adopts the more widely used term ‘orthographic unit’.

<sup>15</sup>“それ以上に細分すると、音列との対応がくずれる最小の文字列” (Kabashima, *ibid.*).

c. 雨 /u/~/ame/ ‘rain’

It will be clear that the three *kanji* graphs behave as separate functional units in (8a-c), whereas the same graphs constitute a single functional unit in (7) above. Such functional units are referred to as orthographic units. Although Kabashima (1977) does not divide orthographic units into subgroups, hereafter multipartite items like that in (7) will be referred to as *complex* and unipartite items like those in (8a-c) as *simplex*.

Kabashima (1977:28ff) applies the notion of orthographic unit in his analysis of the *kana* systems. One of Kabashima’s examples is shown in (9); the original transcription has been partly modified to conform to the conventions used in this paper.

(9) ちょうちょうがとんだ /tyoRtyo ga toNda/ ‘A butterfly flew.’

(Kabashima 1977:38)

Kabashima divides this written representation into six orthographic units as in (10).

(10) a. ちょう    b. ちょ    c. が    d. と    e. ん    f. だ  
      /tyoR/    /tyo/    /ga/    /to/    /N/    /da/

The first two orthographic units in (10a,b) are complex and the rest in (10c-f) are simplex. Of these, the ちょう /tyoR/ in (10a) will be discussed in section 4.2. Regarding the ちょ /tyo/ in (10b), this CyV combination is treated as a single complex orthographic unit based on the following reasoning. It will be recalled from section 2.1 that C(y)V combinations consist of two elements: a large graph, which can represent a CV sequence by itself, and a small graph, which cannot represent anything on its own. As for ちょ /tyo/, it consists of the large ち /ti/ and the reduced version of よ /yo/. As far as the sound represented is concerned, one cannot say that this combination is a simple sum of its constituent elements. Therefore, it is treated as a complex orthographic unit.

Based on this consideration, Kabashima (1977:29) argues that the relation of writing to language in a given writing system is best understood in terms of orthographic units rather than individual graphs. The present paper follows this line of argument and suggests that *kana* should be characterised as moraic systems comprising simplex and complex orthographic units. Both classes of orthographic units represent individual morae, and the only difference lies in their structure, not function.

#### 4.2 *Hiragana* representation of long vowel syllables

Before closing this section, it is necessary to briefly discuss Kabashima’s (1977) treatment of *hiragana* representations of long vowel syllables. In (10a) in the previous subsection, it was shown that Kabashima analyses ちょう /tyoR/ as a single complex

orthographic unit. A similar analysis is given in another example, a slightly modified version of which is given in (11).<sup>16</sup>

- (11) ちょうとつきゅうひかりごう /tyoRtoQkyuR hikarigoR/ ‘Superexpress Hikari’  
(Kabashima 1977:28)

Kabashima divides the written representation into eight orthographic units as in (12).

- (12) a. ちょう    b. と    c. つ    d. きゅう    e. ひ    f. か    g. り    h. ごう  
      /tyoR/    /to/    /Q/    /kyuR/    /hi/    /ka/    /ri/    /goR/

Again, the items in (12a), (12d) and (12h) are treated as complex orthographic units representing /tyoR/, /kyuR/ and /goR/, respectively. This way, *hiragana* representations of long vowel syllables are consistently analysed as complex orthographic units.

Kabashima (1977:28) explains this analysis as follows. At first sight, it appears that these representations can be divided into a CV graph or CyV combination followed by a V graph. To take ちょう /tyoR/ as an example, it might be considered as a string of ちよ /tyo/ and う /u/. However, Kabashima rejects this analysis because ちょう represents /tyoR/ and not /tyou/, which involves a qualitative alteration in the sound represented by う (section 2.1). This observation leads Kabashima to treat ちょう as a single complex orthographic unit rather than a string of two separate orthographic units. This seems to implicate that two kinds of complex orthographic units must be distinguished in *kana*, one representing monomoraic C(y)V (e.g., ちよ /tyo/) and the other representing bimoraic C(y)VR (e.g., ちょう /tyoR/).

While this distinction is needed to deal with the kind of qualitative alteration mentioned above, it is nonetheless irrelevant to the present discussion. To see this, it will be useful to recall Ratcliffe’s (2001:5) remark that “the number of graphs represents the number of morae, while the qualitative differences in the shape of the graphs represent qualitative differences in the stream of speech” (section 2.3). Here, a clear line is drawn between the quantitative and qualitative aspects of written representations in *kana*. Replacing Ratcliffe’s ‘graphs’ with our ‘orthographic units’, it becomes clear that the qualitative alteration in う is unrelated to the quantitative value associated with this graph. That is, regardless of whether う represents /u/ or /R/, it always represents a

---

<sup>16</sup>Kabashima’s (1977) original example is とうきょうはつちょうとつきゅうひかりごう /toRkyoRhatu tyoRtoQkyuR hikarigoR/ ‘Superexpress Hikari departing from Tokyo’. The first part of this example has been omitted to avoid superfluity, even though it includes the C(y)V combinations とう /toR/ and きょう /kyoR/. Incidentally, it would be more conventional to write the word in question using *kanji* and *hiragana*, i.e. (東京発)超特急ひかり号. Needless to say, this point is irrelevant to the point under discussion.

single mora. Viewed this way, ちょう and other similar items do not undermine the notion that the *kana* systems are based on individual morae.

## 5. Phonology, economy or both?

Before concluding this paper, it may be worth mentioning the following observation. In the foregoing discussion, *kana* have been analysed from a purely phonological perspective. Needless to say, this kind of phonology-based approach is common in the linguistic analysis of *kana* and other writing systems. At the same time, it may be also possible (and perhaps meaningful) to discuss *kana* from the viewpoint of economy.

As a thought experiment, let us imagine that the *kana* systems were not possessed of X graphs, and that they instead employed distinct graphs for representing heavy syllables. In that case, there would be a large set of heavy syllabic graphs representing /ka-N/, /ka-R/, /ka-Q/, /sa-N/, /sa-R/, /sa-Q/, /ta-N/, /ta-R/, /ta-Q/, and so on. This is by no means inconceivable because many writing systems, in particular Modern Yi, actually make use of a smaller or larger number of heavy syllabic graphs (Daniels & Bright 1996; Kōno, Chino & Nishida 2001).<sup>17</sup> Yet, the downside is that this would greatly increase the total number of graphs used in the *kana* systems. In that case, *kana* would be far less economical, even if still manageable, than they actually are. Following this consideration, it might be possible to say that the real *kana* systems employ X graphs because it is more economical than to use the hypothetical heavy syllabic graphs. One might even argue that it is merely a matter of coincidence that graphs and graph combinations correspond to morae in *kana*.

Seen this way, it appears worth considering the possible influence of economy on top of that of phonology over the organisation of the *kana* systems. Having said this, it should be added that it is difficult to try to explain the real ‘motivation’ behind the way a writing system is organised. Usually a writing system is the product of linguistic activities involving an intricate interplay between social, cultural and historical factors. This holds true with *kana*, and at present it is not clear to the author of this paper how to analyse and formulate the ‘motivation’ behind their organisation – if there is such a thing. By contrast, the discussions in the previous sections have shown that it is both possible and reasonable to analyse *kana* in terms of the mora. The orthographic units correspond with individual morae either by design or by accident.

## 6. Conclusion

This paper has discussed the type of phonological unit represented in the *kana* systems. Through an examination of the previous accounts, it was shown that the mora-based

---

<sup>17</sup>This should not be confused with the situation in Korean *hankul* (한글). There, a heavy syllable is represented by a single graph (e.g., 한 *han*). However, *hankul* is not syllabic but phonemic because each element of a given graph corresponds to a phoneme in the syllable represented (e.g., ㅎ, ㅏ and ㄴ correspond to *h*, *a* and *n*, respectively) (King 1996:219).

account provides a more adequate generalisation than the traditional syllable-based account does. This was supported by the fact that mora plays an important role in Japanese phonology. Further support was obtained from recent discussions on writing systems typology, in which greater importance has been attached to the role of the mora in many writing systems. Although some previous studies have pointed to the problem of fixed graph combinations, it was argued that the mora-based account can deal with it by using the notion of orthographic unit. Finally, some remarks were made about the possibility of analysing *kana* from the viewpoint of economy. Following these arguments, this paper concludes that *kana* should be treated as moraic systems, which represent individual morae using simplex and complex orthographic units.

### **【References】**

- Buckley, Eugene. 2006. *Syllabic writing and syllable structure*. Paper presented at Linguistics Colloquium, State University of New York at Stony Brook, April 21<sup>st</sup>.
- Coulmas, Florian (2003) *Writing systems: An introduction to their linguistic analysis*. Cambridge: Cambridge University Press.
- Daniels, Peter T. & William Bright (eds.) (1996) *The world's writing systems*. New York: Oxford University Press.
- DeFrancis, John (1989) *Visible speech: The diverse oneness of writing systems*. Honolulu: University of Hawai'i Press.
- Faber, Alice (1992) Phonemic segmentation as epiphenomenon: Evidence from the history of alphabetic writing. In Pamela Downing, Susan Lima & Michael Noonan (eds.), *The Linguistics of literacy*. 111-134. Amsterdam: John Benjamins.
- Gelb, I. J. (1963/1952) *A study of writing*, 2<sup>nd</sup> ed. Chicago & London: University of Chicago Press.
- Hill, Archibald A. (1967) The typology of writing systems. In William M. Austin (ed.), *Papers in linguistics in honor of Leon Dostert*. 92-99. The Hague & Paris: Mouton.
- Itô, Junko & Armin R. Mester (1995) Japanese phonology. In John A. Goldsmith (ed.), *The handbook of phonological theory*. 817-838. Cambridge & Oxford: Blackwell.
- Kenstowicz, Michael (1994) *Phonology in generative grammar*. Cambridge: Blackwell.
- King, Ross (1996) Korean writing. In Peter T. Daniels & William Bright (eds), *The world's writing systems*. 218-227. New York: Oxford University Press.
- Kubozono, Haruo (1999) Mora and syllable. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*. 31-61. Malden: Blackwell Publishers.
- Martin, Samuel E. (1972). Nonalphabetic writing systems: Some observations. In James F. Kavanaugh & Ignatius G. Mattingly (eds.), *Language by ear and by eye: The relationships between speech and reading*. 81-102. Cambridge: MIT Press.
- Miller, D. Gary (1994) *Ancient scripts and phonological knowledge*. Amsterdam: John Benjamins.
- Miller, Roy Andrew (1967) *The Japanese language*. Chicago & London: The University of Chicago Press.

- Poser, William J. (1992) The structural typology of phonological writing. In the meeting handbook for the 66th Annual Meeting of the Linguistic Society of America, 1992. 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*. Malden, Oxford & Victoria: Blackwell Publishing.
- Sampson, Geoffrey (1985) *Writing systems*. Stanford: Stanford University Press.
- Sansom, George (1928) *An historical grammar of Japanese*. Oxford: Oxford University Press.
- Shibatani, Masayoshi (1990) *The languages of Japan*. Cambridge, New York & Melbourne: Cambridge University Press.
- Smith, Janet S. (Shibamoto) (1996) Japanese writing. In Peter T. Daniels & William Bright (eds.), *The world's writing systems*. 209-217. New York: Oxford University Press.
- Sproat, Richard (2000) *A computational theory of writing systems*. Cambridge: Cambridge University Press.
- Taylor, Insup & M. Martin Taylor (1995) *Writing and literacy in Chinese, Korean and Japanese*. Amsterdam & Philadelphia: John Benjamins.
- Trubetzkoy, N. S. (1958/1938) *Gründzuge der Phonologie*, 2<sup>nd</sup> ed. Göttingen: Vandenhoeck and Ruprecht in Gottingen.
- Vance, Timothy J. (1987) *An introduction to Japanese phonology*. Albany: State University of New York Press.
- 樺島忠夫(1977)「文字の体系と構造」大野晋・柴田武(編)『文字』23-60. 東京: 岩波書店.
- 窪菌晴夫(1994)「日本語の音節量について」『国語学』178. 7-17.
- 窪菌晴夫・大田聡(1998)『音韻構造とアクセント』東京: 研究社出版.
- 小泉保(1978)『日本語の正書法』東京: 大修館書店.
- 河野六郎・千野栄一・西田龍雄(編)(2001)『世界文字辞典』東京: 三省堂.
- 斎藤純男(2006)『日本語音声学入門』(改訂版)東京: 三省堂.
- 佐竹秀雄(1989)「現代日本語文字・表記」亀井孝・河野六郎・千野栄一(編著)『言語学大辞典』(第2巻, 世界言語編, 中, さ-に) 1716-1723. 東京: 三省堂.
- 柴田武(1962)「音韻」国語学会(編)『方言学概説』137-161. 東京: 武蔵野書院.
- 中村完(1975)「文字論」田中春美・家村睦夫・五十嵐康男・倉又浩一・中村完・樋口時弘『言語学入門』171-185. 東京: 大修館書店.
- 服部四郎(1979)『音韻論と正書法』東京: 大修館書店.
- 早田輝洋(1977)「日本語と表音文字」林大・林四郎・森岡健二(編)『文字と表記』135-159. 東京: 明治書院.
- 福盛貴弘・池田潤(2002)「文字の分類案——一般文字学の構築を目指して——」『一般言語学論叢』4. 33-56.
- ロドリゲス・ジョアン(1995/1620)『日本語小文典』(池上岑夫, 訳)東京: 岩波書店.