

Radosław M. NOWAK

Adam Mickiewicz University in Poznań

 <https://orcid.org/0000-0003-1691-4694>
e-mail: radoslaw.nowak@amu.edu.pl

Submitted: 15.03.2025

Accepted: 22.04.2026

Published: 15.07.2026

The ambiguous descriptions of Japanese morphography and agglutination: a critical analysis of relations between graphemics and morphology in the Japanese language

ABSTRACT

In a *morphographic script*, graphemes represent morphemes; in an *agglutinative language*, monosemous morphemes attach onto lexical ones to convey grammatical meanings. Although these definitions are generally accepted, the former lacks the nuance needed to accurately describe grapheme–morpheme relations within the Japanese writing system, and the latter is often overlooked in accounts of Japanese morphosyntax. This paper identifies these mismatches and draws connections between Japanese morphography and descriptions of Japanese morphology. Sections 1 and 2 examine how *kanji* characters may correspond to exactly one morpheme, more than one morpheme, or less than one morpheme; and how different scripts represent different morpheme types. Section 3 outlines logical and terminological inconsistencies in Japanese grammars. Section 4 links features of the writing system to the ways Japanese morphosyntax is analyzed. The paper concludes that treating certain inflectional morphemes as lexical items may stem from the “division of labor” between morphographic *kanji* and phonographic *kana*.

KEYWORDS: agglutination, Japanese grammar, Japanese language, Japanese writing system, *kanji*, morphography



Creative Commons Attribution 4.0 International (CC-BY):
<https://creativecommons.org/licenses/by-nc/4.0/>

Introduction

This paper has two main objectives: to establish what *kanji* 漢字 (and, to a lesser extent, *kana* 仮名) characters actually represent within the Japanese writing system and to highlight the connections between the linguistic units represented by units of writing and the ways Japanese grammars¹ describe agglutinative phenomena within the language. Unfortunately, existing accounts on both Japanese morphosyntax as well as the typology of its writing system are inconclusive and fragmented, leaving the academic discourse with an imperfect understanding of these areas of the language. This paper argues that at least some of the underlying causes of this situation stem from the Japanese writing system itself, highlighting the numerous ways that *kanji* characters can relate to units of the Japanese language, and examines how the mixed, morphomoraic nature of the Japanese writing system as a whole has influenced the interpretation of Japanese nominal and verbal inflection.

Several trends can be observed in the descriptions of Japanese as a language, including its writing system. Morphologically, it is widely regarded as agglutinative (Coulmas 2018a: 117; Hasegawa 2015: 3; Hasegawa 2018: 3; Horie 2018: 65; Huszcza et al. 2003: 115; Jabłoński 2021a: 99–100, among others), while its writing system is described as mixed as a whole (Coulmas 2003: 182; Handel 2019: 6, 196; Hasegawa 2015: 50; Rogers 2005: 275; Seeley 1991: 1) with the subset of *kanji* characters typically considered morphographic (Daniels 2018: 85; Joyce and Masuda 2019: 249; Rogers 2005: 272; Wasilewska 2019: 201; for other descriptors, see 1.2.). This paper contrasts how these terms are usually understood with actual graphemic and morphological phenomena in the Japanese language. Common definitions of these terms are presented below as a basis for this analysis.

Inflection is the process of adding affixes (bound morphemes² which can also be called *suffixes* or *endings* in the case of Japanese) to lexical

¹ The term *grammar* is understood in this paper as “a comprehensive description of a language, focusing on, but not limited to, its morphology and syntax”. Similarly, a *grammarian* is a person concerning oneself with such a description. *Descriptive grammar* means “the description of a language as it is”, as opposed to *prescriptive grammar* meaning “the description of a language as it should be”. The term *grammar* is sometimes also used to jointly refer to morphology and syntax, as opposed to other areas of language such as phonology or semantics (Crystal 2008: 218). Such use is avoided in this paper, barring the term *grammatical* adopted as a commonly used synonym for *morphosyntactic*.

² A *bound morpheme* is one that cannot occur on its own (Crystal 2008: 59). This paper adopts the term chiefly in order to highlight the non-independent status of these units, a fact that will be crucial in Section 4.

morphemes (*stems*) in order to convey grammatical meanings without changing the grammatical class of the morphemes to which they are attached (Crystal 2008: 243–244; Fromkin et al. 2017: 46–48). Moreover, this process is regular categorically, semantically, and systemically³ (Grzegorzczkova 2008: 95–96). This is opposed to *derivation* which creates new lexical units – often belonging to a different grammatical class than the original morpheme – in an irregular fashion⁴ (Fromkin et al. 2017: 44; Grzegorzczkova 2008: 95–97).

Inflection can be further divided into *agglutinating inflection* and *fusional inflection*. In the former, inflectional morphemes are monosemous (each has only one meaning), the boundaries between morphemes are usually clear-cut, and each grammatical function is typically expressed with a single morpheme (Comrie 1989: 43–45; Grzegorzczkova 2008: 187). In languages exhibiting fusional inflection, one inflectional morpheme can encode multiple meanings; functionally distinct morphemes may be homophonic; and a single grammatical meaning can in some cases be expressed with more than one morpheme (Grzegorzczkova 2008: 187–188). The terms *agglutination* and *fusion* are used in this paper in this stricter sense, referring to types of inflection and excluding derivational processes; this narrow understanding of *agglutination* can be also found in other sources (Grzegorzczkova 2008: 187; Jabłoński 2022: 12; Polański 1999: 21; Tanaka et al. 1988: 19–20; Velupillai 2012: 458).

Morphography, on the other hand, is generally defined as the use of distinct graphemes to represent different morphemes. The term also applies to primarily phonographic systems where alternate spellings are used to differentiate homophonic morphemes. It should not be confused with orthographic depth, which concerns the consistent spelling of a single morpheme across its allomorphs (Rogers 2005: 275). As previously noted, *kanji* are generally considered morphographic, and thus may be termed as

³ *Categorical regularity* means that there exist different but regular sets of rules for various *classes* of lexemes such as nouns, verbs, adjectives, or their subclasses. *Semantic regularity* signifies that the change in meaning of an inflected words is regular, for example the *-ing* morpheme attached to the stems of English verbs always points to the progressive aspect or gerund. Finally, *systemic regularity* means that all slots in the inflectional *paradigm* of a lexeme can be filled regardless of whether such word-forms are pragmatically useful. For example, every noun in Polish can be inflected in the vocative case used mostly to directly address people – even if the noun in question signifies an object that is unlikely to be conversed with directly (Grzegorzczkova 2008: 95–97).

⁴ For this reason, bound morphemes can be further classified as either *derivational* or *inflectional*. *Derivational morphemes* are generally semantic in nature, combining with lexical morphemes in irregular ways that cannot be presented as a consistent paradigm. *Inflectional morphemes*, on the other hand, are mostly syntactic in nature, attaching to lexical morphemes in a regular, systematic manner to form inflectional paradigms (Bloomfield 1973: 159–160; Grzegorzczkova 2008: 92–93).

morphograms or, less commonly, *morphemograms*. The typologically neutral term *sinogram*, highlighting their genealogical link to the Chinese script, is also adopted throughout this paper as a synonym for *kanji character* in line with its usage in recent literature (Handel 2019: 2; Jabłoński 2022: 19; Wasilewska 2019: 17; Wnuk 2023: 74), and for convenience.

Several other basic terms are also used in this paper. A *lexeme* is understood as an abstract unit of the lexicon of a language with distinct lexical (i.e. not morphosyntactic) meaning. One lexeme can contain within itself various *word-forms*—iterations of the same *lexeme* differentiated by the morphosyntactic information attached to it. For example, *die*, *dies*, and *died* are three word-forms of the same English lexeme *die* (Matthews 1991: 26–31). The term *word-form* is furthermore often shortened to *form* in this paper to avoid redundancy.

The following sections of this article examine this terminology in detail in the context of the Japanese writing system, language, and their descriptions. Sections 1 and 2 outline the typology of the Japanese writing system, focusing on regularities and irregularities in grapheme-morpheme correspondence. When analyzing grapheme-morpheme relations, this paper takes an analytic approach rather than a holistic approach, focusing on individual examples of *kanji* use in order to introduce nuance into the discussion of Japanese morphography. Section 3 addresses various approaches to Japanese nominal and verbal morphology, highlighting inconsistencies in traditional accounts. Section 4 juxtaposes these descriptions of Japanese morphology with the features of the writing system discussed in Sections 1 and 2. The paper concludes that these features may partly account for the incongruences identified in Section 3.

Block examples, separated from the main text, are numbered and consist of five lines of text: a romanized morphemic transcription; a morphemic breakdown of the example; a graphemic transcription; the original conventional spelling; and a translation. In the graphemic transcription, the parts written in *kanji* are indicated with small caps (EXAMPLE) while sinogram and script boundaries are indicated with vertical lines (|). In one case (Example 1-1), an additional line is included between the morphemic breakdown and graphemic transcription, providing a lexical breakdown of the compound noun analyzed as the noun in question consists of otherwise independent nominal elements. Non-existent phrases and spellings are marked with an asterisk (*). Inflectional morphemes are glossed following the terminology of Huszcza et al. (2003). Finally, all example words and phrases are provided by the author of this paper according to his best linguistic knowledge, unless stated otherwise. As Japanese often allows for

multiple spellings of the same word-form, to ensure consistency all example spellings come from the two dictionaries listed at the end of the reference section – where the full list of entries referenced in this paper can also be found.

1. General typology of the Japanese writing system – the moraic *kana* and (mono)morphographic *kanji*

As mentioned in the introduction, the Japanese writing system is widely described as mixed, with the subset of *kanji* characters labeled *morphographic* (Daniels 2018: 85; Rogers 2005: 272; Wasilewska 2019: 201), *logographic* (Frellesvig 2010: 13), or even *ideographic* (Hasegawa 2015: 44; Jabłoński 2021a: 13), and the subset of *kana* characters as *moraic* (Handel 2019: 221; Honda 2009: 215) or *syllabic* (DeFrancis 1989: 58; Hasegawa 2015: 47). This paper advocates for the consistent use of the terms *morphographic* and *moraic* respectively, as they capture the nature of the Japanese scripts the most accurately. The rationale is discussed in the following subsections.

1.1. The simple: the moraic *kana* scripts and basic conventions of Japanese text composition

Although the typology of *kana* scripts is not the main interest of this paper, a brief discussion of the topic is necessary to provide context for the following sections. A *mora* is a unit of rhythm that underpins the Japanese phonological system equal in length to a short open syllable; a consonantal coda, a long vowel, or a long consonant lengthen the syllable by one mora each. The Japanese pitch accent is mora-based, as pitch shifts occur only between moras, even within a single syllable (Honda 2009: 215; Huszcza et al. 2003: 62). This phonological structure likely shaped the development of the native *hiragana* 平仮名 and *katakana* 片仮名 scripts, in which each grapheme typically represents one mora. As a result, syllables may be written with one or more graphs depending on their moraic length (Coulmas 2018a: 115). For instance, the monosyllabic onomatopoeia *fūn* フーン ‘hmm’ standing for the sound of vocal deliberation or mild surprise is written with three graphs, reflecting its three moras. Therefore, *hiragana* and *katakana* are best described as *moraic scripts* (Honda 2009: 215; Rogers 2005: 61), or perhaps *moraries*—a rare term in English literature but attested in other languages (cf. Polish *morariusz*) (Wasilewska 2019: 203).

This convention is broken only in the case of digraphs used to write moras with palatalized consonants followed by *a*, *u*, or *o*⁵, such as *kya* きゃ／キヤ, *shu* しゅ／シュ, or *myo* みょ／ミヨ (Hasegawa 2015: 54). Even then, the special non-moraic status of the graphemes *ya* や／ヤ, *yu* ゆ／ユ, and *yo* よ／ヨ is highlighted by writing them in lowercase. The exception to this exception is the lowercase *tsu* つ／ツ⁶, which is used to signal that the following consonant is long (geminated) and constitutes a standalone mora (Huszcza et al. 2003: 62).

The trends of script usage in modern Japanese are as follows. *Hiragana* is, so to speak, the default *kana*. In texts intended for adult, fluent users of Japanese, it is mostly used to write inflectional affixes (while lexeme stems are typically written in *kanji*), many of the lexical adverbs, as well as other units that either have no *kanji* spelling(s) or whose *kanji* would be too difficult for an average reader (Seeley 1991: 1–2). It is also used in materials intended for readers whose command of *kanji* characters is not sufficient to read texts in the conventional orthography. *Katakana*, on the other hand, is used to write “scientific terms, expressive vocabulary, names and loan words from languages other than Chinese, and for emphasis” (Hasegawa 2015: 54). This method of composing written texts using a mixture of *kanji* and *kana* where each script has its own distinct function is called *kanji kana majiribun* 漢字仮名交じり文 ‘writing with a mix of *kanji* and *kana* characters’ (Frellesvig 2010: 157, 266). Additionally, the final mora(s) of a mixed *kanji-kana* lexeme written in *hiragana* are traditionally called *okurigana* 送り仮名. Beside representing inflectional endings, the role of *okurigana* is to clarify the pronunciation of the *kanji* character it is attached to (Kindaichi et al. 1988: 315). For example, in the word-form *deru* 出る ‘[one] leaves’, the first *kanji* grapheme 出 represents the stem of the verb *de-* and the second *hiragana* grapheme る – the *okurigana* portion of the word – represents the inflectional affix of the non-past tense. At the same time, it disambiguates the reading of the *kanji* 出 as *deru* is the only word-form written with the combination of 出 and る in this order. This disambiguation is necessary as many Japanese morphemes are homographic, i.e. they are written with the same *kanji* character. The character 出 itself may stand for several morphemes and can be read in multiple different ways depending on the

⁵ Due to space constraints, novel *katakana* character digraphs used to write non-native moras are omitted from discussion.

⁶ Graphs in this paragraph are organized as follows: *hiragana* / *katakana*.

context, such as *da* in *dasu* 出す ‘to take out’ or *gai* in *gaishutsu* 外出 ‘[the act of] going out’.

However, this disambiguating function is mostly limited to the conjugated forms of verbs and adjectives, as well as some select nouns and adverbs. It is not typically used to disambiguate nominal morphemes of both Sino-Japanese and native Japanese origin such as *kō* 行 ‘to go; to carry out’ and *gyō* 行 ‘to go; to carry out; line’; *mon* 門 ‘gate’ and *kado* 門 ‘gate’; or *in* 印 ‘seal; mark’ and *shirushi* 印 ‘mark’; as well as some verbs and their inflected forms. This exacerbates the problem of homography within the Japanese writing system (Honda 2009: 226–231).

Despite the existence of digraphs, the typology of the *kana* scripts and the overarching Japanese writing system is relatively clear. While some terminological variation exists across sources, most scholars agree on the basic grapheme-linguistic unit relationship in *kana*, which is best described as *moraic*. The Japanese writing system is mixed, employing several different scripts. The “division of labor” between *kanji*, *hiragana*, and *katakana* described in the paragraphs above, however, is a simplification of the much more complex reality of *kanji* and *kana* usage; a few examples of the intricacies thereof are described below.

1.2. The regular: one-to-one grapheme-morpheme correspondence in basic nouns and adjectives written with *kanji*

As previously noted, *kanji* characters are described in varying ways across the literature. One of the oldest – and most problematic – labels is *ideographic*. Once common in Western accounts of Chinese characters (Coulmas 1999: 82, 224–225; Wnuk 2024: 44), the term is now controversial as it implies that such characters encode ideas directly, bypassing linguistic representation. This notion undermines the very definition of writing as inherently tied to language⁷ (Handel 2019: 4, 41).

Instead, the terms *morphographic* and *logographic* are generally advised, as both emphasize that graphemes described as such correspond to meaningful linguistic units (ibid., 8). The term *logographic* is contested however, as no known script consistently uses graphemes to represent entire words as the term would imply (Osterkamp and Schreiber 2021: 48; Rogers 2005: 14; Sampson 1985: 39). Moreover, while some graphemes may indeed represent

⁷ Technically, the term “ideogram” might be applied to a category in Xu Shen’s (c. 58–c. 148 CE) traditional, etymological classification of sinograms, the *shiji* 指事 ‘characters pointing to abstract things’, i.e. ideas. Even then, scholars avoid this term due to its association with non-linguistic symbolism, preferring alternatives such as *characters denoting concepts* (Wasilewska 2019: 44) or simply *iconic* (Handel 2019: 48).

multiple morphemes (Handel 2019: 61), the term *word* itself remains contentious as there is no universally accepted definition that one could use to objectively determine what constitutes a single word and what does not (Elbourne 2011: 1; Polański 1999: 645–648). As Joyce (2011: 71) points out, many compound lexemes in Japanese that would traditionally be considered independent *words* – such as *yamazakura* 山桜 ‘mountain cherry’, *tozan* 登山 ‘mountain climbing’, or *hitobito* 人々 ‘people’ – can be divided into constituent morphemes, each written with a single character.

Although the definition of a *morpheme* as the smallest meaningful unit of language (Fábregas and Scalise 2025: 1, 26; Matthews 1991: 103) can still be quite vague, as what qualifies as a morpheme may differ from a synchronic and a diachronic perspective (see subsections 2.1. and 2.2. for examples from the Japanese language), it seems a more appropriate terminological base than a term with no clear definition. This paper therefore endorses *morphography* and *morphographic* as the standard terms for scripts where morphemes are the primary linguistic unit represented. Admittedly, the number of morphemes represented by a single grapheme can vary (Handel 2019: 61). When individual graphemes represent more than one morpheme, the term *polymorphemic* is recommended; when they represent exactly one, *monomorphemic* should be used⁸.

The one-to-one correspondence between graphemes and morphemes holds up for many simple (i.e. not compound) nouns and predicative adjectives⁹. Most basic Japanese nouns – such as *mizu* 水 ‘water’, *kaze* 風 ‘wind’, *oya* 親 ‘parent’, *ie* 家 ‘home, house’, *saka* 坂 ‘slope’, *machi* 町 ‘town’, *kasa* 傘 ‘umbrella’, or *koi* 恋 ‘romantic love’ – are monomorphemic and written with exactly one grapheme in their uninflected forms. Hence, the graphemes themselves can also be considered monomorphemic in these contexts (or, perhaps more precisely, *monomorphographic*¹⁰).

⁸ These terms themselves are rare but not novel, being attested, for instance, in Handel (2019: 9, 21) or Osterkamp and Schreiber (2021: 48).

⁹ This part of speech is traditionally known as the *i-adjective* (Hasegawa 2015: 66), referencing its dictionary form ending in *i* い. This paper, however, adopts the term *predicative adjective* from Huszcza et al. (2003: 146), which highlights its ability to function as a standalone predicate. A similar term, *inflected adjective* (Jabłoński 2022: 164), underscores its inflectional properties, as opposed to the other adjective type, often called *na-adjective*, *nominal adjective*, or *adjectival noun* (Hasegawa 2015: 66; Jabłoński 2022: 164), which requires the copula in order to function as a predicate.

¹⁰ This term would emphasize that the grapheme does not *itself* consist of a single morpheme but rather *represents* one. However, in papers focusing specifically on graphemics such as this one, such a distinction may appear verbose; thus, the simpler term is adopted. The same rationale applies to *polymorphemic* and *semimorphemic*.

Similarly, although basic Japanese predicative adjectives in their so-called citation forms consist of two morphemes, one lexical and one inflectional marking the non-past tense, a clear one-to-one grapheme-morpheme correspondence remains observable. For example, in *akai* 赤い ‘[is] red’, the lexical morpheme *aka-* is written with the *kanji* character 赤, while the inflectional morpheme is separated from the stem through the use of the *hiragana* moragram *i* い. The same holds true for many more native predicative adjectives such as *aoi* 青い ‘blue’, *samui* 寒い ‘cold [of weather]’, *atsui* 熱い ‘hot [to touch]’, *yoi* 良い ‘good’, *warui* 悪い ‘bad’, *furui* 古い ‘old [of a thing]’, *karai* 辛い ‘spicy [of food]’. There are, however, many exceptions to this rule (see 2.1. for further discussion).

The most salient and regular one-to-one grapheme-morpheme relationship can be seen in so-called Sino-Japanese vocabulary. These are lexemes borrowed from Chinese throughout Japanese history, as well as words created in Japan using morphemes borrowed from Chinese (Kindaichi et al. 1988: 423). As can be seen in 1-1, they can be quite easily divided into constituent morphemes.

1-1	<i>koku-ritsu</i>	<i>koku-go</i>	<i>ken-kyū-jo</i>
	country-establish	country-language	to.hone-to.seek-place
	‘national’	‘national language’	‘research institute’
	KOKU RITSU	KOKU GO	KEN KYŪ JO
	国立	国語	研究所
	‘National Institute for Japanese Language and Linguistics’		

This complex compound noun consists of three otherwise independent lexemes, *kokuritsu* 国立 ‘national’, *kokugo* 国語 ‘national language’, and *kenkyūjo* 研究所 ‘research institute’¹¹, each of which is built from two or three bound Sino-Japanese morphemes. Morpheme boundaries neatly coincide with character boundaries, forming a perfect one-to-one relationship. This is an almost universal pattern in Sino-Japanese words, mimicking the grapheme-morpheme relationships in Chinese (Handel 2019:

¹¹ Although these lexemes can indeed be considered independent units of the Japanese lexicon overall, when compounded in this way, without any inflectional markers that would typically indicate syntactic relationships between them, they can also be viewed as a single lexical unit, *Kokuritsu-kokugo-kenkyūjo* 国立国語研究所 ‘National Institute for Japanese Language and Linguistics’, and in fact, they frequently function as such. Combining compound Sino-Japanese lexemes to create new, even more complex lexical units is a practice widespread in the Japanese language (Kageyama and Saito 2016: 23, 32).

31), and I suggest the term *monomorphemic* be used when describing such characters as well. Despite not being part of the original Japanese morphemic system, these Sino-Japanese units can be considered morphemes because of their relative regularity of meaning across lexemes in which they appear. Additionally, they remain productive and can be used to create new pseudo-Sino-Japanese lexical compounds. Although it is rather uncommon in modern Japanese as direct loanwords are favored instead, new Sino-Japanese compounds are still created *ad hoc* quite often in art and literature. However, the most important reason why these morphemes retain their morphemic status may be their close association with sinograms, which are widely considered to be meaningful units. A salient proof of this is the fact that Sino-Japanese morphemes are often understood and taught as Sino-Japanese *readings* of *kanji* characters (or *on'yomi* 音読み ‘sound readings’), as opposed to native Japanese readings (or *kun'yomi* 訓読み ‘learned readings’) (Kobayashi et al. 2016: 94; Rose 2017: 21–22; Taylor and Taylor 2014: 276–277). In other words, when learning to write in Japanese you start from the meaningful character and not its corresponding unit of speech; the principal question is not *how to write* a lexeme, but rather *how to read a kanji* – it is implicitly assumed that *kanji* do have meanings. This creates a kind of feedback loop that causes the association between graphemics and semantics to persist even when the unit of speech assigned to a character has been desemanticized. This is why the term *graphomorpheme* is sometimes used to refer to a single (Sino-)Japanese morpheme and its assigned sinogram simultaneously, highlighting their functional unity (Jabłoński 2021a: 125) and monomorphemicity.

2. In-depth typology of the Japanese writing system: deviations from the monomorphemic principle in *kanji*

Although the monomorphemic *kanji* and clearly phonographic *kana* constitute the basis of the Japanese writing system, sinograms can be used in many more ways than the simple grapheme-morpheme relationship discussed in (1.2.). These include writing multiple morphemes with a single character; incorporating part of a lexical morpheme into *okurigana*; and the use of polygraphs. Such cases are discussed in this section.

2.1. The irregular: semimorphemic graphemes, polymorphemic graphemes, and polygraphs

The first significant deviation from the monomorphemic principle is in how parts of the stem of an inflected lexeme may be incorporated into *okurigana*. Despite the reforms introduced shortly after World War II with the intention

of aligning spelling conventions with Japanese speech (Seeley 1991: 153), some etymological features of the former Japanese orthography have been retained to this day. For this reason, there exists a subclass of predicative adjectives ending in *-shii* whose stems are not fully written in *kanji* (see 2-1).

- 2-1 *kanashi-i*
 sad-NPST
 KANA|shii
 悲しい
 ‘[is] sad’

One of the word-forms of this adjective type used to be the bare stem, without any inflectional ending attached (for example: *kanashi* 悲し ‘[is] sad’) (Frellesvig 2010: 90–93). As such, the last mora of the stem came to be written as *okurigana*, most likely as a part of the disambiguating principle discussed in (1.1.); this writing convention has made its way into modern orthography. The many other adjectives of this kind include *atarashii* 新しい ‘new’, *kuyashii* 悔しい ‘regrettable’, *ureshii* 嬉しい ‘joyous, happy’, *munashii* 空しい ‘empty, hollow’, *sabishii* 寂しい ‘lonely’.

Another deviation from the one-to-one grapheme-morpheme correspondence in adjectives stems from lexicalization of former inflectional markers. For instance, the adjective *ōkii* 大きい ‘big’ is derived from the former adnominal form *ōki* 大き of the archaic adjective *ōshi* 大し ‘big’ (Daijirin 1995). The modern spelling reflects this etymology, even though synchronically, the mora *ki* can be considered part of the stem. Similarly, derived adjectives such as *urayamashii* 羨ましい ‘envious’ or *natsukashii* 懐かしい ‘nostalgic’ – created from the verbs *urayamu* 羨む ‘to envy’ and *natsuku* 懐く ‘to become familiar with’ respectively (*ibid.*) – also have traces of their etymology visible in their spellings. Although the morpheme *-(a)shi-* is no longer productive and as such could be considered part of the stem in modern Japanese, the current spelling of these adjectives retains the etymological relationship, leading to an ambiguous grapheme-morpheme relationship. This spelling convention does, however, help to keep the *readings* of characters consistent across different lexemes using the same sinogram. In both *natsuku* 懐く ‘to become familiar with’ and *natsukashii* 懐かしい ‘nostalgic’, it is the *natsu-* portion of the word that is written in *kanji*. At the same time, it limits the already significant number of possible readings ascribed to a single character.

I have not identified any publications discussing this phenomenon in terms of grapheme-morpheme relations, which means the academic community is left without proper terminology to describe it. I propose the term *semimorphemic* as a descriptor for such situations, as it is only part of a morpheme that is represented by its assigned grapheme in these cases. On the surface, this is a deviation from the very definition of morphography as a feature of a script in which graphemes correspond to (whole) morphemes. However, it is undeniable that the specific part *natsu-* of the morpheme *natsuk-* can only be represented by this specific *kanji* character (excluding an entirely phonographic representation with *kana*), and that this character has been assigned to it on the basis of the meaning of the morpheme it is a part of. On the other hand, the character 懷 is never used to represent the sequence *natsu-* in lexemes that are morphologically and semantically unrelated to the morpheme *natsuk-*¹². In this way, the core notion of morphography – that graphemes of a morphographic script should in some way represent meaningful units of language – persists.

The opposite situation occurs when one grapheme is used to represent two or more morphemes. This includes lexicalized compounds of otherwise free lexical morphemes, as well as combinations of lexical and derivational morphemes, usually frequent and highly lexicalized, in which the latter was “incorporated” into the grapheme for ease of writing. Examples of such cases are provided below.

2-2	<i>kokoro-</i>	<i>-zash-i</i>
	heart-	-to.point.to-NMN
	KOKOROZASHI	
	志	

‘aspiration, goal, a thing one strives for’

2-3	<i>mizu-</i>	<i>-umi</i>
	water-	-sea
	MIZUUMI	
	湖	

‘lake’

¹² Alternatively, one could interpret the *kanji* character and the *kana* characters following it as representing the morpheme *natsuk-* (or *natsukashi-*) jointly, making *natsukashi-* 懷かし a trigraph. Nonetheless, the *kana* characters can be assigned their respective moras regardless of context. Thus, if one maps parts of the morpheme onto singular graphemes prioritizing the obvious, i.e. the *kana* characters, it is hard not to come to the semimorphemic interpretation.

- 2-4 *aka-* *-gane*
 red- -metal
 AKAGANE
 銅
 ‘copper’
- 2-5 *hanash-i*
 to.talk-NMN
 HANASHI
 話
 ‘conversation, story’
- 2-6 *shirush-i*
 to.point.to-NMN
 SHIRUSHI
 印
 ‘mark, symbol’

The lexicalization of the compounds above is accentuated by a complete graphemic separation from their etymology (barring 2-5 and 2-6); this can be seen clearly when presented with their etymological but non-normative spellings, which would be *kokorozashi* *心指し, *mizuumi* *水海, and *akagane* *赤金 respectively (Daijirin 1995). Moreover, the lexemes in examples 2-2 through 2-6 are not just lexicalized compounds of archaic morphemes. All these morphemes are still widely used in modern Japanese, leaving no doubt about their morphemic status. I suggest that the term *polymorphemic* be used when describing such characters. Although this term has sporadically been used in the literature (Handel 2019: 61; Wasilewska 2019: 222), it remains uncommon. This term can also be used to describe cases in which a character represents more than one morpheme but less than two. For example, in the diachronically compound verb *shirizoku* 退く ‘to retreat’, 退 represents the morpheme *shiri* ‘back’ and the *zo-* portion of *zok-* ‘to move away’¹³:

¹³ Ambiguities regarding verb stem representation are explored further in Subsection 2.2.

2-7 *shiri-* *-zok-u*
 back- -to.move.away-NPST
 SHIRIZO|ku
 退く
 ‘to retreat’

Finally, there are the so-called *jukujikun* 熟字訓. These can be defined as spellings¹⁴ of native Japanese words using *kanji* characters selected solely for their combined semantic value (Sato 2018: 314, 320). No single morpheme, or even part of a morpheme, can be assigned to individual characters in these lexemes (Honda 2021: 191). A non-exhaustive list of these is included in the *jōyōkanjihyō* 常用漢字表, the official list of characters and readings to be used in public documents and newspapers (Bunkachō 2010a). Examples include *kesa* 今朝 ‘this morning’, *ototoi* 一昨日 ‘the day before yesterday’, *asatte* 明後日 ‘the day after tomorrow’, *tsuitachi* 一日 ‘the first day of the month’, *Tanabata* 七夕 ‘the Tanabata festival’, *itoko* 従兄弟 ‘cousin’, and *Yamato* 大和 ‘Japan’¹⁵.

¹⁴ The Japanese term approaches the subject from the opposite perspective: *jukujikun* 熟字訓 literally means ‘native Japanese *readings* of compound characters’, marking the *pronunciations* of these words as the irregular element, while from the speech-comes-first point of view, it is their *spellings* that are irregular. This convention is understandable, however, as it reflects the view that it is the *kanji* that have readings and not *morphemes* that have assigned characters, as noted above. This in turn results from the process by which Chinese characters were adapted to the Japanese language: scribes were initially more concerned with how to *read* sinograms in Chinese texts than with how to *write* Japanese words. A salient proof of this approach is the graphemic overdifferentiation of certain lexemes. For instance, the lexeme *oba* ‘aunt’ can be written as 伯母 if the aunt in question is an older sibling of one’s parent and as 叔母 if she is a younger sibling. In Japanese, the differentiation is purely graphemic and results from assigning the same reading to two phonetically and graphemically distinct Chinese lexemes. Notably, *oba* is also a case of *jukujikun* (Frellesvig 2010: 263–266, 268).

¹⁵ To add more nuance to this discussion, some of these readings can be traced back to *kambun kundoku* 漢文訓読 ‘[the practice of] reading Chinese texts in Japanese’, a type of simultaneous translation of classical Chinese texts into Japanese using conventionalized translations of Chinese characters (Frellesvig 2010: 258–274). However, it is not a universal rule. Although rare, *jukujikun* also include modern loanwords, such as the lexeme *tabako* 煙草 ‘cigarette’ from Portuguese *tabaco* (Daijirin 1995) written with the two characters 煙 ‘smoke’ and 草 ‘grass’ (or, more commonly, in *katakana* as タバコ). Moreover, some scholars argue that *kanji* in *jukujikun* directly represent “meanings”, closely resembling the notion of ideography mentioned earlier. For example, Sato (2018: 320) claims that “the meanings of the two characters, *now* (今) and *morning* (朝), represent the word meaning *this morning*”. However, it remains uncertain to what extent it can be argued that a specific unit of meaning is really assigned to each character, or what “units of meaning” the lexeme

Some of the so-called *jukujikun* can be analyzed into constituent morphemes diachronically, making it somewhat plausible to connect those morphemes to single *kanji* characters. However, synchronically, these words function as single, indivisible units. Therefore, it seems much more reliable to explain this phenomenon as involving polygraphs that refer to their assigned morphemes collectively, a convention not uncommon in other scripts.

2.2. The chaotic: verbs

Although all possible types of morphography, monomorphemic, polymorphemic, and semimorphemic, have been exemplified in the previous sections, the part of speech in which semimorphemic use of *kanji* is the most common is verbs. Japanese regular verbs can generally be divided into two conjugation groups: consonant verbs (also called *u*-verbs) that have stems that (typically) end on a consonant, and vowel verbs (also known as *ru*-verbs) that have stems that end on a vowel¹⁶ (Hasegawa 2018: 65). As Honda (2021: 190) notes, inflected word-forms of Japanese lexemes are usually written with a mixture of *kanji* and *kana*. What he fails to mention is that the sinograms used to spell consonant verbs almost always omit parts of the stem (see 2-8 to 2-10), while the sinograms used in spelling vowel verbs may represent either the entire stem or just part of it – with the latter case being more common (see 2-11 to 2-14).

2-8 *kaer-u*
 to.go.home-NPST
 KAE|ru
 帰る
 ‘[one] goes home’

2-9 *atatamar-u*
 to.get.warm-NPST
 ATATA|maru
 温まる
 ‘[it] gets warm’

in question consists of in the first place. One would also likely struggle to explain how the concepts of *great* 大 and *harmony* 和 relate to the morpheme *Yamato* 大和 ‘Japan’.

¹⁶ Another set of terms for these conjugation classes is *godan* 五段 ‘quinguirade’ and *ichidan* 一段 ‘monograde’ respectively (Frellesvig 2010: 392). However, these terms are not used in this paper, as they are based on the five-stem theory disputed in (3.2.)

2-10 *kawar-u*
 to.change-NPST
 KA|waru
 変わる

‘[one] changes’

2-11 *kae-ru*
 to.change-NPST
 KA|eru
 変える

‘[one] changes [something]’

2-12 *atatame-ru*
 to.warm-NPST
 ATATA|meru
 温める

‘[one] warms [something]’

2-13 *mi-ru*
 to.look-NPST
 MI|ru
 見る

‘[one] looks’

2-14 *tabe-ru*
 to.eat-NPST
 TA|beru
 食べる

‘[one] eats’

Among these, only the verb in 2-13, *miru* 見る ‘to look’, contains a monomorphemic grapheme. In other cases, at least a phoneme, or sometimes an entire syllable of the stem, is written phonographically with *hiragana*. Some of this can be explained by etymology. For example, in the verbs *atatamaru* 温まる ‘to get warm’ and *atatameru* 温める ‘to warm’, the *-mar-/-me-* parts can be considered diachronically as allomorphs of a verbalizing morpheme and as such not as a part of the stem (Kiyose 1995:

13). In this case, the grapheme 温 could be considered monomorphemic. However, as the morpheme *-mar-/me-* is no longer productive in modern Japanese, a semimorphemic interpretation is also possible.

In other cases, the semimorphemic use of *kanji* comes from the moraic nature of *kana* scripts. The stem of a consonant verb typically ends in a consonant, and this consonant forms a single mora with the following vowel. Since dividing a single mora across multiple graphemes is generally disabled within the Japanese writing system, the final consonant of the stem is incorporated into the *kana* character which simultaneously stands for the first vowel of the inflectional ending (see 2-8).

Overall, the Japanese spelling conventions seem to be primarily concerned with keeping the *reading* of the character consistent across different lexemes instead of accurately mapping their morphemic structure, as mentioned above. The underlying rule for spelling vowel verbs in the *jōyōkanjihyō* (Bunkachō 2010a) appears to be that, when possible, the final mora of the stem should also be written in *kana* (2-14). Moreover, for transitive-intransitive verb pairs, the spelling conventions in the *jōyōkanjihyō* aim to retain consistent character readings by moving the *kanji-kana* boundary a mora back from where it would typically be located in consonant verbs (2-10 and 2-11). This preference for consistency is sometimes extended to adjectives as well (cf. *atataakai* 温かい ‘warm’ with the verbs *atatamaru* 温まる ‘to warm’ and *atatameru* 温める ‘to warm up’ discussed above, or cognate sets such as *kemuru* 煙る ‘to smoke’ and *kemutai* 煙たい ‘smoky’). Although some of these lexemes may contain derivational morphemes from a diachronic perspective, these spelling conventions yield many synchronically semimorphemic graphemes.

On the other hand, polymorphemic graphemes are also quite common in spellings of compound verbs and their derivations. Many alternative spellings are widely used and accounted for by normative sources like dictionaries (Daijirin 1995, Shimmeikai 1999), further blurring the grapheme-morpheme relationships in Japanese writing. In the examples below, all spellings included in the two referenced dictionaries are listed:

- 2-15 *tach-i-agar-u*
 to.stand-NFIN-to.rise-NPST
 TA|chi|A|garu, TACHI|A|garu, TACHI|AGA|ru
 立ち上がる、立上がる、立上る
 ‘to stand up’

- 2-16 *mōsh-i-de-ru*
to.say-NFIN-to.step.out-NPST
MŌ|shi|DE|ru, MŌSHI|DE|ru
申し出る、 申出る
‘to offer’
- 2-17 *mōsh-i-de-∅*
to.say-NFIN-to.step.out-NMN
MŌ|shi|DE, MŌSHI|DE
申し出、 申出
‘to offer’
- 2-18 *osh-i-ire-ru*
to.push-NFIN-to.insert-NPST
O|shi|I|reru, OSHI|I|reru
押し入れる、 押入れる
‘to shove in’
- 2-19 *osh-i-ire-∅*
to.push-NFIN-to.insert-NMN
O|shi|I|re, OSHI|I|re, OSHI|IRE
押し入れ、 押入れ、 押入
‘[Japanese style] closet, wardrobe’
- 2-20 *tach-i-ir-i*
to.stand-NFIN-to.enter-NMN
TA|chi|I|ri, TACHI|I|ri, TACHI|IRI
立ち入り、 立入り、 立入¹⁷
‘[the act of] entering, entry’

¹⁷ To make things even more convoluted, this last spelling is context-dependent, being used chiefly in the phrase *tachiiri kinshi* 立入禁止 ‘no entry’.

As can be seen by comparing the morphemic and graphemic transcriptions, some sinograms in this type of spellings can be semimorphemic (e.g. 立 in *tachiiri* 立ち入り ‘[the act of] entering, entry’ representing only part of the verb stem *tach-*), some polymorphemic (e.g. 立 in *tachiiri* 立入 ‘[the act of] entering, entry’ standing for both the stem *tach-* and the bound morpheme *-i*), or unambiguously monomorphemic (e.g. 出 in *mōshideru* 申し出る ‘to offer’ encoding precisely the single morpheme *de-* ‘to step out’). The choice of spelling depends on the context and, even more so, on the personal preferences of the writer, making Japanese orthography – especially when it comes to verbs – unpredictable and chaotic.

The core argument of this section is thus perfectly illustrated by the character 立. In 2-15, it can be interpreted as either semimorphemic or polymorphemic depending on the spelling chosen by the writer. However, it can also be used monomorphemically, as for instance in the word *kiritsu* 起立 ‘raising, standing up’ where it represents the morpheme *ritsu* 立 ‘to stand; to establish’. In this way, one *kanji* character can be polymorphemic, semimorphemic, or monomorphemic depending on lexical and graphemic context. Admittedly, cases of poly- and semimorphography are mostly observed within the native Japanese stratum of vocabulary or, in other words, in *kun’yomi*.

Despite all of the above, it is important to reiterate that describing *kanji* characters as *morphographic* is an accurate description. As mentioned earlier in the paper, even when the grapheme and morpheme boundaries do not coincide, sinograms still consistently represent specific parts of morphemes or specific combinations of morphemes; their use is not arbitrary. That up to 50% of Japanese vocabulary is of Sino-Japanese origin (Kageyama and Saito 2016: 28; Taylor and Taylor 2014: 260) and graphemes in this type of lexemes are almost uniformly monomorphemic is also not to be understated. Nonetheless, when analyzing specific cases of *kanji* use, a more nuanced terminology than just *morphographic* becomes necessary, as was indicated in the current section.

One more key observation that becomes clear upon closer examination of the examples presented above is the tendency for lexical morphemes to be written fully or in part in *kanji* while inflectional morphemes are almost exclusively represented by *kana* characters. This fact, as well as the irregularities described in this section, bear some unique consequences for descriptions of Japanese; they are presented in the sections (3.) and (4.).

3. Japanese morphology: inconsistent accounts on Japanese morphology and syntax

Even though most, if not all, scholars agree that Japanese is agglutinative, there remains ongoing debate regarding the true nature of its agglutination. For instance, many inflectional morphemes are treated as derivational by traditional grammarians: case endings such as *ga* (subject marker) or *o* (direct object marker) are described as *joshi* 助詞 ‘particles’, standalone parts of speech that are combined with nominal elements through compounding (Jabłoński 2021a: 10, 60). This approach is disputed by other scholars (Jabłoński 2022: 68–69). Similarly, many verb and adjective forms are considered compounds of multiple lexical units called *auxiliary verbs* (Hasegawa 2015: 78; Shibatani 1990: 224–225) even though they are bound inflectional morphemes, unable to convey any linguistic meaning without being attached to a lexical morpheme (Huszcza et al. 2003: 160–165). These issues are explored in this section.

3.1. Nominal elements and their inflection

Japanese sources generally do recognize the distinction between inflectional and derivational morphemes (Tanaka et al. 1988: 304; Tsujimura 2014: 143), though some of them accept the addition of affixes such as *ga* or *no* to noun stems as examples of agglutination (Tanaka et al. 1988: 19). Definitions of these phenomena are also in line with those found in general linguistic sources. For instance, Tanaka et al. (*ibid.*, 18–20, 303) define inflection as the process of communicating grammatical meaning by altering a word’s form through affixation, among other methods, providing examples of this phenomenon in non-Japanese nouns that do not require changes in the stem (e.g. German *Mensch* ‘man’ and *Menschen* ‘men’ or English *tree* and *trees*). Tsujimura (2014: 144) writes: “Inflectional morphemes do not generate different words, unlike the way in which derivational morphemes do; instead, inflected forms are variants of the same word, and for this reason inflectional morphemes do not change the category of the word”. By this definition grammatical phenomena within Japanese agglutination, such as the act of adding the aforementioned grammatical markers *ga* or *no* to nominal elements¹⁸, should be widely accepted as examples of inflection. Peculiarly, this is not the case. A significant portion of Japanese grammarians assert that Japanese nouns *do not undergo inflection*

¹⁸ Nominal elements are understood here as not only nouns proper but also other elements that function like nouns in most or some contexts, such as nominalized forms of verbs, adjectives, numerals, numeral classifiers, or even entire clauses.

(Kindaichi et al. 1988: 118; Tanaka et al. 1988: 77). This is perhaps best captured by the definition of agglutination provided by Iori (2017: 184): “an agglutinating language [...] has the characteristic that morphemes expressing grammatical categories (sentence constituents) are attached following the stem of a predicate to form a string”. Under this approach, a morphological analysis of agglutination is limited to elements that can constitute a predicate, and it is done so on syntactic grounds, conflating the two levels of analysis.

The above is the traditional Japanese approach to nominal morphology and many attempts to solve this logical inconsistency have been made within its framework. The most common solution is as follows: morphemes such as *ga* or *no* are not case *markers*, they are case *particles* and as such “word-external elements” (Honda 2009: 219). It is not the noun that is inflected for case, but rather particles “intervene” between nouns and verbs in order to indicate their syntactic relations. These *particles*, then, are considered to be an entirely independent part of speech. One argument supporting this view is that “other elements” can be placed between the stem of the noun and the *particle* or – another term at times used to refer to these markers – the *phrasal clitic* (Hasegawa 2015: 91; Nakamura 2018: 249; Tsujimura 2014: 126), as can be seen in 3-1 (after Nakamura 2018: 249, original glosses retained).

3-1	<i>Tarō-to</i>	<i>Hanako-dake-ga</i>	<i>eki-kara</i>	<i>arui-ta.</i>
	Tarō-and	Hanako-only-NOM	station-ABL	walk-PAST
	TA RŌ to	HANA KO dakega	EKI kara	ARU ita.
	太郎と	花子だけが	駅から	歩いた。

‘Only Tarō and Hanako walked from the station.’

Firstly, treating the so-called case particles in this way challenges the definition of an *inflectional morpheme* itself, which is described as an affix (bound morpheme) added to a lexical morpheme in order to convey grammatical meaning without changing the grammatical class of the morpheme to which it is attached (Crystal 2008: 243–244; Fromkin et al. 2017: 46–48). For comparison, under such an approach the Polish morpheme *-by* indicating the conditional mood (Bielec 2012: 78; Brooks 1975: 170; Sadowska 2012: 404) would be considered a *particle* as it can be freely separated from the verb (compare *zrobilby to od razu* ‘he would do it immediately’ and *od razu by to zrobil* ‘he would do it immediately’), sometimes alongside other inflectional morphemes such as the first person singular marker *-(e)m* (*zrobilbym to od razu* ‘I would do it immediately’ vs.

od razu bym to zrobił ‘I would do it immediately’). There are even instances where this separation is mandatory (compare *on poprosił, abym wyszedł* ‘he asked me to leave’ and the incorrect **on poprosił, aby wyszedłem*). However, it does not deter Polish grammarians from treating morphemes such as *-by* (conditional mood) or *-(e)m* (first person singular) as inflectional, agglutinative markers (Bańko 2007: 49–52).

Secondly, it is important to state that these *particles* are still bound morphemes that cannot act on their own; they have no meaning without the nominal elements they are attached to. Moreover, there is a finite set of them and they *can* be organized into a declensional paradigm, one that is admittedly extensive (Jabłoński 2022: 68–69). They never change the grammatical class of the morphemes to which they are attached, are phonologically dependent on the stems they are attached to, and cannot be cut off from the stem by anything but another inflectional marker. This is unlike, for instance, English prepositions, which can be separated from nouns by entire phrases. All of this points to the conclusion that they should not be analyzed as lexical items.

Nakamura’s argument can be further challenged by analyzing *dake* and similar items as case markers (Jabłoński 2022: 68–69), though this discussion lies beyond the scope of the present paper. A clearer understanding of the Japanese declensional paradigm and how it compares to those of other languages requires resolving the definitional inconsistencies outlined earlier. The failure to acknowledge the inflectional nature of these so-called “particles” undermines any attempt to describe them systematically within a paradigm. The exact functions of these markers are still up for debate, but for such a debate to start they need to be recognized as primarily grammatical.

3.2. Verbal elements and their inflection

Similarly to nominal elements, verbal elements¹⁹ are also described in two fundamentally different ways. One approach is the classical Japanese perspective (hereinafter: the traditional (Japanese) approach) taught in Japanese schools, where most inflectional markers, if not all, are treated as *jodōshi* 助動詞 ‘auxiliary verbs’. These are referred to as independent items that attach to the inflected stems of verbal elements and undergo conjugation themselves (Shibatani 1990: 221). The other approach, proposed mainly by

¹⁹ Verbal elements are understood here as both verbs proper and predicative adjectives since they also can be conjugated and function as independent predicates, even if their conjugational pattern is more limited than that of verbs.

non-Japanese scholars (ibid., 227), is the synthetic approach, which views these so-called *jodōshi* as purely inflectional, bound morphemes that attach to lexical morphemes (Huszcza et al. 2003: 160–165).

The traditional approach is as follows: the verb is conjugated into five different forms (except the copula, which has six) with different endings; the root (the unchanging part of the verb understood in moraic terms) and the ending *together* form the stem. To this stem, various auxiliary verbs are appended²⁰ (Kindaichi et al. 1988: 122; Shibatani 1990: 224–225) (see Table 1).

Word-form name	‘to die’	‘to look’	Auxiliary verbs
<i>mizen</i> (irrealis)	<i>shi-na</i>	<i>mi</i>	<i>nai, n(u), seru, saseru, reru, rareru</i>
<i>ren’yō</i> (adverbial)	<i>shi-ni</i>	<i>mi</i>	<i>ta, masu, tai, sō, tari, te, tsutsu, nagara</i>
<i>shūshi</i> (conclusive)	<i>shi-nu</i>	<i>mi-ru</i>	<i>rashii, sō, to, kara, ga</i>
<i>katei</i> (hypothetical)	<i>shi-ne</i>	<i>mi-re</i>	<i>ba</i>
<i>meirei</i> (imperative)	<i>shi-ne</i>	<i>mi-ro/mi-yo</i>	

Table 1. Conjugated forms of Japanese verb “stems” with example auxiliary “verbs” that can be attached to them (Shibatani 1990: 224).

This method of describing verb inflection raises questions about the definition of the *stem* as the (mostly) unchanging part of the verb. The term *auxiliary verb*, on the other hand, implies that these elements appended to verbs to indicate tense, voice, and other grammatical categories possess lexical status. An alternative would be to classify markers such as *-(r)are* (passive voice marker), *-te* (gerund marker, among other uses), or *-tari* (exemplificative marker) as inflectional morphemes, and such descriptions of Japanese exist. The argument against this synthetic approach is as follows: elements such as *-(r)are* should be regarded as *verbs*, or at least *derivational suffixes*, because they themselves inflect, for instance, by taking tense markers. The status of elements such as *-te* or *-tari* as inflectional is questioned on the basis that “they all connect two clauses, the function not normally associated with the inflectional endings” (Hasegawa 2015: 152; Kiyose 1995: 6; Shibatani 1990: 227–228). This argumentation, however, is undermined by terminological inconsistency. For example, Honda (2009:

²⁰ These “verbs” are sometimes understood more in terms of script than morphology, being described as “*kana* characters that, in the case of verbs, are appended after the *kanji* to indicate tenses and other grammatical categories” (Hasegawa 2015: 52).

219) and Hasegawa (2015: 152) regard words such as *mirareru* 見られる ‘to be seen’ as *derivations* from the original verb (in this case *miru* 見る ‘to see’) – a *derivation* being by definition a new lexeme (Crystal 2008: 138) – while at the same time dubbing it “the passive *form*” of the verb. In this way, the fundamental differentiation between derivation and inflection is lost.

This line of reasoning, if accepted, could have tremendous implications for grammars of the world’s languages. For comparison, the Polish past tense marker *-l/-l-* (as in: *zrobił* ‘he did’), a bound morpheme unable to be used on its own, would also need to be described as an *auxiliary verb* rather than as an *inflectional ending* since it can “inflect further” by taking markers of mood (*zrobiłby* ‘he would do’), person (*zrobiłem* ‘I did’), or both (*zrobiłbym*, ‘I would do’). In fact, this would be a perfect parallel with the traditional Japanese approach, in which the past tense marker *-ta* is considered a lexical item, i.e. an auxiliary verb that can itself inflect further (Shibatani 1990: 228). However, this not the case since these patterns in Polish are usually recognized as pure inflection or, more specifically, as examples of agglutination, a subcategory of inflection (although Polish on the whole is usually classified as fusional; Bańko 2007: 48–52). This approach is in line with the linguistic definitions of inflection and inflectional morphemes, while the traditional Japanese approach seems to prioritize erroneous assumptions about morphosyntax, such as the claim that inflectional morphemes cannot fulfill syntactic functions.

As mentioned above, Japanese word-forms can also be analyzed as synthetic, as in 3-2. The ability to attach additional inflectional markers – such as the past tense morpheme *-ta* – onto the *-(r)are-* passive morpheme does not necessarily mean that it is inflected; it may simply be regarded as an instance of agglutination of the main, lexical verb. After all, semantically, it is not the “passiveness” expressed by the morpheme *-(r)are-* that happens in the past, rather the past tense remains a feature of the action – for example of eating (see 3-2 and 3-3) – expressed by the lexical stem. It is also not uncommon for agglutinative word-forms to contain multiple inflectional morphemes (Comrie 1989: 43–44). One should not let the non-final status²¹ of the morpheme obscure the fact that *it cannot be used independently* (see 3-3) and as such does not fulfill the definition of a lexical item.

²¹ That is to say: the fact that it has to take another inflectional marker for the word-form to be complete, such as the passive *-(r)are-* marker that has to connect to at least a tense or finiteness marker such as *-(r)u*, *-ta* or *-te*.

3-2 *tabe-rare-ta*
 to.eat-PASS-PAST
 TA|berareta
 食べられた
 ‘[it] was eaten’

3-3 **rare-ta*
 PASS-PAST
 rareta
 られた
 ‘???’

Elements such as *-te* or *-tari*, on the other hand, are denied inflectional status on functional grounds as the function of connecting clauses is deemed atypical for inflectional endings. Again, if such a view on language were accepted, the grammatical status of elements such as the Polish and English participle markers would also need to be reconsidered. Especially in Polish, the participial morpheme *-ąc* serves solely the purpose of introducing another clause (*Uczył się, słuchając muzyki*. ‘He studied [while] listening to music.’)²². Function cannot therefore be the primary ground for morphological classification. Additionally, the *-te* or *-tari* markers are bound morphemes that lack meaning outside of their combinations with verbal elements, just like *-(r)are*.

The problems described above disappear when one assumes that *any bound morpheme attached to the stem of a nominal or verbal element, whether directly or following after other bound morphemes, is inherently inflectional in nature* as long as it does not change the grammatical class of the lexeme and the process of attaching it to the stem is regular and paradigmatical. While a distinction between semantic (e.g. tense) and syntactic (e.g. finiteness²³) functions of inflectional morphemes should be acknowledged (Polański 1999: 186, 377), such differentiation presupposes an initial recognition of their fundamentally inflectional status. Moreover, viewing

²² It is worth noting that Japanese verb forms ending in *-te/-de* can also be described as participles (Jabłoński 2021c: 196–199). The corresponding English *-ing* form is somewhat more ambiguous and functionally diverse, but it is likewise non-finite, unmarked for tense or person (Crystal 2008: 189–190). Nonetheless, it is usually listed as an inflected form of the verb (Castairs-McCarthy 2002: 39–40; Don 2014: 23, 80).

²³ Some sources consider finiteness to be a proper category of the verb, despite its often syntactic function (Cristofaro 2007: 92).

(most) Japanese verbal suffixes as purely inflectional has no negative implications for the understanding of Japanese morphology and syntax. Adopting such an understanding has led primarily non-Japanese (Shibatani 1990: 227–228) grammarians to devise comprehensive tables of potential interconnectivity between Japanese inflectional morphemes (Huszcza et al. 2003: 248; Jabłoński 2021b: 396–403), thereby facilitating a better understanding of the underlying morphological mechanisms. Morphemes such as *-te*, *-tari*, or *-rare-* are treated in these tables as parts of the conjugational paradigm, the first two as non-finite markers, and the last one as the passive marker (Huszcza et al. 2003: 204, 237, 242; Jabłoński 2021b: 396–403).

Gaps in descriptions of Japanese morphology become apparent when consulting sources based in the traditional Japanese approach to morphology described above. Some grammars, such as Cipris and Hamano (2002), entirely omit systematic presentations of inflectional markers, offering no tables of word-forms or paradigms. Even comprehensive works like Kindaichi et al. (1988) fail to include declension paradigms and often exclude morphemes traditionally labeled as “auxiliary verbs” from their conjugation charts. They also lack accounts of how verbal markers interconnect.

In some cases, internal inconsistencies arise. For example, Yamada’s grammar treats verbal markers as part of the conjugational paradigm but regards nominal markers as independent lexemes (Kindaichi et al. 1988: 137). Hasegawa (2015: 75–78), on the other hand, treats certain word-forms of verbs as inflections while describing others – also formed in a paradigmatical way – as derivational. While some newer works present adnominal markers in lists resembling declensional paradigms, they still treat these as standalone *postpositions*, likened to English prepositions such as *to* or *along*, invariably favoring semantics over morphology and syntax (Katsuki-Pestemer 2003: 13–14). Overall, despite frequent references to agglutination, few scholars analyze Japanese morphology systematically from an agglutinative perspective. This is a serious problem, as it obscures the mechanisms by which grammatical markers combine as *they cannot do so freely as they are not lexical units*.

4. Parallels between Japanese graphemics and descriptions of Japanese morphology

Sections 1 and 2 described the ambiguous relationships between the graphemes and morphemes of the Japanese language that arise within the subset of *kanji* characters due to a variety of factors. Firstly, the moraic

nature of Japanese phonography as well as the tendency to prefer general legibility over a clear grapheme-morpheme correspondence often make it impossible for morpheme-morpheme and grapheme-grapheme boundaries to coincide. Secondly, lexical morphemes exhibit a tendency to be written with *kanji* while inflectional morphemes are typically represented with *kana* characters with some notable exceptions. Section 3 briefly highlighted some of the most prominent inconsistencies in grammars of the Japanese language, such as the classification of inflectional morphemes as independent parts of speech. This section aims to juxtapose these features and point out a few similarities between the two.

4.1. *Okurigana* and descriptions of Japanese agglutination

The intersection of Japanese graphemics and morphology takes the form of *okurigana* and the rules governing its usage. As mentioned in (1.1.), its primary role in *kanji kana majiribun* is representing inflectional endings, and clarifying the pronunciation of the *kanji* character it is attached to (Kindaichi et al. 1988: 315). Hasegawa (2015: 52) specifically defines *okurigana* as “*kana* characters that, in the case of verbs, are appended after the *kanji* to indicate tenses and other grammatical categories”. Two things become evident in this statement, the most salient one being the mix-up of morphology and graphemics. A more accurate description of the function of *okurigana* would be to specify that it is *inflectional morphemes* that indicate grammatical categories, and *okurigana* characters merely represent them (as well as, quite commonly, parts of the stem). Indirectly, however, this definitional conflation of graphemics and morphology betrays a tendency to view language through the lens of writing that becomes even more obvious upon a closer inspection of the rules governing the usage of *okurigana*.

The rules of adding *okurigana* to *kanji*, as defined by the Japanese Agency for Cultural Affairs (*Bunkachō* 文化庁), are vague and full of exceptions (Bunkachō 2010b). The general guidelines are as follows (Kindaichi et al. 1988: 315; emphasis added by the author of this paper).

1. *For words that conjugate, conjugational endings are appended.*
2. For derivations and other related words, their mutual relationships are taken into account and the part before the *conjugational ending* is also appended. This also applies to words that could otherwise be misread.
3. *No okurigana is added to nouns.*
4. For adverbs, adnominal elements, and conjunctions, the last syllable of the word is written as *okurigana*.

5. *Okurigana* should be used to minimize the possibility of words being misread or difficult to read.
6. Traditional uses of *okurigana* should be respected.

A few things are of note here. Firstly, conjugation is the only type of inflection mentioned here, with the implication being that it is the only type of Japanese inflection overall. Secondly, the definition of *okurigana* seems to imply that if a lexeme is written with a mix of *kanji* and *kana*, only the *kanji* and the *okurigana* can be considered integral parts of it. Thus, if by definition nouns cannot possess *okurigana* (and they cannot, as declension is clearly absent from the ruleset), whatever is added after them is not part of the same lexeme. If it is not part of the same lexeme then it cannot be an inflectional marker and must thus be considered a separate part of speech – namely, the *particle* – and as such defined in lexical terms. This line of reasoning seems to underlie the peculiar similarity between the definition of *okurigana* and the classification of so-called “particles”.

Of course, one might argue that the causality here is reversed and it is the preexisting concept of “particles” as an independent part of speech that led to this definition of *okurigana*. However, the tradition of representing the stem of a lexeme and the inflectional suffixes attached to it in a morphographic and a phonographic manner respectively dates back to at least the 8th century CE, in the *senmyō* 宣命 ‘imperial edicts’ and *norito* 祝詞 ‘liturgies’. These texts exhibited a feature atypical for the time they were composed in; although mostly written morphographically, some inflectional markers attached to lexical stems were represented phonographically, with *kanji* used for their phonetic value. Additionally, the *kanji* representing these markers were written in a smaller size to further differentiate them from the stems. A similar orthography can be observed in another 8th century text, the *Man’yōshū*, although without the script size differentiation (Seeley 1991: 54; Vovin 2020: 8–16, 28). From the middle of the Heian period (794–1185), *kanji kana majiribun*, which is the style of writing used to write Japanese today, started gaining popularity. It is in this style that inflectional suffixes are written in *kana* while lexical stems are represented by *kanji* characters (Frellesvig 2010: 157, 266). The academic study of Japanese, on the other hand, began much later, with one of the first extensive grammars of Japanese being published as late as in the 17th century (*ibid.*, 302), while the modern discipline of linguistics developed even later – the first chair of modern linguistics in Japan was created in 1898 (Coulmas 2016b: 170). It is thus logical to assume that the preexisting writing conventions might have influenced the newly forming discipline.

This is not an uncommon occurrence. Coulmas (2016a: 39–40) notes that “laying down rules for language is entirely reliant on writing” and that the existence of written texts, especially those held in high regard like religious and governmental ones, necessarily influences how languages are described and used (including speech). He posits that the primary tool of linguistic analysis is writing to this day, even in fields such as phonetics and phonology where precise transcription systems (such as IPA) are an essential tool. The most basic linguistic concepts such as the phoneme, sentence, or word also hinge on writing to some extent (Coulmas 2018b: 26–28). Again, it is therefore not unreasonable to assume that writing conventions have influenced descriptions of both Indo-European languages and Japanese, rather than the other way around. The repercussions of this influence are still observable today.

4.2. Verbal elements

As discussed in 3.2., inflectional forms of Japanese verbal elements are traditionally described in terms of auxiliary verbs attached to inflectable “stems”. This analysis is a direct consequence of the Japanese writing system, or more precisely, of the *kana* scripts. As *kanji* and *kana* characters cannot stand for units smaller than moras, a problem arises when the boundary between lexical and inflectional morphemes cuts through a mora (Shibatani 1990: 225). Thus, in the case of verbal elements such as *kaer-u* 帰る ‘to come home’ (see also 2-8), it is impossible to separate the last consonant of a stem and the first vowel of an inflectional ending in writing. This is one of the reasons why the five stem theory arose and accounts for the semimorphemic character of many *kanji* characters in certain contexts. On the other hand, looking back (4.1.) to the rules of adding *okurigana* to *kanji*, one might notice that, unlike declension, the existence of conjugation is in fact recognized. However, the rules also state that whatever is written after the *kanji* of a conjugatable word is an inflectional ending. This understanding aligns with the original views on Japanese conjugation that treated the last consonant of the stem of consonant verbs as a part of the traditionally understood inflectional ending. In this way, the five-stem theory (see 3.2.) distorts the definitions of both the stem and the ending (see Introduction).

The problem of why inflectional markers attached to verbs are considered to be “auxiliary verbs” (see 3.2.), however, is slightly more complex. One possible explanation for this could lie in the fact that inflected forms of Japanese verbal elements are notoriously long, usually consisting of a single lexical morpheme and multiple inflectional ones. The rules of *okurigana*,

however, do not specify what exactly constitutes a *word that conjugates* and *where exactly okurigana ends*. Moreover, the rules imply that it is the *kanji* that is the core of the lexeme, with no reference to actual word-form morphology. It is then quite possible to treat anything other than the *kana* characters closest to this “core” as separate entities, especially since agglutinative word-forms can usually be divided into constituent segments relatively easily, with just one function assigned to each. This is further exacerbated by the fact that many of the Japanese inflectional morphemes such as *-(r)are-* (passive), *-(s)ase-* (causative), or *-e-/rare-* (potential) end in a vowel, making it easy to separate them from any other morpheme following them both phonetically and graphemically. Such a segment can then be ascribed a lexical meaning.

Admittedly, there also exist non-graphemic reasons for why one might interpret some inflectional suffixes as lexical units. For instance, some word-forms of Japanese verbs include morphemes characteristic for adjectival conjugation. This seemingly violates the rule that inflection should not change the grammatical class of the lexeme²⁴ and might lead to the conclusion that such markers should not be considered part of a synthetic word-form. Another factor might be the influence of foreign grammars (cf. Kindaichi et al. 1988: 117–118, among others); a detailed discussion of these questions, however, falls beyond the scope of the paper.

Considering all of the above it can be concluded that treating verbal markers as independent lexical items and separate parts of speech goes against the definition of inflection. There is nothing lexical about morphemes such as *-(r)are-* or *-(s)ase-*; one cannot use them in a sentence without an accompanying lexical morpheme. They carry no meaning other than grammatical ones, and they ultimately do not change the grammatical class of the morphemes to which they are attached in a regular manner as part of a paradigm, disqualifying them as derivational. There is, then, little reason to treat them as anything else but inflectional.

Conclusions

This paper focused on inflected parts of speech, namely nominal and verbal elements, with particular emphasis on nouns, predicative adjectives, and verbs. Due to space and scope constraints, non-inflected parts of speech such as lexical adverbs and non-predicative adjectives were excluded. For the

²⁴ At the same time, it is important to note that the distinction between verbs and adjectives may be ambiguous (Wierzbicka 1995; Pustet 2000) and the classes may overlap in some cases, as with Japanese adjectives inflecting for tense and other typically verbal categories.

same reasons, many nuances regarding grapheme-morpheme relationships *within* the stems of nominal and verbal elements have been omitted (for a discussion of this topic, see: Honda 2019). A future study could examine grapheme-morpheme relationships across all Japanese parts of speech in greater depth, as various kinds of morphography can be observed in adverbs as well. A statistical analysis of grapheme-morpheme relationships within a representative sample of the Japanese lexicon could also offer valuable insights into the nature of Japanese morphography.

Similarly, this study did not seek to redefine Japanese morphology, settling for juxtaposing the existing frameworks. Whether one views Japanese agglutination in terms of inflectional paradigms or independent lexical elements expressing grammatical meaning through compounding, it *is* possible to describe Japanese morphology and syntax comprehensively and wholly. However, the latter approach hinders grammarians from doing so in a systematic way. If inflectional endings are listed as separate entries in dictionaries – as they often are (Daijirin 1995, Shimmeikai 1999) – and explained primarily in terms of their semantics, then it effectively undermines the description of consistent, unified inflection paradigms. Such paradigms would not only aid in understanding Japanese as a language, especially its morphology and syntax, but also help foreign learners by presenting Japanese inflection in a more regular, systematic way.

The analysis of various types of graphemic representation of verbal elements specifically has revealed the three basic ways in which graphemes can be related to morphemes within the Japanese writing system. In monomorphemic representation, the boundaries between morphemes and graphemes overlap. In semimorphemic representation, only part of the morpheme in question can be unambiguously related to the character with which it is written. Finally, polymorphemic representation occurs when more than one morpheme is written with a single character. I demonstrated how these inconsistent relationships between graphemes and morphemes may influence morphological analysis of Japanese lexemes, leading to treating certain parts of the stem as inflectional endings. On the other hand, the “division of labor” between *kanji* and *kana*, in which inflectional morphemes are almost exclusively written in the latter script, influences what elements are considered parts of the same word-form. Since inflectional endings are visually distinct from the stems to which they are attached it is easier to come to the conclusion that they are independent lexical items – a conclusion which, upon closer examination of Japanese morphology using widely accepted definitions of inflection and agglutination, appears erroneous.

It is evident that Japanese descriptive grammar requires paradigmatic approaches that apply rigorous linguistic criteria in order to formulate internally and externally coherent explanations of grammatical phenomena. While formulating language-specific theories can be helpful, they should always be based on consistent linguistic frameworks and terminology. On the other hand, the field of writing system studies necessitates a more extensive terminology to accurately describe the morphography of the Japanese writing system. Scholars need to be aware of the ambiguities in the relationships between graphemes and linguistic units in Japanese writing, as well as how the mixed nature of the system and the dissonance between character and morpheme boundaries can influence grammars of the Japanese language. Moreover, it is crucial to remember that speech should always take precedence over writing when examining grammatical structures; writing can be a valid object of study in its own right but should never serve as the basis for explaining speech. Only when the biases that stem from writing conventions are recognized can they be actively countered.

Abbreviations

ADN	Adnominal form
NFIN	Non-finite form
NMN	Nominal form
NPST	Non-past tense
PASS	Passive voice
PAST	Past tense

Symbols and typographic conventions

*	Non-existent phrases or spellings
-	Morpheme boundary (morphemic transcription)
	Script or sinogram boundary (graphemic transcription)
SMALL CAPS	Part of a phrase written in <i>kanji</i> (graphemic transcription)
regular font	Part of a phrase written in <i>kana</i> (graphemic transcription)

Author Contributions

The author confirms the sole responsibility for the conception of the study, presented results, and manuscript preparation.

Conflicts of Interest

The author declares that they have no conflict of interest.

References

Bańko, Mirosław 2007. *Wykłady z polskiej fleksji* [lectures on Polish inflection]. Warszawa: Wydawnictwo Naukowe PWN.

Bielec, Dana 2012. *Polish. An essential grammar*. Abingdon: Routledge.

Bloomfield, Leonard 1973. *Language*. London: George Allen & Unwin LTD.

Brooks, Maria Zagorska 1975. *Polish reference grammar*. The Hague: Mouton & Co. N.V. Publishers.

Bunkachō [Agency for Cultural Affairs] 2010a. *Jōyō kanji hyō* [list of kanji for general use]. URL:

https://www.bunka.go.jp/kokugo_nihongo/sisaku/joho/joho/kijun/naikaku/pdf/joyokanjihyo_20101130.pdf [access date: 16 April 2026]

文化庁. 『常用漢字表』.

Bunkachō [Agency for Cultural Affairs] 2010b. *Okurigana-no tsukekata* [how to append *okurigana*]. URL:

https://www.bunka.go.jp/kokugo_nihongo/sisaku/joho/joho/kijun/naikaku/okurikana/index.html [access date: 16 April 2026]

文化庁. 『送り仮名の付け方』.

Castairs-McCarthy, Andrew 2002. *An introduction to English morphology. Words and their structure*. Edinburgh: Edinburgh University Press.

Cipris, Zeljko, Shoko Hamano 2002. *Making sense of Japanese grammar. A clear guide through common problems*. Honolulu: University of Hawai'i Press.

Comrie, Bernard 1989. *Language universals and typology*. Second edition. Chicago: The University of Chicago Press.

Coulmas, Florian 1999. *The Blackwell encyclopedia of writing systems*. Malden, Oxford, Carlton: Blackwell Publishing.

Coulmas, Florian 2003. *Writing systems. An introduction to their linguistic analysis*. Cambridge: Cambridge University Press.

Coulmas, Florian 2016a. “Prescriptivism and writing systems”. In: Ingrid Tieken-Boon van Ostade, Carol Percy (eds.). *Prescription and tradition in language. Establishing standards across time and space*. Bristol/Blue Ridge Summit: Multilingual Matters. 39–56.

Coulmas, Florian 2016b. *Guardians of language. Twenty voices through history*. Oxford: Oxford University Press.

Coulmas, Florian 2018a. “Writing and literacy in modern Japan”. In: Yoko Hasegawa (ed.). *The Cambridge handbook of Japanese linguistics*. Cambridge: Cambridge University Press. 114–132.

Coulmas, Florian 2018b. “Revisiting the *Tyranny of Writing*”. In: Constanze Weth, Kasper Juffermans (eds.). *Tyranny of writing. Ideologies of the written word*. London/Oxford/New York/New Delhi/Sydney: Bloomsbury Academic. 19–30.

Cristofaro, Sonia 2007. “Deconstructing categories: finiteness in a functional-typological perspective”. In: Irina Nikolaeva (ed.). *Finiteness. Theoretical and empirical foundations*. Oxford/New York: Oxford University Press. 91–114.

Crystal, David 2008. *A dictionary of linguistics and phonetics*. Sixth edition. Malden: Blackwell Publishing.

Daniels, Peter T. 2018. *An exploration of writing*. Sheffield/Bristol: Equinox Publishing Ltd.

DeFrancis, John 1989. *Visible speech. The diverse oneness of writing systems*. Honolulu: University of Hawai’i Press.

Don, Jan 2014. *Morphological theory and the morphology of English*. Edinburgh: Edinburgh University Press.

Elbourne, Paul 2011. *Meaning. A slim guide to semantics*. Oxford: Oxford University Press.

Fábregas, Antonio, Sergio Scalise 2025. *Morphology. From data to theories*. Second edition. Edinburgh: Edinburgh University Press.

Frellesvig, Bjarke 2010. *A history of the Japanese language*. Cambridge: Cambridge University Press.

Fromkin, Victoria, Robert Rodman, Nina Hyams 2017. *An introduction to language*. Eleventh edition. Boston: Cengage.

Grzegorzczkova, Renata 2008. *Wstęp do językoznawstwa* [an introduction to linguistics]. Warszawa: Wydawnictwo Naukowe PWN.

Handel, Zev 2019. *Sinography. The borrowing and adaptation of the Chinese script*. Leiden: Brill.

Hasegawa, Yoko 2015. *Japanese. A linguistic introduction*. Cambridge: Cambridge University Press.

Hasegawa, Yoko 2018. "Introduction". In: Yoko Hasegawa (ed.). *The Cambridge handbook of Japanese linguistics*. Cambridge: Cambridge University Press. 3–14.

Honda, Keisuke 2009. "Homographic kanji, their ambiguity and the effectiveness of okurigana as a device for disambiguation". *Written Language & Literacy* 12: 213–236.

Honda, Keisuke 2019. "What do kanji graphs represent in the current Japanese writing system? Towards a unified model of kanji as written signs". In: Yannis Haralambous (ed.). *Graphemics in the 21st century. Brest, June 13-15, 2018. Proceedings*. Brest: Fluxus Editions. 185–208.

Honda, Keisuke 2021. "A modular theoretic approach to the Japanese writing system: possibilities and challenges". In: Yannis Haralambous (ed.). *Grapholinguistics in the 21st Century 2020. Proceedings*. Brest: Fluxus Editions. 621–643.

Horie, Kaoru 2018. "Linguistic typology and the Japanese language". In: Yoko Hasegawa (ed.). *The Cambridge handbook of Japanese linguistics*. Cambridge: Cambridge University Press. 65–86.

Huszcza, Romuald, Maho Ikushima, Jan Majewski 2003. *Gramatyka japońska. Podręcznik z ćwiczeniami. Tom 1* [Japanese grammar. Textbook

and workbook. Volume 1]. Cracow: Wydawnictwo Uniwersytetu Jagiellońskiego.

Iori, Isao 2017. “The layered structure of the sentence”. In: Masayoshi Shibatani, Shigeru Miyagawa, and Hisashi Noda (eds.). *Handbook of Japanese syntax*. Boston/Berlin: De Gruyter Mouton. 157–186.

Jabłoński, Arkadiusz 2021a. *Japanese nominal elements as abandoned parts of speech*. Dąbrowka: Wydawnictwo Rys.

Jabłoński, Arkadiusz 2021b. *Polski leksykon japońskich terminów gramatycznych. Tom 1: A–K* [the Polish lexicon of Japanese grammar terms. Volume 1. A–K]. Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.

Jabłoński, Arkadiusz 2021c. *Polski leksykon japońskich terminów gramatycznych. Tom 2: L–R* [the Polish lexicon of Japanese grammar terms. Volume 2. L–R]. Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.

Jabłoński, Arkadiusz 2022. *Case in Japanese – a morphological approach*. Dąbrowka: Wydawnictwo Rys.

Joyce, Terry 2011. “The significance of the morphographic principle for the classification of writing-systems”. *Written Language & Literacy* 14: 58–81.

Joyce, Terry, Hisashi Masuda 2019. “On the notions of graphematic representation and orthography from the perspective of the Japanese writing system”. *Written Language & Literacy* 22: 247–279.

Kageyama, Taro, Michiaki Saito 2016. “Vocabulary strata and word formation processes”. In: Taro Kageyama, Hideki Kishimoto (eds.). *Handbook of Japanese lexicon and word formation*. Boston/Berlin: Walter de Gruyter. 11–50.

Katsuki-Pestemer, Noriko 2003. *Japanese postpositions: theory and practice*. Munich: LINCOM Europa.

Kindaichi Haruhiko, Hayashi Ōki, Shibata Takeshi (eds.) 1988. *Nihongo hyakka daijiten. An Encyclopaedia of The Japanese Language*. Tokyo: Taishūkan Shoten.

金田一春彦、林大、柴田武編集. 『日本語百科事典』. 東京都: 大修館書店.

Kiyose, Gisaburo N. 1995. *Japanese grammar. A new approach*. Kyoto: Kyoto University Press.

Kobayashi, Hideki, Kiyo Yamashita, Taro Kageyama 2016. “Sino-Japanese words”. In: Taro Kageyama, Hideki Kishimoto (eds.). *Handbook of Japanese lexicon and word formation*. Boston/Berlin: Walter de Gruyter. 93–131.

Matthews, Peter Hugoe 1991. *Morphology. Second edition*. Cambridge: Cambridge University Press.

Nakamura, Wataru 2018. “Case”. In: Yoko Hasegawa (ed.). *The Cambridge handbook of Japanese linguistics*. Cambridge: Cambridge University Press. 249–275.

Osterkamp, Sven, Gordian Schreiber 2021. “Challenging the dichotomy between phonography and morphography: Transitions and gray areas”. In: Yannis Haralambous (ed.). *Grapholinguistics in the 21st Century 2020. Proceedings*. Brest: Fluxus Editions. 47–82.

Polański, Kazimierz (ed.) 1999. *Encyklopedia językoznawstwa ogólnego* [an encyclopedia of general linguistics]. Wrocław: Ossolineum.

Pustet, Regina 2000. “How arbitrary is lexical categorization? Verbs vs. adjectives”. *Linguistic Typology* 4: 175–212.

Rogers, Henry 2005. *Writing systems. A linguistic approach*. Malden: Blackwell Publishing.

Rose, Heath 2017. *The Japanese writing system. Challenges, strategies and self-regulation for learning kanji*. Bristol/Blue Ridge Summit: Multilingual Matters.

Sadowska, Iwona 2012. *Polish. A comprehensive grammar*. Abingdon: Routledge.

Sampson, Geoffrey 1985. *Writing systems. A linguistic introduction*. Stanford: Stanford University Press.

Sato, Eriko 2018. "Sociocultural implications of the Japanese multi-scripts. Translanguaging in translation." In: Hye K. Pae (ed.). *Writing systems, reading processes, and cross-linguistic influences. Reflections from the Chinese, Japanese and Korean languages*. Amsterdam/Philadelphia: John Benjamins Publishing Company. 313–334.

Seeley, Christopher 1991. *A history of writing in Japan*. Leiden: E.J. Brill.

Shibatani, Masayoshi 1990. *The languages of Japan*. Cambridge: Cambridge University Press.

Tanaka Harumi (ed.) 1988. *Gendai gengogaku jiten. Seibido's Dictionary of Linguistics*. Tokyo: Seibidō.

田中春美集. 『現代言語学辞典』. 東京都: 成美堂.

Taylor, Insup, M. Martin Taylor 2014. *Writing and literacy in Chinese, Korean and Japanese*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Tsujimura, Natsuko 2014. *An introduction to Japanese linguistics*. Third edition. Malden/Oxford/Chichester: John Wiley & Sons.

Velupillai, Viveka 2012. *An introduction to linguistic typology*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Vovin, Alexander 2020. *A descriptive and comparative grammar of Western Old Japanese. Revised, updated and enlarged 2nd edition*. Leiden/Boston: Brill.

Wasilewska, Halina 2019. *Adaptacje i metamorfozy – znaki chińskie w wybranych sinograficznych systemach pisma Azji Wschodniej* [adaptations and metamorphoses – Chinese characters in several East Asian writing systems]. Poznań: Wydawnictwo Poznańskiego Towarzystwa Przyjaciół Nauk.

Wierzbicka, Anna 1995. “Adjectives vs. verbs: The iconicity of part-of-speech membership”. In: Marge E. Landsberg (ed.). *Syntactic iconicity and linguistic freezes*. Berlin/New York: Mouton de Gruyter. 223–245.

Wnuk, Zuzanna 2023. *Rozwój metajęzykowej świadomości ortograficznej w kontekście nauczania i uczenia się języka chińskiego jako języka obcego* [the development of metalinguistic orthographic awareness in the context of teaching and learning Chinese as a foreign language]. Doctoral dissertation. Poznań: Adam Mickiewicz University in Poznań, Faculty of Modern Languages and Literatures.

Wnuk, Zuzanna 2024. “Graphemic awareness development of Polish learners of Chinese as a foreign language”. *Theory and Practice of Second Language Acquisition* 10(1): 1–27.

Referenced dictionary entries, grouped by dictionary, in *gojūon* order

Daijirin 1995 = Sanseidō (ed.) 1995. *Sūpā Daijirin* [Super Daijirin]. Second edition. Tokyo: Sanseidō.

三省堂編．『スーパー大辞林』．第2版．東京：三省堂．

Entries from this dictionary are referenced in the paper as “(Daijirin 1995)”.

あか-がね [0] 【銅】

うら-やまし・い [5] 【羨ましい】 (形) [文]シク うらやま・し

うら-や・む [3] 【羨む】 (動マ五 [四])

おおき オホキ 【大きい】

おおき・い オホキイ [3] 【大きい】 (形)

おお・し オホシ 【大し】 (形ク)

おし-いれ [0] 【押(し)入れ】

おし-い・れる [4] 【押(し)入れる】 (動ラ下一) [文]ラ下二 おしい・る

こころ-ざし [0] 【志】

こころ-ざ・す [4] 【志す】 (動サ五 [四])

しり-ぞ・く [3] 【退く】

そ・く 【退く】

たち-あが・る [0][4] 【立(ち)上がる】 (動ラ五 [四])

たち-いり [0] 【立(ち)入り】

たちいり-きんし [0] 【立(ち)入り禁止・立入禁止】

た・べる [2] 【食べる】 (動バ下一) [文]バ下二 た・ぶ
なつかし・い [4] 【懐かしい】 (形) [文]シク なつか・し
みず-うみ ミヅー [3] 【湖】
もうし-で マウシー [0] 【申(し)出】
もうし・でる マウシー [4] 【申(し)出る】 (動ダ下一)

Shimmeikai 1999 = Sanseidō (ed.) 1999. *Shimmeikai kokugo jiten dai-5-han*
[Shimmeikai dictionary of Japanese, 5th edition]. Tokyo: Sanseidō.

三省堂編. 『新明解国語辞典 第5版』. 東京：三省堂.

Entries from this dictionary are referenced in the paper as “(Shimmeikai 1999)”.

あがる 【上がる】 【上る】
おしいれ 【押し入れ】 【押入】 [0]
が [一] (格助)
たちあがる 【立ち上がる】 【立上る】 [4][5][0]:[4][0]
もうしでる 【申し出る】 【申出る】 マウシデル[4]:[4]