Introduction

This book describes the development of the Japanese language from its written beginning until the present day as it is reflected in the written sources; that is to say, its internal history. This is accomplished by first giving an overall description of the oldest attested stage of the language, Old Japanese, and then tracing changes since then, as they are reflected in the written sources and in the present-day language. The possible cognation of Japanese to other languages, its external pre-history, is not considered. Nor is its internal pre-history discussed to any significant extent, except where relevant to understanding its attested history. And nothing is said about dialects, except where they are prominently reflected in the written sources and where they have contributed to the formation of the modern standard language.

The periodization of Japanese shown in (1) is adopted, which overlaps with the main political periods.

(1)   Linguistic periods   Political periods
        Old Japanese (OJ)   700–800   Nara, 712–794

        Early Middle Japanese   800–1200   Heian, 794–1185
           (EMJ)

        Late Middle Japanese   1200–1600   Kamakura, 1185–1333
           (LMJ)   Muromachi, 1333–1573

        Modern Japanese (NJ)   1600–   Edo, 1603–1868
                                              Meiji, 1868–1912
                                              Taishō, 1912–1926
                                              Shōwa, 1926–1989
                                              Heisei, 1989–

Where necessary, early is distinguished from late within both Early Middle Japanese (early: 794–1086; late: the *Insei* period, 1086–1185) and Late Middle Japanese (early: the Kamakura period; late: the Muromachi period). Modern Japanese is abbreviated as 'NJ' (for 'new Japanese') to avoid

1

confusion with Middle Japanese (MJ). By 'cNJ' is meant contemporary NJ, the Japanese of the twentieth century. 'Classical Japanese' is the fossilized, relatively fixed written norm which arose largely out of the language of the twelfth century and which thereafter remained the dominant form of writing in Japanese until the beginning of the twentieth century. This book is not to any great extent concerned with the fossilized Classical Japanese written norm.

Extensive attestation of Japanese goes back to the eighth century. Old Japanese is mainly the language of the Nara period, although it also comprises earlier texts which are included in sources compiled or completed in the eighth century, but their age is difficult to assess. It is sometimes said that Japanese has not changed greatly since Old Japanese. However, it is possible to identify two large sets of internal change, largely coinciding with EMJ (phonological change) and LMJ (grammatical change), respectively, as well as two waves of contact induced change which took place during OJ/EMJ (sinification) and late NJ (westernization), which together transformed the language from its Old Japanese to its modern form.

In the transition between OJ and EMJ and within EMJ, the language underwent significant phonological changes, both in syllable structure and in segmental phonology. At the end of the EMJ period, the phonological structure of Japanese was largely as it is today and phonological changes since then have by comparison been minor. Most major phonological changes were complete by the end of the eleventh century and some scholars do not include the last century of the Heian period, the so-called *Insei* period (1086–1185), as a part of EMJ, but instead as part of LMJ. It was also during the EMJ period that the contact with Chinese manifested its influence in the texts, although we suspect that this influence was present already in OJ.

During the LMJ period by contrast, major grammatical changes took place, which affected both morphology and syntax. Some of these changes are initiated or anticipated in the twelfth-century materials, but they are mainly reflected in the written sources of the LMJ period. The dating of the end of LMJ is difficult to determine on linguistic grounds. The main issue is whether to include the language stage reflected in the Christian materials from the end of the sixteenth century and early years of the seventeenth century in LMJ or in NJ. The sweeping grammatical changes which took place during LMJ were complete by then and in many respects the language of those sources is similar to NJ. However, it is also very similar to the language in an earlier set of sources, the so-called *shōmono* which date from the middle of the fifteenth to the middle of the sixteenth century. The Christian sources, alongside the *shōmono*, are accordingly included in the LMJ period, and we will consider them to constitute the end of LMJ.

NJ is thus the language from then on. Once the phonological changes of EMJ and the grammatical changes of LMJ were complete, Japanese did not

change very much structurally. However, NJ was affected by significant external changes brought about by extensive contact with European languages in the course of the modernization of Japan and Japanese in the late nineteenth and early twentieth centuries.

Attestation of forms

OJ, EMJ and LMJ are dead languages and the sources are closed and limited text corpora. Naturally, many specific forms of individual words are not attested in those texts. The facts and state of attestation are of paramount importance when investigating the system of a language, but of less interest when explaining or exemplifying the system. Recall Winston Churchill's famous experience, recounted in the first chapter of his *My Early Life*, of being set the task upon arrival at his boarding school at the age of seven of learning the singular of the noun of the first declension in Latin. He managed to memorize the paradigm and reproduce it to the satisfaction of his teacher, but he did not understand what it meant and asked about it:

'But,' I repeated, 'what does it mean?'
'Mensa means a table,' he answered.
'Then why does mensa also mean O table,' I enquired, 'and what does O table mean?'
'Mensa, O table, is the vocative case,' he replied.
'But why O table?' I persisted in genuine curiosity.
'O table, – you would use that in addressing a table, in invoking a table.' And then seeing that he was not carrying me with him, 'You would use it in speaking to a table.'
'But I never do,' I blurted out in honest amazement.
'If you are impertinent, you will be punished, and punished, let me tell you, very severely,' was his conclusive rejoinder.

An inflectional paradigm is a pattern of relations; the combination of mutually exclusive inflectional endings with stems of words; or, put differently, the morphologically possible forms of a word. It is possible that no one ever produced the vocative of *mensa* in actual speech or writing in classical Latin. But if anyone wished to address a table, the vocative case was available for that purpose and even if it was never spoken, the vocative form of that noun existed in Latin as a systemic possibility. In OJ, the imperative of the verb *kog-* 'row' is not attested. There is no systemic reason that *kogye* 'row!' should not exist, much less so than for the vocative of Latin *mensa*. The non-occurrence of *kogye* may be regarded as an accident of attestation. Given the existence of OJ *kog-*, which is amply attested in various forms, we know that its imperative would be *kogye*. On the other hand there are restrictions on the use of the imperative of certain types of verbs, e.g. those which signify spontaneous, nonvolitional action. It is debatable whether this is a morphological restriction (that form does not exist) or a syntactic/pragmatic restriction

4   Introduction

(that form cannot be used, it does not make sense for that form to be used), but it is a grammatical restriction in the language, not an accident of attestation. Throughout the book morphological paradigms, particularly verbs, are exemplified. Some of the forms of some of the verbs are not attested, but this is generally not remarked, except for small, irregular verb classes or grammatical auxiliaries where the situation of attestation may be of significance.

Conventions

Throughout the book a phonemic transcription of cited forms is generally employed, appropriate to the period of citation. For example, the word for 'front', which is NJ *mae*, will be written as shown in (2), reflecting its phonemic shape at different stages of the language (exemplified by texts from those periods):

(2)   OJ   early EMJ   mid EMJ   late EMJ onwards
         Man'yōshū   *Tosa nikki*   *Genji monogatari*   *Sarashina nikki*
         mapye   mape   mawe   mae

Japanese editions of pre-modern texts, by contrast, employ a historical spelling and will spell 'front' as まへ ('mahe') regardless of the period from which the text dates. From late EMJ until the second half of NJ this word was, like all words of the shape /Ve/, pronounced with a palatal onglide before the /e/, [maʲe], but that will not be noted in the phonemic transcription. Nor will other allophonic features of pronunciation, such as the prenasalization of /b, d, g, z/ (which from OJ into the LMJ period were pronounced [ᵐb, ⁿd, ᵑg, ⁿz]). However, when transcribing forms with a moraic consonant (which in EMJ and early LMJ was phonemically underspecified for nasality in morpheme internal position), we will use a semi-allophonic transcription and write *punde* 'brush', not *puCde*, although strictly speaking the phonemic shape was /puCde/; see further 7.1.3 about this.

When citing words or passages from OJ texts italics are used for phonographically written text portions and normal type for logographically written text. For example, when citing the word *suru*, the adnominal form of *se-* 'to do', it is transcribed in accordance with the writing in the source as exemplified in (3), with different writings of *suru* cited from different poems in the *Man'yōshū*. In (3a), 須 and 流 are used phonographically for the syllables /su/ and /ru/. In (3b), 為 is used logographically to write part of a form of a verb meaning 'to do' and 流 is used phonographically for the syllable /ru/. Finally, in (3c), 為 is used logographically for a form

of a verb meaning 'to do' which reading tradition interprets as the adnominal form.

(3)   *MYS* poem no.   Writing   Transcription
   a.   17:3932   須流   *suru*
   b.   19:4198   為流   su*ru*
   c.   2:164   為   suru

Words from cNJ will be noted in the modified Hepburn transcription, for example *fuji* '(Mount) Fuji' or *chōshoku* 'breakfast'. Japanese script will generally not be used, except when discussing script and writing. However, where relevant the Chinese characters (*kanji* 漢字) used to write Chinese or Sino-Japanese words are given.

Japanese proper names will usually be given in their usual cNJ form, e.g. *Hōjōki* (not the shape this word would have had at the beginning of the thirteenth century: *Paŭdyaŭ-ki*) and personal names are given in the traditional Japanese order (surname, given name), e.g. Hattori Shirō. Modern Chinese words will be transcribed in *pinyin*. Early Middle Chinese reconstructions follow Pulleyblank 1991.

When citing verbs, the basic stem is used, rather than the OJ/EMJ/ early LMJ conclusive and the late LMJ/NJ nonpast form, which are the forms usually used as citation forms, i.e. the forms used to cite or talk about a verb, for example in dictionaries.[1] Citing the basic stem in the majority of cases unambiguously identifies the conjugation class of a verb, especially when comparing quadrigrade and bigrade verbs, e.g. OJ *ok-* 'put' (quadrigrade), *ake-* 'dawn' (lower bigrade), *okwi-* 'arise' (upper bigrade), whereas that is not the case with the OJ/EMJ/early LMJ conclusive: *oku*, *aku*, *oku*.

Almost all verb suffixes attach directly to the basic stem of regular vowel base verbs, but for some irregular vowel base verbs and for all consonant base verbs, most verb suffixes attach to one of several *derived stems* (see further 3.4.4 and 8.1.4). When citing verbal suffixes a morphophonemic notation is used which shows which, if any, derived stem the suffix attaches to, by using bracketed prefixes: (a), (i), (e), (I). This is exemplified in (4), using the verb *sak-* 'to come into bloom'. The *a-* stem corresponds to the *mizenkei* of Japanese school grammar, the infinitive to the *ren'yōkei*, and the exclamatory to the *izenkei* (see 3.4.6).

---

[1] Notable exceptions, which list verbs under their infinitive, are Ohno's dictionary of pre-modern Japanese (1990) or dictionaries published by the Jesuit missionaries in the early seventeenth century, e.g. Rodrigues (1603–4).

6   Introduction

(4)   Prefix   Stem   Verb   Suffix   Verb + suffix
      (a)   *a-* stem   *saka-* *-(a)n-* (Negative) *saka-n-*
      (i)   Infinitive   *saki-* *-(i)n-* (Perfective) *saki-n-*
      (e)   Exclamatory   *sake-* *-(e)do* (Concessive) *sake-do*
      (I)   *onbin* stem   *sai-*   *-(I)ta* (Past)   *sai-tar-*

This notation does not mean that the prefixes are part of the (synchronic) phonemic shape of the suffixes. The prefixes only show which stem a suffix selects. For example, for both the negative, *-(a)n-*, and the perfective, *-(i)n-*, the basic stem shape of the suffix is /-n-/, but they select different stems, as shown in (4).

In examples, verb forms will be segmented as in (5), i.e. noting (by '-') morpheme boundaries between verb stems and auxiliaries, but not between flectives and the immediately preceding verb or auxiliary stem. In glosses, the inflected form will be noted as part of the gloss for a verb or auxiliary, separated by '.', showing that *saku* is the conclusive form of the verb *sak-* 'bloom', *sakedo* is the concessive form of that verb, and *-kyeri* is the conclusive form of the modal past auxiliary *-(i)kyer-*. This notation shifts the focus away from individual morphemes to the actual inflected word forms. See sections 3.1 and 3.4 for the analysis underlying this notation.

(5) a.   *saku*
         bloom.CONCL
         'it blooms'

    b.   *sakedo*
         bloom.CONC
         'although it blooms'

    c.   *maywopi-ki-ni-kyeri*
         fray-come-PERF-MPST.CONCL
         'it had become frayed'

Where possible, examples from Japanese texts are cited from Iwanami's critical edition *Nihon koten bungaku taikei* (Iwanami 1957–69). Examples from Nara period *Senmyō* (Imperial edicts, see 1.2.3.2) are cited from Kitagawa (1982) and texts originally published in print, such as the Christian materials from the end of the sixteenth and beginning of the seventeenth centuries (see 10.2.2), are cited from the originals (or photographic reproductions). References to Rodrigues's *Arte da lingoa de Iapam* (1604–8) are to Doi's (1955) translation into Japanese, which is more easily accessible for most readers and which gives page references to the original.

Introduction   7

   As few references as possible are given in the body of the text, but a short
list of references is provided at the end of most chapters; where possible refer-
ences are given to scholarship in English. Overall, factual information or
generally accepted descriptions which are available in common handbooks,
overviews and dictionaries, or by looking at the texts, are not referenced.
Handbooks frequently consulted include: *The Japanese language through time*
(Martin 1987), *Jidai-betsu kokugo daijiten: Jōdaihen* (Omodaka *et al.* 1967),
Kokugogaku daijiten (Kokugogakkai 1980), *Kokugogaku jiten* (Kokugogakkai
1955), *Nihon kokugo daijiten* (Shōgakukan 2000–2), *Kokugogaku kenkyūjiten*
(Satō 1977), *Nihon bunpō daijiten* (Matsumura 1971), *Nihongo bunpō daijiten*
(Yamaguchi and Akimoto 2001), *Nihongogaku kenkyūjiten* (Hida *et al.* 2007),
Nihongo hyakka daijiten (Kindaichi *et al.* 1988). Of these, the *Nihongogaku
kenkyūjiten* has an exhaustive listing and description of available textual
sources (pp. 629–1129).

Part I

Old Japanese

1   Early writing in Japan and Old Japanese sources

1.1   Writing

1.1.1   Introduction of writing in Japan

The Japanese were exposed to written matter as early as the late Yayoi period (*c.* ?1000 BC – 300 AD). Thus, inscribed Chinese coins have been unearthed in excavations of sites dating from the first century AD. There is no evidence of any awareness of the function of writing at that early stage, however, and it is likely that the characters which appear on mirrors and other artefacts produced in Japan through the third and fourth centuries were also simple ornaments, in imitation of those found on articles from the continent. To all appearances, writing as such, in the form of Chinese Classics, was introduced into Japan early in the fifth century as part of the great cultural influx from Paekche. The *Kojiki* and the *Nihon shoki* recount this event as the advent of the scribes Wani and Akichi in the years Ōjin 15 and 16 (thought to be early in the fifth century, possibly 404–5; the traditional dating puts this at as early as 284–5, two 60-year cycles earlier). For some time, writing remained in the hands of hereditary professional scribes (*fubito*) who were of continental herit-age. Through the sixth and seventh centuries Sinitic culture, including Chinese Buddhism, flowed into Japan through Paekche. In the course of this, written Chinese assumed enormous importance in matters of state, philosophy and religion. Any serious engagement with such matters required knowledge of written Chinese and for some time writing was equivalent with writing in Chinese. Also composition of Chinese poetry became highly regarded and remained so long into the medieval period. Thus, the oldest surviving poetry anthology in Japan is the *Kaifūsō* (懷風藻; *c.* 751) a compilation of Chinese poetry written in Japan. Reading and interpretation of Chinese canonical texts came to assume great importance, both within Buddhism and in government administration. Chinese texts were read in two ways: either (a) reading them out in a form of Chinese (*ondoku* 音読 'sound reading'); or (b) translating or rendering the texts into Japanese (*kundoku* 訓読 'gloss reading'). These two practices have exerted great influence both on the Japanese language itself and on the way it is written. This will be discussed in detail in 9.1.
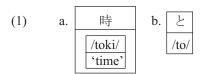
11

1.1.2   *Writing in Japanese*

The earliest attestation of writing in Japanese dates from the fifth century, but it is not until the middle of the seventh century that writing in Japanese became widespread. Throughout the OJ period Japanese was written entirely in *kanji* which were used logographically or phonographically.

1.1.2.1   *Logographic versus phonographic writing*
Writing is a *representation* of language: elements of writing represent elements of language. Depending upon the nature of the linguistic elements that elements of writing stand for, there are in principle two types of writing. First, writing which represents those elements of language which *carry* meaning: words or morphemes. This is *logographic* writing. Second, writing which represents those elements of language which *distinguish* among elements carrying meaning: phonemes or phonological units of greater or smaller extent. This is *phonographic* writing. Below, these two types of writing are illustrated with examples from NJ. (1a) shows logographic writing, with 時 standing for the word which has the sound shape /toki/ and the meaning 'time'; 時 does not stand primarily for the meaning or the sound shape, but for the word, the linguistic sign, which comprises both. (1b) exemplifies phonographic writing, with と standing for the syllable /to/; と can thus be used to write any occurrence of /to/ regardless of the word of whose sound shape /to/ forms part.

(1)   a.   時   b.   と
              /toki/   /to/
              'time'

Actual orthographic systems and practices rarely, if ever, limit themselves to one of these types of writing. For example, Chinese, which is the stock example of logographic writing, has a strong phonographic element, DeFrancis (1984) arguing that this is more prominent than the logographic element. Conversely, most writing systems have a logographic element. This includes alphabet writing as used to write English; for example, *red* and *read* (past tense of the verb 'to read') are written differently although they are homophonous. Also spaces between words, capitalization of some words, and punctuation all contribute an element of logography to alphabet writing.

1.1.2.2   *Adaptation of Chinese script*
It is not known specifically when or how the Chinese script began to be used to write Japanese. Nor is it clear by what stages this took place. It *is* clear, though, that it makes little sense to consider this from a purely Japanese per-