

# Productivity and Constraint on Multi-Verb Constructions in Old Japanese

Wenchao Li

Department of Japanese Studies, Zhejiang University

Zhejiang, China

E-mail: [widelia@zju.edu.cn](mailto:widelia@zju.edu.cn)

Received: June 30, 2019

Accepted: July 18, 2019

Published: August 6, 2019

doi:10.5296/ijl.v11i4.15002

URL: <https://doi.org/10.5296/ijl.v11i4.15002>

## Abstract

Old Japanese is a dead language from the Asuka and Nara periods (7th - 8th century AD). Its writing system, case system and word order make it distinct from Modern Japanese in many respects. This study presents a quantitative linguistic analysis to the patterns of multiple verb combining in Old Japanese. To this end, two databases were built: multi-verb construction in the Early Nara Period written in *variant Chinese* (AD. 712) and *purely classical Chinese* (AD. 720), and multi-verb construction in the Late Nara Period written by *man'yōgana* (AD. 759). The findings reveal that, in the Nara period, the formation of multi-verb constructions is an issue of verb serialising and is facilitated at a syntactic level. Grammaticalisation of unaccusative change-of-state verbs and motion verbs results in tighter integrity of multiple verbs, which, in turn, inspires the device of verb compounding. The entropy of the  $V_{\text{final}}$  unaccusative reveals that the formation via verb serialising is more productive than the formation via verb compounding.

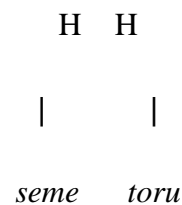
**Keywords:** Japanese, Multi-verb construction, Word formation, Historical linguistics

## 1. Introduction

Typologically speaking, Modern Japanese is alleged to be an Altaic language due to the fact that it is agglutinative and employs a versatile inflection system to convey tense, aspect, voice and modality. Old Japanese, on the other hand, is a dead language from the Asuka and Nara periods (7th - 8th century AD) and presents a rather different picture. Its writing system, case system and word order distinguish it from Modern Japanese in many respects. Early Old Japanese script is comprised of three forms: *Junsei-kanbun* 'purely classical Chinese', *hentai-kanbun* 'variant Chinese' and *man'yōgana*. *Man'yōgana* laid the foundations for the pure phonetic script that was invented in the following period, i.e. the Heian Period

(794-1185). Against this backdrop, the lexical features of Old Japanese ought to be tackled in more depth. Through an examination of word formation, we may come to a better understanding of the Japanese language in the Nara period before the *kana* script was developed. Synchronically, multi-verb construction in Modern Japanese is deemed being a matter of verb compounding (Kageyama 1993; Yu Moto 2005, 2008, 2011, 2012; Nishiyama and Ogawa 2014, etc.). The combination of multiple verbs in Old Japanese, however, raises different opinions. In earlier times, most linguists denied the existence of verb compounds. Kindaichi (1953) is one of the supporting linguists who argues that each constituent in a multiple verb construction in Old Japanese has its own accents, and that it is therefore not proper to consider them as being assigned to a compounding relation. (1) is an illustration of Kindaichi's (1953:334) accent-based hypothesis.

(1) The accent of 攻め取る *seme-toru* 'attack-confiscate'.



Kindaichi's accent-independence theory is challenged by Seki (1977) and Komatsu (2000), who argue that the accent of Old Japanese lexicons can be affected by Chinese cadence (Note 1) due to the fact that Old Japanese was written in classical Chinese characters.

Kindaichi (1953) proposed the 'kakarijoshi (Note 2) insertion'- criterion to support the idea that multi-verb construction cannot be composed via verb compounding. He demonstrates that 咲き初む *saki-somu* 'start blooming' and 降りまさる *ori-masaru* 'snow-increase' can sometimes be 咲きか初むらん *saki-ka-somuran* 'It is just about to bloom'; 降りぞまされる *ori-zo-masareru* 'snowing, snowing, snowing, it does not seem to stop' (1953:335).

Nakamura (1971) and Seki (1977) do not approve of this idea. They claim that *kakarijoshi* merely emphasises the action/topic and thus cannot impede the formation of a verb compound. There are plenty of V-Vs carrying *kakarijoshi* but are well-formed. Instances include 思ひは過ぎず *omoi-wa-sugiru* 'think-keep thinking' and 思ひぞ焼くる *omoi-zo-sakuru* 'to be vexed by continuously missing somebody'. Seki (1977) proposes three types of verb compounds in Old Japanese: modifier-relation V-V, auxiliary-relation V-V and accordance-relation V-V.

When tackling Old Japanese linguistic phenomena, due attention should be paid to the various significant ways in which Old Japanese differs from Modern Japanese: the writing system, the word order, the case system.

### 1.1 Writing System in Old Japanese

Modern Japanese is made up of two moraic scripts (*hiragana* script and *katakana* script) and characters (*kanji*). In the Nara Period (710-794), Chinese characters were borrowed to represent vernacular Japanese on paper. In Early Old Japanese, the *hentai-kanbun* ‘variant Chinese’ script was used, as is illustrated in *Kojiki*. This work is the oldest extant chronicle in Japan, recording myths as well as the history of Mikado, the Emperor of Japan. An example can be seen in (2) below.

(2) 故志 能久途 途... 佐加志 賣 遠 阿理登 岐加志 呂 佐用婆比 爾

Kwosi	no	kuni	ni	sakasi	mye	wo	ani	to	kikosite	sa-ywobapi	ni	
Kosi	COP	Land	DAT	ADJ.COP-ADN	maiden	ACC	exist	COMP <sup>2</sup>	hear-respect	PREF	courtship	DAT
阿理	多多斯	用婆比	途	阿理	加用婆	勢						
ani-	tata-si	ywobapi	ni	ani-	kaywopa-	se						
PREF	set out	courtship	DAT	PREF	go out-respect	EXCL						

‘Hearing that there was a pretty maiden, (the emperor) set out to woo her, went out to propose to her’

(KK 2)

In *Kojiki*, the conjugations are rendered by Chinese characters: the gerund form is conveyed by 呂 ‘て’ (3), the provisional form is denoted by 婆 ‘ば’ (4), and the adnominal form of *ichi dan* conjugation is rendered by the Chinese character 流 ‘る’ (5) etc.

(3) Gerund form conveyed by Chinese characters 呂 ‘て’

阿佐比 能 惠美 佐迦延 岐呂

asa pi no wemi sakaye kite

morning sun COP radiantly smiling come.GER

‘come smiling radiantly as the morning sun’

(KK 4)

(4) Provisional form conveyed by Chinese characters 婆‘ば’

伊刀古 夜 能 伊毛 能 美許等 牟良 登理 能和 賀 牟禮 (Note 3) 伊那 婆

itwokwo ya o inwo no mikoto mura tori no wa ga nure ina ba

beloved.person Excl COP wife COP honourable.person flock bird COP I GEN flock-go. away when

比氣 登理 能和 賀 比氣 伊那 婆

pika tori no wa ga pika ina ba

company bird COP I GEN be.accompanied go.away when

‘My beloved wife, when I go off with my men, flocking like flocking birds, when I go off, with my men, accompanied like birds of a company.’

(KK. 4)

(5) The adnominal form of *ichi dan* conjugation is rendered by the Chinese character 流‘る’.

伊夜 佐岐 陀弓流 延袁 斯 麻加 牟

iya saki- dat-eru ye wo si maka- mu

PREF first arrive.ADN older ACC do.INF many.CONJ

‘Many the older, and the first person who reaches the top of the mountain’

(KK. 16)

Another script in the Early Nara Period is *Junsei-kanbun* ‘purely classical Chinese’, which is deemed the official language of the Nara Period. *Nihon shoki kayō*, the second oldest book of classical Japanese, is representative of this. An illustration is given in (6).

(6)	夜句茂	多菟	伊弩	毛夜霸餓岐	菟磨語味	爾
	ya-kumo	tatu	idu	moya-pyegaki	tumagome	ni
	many clouds	rise	Idumo	many-fenced palace	spouse dwell	DEM
	夜霸餓枳		菟俱盧 迺	夜霸餓岐	迺	
	ya-pyegaki		tukuru sono	ya-pyegaki	wo	
	many-fenced palace	build	DEM	many-fenced palace	ACC	

(NSK.1)

The script *man'yōgana* is employed in Late Old Japanese, which borrows Chinese characters in three ways: (a) solely borrowing semantic meaning (logographic writing); (b) solely borrowing phonological value (phonographic writing); and (c) borrowing both semantic and phonological values. Crucially, these patterns are not separated, but mixed. Generally, phonographic writing is likely to render nouns, verbs, and adjectives, while logographic writing tends to convey case particles. *Man'yōshū* is written in this script. In (7), both logographic writing and phonographic writing are employed.

(7) 霞 立 長 春 日 乃 晚 家 流

kasumitatu nagaki paru pi no kure-ni-kyeru

mist trail long.ACOP.ADN spring day GEN twilight COP turn

'The Mist rises, the long spring day has turned to twilight'

(MYS.15)

霞 (noun), 立 (verb), 長 (adjective), 春 (noun), 晚家 (noun) and 流 (verb) comprise logographic writing, while 日 (noun) and 乃 (genitive case particle) comprise phonographic writing. *Man'yōgana* laid the foundations for *hiragana* and *katakana* in the Heian period.

### 1.2 Case System in Old Japanese

The case system presents a rather different picture. が *ga*, a nominative case particle in Modern Japanese, functions as a genitive particle in Old Japanese, as can be seen in (8). が's function of nominative case particle existed in the database but was limited to conditional sentences, as can be seen in (9).

#### (8) が's main role in Old Japanese: genitive case particle

和 何 許許呂

wa ga kokoro

I GEN heart

'My heart.'

(KK.2)

(9) が's function as a nominative case particle

宇美 賀 由氣 婆  
 umi ga yuke ba  
 sea NOM go. PROV.

(KK36)

The accusative case particle を o in Old Japanese was an exclamatory particle, as can be seen in (10).

(10) 夜久毛 多都 伊豆毛 夜弊賀岐 都麻暮微 (Note 4) 爾 夜弊賀岐 都久流  
 ya-kunwo tatu idumwo ya-pye-gaki tumagomwi ni ya-pye-gaki tukuru  
 many clouds rise Idumo many-fenced palace spouse dwell DAT many-fenced palace build  
 曾能 夜弊賀岐 袁  
 so no ya-pye-gaki wo  
 DEM many-fenced palace EXCL

'In the many-fenced palace of Idumo, many clouds rise. To dwell there with my spouse, I build a many-fenced palace. Oh, that many-fenced palace!'

(KK.1)

### 1.3 Word Order in Old Japanese

This study assumes that, when transmitting Chinese characters to Japan, the diplomats to the Tang Dynasty (*kentooshi*) went back to Japan by way of the Korean Peninsula. As a result, the grammar of Old Paekche had a great deal of influence on Japanese grammar. An immediate example that springs to mind is the agglutinative property; for instance, the position of dative case particle に (written as Classical Chinese character 迺).

(11) 阿遠夜麻 迺 比賀 迦久良婆 奴婆多麻 能 用 波 伊傅 那牟.  
 awo yama ni pi ga kaku raba nubatama no ywo pa ide namu  
 green mountain DAT sun GEN hide. COND black GEN night TOP emerge-PERF-CONJ

'as soon as the sun hides behind the verdant mountains, the jet-black night comes.'

(KK. 4)

Another striking property that resulted from Old Paekche is the word order. Two orders coexisted in Old Japanese: SOV, as used in Altaic languages (12), and SVO, as is the case in Chinese (13).

**(12) SOV order in Old Japanese**

君 乎 待出牟.

kimi wo mati-ide-mu

you ACC wait-come out. CONJ

‘(I) will wait for you to come out’.

(MYS.11.2484)

**(13) SVO order in Old Japanese**

穗 庭 開不出.

po ni pa saki-zu-de

ear of rice garden bloom-NEG-come

‘The ear of rice cannot bloom in a garden’.

(MYS. 10.2283)

The negation 不 ‘zu’ consistently appears before the verb, which clearly is a feature of Chinese. This is further illustrated in MYS. (10.2307).

(14) 不 出 跡 念 者.

zi ide to omope ba

NEG come COMP think.PROV.

(MYS. 10.2307)

Intriguingly, SVO order and SOV order can both appear in the same poem. (15) is an illustration of this taken from *Manyoshu*, where the prepositional phrases 従山出流 and 妹待吾 (in bold) present an SVO order, while 月待登 and 人尔波言而 (underlined) display an SOV order.

(15) 足日木乃 (Note 5) 從 山 出流 月待 登

*asipikwi no ywori yama iduru tukwi matu to*

Asihiki COP PREP mountain flow.CONCL moon wait.CONCL.CONCL

人 尔 波言而 妹 待 吾 乎。

*pito ni pa ipite imo matu ware wo*

people DAT fairg.GER lover wait me Q

'I look at the moon, which rises from the mountain and say to the local people, would my lover wait for me.'

(MYS.12.3002)

Indeed, prepositions are rather favoured by Old Japanese. Another illustration is given in (16).

(16) 從 山 出流

*ywori yama iduru*

from mountain flow.CONCL

'(The stream) flows from mountain'.

(MYS.12.3002)

These are, however, preliminary illustrations, and Old Japanese has a great deal more that should be examined in detail. The present study calculates the combinatorial possibilities of different types of roots and delves into the entropy and complexity of various combinations. Against this backdrop, we might arrive at an overall picture of multiple verb combinations in Old Japanese. The linguistic periodisation of the development of the Japanese language is as follows:

- (a) Old Japanese: AD 700 – AD 800
- (b) Early Middle Japanese: AD 800 – AD 1200
- (c) Late Middle Japanese: AD 1200 – AD 1600
- (d) Modern Japanese: AD1600 – present

This paper is structured as follows. Section 2 provides insight into the concepts relevant to this study, namely verb serialising and verb compounding. Section 3 presents quantified data and explores the formation of multiple verbs. Section 4 highlights the results and concludes the paper.



## 2. Methodology

The purpose of this study is to examine how multiple verbs are combined in Old Japanese and how causative/inchoative compound verbs are altered. This obliges us to touch upon some general concepts that are relevant to this study: verb serialising and verb compounding. Furthermore, to pin down the semantic compositionality and productivity of multi-verb constructions, this study incorporates quantitative linguistic insights for data analysis purposes.

### 2.1 *Verb Serialising and Verb Compounding*

Serial verb construction is particularly prominent in East Asian and African languages. The term ‘serial verb construction’ (SVC) was initially put forward by Westermann (1930: 126) as designating “A row of verbs one after another [in which] the verbs stand next to each other without being connected”. SVC has been studied intensively in linguistic typological work ever since; see, for instance, Andrews (1997); Aikhenvald (1999, 2006); Baker (1989); Foley (2009); Foley and Olson (1985); Stewart (1963). Linguists from various different institutions have all attempted to define SVCs. As yet, however, no definition has been satisfactorily accepted. In earlier times, an SVC was deemed “a surface string of verbs or verb-like items that occurs within what appears to be a single clause” (Bradshaw 1982). A different view comes from Griffiths (1991), who demonstrates that serial verbs do not require complementisers to link multiple verbs. Collins (1993: 91) holds a similar position, arguing that an SVC is a succession of verbs and their complements (if any), with one subject and one tense value that are not separated by any overt marker of coordination or subordination.

Another pathway comes from Baker (1989) and Stewart (1963, 2001), who propose a criterion for forming an SVC known as ‘object-sharing’. However, this view is disputed by Aikhenvald (1999, 2006) and Crowley (2002). They argue that, in Edo language, some multi-verb constructions are composed of ‘intransitive verb + transitive verb’ where the verbs apparently do not share an object. Multi-verb constructions are extensively seen in Altaic languages (Korean, Japanese, Uzbek, etc); South Asian languages (Thai, India, Pakistan, etc); Caucasian languages (Tsez, etc) and Chinese. Li (2018) draws together the evolution of the combinatorial possibilities of multiple verbs in Chinese from three different periods (Old Chinese, Pre-Middle Chinese and Middle Chinese) and pins down the distinctions between verb serialising and verb compounding as follows.

Verb serialising is a phenomenon in which,

- (a) Semantically, the events rendered by the multiple verbs are either successive or co-ordinate. Each verb is independent and conveys an individual event.
- (b) Syntactically, the multiple verbs all function as heads. ‘Argument-sharing’ may be necessary but is not obligatory. Essentially, there must not be an insertion between multiple verbs.
- (c) The combinations are not arbitrary, but nor are they strict. Crucially, multiple verbs are not necessarily of different types.

Verb compounding is a phenomenon in which,

- (a) Semantically, the two verbs jointly constitute a single event. In this regard, the successive relation of multi-verb constructions no longer exists.
- (b) Syntactically, the formation is constrained in two respects. First, ‘argument-sharing’ is obligatory: verbs in the series should be of the same type and share either the external argument or the internal argument. Moreover, one verb acts as the head and the other functions as the complement. The result of this is that only two types of V-V can be licensed, namely modifier-predicate V-V and predicate-complement V-V.

This study will now follow Li’s insight and carry out a corpus-based analysis to account for the ways in which multiple verbs are combined in Old Japanese from a diachronic perspective.

## 2.2 The Data

To approach a better understanding of how multiple verbs are combined in Old Japanese, this study examines one example of spoken Japanese material, 続日本紀宣命 *Shoku nihongi Senmyō*, and three documentary examples of written Japanese, i.e. :

### (17) Data for Old Japanese

古事記 *Kojiki* (AD. 712): 112 poems; approx. 2527 words (KK)

日本書紀 *Nihon shoki* (AD.720): 133 poems; approx. 2444 words (NSK)

万葉集 *Manyoshu* (AD. 759): 4685 poems; approx. 83706 words (MYS)

続日本紀宣命 *Shoku nihongi Senmyō* (AD. 797): approx. 14,000 words

*Senmyō* is made up of imperial edicts and *senmyō-gaki* writing is mostly rendered logographically, with some grammatical items rendered phonographically. There are 62 edicts of *Senmyō* recorded in *Shoku Nihongi*. Below is *Senmyō* 1.

(18) *Senmyō* 1. Monmu 1 8th Month (697), Emperor Monmu

1. 現御神止大八嶋国所知天皇大命良麻止詔大命乎集侍皇子等王等百官人等天下公民諸聞食止詔。

2. 高天原尔事始而遠天皇祖御世中今至 麻弓尔天皇御子之阿礼坐牟弥繼繼尔大八嶋国将知次止天都神乃御子随母天坐神之依之奉之随此天津日嗣高御座之業止現御神止

大八嶋国所知倭根子天皇命授賜比負賜布貴支高支広支厚支大命乎受賜利恐坐弓此乃  
食国天下乎調賜比平賜比天下乃公民乎惠賜比撫賜牟止奈母随神所思行佐久止詔天皇  
大命乎諸聞食止詔。

3. 是以天皇朝廷敷賜行賜幣留百官人等四方食国乎治奉止任賜幣留国々宰等尔国法乎  
過犯事無久明支浄支直支誠之心以而御称々而緩怠事無久務結而仕奉止詔大命乎諸聞  
食止詔。

4. 故如此之状乎聞食悟而欵将仕奉人者其仕奉礼良牟状随品々讚賜上賜治将賜物曾止  
詔天皇大命乎諸聞食止詔。

The above data can be extracted from *The Oxford Corpus of Old Japanese* <http://vsarpj.orinst.ox.ac.uk/corpus/>. The gloss is made by the author. The author manually marked the characteristics of the verbs (e.g. unergative, unaccusative, transitive (Note 6)) that contributed to multi-verb formations. A list of gloss abbreviations is provided at the end of the paper.

In addition, *Senmyō* are imperial edicts and 日本書紀 *Nihon shoki* is the official documentary history of the dynasty. The lexicons in the two are therefore expected to be simple and elegant, leading to an assumption of a low entropy and a high redundancy in *Senmyō* and *Nihon shoki*. On the other hand, 古事記 *Kojiki* and 万葉集 *Manyōshū* are poems, meaning that their expressions might be complex and lead to a high entropy.

Two databases were assembled for the purposes of this study:

(a) Multi-verb construction in the Early Nara Period written in *variant Chinese*, *purely classical Chinese* and *senmyō-gaki*. The data was randomly drawn from *Kojiki* (AD. 712: *variant Chinese*), *Nihon shoki* (AD.720: *purely classical Chinese*) and 続日本紀宣命 *Shoku nihongi Senmyō* (AD. 797: *senmyō-gaki*).

(b) Multi-verb construction in the Late Nara Period written in *man'yōgana*. The data was randomly extracted from 万葉集 *Manyōshū* (AD. 759).

### 3. Productivity and Entropy

The term  $P$  ‘productivity’ was put forward by Baayen and Lieber (1991); Baayen (1992). The formula is shown in (19).

$$(19) P = n_1 \div N$$

$N$  refers to the token number of V-Vs and  $n_1$  refers to hapax legomena, i.e. the V-V that only appears once.

The term  $E$  ‘entropy’ was first introduced by Shannon (1948) in his book *A Mathematical Theory of Communication*. Tamaoka et al. (2004) introduce this concept into the study of Japanese verb compounding. Entropy may help us to calculate a certain  $V2/V1$ ’s variations in combining different  $V1$ s / $V2$ s. The formula is presented in (20).

$$(20) H = -\sum_{j=1}^j p_j \log_2 p_j$$

In the formula,  $p$  stands for the probability of appearance for a specific compound verb among all the compounds created with the  $V2$  or  $V1$ . For instance, suppose a  $Vx$  is a partially grammaticalised unaccusative verb. It appears as a  $V2$  and may form a verb compound with a transitive verb, an unergative verb or an unaccusative verb. The type’s number is 20. In our database, 50 tokens of V-Vs were detected to be formed by the  $Vx$  with 20 different types of  $V1$ . Among them, the  $V1$  ‘ $Vy$ ’ has the highest frequency, i.e. 9. Dividing 9 by 50, we will have:

$$(21) \quad ‘p’: 0.18.$$

Building on this,

$$(22) \text{Log}_2 0.18 = -2.47.$$

$$(23) \text{For } Vx, p_j \log_2 p_j: 0.18 \times -2.47 = -0.445.$$

To arrive at  $H$ , we calculate the remaining 19 types of verbs; add them up and then divide by -1.

#### 3.1 Multiple Verb Combinations in the Nara Period

With the methodology outlined, we may now proceed to the analysis of the data. It should be pointed out that the written system in the Early Nara Period was *hentai-kanbun* ‘variant Chinese’ and classical Chinese. In the Late Nara Period, Chinese characters were further cursivised, which resulted in *man’yōgana*. Essentially, a sign of verb weakening arises, giving rise to a verb compounding strategy in combining multiple verbs. With this in mind, it would be proper to divide this discussion into two parts: multiple verb combinations in the Early Nara Period, and multiple verb constructions in the Late Nara Period.

The Early Nata Period database shows that the most frequent option in terms of multiple verb combinations is to combine three verbs (227 tokens). The second largest number of tokens was attributed to two-verb constructions (225 tokens). 89 tokens were applied to four-verb combinations, and 61 went to the combination of five verbs. The various formation possibilities are illustrated in (24).

(24) **Three-verb construction** (227 tokens)

a. [Transitive + Transitive + Unacc.]

e.g. 宇知夜米許世 *uti-yame-kose* ‘hit-stop-come’ (KK.2)

b. [Tran. + Tran. + Tran.]

e.g. 設備待 *mooke-sonae-matsu* ‘supply-build up-wait’

c. [Tran. + Unacc. + Unacc.]

e.g. 持参入来 *moti-mawiri-ku* ‘carry-Hon.entre-come’

**Four-verb construction** (89 tokens)

d. [Tran. + Tran. + Unerg. + Unacc.]

e.g. 尋覓上往 *tazune-motome-nobori-yuku* ‘look for-search-climb-go’

**Five-verb construction** (61 tokens)

e. [Tran. + Unerg. + Unacc. + Unacc. + Unerg.]

e.g. 飲醉留伏寝 *nomi-yoi-todomari-fushi-ineku* ‘supply-build up-wait’

**Two-verb construction** (225 tokens)

f. [Unacc. + Unacc.]

e.g. 伊傳多知弓 *ide-tatite* ‘exit-stand’; 和多利涅 *watari-die* ‘cross-exit’;

進入 *susumi-iru* ‘go forward-entre’

g. [Tran. + Tran.]

e.g. 宇知和多須 *uti-watasu* ‘hit-ferry’; 宇知微流 *uti-mwiru* ‘hit-see’

奴須美斯勢牟 *nusumi-sise-mu* ‘steal-kill.causative’

h. [Unacc. + Unerg.]

e.g. 伊提摩栖 (HON.)古 *ide-mase-kwo* ‘exit-Hon.-come’

i. [Unerg. + Unerg.]

e.g. 岐伊理袁理 *ki-iri-wori* ‘come-be.INF’

The combination of multiple verbs in the Early Nara Period appears to be arbitrarily composed: one construction can be composed of a few of different types of verbs. Each verb is independent and conveys an individual event. For instance, the first constituent renders a motion event, the second conveys a change-of-state event, the third denotes a motion event, etc. Syntactically, they all behave as the head. Morphologically, the compositionality of each constituent is rather loose. Semantically, the events rendered by multiple verbs are either successive or co-ordinate. This inspires us to deduce that multiple verb construction in the Early Nara Period is an issue of verb serialising. A closer examination of the data regarding the composition patterning, transitivity, headedness and frequency is provided in Table 2.

Table 2. Multiple verb combinations in the early Nara Period

Composition pattern	Tokens	Transitivity	Headed by
<b>Tran. + Tran. + Unacc.</b>	67	unaccusative	Equal
<b>Tran. + Tran. + Tran.</b>	99	transitive	Equal
<b>Tran. + Unacc. + Unacc.</b>	61	unaccusative	Equal
<b>Tran. + Tran. + Unerg. + Unacc.</b>	89	unaccusative	Equal
<b>Tran. + Unerg. + Unacc. + Unacc. + Unerg.</b>	61	unergative	Equal
<b>Unacc. + Unacc.</b>	79	unaccusative	Dual-beaded
<b>Unacc. + Unerg.</b>	18	unergative	V2
<b>Tran. + Tran.</b>	105	transitive	Dual-beaded
<b>Unerg. + Unerg.</b>	23	unergative	Dual-beaded

To shed more light on the complexity of these variations, the entropy should be calculated.

(25) The total amount of tokens of V-Vs with an unaccusative final constituent is:

$$N = 67+61+89+79 = 296.$$

The variations of V-Vs with an unaccusative final constituent are: [Tran. + Tran. + Unacc.] (type number: 67), [Tran. + Unacc. + Unacc.] (type number: 61), [Tran. + Tran. + Unerg. + Unacc.] (type number: 89), [Unacc. + Unacc.] (type number: 79). Thus,

(26) Regarding [Tran. + Tran. + Unacc.]

$$P_{[\text{Tran.} + \text{Tran.} + \text{Unacc.}]} = 0.23; \text{Log}_2 0.23 = -2.12; p_j \log_2 p_j : 0.23 \times -2.12 = -0.488$$

(27) Regarding [Tran. + Unacc. + Unacc.]

$$P_{[\text{Tran.} + \text{Unacc.} + \text{Unacc.}]} = 0.21; \text{Log}_2 0.21 = -2.25; p_j \log_2 p_j : 0.21 \times -2.12 = -0.473$$

(28) Regarding [Tran. + Tran. + Unerg. + Unacc.]

$$P_{[\text{Tran.} + \text{Tran.} + \text{Unerg.} + \text{Unacc.}]} = 0.30; \text{Log}_2 0.30 = -1.74; p_j \log_2 p_j : 0.30 \times -1.74 = -0.557$$

(29) Regarding [Unacc. + Unacc.]

$$P_{[\text{Unacc.} + \text{Unacc.}]} = 0.27; \text{Log}_2 0.27 = -1.89; p_j \log_2 p_j : 0.27 \times -1.89 = -0.510$$

The entropy of the  $V_{final}$  unaccusative is established by adding the above four  $p_j \log_2 p_j$  scores and dividing by -1. Thus,  $H = [(-0.488) + (-0.473) + (-0.557) + (-0.510)] \div (-1) = 2.03$ .

Likewise, the total amount of tokens for V-Vs with a transitive final constituent is  $N = 204$ . Two variations are confirmed, i.e. [Tran. + Tran. + Tran.] (type number: 99), [Tran. + Tran.] (type number: 105). The entropy of the  $V_{final}$  transitive is thus  $H = 0.99$ . The total amount of

tokens of V-Vs with an unergative final constituent is  $N = 102$ . The entropy of the  $V_{final}$  unergative would be  $H = 1.06$ . The entropy of multi-verb constructions with different final constituents is summarised in Table 3:

Table 3. Entropy of multi-verb constructions with different final constituents in the early Nara Period

$V_{final}$	N	H	Transitivity
<b>Unaccusative</b>	296	2.03	unaccusative
<b>Transitive</b>	204	0.99	transitive
<b>Unergative</b>	102	1.06	unergative

As suggested in Table 3, ‘accusative V-ended’ verb compounds are likely to be higher in entropy than ‘transitive V-ended’ verb compounds and ‘unergative V-ended’ verb compounds.

### 3.2 Multiple Verb Combinations in the Late Nara Period

Let us now move on to multiple verb combinations in the Late Nara Period. Before attempting to delve into the formation principle, perhaps it is not unsound to reiterate that the loan of Chinese characters in the Late Nara Period shifted into three patterns: (i) solely borrowing semantic meaning, (ii) solely borrowing phonological value and (iii) borrowing both semantic and phonological values. The writing of this period is known as *man'yōgana*.

A representative work is 万葉集 *Manyōshū* (AD. 759). The formation of multi-verb constructions in the Late Nara Period presents the following features:

Table 4. Multiple verb combinations in the late Nara Period

Composition	Pattern	Transitivity	Headedness	Tokens	Examples
<b>Tran. + Unacc.</b>	Verb compounding	Unacc.	V2	123	打摩 uchi-nabiku
<b>Tran. + Unacc. + Tran.</b>	SVC: successive	Tran.	Equal	23	打越見 uti-kwoye miru
<b>Tran. + Tran.</b>	SVC: successive	Tran.	Dual	210	打和多思 uti-watasu
<b>Unacc. + Unacc. + Unerg.</b>	SVC: successive	Unerg.	Equal	14	出立偲 ide-tati-sinwopi
<b>Unacc. + Unerg.</b>	SVC: successive	Unerg.	Dual	57	出波之利 ide-pasiru
<b>Unerg. + Unacc. + Unerg.</b>	SVC: successive	Unerg.	Equal	11	産礼出有 umare-ide-taru



<b>Unerg. + Unacc.</b>	SVC: successive	Unacc.	Dual	19	潜出 kaduki-idu
<b>Unacc. + Unacc.</b>	SVC: successive	Unacc.	Dual	111	入流 iri-nuru
<b>Unacc. + Unacc. + Tran.</b>	SVC: successive	Tran.	Equal	34	入出見 iri-idumiru
<b>Unacc. + Tran.</b>	SVC: successive	Tran.	Dual	22	入卷 ira-maku
<b>Unacc. + Unacc. + Unacc.</b>	SVC: successive	Unacc.	Equal	46	入通來 iri kaywopi-ko
<b>Unerg. + Unacc. + Unacc.</b>	SVC: successive	Unacc.	Equal	14	榜入來 kogi-ri-ko
<b>Unerg. + Unerg.</b>	Verb compounding	Unacc.	Dual	45	舞踊 mai-odoru 'dance-dance'
<b>Unacc. + Unacc.</b>	Verb compounding	Unacc.	V2	91	乱出 midare-iduru
<b>Unerg. + Unacc.</b>	Verb compounding	Unacc.	V2	21	步疲 aruki-tsukareru 'walk-tired'

The combining of serial verbs is reduced to maximum three verbs. A common option is to combine two constituents of the same kind, e.g. the combinations of [Transitive + Transitive] (210 tokens), [Unergative + Unergative] and [Unaccusative + Unaccusative] (161 tokens). These formations share the following similarities: (a) all the constituents behave as the head, and (b) a conjunct (而 *te*) can be inserted between the multiple verbs, as seen in 開而出 *saki-te-deru* 'open-exit'; 出而入 *ide-te-hairu* 'exit-entre'. The insertion of conjunct (而 *te*) is valid evidence that multi-verb construction is formed via verb serialising. This would

suggest that verb serialising remains the main option for multiple verb construction during the Late Nara Period. There are 11 combinatorial possibilities for building an SVC, as detailed in (30):

- (30) (a). [Tran. + Unacc. + Tran.], (b). [Tran. + Tran.], (c). [Unacc. + Unacc. + Unerg.],  
 (d). [Unacc. + Unerg.], (e). [Unerg. + Unacc. + Unerg.], (f). [Unerg. + Unacc.],  
 (g). [Unacc. + Unacc.], (h). [Unacc. + Unacc. + Tran.], (i). [Unacc. + Tran.],  
 (j). [Unacc. + Unacc. + Unacc.], (k). [Unerg. + Unacc. + Unacc.]

The entropy of the  $V_{\text{final}}$  unaccusative (verb serialising strategy) is  $H= 1.58$ , as detailed in Table 5.

Table 5. The entropy of the  $V_{\text{final}}$  unaccusative (verb serialising patterning)

Composition	Pattern	N	P	Log2	$p_j \log_2 p_j$
Unerg. + Unacc.	SVC	19	0.1	-3.32	-0.332
Unacc. + Unacc.	SVC	111	0.584210526	-0.79	-0.461526316
Unacc. + Unacc. + Unacc.	SVC	46	0.242105263	-2.06	-0.498736842
Unerg. + Unacc. + Unacc.	SVC	14	0.073684211	-3.84	-0.282947368
<b>Total</b>		190			-1.575210526

The entropy of the  $V_{\text{final}}$  unaccusative (verb compounding patterning) is  $H= 1.33$ , as detailed in Table 6.

Table 6. The entropy of the  $V_{\text{final}}$  unaccusative (verb compounding patterning)

Composition	Patterning	N	P	Log2	$p_j \log_2 p_j$
Unacc. + Unacc.	V-V	91	0.387234043	-1.36	-0.526638298
Unerg. + Unacc.	V-V	21	0.089361702	-3.47	-0.310085106
Tran. + Unacc.	V-V	123	0.523404255	-0.94	-0.492
<b>Total</b>		235			-1.328723404

The entropy of the  $V_{\text{final}}$  unaccusative (verb serialising strategy) is 1.58, which is higher than the entropy of the  $V_{\text{final}}$  unaccusative (verb compounding patterning). Given this, we cannot but contend that the formation of ‘verb serialising’ is more productive than the formation of ‘verb compounding’ in multiple verb constructions.

With the quantitative analysis presented, a further look into formation patterning will shed more light on multiple verb combinations in Old Japanese. Morpho-syntactically speaking, the 11 formations in (30) can be classified into two groups:

**Group I successive SVCs:** [Tran. + Unacc. + Tran.], [Unacc. + Unacc. + Unerg.], [Unacc. + Unerg.], [Unerg. + Unacc. + Unerg.], [Unerg. + Unacc.], [Unacc. + Unacc. + Tran.], [Unacc. + Tran.], [Unerg. + Unacc. + Unacc.]

**Group II co-ordinate SVCs:** [Tran. + Tran.], [Unacc. + Unacc.], [Unacc. + Unacc. + Unacc.]

In *successive* SVCs, multiple constituents share the agent and, syntactically, all the constituents behave as the head. The events represented by each constituent are in a successive relation. *Successive* SVCs are confirmed to have eight variations: [Tran. + Unacc. + Tran.], [Unacc. + Unacc. + Unerg.], [Unacc. + Unerg.], [Unerg. + Unacc. + Unerg.], [Unerg. + Unacc.], [Unacc. + Unacc. + Tran.], [Unacc. + Tran.], [Unerg. + Unacc. + Unacc.]. A typical illustration comes from Man’yōshū.2.182.

(31) 檀 岡 尔 飛 反 来 年  
 mayumi no woka ni tobi-kapyeri-ko ne  
 spindle tree GEN hill DAT fly-return-come III  
 ‘Birds! (when you grow up), fly back to the hill.’

(MYS.2.182)

The three constituents, i.e. *tobu-kaeru-kuru*, are syntactically equal. The first constituent *tobu* ‘fly’ renders the MANNER ‘fly’, the second constituent *kaeru* ‘come back’ denotes the MANNER ‘return’, and the third constituent *kuru* ‘come’ succeeds the previous two actions.

Another illustration comes from Man’yōshū.1.58.

(32) 安礼乃埼 榜多味 行之 棚 無 小舟.  
 are no saki kogi-tamwi-yuki-si tana na-si wo-bune  
 Are GEN cape row-turn around-go.SPST tana have-NEG boat  
 ‘That little tana-less boat coasted round the cape of Are.’

(MYS.1.58)

In (32), the complex construction is conveyed by serial verbs, i.e. *kogu* ‘row’, *tamu* ‘turn around’, and *yuku* ‘go’. The former two morphemes conflate the MANNER and the rest conveys the PATH.

It is also important to notice that, in addition to V-V-V verb serialising, another form of SVC is employed: V-V serialising. A typical illustration would be (33).

- (33) 秋芽 之 咲 而 散 去 流 花 尔 有 猿 (Note 7) 尾
- akipagwi no saki te tiri-nu ru pana ni ara-masi wo
- Fall hagi GEN bloom-GER-scatter-PERF.AND flower DAT be.CONJ
- ‘I wish it were the beautifully blooming and beautifully scattering Hagi flower.’

(MYS.2.120)

The compositionality of the multi-verb construction is loose, for 而 *te* is inserted between the constituents. It is composed of two ACTION verbs, i.e. V1 *saku* ‘bloom’ and V2 *chiru* ‘scatter’. The two morphemes have opposite meanings, but the events represented by the V1 and V2 are classified as belonging to the same categories (i.e. both of them are change-of-state events). Given this, syntactically and semantically, the two morphemes seem to receive an equal weight. Therefore, successive verb serialising is tentatively suggested.

The multi-verb constructions we have discussed so far are related to verb serialising. Multiple verbs can also be combined via verb *co-ordinating*. The database confirmed three variations: [Tran. + Tran.], [Unacc. + Unacc.], [Unacc. + Unacc. + Unacc.]. A typical example would be (34):

- (34) 皇子 乃 御門 之 荒 卷 惜 毛
- mikwo no mi kadwo no are-maku wosi mo
- prince GEN HONpalace GEN dilapidated-rough pity FOC
- ‘pity for the prince’s palace turning dilapidated.’

(MYS.2.168)

The two morphemes, i.e. *areru* ‘dilapidated’ and *maku* ‘rough’, have the same meaning. They are considered a synonymous lexicon and receive equal syntactic weight. The events represented by the V1 and V2 are classified by the same categories.

The option of verb coordinating is also seen outside the *Man’yōshū*. There is one token in *Bussokuseki-ka*:

(35) 己乃美阿止夜与吕豆比賀利乎波奈知伊太志

ko no mi- ato ya-yorodu- pikari o panati- idasi

PROXGENHON footprint myriad lights ACC send forth

毛呂毛呂 須久比 和多志 多麻波奈

moromoro sukupi watasi- tamapana

many save-INF cross-INF honorific-OPT

(Bussokuseki-ka. 1)

#### 4. Summary

This study explored how multiple verbs are combined in Old Japanese by building two databases: multi-verb construction in the Early Nara Period, and multi-verb construction in the Late the Nara Period. The findings revealed that:

(I) In the Early Nara period, the formation of multi-verb constructions is facilitated at a syntactic level and pertains to an issue of verb serialising. Semantically, multiple verbs (either two-word construction, three-word construction or four-verb construction) are assigned to a successive relation or coordinate relation. Morphologically and syntactically, each constituent is equally weighed. Nine combinatorial possibilities are confirmed: (a) [Transitive + Transitive + Unacc.]; (b) [Tran. + Tran. + Tran.]; (c) [Tran. + Unacc. + Unacc.]; (d). [Tran. + Tran. + Unerg. + Unacc.]; (e) [Tran. + Unerg. + Unacc. + Unacc. + Unerg.]; (f) [Unacc. + Unacc.]; (g) [Tran. + Tran.]; (h) [Unacc. + Unerg.]; (i). [Unerg. + Unerg.].

(II) In the Late Nara Period, SVC remains the main option for combining multiple verbs, accounting for 11 combinatorial possibilities: (a) [Tran. + Unacc. + Tran.], (b) [Tran. + Tran.], (c) [Unacc. + Unacc. + Unerg.], (d) [Unacc. + Unerg.]; (e) [Unerg. + Unacc. + Unerg.], (f) [Unerg. + Unacc.], (g) [Unacc. + Unacc.], (h) [Unacc. + Unacc. + Tran.], (i) [Unacc. + Tran.], (j) [Unacc. + Unacc. + Unacc.], (k) [Unerg. + Unacc. + Unacc.]. Crucially, in this period, grammaticalisation takes place. Grammaticalisation of unaccusative change-of-state verbs, as well as motion verbs, leads to two different lexical categories: (a) partial unaccusative change-of-state verbs turned into resultative complements, and (b) partial motion verbs grammaticalised into directional complements. Such transitions result in tighter integrity of multiple verbs, which, in turn, inspires the device of verb compounding. The following four constructions, i.e. [Tran. + Unacc.], [Unergative + Unacc.], [Unerg. + Unerg.] and [Unacc. + Unacc.], are the candidates of verb compounds.

A possible timeline for the evolution of multiple verb combinations is provided in Figure 1.

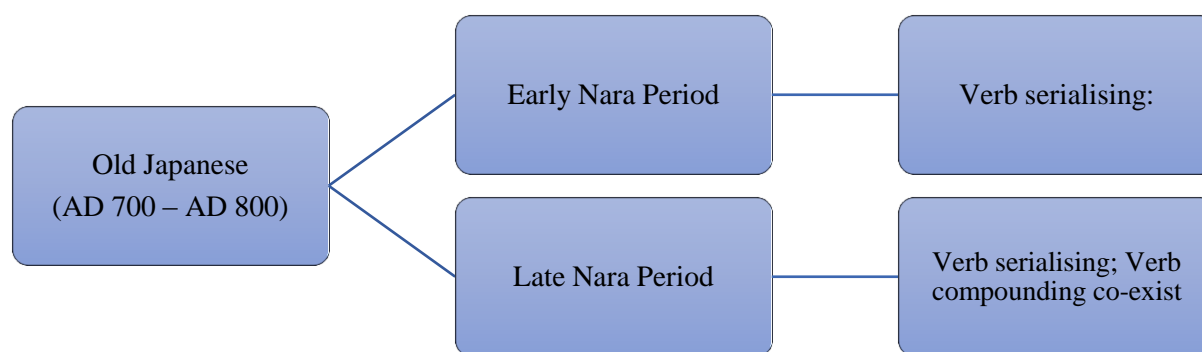


Figure 1. Timeline for the evolution of multiple verb combinations in Old Japanese

(III) A further calculation on the entropy of the  $V_{\text{final}}$  unaccusative suggests that formation via verb serialising (cf.  $H= 1.58$ ) is more productive than formation via verb compounding (cf.  $H=1.33$ ).

### Abbreviations

ACC	accusative
ACOP	adjectival copula
AND	adnominal
ADV	adverb
CONCL	conclusive
COND	conditional
CONJ	conjunctural
CONT	continuous
COP	copula
COS	change-of-state
DAT	dative
DEM	demonstrative
EMPH	emphatic
ETOP	emphatic topic
EXCL	exclamatory
FOC	focus
GER	gerund
HON	honorific

---

IMP	imperative
INF	infinitive
IJP	interjectional particle
MPST	modal past
NEG	negative
OPT	optative
PASS	passive
PART	participle
PLUR	plural
PREF	prefix
PREV	preverb
PROG	progressive
PRON	pronoun
PROV	provisional
RESP	respect
SPST	simple past
TOP	topic

### **Acknowledgement**

This paper is based on work that was supported by the *National Foundation of Social Sciences* (15CYY002) as well as the *Humanities and Social Sciences Research Programme of Zhejiang University* to Wenchao Li. The manuscript was presented at international symposium of studies of medieval and Renaissance European literature (University of Zhejiang). I would like to thank Professor Hao Tianhu, Cong Cong, Zhu Zhenyu and Jiang Wentao, and two anonymous reviewers for helpful comments.

### **References**

- Baayen, R. H. (1992). Quantitative aspects of morphological productivity. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology 1991* (pp. 109-149). Dordrecht: Kluwer.
- Baayen, R. H., & Lieber, R. (1991). Productivity and English derivation: a corpus-based study. *Linguistics*, 29, 801-843.
- Kageyama, T. (1993). *Bunpou to gokeisei* [Grammar and word formation]. Hitsuji Syobo Publishing.

Kindaichi, H. (1953). Kokugo Akusentoshi No Kenkyu ga Nani ni Yakudatuka (The role of studies on Japanese accent): *Proceedings of language and folkloristic studies, in memory of Professor Kindaichi Kyosuke's seventieth birthday*, pp. 329-354. Sanseido Publishing.

Nakamura, Y. (1971). Jodai Fukugodooshi no Kinmitsusei ni Tsuite (The compositionality of verb compounds in Old Japanese). *Memoir of Kokugakuin Junior High School*, 21, 3-47.

Nishiyama, K., & Ogawa, Y. (2014). Auxiliation, atransitivity and transitivity harmony in Japanese V-V compounds. *Journal of Interdisciplinary Information Sciences*, 20(2), 71-101.

Seki, K. (1977). *Kokugo Fukugodooshi no Kenkyu* (Japanese verb compounds). Kasama Publishing.

Shannon, C. C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379-423 (Part I) and 623-656 (Part II).

Tamaoka, K., Lim, H., & Sakai, H. (2004). Entropy and redundancy of Japanese lexical and syntactic compound verbs. *Journal of Quantitative Linguistics*, 11(3), 233-250.

Yumoto, Y. (2005). *Fukugoo dooshi to hassei dooshi no imi to toogo – Mojuru Keitairon Kara Mita Nichieigo no Dooshi Keisei*. [Lexical semantics and formation of compound verbs and derived verbs: a module morphology based analysis on verb formation in Japanese and English]. Hitsuji Publishing.

Yumoto, Y. (2008). Fukugoo Dooshi ni Okeru Koo Jitsugen – Togoteki Fukugoo to Goitekifukugoo no Sai –. [Argument realisation in verb compounds: the distinction between syntactic verb compounds and lexical verb compounds]. *Morphology and Lexicon Forum*, 4, 1-13.

Yumoto, Y. (2011). Nichieigo ni Okeru Kuoria Koozoo o Riyooshita Gokeisei. [Qualia structure-based analysis of word formation in Japanese and English]. *Research report on theoretical approach towards natural language*, 1, 91-100.

Yumoto, Y. (2012). Goiteki Fukugodooshi no Imikaishaku Saikoo. [Revisiting the semantic interpretation of lexical verb compounds]. *Research report on theoretical approach towards natural language*, 1, 89-99.

## Corpora

*The Japanese historical linguistic corpus 2.4.2* (by National Institute for Japanese Language and Linguistics): <https://chunagon.ninjal.ac.jp/>

*The Oxford Corpus of Old Japanese*: <http://vsarpj.orinst.ox.ac.uk/corpus/>



## Notes

Note 1. Chinese is a tonal language. In order to differentiate meaning, the same syllable can be pronounced with different tones.

Note 2. In Old Japanese, there are five ‘kakarijoshi’, i.e. 「ぞ」 *zo*, 「なむ」 *namu*, 「や」 *ya*, 「か」 *ka*, 「こそ」 *koso*. They appear in sentences and require different verb endings, as shown in Table 1.

Table 1. *Kakarijoshi* and their corresponding endings

Kakarijoshi	Ending
ぞ	adnominal form
なむ	
や	
か	
こそ	perfective form

Note 3. 牟禮 *mure* and 比氣 *piki* are in infinitive forms, modifying the manner of the actions, i.e. flocking and accompanied. The paths are encoded into the verb roots, 伊ぬ *inu*, which is in a conditional form.

Note 4. *tumagomwi*: a lexicalised noun.

Note 5. 足日木乃 (*asipikwi no*) is a ‘makurakotoba’ (枕詞: pillow words). Makurakotoba are often used at the beginning of a poem, introducing a noun. Most makurakotoba have no substantial meaning. Here, 足日木乃 *asipikwi no* refers to 山 *yama* (mountain).

Note 6. The analysis of verbs in this study follows Kageyama (1993): transitive verbs have an internal argument as well as an external argument, as described in (1).

(1) **Transitive verb:** [*x* ACT ( ON *y* )] CONTROL [BECOME [ *y* BE AT-*z*]], e.g. *hit, naguru*

Unergative verbs only have an external argument, cf. (2).

(2) **Unergative verb:** cause change-of-state events or motion events with volition.

[*x=y* CONTROL [ *y* BECOME [ *y* BE AT *z* ]]], e.g. *dance, cry, odoru*

Unaccusative verbs only have an internal argument, cf. (3).

(3) **Unaccusative verb:** change-of-state or motion occurs spontaneously.

[BECOME [ *y* BE AT *z* ]], e.g. *die, shinu*

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>)