Analysis and Value of *Hentai Kanbun* as Japanese

EDITH ALDRIDGE

*University of Washington*

1. Introduction

This paper proposes a generative syntactic analysis of word order patterns in *hentai kanbun* (変体漢文). *Hentai kanbun* is an early Japanese writing system for recording Japanese using Chinese characters for their semantic and morphosyntactic value rather than their phonetic value. Consequently, *hentai kanbun* texts bear surface resemblance to Classical or Middle Chinese, due to the lack of Japanese syllabic symbols for functional categories and the tendency for VPs and PPs to be head-initial, as they are in Chinese. This paper primarily considers the 8th century historical chronicle *Kojiki* (古事記). (1) shows SVO word order, which is a consistent feature of Chinese from classical to modern times. Japanese, on the other hand, has been an SOV language throughout its attested history. When the sentence is read (indicated by ‘=>’), the word order is translated into head-final Japanese order.
Adpositional phrases are likewise head-initial, with the adposition preceding its complement DP.

However, not all word orders found in the *Kojiki* are of a type attested in Chinese. In ditransitive clauses in which both internal arguments are overtly expressed, the verb appears between the two objects.

At least since Motoori (1798), there has been an awareness in the field that *hentai kanbun* texts were intended to be read in Japanese. As for the composition of the texts themselves, the general assumption seems to be that *hentai kanbun* is some sort of hybrid, containing elements of both Chinese and Japanese (Miller 1967; Minegishi 1986; Feng 1995; Uchida 1995; Rabinovitch 1996). Others treat *hentai kanbun* texts as primarily Japanese (Tsukishima 1963; Nishimiya 1993; Yamaguchi 1995; Sema 1999). However, systematic analysis of the word order in these texts has by and large remained elusive, Nakagawa (1995) and Aldridge (2001) being the only exceptions that I am aware of.

In this paper, I propose a syntactic analysis which accounts for both the Chinese type orders as in (1) and (2) and unexpected word order types like (3). The word order derivation I propose consists of a two-step encoding process from surface Japanese word order to the word order of the text. The first step simply reverses the order of head and complement within a phrase,

---

1 Gloses conform to the Leipzig MPI rules with the following exceptions: ADN (‘adnominal’), C (‘complementizer’), CNJ (‘conjunction’), HON (‘honorific’), MOD (‘modal’), STAT (‘stative’).

2 Transcriptions for Old Japanese are based on reconstructions in Frellesvig & Whitman (2008).
turning head-final order in Japanese into the head-initial order in the text. This process suffices to derive cases like (1) and (2). For less straightforward cases, I propose the addition of a second step. After the head-complement order has been set, part of the complement can be raised into the specifier of the head. I further suggest that the availability of this second step to the author of the text relies on the existence of an object raising transformation which was commonly employed in natural Japanese of the time. Consequently, we can conclude not only that *hentai kanbun* texts are Japanese (and not Chinese), but also that their apparent deviation from both languages offers clues to the syntax of natural Japanese of the time.

2. Reordering

The orders in (1) and (2) can be straightforwardly accounted for by reversing head-final Japanese order to produce head-initial Chinese type word order.

(4) **Hentai Kanbun Word Order Derivation (general)**

Hentai kanbun word order is derived from surface Japanese by reversing head and complement order within every XP whose head position in Chinese differs from Japanese.

In the case of a verb and its complement direct object like in (1), the verb is reordered from its Japanese position following the object to a position immediately to the left of the object. The arrow in this example indicates the order in the text.

\[
\begin{align*}
\text{VP} & \quad \text{[V'} [soko=no \text{ awona}=\text{wo tumu}] \\
& \quad \text{there=GEN vegetable=ACC pick} \\
\Rightarrow & \quad \text{採 其地之青菜} \\
& \quad \text{pick [DEM place GEN vegetable]} \\
& \quad \text{‘pick the vegetables of that place’}
\end{align*}
\]

The same process applies to the PP in (2), with the result that a postposition in Japanese is reordered as a preposition in the text.

Ditransitive clauses, however, are not quite so easily dispensed with. As shown in (3) above, the verb appears between its two internal arguments in ditransitive clauses. Additional examples are given in (6). Regardless of the relative order or the dative and accusative objects, the verb surfaces between the two internal arguments.

---

3 The reversal only applies to XPs whose head position is different in Japanese and Chinese. Hence, nominal categories, which are head-final in both Japanese and Chinese, do not undergo the word order reversal.
This order is unexpected in Old, Middle, or Modern Chinese. The corresponding Chinese ditransitives are schematized in (7). In both the dative construction (7a) and the double object construction (7b), the verb precedes both of its internal arguments.

(7) Chinese

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>V DP PP</td>
<td>(give something to someone)</td>
</tr>
<tr>
<td>b.</td>
<td>V DP DP</td>
<td>(give someone something)</td>
</tr>
</tbody>
</table>

The two types of ditransitive can be observed in the following Classical Chinese example. The first of the two conjoined clauses employs a dative construction and packages the goal as a PP. The second clause is a double object construction in which the goal appears as a DP in immediate post-verbal position.

(8) Tianzi neng [jian ren yu tian]  
ruler can recommend person to heaven  

bu neng shi tian [yu zhi tianxia]  
not can make heaven give 3.OBJ world  

‘The ruler can recommend someone to heaven, but (he) cannot make heaven give him the world.’   

(Mencius 9)

Aldridge (2001) accounts for the medial positioning of the verb in Kojiki ditransitives by claiming that word order in this text reflects underlying head-initial word order in which the verb surfaces in its base position in VP. Following Miyagawa (1997), Aldridge assumes that the two objects can be base merged in either order, one as the verb’s complement and the other in the specifier position.
We can avoid the controversial position of assuming underlying head-initial order for Japanese and still easily derive the verb-medial order with the linearization rule in (4) by switching the order of head and complement within VP. However, this approach still relies on the assumption that the input to the reordering process is underlying Japanese structure, in which the verb is in V and has not moved to v or higher. This is a problematic assumption, because there is in fact evidence in the text that verb-movement has taken place. Specifically, verbs appear fully inflected in the text. (10) shows an example of the modal suffix –mu, which is written with the Classical Chinese character expressing future modality 將.

(10) 將 入 海 時 (Kojiki, Keikō)
MOD enter sea time
⇒ umi=ni iri-tamapa-mu-to-su-ru toki=ni
sea=DAT enter-HON-MOD-C-do-ADN time=DAT
‘when (she) was about to enter the sea’

(11) shows an example of this modal in Classical Chinese. As in the Kojiki example in (10), the modal immediately precedes the VP.

(11) 吾 將 問 之。 (Analects 1)
Wu jiang wen zhi.
I MOD ask 3.OBJ
‘I will ask him.’

However, in the Kojiki, this modal always appears in immediate preverbal position and is never separated from the verb.

(12) 自 出雲 將 上 坐 倭國 而
from Izumo MOD go.HON Yamato CONJ
⇒ Izumo=from Yamato=GEN country=DAT
nobori-masa-mu to si-te (Kojiki, Ōkuninushi)
go-HON-MOD COMP do-CONJ
‘As (he) was about to go to Yamato from Izumo, …’

Note in particular that the modal follows a PP in (12). (13) shows an early Middle Chinese example with a similar type of PP. In both Middle and Classical Chinese, this modal was a free morpheme and occupied a specific
syntactic position high within the TP structure (Wei 1999 and Aldridge 2010). Crucially, for the purposes at hand, this modal precedes a PP of the type seen in (12).

(13) 公  將  自  東  征  備。
    Gong  jiang  zi  dong  zheng  Bei.  
    lord  MOD  from  east  attack  Liu  Bei
    ‘Our lord was about to attack Liu Bei from the east.’
    (Sanguozhi 1.1)

The contrast between (12) and (13) can be accounted for if we assume that the verb appears in the **hentai kanbun** texts with inflections like the modal suffix attached to it. This entails that the verb has moved at least high enough in the structure where it is visible to the probe which can value inflectional features. I assume, then, that verbs in the Kojiki have moved to v, where they are located in the edge of the vP phase and are therefore accessible to probes in the inflectional domain immediately dominating vP.

However, this presents a problem for deriving word order in ditransitive construction by means of the encoding process in (4). On the assumption that both internal arguments in a ditransitive construction reside within VP, it cannot be the case that the word order in the text is the result of merely reordering the verbal complex in v around its complement VP. This is because reordering V+v around its complement VP would result in the verb preceding both internal arguments. This problem is eliminated, however, if one of the objects moves into the edge of vP, as I show in the next section.

3. Raising

Yanagida (2006) and Yanagida and Whitman (2009) argue that definite or specific objects and raise obligatorily to the edge of vP in Old Japanese. A bare object with a nonspecific interpretation appears in immediate preverbal position. Examining (14), it can be seen that the object follows the genitive subject in (14a), while the definite object raises to a position preceding the genitive subject in (14b). Yanagida and Whitman (2009) point out that unraised objects are always bare X0-level categories and never appear with the case-marker wo. They propose that these objects are incorporated to the verb. Dislocated objects can be phrasal and generally take the case-marker wo. These are analyzed as undergoing object shift to the edge of vP.

(14) a. 佐欲比賣能 故何 比列 布利斯 夜麻
    Sayopimye=no  kwo=ga  pire  puri-si  yama
    Sayohime=GEN  child=GEN  scarf  wave-PST.ADN  hill
    ‘the hill where the girl Sayohime waved her scarf’ (MYS 5:868)
Given the above considerations, ditransitives in the *Kojiki* in which the accusative object precedes the dative can easily be incorporated into the analysis developed so far. Given the evidence that Japanese surface strings provide input to the encoding process, I assume that the verbal complex resides in $v$. The encoding process in (4) will result in $V+v$ being linearized to the left of $VP$. In an example like (15a), which has a phrasal direct object, this object will have undergone object shift to the edge of $vP$. Since it is located in a specifier of $vP$, its position in the text will be just to the left of the reordered verb.

(15) a. 多 禄 給 其 老 女 (*Kojiki*, Ōryaku)  
many thing $\text{give}$ DEM old woman  
=> Amatano $\text{mono}=\text{wo}$ $\text{sono}$ $\text{omina}=\text{ni}$ tamapi-te  
many thing=$\text{ACC}$ DEM old.woman=$\text{DAT}$ give-CNJ  
‘He gave many things to the old woman, and….’

b.  
```
  vP
 /\  
v'  
 /   \ 
 tSubj  V' 
 /     \  
 V+v  V''  
 /         \  
 tDP  V  <V>  DP
```

Another possible source for verb-medial ditransitive order in the text is incorporation of bare objects. This is a possible derivation of the following, in which the preverbal object appears to be just a bare noun. If we analyze this as incorporation, the preverbal position is still accounted for, since the object will move with the verb to $v$.

(16) a. 名 賜 曙 立 王 (*Kojiki*, Suinin)  
name $\text{give}$ Aketatu prince

---

4 The character in the text is slightly different, composed of a mouth and a blade.
Given the restrictions on object position in Old Japanese, word order in ditransitive clauses in the *Kojiki* falls out naturally from the encoding process in (4). However, there are some anomalous orders which are not derived so straightforwardly. There are a number of examples in the text in which a verb or adposition appears within the complement it selects. In (17), an adposition appears to be inside of its complement DP, following a demonstrative in (17a) and a possessor in (17b).

(17) a. 是 於 河 下 如 青 葉 山

| [DEM on river bank] be.like green leaf hill |
| TOP appear hill not.be hill |

=> [kono kapa shimwo ni] awo ba=no yama=no this river bank on green leaf=GEN hill=GEN goto-ki]=pa yama=to miye-te yama=ni ara-zu. be.like-ADN=TOP hill=C seem=CNJ hill=DAT be-NEG

‘What looks like a hill of green leaves on this river bank is not a hill.’

b. 所殺 迦具土 神 之 於 頭

| [[[PASS.kill K god] GEN] from head] |
| PASS.make god name Masakayamatumi god |

masaka yamatumi=no kamwi.
M=GEN god
‘The name of the god born [from the head of the slain Kagutsuchi god] was Masakayamatumi.’ (Kojiki, Izanagi)

What I propose here is that there was a second step in the encoding process for the Kojiki which borrowed from the object shift transformation. Specifically, in addition to the reversal of head and complement order in the relevant XPs, part of the complement in the XP could be raised to the specifier of that XP.

(18) Hentai Kanbun Word Order Derivation (Kojiki only)
Word order in the Kojiki is derived from surface Japanese by:

1. Reversing head and complement order within every XP whose head position in Chinese differs from Japanese.
2. Optionally raising a specifier or adjunct in the complement to the specifier of the head.

The anomalous orders seen in (17) can be accounted for with this two-step process, as shown in (19). The order of P and DP is reversed. Then the specifier or adjunct in the complement DP is moved into the specifier of the PP. This will be the demonstrative in (17a) and the possessor in (17b).

(19)

This proposal goes beyond Nakagawa’s (1995) analysis of word order in the Kojiki. Nakagawa accounts for verb-medial order in ditransitives by claiming that verbs are placed between the two objects in order to distinguish them. However, this approach has nothing to say about cases like (17), in which adpositions appear inside their complements.

An alternative analysis of the placement of the adpositions in (17) might be to claim that the P is a bound form which simply reorders around the noun it attaches to. However, this is not a possible explanation for the examples in (20). In these examples, we see the verb in a higher clause appearing inside a relative clause modifying the complement of the higher
The relative clause in (20a) contains a coordinate structure. The first of these conjuncts appears before the matrix verb. In (20b), the subject of a relative clause precedes the matrix verb.

(20) a. 作 篦 有 取 魚 人。 (Kojiki, Jimmu)

[[set trap exist take fish] person]  
=> [[Upede=wo puse-te, uwo=wo tor-u]  
trap=ACC make-CNJ fish=ACC take-ADN  
pito]  ari-ki.  
person exist-PST

‘There was a person laying fish traps and fishing.’

b. 天 下 者 汝 非 應 知 国。

world TOP [[you not.be should govern] realm]  
=> Ame=no sita=pa [[imasi=no sirusu  
Heaven=GEN below=TOP you=GEN govern  
be-ki] kuni]=ni ara-zu. (Kojiki, Chūai)  
should-ADM realm=COP be-NEG

‘The world below is not a realm which you should govern.’

The placement of these verbs cannot be accounted for in terms of reordering within an X₀-level category, since the reordering does not target morphemes within a single word. These orders can, however, be accounted for on the current proposal if we assume that the preverbal constituent can be raised into the edge of the vP where the matrix verb resides. First note that relative clauses appear in the text preceding the head they modify. Since relatives in both Chinese and Japanese are prenominal in surface order, the reordering process will not apply in deriving the order in the text. I further assume the raising analysis of relative clauses, as proposed and developed by (Kayne 1994, Bianchi 1999, Bhatt 2002, and others). (21) is Honda’s (2002) proposal for modern Japanese. The head nominal is located (base merged for Honda, but raised for Kayne) in a specifier in the CP layer of the clause. The TP portion of the relative clause raises to [Spec, DP], with the result that the head nominal follows it in surface order.

(21) [DP [TP Taro-ga tabeta proj] [D [CP ringo t₁P]]]

Taro-Nom ate apple

‘the apple that Tarō ate’

The raising approach to relative clause structure allows a straightforward analysis of the aberrant word orders in the Kojiki. In (20a), the embedded TP, which is a coordinate structure, has been raised to [Spec, DP]. The specifier within this constituent can subsequently be raised to [Spec, vP] in the matrix clause, according to the second step in the Kojiki word order derivation process in (18).
In (20b), it is the embedded subject which raises to the edge of matrix vP. Admittedly, these movements violate the ban on extraction from moved categories proposed by Wexler and Culicover (1980), Nunes and Uriagereka (2000), and others. However, note that in both cases, it is the specifier of a specifier which moves, which has been argued to take place in cases of possessor raising (Szabolcsi 1984, Landau 1999, and others). I assume, then, that this type of movement is an option made available by Universal Grammar.

Another crucial assumption I make is that the raising step in the hentai kanbun encoding process was the result of co-opting an existing transformation in spoken Japanese of the time. In the next section, I show that the loss of that transformation in the natural language correlates with the unavailability of the raising process in hentai kanbun word order derivation.

4. Loss of Raising

By Early Middle Japanese of the 10th century, specific objects were no longer required to move to [Spec, vP]. In (23), an object marked with the case-marker wo follows a subject marked with ga.

(23) Kaguyapime tepu  

\[ \text{Kaguyahime be.called big villain=GEN bitch=GEN} \]
\[ \text{hito=wo korosa-mu to suru nari-keri.} \]
\[ \text{person=ACC kill-MOD C do COP-PST} \]

‘That bitch of a villain Kaguyahime was trying to kill people.’

(Taketori)
I argue in this section that this change in the natural language is also reflected in the word order of the 10th century *sentai kambun* text *Shōmonki* 「将門記」. At first glance, it appears that the *Shōmonki* is not a fully *sentai kambun* text. It seems to include some of the anomalous word order types found in the *Kojiki*, in which a PP or dative constituent follows the verb but a direct object precedes the verb. This is reminiscent of the word order of ditransitives in the *Kojiki*, in which the verb never precedes more than one internal argument.

(24) 具 由 開 於 京都。 (Shōmonki)

=> Tubusani yosi=wo Kyōto=ni kiku.

‘(He) heard all the news in Kyōto.’

On the other hand, there are also orders like (25) which appear to follow the Chinese model. In other words, the verb precedes both the direct object and the PP.

(25) 玄明 試 閲 此 由 於 将門。 (Shōmonki)

=> Paruaki kokoromin kono yosi=wo Masakado=ni

kiku.

‘Haruaki tentatively put this matter to Masakado.’

This apparent inconsistency is easily resolved, however, in light of the word order change which has taken place from Old to Early Middle Japanese. Specific objects were no longer required to undergo object shift. Likewise, bare nonspecific objects were presumably also not required to incorporate to the verb. In short, the positioning of objects was much freer in Middle Japanese, as it is in Modern Japanese. I assume then that the preverbal object in (24) has achieved its position by scrambling to the edge of *vP*. Thus, when *V+v* is reordered to the left of *VP*, the verb will follow an object which has been scrambled. But it will precede a constituent which remains in *VP*. Note further that the difference between DP-*V*-PP and *V*-DP-PP orders in the *Shōmonki* cannot be due to definiteness or specificity. The object in both (24) and (25) is definite and refers to a situation described in the immediately preceding context.

I would further like to suggest here that the loss of object shift in the spoken language correlates with the loss of the raising step in *sentai kambun* word order derivation. Interestingly, I have found no examples of argument taking heads appearing inside their complements. For example, the
adposition in a PP always precedes its entire complement DP. This suggests that the raising step is not applied in deriving word order in the Shōmonki text.

Another asymmetry between the Kojiki and the Shōmonki is in passive constructions. In the Kojiki, agents in passive constructions always precede the verb.

(26) 汝 者 我 見欺。 (Kojiki, Ōkuninushi)
you TOP me PASS.deceive
=> Na=pa ware=ni azamuka-ye-tu.
you=TOP me=DAT deceive-PASS-PRF
‘You have been deceived by me.’

Classical and Middle Chinese also formed passives with the morpheme 見 jian. But in Chinese, the agent always follows verb.

(27) 吾 長 見 笑 於 大方 之 家。
Wu chang jian xiao yu dafang zhi jia.
I MOD PASS laugh by enlightened GEN person
‘I would have been laughed at by an enlightened person.’
(Chuangzi, Qiushui)

In the Shōmonki, the passive agent also follows the verb.

(28) 將門 被 摺 度度 之 敵
Masakado PASS [break [many GEN enemy]]
=> Masakado=pa tabi-tabi=no teki=ni kujik-are
Masakado=TOP many=GEN enemy=DAT break-PASS.CNJ
‘Masakado was broken by enemies on many occasions….’
(Shōmonki)

The difference between the Kojiki and Shōmonki derivations can be accounted for as follows. In the Kojiki, the passive agent is raised into the vP layer, as per the second step in the encoding process in (18). In the Shōmonki, on the other hand, this raising does not take place and the PP agent remains in VP. In the text, the PP will precede the verb in the Kojiki but follow the verb in the Shōmonki.

(29)a. [TP DP [vP PP V+V [vp tP [vP tV tDP ]]]] (Kojiki)

b. [TP DP [vP V+V [vp PP [vP tV tDP ]]]] (Shōmonki)

In short, word order in the Shōmonki text is derived only by head-final to head-initial linearization. There is no subsequent step of raising.
5. Conclusion

In this paper, I have proposed a syntactic analysis of how *hentai kanbun* word order is encoded from Japanese. I have argued for a single head-complement linearization process which is common to both the 8th century *Kojiki* and the 10th century *Shōmonki*. I have argued for an additional step for the *Kojiki* which involves raising. I have further suggested that the raising step was co-opted by the author of the text from a raising transformation in his native Japanese. If this analysis is on the right track, it indicates not only that *hentai kanbun* texts can be used in an analysis of Japanese but that the rules which derive the peculiar characteristics of their word order also provide a window into syntactic processes of the language of the time.

**Primary Sources (Japanese)**


**Primary Source (Chinese)**

*Hanji Dianzi Wenxian* [Electronic Corpus of Chinese Texts]

http://hanji.sinica.edu.tw/index.html?

Institute of Linguistics, Academia Sinica, Taipei, Taiwan

**References**


