AN OPTIMALITY-THEORETIC ALTERNATIVE

TO THE APPARENT *WH*-MOVEMENT IN OLD JAPANESE

Chiharu Uda Kikuta

Department of English, Doshisha University, Japan

Abstract

This paper claims that the word order restriction observed in Old Japanese (OJ) does not indicate that OJ was a *wh*-movement language, counter to Watanabe's (2002) Minimalist analysis. The apparent *wh*-movement effect is epiphenomenal of the interplay of several constraints sensitive to the profiles of case morphemes. More specifically, the word order restriction reflects a kind of mismatch where the nominal case frame is imposed upon a clause. This curious OJ word order restriction and its loss in Middle Japanese (MJ) simply reflect the gradual and dynamic development of the morphological case system, rather than a change in the *wh*-parameter. The change in the profile of case particles is partly lexical, but it has systematic consequences on the surface syntactic structure. With the mixed category analysis, adopted from Malouf (2000), and Boersma-type OT allowing optionality, the proposed analysis gives a comprehensive picture of the diachronic facts of the language.

1. Introduction[*]

The purpose of this paper is to present an Optimality-Theoretic LFG alternative to the *wh*-movement analysis of the word order restriction observed in Old Japanese. Diachronic aspects of Japanese have recently aroused strong interest in theoretical linguistics. Among the most remarkable works is Watanabe's (2002) Minimalist analysis, which claims that OJ was in fact a *wh*-movement language. Since Japanese has always been a head-final, scrambling language, it was never expected that Japanese had overt *wh*-movement. And yet, the clean and powerful UG-based account had a tremendous amount of impact, and the *wh*-movement analysis very quickly gained ground among Japanese theoretical linguists.

However, I would like to argue that the *wh*-movement analysis is wrong, and that the apparent *wh*-movement effect is epiphenomenal of the interplay of several constraints sensitive to the profiles of case morphemes. With the mixed category analysis, adopted from Malouf (2000), and Boersma-type OT allowing optionality, the word order restriction and its loss can be obtained quite naturally, and a more comprehensive analysis is possible. The wh-movement analysis claims that the Japanese language underwent a change in syntactic parameter setting between OJ and MJ. I claim instead that what changed between the two stages is not the fundamental "syntactic" property, but the system of morphological case markers.

This paper is organized in the following way. Section 2 describes the facts that motivated the *wh*-movement analysis, as well as the shortcomings of the analysis. Section 3 calls attention to other facets of OJ structure, crucial to my proposal. After all the relevant facts are laid out, I will turn to my analysis in Section 4. Section 5 concludes the paper.

---

2. The wh-movement analysis and its problems
2.1. Word order restriction in OJ and the wh-movement analysis

Watanabe's claim (2002) that OJ had overt wh-movement is based on the word order restriction in OJ, which disappeared in early Middle Japanese (MJ), around 9-10C(entury). By carefully examining data in *Man'yosyu*, which is written in OJ,[1] Nomura (1993) and Sasaki (1992) observe that the order among the interrogative particle [ka] (and [ka-mo]), the topic marker [wa], and the subject marker [no/ga] was not free. The schematic description in (1) shows the pattern, and the number on the right end indicates the number of attested data in *Man'yosyu*. The apparent rule is that [ka] almost always preceded [no/ga], while [wa] almost always preceded [ka]; the reverse is exceptional:

(1)   I.   Nominative subject: XP [ka] . . . Subj [no/ga] . . .   approximately 90
                                 Subj [no/ga] . . . XP [ka] . . .   4 (or 5)
      II. Topic: XP [ka] . . . XP [wa] . . .   2 (or 3)
                                 XP [wa] . . . XP [ka] . . .   approximately 50

From this, one can infer the possible order of the three constituents is as in (2), when all the three appear in a single sentence:[2]

(2)   XP [topic: wa]. . . XP [ka] . . . NP [subj: no/ga]

Based on (2), Watanabe (2002) proposes the clause structure with split C system in (3), which follows the insight of Rizzi (1997), and claims that XP[ka] is obligatorily moved out of IP to the Spec of FocP in OJ, as an instance of wh-movement. This is triggered, he claims, by the [-Interpretable] feature of [ka]:

(3)   [TopP Spec Top [FocP Spec Foc [IP Subj VP I]]]

Notice at this point that this wh-movement is in fact Focus-movement, and the presence of a wh-phrase is not relevant at all. *Wh-movement* is a misleading term, since the trigger is the interrogative focus particle [ka], which attaches to either a wh- or a non-wh-word (or phrase, or clause.) When it is attached to a non-wh-expression, it turns the sentence into a polar question, and the constituent with [ka] indicates the Focus of interrogation. When it is attached to a

---

[1] *Man'yosyu* (c760), a collection of over 4000 verses, is the oldest written record of Japanese. The literature in Japan predating *Man'yosyu* is written in Chinese. *Man'yosyu* employs a unique writing system, *Man'yo-gana,* which marks the inception of the Japanese writing system based on the Chinese one. More specifically, *Man'yo-gana* is a set of Chinese characters used as syllabaries (i.e., only as phonological symbols). *Man'yo-gana* is further simplified and modified to derive *hira-gana,* the system of Japanese syllabary used today.

[2] There is only one actual data in *Man'yosyu* representing this very order, but it does not necessarily question the plausibility of (2). The scarcity could partly be because the majority of the data are short verses, and also because the topic [wa] is often a topicalized subject NP.

wh-expression, its function is mostly redundant.

Be this as it may, Watanabe's (2002) analysis had a special appeal because the interrogative particle [ka] has long attracted attention in Japanese philology for a different reason. Namely, when [ka], as well as other focus-marking particles such as polar interrogative [ya] and non-interrogative [so], appear somewhere in a sentence, the predicate takes not the conclusive (or finite) form (*Syusi-kei*) but takes what is called the attributive form (*Rentai-kei*). This "agreement" phenomenon is called *kakarimusubi*, which is virtually the longest-standing conundrum of Japanese grammar.

Against this, Watanabe (2002) presents a very simple answer that *kakarimusubi* is a sign of wh-agreement, which indicates that the IP has a wh-trace in it. So there is really nothing special about *kakarimusubi* if one takes UG into consideration. Incidentally, the restriction in (1)-(2) applies to all the focus particles, including [ya] and [so] (Nomura 2002). Thus the wh-movement analysis simultaneously gives a clever UG-based solution both to the word order restriction and to *kakarimusubi*. For this reason, it is taken almost like a new "discovery" of the unknown "fact" of the Japanese language.

2.2.   Problems with the wh-movement analysis

However, it is indeed doubtful if the *wh*-movement analysis is motivated. For one thing, the fronting of a focus phrase does not necessarily require *wh*-movement in Japanese, given that virtually any type of phrase can be fronted by scrambling in the language. Even in Modern Japanese, which is not a wh-movement language in the usual sense, the focused phrase (of various forms) is often fronted for functional reasons, without a [-Interpretable] feature.

Another piece of evidence that obligatory fronting may not be an indication of syntactic movement comes from the cases of base-generated "fronting," like the ones in (4)-(5). Namely, in (4a) and (5a), a whole subordinate (causal) clause with particle [ka(-mo)] precedes the matrix clause. The [ka] clause is supposed to be moved out of the matrix clause. However, the same order, causal clause preceding the matrix clause, is found in (4b) and (5b), which have no focus particle [ka]. This order is actually the norm, since Japanese is a head-final language after all. If the order in (4a) and (5a) were an instance of wh-movement, what was behind the analogous order of the subordinate clause in (4b) and (5b)?

(4) a. [Nubatamano yo -o naga-mi <u>ka</u>-mo] waga seko -ga yume-ni-yume-ni-si
        [(epithet)   night -o long-because Q]   my   husband nom   dream-dream-in
        mie-kaheru-ramu
        see-return[attri.]
        'Is it because the night is long that my husband keeps coming back in my dream?'
    b. [Akiyama -no momizi -o sige-mi] matohinuru imo -o motomemu yamadi
        [autumn mountain -gen maple leaf -o thick-because] lost wife -acc search   path
        sirazumo
        do-not-know
        'Because maple leaves grow thick in the mountain, I cannot see the path to search for my wife who is lost.

(5) a. [Waga seko-ni mataha   awa-zi-kato omohe-ba-ka]   kesano wakare-no subenakarituru
        my wife-dat   again see-not-C   think-cond-Q   this morning parting-gen sad[attri.]
        'Is it because I know I may not see my wife again that the parting this morning is particularly
        sad?'
   b. [Yama-tooki   miyako-nisi-are-ba] saozika-no   tuma yobu koe wa tomosikumo aru-ka
        mountain-far   town-loc-be-cond. deer-gen   wife call voice top   scarece-be-Q
        'Is it because I'm in town far away from the mountain that the deer's calling his mate can
        scarcely be heard?'

        One may argue that (4a) and (5a) do involve (string-vacuous) wh-movement because they
take the attributive form, showing wh-agreement, while (4b) and (5b) do not.   However, this
argument is circluar, because it is based on the hypothetical claim that the attributive form is a sign
of wh-agreement.   In fact, this claim crucially draws on the analysis that *kakarimusubi* involves
wh-movement, without external evidence.   The fact is that the attributive form is not limited to the
context of *kakarimusubi*; it is the predicate form of a nominal clause and noun modificational clause
in general.[3]   It was also common in ancient Japanese that the nominal clause occurred
independently, expressing some kind of non-assertiveness.[4]   Most of the researchers in Japanese
philology have suggested that the association between the focus particle and the attributive form in
kakarimusubi has to do with semantics such as information structure rather than syntax (cf. Nomura
1995, Handou 2003).   That is, the attributive form is not syntactically triggered by the presence of
the focus particle (i.e., showing wh-agreement).[5]   If the particle indicates the focus of question (or
assertion), the rest conveys the de-focused (presupposed) part of the proposition.   The attributive
form of the predicate is used to reflect this informational structure.[6]   This line of explanation seems
particularly feasible, for instance, in view of the example in (6).

(6)   Tahagoto-<u>ka</u>, oyodore-<u>ka</u>,   komorikuno Hatuse-no   yama-ni   komori-seri-to-ihu.
       insane-word-Q, false-word-Q,   (epithet)   Hatuse-no   mountain-loc   hide-do-C-say[attri.]

---

[3] This is why it is called the "attributive" form.
[4] The decline of *kakarimusubi* during 14-15C coincides with the loss of the morphological distinction
between the conclusive form and the attributive form.   Interestingly, it is not the diffusion of the
conclusive form but of the attributive form that resulted in the unification of the two conjugational forms.
[5] This, of course, is largely because Japanese philology does not assume syntactic operations analogous to
the ones in the generative tradition.   And yet, it is significant that a semantic motivation is available for
kakarimusubi.
[6] Among Japanese philologists, Ohno (1993) proposes the Inversion Hypothesis, by which *kakarimusubi*
is considered as a type of displacement.   More specifically, the focus [ka] phrase is the fronted predicate,
and the following attributive clause is the clausal subject (or topic), from which the topic marker [wa],
which ought to occur at the end of the clause, is dropped for some reason.   This is a kind of
"movement," to be sure, but it is distinct from the wh-movement in the technical sense.   Importantly, (i)
the example (6), which is problematic for the wh-movement analysis, is perfectly fine for the Inversion
Hypothesis, and   (ii) the Inversion Hypothesis crucially assumes that the attributive clause is the inverted
subject (topic), which is a nominal clause by definition.   The latter point is in conformity with my
assumption.   Therefore, although the present analysis does not assume the Inversion Hypothesis, it is
actually compatible with it.

'(lit) A joke? Or a lie?   They say that he has hidden himself in Mt. Hatuse.'

Here, in (6), the [ka] phrases seem to be independent, fragmental phrases.   It is very unlikely that the matrix clause includes a wh-trace for them, although the predicate takes the attributive form.

Given the limitation of OJ data, such syntactic tests as unbounded dependency and crossover are not available,[7] and the only solid motivation for the analysis is that the word order pattern in (2) is (almost) obligatory.   However, being obligatory is not enough because the word order pattern may turn out to be obligatory when other patterns are all blocked for some other reason, which I believe is really the case.

Another serious shortcoming of the wh-movement analysis is found in its implication towards the history of *kakarimusubi*.   If Watanabe's wh-movement analysis were correct, we would expect that when the word order restriction is lost, *kakarimusubi* (or wh-agreement) disappears around the same time.   However, this is not the case.   To be sure, when the word order restriction disappears in MJ, the genuine interrogative function of [ka] is rapidly lost, becoming confined to rhetorical questions.   More and more wh-words occur without [ka], and the particle [ka] no longer triggers *kakarimusubi*.[8]   This is taken to lend support to his analysis, since the loss of *wh*-movement (word order restriction) coincides with the decline of the interrogative particle [ka], suggesting the loss of the effect of [-Interpretable] feature.

However, such change in the behavior of [ka] may very well be lexical, for the above scenario does not seem to cover the whole class of *kakarimusubi*.   For one thing, it is an established fact that the *kakarimusubi* completely disappeared around 14-15C, while the word order restriction was lifted around 9-10C.   Even after the decline of [ka], *kakarimusubi* apparently flourished in MJ with relatively free word order and with different membership: i.e., [namu], [so], [ya], and bare *wh*-word.[9]

(7) *Kakarimusubi* (with attributive form) in OJ and MJ:

	Word Order	Trigger Membership
OJ (-9C):	all observe restriction (1)	[ka], [ya], [so]
MJ (9-14C):	(more or less) all free	[ya], [so], [namu], bare wh-word

Thus, although the OJ word order restriction (and its loss in MJ) covers all focus phrases as a class,

---

[7] No one has the native intuition of OJ, of course.   Watanabe (2003) claims that there is one instance of an unbounded dependency.   However, as Tonoike (2003) contends, the structure of the data is not clear. More importantly, given that even the word order restriction in (1) allows for a few exceptions, the sole instance will not be reliable evidence.   The scarcity may rather indicate its exceptional status.

[8] The focus particle [ka] can occur in two positions: clause-internally and clause-finally.   The word order restriction concerns the clause-internal ones, although the clause-final ones outnumbered the clause-internal ones in the actual data.   During MJ, the clause-internal [ka] with wh-word is replaced by bare wh-word, and the one with non-wh-word is replaced by [ya]. Thus, [ka] does not participate in *kakarimusubi* in MJ.   On the other hand, the clause-final [ka] remains as the interrogative marker until today.

[9] *Kakarimusubi* certainly did not start all at once.   It is plausible that [ka] preceded other focus particles in diachronic change, and [namu] is evidently a late comer.   But the difference of 500 years reduces the credibility.

the proposed scenario fits only the history of [ka]. Against this, Watanabe (2002) suggests that only the *kakarimusubi* in OJ is the "real" syntactic one, while the one in MJ is merely a stylistic imitation, which somehow lasted very long. This distinction, however, seems arbitrary, with no support outside the theory.[10]

In sum, although the wh-movement analysis has a strong appeal in its clarity, the actual data is much more complex and reveals serious shortcomings of the analysis.[11]

3.   Word order restrictions in OJ and Case Morphology

Now, what is really behind the word order restriction in OJ?   A key I believe lies in the attributive form of the predicate.

There are good reasons to believe that the attributive form is a nominal/verbal, mixed category.[12]   For one thing, the attributive predicate can directly head a nominal clause without any nominalizing particle.   Secondly, [no] and [ga] are in fact markers of the genitive as well of the subject (nominative).   Besides, curiously, [no] and [ga] mark subjects only in a clause with the attributive form.   Otherwise, the subject of the clause with conclusive (or finite) form is marked either with topic marker *wa* or without any marker at all.   Markers [no] and [ga] are not even a choice.   This unique distribution of [no] and [ga] has been traditionally ascribed to the nominal nature of the attributive form.[13]

---

[10] If the *kakarimusubi* with free word order were merely a stylistic issue, imitated and maintained for 500 years, the one with fixed word order could also be a matter of style.   It sounds very arbitrary to say that the one which conforms to the "putative" UG prediction is syntactic while the one which does not is stylistic.

[11] Another piece of historical evidence Watanabe provides is that the Internally-Headed Relative Clause, which is generally limited to wh-in-situ languages, gradually developed during MJ; thus, the language turned into the wh-in-situ type at the start of MJ.   However, this is only very indirect evidence at best. From the historical fact one can only logically infer that MJ was a wh-in-situ language.   It cannot prove that OJ was a wh-movement language.

[12] Japanese presents several different types of verbal-nominal mixed category phenomena besides the attributive predicate clause; i.e., the ones involving *-sa* nominalization (Morimoto 1996) and Sino-Japanese deverbal nouns.   The latter are further divided into three types, appearing in Temporal Affix Constructions (Iida 1987; Horiuchi 2004, this volume), Purpose Expressions (Miyagawa 1987), and Light Verb Constructions (Grimshaw and Mester 1988).   The relation among these putative mixed category constructions is not clear.   The attributive predicate clause is different from all others in that the verb inflection alone is responsible for the nominal projection, while the others involve a deverbal noun or a deverbalizing morpheme such as *-sa* followed by a morpheme imparting the verbal character. Horiuchi (2004) presents a head-sharing analysis of the Japanese mixed category constructions along the line of Bresnan (1997).   Although the attributive predicate clause could similarly be analyzed as the case of head-sharing construction, the present analysis adopts Malouf's (2000) style because it better captures the historical change of case marking morphemes.   I appreciate the editors of this volume for reminding me of this issue.

[13] Through rigorous examination of data in *Man'yosyu*, Nomura (1993) claims that the traditional nominal hypothesis is not empirically supported.   The critical point is that there are two types of context which show systematic irregularities; i.e., when the subject is fronted before the focus phrase in *kakarimusubi*, the attributive predicate allows zero-marking and [wa] on the subject (cf. (1)-II above), and [no/ga] can also appear in a conditional clause which is not headed by the attributive predicate.   Nomura (1993) proposes that [no/ga] in OJ marks a phrase within a constituent with "strong unity."   Kikuta (2003b) rejects the proposal as vague and argues that the traditional wisdom is tenable within a Stochastic OT

Moreover, in addition to the wh-word order restriction, rephrased in (8a), another restriction in (8b), which involves the direct object, has recently been observed by Kinsui (2001) and Yanagida (2003). Namely, it seems that the Subj with [no/ga] does not allow the object with the accusative marker [o] to come before the attributive predicate.

(8) a.　　Word Order Restriction [1]　　[NP[no/ga]. . . *wh([ka]). . . predicate(attributive)]
　　 b.　　Word Order Restriction [2]　　[NP[no/ga]. . . *NP[o](=[acc]). . . predicate(attributive)]

Now the question is: are (8a) and (8b) separate? They are separate for Watanabe (2002), since (8b) is observed either with or without [ka], and has nothing to do with wh-movement. And yet their similarity is striking; it is basically a ban on the intervention of the "unity" of [no/ga] and the attributive predicate. Besides, crucially, both of them hold only in OJ, and disappear in MJ. This similarity is unlikely to be an accident.

Another important point is that (8a-b) restrict the surface order of case morphology rather than the possibility of argument realization, given the distribution summarized in (9). The attested patterns are described more schematically in (10). Thus, as in (10a), a direct object with [o] can appear if it is fronted, and the [o]-marked object can occur adjacent to the predicate if [no/ga] are not also present. Moreover, as in (10c) bare NPs are apparently exempt; a bare subject does not induce these restrictions, and, as in (10d), a direct object can occur between [no/ga] and the predicate if it is zero-marked:

(9)　a. Direct object can appear with [o], if it is placed before [no/ga].
　　　b. Direct object with [o] can appear adjacent to the predicate, if [no/ga] is not present.
　　　c. Bare NPs are exempt; bare subject appear before or after wh-phrase, and bare direct object
　　　　　can occur between [no/ga] and the predicate.

(10) a. NP[o]. . .　[NP[no/ga] . . . predicate(attributive)]
　　　b. NP[o]　predicate(attributive)
　　　c. [NP[subj:zero] . . . wh[ka]. . .　predicate(attributive)]
　　　d. [NP[no/ga] . . . NP[obj:zero]. . .　predicate(attributive)]

If the case markers [no/ga] and the attributive form share some nominal nature, it is not very surprising, intuitively at least, that accusative marking [o] and Focus marking [ka] do not intervene. In general, endocentric nominal phrases do not allow non-modificational phrases to occur between a genitive phrase and the nominal head. So the first approximation is that the OJ word order restriction reflects the nominal case frame, which is curiously imposed on arguments licensed by a predicate.

A note is in order, before we proceed, concerning the status of the particle [o]. Although it is the accusative marker in ModJ, its function was rather vague in OJ. For one thing, (i) bare NP was common for the object (as well as for the subject); secondly, (ii) the particle [o] often carried an

---

framework. Importantly, the systematic irregularities, which motivate Nomura's rejection of the nominal hypothesis, fall out automatically as the result of competition among different constraints.

emphatic overtone;[14] and (iii) most curiously, [o] can marks a non-object, and even the subject in a special causal construction, called the *mi*-construction, which is illustrated in (4a-b) above.[15] Marker [o] becomes established as accusative in MJ, and along with it, the *mi*-construction also disappeared.

Finally, Modern Japanese (ModJ) shows a similar but distinct word order restriction. Although [no] in ModJ is unequivocally genitive, and [ga] is nominative, the subject of a nominalized and noun-modificational clause can be marked either with [ga] or [no]. This Ga-No Conversion is a peculiar, but very regular, phenomenon. However, as shown in (11b), genitive [no] subject is not possible for many speakers when it occurs with the accusative [o] object (Harada 1971):

(11)a.　Hirosi –ga nikki -o　yonda　sei –de,　Yuko　-wa　kizutuita.
　　　　　　　 nom　diary　acc　read　because-of　　　top　 was-hurt
　　　　'Yuko was hurt because Hiroshi read her diary.'
　 b.　* Hirosi –no　nikki　-o　 yonda　sei –de,　Yuko　-wa　kizutuita.
　　　　　　 gen　diary　acc　read　because-of　　　top　 was-hurt

This co-occurrence restriction of [no] and [o] is strikingly similar to the OJ restriction in (8b), above.[16] And yet, unexpectedly, neither the preposing of the [o]-marked object as in (12a) nor the dropping of case marker in (12b), improves the sentence at all. This contrasts with the case in OJ, when (10b) and (10c) were just fine:

(12)a.　* Nikki　-o Hirosi –no　　yonda　sei –de,　Yuko　-wa　kizutuita.
　　　　　　 diary　acc　　 gen　read　because-of　　　top　 was-hurt
　 b.　* Hirosi –no　nikki　　　　yonda　sei –de,　Yuko　-wa　kizutuita.
　　　　　　 gen　diary (-zero)　read　because-of　　　top　 was-hurt

I will later show that this follows naturally from my proposal.

4.　Proposal:

---

[14] If [o] in OJ is not purely a case marker but an emphatic (focus?) marker, one may be tempted to claim that (8a) and (8b) are identical and that the pattern (10a) is a type of wh(Focus)-movement after all. This claim, however, is not valid. While the intended wh(Focus)-movement is triggered by the [-interpretable] feature of [ka], the inhibited pattern in (8b) is observed either with or without [ka]. Similarly, the [o]-marked fronting in (10a) does not require the focus particle. This could also cast doubt on the feature-based Focus movement.

[15] *Mi* is a predicate ending particle which attaches to either a verb or an adjective to show nominalization, imperfectivization, among others. *Mi*-construction is a unique structure which forms a causal clause. The subject of the *mi*-attached causal clause is marked either with [o] or zero.
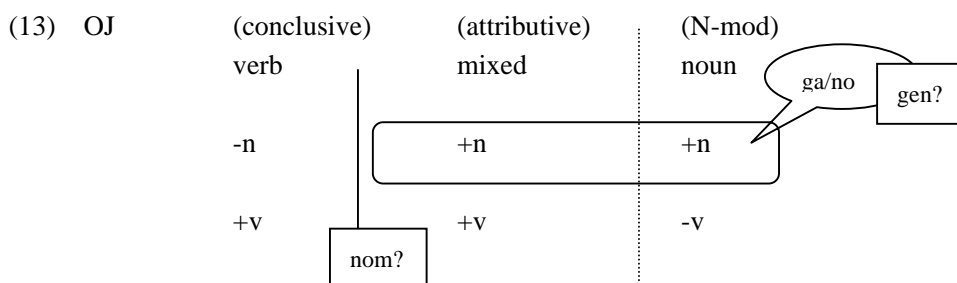
[16] The judgment of (11b) is known to be subject to individual variation. Watanabe (1996) calls this the "Transitivity restriction" and proposes a wh-movement analysis to explain why (11b) is bad. Kikuta (2003a) shows, on the other hand, that this has to do with more surface case-marking rather than the type of the predicate.

To repeat, the key insight is that the OJ word order reflects some kind of restriction on nominal phrases. The solution is sought in the interaction of the profile of case markers in each period and the ranking of OT constraints.

4.1   Abstract Case and Profile of Morphological Case Markers

From the observation in (8) through (10), we saw that the morphological realization of case, and the syntactic licensing of arguments are separate. Specifically, arguments are licensed in f-structure, by virtue of being an argument of a certain grammatical function of a certain head with appropriate case assigning properties. I tentatively use the term "*abstract case*" in the sense that an argument is syntactically licensed. On the other hand, case particles, such as [no] [ga] [o], are a phonological (and morphological) realization of the abstract case, i.e, case-related properties of an argument. Each case particle has a certain profile, and the particle is used only when its profile is compatible with abstract case, or the type of licensing in the above sense. The profile of each morphological case changes through time.

The figure in (13) shows the case profile of [no] and [ga] in OJ, classified according to the categorial type of the head.[17] Notice that I adopt the idea of Malouf (2000) here in assuming that the attributive form predicate is a [+n, +v] mixed category. The profiles in (13) indicate that, in OJ, [ga] and [no] are simply markers of a dependant of a [+n] category. The value of [+/-v] is irrelevant. This is why they appear as both genitive and nominative, and this is also why the subject of the conclusive form does not occur with [no] or [ga]. The irrelevance of [v] for [no/ga] I claim reflects the overall delay in the establishment of [+v] case morphemes, which is also related to the delay in the establishment of the accusative marker [o]:[18]

(13)   OJ   (conclusive)   (attributive)   (N-mod)
          verb   mixed   noun
                                                                      ga/no   gen?

          -n   +n   +n

          +v   +v   -v
                    nom?

Now the case profile gradually changed. As shown in (14), in MJ, the [+v] feature started to count, since the OJ case system was clearly defective. Although [no] and [ga] were still used for both subjects and noun-modifiers; I claim that they are now ambiguous rather than being underspecified: there are both [+v] version and [-v] version for them:

---

[17] The other properties such as grammatical function are omitted for expository purposes.

[18] Thus, it does not make much sense to argue if [ga] and [no] in OJ are genitive or nominative. If "nominative" is defined a [-n, +v] case, and "genitive" a [+n, -v] case, no case markers are available in OJ to genuinely represent nominative or genitive case.

(14) MJ   verb   mixed   ga/no   noun   ga/no   gen

      -n   nom   +n   +n

              +v   ◀┈┈┈┈ +v   -v
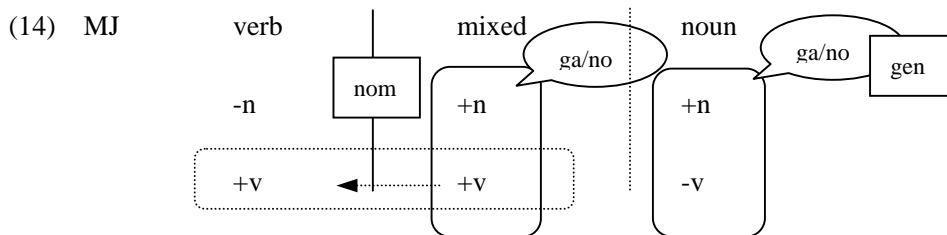
Further down in history, as indicated by the arrow and broken line in the figure (14), [+v] for nominative extended its region even more to cover [-n, +v] as well as [+n, +v], and finally the case profiles developed into a clean two-way system in Modern Japanese, shown in the figure (15). In ModJ, [no] is the [+n] case, and [ga] is the [+v] case.[19]   There is no overlap in their profiles:

(15) ModJ   verb   mixed   noun   no   gen

      -n   +n   +n

    ga

  nom   +v   +v   -v
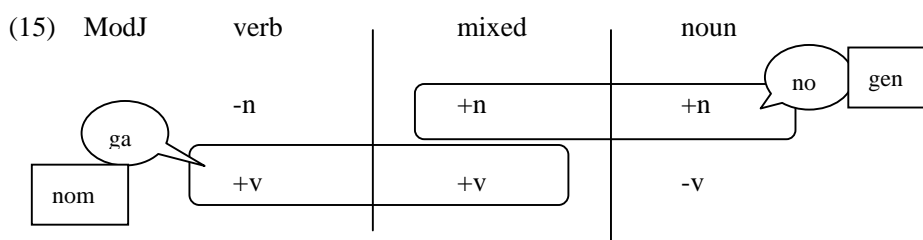
4.2   OT-LFG: Constraints and Their Ranking

The case profiles in the previous section interact with the constraints I propose in (16).

(16) (A) **Appropriate Reflection of Argument Licensing (Abstract Case)**
   (B) **Nominal Phrase Tightness:** $*[_{YP}$ NP[Ncase] XP Y[+N]]   where XP = XP[+mark]
   (C) **Morphological Case Type Consistency**: Case type of coarguments should be consistent (+N-type case, +V-type case).

   The first constraint, (16A) is a simplified form of a Faith constraint.   As I said above, each licensed argument bears "case" in an abstract sense, which is equivalent to the type of argument licensing based on the grammatical function and the like.   And it is essential in determining the appropriateness of the choice of each morphological case marker.   The constraint (16A) checks the basic compatibility in this regard.[20]

---

[19] How and when [ga] and [no] diverged is another long-discussed issue (cf. Nomura 1996).   It is widely known that [ga] and [no] are not completely interchangeable even in OJ-MJ, but they are distinguished in terms of animacy and referentiality, among others.   It is also the fact that the rhetorical (metaphorical) noun modification called *makura-kotoba* (as included in (4a) and (6) above) is limited to [no].   The two, however, are not differentiated in (13) and (14) because (i) their difference is hard to make precise in OJ, when [ga] and [no] are often represented with the same character, and (ii) their difference seems mostly semantic rather than categorical in OJ at least.   Admittedly, this latter situation gradually changes during the course of MJ, and the change marked by the dotted line in (14) applies only to [ga].

[20] Admittedly, this is an over-simplified assumption.   As pointed out by Peter Sells (p.c.), the correspondence between thematic roles (and grammatical functions) and the abstract case is no simple matter but has long challenged theoretical linguists.   This constraint is not meant to ignore the

(16B) encodes the "tightness" of a nominal phrase, which means that in an endocentric nominal phrase, the link between the Spec and head should not be interrupted by a noun-modificational phrase. Note that this constraint is very natural for an NP. Reflecting the case profile, I assume that [Ncase] for NP-tightness is [+n] case in OJ and ModJ, while it is [+n, -v] case in MJ.

Constraint (16C) prescribes that coarguments should share the same case type. Generally, case marking falls into nominal (N) type and verbal (V) type: the former includes the typical genitive, while the latter includes nominative, accusative, dative, and oblique, although there are mixed cases as well. The distinction of case type is quite strictly observed in Japanese. For instance, prenominal modifiers must occur with genitive [no] in ModJ, basically never with verbal cases or with postpositions alone. This is reflected in the contrast in (17) of the English preposition "from" and the Japanese counterpart "kara"; a "kara"-phrase can never occur alone as a noun-modifier as shown in (17b):

(17) a. This wine comes <u>from</u> California vs. Kono wain -wa California -<u>kara</u> kiteiru
     b. The wine <u>from</u> California   vs. California *<u>kara / no / kara-no</u> wain

This principle of case-type consistency obviously lies behind the restriction on [ga/no] conversion in ModJ observed above:

(18) a. John eats bread.   vs. John –ga pan -o taberu.
     b. John's eating bread   vs. John-ga pan –o taberu-koto
                                          *?John-no pan-o taberu-koto

Now I propose that the constraints are ranked as in (19).

(19)   OJ: (A) > (B) >> (C)
       MJ: (A) > (C) >> (B)

I basically follow the idea of Boersma (1997) in assuming constraints are not only ordered but are separated specific numbers apart in weight, although I do not give the specific value of the constraint weight here.

4.3 Evaluation

The tableaux in (20) – (25) show the schematic evaluation in each period. The input is an f-structure of a [+nominal] predicate (attributive form) with two arguments. I take XP[ka] as being a verbal dependant, and [ka] as [+v] particle.

---

importance of the issue in a lexicalist framework. The point, rather, is to show that there can be another layer of correspondence. Most of the previous studies have implicitly assumed that morphological case marking directly reflects the type of abstract case, since their correspondence is often very tight as in ModJ. However, the data in OJ calls for a system which separates the case morpheme from the abstract case.

(20) OJ

GF, GF[Foc], Pred[+v,+n]	AbstCase	NP-tight	Case-type
a. [NP[ga:+n] XP [ka:+v] Pred[+v, +n]]		*	*
b. [NP[no:+n] XP [ka:+v] Pred[+v, +n]]		*	*
c. [NP[ni:+v] XP [ka:+v] Pred[+v, +n]]	*		
☞d. XP [ka:+v] [NP[ga:+n] Pred[+v, +n]]			*
☞e. XP [ka:+v] [NP[no:+n] Pred[+v, +n]]			*
f. XP [ka:+v] [NP[ni:+v] Pred[+v, +n]]	*		

(21)

GF, GF, Pred[+v,+n]	AbstCase	NP-tight	Case-type
a. [NP[ga:+n] NP [o:+v] Pred[+v, +n]]		*	*
b. [NP[no:+n] NP [o:+v] Pred[+v, +n]]		*	*
c. [NP[ni:+v] NP [o:+v] Pred[+v, +n]]	*		
☞d. NP [o:+v] [NP[ga:+n] Pred[+v, +n]]			*
☞e. NP [o:+v] [NP[no:+n] Pred[+v, +n]]			*
f. NP [o:+v] [NP[ni:+v] Pred[+v, +n]]	*		
☞g. [NP[ga:+n] NP [φ] Pred[+v, +n]]			
☞h. [NP[no:+n] NP [φ] Pred[+v, +n]]			
☞i. NP [φ] [NP[ga:+n] Pred[+v, +n]]			
☞j. NP [φ] [NP[no:+n] Pred[+v, +n]]			

Among the candidates in (20), (20c) and (20f), violate the strongest constraint AbstCase, since the subject argument cannot normally be marked with the particle [ni], which is roughly dative. Among the rest, (20a) and (20b), where the Focus phrase XP[ka] intervenes between NP[no/ga] and the predicate, are blocked since they violate the nominal tightness constraint.   The candidates (20d) and (20e), in which the Focus phrase is preposed outside the nominal domain, survive because they violate only the low-ranked Case-type consistency constraint.   Notice that what we have here are the cases of apparent wh-movement.   In other words, what caused the word order restriction in OJ is the dominance of the constraint (B). The evaluation of (21), which involves the direct object, proceeds basically the same way.

The evaluation in (20) and (21) may appear incomplete in that it allows too many "optimal" ones.   In particular, it gives no prediction as to which candidate will be chosen in a given instance, nor is it clear how the candidates which violate the low-ranked constraints come to survive after all.[21]   The evaluation in this section is indeed schematic; it is meant to demonstrate the effect of the word order restriction in OJ and its loss in MJ on the whole class of clauses headed by an

---

[21] This problem was pointed out by Joan Bresnan.   The candidates in (21d-e) violate the NP-tightness constraint, while those in (21g-j) violate none in the evaluation.   Nevertheless, the former will survive after all when the object NP bears semantic and thematic weight, which is ignored here.   As briefly mentioned in the text, it is generally agreed that the particle [o] indicates some kind of emphasis in OJ.

attributive predicate, rather than to explain each attested pattern. In other words, the evaluation shows all syntactically possible candidates. Undoubtedly, the ultimate choice in each actual instance among viable candidates must refer to other factors ignored here. The actual order is determined by considering at least the following factors: (1) informational/thematic prominence of each constituent, (2) phonological weight (or syllable length) of the constituent. The choice between [ga] and [no] must also consider lexical semantics.[22] I also assume that this system follows Boersma's mechanism, and it allows optionality. So when all factors of two viable candidates are on a par, for instance, their occurrence will be about 50% each.

Now the apparent loss of wh-movement in MJ results from the relative demotion of the constraint (B), which is caused by the promotion of (C), which reflects the development of the morphological V-case system in MJ. The evaluation at this stage is shown in (22). Recall that at this point, the case profiles of [no] and [ga] have changed; they are no longer simply [+n] case morphemes, but are ambiguously [+n,+v] and [+n,-v] morphemes:

(22)   MJ

GF, XP[Foc], Pred[+v,+n]	AbstCase	Case-type	NP-tight
☞a. [NP[ga:+v,+n] NP [ka:+v] Pred[+v, +n]]			*
☞b. [NP[no:+v,+n] NP [ka:+v] Pred[+v, +n]]			*
c. [NP[ga:-v,+n] NP [ka:+v] Pred[+v, +n]]		*	*
d. [NP[no:-v,+n] NP [ka:+v] Pred[+v, +n]]		*	*
e. [NP[ni:+v] NP [ka:+v] Pred[+v, +n]]	*		
☞f. NP [ka:+v] [NP[ga:+v,+n] Pred[+v, +n]]			
☞g. NP [ka:+v] [NP[no:+v,+n] Pred[+v, +n]]			
h. NP [ka:+v] [NP[ga:-v,+n] Pred[+v, +n]]		*	
i. NP [ka:+v] [NP[no:-v,+n] Pred[+v, +n]]		*	
j. NP [ka:+v] [NP[ni:+v] Pred[+v, +n]]	*		

---

[22] For instance, it has been observed that the nominative [no] is used for someone that deserves more respect. The exact semantic nature of the two nominative markers, however, is rather controversial.

(23)

GF, GF, Pred[+v,+n]	AbstCase	Case-type	NP-tight
☞a. [NP[ga:+v,+n] NP [o:+v] Pred[+v, +n]]			*
☞b. [NP[no:+v,+n] NP [o:+v] Pred[+v, +n]]			*
c. [NP[ga:-v,+n] NP [o:+v] Pred[+v, +n]]		*	*
d. [NP[no:-v,+n] NP [o:+v] Pred[+v, +n]]		*	*
e. [NP[ni:+v] NP [o:+v] Pred[+v, +n]]	*		
☞f. NP [o:+v] [NP[ga:+v,+n] Pred[+v, +n]]			
☞g. NP [o:+v] [NP[no:+v,+n] Pred[+v, +n]]			
h. NP [o:+v] [NP[ga:-v,+n] Pred[+v, +n]]		*	
i. NP [o:+v] [NP[no:-v,+n] Pred[+v, +n]]		*	
☞j. [NP[ga:+v,+n] NP [φ] Pred[+v, +n]]			
☞k. [NP[no:+v,+n] NP [φ] Pred[+v, +n]]			
l. NP [o:+v] [NP[ni:+v] Pred[+v, +n]]	*		
m. [NP[ga:-v,+n] NP [φ] Pred[+v, +n]]		*	
n. [NP[no:-v,+n] NP [φ] Pred[+v, +n]]		*	
☞o. NP [φ] [NP[ga:+v,+n] Pred[+v, +n]]			
☞p. NP [φ] [NP[no:+v,+n] Pred[+v, +n]]			
q. NP [φ] [NP[ga:-v,+n] Pred[+v, +n]]		*	
r. NP [φ] [NP[no:-v,+n] Pred[+v, +n]]		*	

Among the candidates in (22), (22e) and (22j) are blocked for the violation of AbstCase, as in OJ.   Among the rest, (22c-d) and (22h-i), which involve "genitive" or [-v] version of [no] and [ga], violate Case-type consistency and are blocked.   However, the homophonous strings in (22a-b) and (22f-g) survive because they are marked with the [+v] version of [no] and [ga].   Among the remaining four, (22a-b), where the Focus phrase occurs in the middle, violate the NP tightness, but the violation does not seem to count now.   The evaluation for the direct object in (23) proceeds similarly.

Thus the apparent wh-movement in OJ and its loss in MJ result from the interaction of case profiles and the ranked constraints.   The dominance of Case Type Consistency reflects the development of the v-case.

4.4 Further Consequences: Ga/No Conversion in ModJ:

Now, given the case profiles I proposed for ModJ, with the same constraint ranking as for MJ, the observation on Ga/No Conversion in (11)-(12) obtains automatically.   The data is repeated in (24), and the evaluation is shown in the tableau (25):

(24) a.   Hirosi –ga nikki -o   yonda   sei –de, yuko   -wa kizutuita.
     b. * Hirosi –no nikki -o   yonda   sei –de, yuko   -wa kizutuita.
     c.   Nikki   -o Hirosi –ga yonda   sei –de, yuko   -wa kizutuita.

d. *Nikki -o Hirosi –no yonda   sei –de, yuko   -wa kizutuita.
   'Yoko was hurt because Hiroshi read her diary.'

(25)   ModJ: Ga/No Conversion

GF, GF, Pred[+v,+n]	AbstCase	Case-type	NP-tight
☞a. [NP[ga:+v] NP [o:+v] Pred[+v, +n]]			
b. [NP[no:+n] NP [o:+v] Pred[+v, +n]]		*	*
☞c. [NP [o:+v] NP[ga:+v] Pred[+v, +n]]			
d. NP [o:+v] [NP[no:+n] Pred[+v, +n]]		*	

Among the candidates, (24d) is blocked as cleanly as (24b) due to the violation of Case-type.   The surface string of (24d), with the [o]-marked direct object preposed before the nominative subject, is similar to the acceptable one in OJ, (21e); however, given the case profile and the ranking in ModJ, it has in fact no room for survival.   Thus the contrast of OJ and ModJ is no mystery after all.

In this way, the interplay of the profile of case morphemes and the ranking of the constraint brings forth the desired effect in a simple manner.

5.   Conclusion

I have shown that we need no wh-movement in order to account for the word order restriction in OJ and its loss in MJ.   They simply reflect the gradual and dynamic development of the morphological case system.   It is an undeniable fact that case morphemes such as [no] and [ga] changed their profile over the years.   The change in the profile of case particles is partly lexical, but it has systematic consequences on the surface syntactic structure.

The proposed analysis gives a comprehensive picture of the diachronic facts of the language.   It captures the fact that *kakarimusubi* flourished both in OJ and MJ, with a difference in the freedom of word order.   We do not need to make an arbitrary distinction between the UG-based (=syntactic) *kakarimusubi* in OJ and the mere stylistic one in MJ, which finds no empirical support outside the particular theory.

One of the issues implicitly addressed is the primacy of constraint ranking in syntax.   The analysis presented in this paper assumes a dual system of a case profile and constraint ranking, both of which are subject to change.   To be sure, there is a conceptual redundancy in this system, but at this point, I believe this duality is necessary to accommodate a syntactic system, so long as the syntax involves lexical items.

The analysis presented here is programmatic in several ways.   It ignores the complexity of syntactic (abstract) case licensing, and it does not discuss how functional and phonological factors interact to derive the optimal word order.   Another caveat, of course, is the universal applicability of the constraints proposed, which has to be tested against more empirical data from different languages.   Nevertheless, I hope to have shown that the OT framework offers a totally new perspective to the diachronic syntactic change of the language.

Selected References

Asudeh, A. 2001. Linking, Optionality, and Ambiguity in Marathi. *Formal and Empirical Issues in Optimality Theoretic Syntax,* ed. P. Sells, 257-312. Stanford, CA: CSLI Publications.

Boersma, P. 1997. How We Learn Variation, Optionality, and Probability. *IFA Proceedings 21*, 43-58. ROA-221-109.

Boersma, P. & Hayes, B. 2001. Empirical Tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45-86.

Bresnan, J. 1997. Mixed Categories as Head Sharing Constructions. *On-line Proceedings of the LFG1997 Conference,* eds. M. Butt & T. H. King. Stanford, CA: CSLI Publications.

Bresnan, J. & Aissen, J. 2001. Optimality and Functionality: Objections and Refutations. *Natural Language & Linguistic Theory* 20: 81-95.

Bresnan, J., Dingare, S., & Manning, C. D. 2001. Soft Constraints Mirror Hard Constraints: Voice and Person in English and Lummi. *On-line Proceedings of the LFG2001 Conference,* eds. M. Butt & T. H. King. Stanford, CA: CSLI Publications.

Choi, H.-W.1999. *Optimizing Structure in Context: Scrambling and Information Structure.* Stanford, CA: CSLI Publications.

Dingare, S. 2001. The Effect of Feature Hierarchies on Frequencies of Passivization in English. Unpublished MA thesis, Stanford University. ROA-467-0901

Grimshaw, J. & Mester, A. 1988. Light Verbs and Theta Marking. *Linguistic Inquiry* 19: 205-232.

Handou, H. 2003. Kakarimusubi-ron no Saikoutiku [*Kakarimusubi* Revisited], *Kokugo Kokubun*, vol. 72, no 5, 33-45.

Harada, S.I. 1971. "Ga-No Coversion and Ideolectal Variations in Japanese," *Gengo Kenkyuu* 60, 25-38.

Horiuchi, H. 2004. Lexical Integrity, Head Sharing, and Case Marking in Japanese. *On-line Proceedings of the LFG2004 Conference,* eds. M. Butt & T. H. King. Stanford, CA: CSLI Publications.

Iida, M. 1987. Case-Assignment by Nominals in Japanese. In *Working Papers in Grammatical Theory and Discourse Structure: Interactions of Morphology, Syntax, and Discourse*, eds. M. Iida et al., 93-138. Stanford, CA: CSLI Publications.

Kikuta, C.U. 2003a. Ga-No Kotai-Gensyu ni Hasei wa Hituyou-ka: HPSG de Meisisei to Zyutugosei o tunagu [Is Derivation Necessary for Ga/No Conversion?: A Mixed-Category Analysis in HPSG].   *JELS 20: Papers from the Twentieth National Conference of the English Linguistic Society of Japan*, 61-70.

Kikuta, C.U. 2003b. Subject Case Marking in Early Classical Japanese; A New Perspective from Stochastic Optimality Theory. In *Japanese/Korean Linguistics* 12, ed. W. McClure, 152-164. Stanford, CA: CSLI Publications & SLA.

Kinsui, S. 2001. A History of Japanese Grammar Examined through Case Markers. A lecture given at Tokyo University, August 26, 2001.

Kuhn, J. 2003. *Optimality-Theoretic Syntax: A Declarative Approach.* Stanford, CA: CSLI Publications.

Malouf, R. P. 2000. *Mixed Categories in the Hierarchical Lexicon.* Stanford, CA: CSLI Publications.

Miyagawa, S. 1987.　Restructuring in Japanese.　In *Issues in Japanese Linguistics*, eds. K. Imai, et al., 273-300.　Dordrecht: Foris.

Morimoto, Y. 1996.　Nominalization or Verbalization?: Analyzing Mixed Case Marking in Japanese. Ms. Stanford University.

Mugane, J. M. 2003.　Hybrid Constructions in Gikuyu: Agentive Nominalizations and Infinitive-gerund Constructions.　In Nominals: Inside and Out, eds. M. Butt & T. H. King, 235-265. Stanford, CA: CSLI Publications.

Nomura, T. 1993. Zyodaigo no *no* to *ga* nituite, [On *no* and *ga* in Old Japanese] *Kokugo Kokubun*, vol. 62, nos 2-3, 1-17 & 30-49.

Nomura, T. 1995. Ka-niyoru Kakarimusubi Siron, [On *Kakarimusubi* by *ka*] *Kokugo Kokubun*, vol. 64, no 9, 1-27.

Nomura, T. 1996. Ga-Syusikei e, [Toward Ga-Conclusive Form] *Kokugo Kokubun*, vol. 65 no 5, 524-541.

Nomura, T. 2002.　Kakarimsubi no Kobun: Zyodai kara Tyuko e [The *Kakarimusubi* Construction: From OJ to MJ].　Paper presented at the Symposium at Kansai Linguistic Society.

Ohno, S.　1993.　*Kakarimusubi no Kenkyu* [*A Study on Kakarimusubi*].　Tokyo: Iwanami.

Sasaki, T. 1992. Joudaigo ni okeru –ka-ha no koubun. [On the construction of –ka-ha in Old Japanese] *Kokugo Kokubun*, vol. 61, no 5, 17-33.

Sigurdsson, H.A. 2003. Case: Abstract vs. Morphological.　In *New Perspectives on Case Theory*, eds. E. Brandner & H. Zinsmeister, pp. 223-267. Stanford, CA: CSLI Publications.

Tonoike, S. 2003. Joudai Nihongo ni Sahou Wh-Idou wa soredemo attaka [Should we still believe that Old Japanese had leftward wh-movement?], *Gengo* 32:1, 146-153.

Yanagida, Y. 2003. Obligatory Movement and Head Parameter: Evidence from Early Old Japanese, *Sophia Symposium on Classical Japanese and Linguistic Theory: Sophia Linguistica* 50, 103-116.

Watanabe, A.　1996.　Nominative-Genitive Conversion and Agreement in Japanese: A Cross-Linguistic Perspective, *Journal of East Asian Linguistics* 5, 247-286.

Watanabe, A.　2002.　Loss of Overt Wh-Movement in Old Japanese. In *Syntactic Effects of Morphological Change*, ed. D.W. Lightfoot, 179-195.　Oxford: Oxford UP.