Yanagida, Yuko. 2026. Left periphery and microparametric word order change in OV languages: Evidence from Japanese. *Glossa: a journal of general linguistics* 11(1). pp. 1–42. DOI: https://doi.org/10.16995/glossa.25426

Open Library of Humanities

Left periphery and microparametric word order change in OV languages: Evidence from Japanese

Yuko Yanagida, University of Tsukuba, yanagida.yuko.ft@u.tsukuba.ac.jp

This paper examines diachronic changes in Japanese focus constructions within the parameter hierarchy framework (Roberts 2019). The analysis adopts Kayne's (1994) Linear Correspondence Axiom, under which head-final orders arise via roll-up movement of a complement to its specifier position. It is argued that Modern Japanese clefts of the form [[...wa] XP da] are monoclausal, derived by focus movement of XP followed by what I call *core-spine roll-up* movement of FinP to Spec,ForceP, and that the particle *wa* and the copula *da* instantiate functional heads (Force^0 and Foc^0) within the split CP layer (Rizzi 1997). This head-initial analysis of clefts resolves several long-standing puzzles: the asymmetry between NP and PP clefts, the amelioration effects of numeral quantifiers, and the contrasts between scrambling and focus movement. These acceptability patterns follow from Rizzi's (2016) Maximality Principle, under which only an element that is maximal with respect to a given label may undergo A′-movement. Turning to Old Japanese, I propose that *kakari-musubi* (focus concord) constructions are the diachronic precursor of the modern cleft construction, with the topic *pa* (Top^0) and the focus particles *so*/*ka* (Foc^0) heading a disharmonic, head-initial left periphery that relies on agreement-based movement rather than roll-up movement. The shift from *kakari-musubi* to clefts is analyzed as microparametric change in the feature specifications of left-peripheral functional heads within the CP phase, yielding new configurations of word order and clause structure.

Glossa: a journal of general linguistics is a peer-reviewed open access journal published by the Open Library of Humanities. © 2026 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/. Articles published by the Open Library of Humanities may also contain other CC-licensed material, or material used or reproduced under an exception or limitation to copyright. Such material is not subject to the article's license. It is either marked with its own license, or marked that it has been reproduced with permission or under fair-use.
OPEN ACCESS

1 Introduction

Traditionally, word order change has been characterized as a macroparametric shift from OV to VO order, as exemplified in the shift observed in the history of English. More recent developments in parametric theory, however—most notably the parameter hierarchy approach developed by Roberts (2019) and collaborators from the mid-2000s onward within the microparametric tradition (cf. Biberauer et al. 2014; Biberauer & Roberts 2015; 2017 among others)—have redirected attention from language-wide macroparameters to finer-grained parametric variation. On this view, cases of word order change need not entail wholesale restructuring of syntactic architecture; rather, they may result from local adjustments targeting specific functional heads within extended projections.

The parameter hierarchy literature has largely concentrated on Germanic and other European languages—predominantly VO word order—while empirical studies of OV languages remain comparatively scarce. To address this gap, this study draws on Japanese data, with particular attention to diachronic changes in focus constructions. This paper adopts the parameter hierarchy model, with a particular focus on the role of Kayne's (1994) Linear Correspondence Axiom (LCA), where head-initial order is taken to be basic, and head-final order is derived via roll-up movement of the complement to a specifier position. Empirically, however, the claim that head-final orders are uniformly derived under the LCA remains contested and largely theory-internal, with relatively few decisive, cross-linguistically robust arguments—particularly from rigorously head-final systems like Japanese—having been advanced to date. Against this backdrop, the paper offers new synchronic and diachronic evidence which, when coupled with microparametric variation, demonstrates that this framework yields a more explicit and explanatory account of Japanese focus constructions and patterns of word order change.

The purpose of this paper is twofold. First, I argue that Modern Japanese clefts of the form [[…*wa*] XP *da*] are monoclausal, derived by focus movement of XP followed by what I call *core-spine roll-up* movement of FinP to Spec,ForceP, instantiating a head-initial left periphery in which *wa* and the copula *da* are functional heads (Force^0 and Foc^0) within the CP layer (Rizzi 1997). Second, clefts in Modern Japanese and so-called *kakari-musubi* (focus concord) constructions in Old Japanese (8th century) are analyzed as diachronically related manifestations of the same left-peripheral focus system, modeled within Rizzi's articulated CP layer. Diachronic changes in these focus constructions are captured as microparametric adjustments of individual functional heads (Force^0, Foc^0, Fin^0), which result in changes in word order and clause structure in Japanese. Given the assumption that the clausal spine is head-initial, we depart from the standard analysis of Japanese particles such as topic *wa*, Q-particle *ka*, which are assumed to be D/Q heads within DP, and instead treat them as functional heads of the extended verbal projection.

The paper is structured as follows. Section 2 outlines the parameter hierarchy theory (Roberts 2019), which I adopt throughout the paper. Section 3 analyzes the syntax of so-called "cleft" focus

constructions in Modern Japanese. Building on Hiraiwa & Ishihara's (2002) split CP analysis of clefts, I argue that adopting a head-initial CP resolves many empirical puzzles in the literature on Japanese cleft constructions. Section 4 turns to diachronic data from Old Japanese and shows that syntactic change is best analyzed as involving microparametric shifts in left-peripheral functional heads. Taken together, these analyses show that, on both empirical and theoretical grounds, the head-initial hypothesis has an advantage over the head-final analysis: clause-structure change can be captured as changes in the localized settings of individual functional heads, with certain particles playing a key role in this process.

2 Theoretical background

2.1 Overview of the parameter hierarchy hypothesis

The notion of the parameter hierarchy—the idea that parameters are organized in a structured, dependent sequence rather than as a flat set of independent switches—was first explicitly formulated by Roberts & Holmberg (2010) in their work on syntactic variation. Their approach is discussed in detail in Biberauer et al. (2014) and further developed in Roberts (2019). The discussion below proceeds on the basis of the framework advanced by Roberts (2019).

Under the parameter hierarchy approach, cross-linguistic variation in word order and syntactic structure is not governed by a single, overarching macroparameter, but rather by a hierarchically organized system of parameters. These parameters range from macroparameters, which affect the entire language, to mesoparameters, microparameters, and nanoparameters, each defined by their scope and granularity. The values of these parameters are determined by the presence or absence of specific formal features on functional heads within the extended projections of phrases.

A crucial theoretical assumption in this model is Kayne's (1994) Linear Correspondence Axiom (LCA). The LCA states that there is no OV/VO parameter; head-initial order (as in English) is the default and OV order (as in Japanese) results from the movement of complements to specifier positions—a process which is called *roll-up* movement, defined by the formal features of functional categories (Biberauer et al. 2014).[1] When this feature is present on a head, it triggers the movement of its complement to the specifier position, and this can apply recursively up the extended projection. Depending on whether extended functional heads contain the roll-up feature, languages are classified into three groups.

[1] Biberauer et al. (2014) use the caret (ˆ) to mark a movement-triggering feature. The ˆ feature is purely formal—essentially an EPP-type requirement—and carries no semantic content. Roberts (2019) proposes an alternative trigger for roll-up movement grounded in Chomsky's (2013; 2015) Labelling Algorithm, rather than an independent ˆ/EPP-style feature.

(1) A. In strictly head-initial (VO) languages (e.g., English, Romance), none of the functional heads in the extended projection have the relevant features, so roll-up movement does not occur.
B. In strictly head-final (OV) languages (e.g., Japanese, Korean), all relevant functional heads possess the features that trigger roll-up movement, resulting in consistently OV order.
C. In mixed or disharmonic languages (e.g., German), only some functional heads (such as T) have these features, leading to more complex word order patterns (e.g., Verb-Second, differing main and subordinate clause orders).

By deriving head-final order through roll-up movement as mandated by the formal features of functional heads, the theory offers a principled explanation for both the typological diversity and the historical change of word order. Rather than requiring a wholesale reconfiguration of syntactic structure (macroparametric reset), the framework allows for gradual, local changes in the feature specification of individual heads.

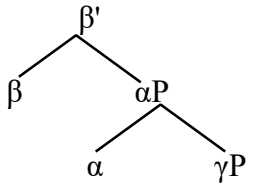
2.2 FOFC and disharmony

Within Roberts's (2019) parameter-hierarchy approach which is primarily based on Biberauer et al. (2014), the Final-Over-Final Condition (FOFC) plays a crucial role in constraining word-order variation. In particular it restricts a disharmonic word order within extended verbal projections.

(2) The Final-Over-Final Condition
A head-final phrase βP cannot dominate a head-initial phrase αP where α and β are heads in the same Extended Projection.
(Holmberg (2000: 124), restated in Biberauer et al. (2014) and Roberts (2019: 103)).

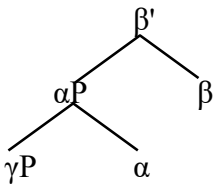
Now consider four possible complementation combinations among head-initial and head-final categories (Biberauer et al. 2014: 171). Among them, only (3d) is ruled out by FOFC.

(3) a.



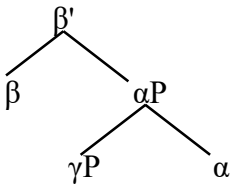
Consistent head-initial (harmonic)

b.



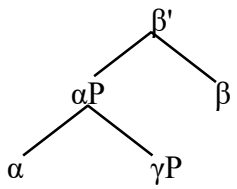
Consistent head-final (harmonic)

c.



Initial-over-final (disharmonic)

d.



Final-over-initial (disharmonic)

iberauer et al. (2014) assume Kayne's (1994) Linear Correspondence Axiom (LCA), under which every phrase within the extended verbal projection is head-initial. That is, the head-initial order in (3a) is the unmarked base order, requiring no further operation. The head-final order in (3b) is derivationally more complex, being derived by roll-up movement: γP, the complement of α, first undergoes roll-up to Spec,αP and subsequently αP undergoes roll-up to Spec,βP, monotonically from the bottom up within the extended projection (EP) of a verb or a noun.[2] The FOFC allows these two harmonic word orders, which are the most common. The FOFC also allows the disharmonic configuration in (3c), which is attested in V2 languages like German (4), where a head-initial complementizer dominates a head-final TP. Given that the finite auxiliary is in T, roll-up would start at *v*P, containing the subject, and raise it to Spec,TP, leaving the trace $<v\text{P}>$ behind.

(4) [C [$_{\text{TP}}$ [$_{v\text{P}}$ Subject Object V] T $<v\text{P}>$]] (=(3c))
… daβ Johann das Buch gelesen hat
that John the book read.PPT have.3SG.PRES
'… that John has read the book.' (Roberts 2019: 114)

They argue that the FOFC accounts for a range of cross-linguistic generalizations, as illustrated in (5):

(5) A. Final Cs are not found in VO languages, whereas initials Cs are found in both VO and OV languages.
B. Final auxiliaries in VO languages are not found; that is, *V-O-AUX orders are non-existent, whereas Aux-O-V type languages are attested.
C. In languages with disharmonic word order, the higher a head is in an extended projection, the less likely it is to be head-final (Biberauer et al. 2014: 172).
D. In diachronic word-order change, the shift from head-final to head-initial starts at the top of the extended projection, whereas the shift from head-initial to head-final starts at the bottom (see (56–57) in section 4).

Biberauer et al. (2014) discuss a potential FOFC violation in Mandarin Chinese which has basic SVO order but exhibits clause-final question particles such as *ma*.

(6) Hongjian xihuan zhe ben shu **ma**?
Hongjian like this CL book Q
'Does Hongjian like this book?'

(6) could be derived in such a way that the head-initial TP raises to Spec,CP, as schematized in (7), yielding a head-final C. This corresponds to the disharmonic configuration in (3d) which is ruled out by FOFC.

[2] In order for roll-up movement to be FOFC-compliant, Roberts (2019) proposes the so-called "Start At The Bottom Generalization (SATBG)": in a given extended projection where roll-up applies, the cycle of roll-up must apply in the minimal XP containing the Root (Roberts 2019: 141(78')).

(7) *[$_{CP}$ [$_{TP}$ Subj V Object] C <TP>]

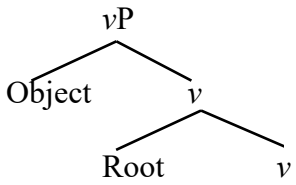
Given that the FOFC constrains only the order of heads within a single extended projection (from V up through *v*, T, and C), Biberauer et al. (2014: 201) propose that clause-final particles as in (6) fall outside the extended verbal projection and thus outside the scope of the FOFC. According to them, Q-particles are "acategorial" elements with a cordinating function that have the structure [TP [OR-NOT ~~TP~~]], where ellipsis of the second TP, identical with the first TP, leaves the negative disjunction as a clause-final particle. The related issue of the Q particle *ka* in Japanese is discussed in detail in section 3.3.

2.3 Labeling

Roberts (2019: 156) adopts Chomsky's (2013; 2015) labeling algorithm (LA) and claims that roll-up is driven by the need to satisfy labeling requirements: Leaving technical details aside, the gist of Roberts's (2019: 160) is that in a structure [$_{\alpha}$ X, YP], how the category of X is determined depends on whether X's categorial feature is weak or strong. When X's categorial feature is strong, X provides the label for α. When X's categorial feature is weak, X can be labelled in several ways. One option is to move some φ-bearing DP into a specifier position, as in subject movement to Spec,TP. Another option is that X is labelled by roll-up, i.e. complement-to-specifier movement. In other words, roll-up is driven by the need for merged objects to be categorially identifiable.[3]

Furthermore, Roberts (2019: 159) assumes that verbs (V) and nouns (N) are Roots and they themselves do not possess categorial features (Embick 2010: 13); each Root merges with a categorizer (*v* or *n*): [$_{v}$ Root *v*] and [$_{n}$ Root *n*], and receives a categorial value. Suppose that [VO] is a default order, and moves to Spec *v*P, yielding [$_{vP}$ VO *v*]. The derivation leads to a violation of FOFC. To avoid a FOFC violation, Roberts (2019: 159) suggests that at the stage of the first merge, OV (Object-Verb) is the basic word order. At the deepest branching node in an EP, the object merges with the complex head [$_{v}$ Root *v*], as shown in (8).

(8)



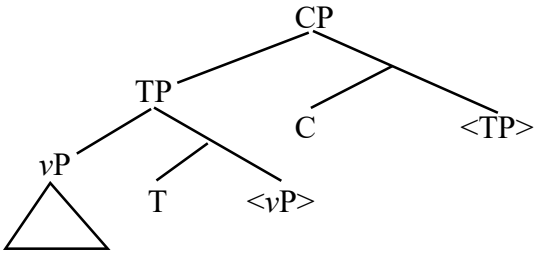
Importantly, in the structure [$_{vP}$ Object [$_{v}$ Root *v*]], the object asymmetrically c-commands the Root, so there is no violation of Kayne's LCA. Roberts (2019: 159) notes that "as long as the

[3] In addition, in symmetrical configurations of the form [XP, YP] Chomsky (2013; 2015) proposes that labeling can be achieved by shared features such as Q-features; that is, the common feature set of the two phrases determines the label of the resulting object.

categorizer is silent, there is no linearization problem." Under this approach, roll-up starts at *v*P and applies cyclically all the way up to CP in harmonic OV languages like Japanese, resulting in [[[OV] T] C]. The derivation of (9) in Japanese is schematized in (10), which corresponds to (3b); the angle brackets < > indicate traces of movement.

(9) Taroo ga hon o ka-tta to (omo-tta)
Taroo-NOM book ACC buy-PST COMP (think-PST)
'(I thought) that Taroo bought the book.'

(10)



In Japanese-type OV systems, if we follow Roberts's labeling mechanism, the categorial features of T and C are weak, so TP and CP cannot be labeled by T/C alone and instead receive their category via successive phrasal roll-up: *v*P raises to SpecTP and TP raises to Spec,CP. Since Spec,TP is occupied by *v*P and Spec,CP by TP for labeling purposes, there is neither space nor need for subject-to-SpecTP or *wh*-to-Spec,CP. By contrast, in English-type VO systems, labeling is achieved by feature sharing rather than roll-up: T bears ϕ/EPP and attracts the subject to Spec,TP to yield a label via ϕ-agreement, and C bears Q/A′-EPP and attracts a wh-phrase to Spec,CP to yield a label via Q-feature sharing. Thus the two systems implement mutually exclusive solutions to the labeling problem: Japanese resolves TP/CP labels by phrasal roll-up (precluding subject/wh movement), whereas English resolves them by subject and wh movement (precluding roll-up).

2.4 The Fukui–Kuroda parametric view of Japanese syntax

Within the framework of generative syntax, the literature has shown that major differences between Japanese and English can be captured in terms of properties such as those in (11):

(11) a. multiple subjects and multiple genitives
b. wh-movement
c. scrambling and relatively free word order
d. overt case marking for grammatical roles

In (11), Japanese and English exhibit opposite patterns with respect to these properties: in Japanese, the relevant properties involve the presence of (11a), the absence of (11b), and the presence of (11c) and (11d), whereas English shows the reverse pattern.

Kuroda (1988; 2007) and Fukui (1986; 1995) analyze this constellation of properties as a cluster effect of the OV/VO macro-parameter, with Japanese instantiating a head-final OV system and English a head-initial VO system. Importantly, Old Japanese (OJ) (8th century) displayed the opposite profile from Modern Japanese (ModJ), aligning more closely with English on the parameters in (11). OJ licensed only a single grammatical subject per clause, required wh-movement (Watanabe 2002), and exhibited relatively rigid word order (Yanagida 2006). In terms of case, Miyagawa (1989; 2012) argues that OJ allowed objects to be licensed by abstract Case (case relations not overtly marked in the morphology), whereas ModJ requires overt morphological case on objects, with the historical shift tied to changes in the verbal inflectional system that eliminated the earlier option for abstract Case licensing.

Kuroda (1988), drawing on Chomsky's (1986) X-bar theory, proposes that languages vary in whether X-Agreement is "forced." In forced-agreement languages such as English, subjects must raise to Spec-TP for abstract Case licensing, expletives are required when no subject is present, and wh-movement is obligatory. In non-forced-agreement languages such as Japanese, subjects may remain inside VP, expletives are unnecessary, and wh-movement does not apply. On the basis of diachronic evidence provided by Watanabe (2002), Kinsui (2002), and Yanagida (2006), Kuroda (2007) further argues that Japanese shifted from a forced agreement system in OJ to a non-forced agreement system in the modern period.

Alternatively, Fukui (1986, 1995) develops a reductionist view on which cross-linguistic parametric variation is derived from the inherent formal properties of functional heads. On Fukui's view, while the functional category I (=T) exists only as a placeholder for the tense morpheme, Japanese lacks the kind of agreement-bearing functional heads that obligatorily attract phrases to their specifier positions, which explains the absence of subject movement and *wh*-movement. At the same time, because movement is not tightly tied to agreement, the language freely permits scrambling, multiple subjects within a single clause, and multiple genitives inside a nominal.[4]

2.5 The position of the subject

A Fukui-Kuroda-style theory in which the subject stays inside *v*P is in principle compatible with an FOFC-compliant roll-up, derived through complement-to-specifier movement of each phrase in the extended projection, cyclically from bottom to top. However, numerous works since Fukui and Kuroda have suggested that a certain subject moves out of *v*P into a higher position (Spec,TP)

[4] Kuroda (1988) and Fukui (1986; 1995) belong to the earliest wave of parametric work that treated the English–Japanese differences in (11) as manifestations of a single OV/VO parameter. The discussion below takes this early parametric perspective as historical context, but does not presuppose a particular version of their analyses. For an overview of the key theoretical differences between Fukui and Kuroda's systems, see Fukui (2022).

(cf. Takezawa 1987; Tateishi 1994; Miyagawa 2010, among others). Miyagawa (2010) argues that discourse features such as topic and focus on C are inherited by an intermediate functional head α between C and T. He takes the subject to be the canonical bearer of these discourse features, so the subject typically raises from its *v*P-internal position into Spec,αP in the left periphery of the TP-domain.

Kayne (1994: 143, fn. 3) in fact proposes that the topic *wa* and the subject marker *ga* in Japanese are clausal heads rather than postpositional particles, schematically [$_{\alpha P}$ John [*wa/ga* [$_{vP}$ …]]] in which the subject moves to Spec,αP, and Whitman (1998; 2001) refines this analysis, by arguing that *wa* and *ga* head CP and TP, respectively. Note, however, that if αP = TP, then TP may not undergo further roll-up to Spec,CP, as this would form the structure [$_{CP}$ [$_{TP}$ John *ga* [$_{vP}$…]] C], which leads to a FOFC violation. Once A or A′ movement occurs in a Spec position, the derivation no longer allows further roll-up due to FOFC. A question then arises as to how FOFC-compliant roll-up produces the clause-final C. This issue will be discussed in section 3.3.

Two main uses of Modern Japanese *ga* are often distinguished: an *exhaustive-listing* use, assumed to be a type of narrow focus, in which *ga* marks a focused phrase interpreted as the unique value among relevant alternatives, and a *neutral-description* use, in which *ga* appears in broad focus, thetic sentences that simply present an event or situation without any implication of exhaustivity (Kuno 1973). While the exact position of *ga* is not within the scope of this paper, it is likely that a subject with an exhaustive-listing interpretation moves out of *v*P, whereas *ga* with a neutral-description use stays inside *v*P (cf. Diesing 1988).

In a series of works such as Roberts (2007; 2019), he argues that diachronic change in syntax is fundamentally driven by changes in the feature composition of functional heads: when the formal specifications of these heads are reanalyzed, the result is new configurations of movement, word order, and clause structure. Adopting this microparametric view of syntactic change, the present paper specifically claims that topic–focus fields within CP are headed by these discourse-related particles and that the feature specification of these functional heads is central to syntactic change in focus constructions in Japanese.

3 Synchronic analysis of focus constructions in Modern Japanese

This section examines the Japanese focus construction traditionally referred to in the literature as a “cleft” (Hoji 1987). Structurally, this construction resembles the English pseudo-cleft construction of the type illustrated in (12).

(12) What [$_{TP}$ John bought yesterday] is [$_{XP}$ this book].

In Japanese, the presuppositional clause is nominalized with *no*, marked with the topic particle *wa*, and followed by the focused constituent (XP) plus the copula *da*. The element immediately preceding the copula functions as the focus, introducing new information, while the nominalized clause sets up the presupposition or background.

(13) […. no] wa XP da
NMLZ TOP Focus COP

This construction has been the subject of extensive research, and a range of issues have been debated: whether it qualifies as a genuine cleft or a pseudo-cleft; whether the clause structure is biclausal or monoclausal; whether the focused XP is derived via movement or base-generated; how the construction connects to the topic–focus partition in information structure; and what role nominalization and particles play in licensing it.

Hoji (1987; 1990) identifies two distinct structures for Japanese clefts: the case-marked cleft (CM cleft) (14) involves movement, while the non-case-marked cleft (non-CM cleft) (15) does not. The latter is instead licensed by the "aboutness condition" (Kuno 1973). Hoji's analysis builds on evidence from subjacency effects: CM clefts display subjacency violations, while non-CM clefts do not.

(14) **Case-marked clefts (CM clefts)**
[Yumi ga ka-tta no] wa [hon o san-satu] da.
Yumi NOM buy-PST NMLZ TOP book ACC 3-CL COP
'It is three books that Yumi bought.'

(15) **Non-case-marked clefts (non-CM clefts)**
[Yumi ga ka-tta no] wa [hon-ø san-satu] da.
Yumi NOM buy-PST NMLZ TOP book 3-CL COP
'It is three books that Yumi bought.'

Another robust diagnostic is multiplicity of focus. CM clefts (16) may host multiple focused constituents, each transparently marked with its own case particle; non-CM clefts (17), in contrast, cannot.

(16) Multiple Clefts with CM clefts
[ka-tta no] wa [Yumi ga sono hon o LA de] da.
buy-PST NMLZ TOP Yumi NOM this book ACC LA in COP
' (Lit.) It is Yumi, this book, in LA that she bought.'

(17) Mutiple Clefts with non-CM clefts
*[ka-tta no] wa [Yumi-ø sono hon-ø LA-ø] da.
buy-PST NMLZ TOP Yumi that book LA COP
'(Lit.) It is Yumi, this book, LA that she bought.'

This paper focuses on CM clefts and investigates the nature of the movement involved. A key question arises here: If CM clefts instantiate focus movement, as Hoji argues, how can this be reconciled with the empirical fact that Japanese lacks overt wh-movement into Spec,CP, and with the theoretical assumption that Japanese clause structure involves roll-up movement of TP into Spec,CP? Addressing this tension requires a detailed examination of the empirical facts surrounding CM clefts, undertaken in the following subsections.

3.1 Empirical facts

A central issue that arises in the analysis of CM clefts concerns the grammatical status of case-marked NPs inside the focus position. Cho et al. (2008) indicate that, across both Japanese and Korean, case-marked NPs are often judged as marginal or degraded by many speakers. Strikingly, however, when the NP is followed by a numeral quantifier (NQ), the cleft becomes fully acceptable. The contrast is illustrated below for Japanese (18) and Korean (19).

(18) Japanese

a. [Yumi ga ka-tta no] wa [hon o san-satu] da.
Yumi NOM buy-PST NMLZ TOP book ACC 3-CL COP
'It is three books that Yumi bought.'

b. [Yumi ga ka-tta no] wa hon (*/?o) da.
Yumi NOM buy-PST NMLZ TOP book ACC COP
'It is a book that Yumi bought.'

(19) Korean

a. [Yumi ka san kes] un [chayk-ul sey-kwen] i-ta.
Yumi NOM bought NMLZ TOP book-ACC 3-CL COP-DEC
'It is three books that Yumi bought.'

b. [Yumi ka san kes] un chayk-(*ul) i-ta.
Yumi NOM bought NOMZ TOP book-ACC COP-DEC
'It is a book that Yumi bought.'

Cho et al. (2008) further point out that the asymmetry becomes even sharper in the case of subject clefts. Clefting a case-marked subject yields complete ungrammaticality in both Japanese and Korean, as shown in (20–21).

(20) Japanese

[kasikoi no] wa kuzira (*ga) da.
intelligent NMLZ TOP whale NOM COP
'It is whales that is intelligent.'

(21) Korean
[ttokttokhan kes] un kolay (*ka) i-ta.
intelligent NMLZ TOP whale NOM COP-DEC
'It is whales that is intelligent.'

Nonetheless, the restriction does not indicate that subjects cannot be clefted at all. When an NP is followed by a numeral quantifier, or when multiple focused elements are involved, the resulting cleft sentences are fully acceptable.

(22) Japanese
a. [mise kara dete-ki-ta no] wa gakusei (*ga) da.
store from out-come-PST NMLZ TOP student NOM COP
'It is a student who came out of the store.'
b. [mise kara dete-ki-ta no] wa [gakusei ga futari] da.
store from out-come-PST NMLZ TOP student NOM two COP
'It is two students who came out of the store.'
c. [dete-ki-ta no] wa [gakusei ga mise kara] da.
out-come-PST NMLZ TOP student NOM store from COP
'(Lit.) It is a student, from the store, who came out.'

(23) Korean[5]
a. [kakey eyse naon kes] un hakseyng (*i) i-ta.
store from came.out NMLZ TOP student (*NOM) COP-DEC
'It is a student who came out of the store.'
b. [kakey eyse naon kes] un [hakseyng i twul] i-ta.
store from came.out NMLZ TOP student NOM two COP-DEC
'It is two students who came out of the store.'
c. [naon kes] un [hakseyng i kakey eyse] i-ta.
came.out NMLZ TOP student NOM store from COP-DEC
'(Lit.) It is a student, from the store, who came out.'

The cross-linguistic parallel suggests that the degraded status of single case-marked focus cannot be dismissed as an idiosyncrasy of Japanese, but reflects a deeper property of the construction in both languages. Cho et al. (2008) argue that the problematic status of case-marked focus and the possibility of multiple focus in clefts undermine the movement-based account. Instead, they propose that case-marked clefts should be analyzed as a kind of "amalgam cleft" (Declerck 1988), essentially a subtype of gapping, which does not involve focus movement.[6] However, this

[5] I am grateful to Hwang Hyun Kyung for providing the example in (23) and for sharing her native-speaker judgments.
[6] Cho et al. (2008) analyze CM clefts as covert amalgam clefts in the sense of Declerck (1988). An amalgam cleft is a specificational copular sentence in which the post-copular constituent is a full clause rather than a single phrase, as

analysis does not directly account for the contrast highlighted by Sadakane & Koizumi (1995) between NP clefting and PP clefting: while clefting of case-marked NPs tends to be unacceptable, clefting of postpositional phrases is fully grammatical. (24–25) are from Sadakane & Koizumi, and the grammaticality judgments are theirs.

(24) **NP clefting**

a. *[kinoo piza o tabe-ta no] wa [$_{NP}$ Mary ga] da.
yesterday piza ACC eat-PST NMLZ TOP Mary NOM COP
'It is Mary who ate pizza yesterday.'

b. ??[kinoo Mary ga tabe-ta no] wa [$_{NP}$ piza o] da.
yesterday Mary NOM eat-PST NMLZ TOP piza ACC COP
'It is pizza that Mary ate yesterday.'

(25) **PP clefting**

a. [John ga tegami o mora-tta no] wa [$_{PP}$ Mary kara] da.
John NOM letter ACC receive-PST NMLZ TOP Mary from COP
'It is from Mary that John received a letter.'

b. [Mary ga keeki o ki-tta no] wa [$_{PP}$ kono naifu de] da.
Mary NOM cake ACC cut-PST NMLZ TOP this knife with COP
'It is with this knife that Mary cut the cake.'

This contrast raises an important puzzle for both the movement approach (Hoji 1987; 1990; Hiraiwa & Ishihara 2002) and the non-movement approach (Hoji & Ueyama 2003; Cho et al. 2008). Two core empirical questions arise:

(i) Why does the presence of a floated numeral quantifier make clefted case-marked NPs fully acceptable?

(ii) Why are NP clefts degraded in acceptability, while PP clefts remain fully acceptable?

At this point, it is worth noting that the contrast between bare case-marked NPs and structurally richer nominals (those containing a numeral quantifier) suggests an internal "size" hierarchy within the DP domain, in contrast to PPs, which uniformly behave as maximal categories. This issue will be taken up again in section 3.5, where it is argued that the relevant patterns follow from the Maximality Principle proposed by Rizzi (2016).

The discussion that follows adopts Hiraiwa & Ishihara's (2002) analysis, under which the clefted constituent moves to Spec,FocP in the CP left periphery. On a Kaynean LCA-based

in *What Yumi bought is [she bought a LEXUS]*. Since the post-copular position hosts a full clause, amalgam clefts can contain multiple foci, as in *What Yumi bought is [YUMI bought a LEXUS in LA]*. On their analysis, CM clefts in Japanese and Korean are derived by gapping of the post-copular clause, in which the verb and non-focused constituents are elided under identity with the presuppositional clause, leaving only the focused element(s).

head-initial approach, the analysis targets the empirical problems outlined above as well as a wider range of issues surrounding case-marked clefts in the literature.

3.2 Left peripheral CP in Japanese

Hiraiwa & Ishihara (2002) argue that CM clefts (which they simply call clefts) are derivationally linked to the so-called *no-da* in-situ focus construction. The *no-da* construction is a common and distinctive grammatical pattern in Japanese, formed by combining the nominalizer particle *no* (realized as *n* in its reduced form) with the copula *da* (or its polite form *desu*), as in (26).

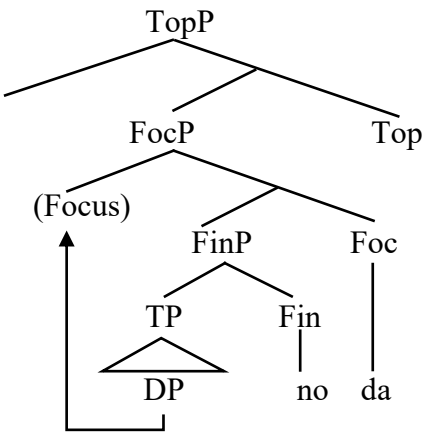
(26) [Yumi ga hon o san-satu ka-tta n(o)] da/desu.
Yumi NOM book ACC three-CL buy-PST NMLZ COP
'(It is that) Yumi bought three books.'

Its connection to clefts, according to Hiraiwa & Ishihara, is both structural and functional. In the context of focus constructions, the *no-da* construction allows any phrase within the clause to be interpreted as focused, depending on which element receives prosodic (intonational) prominence. The *no-da* construction shares syntactic properties with clefts, such as resisting certain substitutions and case alternations. They argue that these constructions can be unified under a single syntactic framework, specifically Rizzi's (1997) articulated CP structure:

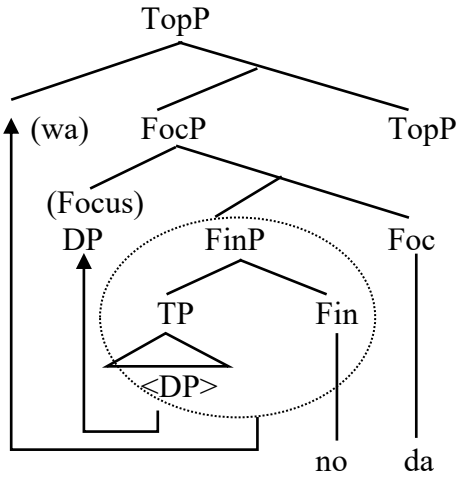
(27) [TopP … [FocP… [FinP… [TP…]]]]

They propose that cleft constructions are derived from the *no-da* in-situ focus construction via two syntactic operations: (i) Focus Movement: The focused element moves to Spec, FocP, headed by the copula *da*. (ii) Remnant topicalization: The remaining Fin(ite)P is topicalized and moved to Spec, TopP, marked by *-wa*. The derivations proposed by Hiraiwa & Ishihara are given in (28a–b).

(28) a. Focus Movement



b. Remnant Topicalization



The important theoretical consequence of Hiraiwa & Ishihara's analysis of cleft constructions is that Japanese clefts are monoclausal, derived by focus movement and subsequent remnant topicalization. This analysis of Japanese clefts as monoclausal focus constructions is central to the argument developed here. I nevertheless retain the term *cleft* for expository continuity.

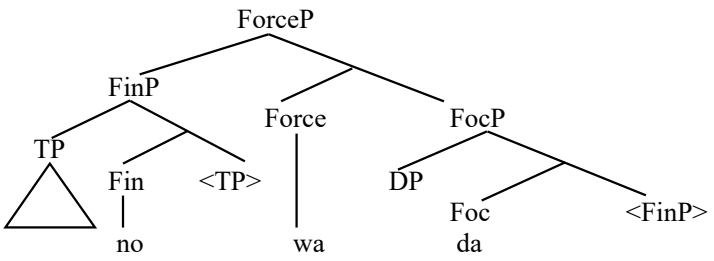
Although Hiraiwa & Ishihara (2002) provide a theoretically elegant account of cleft constructions, their treatment of multiple clefts is problematic. They assume that multiple clefts involve successive movements of focused elements to Spec,FocP. However, the empirical consensus is that a clause cannot host more than one structurally represented focus. Indeed, Rizzi (1997) argues that while TopP is cyclically recursive, FocP is not.

By contrast, Koizumi (2000) proposes that multiple clefts are instances of VP-clefting. On his analysis, the verb first raises out of the VP within the presuppositional clause. Once the verb has vacated the VP, the remnant VP—now containing multiple focused elements—undergoes clefting as a single constituent.[7] This approach is in fact compatible with Hiraiwa & Ishihara's broader analysis of clefts. A plausible derivation is that the elements in the presuppositional clause scramble out of VP/*v*P and adjoin to TP, after which the remnant VP/*v*P raises to Spec,FocP. On this view, the clefted focus constituents are interpreted within VP on a par with *no-da* constructions.

3.3 The head-initial CP structure

This section adopts Kayne's (1994) LCA-based hypothesis as the theoretical foundation for the account of single-constituent clefts outlined in Section 3.1, situating the discussion within the parameter hierarchy framework. Combining the FOFC compliant roll-up hypothesis with Hiraiwa & Ishihara's proposal for cleft constructions, this section hypothesizes that the left-peripheral CP has the structure shown in (29) and argues that this reanalysis offers a more theoretically precise account of empirical issues in Japanese cleft constructions. A head-initial reanalysis of a head-final CP like (28) more or less entails inverting the positions of the relevant heads posited by Hiraiwa & Ishihara, while relocating the landing site of FinP from TopP to ForceP.

(29) Cleft Structure



[7] Koizumi (2000) uses cleft constructions to argue for verb raising rather than to develop a general analysis of clefts, and does not provide a separate analysis of single-constituent clefts.

Within Rizzi's layered CP, the obligatory core encodes clause type and finiteness—namely, ForceP and FinP—whereas the discourse-related projections TopP and FocP are optional and are projected only when required by discourse conditions. Since topicalization typically involves optional movement, the complement-to-specifier movement within the articulated CP layer proposed by Hiraiwa & Ishihara is better analyzed not as topicalization, but as a more general type of roll-up movement triggered by Force^0, yielding a surface head-final CP, as illustrated in (29). On this view, Fin^0 bears a roll-up feature that triggers movement of TP to Spec,FinP. The focused DP then moves to Spec,FocP, headed by the copula *da*. Force^0, in turn, triggers movement of FinP to Spec,ForceP.[8]

Although roll-up is typically defined as complement-to-specifier movement within a single projection, I treat the movement of FinP to Spec,ForceP as a form of roll-up. This requires an extension of the canonical notion of roll-up. In Roberts's canonical cases, the operation is confined to a single projection. In an articulated domain, however, the relevant dependency between Force^0 and Fin^0 is computed over the obligatory core of that domain. In the cleft construction, the intervening FocP does not disrupt the dependency between the two core heads. The successive movements of TP to Spec,FinP and of FinP to Spec,ForceP are therefore analyzed as what I call *core-spine roll-up*: a sequence of complement-to-specifier movements that relates the lower and higher heads of the obligatory core spine. This operation is formally motivated by the labeling requirement of ForceP and derives the order $[_{\text{ForceP}}$ FinP Force^0 $<$FinP$>$].

Heycock (2008) argues that Japanese *wa* cannot be treated as a uniformly topical morpheme: non-contrastive topics, contrastive topics, and "categorical judgments" in the sense of Kuroda (2005) all show that the interpretation of *wa* depends on clausal structure, predicate type, and discourse configuration. Building on this view, I assume that *wa* is not lexically specified as "topic." *Wa* acquires different readings depending on which left-peripheral head it realizes in a split CP. In a simple sentence such as "As for John, he bought a book," where *wa* appears with the subject, *wa* can be analyzed as realizing an ordinary Topic head and marking John as an aboutness topic. In contrast, in a Japanese counterpart of the English cleft "It was the book that John bought" (literally, a structure of the form [John bought] WA the book COP), the *wa*-marked clause does not function as an aboutness topic; rather, it serves as an operator in the highest Force-related projection which types the clause as a specificational (identificational) sentence. Such clefts presuppose an implicit question—roughly, "what did he buy?"—and the cleft asserts

[8] As one of the reviewers points out, the extraction of a DP from a TP that has moved to Spec,FinP can indeed be characterized as counter-cyclic. I accept this consequence of my analysis. A comparable pattern has been reported in connection with wh-movement: after a wh-phrase reaches an intermediate CP edge, one of its constituents may undergo further movement while other material is stranded at that edge (Davis 2020). Nevertheless, the extraction proposed here is compatible with the phase-based architecture I adopt, since phase edges remain accessible under the Phase Impenetrability Condition.

an exhaustive identification of the value of that variable ("the thing John bought is the book"). On the present proposal, different structural positions in the left periphery correspond to different clause types and hence to the various interpretations associated with *wa*.

In contrast, in Hiraiwa & Ishihara's head-final configuration in (28b), no structural position is available for the topic *wa*: if the presuppositional clause were to move to Spec,TopP, the only remaining option would be to insert *wa* at PF. As discussed in more detail in section 3.4, however, this strategy is problematic.

Moreover, while non-finite copular particles with focus functions are widely attested cross-linguistically, the Japanese copula *da* is finite, as shown by the existence of the past-tense counterpart *da-tta 'be-past'*, which suggests that the copula originates inside TP. Movement of the copula to the head of FocP thus involves head movement and is subject to the Head Movement Constraint (HMC). Under the head-final structure in (28b) assumed by Hiraiwa & Ishihara, the presence of the nominalizer *no* in Fin^0 blocks this movement: moving the copula directly from T^0 to Foc^0, bypassing Fin^0, would violate the HMC. By contrast, in the proposed head-initial structure in (29), roll-up of TP to Spec,FinP precedes the head movement of *da* to Foc^0, so the derivation never bypasses any head and thus does not incur an HMC violation.

A problematic case that requires explanation is how to derive a clause-final complementizer. Specifically, the cleft construction in (29) appears embedded under a matrix verb taking the complementizer *to* or the Q-particle *ka,* as in (30).

(30) a. [[[Yumi ga ka-tta no] wa kono hon-ø da] **to**] omou.
Yumi NOM buy-PST NMLZ TOP this book COP COMP think
'(I) think that it is this book that Yumi bought.'

b. [[[Yumi ga ka-tta no] wa kono hon-ø da] **ka**] sira-nai.
Yumi NOM buy-PST NMLZ TOP this book COP Q know-not
'(I) don't know if it is this book that Yumi bought.'

In Rizzi's CP-layer system, the complementizer is assumed to occupy $Force^0$. It is therefore tempting to reserve $Force^0$ for *to/ka* in (30), in which case the categorial status of clefts would be TopP, as in Hiraiwa & Ishihara (2002), and TopP would then undergo roll-up movement to Spec,ForceP, yielding clause-final *to/ka,* as shown in (31).

(31) $[_{\text{ForceP}}\ [_{\text{TopP}}$ FinP wa XP da $<$FinP$>]$ $[_{\text{Force}}$ **to/ka**$]$ $<$TopP$>$ $]$

Note that if the Top-FocP field is not projected, monotonic complement-to-specifier stacking established lower in the clause is FOFC-compliant, and result in the consistently harmonic head-final structure within the CP layer. However, the configuration in (31), which corresponds to the final-over-initial configuration in (3d), is excluded by FOFC. Now how can we resolve this

problem? An important point is that, in the system of Biberauer et al. (2014), the FOFC constrains only the order of heads within a single extended projection (from V up through *v*, T, and C), all of which share the categorial features of the lexical head V.

Fukui (1986; 1995) argues that clause-final *ka/to* are not Cs: *ka* is a noun bearing the feature [+Q], and *to* is a simple postposition. According to Fukui, the difference in categorial status between these particles is evidenced by their (non-)cooccurrence with the case particle (Fukui 1995: 117–119): *ka* can be followed by a case particle, whereas *to* cannot.

(32) a. boku wa [[John ga nani o ka-tta] ka] **o** siri-tai.
I TOP John NOM what ACC buy-PST Q ACC know-want
'I want to know what John bought.

b. *boku wa [[John ga hon o ka-tta] to] **o** kangae-ta.
I TOP John NOM book ACC buy-PST COMP ACC think-PST
'I think that John bought the book.'

Given the standard assumption in generative grammar that case is assigned to nominal phrases but not to clauses, Fukui (1986; 1995) takes *ka* to be a nominal element. If this view is correct, then both *ka* and *to* are located outside the extended verbal projection and therefore lie beyond the domain of application of the FOFC. While Fukui maintains that Japanese lacks C-elements, subsequent research that assumes the existence of functional categories in Japanese has generally treated *ka/to* as C-elements. In the head-initial framework adopted in this paper, however, the question arises of how the FOFC violation just discussed can be avoided. The aim of the present paper is not to settle this issue, but rather to put forward some possible proposals that might address it.

Recall, as stated in section 2.2, that Mandarin Chinese (6), repeated in (33), has a basic SVO order but also clause-final question particles such as *ma*, which would lead to a FOFC violation. This type of clause-final particle provides a prominent class of potential counterexamples to the FOFC, as discussed by Biberauer et al. (2014).

(33) Hongjian xihuan zhe ben shu **ma**?
Hongjian like this CL book Q
'Does Hongjian like this book?'

They argue that clause-final particles marking mood, discourse status, interrogation do not bear the categorial feature V and therefore fall outside the verbal extended projection and the scope of FOFC. They further suggest that Q-particles pattern like coordinating conjunctions in being "acategorial" elements, stating that "apparently FOFC-violating final question particles may actually be initial negative disjunctions of an elided disjunct clause. The structure of these yes/no questions would be [TP [OR-NOT ~~TP~~]], where ellipsis of the second TP, identical with the first TP, leaves the negative disjunction as an apparently clause-final particle" (Biberauer et al. 2014: 201).

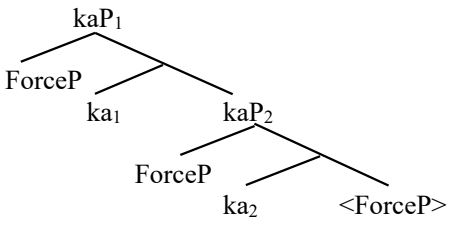
A related issue is addressed by Mitrović (2014; 2021), who develops the notion of "superparticles" as morpho-syntactic atoms that encode complex logical meanings (e.g. polarity, disjunction, scalarity). These elements systematically recur across different environments such as coordination, interrogatives, and focus constructions. He treats Japanese *ka* as a κ-type superparticle which uniformly underlies its disjunctive (such as *John ka Mary (ka)*), interrogative, and indefinite uses, rejecting a homophony-based analysis. Building on den Dikken's (2006) work on Junction Phrase (JunctionP), he argues that coordination is mediated by a dedicated functional projection hosting a junction head (J).

On this view, what surface as "particles" are instances of such junction heads or related functional material, whose distribution reflects properties of the coordination structure. In fact, the cleft focus construction can be conjoined and embedded under a matrix verb selected by the complementizer *ka*, in which case *ka* conveys both interrogative and disjunctive interpretations.

(34) Yumi ga ka-tta no wa ringo da **ka** (Yumi ga ka-tta no wa)
Yumi NOM buy-PST NMLZ TOP apple COP Q/or Yumi NOM buy-PST NMLZ TOP
mikan da **ka** sira-nai
orange COP Q/or know-not
'(I) don't know whether what Yumi bought was apples or (what Yumi bought was) orange.'

In (34), the parenthetical presuppositional clause in the second disjunct is identical to its counterpart in the first disjunct and can therefore undergo PF deletion under identity. Japanese clausal complements with *ka* instantiate a structure of the form [$_{\text{kaP1}}$ ForceP [**ka** [$_{\text{kaP2}}$ ForceP [**ka** [ForceP]]]]], where kaP corresponds to a JunctionP in the sense of Mitrović (2021), hosting an acategorial junction or superparticle head. In this configuration, I take the lowest ForceP embedded under ka$_2$ to undergo roll-up to its specifier position, yielding the surface order in which the particle *ka* appears clause-final, and the matrix verb *siru* 'know' in (35) selects this junction phrase kaP in (35).

(35)



This analysis captures the disjunctive flavor of *ka* while leaving *ka* outside ForceP, thereby maintaining that the derivation is FOFC-compliant. Empirically, the particle *to* can also be acategorial, acting as a conjunctive coordinator that can conjoin DPs (e.g. John *to* Mary), and, according to Frellesvig (2001), both its conjunctive use and its complementizer use are historically related and derived from the copula, whose primary function is to link phrases and clauses. This seems to suggest that the core status of *to* is likewise that of a junction element, rather than that

of a purely C-related complementizer head in the extended verbal projection. A more detailed theoretical and empirical discussion of the status of *to* is left for future research.

In what follows, I adopt the structure proposed in (29) for cleft focus constructions, and Section 3.4 develops the discussion on this basis to address the empirical issues associated with cleft constructions.

3.4 Nominative/genitive conversion

Nominative-genitive conversion (NGC) in Japanese—also known as *ga-no conversion*—is a well-studied phenomenon in Japanese syntax. In certain adnominal environments, such as relative clauses (36a) and noun-complement clauses headed by *koto* 'fact' (36b), the nominative case marker *ga* on a subject can alternate with the genitive marker *no*.

(36) a. [kino John ga/no ka-tta] hon
yesterday John NOM/GEN buy-PST book
'the book which John bought yesterday'

b. John wa [$_{CP}$ kinoo Mary ga/no ki-ta koto] wo sira-naka-tta.
John TOP yesterday Mary NOM/GEN come-PST FN ACC know-NEG-PST
'John didn't know that Mary came yesterday.'

Miyagawa (2001; 2011) proposes a D-licensing analysis for nominative–genitive conversion: the genitive subject is licensed by the D-head of the host DP: in a prenominal clause, the subject gets nominative when T probes, but genitive when the higher D probes and assigns Case to the subject from outside the clause (an ECM/raising-style configuration). Hiraiwa (2002) argues that Miyagawa's claim is empirically and theoretically inadequate: Genitive subjects occur in many adnominal subordinate clauses without any external D at all, so D cannot be the licenser in those cases. Examples (37) are taken from Hiraiwa (2002).

(37) a. [sen-getsu ikkai denwa **no** a-tta **kiri**] John kara nani-mo renraku
last-month once call GEN BE-PST.ADN since John from any-FOC call
ga nai.
NOM not
'There has been no call from John since he called me up once last month.'

b. John wa [ame **no** yamu **made**] office ni i-ta.
John TOP rain GEN stop.ADN until office at be-PST
'John was at his office until the rain stopped.'

c. [boku **no** omou **ni**] John wa Mary ga suki ni tigainai
I GEN think.ADN as John TOP Mary NOM like DAT must
'As far as I think, John must like Mary.'

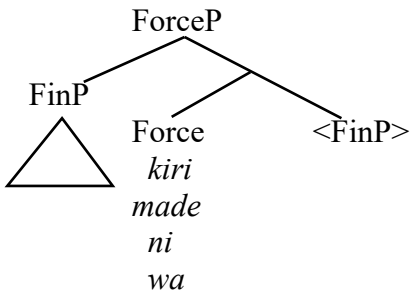
d. John wa [Mary **no** yon-da **yori**] takusan no hon o yon-da.
John TOP Mary GEN read-PST.ADN than many GEN book ACC read-PST
'John read more books than Mary did.' (Watanabe 1996: 396)

Hiraiwa proposes that genitive subjects in nominative–genitive conversion arise when the clause has a special inflectional configuration on the verbal spine, corresponding to the classical *adnominal* clause in older Japanese. This configuration is represented as a head amalgam V–*v*–T–C, and it is the highest head C in this amalgam that licenses genitive Case on the subject. More concretely, when an affixal C is merged, it forms a V–*v*–T–C complex via AGREE or head movement. Hiraiwa, in fact, gives the following non-cleft example in which *wa* licenses a genitive subject.

(38) [John **ga/no** kuru to ko-nai to de **wa**] oo-chigai da.
John NOM/GEN come and come-not and COP TOP great-difference COP
'It makes a great difference whether John comes or not.'

Note that the subordinate clauses in (37–38) do not project to the Topic-Focus field, and thus the roll-up of FinP to Spec,ForceP is FOFC compliant. The subordinators *kiri* 'since', *made* 'until', *ni* 'when/as', and *wa* reside in Force^0 as in (39), and the genitive subject in these clauses is licensed by these particles. (39) follows from Hiraiwa's (2002) C-licensing analysis.

(39)



Hiraiwa & Ishihara (2002), however, observe that *no-da* constructions disallow a genitive subject as shown in (40), suggesting that the nominalizer *no* in Fin^0 does not license the genitive subject.

(40) a. [John ga/*no kono ringo o tabe-ta **no**] da.
John NOM/GEN this apple ACC eat-PST NMLZ COP
'(It is that) Taro ate this apple.'

b. [John ga/*no koen de hasi-tta **no**] da.
John NOM/GEN park LOC run-PST NMLZ COP
'(It is that) John run at the park.'

According to Hiraiwa & Ishihara,CM-clefts and non-CM clefts diverge in their compatibility with NGC. Specifically, genitive subjects are degraded in CM-clefts (41a) but acceptable in non-CM clefts (41b):

(41) a. [Taro ga/??no tabe-ta **no**] wa kono ringo o 3-tu da.
Taro NOM/GEN eat-PST NMLZ TOP this apple ACC 3-CL COP
'It is three of these apples that Taro ate.'

b. [Taro-ga/no tabe-ta **no**] wa kono ringo-ø 3-tu da.
Taro-NOM/GEN eat-PST NMLZ TOP this apple-ø 3-CL COP

They take this as evidence for a derivational relation between CM-clefts and *no-da* constructions. Yet a puzzle remains: while genitive subjects are entirely excluded in (40), they are at least marginally possible in CM-clefts (41a).[9] If CM-clefts are derived from *no-da* constructions, what accounts for this sharp contrast?

The above observation supports the view that *wa* in Force^0 is a licenser for genitive subjects. The *no-da* construction lacks such a licensing head. This contrast, in turn, highlights a problem for the head-final cleft structure (28b) proposed by Hiraiwa & Ishihara: since the articulated CP has no position for *wa*, it incorrectly predicts that both clefts and *no-da* constructions should equally reject genitive subjects.

3.5 Articulated DP structure and placement of numerals

As noted in Section 2.4, Fukui (1986; 1995) develops the Functional Parametrization Hypothesis, according to which cross-linguistic variation is determined by the inventory of functional categories and features available in a given language. Japanese nominals have the following properties:

(42) a. Article-type determiners are unattested.
b. Number agreement does not occur.
c. Prenominal elements are not subject to rigid ordering restrictions.

Unlike English, which exhibits number agreement between demonstratives and nouns, as in *this book* versus *these books*, Japanese does not exhibit such agreement. Moreover, while English allows *this red book* and *my red book*, but not *red this book* or *red my book*, Japanese allows demonstratives and genitives to precede or follow adjectives. On the basis of these observations, Fukui argues that Japanese lacks both a DP layer and φ-agreement features, with demonstratives and genitives functioning as NP-internal modifiers.

Watanabe (2008), in contrast, contends that Japanese nominal phrases are indeed DPs with a richly articulated functional structure consisting of #P, CaseP, QP, and DP.

(43) [DP…[QP…[CaseP…[#P …[NP…]]]]]

[9] In my judgment, (41a) and (41b) do not differ in acceptability.

With this multilayered DP, Watanabe attempts to account for the different word orders within DP. In (44a) the numeral quantifier appears between the noun and case marker In (44b), it precedes the noun, and in (44c) it follows the case-marked NP.

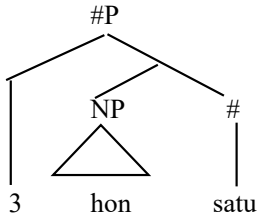
(44) a. hon **san-satu** o ka-tta
book 3-CL ACC buy-PST

b. **san-satu** hon o ka-tta
3-CL book ACC buy-PST
'bought three books'

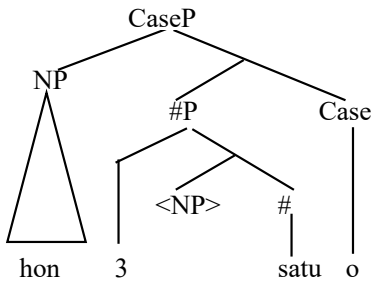
c. hon o **san-satu** ka-tta
book ACC 3-CL buy-PST

In Watanabe's proposal, agreement features involving number, case, and specificity on D trigger movements of the lower projections, which account for the surface distribution of numeral quantifiers (NQs). The alternations illustrated in (44) are captured through successive remnant movements. The lowest layer, #P, has a structure whose head is the locus of classifiers, as in (45a). (44a) is derived by movement of NP to Spec,CaseP headed by *o* in (45b), and (44b) is derived by movement of #P to Spec,QP (45c). Finally, (44c) is derived by movement of CaseP to Spec,DP (45d) (Watanabe 2008: 517).

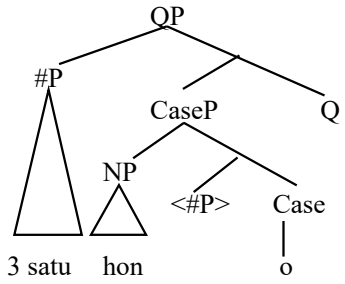
(45) a.



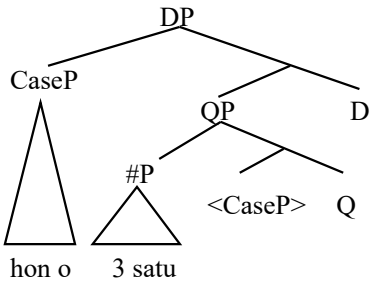
b.



c.



d.

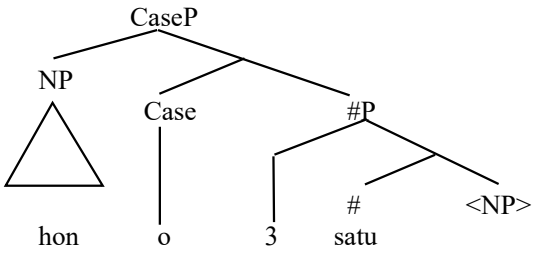


Watanabe's derivations involve stepwise complement-to-specifier movements within DP. Such remnant movements, similar to those posited in Hiraiwa & Ishihara's analysis of layered CP, are proposed in Watanabe's account to be driven by agreement features on functional heads. However, if, as Fukui (1986; 1995) argues, the core syntactic properties of Japanese nominal structure

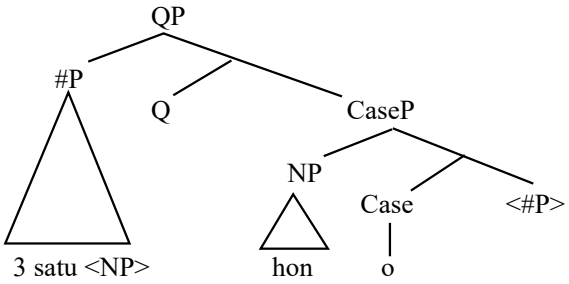
summarized in (42) follow from the absence of agreement features, then the complement-to-specifier movements within the DP layer, parallel to those in the CP layer, are better analyzed not as agreement-driven movement but as instances of roll-up triggered by labeling requirements associated with functional heads in an articulated structure.

Watanabe's derivations can now be reproduced in a head-initial configuration, as shown in (46–48).

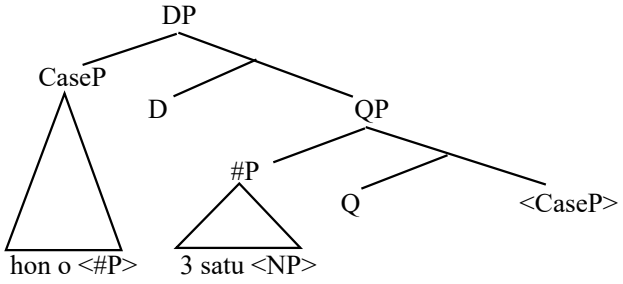
(46) NP movement to Spec,CaseP



(47) #P movement to Spec,QP



(48) CaseP movement to Spec,DP



Note that (46) appears to correspond to the surface word order in (44c), but the NP and the case marker *o* at this stage do not form a constituent. The sequence *hon o* 'book ACC' forms a constituent only after #P moves out of CaseP in (47). CaseP then moves to Spec,DP (48).[10]

[10] There are a number of pieces of evidence that the NP and the case particle form a constituent. They can be moved out of the DP, as in (i).

(i) $[_{CaseP}$ hon o$]_i$ boku ga kinoo $[_{DP}$ t_i sansatu] ka-tta.

As noted in Section 3.3, roll-up in the articulated CP domain is computed over the obligatory core spine defined by Force^0 and Fin^0, even when discourse-related projections intervene between them. This core-spine roll-up is independently motivated in the nominal domain. Following Watanabe's (2008) articulated structure of Japanese DP, I take D^0 and Case^0 to define the obligatory core spine of the nominal domain. The successive movements of NP to Spec,CaseP and of CaseP to Spec,DP are therefore analyzed as *core-spine roll-up*: a sequence of complement-to-specifier movements that relates the lower and higher heads of the obligatory core spine. The intermediate movement of #P to Spec,QP is an optional instance of roll-up, arising when a numeral is overtly realized. The CP and DP derivations thus instantiate the same generalization: in a head-final system, core-spine roll-up relates the lower and higher heads of the obligatory core, while additional articulated projections may host independently motivated instances of roll-up.

This multilayered structure of the DP provides a straightforward account of the empirical puzzles introduced in section 3.1, namely: (i) why the presence of a floated numeral quantifier ameliorates the acceptability of clefted case-marked NPs, and (ii) why NP clefts are degraded while PP clefts are fully acceptable. Consider the relevant examples in (49), as discussed in section 3.1.

(49) a. ?*[Yumi ga kino ka-tta no] wa hon o da.
Yumi NOM yesterday buy-PST NMLZ TOP book ACC COP
'It is the book that Yumi bought yesterday.'

b. [Yumi ga kino ka-tta no] wa hon o san-satu da.
Yumi NOM yesterday buy-PST NMLZ TOP book ACC three-CL COP
'It is three books that Yumi bought yesterday.'

c. [Yumi ga kino hon o ka-tta no] wa kono mise de da.
Yumi NOM yesterday book ACC buy-PST NMLZ TOP this store in COP
'It is in this store that Yumi bought a book yesterday.'

Under Watanabe's (2008) articulated DP structure, the restriction on clefting in (49) follows from Rizzi's (2016) Maximality Principle, grounded in Chomsky's (2013) labelling theory, which requires that only maximal objects with a given label be eligible for movement:

(50) **Maximality Principle** (Rizzi 2016: 116)
Only maximal objects with a given label can be moved.

---

book ACC I NOM yesterday three buy-PST
'(Lit.) The book I bought three yesterday.'

The movement of CaseP out of DP is possible since, under Chomsky's (2001) PIC, the specifier position at the phase edge remains accessible for movement. Furthermore, the order in (44a) may be derived by roll-up movement of #P to Spec,CaseP, with NP subsequently raising to Spec,QP.

The Maximality effect is induced by criterial Spec-head feature sharing. In clefts, focus movement creates a criterial configuration in Spec,FocP: the moved XP enters into a local Spec–head relation with Foc^0, and the resulting configuration is labelled through the focus feature shared by the XP and Foc^0. In (49a), CaseP is not the maximal projection in the nominal domain, but an intermediate projection embedded within DP. It is therefore not an eligible movement target for the criterial Spec,FocP position. By contrast, in (49b), once CaseP has raised to Spec,DP, the NQ constituent is a maximal DP; the PP in (49c) is likewise maximal. Both are therefore eligible for focus movement. The same principle also derives the familiar effect of "criterial freezing" (Rizzi 2006; Rizzi & Shlonsky 2007): once an XP occupies a criterial position and satisfies the relevant criterion, it cannot undergo further movement. For further consequences of the interaction between clefting and the Maximality Principle, see Section 3.6.

Note that clefting differs from long-distance scrambling with respect to the Maximality Principle: CaseP can freely undergo long-distance A′-scrambling.

(51) a. **sono hon o** boku wa [Yumi ga ka-tta] to omo-tta.
that book ACC I TOP Yumi NOM buy-PST COMP think-PST
'That book, I thought Yumi bought.'

b. **sono mise de** boku wa [Yumi ga hon o ka-tta] to omo-tta.
that store LOC I TOP Yumi NOM book ACC buy-PST COMP think-PST
At that store, I thought Yumi bought the book.'

Saito (1985; 1989; 1992) establishes that, whereas clause-internal scrambling in Japanese is neither A-movement nor A′-movement (non-A/non-A′), long-distance scrambling across a finite CP boundary is unambiguously A′-movement, as diagnosed by Weak Crossover effects and Condition C reconstruction. Crucially, however, under the traditional Saito-style analysis, long-distance A′-scrambling differs from wh-movement in that it is not attracted to a criterial functional Spec through feature sharing. Rather, the scrambled phrase occupies a TP-adjoined, non-criterial A′-position rather than entering into a local Spec–head configuration with a matrix C/Foc head. Because it lacks the feature-sharing criterial configuration in which the relevant Maximality effect arises, long-distance scrambling does not yield that effect.[11]

In sum, Watanabe's finely layered DP hypothesis, Hiraiwa & Ishihara's analysis of clefting as monoclausal focus movement, and the Maximality Principle together provide a principled explanation for the puzzles raised earlier. This unified account captures the complex distribution

[11] Adjunction here is used in the traditional descriptive sense of a non-criterial A′-position, following the analysis of long-distance scrambling developed by Saito (1985; 1989; 1992). I do not assume that this terminology by itself settles how such positions should be implemented in a strictly Kaynean antisymmetric syntax. The compatibility of scrambling with a fully LCA-compliant theory of phrase structure is an independent issue. Nothing in the present argument depends on a particular implementation of long-distance scrambling; the crucial point is that the movement in (51) does not target a criterial Spec–head configuration involving feature sharing with a matrix C/Foc head.

of numeral quantifiers, clarifies the asymmetry between NP and PP clefts, and explains the difference between long-distance A′-scrambling and focus movement in clefts.

3.6 Negative polarity items

Another important issue in the literature on Japanese clefts concerns their interaction with negative polarity items (NPIs), in particular XP-*mo* expressions. Such expressions are subject to a clause-mate licensing condition, as illustrated by the contrast in (52). In (52a), the XP-*mo* expression and the negative predicate occur in the same embedded clause. In (52b), by contrast, the XP-*mo* expression occurs in the embedded clause, whereas negation is located in the matrix clause. The latter sentence is consequently unacceptable.

(52) a. boku wa [John ga paatii de **biiru o ip-pon mo** noma-**naka**-tta]
I TOP John NOM party LOC beer ACC one-CL FOC drink-not-PST
to omo-tta.
COMP think-PST
'I thought John didn't drink even one bottle of beer at the party.'

b. *boku wa [John ga paati de **biiru o ip-pon mo** non-da] to
I TOP John NOM party LOC beer ACC one-CL FOC drink-PST COMP
omowa-**naka**-tta.
think-not-PST
'I did not think that John drank even one bottle of beer at the party.'

Hoji & Ueyama (2003) observe that an XP-*mo* expression licensed by negation inside an embedded clause can undergo long-distance A′-scrambling, as in (53a), but cannot occur as the focused constituent of a case-marked cleft, as in (53b).

(53) a. **biiru o ip-pon mo** boku wa [John ga paatii de noma-**naka**-tta]
beer ACC one-CL FOC I TOP John NOM party LOC drink-not-PST
to omo-tta.
COMP think-PST
'Even one bottle of beer, I thought John didn't drink at the party.'

b. *[John ga paati de noma-**naka**-tta no] wa **biiru o ip-pon mo** da.
John NOM party LOC drink-not-PST NMLZ TOP beer ACC one-CL FOC COP
'It is even one bottle of beer that John did not drink at the party.'

Hoji & Ueyama take the pattern in (53b) to argue that case-marked clefts involve movement of a null operator within the presuppositional *no*-clause, as schematized in [[OP$_i$ … t$_i$ … no] wa FOCUS$_i$ da]. Their null-operator analysis differs from the analysis originally proposed by Hoji (1987; 1990), under which the focused phrase itself moves directly. On Hoji & Ueyama's analysis,

the unacceptability of (53b) follows from the clause-mate condition: the NPI and negation do not occur in the same clause.

The contrast in (53) is central to the present analysis. As discussed in Section 3.3, I adopt the view that cleft-focus constructions are monoclausal, derived by focus movement followed by remnant FinP movement, as in the structure in (29). Both examples contain an NPI that is licensed by sentential negation in the same clause and therefore cannot be reduced to the clause-mate requirement on NPI licensing. What differs, however, is the type of movement: in (53a), the NPI undergoes long-distance scrambling, whereas in (53b) it undergoes feature-driven movement to the criterial focus position of the cleft. I argue that this difference follows from the Maximality Principle formulated in (50).

Following Watanabe (2004), the analysis pursued here treats Japanese XP-***mo*** expressions as negative concord items (NCIs), rather than as English-type NPIs. Watanabe (2004) proposes that Japanese NCIs such as *nani-mo* bear an uninterpretable focus feature [*u*Focus], morphologically realized by *mo*, in addition to their inherent negative feature. While the negative feature of the NCI is licensed through Agree with Neg^0, the [*u*Focus] feature must likewise be checked and eliminated through its relation to a Focus feature in the negative domain. Crucially, Watanabe claims that this licensing does not involve overt movement of the NCI to Spec,NegP: Neg^0 lacks an EPP feature in Japanese, and the NCI may therefore remain in situ:[12]

(54) $[_{NegP}\ [_{VP}$ nani-mo [Neg/Focus] V] Neg^0 [Neg]]

This allows the NCI in (53a) to be licensed through Agree with Neg^0 within the embedded clause and subsequently undergo long-distance scrambling. Under an adjunction analysis of long-distance scrambling, the scrambled XP is merged in a TP-adjoined position rather than in the specifier of a functional head. Consequently, because it does not form the Spec–head configuration associated with criterial movement, the Maximality Principle is not implicated. The cleft in (53b), by contrast, involves feature-driven focus movement to Spec,FocP and is therefore subject to the Maximality Principle.

This principle captures the core phenomenon known as “criterial freezing” (Rizzi 2006; Rizzi & Shlonsky 2007). The logic can be illustrated with a wh-phrase in Spec,Q. A wh-phrase and Q^0 share a Q-feature, so the configuration in (55a) is labeled Q and constitutes a maximal Q-projection. Further movement of the wh-phrase alone, as in (55b), extracts a non-maximal subpart of the Q-labeled constituent and therefore violates the Maximality Principle.

---

[12] Watanabe (2004) proposes that for languages like West Flemish, Neg^0 does have an EPP feature, so once Agree applies, the NCI is forced to move to Spec, NegP. So his view is that NCI-Neg relations are handled uniformly by Agree, but whether NCIs move to Spec,NegP depends on the presence of an EPP feature on Neg^0, which is language-specific.

(55) a. $[_{\alpha}$ [which$_{Q}$ book]$_{i}$ [Q^{0} [Bill read t$_{i}$]]]
α = maximal Q-projection

b. *$[_{\beta}$ [which$_{Q}$ book]$_{i}$ [Q^{0} [John wonders $[_{\alpha}$ _ Q^{0} [Bill read t$_{i}$]]]]]

I extend this reasoning to the NCI facts in (53b). The proposal is not that every NCI in Japanese must raise to Spec,NegP; as noted above, ordinary NCI–Neg licensing can be established in situ by Agree. Rather, I propose that when an NCI is targeted by feature-driven focus movement, its [Neg] and [uFocus] requirements must first be resolved in a local configuration with Neg^{0}. In this configuration, XP-mo and Neg^{0} share the relevant negative and focus-related features, thereby forming a maximal Neg/Foc projection. Movement of XP-mo alone to Spec,FocP would then extract a non-maximal subpart of this projection and thus violate the Maximality Principle, just as further wh-movement out of the Q-labeled configuration in (55b) does. The NCI is consequently frozen in the Neg-related configuration and cannot function as the cleft pivot.[13]

The contrast between (53a) and (53b) thus follows. In (53a), the NCI is licensed in situ by Agree with Neg^{0} and then undergoes non-feature-driven long-distance scrambling. Since it does not enter into a Spec-head relation with Neg^{0}, no labeling-based freezing effect arises. In (53b), by contrast, feature-driven focus movement requires the NCI to enter the local Neg/Foc configuration, and the Maximality Principle blocks extraction of the NCI alone from that configuration to Spec,FocP. On this view, the additional NegP configuration is required by feature-driven focus movement, not by NCI licensing alone. The inability of XP-*mo* to appear as a cleft focus is thus compatible with the monoclausal analysis of Japanese clefts and follows from the distinction between long-distance A′-scrambling and feature-driven focus movement.

To summarize section 3, Japanese clefts are analyzed as a monoclausal focus construction in which a focused XP moves to Spec,FocP and the remnant FinP moves to Spec,ForceP via core-spine roll-up, yielding the surface head-finality of the CP domain. This head-initial CP analysis offers a precise account of long-standing puzzles, including the acceptability contrast for genitive

[13] Under the head-initial analysis adopted here, *v*P must first undergo roll-up to Spec,NegP to derive the surface order V–Neg. The relevant structure is therefore $[_{NegP}$ *v*P [Neg^{0} [Neg/Foc] <*v*P>]], where <*v*P> represents the lower copy of *v*P. When an NCI is targeted by feature-driven focus movement, the Focus-related feature of Neg^{0} further requires a local feature-sharing relation with XP-*mo*. XP-*mo* therefore raises from within the rolled-up *v*P to an outer Spec,NegP, yielding $[_{NegP}$ XP-mo $[_{NegP}$ *v*P [Neg^{0}[Neg/Foc] <vP>]]]. The inner Spec,NegP is thus occupied by *v*P as a result of labeling-driven roll-up triggered by the weak categorial feature of Neg^{0}, whereas the outer Spec,NegP hosts XP-*mo* in the Neg/Foc feature-sharing configuration. This configuration forms the maximal Neg/Foc projection relevant to the Maximality effect discussed in the text. Two aspects of this implementation are not independently established here. First, the analysis requires two specifier positions in NegP and assigns two distinct syntactic functions to a single head, Neg^{0}: labeling-driven roll-up of *v*P and Neg/Foc feature sharing with XP-*mo*. Second, how the requirement imposed by focus movement to the higher Spec,FocP gives rise to the lower NegP configuration in a strictly cyclic derivation remains to be formally worked out. These points constitute limitations of the present analysis and issues for future research.

subjects in clefts vs. *no-da*, NP/PP cleft asymmetries, NQ-based amelioration via an articulated DP, and restrictions on the movement of NPIs. Building on these results, section 4 develops a diachronic account of changes in clausal structure, focusing on the left peripheral CP.

4 Old Japanese as a disharmonic OV language

The parameter hierarchy approach to word order variation represents a significant shift in how parameters are conceptualized within generative grammar, especially under the Minimalist Program. Under this approach, word order variations are best understood in terms of microparametric differences—small, local settings tied to the features of specific heads within extended verbal projections, rather than global macroparameters. Languages are classified as "harmonic" if they consistently display either head-final or head-initial order across phrase types. However, a significant number of languages display "disharmonic" properties, in particular, V2 languages like German, where the expected harmony between phrase structures is disrupted by mixed head-directionality.

This approach not only explains the synchronic diversity of word order across languages but also captures the diachronic word order change. The FOFC restricts possible combinations of head-directionality across projections and shapes the trajectory of word order change: transitions from head-final to head-initial structures typically proceed in a top-down fashion, involving stepwise reanalysis of higher projections first; conversely, shifts from head-initial to head-final proceed in a bottom-up fashion, with lower projections reanalyzed first (see Roberts 2019: 131–132). This is illustrated in (56–57).

(56) Head-final to head-initial (Top-down)
a. [[[O V] T] C] → b. [C [[O V] T]] → c. [C [T [O V]]] → d. [C [T [V O]]]

(57) Head-initial to head-final (Bottom-up)
a. [C [T [V O]]] → b. [C [T [O V]]] → c. [C [[O V] T]] → d. [[[O V] T] C]

A disharmonic word order arises in change when roll-up applies in some, but not all, domains within the extended projection of VP, resulting in a mixture of head-initial and head-final structures within a clause. The FOFC allows the disharmonic configuration in (56b–c), instantiated in the Old English (OE) examples in (58) taken from van Kemenade (1987: 17–19).

(58) a. [$_{CP}$ O V [$_{TP}$ S…]]
eall ðis aredað se reccere swiðe ryhte (CP, 168,3)
all this arranges the ruler very rightly
'The ruler arranges all this very rightly.'

b. [C [$_{TP}$ [$_{vP}$ SOV] T <*v*P>]]
…þæt hie þæt to his honda healdan sceoldon (Parker 887)
that they it from his hand hold should
'that they ought to hold it from his hand.'

In (58a), the main clause shows Verb Second (V2) effects, with the finite verb moving to C, while in the embedded clause (58b) a head-initial complementizer prevents V-to-C movement, and C dominates a head-final TP after roll-up of *v*P to Spec,TP, assuming that the finite auxiliary occupies T (see section 2.2). OE exhibited an attested disharmonic structure in (56b) that corresponds to (58b) and developed into the harmonic word order shown in (56d).

Work within the parameter hierarchy framework, however, has primarily focused on predominantly VO languages. Existing treatments of OV languages remain largely theory-internal and lack systematic empirical validation. This section therefore offers an empirical assessment of this framework, drawing on evidence from Japanese, an OV language. It is shown that while Modern Japanese (ModJ) is a harmonic head-final language, Old Japanese (OJ; 700–800) does not fully conform to this harmonic pattern.

The discussion in this section focuses on focus constructions in OJ as the diachronic source of the ModJ cleft. I argue that the development of the ModJ cleft reflects a local microparametric change within the articulated CP domain, from the disharmonic structure in (57c) to the harmonic head-final configuration in (57d). This change involved the emergence of core-spine roll-up whereby FinP raises to Spec,ForceP. ModJ clefts are thus analyzed as the historical outcome of the OJ focus construction, resulting from the upward extension of roll-up to Force^0, the highest head of the CP domain.

4.1 Nominalization

OJ exhibits a distinctive system of clause types that is crucial for understanding its syntax and case marking. The major clause types are traditionally labeled as *shūshi* (conclusive) and *rentai* (adnominal/nominalized), each with its own syntactic and morphological properties. Yanagida & Whitman (2009) follow earlier Japanese tradition such as Konoshima (1962) that the adnominal form is fundamentally a *nominalized* clause type, and an NP modifying (adnominal) function is secondary. A nominalized clause type appears in various embedded environments (such as relative clauses, conditionals, and complement clauses) endowing the clause with [+N] properties and enabling it to function syntactically as a nominal (see also Miyagawa 1989). One of the important main clause properties of nominalized clauses in OJ is their central role in licensing the *kakari-musubi* (KM) (focus concord) construction; specifically, KM arises when focus particles trigger the use of the adnominal form, reflecting the nominalized status of the clause.

Section 4 argues that a head-initial analysis of the KM construction provides evidence for its connection to the seemingly unrelated cleft constructions in ModJ. This analysis demonstrates—on both theoretical and empirical grounds—that changes in the formal features of a specific functional head can induce a global restructuring of clause structure within a language.

4.2 Wh/focus movement

In OJ, KM is defined by a unique syntactic pattern in which the presence of certain focus particles—such as *so* (later became the voiced *zo*), *namu, ya, ka*, and *koso*—requires the predicate to appear in a nominalized form, rather than the standard conclusive (*shūshi*) form. The construction serves to highlight, emphasize, or question the constituent marked by the particle, and the nominalized verb form often imparts a presuppositional or backgrounded interpretation to the clause. The traditional view links the loss of KM to the morphological merger of the adnominal and conclusive forms in the 12th–15th centuries. As these forms converged, the morphosyntactic cue necessary for focus concord was eliminated, thereby eroding the structural foundation of the construction. Consequently, the loss of KM reflects a broader typological shift in Japanese, from a system in which nominalization played a central role in clause structure and information structure, to one in which such distinctions were no longer morphologically encoded. (For comprehensive descriptive discussions of KM in OJ, see Narrog 2019).

While an extensive body of prior research exists on KM, the discussion here is limited to the prototypical focus concord constructions that involve the particles *ka* and *so* because *ka* and *so* have functionally and syntactically identical properties. While in ModJ, the Q particle *ka* occurs in clause-final position, it is directly attached to the wh-phrase in OJ.

Based on Nomura's (1993) observation that *ka*-marked wh-phrases invariably precede the subject in OJ, Watanabe (2002) argues that OJ exhibits wh-movement, with wh-phrases moving to FocP in the articulated CP domain.[14] The examples in (59) are cited from Watanabe (2002:182), with minor modifications to the glosses and romanization.[15]

(59) a. **iduku yu ka** imo ga iriki-te ime ni miye turu. (MYS 3117)
where from Q wife-SUBJ enter-CONJ dream-LOC appear PERF.ADN
'From where did my wife come and appear in my dream?'

[14] Aldridge (2018) advances an alternative movement account of KM, proposing that the focused constituent raises to Spec,TP, i.e., to a position lower than CP. This is based on the observation that in OJ KM clauses a bare nominative subject consistently precedes the focused constituent. In line with the present hypothesis, unmarked subjects can be taken to be base-generated topics in a left-peripheral topic position.

[15] In this paper, OJ examples are taken from *Man'yōshū* (abbreviated as MYS), the oldest collection of Japanese verse compiled in the mid-8th century A.D. The examples are romanized following the Frellesvig–Whitman system (Frellesvig & Whitman 2008).

b. **nani-wo ka-mo** mikari-no pito-no wori-te kazasa mu? (MYS 1974)
what-OBJ Q-FOC hike-GEN person-GEN.SUBJ pick-CONJ wear AUX.ADN
'What should hikers pick and wear on the hair?'

In OJ, the complex particle *ka*+*mo* could directly attach to wh-phrases, as in (59b), and function as the Q-marker of constituent questions. In contrast, in ModJ *ka* alone serves as the clause-final Q-particle licensing both *wh*- and yes-no alternative questions, but *ka*+*mo* has been reanalyzed as a clause-final epistemic modal that only takes propositional (yes–no) inputs (60).

(60) Modern Japanese
a. John ga wain o kau **kamo** (sire-nai).
John NOM wine ACC buy Q know-not
'It may be that John will buy wine'

b. *John ga nani o kau **kamo** (sire-nai).
John NOM what ACC buy Q know-not
'*It may be that John will buy what?'

In Mitrović's (2021) superparticle framework (see section 3.3), κ corresponds to the *ka*-type particle and μ to the *mo*-type particle. Specifically, κ is existential/inquisitive in nature, opening up alternatives (yielding disjunction, existential quantification, or questions), while μ is universal/exhaustive, closing off alternatives (yielding conjunction, universal quantification, or additive meanings). Mitrović proposes that meanings of superparticles change diachronically through structurally constrained grammaticalization: nominal quantifier particles are reanalyzed as clausal Focus/Junction heads. The diachronic change can then be captured, in Mitrović's terms, by saying that an earlier [κ+μ] configuration (*ka*+*mo*) functioned as the Foc head for wh-questions. In the modern system, by contrast, [κ+μ] has been lexicalized as *kamo*, a sentential modal particle that instantiates such a Junction head over propositional alternatives, lacks Q/Foc-features, and therefore cannot directly combine with wh-phrases.

According to Watanabe (2007) *so* —which occurs with both a *wh*-phrase and with a non-interrogative focus phrase—patterns with *ka* in that it also precedes the subject. (61) are taken from the Corpus of Historical Japanese (CHJ).[16,17]

[16] The Corpus of Historical Japanese (CHJ) is a large, diachronic corpus developed by the National Institute of Japanese Language and Linguistics to support research on the history of Japanese from Old Japanese through the early 20th century.

[17] One of the reviewers rightly observes that, if OJ *wo* in (61b) were analyzed as a DP-internal CaseP on a par with ModJ *o*, the example should be ruled out by the Maximality Principle that excludes ModJ sequences like [$_{\text{CaseP}}$ *piza ga/o*] *da*, as discussed in section 3.5. (Under the same analysis, (59b) should likewise be ruled out.) However, OJ *wo* differs significantly from its ModJ counterpart: *wo* appears only with definite/specific objects, and indefinite/non-specific objects are morphologically unmarked (Yanagida & Whitman 2009). This suggests that the status of *wo* in OJ is that of the highest D, so movement of the *wo*-marked object as a DP does not lead to a violation of the Maximality Principle. A more detailed investigation of the categorial status and distribution of OJ *wo* is left to future work.

(61) a. **kyepu so** wa=ga ko-si. (MYS 2216)
today FOC I SUBJ come-PST.ADN
'It is today that I have come.'

b. kata-yori ni **ito wo so** wa=ga yoru. (MYS 1987)
single-ply be thread OBJ FOC I SUBJ twist.ADN
'It is the thread with a single ply that I am twisting.'

When the sentence contains the topic phrase marked by *pa*, all these three can appear in the designated order as illustrated in (62) (Watanabe 2002: 183).

(62) patuse no kapa **pa** ura nami **ka** pune **no** yori-ko-nu. (MYS 3225)
Hatsuse GEN river TOP shore absentness Q boat GEN.SUBJ approach-come-NEG.ADN
'Is it because the Hatsuse River has no shore that no boat comes near?'

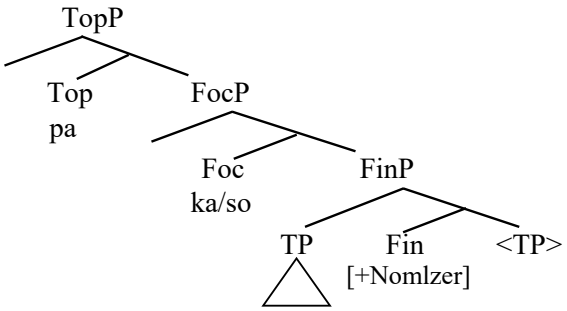
In Watanabe's (2002; 2007) framework, *pa*-marked topic moves or base-generated in TopP, the wh/focused phrase moves to Spec,FocP and the subject moves to Spec,IP, suggesting the OJ left peripery is structured in (63).

(63) Left peripheral structure in OJ (Watanabe 2002; 2007)
[$_{\text{TopP}}$ Topic **pa** [$_{\text{FocP}}$ Wh/focus-phrase **so/ka** [$_{\text{IP}}$ Subject **ga/no** [VP…]]]]

Drawing on his analysis of the *Tale of Genji* (ca. 1010), Watanabe (2002) observes that in Early Middle Japanese (EMJ; 800–1200) wh-phrases frequently appear without *ka* and show no word-order restrictions, as found in OJ, leading him to conclude that overt wh-movement had been lost by this period. According to Watanabe, this syntactic change—the loss of wh-movement—was the primary factor driving the subsequent morphological collapse of the adnominal and conclusive forms. Watanabe situates these changes within broader typological generalizations (e.g., Bach's generalization that SOV languages tend to have wh-in-situ rather than overt wh-movement). OJ, being SOV, originally had the marked option of overt wh-movement, but its loss in EMJ reflects a shift toward the typologically unmarked OV pattern found in ModJ.

Given that the left-peripheral particles occupy fixed positions in (63), and assuming an LCA-based clause structure in the spirit of Kayne (1994) and Whitman (1998; 2001), I propose that *pa/ka/so* function as heads within the CP layer, each triggering feature sharing movement of a phrase to its specifier position. Recall that in ModJ cleft focus constructions, the overt nominalizer *no* is required in Fin^0, whereas in OJ the adnominal forms themselves have a nominalizing function; I therefore take Fin^0 in OJ to host an overt or covert suffixal nominalizer. In other words, Fin^0 in OJ provides the nominalizing morphology that licenses the KM clause as the complement of the left-peripheral focus system.

(64) KM focus construtions in OJ



Importantly, in OJ, both *ga* and *no* are genitive markers used for NP possessors and for the subjects of *nominalized* clauses. *Ga* is reanalyzed as nominative case in ModJ, but *no* remains genitive throughout the history of Japanese. Recall that in ModJ, the nominalizer *no* in Fin^0 does not license a genitive subject (see (40) in section 3.4). In OJ, by contrast, such a subject can be licensed by Fin^0.

A head-initial analysis makes clear that the apparent heterogeneity of the KM construction and the ModJ cleft (section 3.3 (29)) in fact collapses into a single diachronic trajectory.[18] In KM constructions, the left periphery is structured around Topic and Focus projections, with clause-typing force encoded in Fin^0 rather than in a distinct $Force^0$. In this system, the CP domain is disharmonic: lower FinP triggers roll-up of its TP complement, but the Topic/Focus field above FinP does not participate in FOFC-compliant roll-up. In ModJ clefts, the particle *wa* is reanalyzed as a Force-related head that types the clause as an identificational specificational sentence. Under this reanalysis, FinP undergoes core-spine roll-up to Spec,ForceP, yielding a head-final CP in which Force and Fin jointly determine both clause type and finiteness.

In the parameter hierarchy approach, the disharmonic word order characteristic of OE is accounted for through roll-up movement: the T head bears a roll-up feature that triggers the movement of *v*P to Spec,TP as in (58b), a mechanism also attested in OJ. In the broader diachronic literature, the change from OV to VO in English has often been linked to the presence and subsequent loss of verb movement to the CP domain. In OE, robust V-to-C movement, together with rich inflection, yielded systems in which OV-like and VO-like orders coexisted, providing learners with overt evidence for a grammar in which the verb raises to higher functional heads.

---

[18] The traditional grammars have long noted a functional similarity between the *kakari*-particle *so* in OJ and the copula *da* in ModJ. The copula *da* emerged via forms such as the OJ assertive auxiliary *nari*; it did not develop directly from *so*, therefore, there is no morphological or etymological continuity between these particles. Nevertheless, Ohno (1993) argues that the OJ pattern [X pa Y so] corresponds to the basic ModJ [X wa Y da]. Although the precise historical point at which the ModJ "cleft" construction emerged remains unclear, I propose that it is diachronically continuous with the OJ KM construction.

As V-movement weakened and inflectional morphology was reduced, this evidence disappeared, and children came to acquire a simpler VO grammar in which the verb remains in a low position and clause structure is labelled by feature sharing rather than by verb movement. In parameter hierarchy terms (Roberts 2019), this development can be seen as a reanalysis of the formal features on heads like *v*,T, C, shifting from an OV system with roll-up and/or verb movement to a VO system in which extended projections are uniformly head-initial and roll-up is inoperative.

By contrast, in Japanese, clause typing and left-peripheral structure are expressed by particles occupying head positions. I propose that these particles block comparable verb movement into the CP domain. Head-final order is instead derived by phrasal roll-up. This difference in the availability of verb movement and in how clause structure is encoded plausibly underlies the divergent diachronic trajectories of English (OV→VO) and Japanese, which remains persistently head-final.

5 Concluding remarks

This paper claims that while the head-final profile of ModJ is derived by labeling-driven roll-up across the clausal spine, OJ was disharmonic, with a head-initial C that triggered A′-movement; the diachrony of OV stability and focus/cleft behavior follows from microparametric changes on left-peripheral heads, not from macroparametric resets.

Synchronically, section 3 analyses Japanese “clefts” as involving focus movement to Spec,FocP in CP layer. This movement, however, is obscured by subsequent core-spine roll-up of FinP to the topmost Spec,ForceP. This paper proposes that core-spine roll-up relates the lower and higher heads of the obligatory core in both CP (Fin–Force) and DP (Case–D). Three longstanding cleft puzzles are unified under Rizz’s (2016) Maximality Principle. First, PP focus is acceptable, single case-marked NP focus is degraded, but a following numeral quantifier (NQ) rescues NP clefts. An articulated DP, as proposed by Watanabe (2008), holds that case-marked NPs correspond to sub-DP CaseP, but NQ configurations build full DPs. The Maximality Principle allows only PP/DP to move to FocP. Second, cleft focus movement targets the criterial Spec,FocP position and is subject to the Maximality Principle, whereas long-distance A′-scrambling involves non-criterial TP-adjunction and can move CaseP without inducing a Maximality effect. Third, the same principle explains why negative concord items of the wh-*mo* type cannot serve as cleft focus: their focus-related feature (Watanabe 2004) requires a local Neg-related configuration that freezes them against further movement to Spec,FocP.

Two structural claims further support the left-peripheral configuration. *Wa* in Force^{0}, which realizes the highest CP head in clefts, can license genitive subjects inside the presuppositional TP, explaining why genitive subjects are barred in *no-da* constructions but possible in clefts. These

claims, together with FOFC considerations, motivate a head-initial CP in which apparent head-finality is produced by roll-up rather than by a base head-final configuration.

Section 4 contends that the apparent heterogeneity between the KM construction in OJ and the ModJ cleft collapses into a single diachronic trajectory. In OJ, the CP domain is disharmonic: the higher Topic and Focus positions are head-initial, realized by *pa* and *so/ka* respectively, but the lowest functional projection Fin^0 triggers roll-up of TP to Spec,FinP, showing that the lowest functional category inside CP is head-final. I argue that Japanese focus constructions underwent microparametric change, involving reanalysis of the feature specifications of left-peripheral functional heads, which triggered the loss of *kakari-musubi* and the emergence of an innovative focus construction in ModJ.

Abbreviations

ACC = accusative, ADN = adnominal, AUX = auxiliary, CL = classifier, COMP = complementizer, CONJ = conjunction, COP = copula, DAT = dative, DEC = declarative, FN = formal noun, GEN = genitive, KM = *kakari-musubi*, LOC = locative, ModJ = Modern Japanese, NEG = negation, NMLZ = nominalizer, NOM = nominative, OBJ = object, OE = Old English, OJ = Old Japanese, PERF = perfect, PPT = past participle, PRES = present, PST = past, Q = question particle, SUB = subject, TOP = topic.

Acknowledgements

I am deeply grateful to the three anonymous reviewers for *Glossa* for their time, careful reading, and many valuable comments and suggestions, which substantially improved this paper. I would also like to thank John Whitman for generously sharing his knowledge and for our many insightful discussions, which have informed this work. Any remaining errors or omissions are, of course, my own. This work was supported by the Japan Society for the Promotion of Science through its Grants-in-Aid for Scientific Research program (JSPS KAKENHI Grant Number JP23K00517).

Competing Interests

The author has no competing interests to declare.

References

Aldridge, Edith. 2018. C-T inheritance and the left periphery in Old Japanese. *Glossa: A Journal of General Linguistics* 3(1): 26. 1–22. DOI: https://doi.org/10.5334/gjgl.254

Biberauer, Theresa & Holmberg, Anders & Roberts, Ian. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45. 169–225. DOI: https://doi.org/10.1162/LING_a_00153

Biberauer, Theresa & Roberts, Ian. 2015. The clausal hierarchy, features and parameters. In Shlonsky, Ur (ed.), *Beyond functional sequence,* 295–313. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780190210588.003.0016

Biberauer, Theresa & Roberts, Ian. 2017. Conditional inversion and types of parametric change. In Los, Bettelou & de Haan, Pieter (eds.), *Word order change in acquisition and language contact: Essays in honour of Ans van*, 57–77. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/la.243.04rob

Cho, Sungdai & Whitman, John & Yanagida, Yuko. 2008. Clefts in Japanese and Korean. *Proceedings of the Chicago Linguistic Society* 44(1). 61–77. Chicago, IL: Chicago Linguistic Society.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press

Chomsky, Noam. 2001. Derivation by Phase. In Kenstowicz, Michael (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/4056.003.0004

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI: https://doi.org/10.1016/j.lingua.2012.12.003

Chomsky, Noam. 2015. Problems of projection: Extensions. In Di Domenico, Elisa & Hamann, Cornelia & Matteini, Simona (eds.), *Structures, strategies and beyond: Studies in honour of Adriana Belletti*, 1–16. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/la.223.01cho

Davis, Colin. 2020. Crossing and stranding at edges: On intermediate stranding and phase theory. *Glossa: A Journal of General Linguistics* 5(1):17. 1–32. DOI: https://doi.org/10.5334/gjgl.854

Declerck, Renaat. 1988. *Studies in copular sentences, clefts, and pseudo-clefts*. Leuven: Leuven University Press. DOI: https://doi.org/10.1515/9783110869330

Den Dikken, Marcel. 2006. Either-float and the syntax of co-or-dination. *Natural Language and Linguistic Theory* 24(3). 689–749. DOI: https://doi.org/10.1007/s11049-005-2503-0

Diesing, Molly. 1988. Bare plural subjects and the stage/individual contrast. In Krifka, Manfred (ed.), *Genericity in natural language* (SNS-Bericht 88–42), 107–154. Tübingen: Universität Tübingen.

Embick, David. 2010. *Localism versus globalism in morphology and phonology*. Linguistic Inquiry Monograph 60. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262014229.001.0001

Frellesvig, Bjark. 2001. A common Korean and Japanese copula. *Journal of East Asian Linguistics*. 10(1). 1–35. DOI: https://doi.org/10.1023/A:1026512817255

Frellesvig, Bjarke & Whitman, John. 2008. *Proto-Japanese*. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/cilt.294

Fukui, Naoki. 1986. *A theory of category projection and its applications*. Cambridge, MA: Massachusetts Institute of Technology doctoral dissertation.

Fukui, Naoki. 1995. *Theory of projection in syntax*. Stanford, CA: CSLI Publications.

Fukui, Naoki. 2022. Nichi eigo hikaku tojiron to parameter no riron [Comparative syntax of Japanese and English and the theory of parameters: On the differences and similarities between the Fukui and Kuroda systems.] *Gengo Kenkyu* 161. 1–33.

Heycock, Caroline. 2008. Japanese *-wa*, *ga*, and information structure. In Miyagawa, Shigeru & Saito, Mamoru (eds), *The Oxford handbook of Japanese linguistics*, 54–83. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780195307344.013.0003

Hiraiwa, Ken. 2002. Nominative-genitive conversion revisited. *Japanese/Korean Linguistics* 10. 546–559. Stanford, CA: CSLI Publications.

Hiraiwa, Ken & Ishihara, Shinichiro. 2002. Missing links: Cleft, sluicing, and "*no da*" construction in Japanese. *MIT Working Papers in Linguistics* 43. 35–54.

Hoji, Hajime. 1987. Japanese clefts and chain binding/reconstruction effects. Handout of a talk presented at the *Sixth West Coast Conference on Formal Linguistics* (WCCFL 6).

Hoji, Hajime. 1990. *Theories of anaphora and aspects of Japanese syntax*. Unpublished manuscript, University of Southern California. (https://www.researchgate.net/publication/248070527_Theories_of_Anaphora_and_Aspects_of_Japanese_Syntax)

Hoji, Hajime & Ueyama, Ayumi. 2003. Resumption in Japanese: A preliminary study. Handout, WECOL 2003, University of Arizona.

Holmberg, Anders. 2000. Deriving OV order in Finnish. In Svenonius, Peter (ed.), *The derivation of VO and OV*, 123–152. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/la.31.06hol

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kinsui, Satoshi. 2002. On the nominative and accusative case markers in *Man'yōshū*. Unpublished manuscript, Osaka University.

Koizumi, Masatoshi. 2000. String vacuous overt verb raising. *Journal of East Asian Linguistics* 9. 227–285. DOI: https://doi.org/10.1023/A:1008311420647

Konoshima, Tadatoshi. 1962. Chūkogo ni okeru yōgen rentaikei no yōhō [The use of the participial adjective in medieval Japanese]. *Kokugogaku* 48. 102–107.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT press.

Kuroda, S.-Y. 1988. Whether we agree or not: A comparative grammar of English and Japanese. *Linguisticae Investigationes* 12. 2–44. DOI: https://doi.org/10.1075/li.12.1.02kur

Kuroda, S.-Y. 2005. Focusing on the matter of topic: A study of *wa* and *ga* in Japanese. *Journal of East Asian Linguistics* 14. 1–58. DOI: https://doi.org/10.1007/s10831-004-2701-5

Kuroda, S.-Y. 2007. On the syntax of Old Japanese. In Frellesvig, Bjarke & Smith, J. C. & Shibatani, Masayoshi (eds.), *Current issues in the history and structure of Japanese*, 263–318. Tokyo: Kurosio.

Mitrović, Moreno. 2014. Deriving *ka* (*karimusubi*) in premodern Japanese. *Acta Linguistica Asiatica* 4(3). 9–28. DOI: https://doi.org/10.4312/ala.4.3.9-28

Mitrović, Moreno. 2021. *Superparticles: A microsemantic theory, typology, and history of logical atoms*. *Studies in Natural Language and Linguistic Theory* 98. Dordrecht: Springer. DOI: https://doi.org/10.1007/978-94-024-2050-0

Miyagawa, Shigeru. 1989. *Structure and case marking in Japanese. Syntax and semantics* 22. New York: Academic Press. DOI: https://doi.org/10.1163/9789004373259

Miyagawa, Shigeru. 2001. The EPP, scrambling and wh-in-situ. In Kenstowicz, Michael (ed.), *Ken Hale: A life in language*, 293–338. Cambridge: MIT Press. DOI: https://doi.org/10.7551/mitpress/4056.003.0012

Miyagawa, Shigeru. 2010. *Why agree? Why move? Unifying agreement-based and discourse-configurational languages*. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/8116.001.0001

Miyagawa, Shigeru. 2011. Genitive subjects in Altaic and specification of phase. *Lingua* 121. 1265–1282. DOI: https://doi.org/10.1016/j.lingua.2011.01.009

Miyagawa, Shigeru. 2012. *Case, argument structure and word order*. New York: Routledge. DOI: https://doi.org/10.4324/9780203126844

Narrog, Heiko. 2019. Origin and structure of focus concord constructions in Old Japanese—a synthesis. Glossa: *A Journal of General Linguistics* 4(1): 108. 1–41. DOI: https://doi.org/10.5334/gjgl.629

Nomura, Takashi. 1993. Jōdaigo no *no* to *ga* ni tsuite [On the particles *no* and *ga* in Old Japanese]. *Kokugo Kokubun* 62. 1–17

Ohno, Susumu. 1993. *Kakari-musubi no kenkyu* [A study of *kakari-musubi]*. Tokyo: Iwanami.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Haegeman, Liliane (ed.), *Elements of grammar: A handbook of generative syntax,* 281–337. Dordrecht: Kluwer Academic Publishers. DOI: https://doi.org/10.1007/978-94-011-5420-8_7

Rizzi, Luigi. 2006. On the form of chains: Criterial positions and ECP effects. In Cheng, Lisa Lai-Shen & Corver, Norbert (eds.), *Wh-movement: Moving on*, 97–133. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/7197.003.0010

Rizzi, Luigi. 2016. Labeling, maximality and the head–phrase distinction. *The Linguistic Review* 33(1). 103–127. DOI: https://doi.org/10.1515/tlr-2015-0016

Rizzi, Luigi & Shlonsky, Ur. 2007. Strategies of subject extraction. In Gärtner, Hans-Martin & Sauerland, Uli (eds.), *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, 115–160. Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110207552.115

Roberts, Ian. 2007. *Diachronic syntax*. Oxford: Oxford University Press.

Roberts, Ian. 2019. *Parameter hierarchies and universal grammar*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780198804635.001.0001

Roberts, Ian & Holmberg, Anders. 2010. Introduction: Parameters in minimalist theory. In Bioerauer, Theresa & Holmberg, Anders & Roberts, Ian & Sheehan, Michelle (eds), *Parametric variation: Null subjects in minimalist theory*, 1–57. Cambridge: Cambridge University Press.

Sadakane, Kumi & Koizumi, Masatoshi. 1995. On the nature of the dative particle *ni* in Japanese. *Linguistics* 33. 5–33. DOI: https://doi.org/10.1515/ling.1995.33.1.5

Saito, Mamoru. 1985. *Some asymmetries in Japanese and their theoretical implications*. Cambridge, MA: Massachusetts Institute of Technology doctoral dissertation.

Saito, Mamoru. 1989. Scrambling as semantically vacuous A′-movement. In Baltin, Mark R. & Kroch, Anthony S. (eds.), *Alternative conceptions of phrase structure*, 182–200. Chicago: University of Chicago Press.

Saito, Mamoru. 1992. Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1(1). 69–118. DOI: https://doi.org/10.1007/BF00129574

Takezawa, Koichi 1987. *A configurational approach to case-marking in Japanese*. Seattle: University of Washington doctoral dissertation.

Tateishi, Koichi. 1994. *The syntax of subjects*. Stanford, CA: CSLI Publications.

van Kemenade, Ans. 1987. *Syntactic case and morphological case in the history of English*. Dordrecht: Foris Publications. DOI: https://doi.org/10.1515/9783110882308

Watanabe, Akira. 1996. Nominative-genitive conversion and agreement in Japanese: A cross-linguistic perspective. *Journal of East Asian Linguistics* 5. 373–410. DOI: https://doi.org/10.1007/BF00132699

Watanabe, Akira. 2002. Loss of overt wh-movement in Old Japanese. In Lightfoot, David (ed.), *Syntactic effects of morphological change*, 179–195. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199250691.003.0010

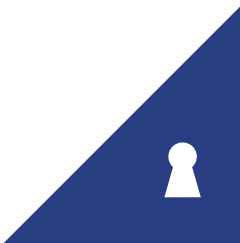
Watanabe, Akira. 2004. The genesis of negative concord: Syntax and morphology of negative doubling. *Linguistic Inquiry* 35(4). 559–612. DOI: https://doi.org/10.1162/0024389042350497

Watanabe, Akira. 2007. Topic-focus articulation in Old Japanese: *So/zo* and *koso*. In Kelepir, Meltem & Öztürk, Balkız (eds.), *Proceedings of the 2nd Workshop on Altaic Formal Linguistics (WAFL 2)* (MIT Working Papers in Linguistics 54), 121–137. Cambridge, MA

Watanabe, Akira. 2008. The structure of DP. In Miyagawa, Shigeru & Saito, Mamoru (eds.), *The Oxford handbook of Japanese linguistics*, 513–540. Oxford: Oxford University Press.

Whitman, John. 1998. Gojyun to kukōzō [Word order and clause structure]. In Nakau, Minoru (eds), *Kaku to gojun to tōgo kōzō: Nichi-eigo hikaku sensho* [*Case, word order and syntactic structure*: *A Comparative series on Japanese and English*] 9. 104–195. Tokyo: Kenkyūsha.

Whitman, John. 2001. Kayne 1994:P. 143. Fn 3. In Alexandrova, Galina & Arnaudova, Olga (eds.), *The Minimalist parameter*, 77–100. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/cilt.192.07whi

Yanagida, Yuko. 2006. Word order and clause structure in early Old Japanese. *Journal of East Asian Linguistics* 15. 37–68. DOI: https://doi.org/10.1007/s10831-005-2165-2

Yanagida, Yuko & Whitman, John. 2009. Alignment and word order in Old Japanese. *Journal of East Asian Linguistics* 18. 101–144. DOI: https://doi.org/10.1007/s10831-009-9043-2