

SYMMETRIC PROJECTION IN JAPANESE AND ENGLISH: A MODIFICATION OF STABLER'S PARSING/GENERATION MODEL FOR MINIMALIST GRAMMAR

Miyoko Yasui

Dokkyo University
1-1 Gakuencho, Soka,
Saitama, 340-0042 JAPAN

Email: myasui@dokkyo.ac.jp

ABSTRACT

The essence of standard X-bar theory is that structure building is asymmetric in the sense that a complex structure inherits properties from only one of its constituents. There are some structures, however, that are best analyzed as reflecting the properties of all their constituents. This kind of symmetric projection should in principle be allowed within the minimalist program if the union of the features of all the constituents contains no incompatible features. This claim is supported by the fact that Japanese wh-phrases marked with *ka* can function as indefinites as well as interrogatives. Under the assumptions that a wh-phrase with *ka* has the same internal structure regardless of its interpretations and that *ka* has no category feature, merging a wh-phrase with *ka* is a case of symmetric projection. The properties of both *ka* and its sister wh-phrase interact with those of the predicate taking the *ka*-phrase as its argument or adjunct, which ensures that an appropriate interpretation will be picked up from the two possible interpretations of the *ka*-phrase.

1. INDEFINITES AND PHRASAL INTERROGATIVES IN JAPANESE

It is well-known that wh-phrases in Japanese, unlike those in English, can be used as indefinite as well as interrogative expressions. For example, each sentence below contains exactly the same strings of words; the first one is interpreted as an indefinite and the second, as an interrogative:

- (1) John-ga [dare ka] nagutta rasi! ga [dare ka] sitteiru hito wa i nai.
NOM someone beat it-seems but who Q know person TOP exist not
'John seems to have beaten someone but nobody knows who.'
- (2) John-ga [dare kara ka] henna tegami-o moratta rasii ga,
NOM someone from strange letter ACC received it-seems but
boku-wa [dare kara ka] sir- anai.
I TOP who from Q know not
'John seems to have received a strange letter from someone but I don't know from whom.'
- (3) John-wa [naze ka] naiteita ga [naze ka]-wa daremo siranai.
TOP some reason was-crying but why Q TOP nobody know
'John was crying for some reason but nobody knows why.'

The same point can be illustrated with (4), which can be interpreted as (5a) or (5b) depending on its context:

- (4) John-wa [naze ka] rikai dekinakatta
TOP why understand couldn't

- (5) a. (Mary didn't come and) John was not able to understand [why she didn't come]
 b. (Mary explained the situation to John but) he was not able to understand it [for some unclear reason]

A wh-phrase to be construed as an interrogative will be called 'phrasal interrogative' in contrast to 'clausal interrogative' exemplified by the bracketed embedded question in (5a).

A standard approach to the indefinite and phrasal interrogative usage of wh-phrases is to assume an empty clause for the latter, while treating the former as categorially the same as the inside wh-phrase (see [15] among others). Under this view, (2) would be analyzed as follows:

- (6) John-ga [PP [PP dare kara] ka] henna tegami-o moratta rasii ga,
 NOM someone from strange letter ACC received it-seems but
 boku-wa [CP [PP dare kara] [[IP e][COMP ka]]] sir- anai.
 I TOP who from Q know not

The idea behind this analysis is that if the same strings of words have distinct meanings, they have distinct structures. This is certainly one of the possible approaches, but a question should be raised whether there is a simpler and more plausible alternative to it. The purpose of this paper is to pursue this possibility.

For ease of exposition, I will adopt Stabler's parsing/generation model presented in [12], [13], and [14], where the basic ideas in the minimalist program are assumed. The analysis, however, can be translated into any theoretical framework as long as it treats lexical items as consisting of syntactic and other features and assumes structure-building operations and principles such as Merge and the Projection Principle in the minimalist program.¹ The main claims will be that (i) lexical categories have category features that will project up to the complex they form, but functional categories do not; (ii) a complex structure is constructed from two inputs by combining the features of both symmetrically if one of them lacks a category feature; (iii) the indefinite/interrogative distinction with Japanese wh-phrases can be made simply under assumptions (i) and (ii); and (iv) two of the major syntactic differences between Japanese and English are also deducible from (i), (ii), and one parameter as to functional categories.

2. SEMANTIC AMBIGUITY WITHOUT STRUCTURAL AMBIGUITY

As discussed at the beginning, it is not unreasonable to assign a clausal structure to a wh-phrase that is to be interpreted as an interrogative 'clause.' In fact, assuming the same internal structure for the indefinite and interrogative usage of wh-phrase, as I will in this paper, might appear to be counterintuitive at first. There are clear cases, however, in which a single phrase can express distinct meanings but can hardly be associated with two distinct syntactic structures.

Made, which is one of the Japanese particles, can be construed as temporal ('until' in English) or spatial ('as far as') depending on the main verb:

- (7) a. John-wa 3-zi made matta. b. John-wa Osaka made itta.
 TOP o'clock until waited TOP as far as went
 'John waited until three o'clock. 'John went as far as Osaka.'

The interpretive mechanism here is simple: *made* can be temporal or spatial; its sister constituent in (7a) denotes a specific time and the temporal interpretation is obtained, whereas its sister in (7b) expresses the name of a place and hence, its spatial interpretation. Replacing the *made*-phrase in (7a) with that in (7b), as in (8), produces some strangeness:

¹ See [2] and [3] for the minimalist program. Stabler's model is a kind of unification-based grammar, but it adopts the guiding ideas in the minimalist program, instead of those like FOOT features. See [5], [6], and [11] for unification-based grammars.

- (12) a. John-wa [PP [IP [α e] Osaka [P e] [V e]] [P made]] mat-ta
 b. *John-wa [PP [IP densha-ga Osaka ni [V e]] [P made]] mat-ta
 c. *John-wa [PP [IP densha-ga Osaka [P e] tuku]] [P made]] mat-ta

(12b,c) show that a preposition and a verb cannot be freely empty. It can be concluded that a *made*-phrase is to be analyzed as a PP consisting of a simple NP and the P *made* under its temporal as well as spatial interpretation.²

The following data from English prove the same point (see also [9]):

- (13) a. That was the best film I've seen since 1980.
 b. That was the best film I've seen since "Ben Hur." (=since I saw "Ben Hur")

Unlike the year 1980 in (13a), the title of the movie "Ben Hur" in (13b) has no intrinsic semantic feature related to time, but the *since* phrase in both expresses a specific time; the *since* phrase in the latter is somehow interpreted as an adverbial clause as indicated above.

In this way, Japanese, English and probably any other languages allow some phrases to have clausal interpretations under specific conditions, and I will argue below that the interpretation of a wh-phrase as an interrogative clause is an instance of this universally available mechanism.

3. A UNIFORM ANALYSIS OF THE TWO INTERPRETATIONS OF WH-PHRASES

It has been demonstrated that a phrase with *made* can specify a spatial endpoint or be construed as an adverbial clause expressing a temporal endpoint. The shared meaning is 'endpoint' and the axis is unspecified, being either temporal or spatial. I would like to deduce two kinds of interpretations for Japanese wh-phrases along this line without assuming distinct internal structures for them. Two questions arise here: (i) what is the shared meaning between indefinite and interrogative wh-phrases with *ka* and (ii) how are they distinguished if they have exactly the same internal structures?

As for the first question, Nishigauchi in [7] makes an important observation on pairs of examples such as (14a,b):

- (14) a. Dare ka kara henna tegami-ga todoita.
 who Q from strange letter NOM arrived 'A strange letter came from somebody.'
 b. Dare kara ka henna tegami-ga todoita.
 who from Q strange letter NOM arrived 'A strange letter came from god knows who.'

The phrase *dare ka* in (14a) can be paraphrased as 'someone' and may be interpreted as having a specific reference in the mind of the speaker. In contrast, the specific interpretation is impossible with *dare kara ka* in (14b); it is associated with the interpretation 'A letter came from someone but I don't know who it is from.' Indefinite expressions we have been concerned with are the second type.³ Nishigauchi concludes that the semantic property of *ka* in (14b) is essentially homogeneous with that of the question morpheme *ka* in interrogative clauses. Let us express that semantic aspect of *ka* with

² Takahashi's analysis in [15] does not face this problem, where it is assumed that a wh-phrase is obligatorily raised into a CP-spec and the elided part is an IP. Wh-movement in Japanese, however, is not obligatory in general. If a wh-phrase to be interpreted as a phrasal interrogative remains in situ, Takahashi's analysis suffers from the problem discussed in the text. See [16] for its other drawbacks.

³ Those of the first type can be treated as lexicalized expressions: *dareka* (someone) and *dokoka* (somewhere) are listed in the lexicon. Note that (ia) is acceptable, while (ib) is unacceptable simply because *dare-no ie ka* is not listed in the lexicon:

- (i) a. John-wa dareka -no ie kara deteki-ta.
 TOP someone GEN house from came
 b. *John-wa dare-no ie ka kara deteki-ta.
 TOP who GEN house Q from came 'John came from someone's house.'

the feature [unspecific]. The unspecified part is whether a wh-phrase with *ka* fails to denote a specific reference or a specific truth value: the former is a case of indefinite and the latter, a phrasal interrogative.

The second question has to do with selectional properties of predicates taking wh-phrases with *ka* as their complements. In the case of phrasal interrogatives, it is presumably *ka* that is selectionally relevant for predicates. An interrogative predicate requires its complement to be marked with *ka*, but the complement does not have to contain a wh-phrase as in (15):

- (15) John-wa Mary-ga koko-ni kita *(ka) siritagatteiru.
 TOP NOM here came Q want-to-know
 'John wants to know if Mary came here.'

In contrast, the selectionally relevant properties of indefinite NP and PP arguments in (1) and (2) are the properties of the phrases before *ka* rather than *ka* itself. The verb *nagur* (beat), for instance, requires a constituent denoting some concrete object that can be beaten, but it need not be an indefinite NP with *ka*; a specific person like John can be the target of beating. Similarly, the indefinite adjunct *naze ka* in (3) can be replaced by a non-indefinite adjunct such as *1-zikan mo* (for one hour), which shows that the first bracketed constituent in (3) need not be a phrase with *ka*. In brief, the properties of *ka* are selectionally relevant in the case of phrasal interrogatives but they are not in the case of indefinites; in the latter, the categories and/or semantic properties of the sister phrases of *ka* are crucial in determining whether they fit into a sentence or not.

Under the assumption that a wh-phrase with *ka* has the same internal structure regardless of its interpretations, it follows that the properties of both *ka* and its sister wh-phrase need to be projected to a complex structure they form. Putting it within the minimalist program, it amounts to saying that the label of a structure constructed from *ka* and a wh-phrase should be the union of their features. This kind of symmetric projection has not been adopted generally in syntactic studies, where either *ka* or its sister phrase is analyzed as deciding the properties of the complex they form. So, if *ka* heads the complex, the selectional properties of phrasal interrogatives can be naturally accounted for, but those of indefinites cannot be. The opposite holds if *ka* is taken as a non-head. It can be said that if grammar allows only asymmetric projection, a unified treatment of indefinite and interrogative wh-phrases is impossible.

In fact, the idea of symmetric projection is not incompatible with the fundamental assumptions in the minimalist program. Specifically, it is argued in [2] that the simplest object constructed from two lexical items α and β is the set $\{\alpha, \beta\}$, where α and β are the sets of their component features. The label of the new object should be made up of either the set α or β . There are three possibilities: (i) the intersection of α and β ; (ii) the union of α and β ; and (iii) one or the other of α, β . Suppose that α is the verb *earn* and β is the noun *money*. *Earn* contains the category feature [verbal], while *money* contains the feature [nominal]. The two lexical items are distinct in value for the category features and possibly for many other features. Thus, (i), which is usually null, is excluded as the label of the structure *earn money*. Neither can the option (ii) be taken in this case. The category features of the two lexical items are incompatible in a single set; the structure *earn money* cannot be both verbal and nominal. Therefore, the structure *earn money* should be labeled in terms of (iii). On the other hand, if there is a case where two lexical items are not distinct in value for any feature, the minimalist framework should offer it a distinct analysis from the one allowed under the X-bar theory, though this possibility is not pursued in [2]. It is mainly due to the category features of two lexical items that make the union of their features incompatible, as in the case of *earn* and *money*. If one of the two lexical items lacks category features, the union of their features contains no contradictory category features and probably no other incompatible features. The label of the object consisting of the two items could be the union of their features in this case.⁴

⁴ Similar ideas have been proposed in [1] and [4]. Note that the unification at stake here is restricted to two sister constituents as its inputs, unlike the more general notion in [11], for instance, which covers satisfaction of argument structure by some constituent.

Going back to (1)-(3), the two bracketed phrases in each should be analyzed as having the same internal structure. If *ka* is assumed to have no category feature, the labels of the two bracketed phrases in (1) repeated below are the unions of the features of the *wh*-phrase and those of *ka*:

- (1) John-ga [NP [N dare] ka] nagutta rasii ga [NP [N dare] ka] sitteiru hito -wa
 NOM someone beat it-seems but who Q know person TOP
 i nai
 exist not
 'John seems to have beaten someone but nobody knows who.'

Specifically, the *ka*-phrases are both categorially NPs due to the category feature of the inside noun *dare*. Moreover, the semantic feature [unspecific], which originates in *ka* and is crucial for the interpretation of the second phrase, projects up to both phrases. Semantic features also come from *dare* (someone/who)> Among them, features such as [human] are selectionally relevant in the interpretation of the first phrase; a human being can be beaten but abstract entities such as an interrogative clause cannot be. (2) and (3) are to be accounted for along this line. A more formal analysis will be presented in the subsequent sections.

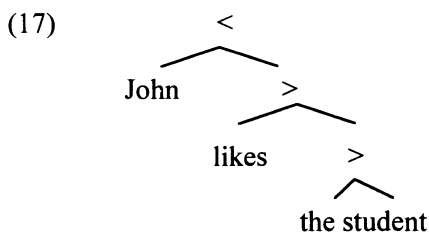
4. INCORPORATION OF SYMMETRIC PROJECTION INTO STABLER'S SYSTEM

4.1 Feature Types

As assumed in most current syntactic theories, Stabler in [13] considers each lexical item as a collection of phonetic, semantic and syntactic features. The syntactic features are classified into four kinds: *category*, *select*, *licensors*, and *licensees*. *Category* is the set of category features such as *v*(erb), *n*(oun), *c*(omplementizer) and *d*(eterminer). *Select* is the set of features of the form =*x*, with *x* ranging over *category*; =*n*, for example, indicates the selection of noun phrase. *Licensees* and *licensors* are features involved in movement. In particular, a noun phrase has a licensee case feature, and a transitive verb has the corresponding licensing feature +*case*. Example lexical entries are given below:⁵

(16)	category	selection	licensor	licensee	semantic	phonetic
the	d	= n		case
John	d			case	[human],
student	n				[human],
likes	v	= d =d	+ case		[stative],

From this limited lexicon, the following binary-branching structure can be derived through the structure-building operations Merge and Move:⁶



⁵ *Category*, here, is what is called *Base* in [13]. Stabler employs -*case* as a licensee feature, but the minus sign is liable to be interpreted as the absence of case feature. Hence, it is not used here.

⁶ In (17), the shell structure with the light verb *v* is omitted, which is irrelevant to the discussion here.

The head is expressed in terms of < (the right constituent is the head) or > (the left constituent is the head).

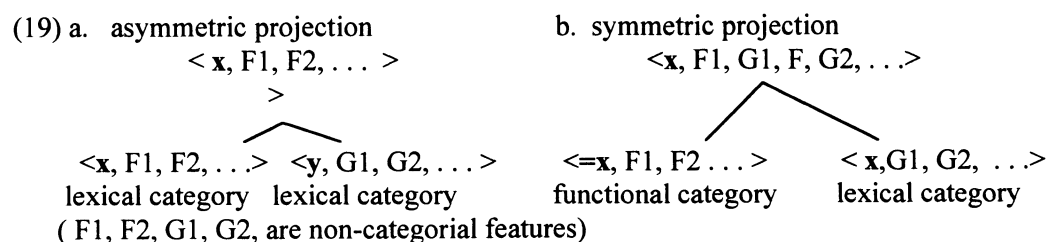
4.2 Category Features

Looking at the lexical entries in (16), we can identify some of Stabler's tacit assumptions in [13]. First, every item is assumed to possess a category feature though it need not have other features: the leftmost columns in (16) are all filled. Moreover, each complex structure made from two inputs is assumed to have one and only one head, which is expressed in terms of either < or >: feature projections are asymmetric in the sense that the complex inherits the properties of either one of the inputs and never both. These two assumptions are shared by most syntactic theories, but they have been given up in the previous section, where the necessity for symmetric projection (or the union as a label of complex structure) has been discussed. More specifically, *ka* has been assumed to have no category feature, and *ka* and a *wh*-phrase form a structure that is labeled in terms of the union of their features. I will make a more general claim: functional categories all lack category features, and the complex they form is labeled symmetrically in terms of their non-categorial features as well as their sister's features.

Of course, asymmetric projection of the kind assumed in X-bar theories is also necessary. How, then, can cases of symmetric projection be distinguished from those of asymmetric projection? A quite simple algorithm can be devised. Suppose that the syntactic and semantic features of each lexical entry are listed in the following order:⁷

(18) *category -- selection -- licensor/licensee -- semantic*

Then, the first feature is a category feature in the case of lexical category, and it is a feature of categorial selection in the case of functional category. If the binary-branching system is taken, a pair of input constituents can be two (projections of) lexical categories as in (19a), or one lexical category and one functional category as in (19b):



In (19a), the first features of the two inputs are category features and they are incompatible in a single set; thus, the complex they form needs to be a case of asymmetric projection headed, in this case, by the left constituent. The complex inherits features only from the head. In (19b), the functional category does not have a category feature that would be in conflict with the category feature of the lexical category; hence, it is a case of symmetric projection, inheriting features from both constituents. Note that the first feature of the functional category is about categorial selection, and it is checked off appropriately by the category feature of the lexical category. A case of merging two functional categories is subsumed under (19b): one of them has inherited a category feature from some lexical category it dominates.⁸

⁷ See [13]. Within each set, features are assumed to be unordered; **case** and **WH**, for instance, are not ordered with each other within the set *licensor/licensee*.

⁸ A complementizer selects the functional category Tense, which in turn selects the lexical category V; this case falls under (19b) with the right constituent being the complex built from Tense and VP.

4.3 Selectional Features

Another characteristic of Stabler's system that calls for a modification resides in selectional features: the selectional feature of each item, if any, expresses its categorial selection since x in the selectional feature $=x$ is assumed to range over the set of category features. As Pesetsky (1982) and Abney (1987) point out, however, selectional properties of lexical categories such as verbs and nouns are essentially semantic; their semantically selected arguments are realized just typically as particular syntactic categories. The following examples illustrate this point:

- (20) a. [DP The boy] killed [DP a cat]
b. [DP John] went [PP to New York]/[AdvP there]
c. John considered [CP that Mary was incompetent]/[IP Mary to be incompetent]/
[AP Mary incompetent]/[PP Mary below the standard]/ [NP Mary a good candidate]
d. John persuaded Mary [CP PRO to leave]/[PP into leaving]

Kill takes two arguments 'killer' and 'his/her victim'; they are typically realized as DPs. The internal argument of *go* is typically realized as a PP but it can be an adverbial phrase, as in (20b). The propositional arguments of *consider* and *persuade* exhibit a much wider variety of categorial realization, as shown in (20c,d).

In Stabler's more recent model outlined in [10], selectional features of the form $=x$ are replaced by hybrid features of semantic and categorial selections such as θd , which presumably means that it takes a DP argument of the thematic type θ . Expressing argument structures of verbs and other lexical categories in terms of these hybrid features seems to be appropriate, but it cannot be extended to functional categories; they have no thematic relations with their complements, and their selections of complements are purely categorial. I will therefore adopt two types of selectional features, $=\theta$ for lexical categories and $=x$ for functional categories. Then, the lexical category *go*, and the determiner *the*, which is a functional category, are characterized as follows:

- (21) a. *go* : $=\theta$ (goal), $=\theta'$ (agent) b. *the* : $=n$

One of the distinctions that are standardly assumed between lexical and functional categories can be captured in terms of these two types of selectional features.

4.4 Features and Structure-Building

Five kinds of features have been assumed: category features such as v and n , features of semantic/categorial selection such as $=v$ and $=\theta$, licensor/licensee features such as $+case$ and $case$, semantic features, and phonetic features. The order in (18) refers to the three kinds of syntactic features as well as semantic features; phonetic features should be set aside as irrelevant to syntactic structure building. Among the four, category and semantic features remain intact throughout syntactic derivation since they are relevant to semantic interpretations at the interface level, but selection and licensor/licensee features are present only for driving structure-building operations in syntax. I will assume, following the standard approach in the minimalist program, that the order specified in (18) controls structure-building processes: syntactic features other than category features will be checked off by appropriate corresponding features in this order through Merge or Move, and the resultant structure is a legitimate output of syntax if it contains semantic and category features, but no other syntactic features.

5. A FORMALIZATION

Let us formalize the first approximation made in Section 3, starting with (1) repeated below:

- (1) John-ga [dare ka] nagutta rasii ga [dare ka] sitteiru hito wa i nai.
 NOM someone beat it-seems who Q know person TOP not-exist
 'John seems to have beaten someone but nobody knows who.'

Dare (who) is a noun, so it should bear the category feature **n**. Moreover, its special status as a wh-phrase is to be expressed with the licensee feature **WH** by analogy to the case feature **case**. As for *ka*, it is assumed to have no category feature but to have the licenser feature **+WH**, which is reasonable. Then, *dare* (who) and *ka* (Q) are analyzed as (22a,b), respectively, where [human] and [unspecific] are semantic features:

- (22) a. *dare*: < **n**, **WH**, **acc. case**, [human], ... >⁹ b. *ka*: < **+WH**, [unspecific], ... >

Merging *dare* with *ka* produces the following structure:

- (23)
- | | |
|--|------------------------------------|
| < n , acc. case , [human], [unspecific], ... > | |
| < n , WH , acc. case , [human], ... > | < +WH , [unspecific], ... > |
| dare | ka |

The topmost label is decided symmetrically by *dare* and *ka*, being the union of their features except for the licensee feature **WH** and the licenser feature **+WH**, which have been checked off through the operation Merge. Crucially, the label inherits the category feature **n** and the semantic feature [human] from *dare* in addition to the semantic feature [unspecific] from *ka*

The first bracketed phrase in (1) is the internal argument of the verb *nagur*, (beat), which is to be analyzed as < **v**, =**θ** (patient), =**θ'** (agent), **+acc. case**, [volitional], ... >. The verb has the category feature **v**: it takes two arguments, which are expressed as =**θ** (patient) and =**θ'** (agent); and it has the licenser feature **+acc(usative) case**. Merging the verb with (23) will produce (24):

- (24)
- | | | |
|--|---|-------|
| < v , = θ' , +acc. case , [volitional], ... > | | |
| < n , acc. case , [human], [unspecific], ... > | < v , =θ , = θ' , +acc. case , [volitional], ... > | |
| dare ka | | nagur |

Dare ka and *nagur* have category features **n** and **v**, respectively; this should be a case of asymmetric projection. Thus, the topmost label inherits the category feature **v** and other features only from the verb *nagur*. The selectional feature =**θ** (patient) of the verb, which means that it requires a patient argument, is checked off when the verb merges with *dare ka*; the latter has the semantic feature [human], and concrete objects like a person can be beaten.

Under the predicate-internal subject hypothesis, the next step is to merge (24) with another complex structure *John-ga* (NOM) as in (25); it is a case of asymmetric projection:

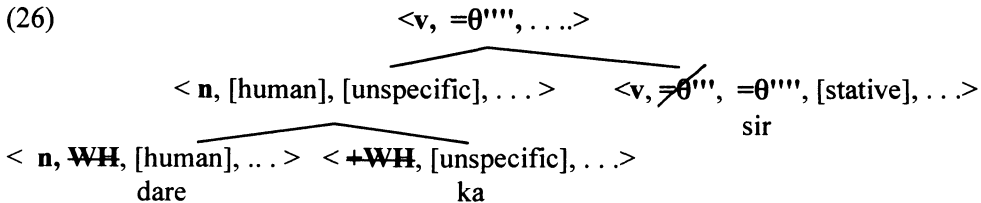
- (25)
- | | | |
|--|--|-------|
| < v , +acc. case , [volitional], ... > | | |
| < n , nom. case , [human], ... > | < v , =θ' , +acc. case , [volitional], ... > | |
| John-ga | | |
| | < n , acc. case , [human], [unspecific], ... > | nagur |
| | dare ka | |

In (25), the case-marker *ga* is assumed to lack a category feature just like *ka*. *John-ga* inherits the category feature **n** and the semantic feature [human] from the noun *John*, while inheriting the

⁹ According to Nishigauchi [7], the first occurrence of *dare ka* in (1) is underlyingly *dare-o-ka*. So, exactly speaking, the case feature comes from the accusative case-marker *o*, which is to be deleted before *ka* and other adverbial features such as *wa*. See also [16].

licensee feature **nom(inative) case** from *ga*. The second selectional feature = θ' (agent) of the verb is properly checked off; *John-ga* denotes a human and a human can be a beater of someone. After the selectional requirements of the verb are all satisfied, the licensee feature (**acc. case**) of *dare ka* and the licenser feature (**+acc. case**) of the verb will check off each other. Finally, the structure formed so far merges with the past tense *-ta*, and the licensee feature (**nom. case**) of *John-ga* is to be checked off through this process.¹⁰

The second bracketed phrase in (1) and the verb *sir*, which takes it as its complement, are analyzed as constituting the structure in (26):¹¹



The feature of semantic selection, = θ''' , is about the interrogative complement of the verb *sir* (know). Remember that the semantic feature shared by indefinites and interrogatives is [unspecific] and that an interrogative lacks a specific truth value. Then, *dare ka* in (26) can satisfy the selectional requirement imposed by the feature = θ''' , and the latter is properly checked off. The remaining steps in deriving the second conjunct of (1) are irrelevant and thus set aside from the discussion here.

As has been discussed in Section 2, phrasal interrogatives of this kind require discourse-level information for their full interpretations. In particular, the second conjunct of (1) will ultimately be interpreted on a par with the full sentence below:

- (27) [John-ga dare-o nagutta ka] sitteiru hito wa i nai.
 NOM who -ACC beat Q know person TOP exist not
 'Nobody knows who John hit.'

The kind of discourse-level interpretive mechanism at stake is highly complex and falls outside the domain of syntax.¹²

Skipping (2), consider (3) repeated below:

- (3) John-wa [naze ka] naiteita ga [naze ka]-wa daremo siranai.
 TOP some reason was-crying but why Q TOP nobody know
 'John was crying for some reason but nobody knows why.'

The second bracketed phrase is to be interpreted as a phrasal interrogative just like the one in (1). One difference is that the first bracketed phrase is an adjunct rather than an argument of the verb *nak* (cry). Merging a projection of the verb *nak* with the first bracketed phrase, which denotes some unspecific reason, calls for no explanation just like merging it with a phrase expressing a specific reason such as *maketa kara* (because he was defeated).

Last, *naze ka* in (4) constitutes a phrasal interrogative complement under its reading (5a). On the other hand, *naze ka* is an adjunct under (5b), and the complement of the verb is not overtly realized, which is generally possible in Japanese. No further explanation is necessary.

¹⁰ Accusative case-checking is assumed to be mediated by the light verb *v*, and the details of nominative case-checking are ignored here.

¹¹ For the absence of the feature **acc. case**, see note 9.

¹² Unlike clauses, tense features are not contained in phrases to be interpreted as time adverbial or as interrogative clauses. Thus, the absence of tense features might trigger discourse-level interpretive operations.

6. SOME CONSEQUENCES: WH-PHRASES IN ENGLISH

It has been claimed in Section 4.2 that functional categories in general lack category features. If it is correct, the English interrogative complementizer, which is to be expressed as [WH], has no category, just like *ka*. Why, then, is the indefinite usage of a wh-phrases impossible in (31a)?

- (28) a. *John seems to have beaten who.
 b. John seems to have beaten [DP[WH][DP who]].

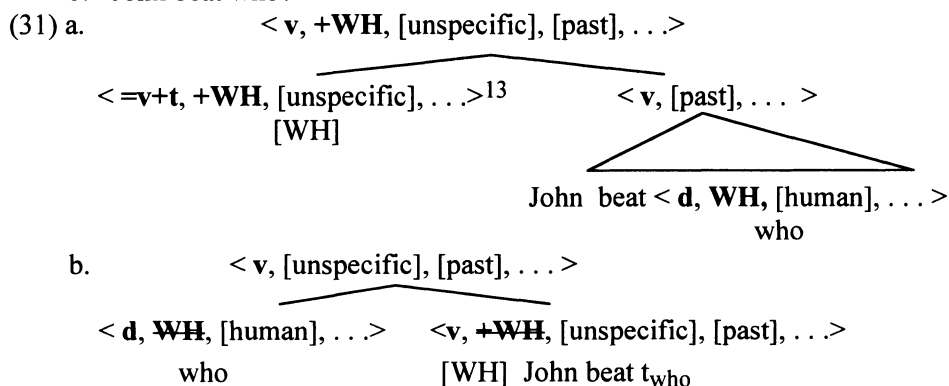
If [WH] in (28b) has no category, merging it with the DP *who* is a case of symmetric projection, and no selectional violation should occur between the verb *beat* and the DP. What appears to be offending is the relation between *who* and [WH]. Remember that *ka* can merge with a variety of categories, as exemplified in (1)-(3). The fact shows that *ka* has no selectional feature. In contrast, [WH] and other complementizers in English select a TP [1]. If so, (28a), with the structure in (28b), fails to satisfy this selectional requirement.

More generally speaking, I will assume one parametric difference as to functional categories: functional categories in some languages have features of categorial selection, as in English, while those in other languages do not, as in Japanese. The main claims are summarized below:

(29)	category	semantic selection	categorial selection	licensor/licensee
lexical category	yes	yes	no	yes
functional category	no	no	yes (English) no (Japanese)	yes

The proposed parameter can account for another major difference: obligatory wh-movement in English and its absence in Japanese. For instance, (30a) can be analyzed as having the intermediate structures in (31a,b) in its derivation:

- (30) a. Who did John beat?
 b. *John beat who?



The right constituent of (31a) is a projection of the verb *beat* with the tense feature [past], so that the selectional feature of [WH] is properly checked off. Suppose that in a given constituent, only one syntactic feature can be checked off by a single application of Merge or Move. Then, the licensor feature +WH of [WH] and the licensee feature WH of *who* are intact. If the derivation stops at this level, it will not converge due to the unchecked features: +WH, and WH. This is why (30b) is not allowed. Raising *who* and merging it with the structure in (31a) will produce (31b), where *who* and the constituent containing +WH become sisters and their unchecked features can properly delete each other.

¹³ [WH] selects a projection of Tense, which is expressed as =v+t. Alternatively, tense inflections, which have semantic content, can be analyzed as lexical categories. Then, they have the category feature t, and the selectional feature =t can be assumed for all complementizers in English.

In contrast, *ka* in Japanese lacks a selectional feature; its only syntactic feature to be checked off is the licenser feature +**WH**. It is reasonable to assume that this feature can be checked off simply by merging *ka* with a constituent containing the feature **WH**. The constituent at stake can be a wh-pronoun such as *dare* (who), a wh-phrase such as *dono hon* (which book), or an entire clause containing a wh-pronoun somewhere. In this way, checking of **WH** and +**WH** in Japanese can be completed by the first application of Merge to *ka* without inducing overt wh-movement due to its lack of selectional feature.

To summarize the discussion in this paper, the indefinite and phrasal interrogative interpretations of wh-phrases in Japanese have been explained under the assumptions that *ka* and functional categories in general lack category features and that a complex structure constructed from two inputs can be labeled in terms of the union of their features as long as one of them lacks a category feature. Two of the major differences between Japanese and English have been shown to be deducible from these assumptions, together with one parameter: functional categories in English have features of categorial selection, while those in Japanese do not.

REFERENCES

- [1] S. P. Abney. The English noun phrase in its sentential aspect, Doctoral dissertation, MIT, Cambridge, Mass, 1987.
- [2] N. Chomsky. The minimalist program, Cambridge, Mass.: MIT Press, 1995.
- [3] N. Chomsky. "Minimalist inquiries: the framework," ms. MIT, 1998.
- [4] A. M. Di Sciullo and E. Williams. On the definition of word, Cambridge, Mass.: MIT Press, 1987.
- [5] G. Gazdar, E. Klein, G. Pullum, I. Sag. Generalized phrase structure grammar, Oxford: Basil Blackwell, 1985.
- [6] Gunji T. Japanese phrase structure grammar, Dordrecht: D. Reidel, 1987.
- [7] Nishigauchi T. Quantification in the theory of grammar. Dordrecht: Kluwer, 1990.
- [8] D. Pesetsky. Paths and categories. Doctoral dissertation, MIT, Cambridge, Mass, 1982.
- [9] R. Quirk, S. Greenbaum, G. Leech, and J. Svartvik. A grammar of contemporary English, London: Longman, 1972.
- [10] C. Retore and E. P. Stabler. "Resource logics and minimalist grammars," INRIA Technical report No. 3780, 1999.
- [11] S. M. Shieber. An introduction to unification-based approaches to grammar, CSLI Lecture Notes, No. 4, Stanford: Center for the Study of Language and Information, 1986.
- [12] E. P. Stabler. "Parsing and generation for languages with movement," In C. R. Berwick, ed. Principle-based parsing: from theory to practice, Boston: Kluwer, 1996.
- [13] E. P. Stabler. "Derivational minimalism," In C. Retore, ed. Logical aspects of computational linguistics. New York: Springer, pp. 68-95, 1997.
- [14] E. P. Stabler. "Acquiring languages with movement," Syntax 1: 72-97, 1998.
- [15] Takahashi D. "Sluicing in Japanese," Journal of East Asian Linguistics 3:265-300, 1994.
- [16] Yasui, M. "Categorial Status of the Question Morpheme *Ka* in Japanese and Bare Phrase Structure," In M. Muraki and E. Iwamoto, ed. Linguistics: in search of the human mind: a Festschrift for Kazuko Inoue, Tokyo: Kaitakusha, 1999.