

Ellipsis of case-markers and information structure in Japanese

Yo Sato^a and Tam Wai Lok^b

^aDepartment of Computer Science, University of Hertfordshire
College Lane, Hatfield, HA10 8AB
satoama@gmail.com

^bNational Institute of Advanced Industrial Science and Technology
Aomi 2-3-26, Koto-ku, Tokyo 135-0064, Japan
tamuwairokku@aist.go.jp

Summary. This paper presents an HPSG formalisation of how the ellipsis of case-marking affects the focus of the clause in Japanese. We restrict our attention to the nominative and accusative markers *ga* and *o*, and in view of the fact that the ellipsis effects on focushood vary between 1) *ga* and *o* and 2) different argument structures of the head verb, develop an account that combines both aspects, i.e. case and argument structure of the verb.

Keywords: Case-marking, Information structure, focus, Japanese

1 Introduction / Background

Postpositions are crucial building blocks of a clause in Japanese. They attach mainly to nominals, indicating their semantic, syntactic or pragmatic properties, and are projected to postpositional phrases (PostPs). Since PostPs become dependents of a verb in a clause, postpositions crucially participate in argument structure realisation. Yet some postpositions, case-marking ones among others, are frequently dropped, giving rise to the situation where some arguments in a clause are headed by phonologically empty items. In this paper we ask under what conditions such an ellipsis occurs, in relation to the *information structure*. More specifically we argue that overt marking is required in a focal environment, although the *default focus* specified in the head verb is exempt from this requirement. We show an HPSG formalisation for this mechanism.

We restrict our target to case-marking postpositions (case-marker in short) in informal discourse, particularly *ga* (nominative, normally subject marker as well) and *o* (accusative, normally direct object marker), where ellipsis frequently occurs. To avoid complication we do not consider the topic marker *wa*, which also tends to be elided, although we will say a word or two at the end about how our analysis may be extended to cover it. We will therefore use examples of embedded clauses, where the topic marking is normally suppressed.

First we consider three types of existing accounts. A first type, the intuitive ‘recoverability’ account (Kuno, 1987), says that a case-marker may be dropped if the thematic role is uniquely identifiable, or, recoverable. According to it, in the following example, it is felicitous to drop *o* in (1a) because thanks to the argument structure of the transitive verb ‘*ijimeru*’ coupled with the presence of the subject-marker *ga*, the objecthood of the case-dropped nominal, *Jiro*, can be recovered (the bracketed variant demonstrates that the word order swap doesn’t affect the omisibility). The infelicity of (1b), then, can be attributed to the impossibility to uniquely identify subject or object (either Taro or Jiro can be subject or object). In (2), though the thematic roles are not recoverable syntactically, they can be argued to be recovered pragmatically, because alcohol cannot be plausibly interpreted to drink Taro.

- (1) Taro-ga Jiro-o ijimeten-no michatta.
 -NOM -ACC bully-COMP saw
 lit: (I) saw that Taro was bullying Jiro. ('I saw Taro bullying Jiro')
- a. Taro-ga Jiro- ϕ ijimeten-no michatta. (Jiro- ϕ Taro-ga ijimeten-no michatta.)
 b. ? Taro- ϕ Jiro- ϕ ijimeten-no michatta.
- (2) Taro- ϕ sake- ϕ nonden-no michatta. ('(I) saw that Taro was drinking alcohol')

However, recoverability cannot account for the apparent asymmetry in omissibility between *ga* and *o* in the following examples. The second example is meant to show that this is not due to pragmatic plausibility either.

- (3) a. ?? Taro- ϕ Jiro-o ijimeten-no michatta.
 b. ?? Taro- ϕ sake-o nonden-no michatta.

Based on such examples Kageyama (Kageyama, 1993) offers a structural constraint: that an internal argument does, while an external argument does not, allow for ellipsis. This structural account may work for most of our examples, but apart from the fact that we already have an exception (2), the main problem with such a binary criterion, which makes the situation seem black-and-white, is that the reality isn't. First, the subject-marker drop is acceptable with an intransitive verb:

- (4) Taro-ga/ ϕ hashitteru-no michatta. ('(I) saw that Taro was running')

Further, under the context where the subject is de-focused, the acceptability improves.

- (5) Taro-ga nani shiten-no mitatte? ('What did you say you saw Taro doing?')
 Taro- ϕ Jiro-o ijimeten-no michatta.

The last example suggests the likelihood of the involvement of information structure, and in fact so does (4), because the intransitive subject tends to be focused than the transitive one. The last account to consider is indeed information-structure oriented: that a case-marking postposition cannot be elided if the nominal that it attaches to receives a 'narrow' focus (Yatabe, 1999).

- (6) a. Dare-ga sake-o nonden-no mitano? ('Who did you see drinking alcohol?')
 Taro- $\{ga/??\phi\}$ sake- $\{o/\phi\}$ nonden-no michatta.

However, there does not seem to be a problem eliding marking with the focused object [(7a)].

- (7) a. Taro-ga nani-o nonden-no mitano? ('What did you see Taro drinking?')
 Taro- $\{?ga/\phi\}$ sake- $\{o/\phi\}$ nonden-no michatta.
 b. Nani-o mitano? ('What did you see?')
 Taro- $\{ga/??\phi\}$ sake- $\{o/\phi\}$ nonden-no michatta.

Furthermore, as (7b) shows, no difference is observed in terms of case-marker omission between the narrow focus (6a) and sentence-focus contexts (7b). Thus the narrow-focus account does not fully capture the situation either. Also notice that (7a)'s relative infelicity (compared with 6a) in overt subject marking. This is a curious case, because it cannot be explained by either account, and in our view, an illuminating one.

We believe that it still is possible to combine the valid insights of both accounts. In what follows we develop an account that says some arguments require case-marking to be focused while others may be focused regardless of it, or put differently, case-marking determines focus for some arguments while it is inert for others. We then relate this difference —what arguments require and does not require case-marking— to information structure. Recall the dictum "*Subject is the unmarked topic*" (Li and Thompson, 1976): even without context, the default information structure can be conventionalised in a way related to argument structure. We say *Object is the*

default focus, although we offer an analysis that accounts for all the data including the object-marker drop.

2 Focus and case-marker elidability

2.1 Overview

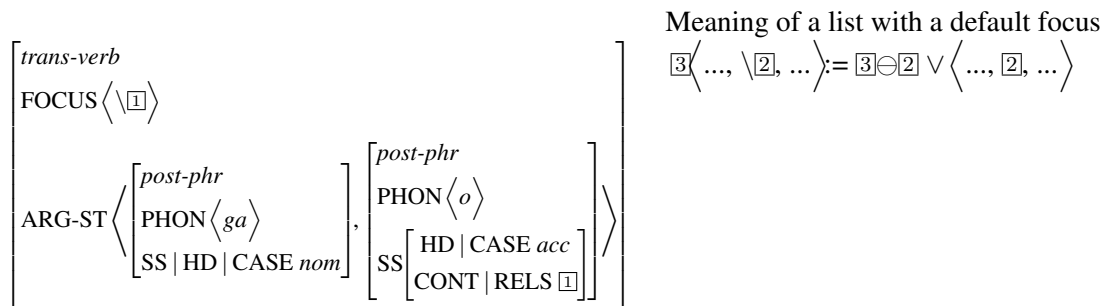
Our overall strategy is to specify one¹ of the arguments of a verb as the ‘default focus,’ and say it does not require a case-marker to be realised as a focus. The corollary of this is that an argument which is not the default focus requires a case-marker to be a focus. From the ellipsis perspective, the claim amounts to this: to be able to elide a case-marker, a PostP needs to be either a non-focus or the default focus.

Before proceeding to the formal analysis, it would be necessary to refer to the existing proposals on focus in general, so as to clarify what we share and do not share with them. First to be noted is the fact that, since the seminal work by Selkirk (Selkirk, 1995), most work on focus centres around *prosody*, i.e. focus-marking (‘F-marking’) by accentual or intonational means. We certainly accept that focus can be influenced by these means in Japanese too (see e.g. (Ohshima, 2006)), but since our subject matter is case-marker ellipsis we will only discuss the syntactic aspect of focus in this connection. Furthermore, whereas the prosodic F-marking is usually considered to identify focus unambiguously (the F-marked constituent automatically gets the focus status), this does not apply to our syntactic analysis. In fact, our analysis will not include ‘focus-marking’ as such, that is, there is no single feature that uniquely determines focus (reflecting the absence of obvious syntactic focus marking in Japanese). Instead it is a combination of argument structure and case-marking that does the work.

Another important issue is what ‘projection’ schema we adopt for focus. On this point, what is common among differing proposals is a rule or schema that lets a mother inherit (at least) some of the daughters’ focus values. We use a straightforward variant of Selkirk’s (Selkirk, 1995) classical ‘vertical’ projection rule, which states that the focus values of the head daughter is inherited to its mother, and add our own mechanism where focus is contributed from non-head daughters.

2.2 HPSG formalization

Now for the analysis, and we start with specifying the default focus in the arguments of a verb, to capture the ‘conventionalised’ nature of a clause’s information structure. The lexical specification will give a flexibility in what argument is to be focused by default, and we will come back to specific advantages of this setup later. For our discussion, we assume the *trans(itive)-verb* subtype to behave uniformly with respect to default focus: the object is the default focus. Along with the past HPSG literature (e.g. (De Kuthy, 2002)), the focus feature is a list ranging over the semantic content values, but the backslash (‘\’) notation points to the default focus.



The meaning of a focus list containing a default element is represented on the right. Simply put, it represents a disjunction of lists, which says the default element may be either realised as a focus or not.

¹ Although in the data we discuss the number of default focus elements is one, there is no theoretical reason why it should always be one. We will allude to the case there is none later. There can be more than one also.

Now, to incorporate the effect of the presence/absence of the overt case-marker, we introduce the binary *focus-marking potential* (FMP) head feature for case-marking postpositions.² We simply say that any overt case-markers have this potential (value *plus*) and phonologically empty ones (zero-markers) do not.

$$\begin{array}{c}
 \left[\begin{array}{l} \textit{post-case-overt} \\ \text{PHON } \textit{nelist} \\ \text{SS | LOC | HD} \left[\begin{array}{l} \text{CASE } \textit{case} \\ \text{FMP } \textit{plus} \end{array} \right] \end{array} \right] \left[\begin{array}{l} \textit{post-case-zero} \\ \text{PHON } \langle \rangle \\ \text{SS | LOC | HD} \left[\begin{array}{l} \text{CASE } \textit{case} \\ \text{FMP } \textit{minus} \end{array} \right] \end{array} \right] \\
 \\
 \left[\begin{array}{l} \textit{phrase} \\ \text{FOCUS } \langle \mathbb{1} \rangle \\ \text{HD-DTR} \left[\text{FOCUS } \langle \mathbb{1} \rangle \textit{list} \right] \end{array} \right] \vee \left(\begin{array}{l} \left[\begin{array}{l} \textit{verbal} \\ \text{FOCUS } \langle \mathbb{3} \oplus \mathbb{1} \rangle \\ \text{HD-DTR} \left[\begin{array}{l} \textit{verbal} \\ \text{ARG-ST } \langle \dots \langle \mathbb{2} \rangle, \dots \rangle, \\ \text{FOCUS } \langle \mathbb{3} \rangle \end{array} \right] \end{array} \right] \\ \left[\begin{array}{l} \text{NHD-DTRs } \langle \mathbb{2} \rangle \left[\begin{array}{l} \textit{post-phr} \\ \text{SS | LOC | CONT | REL } \langle \mathbb{1} \rangle \\ \text{FMP } \textit{plus} \end{array} \right] \end{array} \right] \end{array} \right) \wedge \left(\mathbb{1} \notin \textit{def-focus}(\mathbb{3}) \right)
 \end{array}$$

Figure 1: Focus Projection Schema

Our focus projection schema is shown in Figure 1. We first adopt a simple schema where the head daughter’s focus value is passed up to the mother (base case, on the left-hand side in the Figure). The crucial step, then, is to add a provision for the interaction of default focus with FMP. The right-hand AVM represents this. There are three cases to consider (or more precisely two subcases with one having further two subcases), depending on (a) whether the PostP in question is the default focus or not, and (b) if it is not, whether it has a positive FMP (b1) or not (b2). However we only need to take care of (b1) in our additional provision, where the focus value is contributed from the non-head daughter. This is because in (a) the default focus is already specified in the head, and in (b2) no focus is contributed from the daughters (*def-focus* simply extracts the default focus from the FOCUS list.)

Example projections are shown in Figure 2, using the ‘Taro bullies Jiro’ example. In the interest of space, the reader will find both overt and ellipsis cases at some nodes: the AVM left to ‘/’ represents the overt case, and the one on the right the elided case. The main point is the contrast between the process combining the overtly marked object PostP (*Jiro-o*) with the verb, and the process combining the overtly marked subject (*Taro-ga*) with the VP. In the former, because it is the default focus, the PostP, though with its positive FMP, does not contribute a focus, leaving the work for the base projection schema. In the latter, in contrast, because it is not the default focus, does contribute its REL value as the additional focus of the clause. Thus there are two possible outcomes for the focus of the clause, shown underneath the tree with their values spelt out. The first one is what obtains for the case of the overtly marked subject, where the focushood of the subject is registered unambiguously, while the second the case of the subject without an overt case-marker. For both cases, the focushood of the object is left ambiguous.

We do not require a special apparatus for the sentence-focus context as far as our subject matter, case-marker ellipsis, is concerned since no difference in this regard is observed in comparison

² The term is deliberately reminiscent of (De Kuthy and Meurers, 2003)’s *focus projection potential*. The role it plays is rather similar, but as we shall see, our ‘potential’ is exerted ‘vertically’ to influence its mother, while DeKuthy and Meurers’s is ‘horizontally’ to its head sister.

with the narrow-focus context. We say what influences case-marking is whether a PostP *participates* in focus, not the range of focus.

It is straightforward to extend the account to intransitive verbs. All that is required is simply to set the subject to the default focus, which annuls the case-marker's potential to contribute focushood, just as in the object in the previous example. Also, regardless of the presence of case-marking, the subject, now a default focus, can be a focus, as in (4).

3 Raminifications and possible extensions

Since the default focus is specified lexically in verbs, this specification can be adapted to various subtypes of verbs. For example Kageyama (ibid.) claims, pushing his account based on the internal/external dichotomy, that ellipsis is allowed for *unaccusative* intransitive verbs but not for *unergative* ones. Apart from its validity, such a possibility can be easily accommodated, by specifying as the default focus its subject for the latter type and none for the former.

An interesting prediction our analysis makes, where an overtly marked PostP that is not the default focus must become a focus, that such a PostP would be infelicitous in a de-focused context. We contest it is the case, as shown in 6a, where overt marking is observed to be less felicitous than the null-marking case. If one pushes this point further, this leads to the view that the absence of case-marking is in some contexts obligatory and hence is not exactly 'ellipsis' but a contextually-driven decision to 'zero-mark' a nominal, as argued by (Shimojo, 2006).

A further issue arises when a language is equipped with explicit and unambiguous syntactic focus-marking just like the accentual F-marking. In Ryukuan, the only language known to historically related to Japanese, has the focus-marker *du*. An analysis of such a focus-marker would perhaps revert us back to the traditional focus-projection debate, but the fact that the language also has case-markers in parallel warrants the interesting investigation as to whether the two sets of postpositions interact to determine focus, for if the default mechanism is still found for case markers, such a mechanism would be independent of explicit focus marking, and if not, it may be rendered unnecessary by explicit focus marking.

Lastly, but most importantly, the possible application of our proposal to the matrix clause should be sought. It may help account for the fact that, in a matrix environment, it is clearly infelicitous to overtly mark a de-focused subject of a transitive verb in Japanese.

- (8) Taro-ga Jiro-o doushitatte? (What did Taro do to Jiro?)
Taro-{??ga/phi/wa} Jiro-o ijimeteru.

Although this is usually considered the case of a topic-marker (*wa*) ellipsis (precisely the reason why we have been avoiding the matrix environment), it may well be related to what we discussed, at least in two ways. First, it may be that the ellipsis of *ga* is a pre-condition for the *wa* marking. This hypothesis is supported by the fact that *ga* and *wa* cannot co-occur. Second, the presence or absence of *wa* may also be accounted for by a similar mechanism, whereby the subject may be registered as the default topic. These hypotheses however require careful treading to confirm, a thorough study in particular on various constructions and contexts.

With this proviso, another important extension would be to extend the analysis to a structurally similar language like Korean. It is well attested that the information-structure interpretation of the subject-marker (*i/ga*) in Korean in some constructions is different from Japanese, in which focus implication is much stronger.

- (9) (To the question, 'what's your name?')
Japanese: ??? Watashi-no namae-ga Taro-desu.
Korean: Jae ireum-i Taro-ibnida.

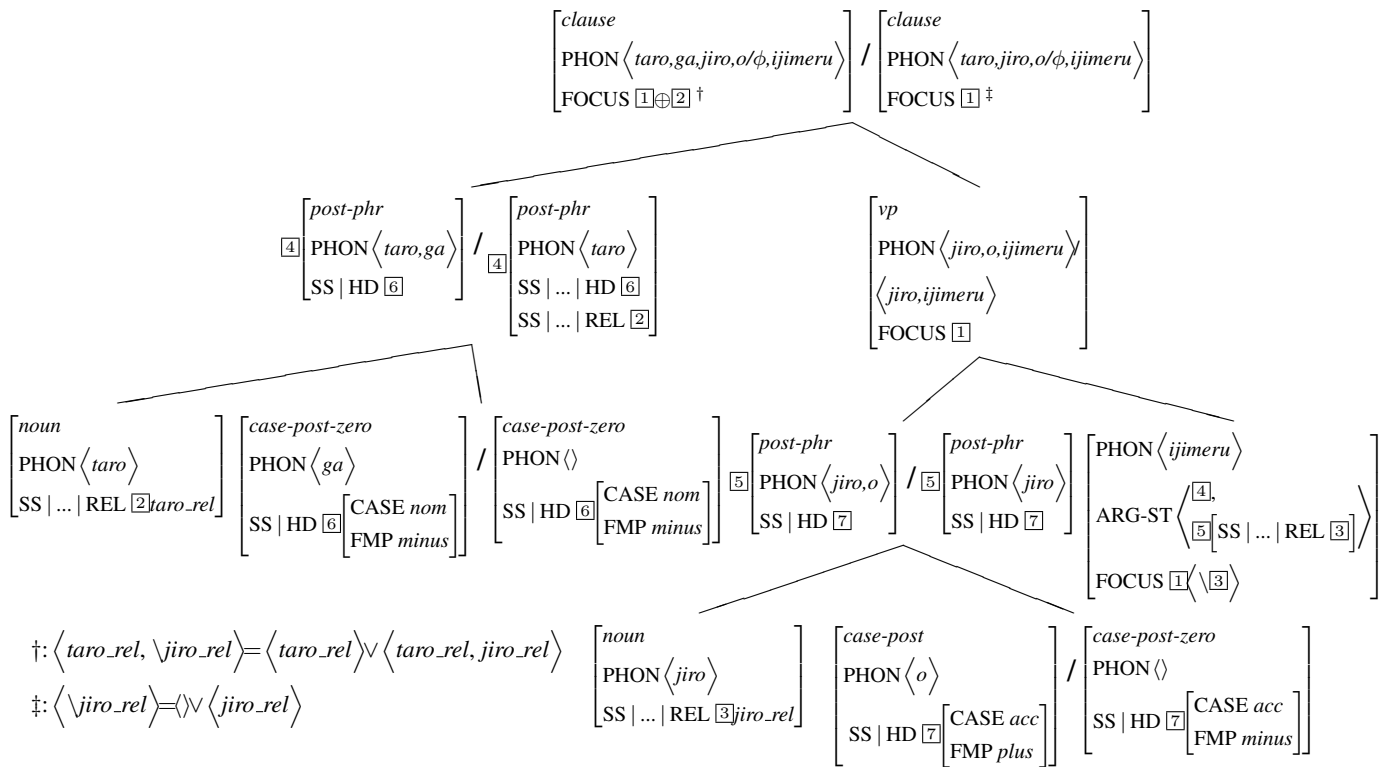


Figure 2: Focus projections for *Taro-ga/φ Jiro-o/φ ijimeru*

Such a difference, if it is the difference in default focushood as it appears to be, can be accounted for by lexically differentiating the default specification for each language.

References

- De Kuthy, K. (2002). Discontinuous NPs in German: A case study of the interaction of syntax, semantics and pragmatics. In *Studies in Constraint-Based Lexicalism*. Stanford: CSLI Publications.
- De Kuthy, K. and Meurers, D. (2003). The secret life of focus exponents, and what it tells us about fronted verbal projection. In *Proceedings of HPSG 03*.
- Kageyama, T. (1993). *Bunpô to Go-keisei (Grammar and word formation)*. Hitsuji Shobo.
- Kuno, S. (1987). *Functional Syntax*.
- Li, C. and Thompson, S. (1976). Subject and topic: a new typology of language. In *Subject and topic*. New York: Academic Press.
- Ohshima, D. Y. (2006). Boundary tones or prominent particles?: Variation in Japanese focus-marking contours. In *Proceedings of Berkeley Linguistics Society*.
- Selkirk, E. (1995). Sentence prosody: Intonation, stress, and phrasing. In *The Handbook of Phonological Theory*. Blackwell.
- Shimojo, M. (2006). Properties of particle omission revisited. *Toronto Working Papers in Linguistics*, 26:123140.
- Yatabe, S. (1999). Particle ellipsis and focus projection in Japanese. *Language, Information, Text*, 6:79–104.