

Constituency, word order and focus projection

Takafumi Maekawa

University of Essex


tmaeka@essex.ac.uk

1 Introduction

Within the standard version of HPSG (Pollard and Sag 1989, 1994), word order is dealt with in terms of a set of linear precedence (LP) rules which state the possible permutations of constituent in a local tree. In this framework, a relatively flexible word order of Japanese illustrated by (1) can be accounted for quite easily.¹

- (1) a. Taroo-wa [_S tukue-ni hana-o oita]
Taro-TOP desk-LOC flower-ACC put
b. Taroo-wa [_S hana-o tukue-ni oita]
Taro-TOP flower-ACC desk-LOC put

If we assume that the constituent marked with the square brackets has the structure like (2a) in which lexical verbal head and one or more complements form a constituent, the general LP rule (2b) can alone give an account for the word order variation in (1), which has been often called ‘scrambling’.

- (2) a.  b. [] ≤ HEAD[LEX +]

LP rule (2b) states that in any phrase containing a lexical head among its immediate daughters, the head has to follow all of its sister constituents. Conforming to this rule, the lexical head daughter *oita* ‘put’ follows all other daughters both in (1a) and (1b). No further rules are needed to predict the word order variation as in (1)

However, recent years have seen the emergence of a view in which linear order is independent to a considerable extent from constituency and is analysed in terms of a separate level of ‘order domains’ (Pollard *et al.* 1993, Reape 1994, and Kathol 2000). The connection between the linear position of a *wa*-marked nominal in a sentence and possible domains of contrastive focus suggest the significance of the latter approach in representing, not only word order, but also the interface of syntax and information structure.

2 The data

The particle *wa* shows that the constituent marked with it carries a ‘contrastive focus’ interpretation.² In (3) the object *hon* ‘book’ is marked with *wa* and it has a contrastive focus reading. Thus, the sentence implies that Taro read books, but he read nothing else, e.g., magazines.³

- (3) Taroo-wa {hon-wa} yonda
Taro-TOP book-CF read ‘Taro read books (but read nothing else).’

As (4) shows, the domain of contrastive focus can be extended beyond the constituent marked with *wa* (Noda 1996; see Choi 1999 for analogous data in Korean).

- (4) Taroo-wa {hon-wa yonda}
Taro-TOP book-CF read ‘Taro read books (but did nothing else).’

In (4) a constituent *hon-wa yonda* ‘read books’, not just *hon-wa*, carries a contrastive focus interpretation, and the sentence implies that Taro read books, but he did not do anything else, such as playing tennis. One might argue that facts like (4) could be accounted for in terms of so-called ‘focus projection’, in which focus can be extended beyond the element that carries pitch accent in such languages as English and German (cf.

¹ Within the standard HPSG, there are at least two alternative constituent structures that might be advanced for a sentence composed of subject, one or more complements and verb. In one the complements and the verb form a constituent, and the latter forms another constituent with the subject. The other alternative is the one in which the subject, complements and verb are all sisters with the single mother (cf. Alexopoulou and Kolliakou 2002; Borsley 1989, 1995; Pollard 1994). One might ask a question: which structure do Japanese sentences have? In this paper we will avoid this question by taking into consideration only the cases where the subject is topicalized, and tentatively represent the mother of the verb and its sisters as S.

² As illustrated in (1), the same particle functions as a topic marker as well, but this aspect of *-wa* is irrelevant to the main subject. In the rest of this paper, the topic marker is glossed as TOP and the contrastive focus marker as CF.

³ Here and throughout, a domain of contrastive focus is marked with braces.

Jackendoff 1972; Selkirk 1995, and many others). Recent HPSG analyses on focus projection are built on the view that, at each level of syntactic combination, a focus-background structure for the mother is computed from the focus-background structure provided by the daughter constituents (Engdahl and Vallduví 1996; Alexopoulou and Kolliakou 2002; De Kuthy 2002; De Kuthy and Meurers 2003). In De Kuthy and Meurers's (2003) system, for example, if a daughter with the focus projection potential (FPP *plus*) is focused, the mother can be in the focus.⁴ Thus, the broad contrastive focus illustrated by (4) might be accounted for in an analogous way: the constituent is given contrastive focus since one of its non-head daughters (*hon* 'book') is marked with *wa*.

However, the position of a *wa*-marked nominal and its interaction with possible domains of contrastive focus pose a challenging problem for an argument along these lines.

- (5) a. Taroo-*wa* [_S tukue-*ni* hana-*wa* oita]
 Taro-TOP desk-LOC book-CF put 'Taro put the flower on the desk.'
 b. Taroo-*wa* [_S hana-*wa* tukue-*ni* oita]
 Taro-TOP flower-CF desk-LOC put 'Taro put the flower on the desk.'

The sentences in (5) each contain a *wa*-marked nominal, but it is immediately preceding the verb in (5a), but it is in the initial position of the S in (5b). If we adopt the analyses of focus projection surveyed above along with the standard HPSG tree-configuration such as (2a), these sentences should have the same possible focus domains: the *wa*-marked nominal and its mother S. The fact is, however, (5a) does not have an S focus interpretation while (5b) does. The focus interpretation of the latter is illustrated by (6).

- (6) Taroo-*wa* [_S hana-*wa* tukue-*ni* oita] -*ga* [_S hon-*wa* tana-*ni* narabeta]
 Taro-TOP flower-CF desk-LOC put -but book-CF shelf-LOC set
 'Taro put the flower on the desk, but he set the books on the shelf'

In (6) the S in (5b) is compared with another S, and hence it is given contrastive focus. The unacceptability of (7) gives evidence that (5a) does not have a S focus interpretation.

- (7) # Taroo-*wa* [_S tukue-*ni* hana-*wa* oita] -*ga* [_S tana-*ni* hon-*wa* narabeta]
 Taro-TOP desk-LOC flower-CF put -but shelf-LOC book-CF set
 'Taro put the flower on the desk, but he set the books on the shelf'

In (7) the S in (5a) is compared with another S and is intended to have contrastive focus. As the infelicity of (7) (marked by #) suggests, the *wa*-marked NP in the middle of the S cannot extend contrastive focus to the whole S. The following fact suggests that what carries a contrastive focus in (5a) is just part of the S.

- (8) Taroo-*wa* [_S tukue-*ni* { hana-*wa* oita }] -*ga* [_S sokoni { hon-*wa* narabe-nak-atta }]
 Taro-TOP desk-LOC flower-CF put -but there book-CF set-NEG-PAST
 'Taro put the flower on the desk, but he didn't set the books there'

It should be noted that *tukue-*ni** 'desk-LOC' and *sokoni* 'there' is coreferential so there is no sense in contrasting them. Therefore, what are really contrasted in (8) are the remaining parts of the Ss, namely *hana-*wa* oita* 'put the flowers' and *hon-*wa* narabe-nak-atta* 'didn't set the books'. Thus, the infelicity of (7) is due to the fact that only a part of the S carries contrastive focus in the context where the whole S focus is intended. (9) is another example to show that part of a constituent can carry contrastive focus.

- (9) Taroo-*wa* [_S { hana-*wa* tukue-*ni* } oita] -*ga* [_S { hon-*wa* yuka-*ni* } oita]
 Taro-TOP flower-CF desk-LOC put -but book-CF floor-LOC put
 'Taro put the flower on the desk, but the books on the floor'

In (9) contrastive focus appears to be given to the Ss, but note that they both include the same verb *oita* 'put'. There is no sense in including this common element in the domain of contrastive focus. Therefore, what are really contrasted in (8) are the remaining part of the Ss marked by the braces.

The available interpretations for sentences (5a) and (5b), therefore, can be summarised in (10a) and

⁴ The FPP feature is assumed for *synsem* objects so that verbs can lexically mark which of their arguments can project focus. See De Kuthy and Meurers (2003) for details.

(10b), respectively.

- (10) a. Taroo-wa tukue-ni {{hana-wa}_{F1} oita }_{F2}
 b. Taroo-wa {{{hana-wa}_{F1} tukue-ni}_{F2} oita }_{F3}

From these observations one general point becomes clear: in the case of broad focus, the *wa*-marked nominal indicates the left periphery of the focus domain. Note also that the focus domain marked by F2 in each sentence in (10) does not correspond to a constituent; it is just part of an S. These facts suggest the necessity of taking word order, rather than constituent structure, into account for establishing a proper treatment of focus projection facts in Japanese. This suggests the relevance of the Linearization HPSG approach, in which linear order is independent to a considerable extent from constituency.

3 A Linearization HPSG analysis

In this framework, linear order is represented in a separate level of ‘order domain’, to which ordering constraints apply (cf. Pollard *et al.* 1993; Reape 1994; Kathol 2000, etc.). It is given as the value of the attribute DOM(AIN). We assume, along with Reape (1994), Donohue and Sag (1999), and Kathol (2000), that an order domain consists of an ordered list of signs.⁵ At each level of syntactic combination, the order domain of the mother category is computed from the information provided by the daughter constituents.

Before turning to our main proposal, we will give an HPSG-style formalization to the constraints we have already stated.

$$(11) \quad \left[\begin{array}{c} \text{word} \\ \text{NP}[\text{wa}] \end{array} \right] \rightarrow [1][\text{INFO-STR} \mid \text{FOCUS} \langle [1] \rangle]$$

It states that if a word is marked with *-wa*, its FOCUS value is the sign itself. The feature FOCUS is appropriate for the INFO-STR (information structure), which in turn is appropriate for signs. Following Engdahl and Vallduví (1996), Alexopoulou (1999) and Alexopoulou and Kolliakou (1999), the FOCUS feature has a list of signs as its value.⁶

Now we propose the following constraint specifying the possible domains of contrastive focus in a phrase headed by a verb.

$$(12) \quad \left[\begin{array}{c} \text{S} \\ \text{H-DTR} \left[\text{DOM} \left\langle \dots [1] \left[\begin{array}{c} \langle \alpha \rangle \\ \text{FOCUS} \langle [1] \rangle \end{array} \right] \dots \left[\begin{array}{c} \langle \beta \rangle \\ \text{XP} \end{array} \right] \dots \right\rangle \right] \end{array} \right] \\ \rightarrow [\text{INFO-STR} \mid \text{FOCUS} \langle [1] \rangle] \vee \left[\begin{array}{c} \text{INFO-STR} \mid \text{FOCUS} \langle [2] \rangle \\ \text{DOM} \left\langle \dots [2] \left[\begin{array}{c} \langle \alpha \oplus \dots \oplus \beta \rangle \\ \text{FOCUS} \langle [2] \rangle \end{array} \right] \dots \right\rangle \end{array} \right]$$

The first disjunct in the consequent of (12) covers the case of narrow focus where the FOCUS value of a phrase inherits that of the *wa*-marked domain element. The second covers the case where contrastive focus projects further, stating that the FOCUS value of the phrase is structure-shared with the FOCUS value of the domain element constructed via partial compaction of the *wa*-marked element with the elements to its right. This partial compaction operation involving *-wa* is defined as follows.

$$(13) \quad [\text{DOM} \langle [\delta_1, \dots, [\delta_i[\text{wa}]], \dots, [\delta_n], \dots] \rangle] \Rightarrow [\text{DOM} \langle [\delta_1, \dots, [\delta_{i-1}], [\delta_i[\text{wa}] \circ \dots \circ \delta_n], [\delta_{n+1}], \dots] \rangle]$$

(13) specifies the case where the domain element marked with *wa* is combined with the other domain elements to its right, and they are turned into a single domain element; remaining elements are left as separate domain elements.⁷

Combining these constraints, we can obtain an analysis that predicts the interaction of the position of a

⁵ This assumption might involve some problems. See Kathol (2001) for details.

⁶ Thus, the values of the FOCUS and DOM feature are list of signs in our system.

⁷ For further details of compaction, see Kathol and Pollard (1995) and Yatabe (1996; 2001)

structure-shared with the FOCUS value of the domain element constructed by this partial compaction operation. The resulting domain object in (17b) corresponds to the lower S as a whole but those in (17a) correspond only to part of the S, namely *hana-wa tukue-ni* ‘... the flower on the desk’.

Recall that (14a), as opposed to (14b), cannot carry an S focus interpretation, as illustrated by (7), which is repeated here for convenience.

- (18) # Taro-wa [_S tukue-ni hana-wa oita]-ga [_S tana-ni hon-wa narabeta]
 Taro-TOP desk-LOC flower-CF put -but shelf-LOC book-CF set
 ‘Taro put the flower on the desk, but he set the books on the shelf’

This sentence has a *wa*-marked element in the middle of the S, and it is intended to have an S contrastive focus. The constraints proposed above may alone be enough to provide a good account of the infelicity of this sentence. In our system, a broad contrastive focus interpretation can be given to a domain element only if the *wa*-marked element is partially compacted with domain elements to its right, as the second disjunct of (12) states. However, in order to get the S focus interpretation in (18), the *wa*-marked element should be combined with the elements to its left, *tukue-ni* ‘desk-LOC’, as well. Our system does not ban this operation, but it does not conform to (12). In (18), therefore, contrastive focus cannot project to the domain corresponding to the whole S.

4 Conclusion

In this paper we have investigated the interaction between the linear position of a *wa*-marked nominal in a sentence and possible domains of contrastive focus, and showed that contrastive focus projection and a relatively flexible word order of Japanese should be analysed independently of constituency. This suggests the significance of Linearization HPSG in representing not only word order but also the interface of syntax and information structure.

References

- Alexopoulou, T. and D. Kolliakou. 2002. ‘On linkhood, topicalization and clitic left dislocation’. *JL* 38, 193–245.
- Borsley, R. D. 1989. ‘An HPSG approach to Welsh’. *JL* 25, 333–354.
- Borsley, R. D. 1995. ‘On some similarities and differences between Welsh and Syrian Arabic’. *Linguistics* 33, 99–122.
- Calcagno, M. 1993. ‘Toward a linearization-based approach to word order variation to Japanese’. In A. Kathol and C. Pollard (eds.), *Papers in Linguistics*, 26–45.
- Choi, H-W. 1999. *Optimizing Structure in Context*. Stanford: CSLI Publications.
- De Kuthy, K. 2002. *Discontinuous NPs in German – A Case Study of the Interaction of Syntax, Semantics and Pragmatics*. Stanford, CA: CSLI Publications.
- De Kuthy, K. and W. D. Meurers. 2003. ‘The secret life of focus exponents, and what it tells us about fronted verbal projections’. In S. Müller (ed.), *Proceedings of the HPSG03 Conference*.
- Donohue, C. and I. A. Sag. 1999. ‘Domains in Warlpiri’. Ms. Stanford University.
- Engdahl, E. and E. Vallduví. 1996. ‘Information packaging in HPSG’. C. Grover and E. Vallduví (eds.), *Edinburgh Working Papers in Cognitive Science, vol 12: Studies in HPSG*, 1–31.
- Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- Kathol, A. 2000. *Linear Syntax*. Oxford: Oxford University Press.
- Kathol, A. and C. Pollard. 1995. ‘Extrapolation via complex domain formation’. In *33rd Annual Meeting of the ACL*, 174–180. San Francisco: Morgan Kaufmann.
- Noda, H. 1996. ‘*Wa*’ to ‘*Ga*’ [‘*Wa*’ and ‘*Ga*’]. Tokyo: Kurosio.
- Pollard, C. 1994. ‘Toward a unified account of passive in German’. In J. Nerbonne, *et al.* (eds.), 273–96.
- Pollard, C. and I. A. Sag. 1989. *Information-Based Syntax and Semantics*. Stanford: CSLI Publications.
- Pollard, C. and I. A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Reape, M. 1994. ‘Domain union and word order variation in German’. In J. Nerbonne, *et al.* (eds.), *German in Head-Driven Phrase Structure Grammar*, 151–98. Stanford: CSLI Publications.
- Selkirk, E. 1995. ‘Sentence prosody: Intonation, stress, and phrasing’. In J. A. Goldsmith (ed.), *The Handbook of Phonological Theory*, 550–569. London: Basil Blackwell.
- Yatabe, S. 1996. ‘Long-distance scrambling via partial compaction’. In M. Koizumi and M. Oishi and U. Sauerland (eds.), *Formal Approaches to Japanese Linguistics 2* (MIT Working Papers in Linguistics 29). Cambridge, MA, 303–317.
- Yatabe, S. 2001. ‘The syntax and semantics of left-node raising in Japanese’. In D. Flickinger and A. Kathol (ed.), *Proceedings of the 7th International HPSG Conference*, 325–344.