Constituency, word order and focus projection

Takafumi Maekawa

tmaeka@essex.ac.uk

Department of Language and Linguistics

University of Essex

Abstract

In this study we show that constituency is of limited importance for a proper treatment of the interaction between the linear position of a *wa*-marked nominal in a Japanese sentence and possible domains of contrastive focus, and that constraints concerning contrastive focus should be represented in terms of linear order and not constituency. Linearisation HPSG, where linear order is independent from constituency, provides a good basis for an analysis. Some constraints are provided in terms of order domains, and it is shown that these constraints can deal with the phenomena in question, and that the cases problematic for the constituency-based analyses can also be accounted for by our analysis.
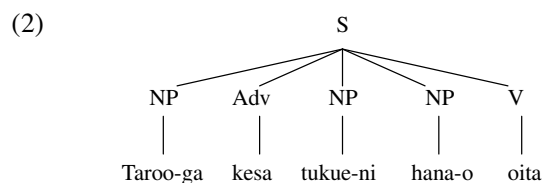
1   Introduction

The most widely adopted view of word order within the framework of HPSG is that a set of linear precedence (LP) rules state the possible permutations of constituent in a local tree (see, e.g., Pollard and Sag 1987, 1994).[*] In such a framework, there are at least two possible analyses of a relatively flexible word order of Japanese illustrated by (1), which has been often called 'scrambling'.

(1) a. Taroo-ga   kesa   tukue-ni   hana-o   oita
         Taro-NOM   this morning   desk-LOC   flower-ACC put
         'Taro put the flower on the desk this morning'

     b. Kesa   tukue-ni   Taroo-ga   hana-o   oita
         this morning   desk-LOC   Taro-NOM   flower-ACC put

First, we might assume a flat structure like (2) in which a lexical verbal head and all of its dependents form a single constituent.[1]
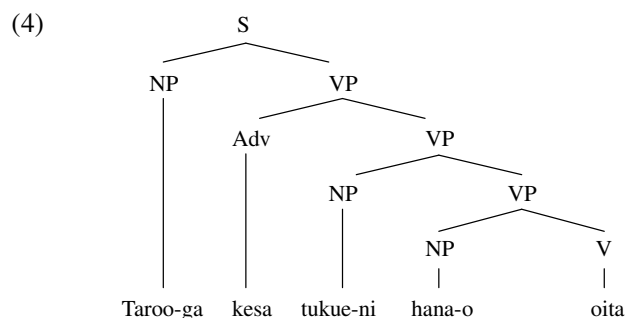
---

[1] There have been alternative ways proposed to build a flat structure in HPSG. See Alexopoulou and Kolliakou (2002), Borsley (1989, 1995) and Pollard (1994) for examples.

(2)

```
                           S
          ┌──────┬─────┼──────┬──────┐
         NP     Adv    NP     NP     V
          │      │      │      │      │
      Taroo-ga  kesa  tukue-ni hana-o oita
```

Then, the general LP rule (3) can alone give an account for the word order variation in (1).

(3)   [ ]   ≤   HEAD[LEX +]

This LP rule states that a lexical head follows any sister. Conforming to this rule, the lexical head daughter *oita* 'put' follows all its sisters both in (1a) and (1b). No further rules are needed to predict the word order variation as in (1). Second, we might assume a binary branching structure like the following.

(4)

```
                  S
          ┌──────────┴────┐
         NP              VP
          │        ┌──────┴────┐
          │       Adv          VP
          │        │      ┌─────┴──────┐
          │        │     NP            VP
          │        │      │        ┌────┴────┐
          │        │      │       NP         V
          │        │      │        │         │
      Taroo-ga   kesa  tukue-ni  hana-o     oita
```

A binary branching constituent structure has been advanced for German clauses in HPSG (see, e.g., Kathol 2000 and Müller 2002), so it might be quite reasonable to assume it for Japanese clauses as well. There are several possible ways to represent the relatively free constituent order in binary branching: to assume a set, rather than a list, of valence information (Gunji 1986; Hinrichs and Nakazawa 1989; Pollard 1996; also see Müller 2004b); to relax the requirement that elements should be removed from the list in order of their obliqueness (as suggested, but rejected, by Müller 2004a); to assume a lexical rule that licenses alternative orders for SUBCAT lists (Uszkoreit 1986).[2] Finally, we might assume 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; Kathol 2000). In this approach, the order variation of a sentence would be represented in the DOM(AIN) list, no matter which constituent structure it has.

---

[2] See Müller (2004a) for the details and problems of these approaches.

(5)
$$\begin{bmatrix} \text{DTRS} & ... \\ \text{DOM} & \langle \textit{Taroo - ga, kesa, tukue - ni, hana - o, oita} \rangle \end{bmatrix}$$

In this study these three approaches will be compared, and we shall argue that the connection between the linear position of a *wa*-marked nominal in a sentence and possible domains of contrastive focus provides support for the third, order domain approach. Some constraints will be provided in terms of order domains, and we will show that they can deal with the phenomena in question, and that the cases problematic for the constituency-based analyses can be accounted for by our analysis.

The organisation of this study is as follows. In section 2 we shall survey the basic data, and see that the phenomena in question seem to be similar to those accounted for in terms of focus projection. In section 3 we shall compare the flat structure analysis and binary branching structure analysis introduced above, and argue that neither of them is satisfactory. Section 4 will present an alternative analysis in terms of order domains. Section 5 is the conclusion.

2   Particle *wa* and an extension of contrastive focus

The particle *wa* indicates that some material containing it carries a 'contrastive focus' interpretation.[3] In (6) the object *sakana* 'fish' is marked with *wa* and it has a contrastive focus reading. Here and throughout, a domain of contrastive focus is marked with braces. Thus, the sentence implies that Taro ate fish but he did not eat anything else.

(6)   Taroo-wa   {sakana-*wa*}   tabeta
         Taro-TOP   fish-CF   ate
         'Taro ate fish (but ate nothing else).'

As (7) shows, the domain of contrastive focus can be extended beyond the element marked with *wa* (Noda 1996; see Choi 1999 for analogous data in Korean).

(7)   Taroo-wa   {sakana-*wa*   tabeta}
         Taro-TOP   fish-CF   ate
         'Taro ate fish (but did nothing else).'

In (7) *sakaana-wa tabeta* 'ate fish', and not just *sakaana-wa*, carries a contrastive

---

[3] As illustrated by the examples below, 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 grossed as TOP and the contrastive focus marker as CF.

focus interpretation, and the sentence implies that Taro ate fish but did not do anything else, such as playing tennis. If there is some other element between the *wa*-marked nominal and the verb, it can be in the focus domain.

(8) a. Taroo-wa   {hana-*wa*   tukue-ni   oita}
       Taro-TOP   flower-CF   desk-LOC   put
       'Taro put the flowers on the desk (but did nothing else).'
    b. Taroo-wa   {sakana-*wa*   resutoran-de   tabeta}
       Taro-TOP   fish-CF   restaurant-LOC ate
       'Taro ate fish at the restaurant (but did nothing else).'

In (8) there is another complement *tukue-ni* 'on the desk' (a) and an adjunct *resutoran-de* 'at the restaurant' (b) between the *wa*-marked complement and the verb, and they can be included in the focus domain, as the translation shows.

The most plausible way to analyse the extension of contrastive focus would be to take it as an instance of 'focus projection'. In languages such as English and German, focus can be extended beyond the element that carries pitch accent. Focus projection, which has been a dominant approach to this phenomenon, argues that in each local tree a focus-background structure for the mother is computed from the focus-background structure of the daughter constituents (see, e.g., Jackendoff 1972 and Selkirk 1995). Recent HPSG analyses are along the same lines (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] The broad contrastive focus as in (7) and (8) might seem to be accounted for in an analogous way by assuming either binary branching or a flat structure: the constituent is given contrastive focus if one of its non-head daughters (e.g., *sakana* 'fish' in (7)) is marked with contrastive *wa*.[5]

In the next section, we shall look at how the binary branching and the flat structure analysis deal with the phenomena in question, and point out that neither of them is satisfactory.

3   Constituency-based analyses

This section will compare the possible constituency-based approaches to order variation, i.e., a flat structure analysis and binary branching structure analysis.

---

[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.

[5] It is assumed here that the topic is combined with the rest of the sentence as a filler, along the same lines as the English topicalisation (Pollard and Sag 1994).

3.1 Flat structure

The position of a *wa*-marked nominal and its interaction with possible domains of contrastive focus pose a problem for the flat structure analysis.

(9) a. [$_S$ Taroo-ga   tukue-ni   hana-*wa*   oita] (-to kiita)
              Taro-NOM desk-LOC flower-CF put   that heard
             '(I hear that) Taro put the flower on the desk.'

      b. [$_S$ hana-*wa*   Taroo-ga   tukue-ni   oita] (-to kiita)
              flower-CF Taro-NOM desk-LOC put   that heard
             '(I hear that) Taro put the flower on the desk.'

The sentences in (9) each contain a *wa*-marked nominal. It is immediately preceding the verb in (9a), but is in the initial position of the S in (9b). If we adopt the approach to focus projection discussed above along with a flat structure, these sentences should have the same possible focus domains: the *wa*-marked nominal and its mother S. The fact is, however, that (9a) does not have an S focus interpretation while (9b) does. The interpretation of the latter is illustrated by (10).

(10)   [$_{S1}$ hana-*wa*   Taroo-ga   tukue-ni   oita] -ga
            flower-CF Taro-NOM desk-LOC put -but
           [$_{S2}$ hon-*wa*   Jiroo-ga   tana-ni   narabeta] (-to kiita)
            book-CF   Jiro-NOM   shelf-LOC set   that heard
           '(I hear that) Taro put the flower on the desk but Jiro set the books on the shelf.'

In (10) the S in (9b) (marked as S1) is connected with another clause (S2) with a disjunctive conjunction *-ga* 'but' so that S1 is in the contrastive relation with S2; both clauses have a *wa*-marked element in its initial position, and they have all different elements. Thus, it would be possible to say that the whole of S1 (and S2) carries contrastive focus. Now let us look at (11), where (9a) is contrasted with another sentence.

(11) # [$_{S1}$ Taroo-ga   tukue-ni   hana-*wa*   oita] -ga
             Taro-NOM desk-LOC flower-CF put -but
           [$_{S2}$ Jiroo-ga   tana-ni   hon-*wa*   narabeta] (-to kiita)
            Jiro-NOM   shelf-LOC book-CF set   that heard
   '(I hear that) Taro put the flower on the desk but Jiro set the books on the shelf.'

In (11) the S in (9a) is contrasted with another S, and is intended to have a

sentential contrastive focus. The infelicity of (11) (marked by #) suggests that (9a) does not have a S focus interpretation, and gives evidence that the *wa*-marked NP in the middle of the S cannot extend contrastive focus to the whole S.

A construction such as the following provides another problem for the flat structure analysis.[6]

(12)   Taroo-ga   akai   {uwagi-wa   kiru} -ga
       Taroo-NOM red   jacket-CF   wear but
       onazi iro-no   {zubon-wa   kiraida}   (-to   kiita).
       same colour-GEN   trousers-CF hate   that heard
       'Taro wears a red jacket, but hates trousers of the same colour.'

The structure for the first clause of (12) is (13).

(13)
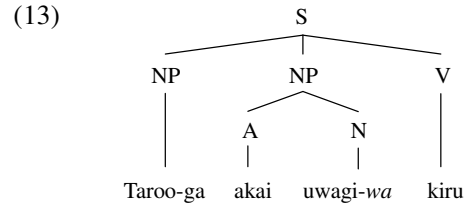
In (12) it is impossible to regard the adjective as a part of the focus domain; the modifiers of the noun in these clauses, *akai* 'red' and *onazi iro-no* 'of the same colour', both refer to a red colour, and there is no sense in contrasting the same colour. Hence they should be excluded from the focus domain: only the V and a part of the NP carry contrastive focus. If we assumed (13), however, the possible focus domains that could be represented would be N, NP and S only, and we could not analyse a case like (12).

3.2   Binary branching structure

In the last subsection, we saw that we could not deal with the extension of focus in (9a). The following fact suggests that what carries a contrastive focus in (9a) is just part of the S.

(14)   [S1 Taroo-ga   tukue-ni   hana-*wa*   oita] -ga
          Taro-NOM   desk-LOC   flower-CF put -but
       [S2 hon-*wa*   narabe-nak-atta] (-to   kiita)
          book-CF   set-NEG-PAST   that heard
       '(I hear that) Taro put the flower on the desk, but he didn't set the

books there.'

As is well known, it is allowed in Japanese to leave unexpressed the element which refers to who or what has been already introduced in discourse. (14) is such a sentence. As indicated by the translation, it is intend that the second S has a subject, a time and place adverbial coreferential to those of the first S. Since there is no sense in contrasting the coreferential things, therefore, it is possible to say that what are really contrasted in (14) are the remaining parts of the Ss, namely *hana-wa oita* 'put the flowers' in S1 and *hon-wa narabe-nak-atta* 'didn't set the books' in S2. Thus, the contrastive focus domain of (9a) is the part marked with braces in (15).

(15)   [$_S$ Taroo-ga   tukue-ni   { hana-*wa*   oita }] (-to kiita)
             Taro-NOM desk-LOC   flower-CF put   that heard

The infelicity of (11), where (9a) is contrasted with another S, 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. Therefore, we need an alternative analysis which can capture the fact that the *wa*-marked element in (9a) does not extend its focus to the whole S and its contrastive focus domain is just part of the S, as marked in (15).

One might argue that a binary branching approach would be such an alternative. We could assume the following binary branching structure for (9a).

(16)
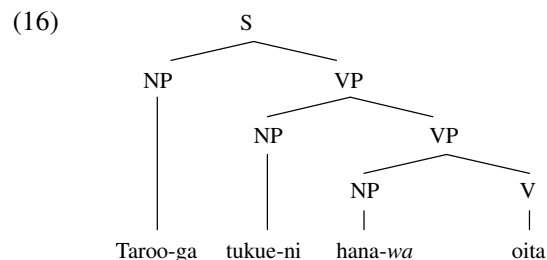
The contrastive focus domain of (9a), shown in (15), corresponds to the lower VP in (16). The broad contrastive focus can be dealt with on the basis of the traditional conception of focus projection introduced in the last section: the lower VP in (16) is given contrastive focus since its non-head daughter (i.e., *hana* 'flower') is marked with *wa*.

However, an example like the following poses a problem for this analysis.

(17)   [$_{S1}$ Taroo-ga   hana-*wa*   tukue-ni   oita] -ga
             Taro-NOM flower-CF desk-LOC put   -but
            [$_{S2}$ hon-*wa*   yuka-ni   oita] (to kiita)
             book-CF   floor-LOC put   that heard

'(I hear that) Taro put the flower on the desk, but the books on the floor'

The binary branching structure for the S1 of (17) would be something like (18).

(18)

```
                    S
            ┌───────┴───────┐
           NP              VP
            │         ┌──────┴──────┐
            │        NP            VP
            │         │        ┌─────┴─────┐
            │         │       NP           V
            │         │        │           │
        Taroo-ga  hana-wa  tukue-ni      oita
```
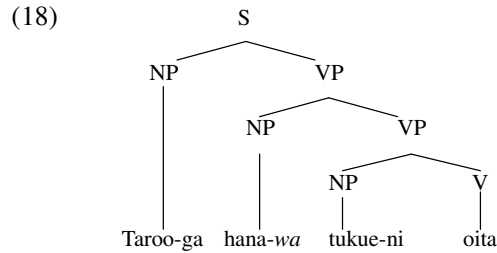
As the translation indicates, sentence (17) is intended to carry an interpretation where Taro's putting the flower on the desk and his putting the books on the floor is compared. Therefore, the contrastive focus domain might seem to correspond to the upper VP in (18). Note, however, that both S1 and S2 include the same verb *oita* 'put'. There is no sense in contrasting the same element, so the verb should be excluded from the domain of contrastive focus. Therefore, what are really contrasted in (17) are the remaining parts of the Ss, which are marked by the braces in (19).
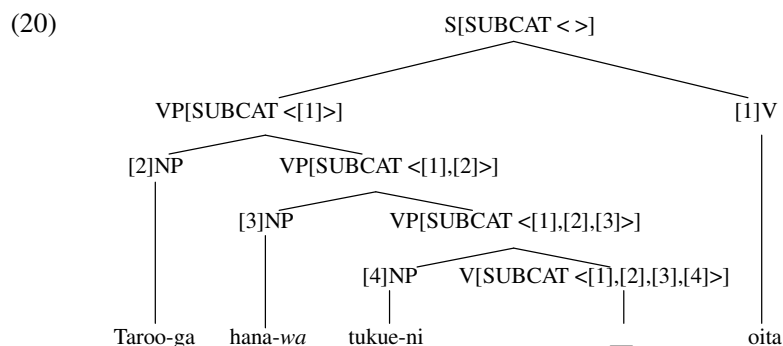
(19) [$_S$ Taroo-ga { hana-*wa* tukue-ni } oita] -ga
           Taro-NOM   flower-CF desk-LOC put   -but
      [$_S$ { hon-*wa* yuka-ni } oita] (-to kiita)
             book-CF floor-LOC put   that heard

Such a domain cannot be represented with a binary branching structure in (18).

One might argue that these cases would be accounted for with binary branching if we assumed an analogue of verb movement.[7,8] A verb movement analysis of (19) would have the representation in (20), where it is assumed that a SUBCAT list of the verbal trace contains the verb and its all arguments (Müller 2004b; Netter 1992, 1998).

---

[7] Some HPSG researches have assumed such a mechanism in order to describe the finite verb position in German (Frank 1994; Jacobs 1986; Kiss and Wesche 1991; Meurers 2000; Müller 2004b; Müller and Kasper 2000; Netter 1992, 1998).
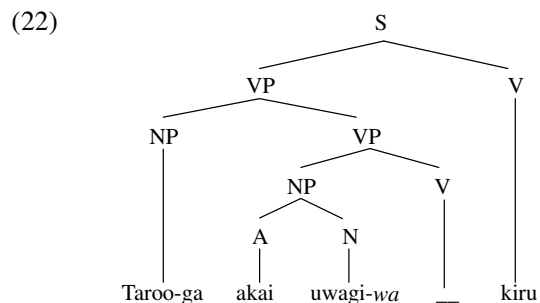[8] This possibility was pointed out to me by Bob Borsley.

(20)

```
                                    S[SUBCAT < >]
                        _____|_____
                       |                                       |
              VP[SUBCAT <[1]>]                               [1]V
           _____|_____                                  |
          |                 |                                  |
        [2]NP        VP[SUBCAT <[1],[2]>]                      |
          |          _____|_____                        |
          |         |                 |                       |
          |       [3]NP      VP[SUBCAT <[1],[2],[3]>]         |
          |         |        _____|_____                |
          |         |       |                 |               |
          |         |     [4]NP    V[SUBCAT <[1],[2],[3],[4]>]|
          |         |       |                 |               |
          |         |       |                 |               |
       Taroo-ga  hana-wa  tukue-ni            __             oita
```

This structure would allow contrastive focus to extend from *hana-wa* to the VP dominating *hana-wa* and *tukue-ni*, which corresponds to the domain marked with the braces in (19). Example (12), however, is problematic for this approach as well. It is repeated in (21)

(21) Taroo-ga   akai   {uwagi-wa   kiru} -ga
      Taroo-NOM red   jacket-CF   wear but
      onazi iro-no   {zubon-wa   kiraida}   (to   kiita).
      same colour-GEN   trousers-CF hate   that heard
      '(I hear that) Taro wears a red jacket, but hates trousers of the same
      colour.'

A binary branching structure for this would be (22).

(22)

```
                              S
                     _____|_____
                    |                   |
                    VP                  V
              _____|_____            |
             |             |           |
            NP             VP           |
             |        _____|_____     |
             |       |             |    |
             |      NP             V    |
             |    ___|___          |    |
             |   |       |         |    |
             |   A       N         |    |
             |   |       |         |    |
          Taroo-ga akai uwagi-wa   __  kiru
```

Recall that the adjective should be excluded from the focus domain of (21). If we assumed constituent structure (22), it would still be impossible to represent the contrastive focus domain marked in (21) since the possible focus domains would be N and NP, *uwagi-wa* and *akai uwagi-wa*.

3.3   Summary

In this section we have seen that none of the constituent-based approaches can handle the interaction of word order variation and contrastive focus projection.

This fact suggests that we need an alternative analysis.

Summarising the observations in the last two sections, we get the following pattern of extension of contrastive focus from the *wa*-marked nominal.

(23) a. $\{\{\{\{$hana-*wa*$\}_{F1}$ Taroo-ga$\}_{F2}$ tukue-ni$\}_{F3}$ oita$\}_{F4}$

    b. Taroo-ga $\{\{\{$hana-*wa*$\}_{F1}$ tukue-ni$\}_{F2}$ oita$\}_{F3}$

    c. Taroo-ga   tukue-ni $\{\{$hana-*wa*$\}_{F1}$ oita$\}_{F2}$

Each sentence in (23) has a *wa*-marked nominal in different linear position: the initial position (a), the second position (b), and the third position just before the verb (c). From this, three general points become clear. Firstly, possible domains of contrastive focus change according to the linear position of the *wa*-marked nominal; (23a) has four possible domains, (23b) has three and (23c) has only two. Second, the domain of contrastive focus extends only rightwards on the basis of linearity. Thirdly, there exist domains that do not correspond to any constituent, whether we assume a binary branching or a flat structure; note the domains marked by F2 and F3 in (23a) and F2 in (23b).

These points suggest that constituency is of limited importance for a proper treatment of extension of contrastive focus; it should be more closely related to linear order. Thus, we need an approach where linear order would be independent from constituency, and constraints concerning contrastive focus could be represented on the same level as the former, not the latter. As we will see in the next section, Linearization HPSG provides a good basis for such an approach, so we shall give an alternative analysis within the framework.

4 A Linearization HPSG analysis

In this section, we will provide an alternative analysis of contrastive focus projection. 4.1 sets the framework for our analysis. In 4.2 we shall provide our proposals. Finally, in 4.3 we shall argue that our proposals can deal with the cases which are problematic in the constituency-based approaches.

4.1 Framework

The analysis to be presented below will be largely based on a version of linearisation-based HPSG. In this framework, linear order is represented in a separate level of 'order domain', to which ordering constraints apply (see, e.g., Pollard et al. 1993; Reape 1994; and Kathol 2000). Order domains are given as the value of the attribute DOM(AIN). At each level of syntactic combination, the order domain of the mother category is computed from the order domains of the

daughter constituents. We assume, along with Reape (1994), Donohue and Sag (1999), Kathol (2000: 101), and Jaeger (2003), that an order domain consists of an ordered list of signs, which we will call 'DOM elements'.[9]

Next, we assume that a sign has information structure, which is represented as a value of its INFO(RMATION)-STR(UCTURE) (Engdahl and Vallduví 1996; Alexopoulou and Kolliakou 2002; De Kuthy 2002; De Kuthy and Meurers 2003). Its feature geometry would reflect a focus-background structure of a sign, and we assume that each of those features has a list of signs as its value (Engdahl and Vallduví 1996; Alexopoulou 1999; Alexopoulou and Kolliakou 1999). Among those appropriate for INFO-STR, however, the only feature mentioned in this study is CONTR(ASTIVE)-FOC(US). Its value is structure-shared with the sign which is the part of a sentence with contrastive focus.[10]

We stated above that a DOM element is a sign, and a sign has information structure. This means that a DOM element include information structure in its internal structure. This latter point has a considerable significance for our analysis in that it enables the information-structural status of DOM elements to be accessible for ordering constraints (see also Jaeger 2003).

4.2　　Proposals

Now we propose the following constraint.

$$(24) \quad \begin{bmatrix} S \\ DOM \left\langle ...,[1] \begin{bmatrix} sign \\ INFO\text{-}STR \mid CONTR\text{-}FOC \left\langle [1] \right\rangle \end{bmatrix}, ... \right\rangle \end{bmatrix} \\ \rightarrow \begin{bmatrix} INFO\text{-}STR \mid CONTR\text{-}FOC \left\langle [1] \right\rangle \end{bmatrix}$$

(24) states that if the CONTR-FOC value of a DOM element of an S is structure-shared with that DOM element itself, the CONTR-FOC value of the S is structure-shared with the CONTR-FOC value of that DOM element.

There are two cases that are covered by this constraint. One is the case where the DOM element in the antecedent of the constraint (24) is the one licensed by (25).

$$(25) \quad \begin{bmatrix} word \\ NP[wa] \end{bmatrix} \rightarrow [1] \begin{bmatrix} INFO\text{-}STR \mid CONTR\text{-}FOC \left\langle [1] \right\rangle \end{bmatrix}$$

Constraint (25) states that if a word is marked with *wa*, its CONTR-FOC value is

---

[9] The assumption that DOM elements are signs might involve some problems. See Kathol (2000) for discussion.

the sign itself. This is an HPSG-style formalization to the constraint which has already been stated: a *wa*-marked element carries contrastive focus. This corresponds to the narrowest domain of contrastive focus in which the CONTR-FOC value of an S inherits that of the *wa*-marked DOM element which is composed of only one lexical sign.

Another is the case where the DOM element in the antecedent of (24) satisfies the following constraint.

$$(26) \qquad \begin{bmatrix} sign \\ \delta_i[wa] \oplus ... \oplus \delta_n \end{bmatrix} \rightarrow [1]\left[\text{INFO-STR} \mid \text{CONTR-FOC} \langle [1] \rangle \right]$$

What (26) says is that if a *wa*-marked element combines with the other DOM element(s) to its right, then the resulting single DOM element should have its CONTR-FOC value structure-shared with that DOM element itself. If there are remaining elements, they are left as separate DOM elements. This is an instance of 'partial compaction' which has been discussed by Kathol and Pollard (1995) and Yatabe (1996, 2001). Note that compaction of a *wa*-marked element with the elements to its left is not prevented. However, the resulting DOM element simply does not satisfy the constraint (26), so its CONTR-FOC value would not be structure-shared with the DOM element. Only the DOM element with a *wa*-marked element on its left periphery can satisfy this constraint. Such a DOM element is in turn entitled to occur in the antecedent of constraint (24), in just the same way as a single *wa*-marked element, discussed above. As we shall see below, this is the case where contrastive focus extends beyond the *wa*-marked element.

4.3   Analysis

Combining the constraints introduced above, we can obtain an analysis which predicts the interaction of the linear position of a *wa*-marked nominal and the possible domains of contrastive focus, which was summarised in 3.3 and is repeated here for convenience.

(27) a. $\{\{\{\{\text{hana-}wa\}_{F1} \text{ Taroo-ga}\}_{F2} \quad \text{tukue-ni}\}_{F3} \text{ oita}\}_{F4}$
    b.   Taroo-ga $\{\{\{\text{hana-}wa\}_{F1} \quad \text{tukue-ni}\}_{F2} \quad \text{oita}\}_{F3}$
    c.   Taroo-ga   tukue-ni $\{\{\text{hana-}wa\}_{F1} \text{ oita}\}_{F2}$

4.3.1   Narrow domain of contrastive focus

Let us start with the cases where the domain of contrastive focus is the narrowest, confined to a *wa*-marked element. Let us look at (28), which represents the S node for the contrastive focus domain marked with F1 of (27a).

---

[10] Thus, the value of the CONTR-FOC and DOM feature is a list of signs in our system.

$$(28) \quad \begin{bmatrix} \text{INFO - STR | CONTR - FOC } \langle [1] \rangle \\[6pt] \text{DOM} \quad \left\langle [1]\begin{bmatrix} \langle hana\text{ - }wa \rangle \\ \text{CONTR - FOC } \langle [1] \rangle \end{bmatrix}, [\langle Taroo\text{ - }ga \rangle], [\langle tsukue\text{ - }ni \rangle] [\langle oita \rangle] \right\rangle \end{bmatrix}$$

Constraint (25) is responsible for the structure-sharing (tagged [1]) between the CONTR-FOC value of the *wa*-marked DOM element and the DOM element itself. Then the CONTR-FOC feature of S inherits the value [1], which is a case covered by (24). This is the process in which a single *wa*-marked element is licensed to carry contrastive focus in the S sign.

The narrowest domain in (27b,c), with a *wa*-marked nominal in a position different from the above, are analysed in an analogous way. (29a) and (29b) are the representations of the S node with F1 domain of (27b) and (27c), respectively.

$$(29) \text{ a.} \quad \begin{bmatrix} \text{INFO - STR | CONTR - FOC } \langle [2] \rangle \\[6pt] \text{DOM} \quad \left\langle [\langle Taroo\text{ - }ga \rangle][2]\begin{bmatrix} \langle hana\text{ - }wa \rangle \\ \text{CONTR - FOC } [2] \end{bmatrix}, [\langle tsukue\text{ - }ni \rangle] [\langle oita \rangle] \right\rangle \end{bmatrix}$$

$$\text{b.} \quad \begin{bmatrix} \text{INFO - STR | CONTR - FOC } \langle [3] \rangle \\[6pt] \text{DOM} \quad \left\langle [\langle Taroo\text{ - }ga \rangle] [\langle tsukue\text{ - }ni \rangle][3]\begin{bmatrix} \langle hana\text{ - }wa \rangle \\ \text{CONTR - FOC } [3] \end{bmatrix}, [\langle oita \rangle] \right\rangle \end{bmatrix}$$

They are different from the previous one in the position of the *wa*-marked element in the DOM list: in (29a) it is in the second position of the list while in (29b) it is in the third position. As in the previous case, however, the INFO-STR|CONTR-FOC value of the S is structure-shared with the value of the CONTR-FOC feature of the *wa*-marked element. Thus, in the narrowest domain cases, the S's CONTR-FOC value inherits that of a *wa*-marked element, in whatever position the latter is.

4.3.2   Broad domain of contrastive focus

Next let us turn to the cases where contrastive focus is extended beyond the *wa*-marked element; that is, the domains marked F2 to F4 in (27).

Let us start with the case which was provided as problematic to a flat structure approach, i.e., (9a). This corresponds to the contrastive focus domain marked as F2 in (27c), which is repeated here.

(30)   Taroo-ga   tukue-ni   { hana-*wa*   oita }$_{F2}$
            Taro-NOM   desk-LOC   flower-CF   put

In our analysis, the S sign of (30) has the following schematic analysis.

$$(31) \quad \begin{bmatrix} \text{INFO - STR | CONTR - FOC} \langle [4] \rangle \\ \text{DOM} \left\langle \left[\langle \textit{Taroo - ga} \rangle \right], \left[\langle \textit{tukue - ni} \rangle \right], [4] \begin{bmatrix} \langle \textit{hana - wa, oita} \rangle \\ \text{CONTR | FOC} \langle [4] \rangle \end{bmatrix} \right\rangle \end{bmatrix}$$

The *wa*-marked element, *hana-wa* 'flower-CF', is partially compacted with the element to its right, *oita* 'put', and they form a single DOM element. The remaining elements are left as separate DOM elements. Due to (26), then, a DOM element constructed by compaction has its CONTR-FOC value structure-shared with the DOM element itself. Finally, following (24), the CONTR-FOC value of the S is structure-shared with the CONTR-FOC value of this DOM element. In this way the F2 domain in (30) is licensed to carry contrastive focus, and this is how our analysis accounts for the case problematic to the flat structure approach.

We provided (17) as a counterexample to the binary branching approach. It corresponds to the contrastive focus domain marked as F2 in (27b) repeated here.

(32) Taroo-ga { hana-*wa*   tukue-ni }$_{F2}$ oita
      Taro-NOM   flower-CF   desk-LOC   put

Its S sign would be analysed in the following way.

$$(33) \quad \begin{bmatrix} \text{INFO - STR | CONTR - FOC} \langle [5] \rangle \\ \text{DOM} \left\langle \left[\langle \textit{Taroo - ga} \rangle \right], [5] \begin{bmatrix} \langle \textit{hana - wa, tukue - ni} \rangle \\ \text{CONTR - FOC} \langle [5] \rangle \end{bmatrix}, \left[\langle \textit{oita} \rangle \right] \right\rangle \end{bmatrix}$$

In (33), *hana-wa* 'flower-CF' is partially compacted with the element immediately to its right, *tukue-ni* 'desk-LOC', and they form a single DOM element, which is tagged [5]. The remaining elements, *Taroo-ga* 'Taro-NOM' and *oita* 'put', are left as separate DOM elements. Following (24), the CONTR-FOC value of the compacted DOM element is [5]; it is structure-shared with the DOM element itself. The CONTR-FOC value of the S is structure-shared with the CONTR-FOC value of this DOM element, and hence it is also tagged [5]. The example problematic for the binary branching approach can thus be given an analysis in our system.

As has been already clear from the above, what we have called a broad domain of contrastive focus corresponds to a DOM element constructed via compaction involving a *wa*-marked element. Constraint (26) entails that a DOM element can satisfy it as long as a *wa*-marked element has been combined with other elements to its right. This means that such a DOM element can extend rightwards until the end of the sentence. This in turn means that (27a) has the three possibilities shown in (34) for DOM elements eligible for constraint (26).

(27a) is repeated in (35). (34) shows only the PHON value of the DOM elements.

(34) a. *\<hana-wa, Taroo-ga\>*

b. *\<hana-wa, Taroo-ga, tukue-ni\>*

c. *\<hana-wa, Taroo-ga, tukue-ni, oita\>*

(35) {{{{hana-*wa*}$_{F1}$ Taroo-ga}$_{F2}$ tukue-ni}$_{F3}$ oita}$_{F4}$

As is clear from these, each possible compacted DOM element shown in (34) corresponds to the broad domains of contrastive focus marked F2 to F4 in (35). Thus, constraints provided in 4.2 can predict these contrastive focus domains, and all other broad domains shown in (27) would be predicted along these lines: F2 to F4 domains are made by combining a *wa*-marked element with other elements to its right one at a time.

The fact observed in 3.1 that (10) is acceptable whereas (11) is not is accounted for along these lines. Example (10) is repeated in (36a), and (11) in (36b).

(36) a. [$_{S1}$ hana-*wa*   Taroo-ga   tukue-ni   oita] -ga

    flower-CF   Taro-NOM   desk-LOC   put -but

[$_{S2}$ hon-*wa*   Jiroo-ga   tana-ni   narabeta] (-to kiita)

    book-CF   Jiro-NOM   shelf-LOC   set   that heard

'(I hear that) Taro put the flower on the desk but Jiro set the books on the shelf.'

b. # [$_{S1}$ Taroo-ga   tukue-ni   hana-*wa*   oita] -ga

    Taro-NOM   desk-LOC   flower-CF put -but

[$_{S2}$ Jiroo-ga   tana-ni   hon-*wa*   narabeta] (-to kiita)

    Jiro-NOM   shelf-LOC   book-CF set   that heard

'(I hear that) Taro put the flower on the desk but Jiro set the books on the shelf.'

In both, S1 is contrasted with S2 so that it is intended that whole of S1 (and S2) has contrastive focus. The contrastive focus domain in (36a) is licensed in the way discussed in the last paragraph. These Ss have a *wa*-marked element in their initial position. This ensures that *hana-wa* and *hon-wa* can combine with all other DOM elements (i.e., total compaction) to be eligible for constraint (26) since they are all to its right. Let us turn to (36b). In this sentence, the *wa*-marked element is not in the initial position. In our system, a broad contrastive focus interpretation can be given to a DOM element only if the *wa*-marked element is compacted with DOM elements to its right. However, in order to get the whole S focus interpretation, the *wa*-marked element in (36b) would have to be combined not

only with the element to its right but also with the elements to its left. Recall that compaction of a *wa*-marked element with the elements to its left is not prevented. However, the resulting DOM element simply does not satisfy constraint (26), so its CONTR-FOC value would not be structure-shared with the DOM element. The infelicity of (36b) is, thus, attributed to the fact that a sentential contrastive focus domain in (36b) would involve *Taroo-ga* and *tukue-ni* which are to its left.

For the same reason as (36b), such contrastive focus domains as shown below, where the domain is extended from the *wa*-marked element to its left, are predicted to be unacceptable.

(37) a. {Taroo-ga hana-*wa*} tukue-ni oita
     b. {Taroo-ga {tukue-ni hana-*wa*}} oita

This is borne out by the following data.

(38) a. # [$_{S1}$ { Taroo-ga hana-*wa* } tukue-ni oita] -ga
          Taro-NOM flower-CF desk-LOC put -but
         [$_{S2}$ { Jiroo-ga hon-*wa* } soko-ni oita] (-to kiita)
          Jiro-NOM book-CF there-LOC set that heard
         '(I hear that) Taro put the flower on the desk but Jiro put the books there.'

     b. # [$_{S1}$ Taroo-ga { tukue-ni hana-*wa* } oita] -ga
          Taro-NOM desk-LOC flower-CF put -but
         [$_{S2}$ { tana-ni hon-*wa* } oita] (-to kiita)
          shelf-LOC book-CF put that heard
         '(I hear that) Taro put the flower on the desk but put the books on the shelf.'

In (38a, b), (a) and (b) of (37) are respectively set in the context where the sequences marked with braces have contrastive focus: the verbs in both clauses are the same, and the locative arguments of S1 and S2 in (38a) and the subject of S1 and the null argument of S2 in (38b) have the same reference. As predicted, (38a, b) are both unacceptable.

We now move on to example (12), which we presented as a counterexample to both the binary and the flat structure analysis. The example, once repeated in (21), is again repeated here for convenience.

(39)   Taroo-ga akai {uwagi-wa kiru} -ga
        Taroo-NOM red jacket-CF wear but
        onazi iro-no {zubon-wa kiraida} (-to kiita).
        same colour-GEN trousers-CF hate that heard

'(I hear that) Taro wears a red jacket, but hates trousers of the same colour.'

The first clause of (39) would be analysed in the following way. The adjective *akai* 'red' would be combined with the noun *uwagi-wa* 'jacket-CF' at some point of combination in constituent structure, but in the order domain they can be separated. This is possible because of the assumption that an order domain is independent of constituency. Instead, *uwagi-wa* combines with the element to its right, *kiru* 'wear', by partial compaction. The S sign of this clause is as follows.

(40)
$$
\begin{bmatrix}
\text{INFO-STR} \mid \text{CONTR-FOC} \langle [6] \rangle \\
\text{DOM} \left\langle \big[\langle Taroo\text{-}ga \rangle\big]\big[\langle akai \rangle\big][6]\begin{bmatrix} \langle uwagi\text{-}wa, kiru \rangle \\ \text{CONTR-FOC} \langle [6] \rangle \end{bmatrix} \right\rangle
\end{bmatrix}
$$

Due to (26), the DOM element constructed via partial compaction has its CONTR-FOC value structure-shared with the DOM element itself. Then, following (24), the CONTR-FOC value of the S is structure-shared with the CONTR-FOC value of this DOM element. This analysis is supported by the fact that the noun and its modifier do not always have to be adjacent.

(41)   Akai, Taroo-ga   { uwagi-wa   kiru }   -ga
       red   Taroo-NOM   jacket-CF   wear   but
       onazi iro-no   { zubon-wa   kiraida} (-to   kiita).
       same colour-GEN   trousers-CF hate   that heard
       '(I hear that) Taro wears a red jacket, but hates trousers of the same colour.'

In (41), which in my judgement is acceptable at least in colloquial speech, *akai* 'red' is in front of the sentence, and is separated from *uwagi-wa* 'jacket-CF', intervened by *Taroo-ga* 'Taro-NOM'. The order domain of (41) would have the following structure.

(42)
$$
\begin{bmatrix}
\text{INFO-STR} \mid \text{CONTR-FOC} \langle [7] \rangle \\
\text{DOM} \left\langle \big[\langle akai \rangle\big]\big[\langle Taroo\text{-}ga \rangle\big][7]\begin{bmatrix} \langle uwagi\text{-}wa, kiru \rangle \\ \text{CONTR-FOC} \langle [7] \rangle \end{bmatrix} \right\rangle
\end{bmatrix}
$$

The order domains in (40) and (42) are just alternatives, the only difference being the position of *akai*; it is separated from *uwagi-wa* in (42), and is in the initial position of the DOM list.

5   Conclusion

We showed that constituency is much less significant than linear order for a proper treatment of the interaction between the linear position of a *wa*-marked nominal in a sentence and possible domains of contrastive focus, and that constraints concerning contrastive focus should be represented in terms of linear order and not constituency. We argued that Linearisation HPSG, where linear order is independent from constituency, provides a good basis for this. Finally, we gave some constraints in terms of order domains that can deal with the phenomena in question, and showed that cases problematic for the constituency-based analyses can also be accounted for by our analysis. If our analysis is on the right track, it suggests that Linearization HPSG is important not only for representing word order but also for the analysis of information structure and its interaction with syntax as well.

References

Alexopoulou, Theodra and Kolliakou, Dimitra. 2002. On linkhood, topicalization and clitic left dislocation. *Journal of Linguistics* 38, 193–245.

Borsley, Robert D. 1989. An HPSG approach to Welsh. *Journal of Linguistics* 25, 333–354.

Borsley, Robert D. 1995. On some similarities and differences between Welsh and Syrian Arabic. *Linguistics* 33, 99–122.

Calcagno, Mike. 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, Hye-Won. 1999. *Optimizing Structure in Context*. Stanford: CSLI Publications.

De Kuthy, Kordula. 2002. *Discontinuous NPs in German − A Case Study of the Interaction of Syntax, Semantics and Pragmatics*. Stanford, CA: CSLI Publications.

De Kuthy, Kordula and Meurers, Walt Detmar. 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*. Stanford: CSLI Publications.

Donohue, Cathryn and Sag, Ivan A.. 1999. Domains in Warlpiri. Ms. Stanford University.

Engdahl, Elisabeth and Vallduví, Enric. 1996. Information packaging in HPSG. In C. Grover and E. Vallduví (eds.), *Edinburgh Working Papers in Cognitive Science, vol 12: Studies in HPSG*, 1–31.

Frank, Anette. 1994. Verb second by lexical rule or by underspecification. Arbeitspapiere des SFB 340 No. 43, IBM Deutschland GmbH, Heidelberg.

Gunji, Takao. 1986. Subcategorization and Word Order. In W. J. Poser (ed.), *Papers from the Second International Workshop on Japanese Synta*x, 1–21. Stanford: CSLI Publications.

Hinrichs, Erhard W. and Nakazawa, Tsuneko. 1989. Subcategorization and VP Structure in German. In *Aspects of German VP Structur*e, SfS-Report-01-93, Eberhard-Karls-Universität, Tübingen.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jacobs, Joachim. 1986. The syntax of focus and adverbials in German. In W. Abraham and S. de Meij (eds.), *Topic, Focus and Configurationality. Paper from the 6th Groningen Grammar Talks, Groningen, 1984*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Kathol, Andreas. 2000. *Linear Syntax*. Oxford: Oxford University Press.

Kathol, Andreas. and Carl Pollard. 1995. Extraposition via complex domain formation. In *33rd Annual Meeting of the ACL*, 174–180. San Francisco: Morgan Kaufmann

Kiss, Tibor and Wesche, Birgit. 1991. Verb Order and Head Movement. In O. Herzog and C.-R. Rollinger (eds.), *Text Understanding in LILO*G, 216–242. Berlin/Heidelberg/New York: Springer-Verlag.

Meurers, Walt Detmar. 2000. Lexical Generalizations in the Syntax of German Non-Finite Constructions. Arbeitspapiere des SFB 340 No. 145, Eberhard-Karls-Universität, Tübingen.

Müller, Stefan. 2004a. Continuous and discontinuous constituents? A comparison between syntactic analyses for constituent order and their processing systems. *Research on Language and Computation* 2, 209–257.

Müller, Stefan. 2004b. An analysis of depictive secondary predicates in German without discontinuous constituents. In S. Müller (ed.), *Proceedings of the HPSG04 Conference*. Stanford: CSLI Publications.

Müller, Stefan and Kasper, Walter. 2000. HPSG analysis of German. In Wahlster (ed.), *Verbmobil: Foundation of Speech-to-Speech Translation*, 238–253. Berlin/Heidelberg/New York: Springer-Verlag.

Netter, Klaus. 1992. On non-head non-movement. An HPSG treatment of finite verb position in German. In Görz (ed.), *Konvens 92.1. 'Konferenz Verarbeitung natürlicher Sprache'*. 218–227. Berlin/Heidelberg/New York: Springer-Verlag.

Netter, Klaus. 1998. *Functional Categories in an HPSG for German.* Saarbrücken Dissertations in Computational Linguistics and Language Technology, No. 3.

Deautsches Forschungszentrum für Künstliche Intelligenze und Universität des Saarlandes, Saarbrücken.

Noda, Hisashi. 1996. *'Wa' to 'Ga'* ['Wa' and 'Ga']. Tokyo: Kurosio.

Pollard, Carl. 1994. Toward a unified account of passive in German. In J. Nerbonne, *et al*. (eds.), 273−96.

Pollard, Carl. 1996. On head non-movement. In H. Bunt and A. van Horck (eds.), *Discontinuous Constituenc*y, 279–305. Berlin, New York: Mouton de Gruyter.

Pollard, Carl, Kasper, Robert and Levine, Robert. 1994. Studies in constituent ordering: Towards a theory of linearization in Head-driven Phrase Structure Grammar. Research Proposal to the National Science Foundation, Ohio State University.

Pollard, Carl and Sag, Ivan A. 1989. *Information-Based Syntax and Semantics*. Stanford: CSLI Publications.

Pollard, Carl and Sag, Ivan A. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Reape, Mike. 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, Elizabeth. 1995. Sentence prosody: Intonation, stress, and phrasing. In J. A. Goldsmith (ed.), *The Handbook of Phonological Theory*, 550−569. London: Basil Blackwell.

Yatabe, Shûichi. 1996. Long-distance scrambling via partial compaction. In M. Koizumi abd M. Oishi and U. Sauerland (eds.), *Formal Approaches to Japanese Linguistics 2* (MIT Working Papers in Linguistics 29). Cambridge, MA, 303−317.

Yatabe, Shûichi. 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.