PHRASE STRUCTURE BUILDING FROM THE PERSPECTIVE OF HUMAN COGNITION: AN ANALYSIS OF 'SYNTACTIC' PHENOMENA IN TERMS OF THE INFORMATION STRUCTURE OF NATURAL LANGUAGES

by

Yuji Hatakeyama

This paper is a critical review of Generative Grammar. In Generative Grammar, it has been assumed so far that phrase structure is constructed from right to left in the V'->VP->I'->IP order through Merge (cf. Chomsky 1995a). In this paper, taking into consideration human cognition, specifically the processing of information, I will argue against this general idea of phrase structure building and show that phrase structures are built up from left to right in the IP->I'->VP->V' order through branching. I will further claim that an item bearing old information is introduced in the phrase structure earlier than one assuming new information. Based on this new interpretation of phrase structure building, I will provide a real account of some 'syntactic' problems Generative Grammar has failed to deal with so far.

1. Introduction

In current research of Generative Grammar, the Minimalist Program, it is assumed that all the working hypotheses (including syntactic objects and syntactic operations) in the system need to be motivated by the Bare Output Condition (Chomsky 1995a, 1999, 2000a). The Bare Output Condition, a licensing condition of grammar which is imposed by the cognitive system external to the grammar system, constrains both the PF (Phonetic Form) and the LF (Logical Form) interfaces. What does the grammar system look like when it is constrained by the Bare Output Condition? In this paper I will answer this question from the perspective of phrase structure building.

This article is organized as follows. Section 2 considers human beings' cognition and parsing of sentences; it tackles the following two questions: (i) How are phrase structures built up? (ii) How are syntactic procedures applied? In section 3 I will show what kind of empirical advantages my approach to phrase structure building has. In this

section I will argue that my phrase structure building can (i) treat Huang's (1982) Condition on Extraction Domain and the Definiteness Effect in a unified way (sec. 3.1), (ii) capture Pesetsky's (1987) Discourse-linking Effects properly (sec. 3.2), and moreover (iii) straightforwardly account for why the *wh*-word *why* does not induce *wh*-island effects (sec. 3.3). In section 4, 'Concluding Remarks', I will show that there exists another functional category, SP (Sentential Phrase), which determines the type of sentences higher than CP.

2. Proposal: How are Phrase Structures Built Up?

In this section I will consider how languages are structured from the perspective of human cognition, especially the parsing of sentences. Human beings speak, hear, and understand languages. How do human beings speak and hear them? Let us consider this matter from the viewpoint of sentence parsing. I will begin with the aspect of hearing. Since a language is constrained by time, human beings cannot produce or pronounce two words at the same time. Thus, a hearer hears a sentence from the initial word to the final word through the intermediate words. Thus, if a sentence is constructed in the IP -> VP order, as I will contend below, a hearer should first hear a word in the IP domain and then one in the VP domain, constructing structures in this order. If not, garden path sentences such as (1), which induce some difficulties in parsing, would not be present in natural languages at all.

(1) a. The horse raced past the barn fell.
     b. The cotton clothing is made of grows in Mississippi.
     c. Without her contributions failed to come in.
     d. Since Jay always walks a mile seems like a short distance to him.

This is because if human beings construct structures from VP to IP and parse sentences in this order, a hearer would not go into a blind alley when he or she hears such garden path sentences as (1). Hence it is possible to argue that human beings parse a sentence from the

sentence-initial word to the sentence-final word through the intermediate words (cf. Just and Carpenter's 1980 Immediacy Principle; for the validity of their principle, see Frazier and Rayer's 1982 report). That is, phrase structures are built up from IP to VP, and not from VP to IP.

Based on their Frequency-based or Probabilistic Constraint-based Theories, MacDonald, Pearlmutter and Seidenberg (1994:700) argue that a sentence's degree of ambiguity and the possibility of 'garden path' sentences depend on to what extent the sentence's words are associated with (or familiar to) our everyday life. Weinberg (1999:312), examining MacDonald, Pearlmutter and Seidenberg's (1994) parsing system, argues that their parsing theory does not replace the grammar system, but rather complements it. My view is a bit different from Weinberg's. As will be shown below, I take parsing theory to be equal to grammar theory in its status and importance. I also regard both theories as closely related. That is, I consider that any linguistic knowledge does not exist without the premise of language processing (i.e. parsing), and that no language processing happens unless linguistic competence (that is, knowledge of language) is premised.[1]

Next, let us consider the aspect of speaking. A sentence or proposition consists of theme and rheme (cf. Halliday 1967-1968 and the works of Vilém Mathesius, Jan Firbas, Petr Sgall and other representatives of a Praguian approach to theme-rheme structure). Roughly speaking, theme corresponds to a subject and rheme to a predicate. Reflecting on myself, human beings seem to first decide to talk about something (theme) and then consider what or how it is (rheme), and not vice versa. If the reverse were the case, we human beings would first think about *Y* of *X is Y* and then consider *X*, which has the property of *Y*, but this seems to go counter to human beings' thinking process. Therefore, we can say that human beings first construct an IP, a domain of theme, and then a VP, a domain of rheme.[2]

According to Greenberg (1963), with regard to the relative order of S(ubject), O(bject) and V(erb), SOV (as in Japanese, Korean, Turkish), the orders SVO (as in English, French, Chinese) and VSO (as in Arabic, Irish, Tagalog) are dominant over other possible word orders. The number of VOS and OVS languages is very small: only

Malagasy and some Caribbean languages belong to VOS and OVS languages, respectively. As for the remaining possibility, OSV languages, to the best of my knowledge, no convincing argument has ever been made for their existence. According to Hosaka (1994), Arabian, a VSO language, is an SVO language at the base. VOS and OVS languages might also be SOV or SVO languages at the base, namely, languages whose S precedes V at the base. Thus, in most languages, S may precede V at the base. This seems to support the proposed supposition (that we human beings first think about theme and then about rheme) as a cross-linguistic generalization.

If the above is right, a sentence will not be constructed in the V' -> VP -> I' -> IP order, as has been generally assumed, but in the IP -> I' -> VP -> V' order, as I am contending here. Concretely, a sentence is constructed as follows. First, an IP node appears as a root, and then it branches into the specifier of IP (XP) and an I' node. Next, the branched I' further branches into the head I and a VP. VP, likewise, branches into the specifier of VP (YP) and a V' node, which further branches into the head V and the complement of V (ZP). Since XP, YP and ZP are all maximal projections, they may branch further. In this way, the structure of a sentence is built up from left to right by branching. Notice that the proposed left-to-right branching phrase structure building contrasts with the right-to-left merging phrase structure building assumed in the Minimalist Program.[3] Notice also that according to Nagao et al. (1999:124-126), Left-corner Parsing is most feasible with regard to psychological reality.

Indeed, from the above discussion, we can see how phrase structures grow with their nodes branching out. But, with respect to the growth of phrase structures, does a node grow upward or downward? Now, as a working hypothesis, let us assume here that branching takes place upward. That is, as has been generally assumed in Generative Grammar, let us suppose here that phrase structures are built up in a bottom-up fashion. But, notice here that the present phrase structure building is rather different from the standard phrase structure building; under the general phrase structure building, a 'tree diagram' is in fact a root diagram as a root grows downward, while under the present phrase structure building a root grows upward to be a tree with its branches branching out, representing a tree diagram.[4]

With this in mind, let us consider how (2) is structured. For ease of exposition, I will here assume that a sentence consists of two categories, IP and VP.

(2)   John ate an apple

First, an IP node appears as a root.

(3)   IP

Next, this IP grows upward, branching into the specifier of IP (XP) and an I' node.

(4)   XP   I'
          \   /
           IP

The I' node further branches into the head I and a VP node.

(5)   I   VP
              \   /
         XP   I'
          \   /
           IP

The VP in (5) branches into the specifier of VP (YP) and a V' node.

(6)   YP   V'
               \   /
           I   VP
            \   /
         XP   I'
          \   /
           IP

The V' further branches into the head V and the complement of V (ZP).

(7)
```
                    V    ZP
                     \   /
              YP    V'
               \   /
           I    VP
            \   /
        XP    I'
         \   /
          IP
```
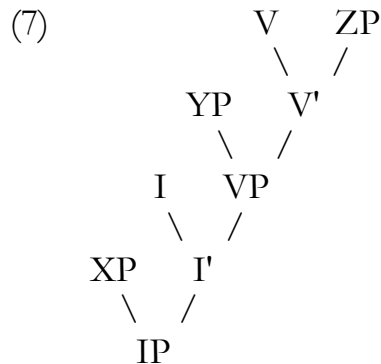
The structure of (2), represented as (8), is produced when we insert *John*, *[+past]*, t$_{John}$, *eat* and *an apple* at *XP*, *I*, *YP*, *V* and *ZP* of (7), respectively.[5]

(8)
```
                    eat    an apple
                      \   /
            t$_{John}$    V'
                \   /
          [+past]  VP
              \   /
          John    I'
            \   /
             IP
```
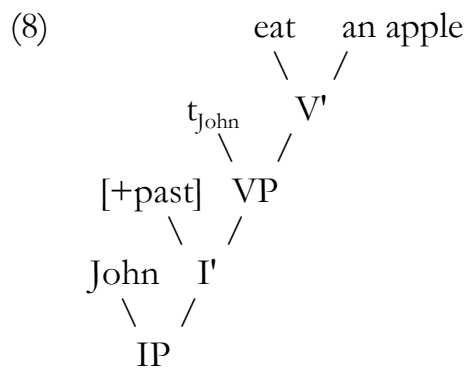
Here, we have a question to ask: under the proposed grammar system, how are lexical items introduced into the phrase structure? In what follows, I will tackle this question and show that an item bearing old information (hereafter an old item) is introduced to the phrase structure earlier than an item bearing new information (hereafter a new item).

Observe the following examples in English and Japanese. (9) - (10) are both adapted from Takami (1997:99).

(9) a. John met a 7-foot-tall girl on the street.
     b. *A 7-foot-tall girl met John on the street.

(10)   a. Taro-ga   aru joyuu to kekkon-shita
               Taro-Nom a   actress   married
               'Taro got married to an actress'
        b. *Aru joyuu-ga   Taro to kekkon-shita
                a   actress-Nom Taro   married
               'An actress got married to Taro'

(9a,b), both of which are equivalent to each other in terms of meaning, do not have any problems with either meaning or syntax. The same observation holds true for (10a,b); (10a,b) have no syntactic or semantic problems. But, as the contrasts in (9)-(10) show, the (a)-examples are acceptable whereas the (b)-examples are unacceptable. Why is there such a difference in acceptability?

As we have seen above, we cannot ascribe the unacceptability of the (b)-examples to their syntax or semantics. Neither do we attribute it to their phonology. Consequently, we are left with only one possibility: discourse. According to Quirk et al. (1985:1357, 1362), English (and other languages) have the following structure of information (see also Clark and Clark 1977; Clark and Haviland 1977; Gundel 1988; Gundel et al. 1993; Halliday 1967-1968:part 2; Kuno 1978:148-149; Prince 1992):

(11)   Information flows from old items to new items.

That is, in a linear order an old item (i.e., an element taking on old information) must precede a new one (i.e., an element taking on new information).[6] As we will see below, (9b) and (10b) are ruled out as they violate (11), a constraint on discourse.

In (9) two persons are present: *John* and *a 7-foot-tall girl*. When a speaker says (9), he or she knows that at this point *John*, representing old information, has already been introduced in the preceding context, whereas *a 7-foot-tall girl*, representing new information, is first introduced to the discourse. The same reasoning holds true for (10): *Taro* assumes old information, while *aru joyuu* 'an actress' assumes new information. In (9a) and (10a), an element assuming old information precedes an element assuming new information. In (9b) and (10b), on the other hand, an item with new information precedes an item with

old information. Thus, the (a)-examples, conforming to (11), are acceptable; the (b)-examples, failing to obey (11), are unacceptable.

As we can see from the above discussion, a speaker first produces an old item and then a new item. Corresponding to this, a hearer first hears what he or she already knows, and then hears what he or she does not yet know. This clearly reflects human cognition, for humans process new information based on old information, but not vice versa – old items are more accessible to both speaker and hearer than are new items (cf. Ariel 1990, Arnold 1998, Chafe 1994, Givón 1983, Gundel et al. 1993, Sperber and Wilson 1995[7], among others).

Now with this in mind, let us consider how lexical items are introduced into the phrase structure (7) we are assuming here. As we have seen above, under the present phrase structure building, which is based on human beings' cognition, an IP first appears and a VP is subsequently created. Further, as a working hypothesis, I am assuming here that a phrase structure is built up in the bottom-up fashion. Concretely, in the structure (8), *John* is first inserted into the specifier of IP, which is the lowest position in the phrase structure, and then *eat* and *an apple* are inserted into the head V and the complement of V respectively, both positions being the highest positions in the phrase structure, at the final stage of the derivation. Thus, with the present phrase structure building and the information structure (11), we can put forth the following Lexical Item Insertion rule:

(12)   Lexical Item Insertion
       Insert an old item into a lower position of the phrase structure; a new item is inserted into a higher position of the phrase structure. That is, insert an old item earlier than a new item when constructing phrase structures.

In the following sections, based on the proposed grammar system, especially on assumption (12), I will try to provide an account for some 'syntactic' phenomena from a functional perspective. I will then show that some facts considered 'syntactic' so far are actually functional, pragmatic, and thus are really explained by resorting to functional notions.

As is shown directly, my proposed grammar system, in contrast with Chomskyan linguistics, largely rests on pragmatics; as (13a,b) show, Chomsky admits that pragmatics is important for the study of natural language, but he denies that pragmatics can be the object of empirical science, as (14a,b) illustrate. (13a) is quoted from Stemmer (1999:398-399), (13b) from Chomsky (1995b:26), (14a) from Chomsky (1995b:27), and (14b) from Chomsky (2000b:69).

(13) a. My own view has always been ... [that] a general linguistic theory must incorporate pragmatics ... as a central and crucial component ... '[P]ragmatics' must be a central component of any linguistic theory that aims to be comprehensive.

b. It is possible that natural language has only syntax and pragmatics.

(14) a. [G]eneral issues of intentionality, including language use, cannot reasonably be assumed to fall within naturalistic inquiry.

b. The study of communication in the actual world of experience is the study of the interpreter, but this is not a topic for empirical inquiry, for the usual reasons: there is no such topic as the study of everything.

In Chomsky (1957:102-103; 1992a,b), as well, a view of pragmatics is presented which is not effectively different from (14a,b). But Chomsky (1975) suggests the existence of pragmatic competence as well as linguistic competence. And Chomsky (2000b:132) further doubts the existence of semantics, and he suspects that the language system consists of only syntax and pragmatics. See also Chomsky (1957:102-103) for the 'absence' of semantics from the language system.

To sum up, in this section I have considered how phrase structures are built up from the perspective of human cognition. It was shown that the generally assumed phrase structure is in fact a root diagram, not a tree diagram, and that it does not correctly reflect the human cognitive process (or, specifically speaking, their parsing of sentences). Thus, as an alternative, I proposed that phrase structures

are actually constructed in a bottom-up fashion from left to right in the IP -> I' -> VP -> V' order through branching. This is contrasted with the generally assumed phrase structure building, where phrase structures are considered to be constructed from right to left in the V' -> VP -> I' -> IP order through Merge. Further, with regard to lexical insertion, I put forward the proposal that an item bearing old information must be introduced to the phrase structure earlier than an item bearing new information. To put it another way, I proposed that an old item must be inserted into the slot nearest IP, whereas a new one must be inserted into the slot nearest VP. We have also observed that the proposed lexical insertion restriction is well supported by the way humans process information and by the cognitive and functional aspects of natural languages.

3. *Empirical Advantages*

In this section I will show some empirical advantages of the phrase structure building presented here. Especially, in the following sub-sections, I will examine several 'syntactic' phenomena from the perspective of human beings' cognition (more specifically speaking, from the perspective of the information structure of natural languages), and argue that the proposed grammar system can straightforwardly account for the 'syntactic' phenomena.

3.1. A Uniform Explanation of Huang's (1982) Condition on Extraction Domain and the Definiteness Effect

In this section, based on the phrase structure building presented above, I will consider Huang's (1982) Condition on Extraction Domain (i.e., subjects and adjuncts, unlike objects, constitute an island for extraction) and the Definiteness Effect (i.e., extraction from definite elements is banned). In what follows, I will first consider (i) why subjects constitute an island for extraction, contrasting with objects, and then (ii) why definite elements also constitute an island for

extraction. And finally, I will consider (iii) why adjuncts, like subjects and definite elements, also form an island.

This section considers these three questions from the perspective of a natural language's function and shows that these 'syntactic' phenomena are actually functional, not syntactic, and that the present grammar system can easily answer them. Now, let us consider the so called Subject Condition. As is well known, subjects and objects show an asymmetry with respect to extraction.

(15)   a. *Who$_i$ is [a picture of t$_i$] on sale?
       b. Who$_i$ did you see [a picture of t$_i$] ?

The above contrast has been provided with various technical accounts in Generative Grammar (though these are not essentially different from each other). Generative Grammar consistently claims that subject positions constitute a barrier (or an island) for extraction, while object positions do not. To put it differently, it is considered in Generative Grammar (more specifically speaking, in the Barriers' system) that since the specifier of IP is not L-marked it resists extraction of an element, whereas the complement of V, being L-marked, allows for extraction of an element (Chomsky 1986b, 2000a). But this structure-based account amounts to saying that subjects form an island for extraction while objects do not. That is, it is just a (technical) description, not an account, of facts.

As long as we refer to structural notions, we cannot really explain the contrast in (15) (and other island phenomena). As we have seen above, if we base ourselves on syntactic structures, we cannot get out of rephrasing facts technically. Let us see this more concretely, by taking some examples. For example, Koizumi (1995:37; fn. 22) proposes the following as a Minimalist Program version of Huang's Condition on Extraction Domain:

(16)   Koizumi's (1993) Minimalist Version of Huang's (1982) CED:
       A domain X is transparent with respect to extraction if (a) there is a head H that selects X, and (b) X is in the minimal domain of H; it is opaque otherwise.

(16) is merely a Minimalist version of Huang's Condition on Extraction Domain; by using the technical terms of the Minimalist Program, Koizumi technically rephrases the generalization that subjects and adjuncts constitute an island, while objects do not. That is, (16) still remains on the level of description and does not reach the level of explanation.

Further, Uriagereka (1999a, 1999b) claims that specifiers (that is, items in the subject position) form an island based on the following reasoning:

(17)　　a.　For a derivation to converge, Spell-Out can be applied several times in a cyclic fashion (Uriagereka 1999a:404, 426). This is the null hypothesis. Additionally, this Multiple Spell-Out is just a version of Halle and Marantz's (1993) Distributed Morphology (Uriagereka 1999a:422).

　　　　b.　Specifiers are formed by Spell-Out (or the Morphological Repair Operation) at some point of a derivation (Uriagereka 1999a:423).

　　　　c.　Overt syntactic operations are unavailable after Spell-Out.

　　　　d.　Thus, specifiers constitute an island or a barrier.

Indeed, as Uriagereka (1999a:404, 421) herself acknowledges, this logic holds under her (1999b) Dynamic Derivational System. But, her 'account' of islands (including Subject Island) in (17), just like Koizumi's 'account', is not an account either, but rather a description of facts, using technical terms.

Takahashi (1994), too, attempts to capture the effects of Huang's Condition on Extraction Domain by his (1994:25) Uniformity Corollary on Adjunction:

(18)　Uniformity Corollary on Adjunction (UCA):
　　　Adjunction is impossible to a proper subpart of a uniform group, where a uniform group is a non-trivial chain or a coordination.

His Uniformity Corollary on Adjunction involves the same difficulties that Koizumi's and Uriagereka's analyses do; also Takahashi's UCA is just a generalization of facts.

As the preceding section has shown, our reasoning is based on the idea that human cognition processes (especially, the parsing of sentences) largely affect the construction of phrase structures. Hence, building on this concept, we necessarily presuppose that 'syntactic' processes are also largely affected by cognitive ones. Below, I will explain the contrast in (15) (and other island phenomena) without resorting to syntactic structures.

(15a), in which extraction from subject takes place, is ungrammatical. On the other hand, (15b), where extraction from object occurs, is grammatical. (15a,b) are repeated below.

(15)   a. *Who$_i$ is [a picture of t$_i$] on sale?
        b. Who$_i$ did you see [a picture of t$_i$] ?

As we have seen in the preceding section, with respect to linear order, an item bearing old information needs to precede an item bearing new information (cf. (11)-(12)). Thus, a subject has old information while an object has new information. That a subject bears old information means that its semantic contents are already known to the hearer and the speaker. On the other hand, that an object bears new information means that its semantic contents are not yet known, at least to the speaker.

To ask a hearer about something which both the speaker and the hearer already know contradicts the principle of normal discourse /conversation. This discourse/conversation principle is effectively equivalent to Grice's (1975) Co-operative Principle and his Conversational Maxims. As we have noted above, the contents of a subject are already known to a speaker and a hearer. Thus, to inquire about (a part of) a subject is contrary to the discourse/conversation principle. Hence, (15a) is ruled out.[8,9]

A topic or topicalized element's content is already known to hearer and speaker. Thus, the proposed analysis of 'Subject Condition' predicts that extraction from within the topic or topicalized element is

banned. As the unacceptability of the following example, adapted from Lasnik and Saito (1993:101), indicates, this prediction is right:

(19) ??Vowel harmony$_j$, I think that [articles about t$_j$]$_i$, you should read t$_i$ carefully.

The above example therefore constitutes strong evidence for the present analysis.

Let us consider next the grammaticality of (15b). In (15b), extraction takes place from object. The object in question bears new information which is not familiar to the speaker. It is not strange that a speaker asks a hearer something the speaker does not understand well. This conforms to the discourse/conversation principle. Thus, (15b), not being confronted with any problems, is acceptable.

As the preceding discussion has shown, our analysis of phrase structure building, being based on the information structure of natural languages and being furthermore reducible to human beings' cognition, can easily capture the extraction asymmetry between subjects and objects.

The proposed grammar system can further correctly capture the ungrammaticality of the following example; in other words, it can explain the Definiteness Effect. The Definiteness Effect was first noticed by Chomsky (1973, 1977), and further discussed by a number of authors including Erteschik-Shir (1973), Oehrle (1974), Fiengo and Higginbotham (1981), Bowers (1988), Karimi (1989), Enç (1991), Mahajan (1992), and Diesing (1992).

(20) *Who$_i$ did you see [those pictures of t$_i$]?

In the above example, just as in (15b), extraction takes place from object; still, the sentence is ungrammatical, in contrast with (15b). A substantial difference between (20) and (15b) is that the object NP of the former is definite while that of the latter is indefinite. It is thus conceivable that this difference materializes as a difference in grammaticality between (20) and (15b).

How is the Definiteness Effect in (20) captured? Several generative linguists have tried to account for the ungrammaticality of

(20), but they fail to derive it from discernable principles. Their 'accounts' go more or less as follows: (i) Definite objects, such as in (20), contain a null operator which blocks extraction of an element from within them. That is, the Definiteness Effect is dealt with as a sort of Relativized Minimality effect. (cf. Endo 1995). (ii) Under the DP analysis, an NP dominated by DP is not L-marked and so constitutes an inherent barrier; D, which is sister to the NP, is not a lexical element. The DP immediately dominating NP inherits a barrierhood from NP. Hence, if we extract an element from within the NP, the relevant sentence runs afoul of the Subjacency Condition and is ruled out; two barriers are crossed. Thus, the Definiteness Effect can be dealt with as a sort of Subjacency Effect (cf. Bowers 1988). (iii) A definite object overtly raises to the specifier of AgrOP. The specifier of AgrSP constitutes an island (Subject Condition). Thus, it is not unreasonable to speculate that the specifier of AgrOP also forms an island. Therefore, if we extract an element from within the definite object, the relevant sentence is excluded. That is, the Definiteness Effect can be dealt with as a sort of island effect (cf. Mahajan 1992; see also Torrego 1998). As for the problems with Bower's (1988) and Mahajan's (1992) analyses of the Definiteness Effect, see Karimi (1999:sec. 2).

   Indeed, the Definiteness Effect seems to be successfully dealt with as a sort of Relativized Minimality, Subjacency, or island effect. But why does the definite object showing the Definiteness Effect have to contain a null operator? The postulation of a null operator within such objects is just a reverse operation of facts. There is no reason to assume a null operator in definite objects, and we cannot find any logical necessity for postulating such operators. Further, why must an element which is not L-marked be a blocking category (and consequently an inherent barrier)? And why should there be such a curious thing as inheritance barrier? There is no compelling reason to postulate such notions. Moreover, why does the specifier of Agr constitute an island? Logically speaking, it might be OK to say that it is only the specifier of Agr that cannot constitute an island. There is no reason to assume islandhood for such a specifier. Such an 'account' is no different from saying that we cannot extract an element from definite objects. That is, it is just a paraphrase of facts, using technical

terms. It would not be possible to really explain the Definiteness Effect if we were to rest on syntactic notions.

Under the present grammar system, the ungrammaticality of (20) can best be dealt with as we have done in the case of (15a). Both (20) and (15a) are repeated here.

(20) *Who$_i$ did you see [those pictures of t$_i$]?

(15) a. *Who$_i$ is [a picture of t$_i$] on sale?

That is, we can take the Definiteness Effect to be an allomorph of the Subject Condition. In (20), the object NP is preceded by the definite article *those*, and thus the object NP in question functions as a topic in the preceding context and is already familiar to the hearer and the speaker. As in the case of (15a), it is contrary to the discourse/conversation principle to ask a hearer about something which the speaker of (20) already knows. Thus, (20) is excluded for the same reasons that (15a) (and (19)) were ruled out. That is, the ungrammaticality of (20), as that of (15a), can only be accounted for by referring to the information structure of natural languages, not to their syntactic structure (see also Tunstall's (1996:309-311) DP Interpretation Hypothesis and Familiarity Corollary for Extraposition).

As we can see from the following contrast, in Japanese it is not allowed to ask a hearer about something that both hearer and speaker already know:

(21) a. nani-ni kansuru hon-o   Taro-wa Hanako-ni age-ta no
        what about   book-Acc Taro-Top Hanako-Dat gave Q
        'What did Taro give Hanako a book about?'
     b. *nani-ni kansuru {sono/ano} hon-o Taro-wa Hanako-ni age-ta no
        what about   the that   book-Acc Taro-Top Hanako-Dat gave Q
        'What did Taro give Hanako {the/that} book about?'[10]

Takami (1995:chap.3, sec.6.2) attempts to account for Definiteness Effects in the framework of Functional Grammar. In passing, Takami

(1995:chap.3, sec.6.3, 6.4) analyzes Extraposition from NP similarly to the present analysis of the Definiteness Effect. Additionally, Erteschik-Shir (1997) analyzes such island phenomena as Definiteness Effects and the Subject Condition from the viewpoint of Functional Grammar. She proposes that there exists an f-structure (focus structure), an independent representation consisting of topic and focus, at the interface where syntax, PF and semantics meet, and that several island effects can be dealt with even within Functional Grammar, given some functional constraints on the f-structure.[11]

Finally, let us consider how the present grammar system copes with the so-called Adjunct Condition. Observe the following examples where the Adjunct Condition is violated. (23a,b) are adopted from Radford (1988:487) and Chomsky (1986b:31), respectively.

(23)   a. *Who$_i$ did the government collapse [after the downfall of t$_i$]?
       b. *Who$_i$ did they leave [before speaking to t$_i$]?
          (*To whom$_i$ did they leave [before speaking t$_i$]?)

As the ill-formedness of (23) indicates, adjuncts form an island, as do subjects. Why do adjuncts constitute an island? In Generative Grammar, the Adjunct Condition has been 'accounted for' by using the following logic: adjuncts appear in an adjoined position. The position constitutes a barrier. They therefore constitute an island for extraction. But, this 'account' is not an account at all. Why does an element in an adjoined position constitute a barrier? Unless this question is answered, the question above – why do adjuncts form an island? – has not been answered either (cf. also Takahashi 1994, Toyoshima 1997, Ishi 1997, among others, for structure-based syntactic analyses of the Adjunct Condition).[12] Generative Grammar, as it stands, cannot account for the question as long as it sticks to formalism. In passing, as Martin and Uriagereka (2000:10) acknowledge, neither can the current framework of Generative Grammar, the Minimalist Program, give a systematic account of the so-called argument-adjunct asymmetries. Chomsky (2002:152) also acknowledges, on the basis of data put forward by Postal (1998), that no real account has been provided for island effects so far.

How about the proposed grammar system which is based on human cognition and its functional aspects? Under the present grammar system, the Adjunct Condition is straightforwardly derived from the information structure of English, as were the Subject Condition and the Definiteness Effect. First, observe (24a,b), corresponding to (23a,b) respectively.

(24) a. The government collapsed after the downfall of Prime Minister Mori.
      b. They left before speaking to John.

According to the information structure of English, the nearer an element is to the sentence-final position, the more important its information (cf. (11)-(12)). Thus, according to this information structure, the adjuncts in (24), occurring in the sentence-final position, are expected to bear important information. But this is not the case. The relevant adjuncts, compared with the preceding clauses, have relatively less important information. To see this, let us consider (24a). When we read (24a) with a natural intonation, it is to be construed as informing us about what happened to the government after the downfall of Prime Minister Mori; it does not tell us following what the government collapsed. In other words, what is asserted in (24a) is the preceding clause; the relevant adjunct is functioning as its premise or, put it another way, its 'background'). An element asserted is generally considered to contain important information, while one functioning as a premise bears less important information. Thus, we can say that in (24), the adjuncts, compared with the preceding clauses, have relatively less important information.

This is supported by the following queries and their corresponding responses. A in (25) corresponds to (24a), A in (26) to (24b).

(25) A: Did the government collapse after the downfall of Prime Minister Mori?
      B: No.

(26) A: Did they leave before speaking to John?
      B: No.

The hearer B of (25) negatively responds to speaker A's inquiry. Which portion of speaker A's utterance does hearer B negate? Of course, we can say that hearer B negates all the contents of the utterance of speaker A. But when we divide the utterance of speaker A into the preceding clause and the following adjunct, which part does hearer B negate? When speaker A says sentence (25) in a natural way, hearer B negates the statement of the preceding clause, but not of the adjunct. That is, hearer B in (25) argues that the government did not collapse, but not that the government collapsed not after, but before the downfall of Prime Minister Mori. To be eligible as a focus of negation, an element must bear important information. Thus, from the above observation, we can say that the adjunct in (24a) bears relatively less important information, compared with the information of the preceding clause. This holds true for (26), too.

There is another piece of evidence for the argument that the adjuncts in question bear relatively less important information. Observe (27).

(27)   a. After the downfall of Prime Minister Mori, the government collapsed.
       b. Before speaking to John, they left.

The above examples, where the adjuncts in (24) are preposed to the sentence-initial position, are acceptable. A sentence-initial element, functioning as the theme of the relevant sentence, takes on less important information. That is, unless an element bears less important information, it cannot be fronted to the sentence-initial position. Hence, we can take the grammaticality of (27) to indicate that the adjuncts occurring in (24) bear less important information.[13]

Based on the above discussion, we can say that the adjuncts in (24) corresponding to those in (23) bear less important information. Now, with this in mind, let us consider the ungrammaticality of (23), reproduced here.

(23)   a. *Who$_i$ did the government collapse [after the downfall of t$_i$]?

b. *Who$_i$ did they leave [before speaking to t$_i$]?
    (*To whom$_i$ did they leave [before speaking t$_i$]?)

In (23), *wh*-phrases are extracted from adjuncts. In a *wh*-question, a fronted *wh*-phrase is considered to function as the focus of the sentence. To put it differently, a speaker produces a *wh*-question sentence in order to have its *wh*-phrase assigned a value by the hearer. This means that for a speaker, the *wh*-phrase takes on the most important information in the relevant *wh*-question sentence. As we have seen above, the information the adjuncts in (23) provide is low with respect to importance. This means that an element contained in the adjuncts also bears relatively less important information, in comparison with the preceding clauses. Thus, *wh*-questioning of a less important element is contrary to the *wh*-question's function or, more generally speaking, to the principle of discourse or conversation. Therefore, (23) is excluded.

Let us summarize the discussion of this section. I have first considered why extraction from objects is permitted, whereas extraction from subjects is not. I have then observed that this asymmetry in extraction possibility between subjects and objects can be straightforwardly accounted for by considering the information structure of natural languages. I have further considered why extraction from objects is not allowed when they are definite. I have shown that this Definiteness Effect can be treated in the same manner as the Subject Condition. I have also considered why adjuncts, like subjects, constitute an island for extraction. I have shown that since adjuncts, like subjects, bear less important information, they form an island for extraction. Thus, insofar as the discussion above is correct, it can be concluded that the proposed grammar system can treat Huang's (1982) Condition on Extraction Domain and the Definiteness Effect in a unified way.

3.2. Discourse-linking Effects

There is another 'syntactic' phenomenon closely related to discourse: Pesetsky's (1987) Discourse-linking Effect. As we can see from (28),

when two *wh*-phrases appear in one clause, a *wh*-phrase in the higher position in the phrase structure is predominantly *wh*-moved over one in the lower position, in standard phrase structure terms. (In present phrase structure terms, we can say it as follows: a *wh*-phrase in the lower position of the phrase structure is predominantly affected by *wh*-movement over one in the higher position).

(28)   a.   Mary asked [who read what]
   b.   *Mary asked [what who read]

But, as (29) indicates, when two *wh*-phrases have been already introduced in the preceding context and are functioning as a topic, no Superiority Effects show up.

(29)   a.   Mary asked which man read which book
   b.   Mary asked which book which man read

Pesetsky accounts for the grammaticality of (29b), where Superiority Effects disappear, as follows. The *wh*-phrase *which book* in the embedded clause is an internal argument of the verb *read*. It is then generated in the complement position of *read* at the base. At some point of the derivation from D-structure to S-structure, the *wh*-phrase *which book* moves to the specifier of CP of the embedded clause. In contrast, the subject *which man* of the embedded clause does not move anywhere in overt syntax, but remains in the specifier of IP, its base-generated position, and undergoes semantic interpretation in this original position.[14] Since the subject of the embedded clause in (29b), unlike that of the embedded clause in (28b), does not undergo covert movement for semantic interpretation reasons, (29b), unlike (28b), does not incur the Empty Category Principle violation. Thus, (29b), unlike (28b), is acceptable. This is Pesetsky's account of the grammaticality of (29b).

 While Pesetsky's account of Discourse-linking Effects, as observed in (29b), is successful, his analysis, in its present state, cannot be incorporated into current Generative Grammar, that is, the Minimalist Program. This is because under the Minimalist Program, due to the Minimal Link Condition (henceforth MLC), which is a

constituent of the definition of Attract, in (29) the *wh*-phrase *which book* is not allowed to move to the specifier of CP across the other *wh*-phrase *which man*. Thus, under the Minimalist Program, as it stands, the grammaticality of (29b) would still remain a problem.

In addition to this, Pesetsky's analysis of Discourse-linking Effects faces another serious problem, as pointed out by Lasnik and Saito (1993). Lasnik and Saito (1993:118) observe that superiority effects disappear in (30).

(30) Who wonders what who bought?

The above example is acceptable as long as *who* of the embedded clause takes matrix scope; that is, as long as it is associated with *who* of the matrix clause, but not with *what* of the embedded clause (Lasnik and Saito 1993:118-119). To capture this fact, Chomsky (1995a:387) reasons as follows: since *who* of the embedded clause and *who* of the matrix clause are not different in form, they can be related to each other. On the other hand, *who* of the embedded clause and *what* of the embedded clause are distinct in form, so they cannot be associated with each other. Chomsky (1995a:387) presents the following example to support his idea:

(31) What determines to whom who will speak?

In the above example, *who* of the embedded clause is similar to *whom* of the embedded clause in form, but dissimilar to *what* of the matrix clause in form. Therefore, according to Chomsky's prediction, *who* will take embedded scope, but not matrix scope; that is, it is associated with *whom*, but not *what*. The correctness of this prediction, in fact, strongly suggests that superiority is not a purely syntactic phenomenon, but rather a pragmatic or semantic phenomenon, as I will show below. In agreement with this, Fukui (1997:57) suggests that superiority itself might be outside of the computational component of the human language system.

Furthermore, Hornstein (1995:130) claims that superiority effects are a sort of weak crossover effects. According to him, (32a) is assigned the LF structure (32b).

(32) a. *What did who buy?
      b. [what$_i$ [[pro$_i$ N] buy t$_i$]]

The subject NP *who* is converted to *[pro N]* in LF, and *pro* in *[pro N]* does not c-command the trace t_i of the raised element *what$_i$*. Thus, the ill-formedness of (32a) can be taken to be a realization of weak crossover effects. This analysis of superiority effects by Hornstein is in agreement with the above-mentioned suggestion: superiority might be a pragmatic or semantic matter, a view I share, as will be seen below. In effect, weak crossover effects are semantic rather than syntactic.

Can the proposed grammar system, based on human cognition (or especially on its processing of information), in any way correctly capture the grammaticality of (29b) (repeated below)?
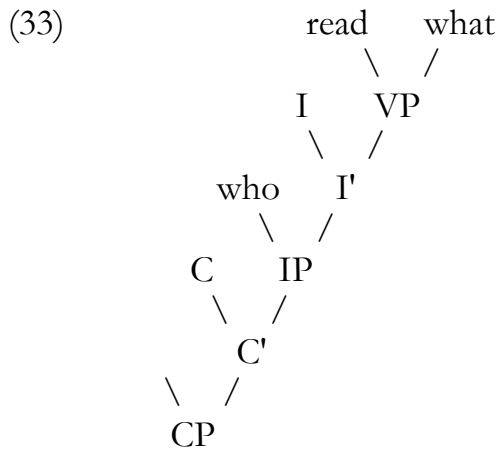
(29) b. Mary asked which book which man read

According to the new grammar system, a discourse-linked element should appear near the clause-initial position by virtue of the information structure (cf. (11)-(12)). Thus, when the phrase structure rises to the embedded CP clause, the discourse-linked *wh*-phrase *which book* is introduced directly into the specifier of CP of that clause. Now suppose that derivation further proceeds to the final stage. When the *wh*-phrase *which man* and the verb *read* are inserted into the specifier of IP of the embedded clause and V of the clause, respectively, the result is (29b). Suppose now that the interpretation of the subject *wh*-phrase is executed by way of Absorption, in the sense of Higginbotham and May (1981); that is, that the subject can be interpreted without moving to the Comp of the embedded clause. In this derivation of (29b), unlike Pesetsky's analysis, the discourse-linked *wh*-phrase *which book* does not move at all. Thus, the violation of the Minimalist Program's MLC does not exclude the availability of the Minimalist Program's MLC in this new grammar system (see section 4 below). Therefore, (29b) is acceptable. Incidentally, under the present grammar system, the generation of (29a,b) is reducible to the question of which *wh*-phrase between *which man* and *which book* is first chosen from the Numeration.

Can the proposed grammar system correctly exclude (28b) (repeated here)?

(28) b. *Mary asked [what who read]

It captures the ungrammaticality of (28b) by the Minimalist Program's MLC (more specifically speaking, Attract). The two *wh*-phrases in (28b) are not discourse-linked. Therefore, neither of them is directly inserted into the specifier of CP of the embedded clause. The one *wh*-phrase *who* and the other *wh*-phrase *what* are introduced into the specifier of IP and the complement position of V, respectively. Thus, at some point of the derivation, (28b) has the following structure of the embedded clause:

(33)   read   what
                            \   /
                    I   VP
                      \   /
              who   I'
                 \   /
          C   IP
            \   /
              C'
        \   /
          CP

Intuitively, the *wh*-phrase *who* occupying the specifier of IP is nearer to the specifier of CP than is the *wh*-phrase *what*, occupying the complement position of V. Thus, the former is required to move to the specifier of CP of the embedded clause. In spite of this, in (28b) not *who* but *what* is moved to the specifier of CP. Hence, (28b) is ruled out as violating the MLC

From the above discussion, the following emerges: the grammar system is first constrained by some cognitive and functional conditions of natural languages, which are reducible to Bare Output Conditions. That is, an element bearing old information is introduced to the phrase structure earlier than one bearing new information (cf. (12)). Then, in order to convert the information structure, syntactic operations,

including Extraposition from NP, Heavy NP Shift, Stylistic Inversion, Passivization, Raising, etc., are invoked internally to the grammar system, within the syntactic, computational component.[15]

As is clear from the above discussion, the structure building presented here is not the same as the one assumed in Generative Grammar. And they are in effect incompatible with each other; in the 'standard' Generative Grammar, it is supposed that a syntactic structure is created not for some functional reason, but for formal/morphological reasons; syntactic structure happens to match up with a certain information structure. That is, syntactic structure is primary, while information structure is secondary or subsidiary. Strictly speaking, Generative Grammar supposes that language computation is 'blind' and therefore does not have any teleological purpose or intention: syntax has no functional purpose. As shown above, the present analysis of phrase structure building is incompatible with Generative Grammar, but this does not refute the present grammar system. It is completely natural and reasonable that the present system does not go along with Generative Grammar in terms of structure building; this is to be expected. The reason is that the present approach to phrase structure building bases itself on language processing, viz. parsing, as well as on linguistic competence. The discussion turns around the refutation of Generative Grammar's core idea that syntax is autonomous.

The present section also considered how Pesetsky's (1987) Discourse-linking Effects can be treated under the proposed grammar system. We humans process new or unknown information based on old or known information. Reflecting this aspect of our cognition, a linguistic element referred to in the preceding context, that is, an element bearing old information, is allocated to the clause-initial position, while an element bearing new information is placed in the rear position (cf. (12)). It is shown that human cognition, more specifically the information structure of natural languages, can properly deal with Pesetsky's Discourse-linking Effect.

3.3. Why Does a *Wh*-phrase *Why* Induce No *Wh*-island Effects?

Just as the Discourse-linking Effects, observed in Pesetsky (1987), could be accounted for on the basis of the information structure of natural languages, this section will introduce another 'syntactic' phenomenon which can also be explained by referring to this information structure.

　　When *wh*-movement takes place within an embedded clause, it is impossible to extract another element from within the embedded clause. To put it another way, a *wh*-phrase or *wh*-movement creates a *wh*-island. (34a,b) are taken from Cinque (1990:1), and (34c) is borrowed from Lasnik and Uriagereka (1988:20).

(34)　a.　??To whom didn't they know when to give their present t?
　　　b.　*How did they ask you who behaved t?
　　　c.　?*What does Bill wonder who [t saw t]?

But, as the grammaticality of the following examples indicates, we can associate a *why wh*-phrase appearing in the matrix clause, with the embedded clause across the intervening *wh*-phrase.

(35)　a.　Why do you wonder who created the world?
　　　b.　Why do you wonder what he's thinking?
　　　c.　Why do you care what he said?
　　　d.　Why do you care who ate the pie?

That is, extraction of *why* seems to be exempted from *wh*-island effects.[16]

　　To capture this fact, Generative Grammar has been so far stipulating that a *why wh*-phrase is base-generated in its surface position (cf. Rizzi 1990, Cinque 1990:92-93). This analysis is merely a postulate, not an explanation at all, so the grammaticality of (35) needs to be explained.

　　The present phrase structure, which is largely based on human cognition, can correctly capture the grammaticality of (35). I will here take (35a) as a case study to explain why (35) is well-formed. In (35a),

that someone created a world is premised and the sentence is construed as questioning the reason why he or she did so. This means that the sentence-initial *why wh*-phrase cannot be meaningful, unless its associated statement (in this case, the statement that someone created a world) is premised. In other words, a *why wh*-phrase which takes its modified sentential statement as its preceding discourse can be used if and only if it is properly linked to the statement. To exemplify this: if someone asks why John got married to Mary, while at the same time the preceding context does not contain a single topic indicating that John got married to Mary, that speaker would not be taken seriously. Thus, a *why wh*-phrase can be related to be a discourse-linked element; *why* is discourse-linked to its modified, preceding sentential statement. It is then possible to locate a *why wh*-phrase in its surface position directly at the base, just as we base-generated *which man* of (29a) and *which book* of (29b) in their surface position. Since a *why wh*-phrase is not extracted from within an embedded clause, no *wh*-island effects (namely, MLC violations) occur. This is my explanation of the grammaticality of (35).

To sum up, this section has considered why a *why wh*-phrase does not induce *wh*-island effects. This question has been 'answered' in Generative Grammar by stipulating that *why* is base-generated in its surface position. But this 'account' by Generative Grammar is not an account, but merely a description of facts. Above, I have offered a principled (more specifically, a natural language information structure-based) answer to this question.

4. Concluding Remarks: The Potentiality and Further Expansion of the Proposed Grammar System

Thus far, I have considered what the mechanism of the grammar of natural languages looks like when we construct a grammar system on the basis of human cognition. I have first considered the construction of syntactic structures, showing that structures rise from left to right in the IP -> I' -> VP -> V' order by branching. I have then shown that an element assuming old information is introduced into the phrase structure earlier than one bearing new information. This lexical
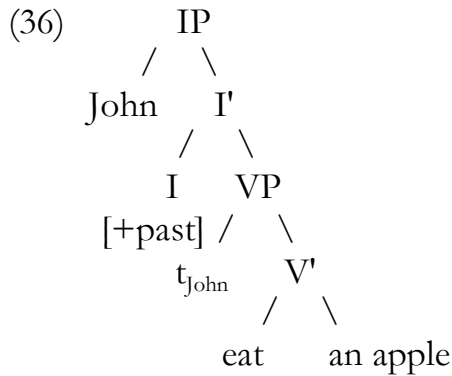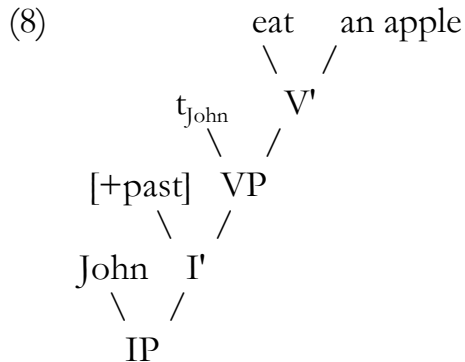
insertion was shown to be well motivated by the way humans process information.

Based on the proposed grammar system, I have given real accounts for several 'syntactic' phenomena, which Generative Grammar had failed to explain. I have first dealt with Huang's (1982) Condition on Extraction Domain and the Definiteness Effects from the functional perspective of natural languages. I have then shown that the so-called Subject Condition, Adjunct Condition, and Definiteness Effects are all the same functional phenomenon and so can be dealt with in a unified way. Further, I have considered Pesetsky's (1987) Discourse-linking Effects, arguing that these are also accounted for straightforwardly if we take into consideration the information structure of natural languages. Specifically, I have claimed that a discourse-linked *wh*-phrase is directly inserted into the surface position at the base; that is, it is not placed there through movement operation. Additionally, I have accounted for why a *wh*-phrase with *why* does not induce a *wh*-island effect. The *wh*-phrase is discourse-linked to its preceding sentential statement and is thus located directly in its surface position, as other discourse-linked *wh*-phrases are. In this way, it is exempted from *wh*-island effects, which are induced only if *wh*-movement actually occurs across another *wh*-phrase.

Finally, before concluding this section, let us consider the theoretical consequences of the proposed grammar system. First, if our suppositions (especially, as to phrase structure building) are correct, all movement should be lowering rather than raising. This is so since under the proposed phrase structure, IP is located lower than I', which is located lower than VP (cf. (7)). Would not, then, this arboreal picture imply some serious consequences for the theories of syntax?

To answer this question, let us compare the proposed phrase structure (8) with the generally assumed phrase structure (36), both of which are tree-type representations of (12), reproduced here:

(12)   John ate an apple

```
(8)                eat    an apple
                     \    /
          t_John      V'
             \       /
         [+past]   VP
              \    /
          John   I'
             \   /
              IP
```

```
(36)        IP
           /  \
       John    I'
              /  \
            I    VP
        [+past] /  \
           t_John   V'
                  /  \
               eat    an apple
```

In (8), IP, a root, grows upward through branching. In contrast, in (36), IP, a 'root', grows downward through branching. Visually, (8) contrasts with (36). Of course, this presupposes that phrase structures are built in a bottom-up fashion (recall the discussion in section 2). If we assume as a working hypothesis that structures are built in a top-down fashion, our phrase structure is not visually different from the generally assumed phrase structure (36) at all. But notice that the proposed grammar system and Generative Grammar are absolutely different from each other in the way phrase structures are built up. Under the present grammar system, whether phrase structures are built in a bottom-up manner or in a top-down manner, phrase structures are created in the IP -> I' -> VP -> V' order, namely, from left to right, by branching, not by Merging, in contrast with the Minimalist Program.

How are (8) and (36) represented if we rewrite them in brackets? (8) and (36) have the same structures in (37). This means that they are also not different from each other even if we rewrite them in terms of set theory, as has been often done in the Minimalist Program; to the

31

best of my knowledge, Chomsky (1957:27-29, 87) would be the first to formalize phrase structure in terms of set theory.

(37) $[_{IP}$ John $[_{I'}$ I $[_{VP}$ t_{John} $[_{V'}$ eat an apple$]]]]$

That is, (8) and (36) are not distinct from each other in terms of the relation of domination between two nodes (or the Minimalist Program's *terms*). Thus, if we take movement to be a change of position between a dominated position and a dominating position, (8) and (36) are not different from each other with respect to movement. This suggests that if the Minimalist Program's MLC is a real syntactic process and can truly explain syntactic phenomena, we can incorporate the procedure into the proposed grammar system with its effects maintained. Actually, I used this procedure in sections 3.2 and 3.3 to explain syntactic phenomena under the presupposition that MLC is real and basically correct.

The preceding discussion might give the impression that the proposed phrase structure building is just a notational variant of the previous, standard phrase structure building. But, there is a big difference between them – a conceptual difference. As I have shown in section 2, the present phrase structure building is based on human cognition, especially on humans' processing of natural language information. To put it differently, the present phrase structure building is largely motivated by the Minimalist Program's Bare Output Condition. In contrast, in Generative Grammar, phrase structure building has not been motivated too well by conditions external to the grammar system (as has the proposed grammar system), but rather, is simply stipulated as such. That is, it seems to me that there are no strong motivations, based on human cognition, to support Generative Grammar's hypothesis that phrase structures are built up from right to left in the V' -> VP -> I' -> IP order through Merging. Thus, even if there are no substantial empirical differences between the present grammar system and Generative Grammar, if all the notions and procedures necessary to a grammar system should be motivated by the Bare Output Condition, as proclaimed in the Minimalist Program, we can still safely say that the proposed phrase structure building is superior to the standard phrase structure building assumed in

Generative Grammar, for conceptual reasons. Actually, as shown in section 3, there is a substantial difference in explanatory strength between the proposed grammar system and Generative Grammar, which indicates that the former system is superior to the latter in empirical respects as well.

Next, I would like to consider how an external argument is assigned its θ-role. In Generative Grammar, a phrase structure builds up in the V' -> VP -> I' -> IP order. In a closer look at a VP-internal domain, V first selects its internal argument (namely, assigns its θ-role to its indispensable element) and rises to V'. Then V' takes an external argument, becoming VP. At this point, V' gives a θ-role (<Agent> or <Causer>) to the external argument (hereafter, <   > stands for a θ-role). It is important here that an external argument be assigned its θ-role not by V, but by V'. Thanks to this external θ-role assignment, we can account for why *John* in (38a) bears a θ-role <Agent> while *John* in (38b) assumes a θ-role <Patient> (cf. Chomsky 1986a:59-60).

(38)   a. John broke the window
          b. John broke his arm

Under the proposed grammar system, phrase structures are built up in the IP -> I' -> VP -> V' order, in contrast with Generative Grammar. In the proposed grammar system, VP first branches into the specifier of VP (an external argument) and V', which is then branched into the head V and the complement (an internal argument). Can this system correctly capture the fact that V', but not V, assigns a θ-role to its external argument? As will be seen below, this problem, i.e., the interpretation of (38a,b), can be dealt in a way that is similar to our treatment of garden path sentences.

Suppose that a hearer hears the sequence of words, *John broke*. At this point, the hearer considers the default θ-grid format of *break*, i.e., (<*Agent*>, <Theme>) (an italicized θ-role stands for an external θ-role). Then he or she assigns *John* a θ-role <Agent>. Now the verb *broke* has another θ-role <Theme>, to be assigned to its internal argument. If the nominal expression *the window* is taken from the Lexicon as an internal argument and is then assigned a θ-role <Theme>, then (38a) is produced. On the other hand, if the nominal

expression *his arm* is chosen from the Lexicon and is then provided with a θ-role <Theme>, the linear order of (38b) appears. But this chain of words, or sentence, is semantically anomalous; it is unusual and unreasonable that John broke his own arm intentionally. In short, when a hearer hears a sentence, he or she checks its semantic validity on the basis of common sense or his or her experience (cf. Kimball 1973). When a hearer feels a sentence represents a semantic anomaly, such as in *John broke his arm* with *John* bearing a θ-role <Agent> and *his arm* <Theme>, he or she will reconsider another θ-grid format for *break*, (<*Patient*>, <Theme>), and then assign the θ-roles to the two arguments, once again in accordance with the θ-grid. This θ-role assignment does not incur a semantic anomaly. Thus, *John* in (38b) bears a θ-role <Patient> and *his arm* takes on a θ-role <Theme>. In brief, θ-role assignment is executed in the same way that garden path sentences are parsed (cf. section 1): humans process a sentence from the initial word to the final word, and if they are confronted with some semantic or syntactic difficulties, they go back to the starting point or the initial word and then try to reanalyze the sentence in a different way.

Finally, I would like to conclude this discussion by presenting the possibility of further extending the proposed grammar system. As is clear from the preceding, the present conception of phrase structure building rests primarily on human cognition. Specifically, it is based on the parsing of sentences. Now, with this in mind, let us observe the following sentences:

(39)   a. John ate an apple. (=(2))
        b. What did John eat?

When humans prepare to say (39a,b) or something corresponding to them, there is a previous intention to say a declarative sentence in (39a) and a *wh*-question sentence in (39b) (cf. Ochs 1979:55, Clark and Wasow 1998). It is then conceivable that before we utter *John* in (39a) and *what* in (39b), some marker has already been placed in our minds, indicating a declarative sentence for (39a) and a *wh*-question sentence for (39b). Here, I will suppose that such markers are actually in our minds, and that they are realized in the phrase structure as SP (an

abbreviation for Sentential Phrase). Now, such an SP is neither original nor new at all. Many linguists have also put forward similar functional categories (see, for example, Katz and Postal (1964), Baker (1970), Bresnan (1972), Chomsky and Lasnik (1977), Pesetsky (1987), Cheng (1991), Aoun and Li (1993), Denham (2000), and others).

Chomsky (1995a:292) argues that declarative sentences such as (39a) are not an IP but a CP on the assumption that Comp determines the type of a clause. My SP is similar to his CP in terms of its function. But, as will be seen below, the former is utterly different from the latter, as it is closely associated with punctuation. Note further that I am here proposing a new functional category SP, which is distinct from CP. As is well known, the Minimalist Program subscribes to the Lexicalist Hypothesis, which assumes that all the elements are introduced to the phrase structure with their formal features already attached to them.[17] This means that in the Minimalist Program, when a speaker produces a sentence, he or she has already determined what type of sentence he or she tries to produce. This indicates that the Minimalist Program also presupposes something akin to my SP as a hidden assumption. Hence, if the discussion so far is on the right track, my phrase structure building can be taken to conform to the spirit of the Minimalist Program.

Suppose further that SP has a specifier and a complement, as do other functional and lexical categories. Then, (39a,b) will have the structures in (40a,b), respectively.

(40) a. $[_{SP}$ Spec $[_{S'}$ S $[_{IP}$ John ate an apple$]]]$
      b. $[_{SP}$ Spec $[_{S'}$ S $[_{CP}$ what did John eat$]]]$

Generally, the head of a functional category has its own formal features. Thus, the question is: What kind of formal features are there in the head S of (40a,b)? (40a) represents the structure of a declarative sentence such as (39a). Before a speaker says the sentence (39a), he or she has already determined to say a declarative sentence. It is then conceivable that in the head S of (40a) there exists a feature [Dec(larative)]. Further, it is reasonable to consider that this formal feature is realized as a period '.' in the PF component. This is because a declarative sentence is marked by punctuation with a period in the

sentence-final position. On the other hand, it is conceivable that in the head S of (40b) the feature [Q(uestion)] is existent, and that the formal feature is materialized as a question mark '?' in the PF component: a question sentence is marked by a question mark in the sentence-final position.[17]

It is generally assumed that the specifier of a functional category and its head enter a Spec-head agreement relation. Suppose now that this also applies to SP. Let us further assume that the complement of S lowers to the specifier of SP in overt syntax. (Recall that in the proposed grammar system, all movements are lowering). (39a, b) then have the structures in (41a,b), respectively, at the point of Spell-Out.

(41)   a. $[_{SP}$ $[_{IP}$ John ate an apple] $[_{S'}$ $S_{[Dec]}$ $t_{IP}]]$
       b. $[_{SP}$ $[_{CP}$ what did John eat] $[_{S'}$ $S_{[Q]}$ $t_{CP}]]$

In (41a), the IP's move to the specifier of SP is in agreement with the head S, as both elements bear the same formal feature [Dec]. Thus, the head S is properly licensed and materializes as a period '.' in the PF component. In (41b), on the other hand, the CP, which has lowered to the specifier of SP, agrees with the head S in terms of formal features, as both of them have the same formal feature [Q]. Thus, the head S is properly licensed and is realized as a question mark '?' in the PF component.

The proposed grammar system can also capture the ungrammaticality of the following examples:

(42)   a.   *John ate an apple? (in the meaning of a declarative sentence)
       b.   *What did John eat. (in the meaning of a *wh*-question sentence)

In (42a), the head S takes on a formal feature [Q], but an IP, which is specified for [Dec], has moved to the specifier of SP. Thus, a Spec-head agreement fails to be executed, causing the relevant sentence to be excluded. In (42b), too, though the formal feature of the head S is [Dec], a CP, whose feature specification is [Q], has moved to the specifier of SP. Therefore, the relevant example is ruled out due to its

failing to form a Spec-head agreement relation. As the above discussion shows, the proposed grammar system can correctly capture the ill-formedness of (42a,b) as well as the well-formedness of (39a,b).[19]

Thus far, I have reviewed some assumptions of Generative Grammar, a grammar system maintaining the autonomy of syntax. This section in a sense criticizes Generative Grammar, but never denies the existence of the phrase structures on which Generative Grammar is crucially based in its account of syntactic facts; phrase structures do exist and are necessary to capture a multitude of linguistic data. However, as I have presented it in section 2, my phrase structure theory is completely different from Generative Grammar's in that it assumes that phrase structures are built from left to right in the IP -> I' -> VP -> V' order through branching.

The current framework of Generative Grammar, or the Minimalist Program, has vigorously promoted its formalism in order to construct a grammar system. Because of its radical approach to C_{HL} (Human Language Computation), some important aspects of natural languages have been neglected in the current system of Generative Grammar. What has been neglected in Generative Grammar (especially in the Minimalist Program) are natural language's cognitive and functional aspects (cf., for example, the papers in Tomasello (1998) and references cited therein).[20] In contrast to Generative Grammar's strategy of leaving aside language's cognitive and functional aspects in constructing a grammar system, recently, in theoretical linguistics, such facets have been highlighted; for example, Cognitive Grammar and Systemic-Functional Grammar have paid much attention to these matters. If we only look at natural language's cognitive aspects, we fail to capture its formal and functional aspects; but if we only look at natural language's functional aspects, we fail to capture its cognitive and formal aspects. The most desirable grammar system, I think, is one which incorporates all aspects of natural languages in a balanced way.

This paper represents an effort to integrate Generative Grammar, Cognitive Grammar, and Functional Grammar into the grammar system I have put forward here. While Generative Grammar, on the one hand, and Cognitive Grammar and Functional Grammar on the other, are completely different, having no points of intersection, still,

37

there are similarities as well as differences between them, as has been argued by authors such as Newmeyer (1998), Croft (1995), and Ackerman (2000), among others. For example, Newmeyer argues that ultimately there is no principled incompatibility between the formalist and the cognitive-functionalist and that formalism and cognitive-functionalism can inform one another in an industrious partnership of equals. Also Jackendoff's (1990, 1997) conceptual semantics can be regarded as bridging formalist and cognitivist concerns, since it is concerned with how the mental representation of the world relates to the language system.

Here, I would like to make one further point on the relation between Generative Grammar and Cognitive Grammar and their interface. The main idea and concept of Cognitive Linguistics already existed in Generative Semantics; Cognitive Linguistics is not a new wave of modern linguistics but just an extension of Generative Semantics (cf. Lakoff 1987:582). As is well known, Generative Grammar battled with Generative Semantics about thirty years ago and won the fight, but its current research program, the Minimalist Program, has assumed the flavor of Generative Semantics, e.g., with regard to the structure of verbs; here, syntactic structure and semantic structure are not independent of each other, and they seem to constitute one system. Similarly, a sentence has its meaning formed step by step in the course of a derivation (cf. Epstein, Groat, Kawashima, and Kitahara 1998; see also Jackendoff 1972, who argues that information is sent to the semantic component every time a derivation reaches a cyclic node). Although Generative Grammar and Cognitive Linguistics are different from each other in several respects (cf., e.g., their concepts for 'language' and 'grammar'), they are consonant with each other on one point: both of them are, in a sense, allomorphs of Generative Semantics.[21]

As I have shown in this paper, although we can define or rephrase 'syntactic' phenomena by using syntactic notions which are based on phrase structures, we cannot really explain them as long as we resort to such structure-based notions. In order to really explain 'syntactic' problems, we need to appeal to non-syntactic notions (such as cognitive and functional ones, as have been used here; see Langacker 1987:53-54, 1990:18-19). As to the technical rephrasing of problems

inherent in Generative Grammar, that may be left to that system's methodology.

To the best of my knowledge, except for something like Merge, which is a syntactic procedure considered conceptually necessary in the Minimalist Program, most of the assumptions of Generative Grammar are either products of theory-internal discussions (i.e., outcomes of the interaction of assumptions) or retroactive calculations on the basis of observed facts (i.e., paraphrases of facts using syntactic notions). To put it differently, the assumptions put forward in Generative Grammar are either generalizations of facts or simply stipulations.

Besides Merge, displacement or movement is also a basic or salient property of natural language. To capture this property, devoted adherents of the Minimalist Program (or in general Chomskyan linguistics/syntax) postulate technical, abstract, and syntactic objects with strong formal features. But, as Chomsky (1995a:233) himself concedes, this is nothing but a technical statement of the facts, i.e., a bare stipulation. Although Chomsky (1999, 2000a) suggests that these strong features may be eliminated, as long as the research program follows the dogma of formalism, the properties in question will not receive a real explanation.[22]

This paper has critically reviewed the methodology of Generative Grammar. Another careful examination of the methodology of Generative Grammar is due to Nakai (1999). This author carefully looks at Generative Grammar in terms of scientific philosophy and considers at length how Generative Grammar can be assessed in the world of natural sciences. Nakai (1999:211) concludes that he cannot decide whether Generative Grammar is a natural science or just a game for the intelligent. (Incidentally, Derwing (1973:47) takes Generative Grammar to be a puzzle for the intelligent). In addition to Nakai, other works which critically review Generative Grammar's methodology comprise those by Gross (1979:879-880), Cook (1988:167-168), Linell (1979:73-74) and Carr (1990:47). Further, Botha (1978:31, 318) casts doubt on the psychological reality of theoretical objects assumed in Generative Grammar (see also Botha 1981:142). Nakai (1999) offers an assessment of all these their criticisms.[23]

Before concluding this article, I would like to quote the following passage from Ikeda (1995:35; originally in Japanese, the translation is mine):

> From the perspective of the completeness of a theory, as long as we explain natural phenomena by means of something visible, the level of completeness is low. This is because the visible thing itself is a natural phenomenon. Thus, explaining natural phenomena by visible things is not different from explaining natural phenomena by natural phenomena.

As is well known, under the recent research program of Generative Grammar, i.e., the Minimalist Program, the system is being constructed and made more sophisticated by referring to a feature, a syntactic object, which we cannot see at all (though we can see its effects). Comparing this current state of Generative Grammar with the above passage quoted from Ikeda, we can say that the Minimalist Program may be regarded as becoming 'complete' as a theory of grammar. But we must maintain that even if a theory is becoming 'complete', this does not necessarily entail that the theory is also powerful in its explanations.

Considering this, the following quotation from Law (2000:430) is suggestive:

> The fact that these constructions [= a variety of data, ranging from general phrasal movement and verb movement in French and English, to scrambling and topicalization in German and Japanese as well as multiple subject construction in Icelandic; Y.H.] require additional, theory-internal auxiliary assumptions, for example, that different features trigger different types of movement or that feature strength varies in certain ways, seems to indicate that the explanatory power of the Minimalist Program is compromised by its empirical scope.

Generative Grammar has attempted to construct a Universal Grammar for over forty years. At the present stage of the enterprise, the Minimalist Program, a Universal Grammar seems about to be revealed

to us. In this sense, Generative Grammar can be considered successful in that it has attained an explanatory adequacy of a very high level. Furthermore, as the Chomskyan syntacticians have shown, the Minimalist Program is also useful in capturing important properties of several constructions. Hence, the Minimalist Program similarly succeeds in satisfying descriptive adequacy at a very high level. However, the 'accounts' of linguistic phenomena which Generative Grammar has thus far offered are but descriptions or generalizations, not actual explanations of such phenomena. To provide a real explanation of such linguistic data, we must examine notions from other fields than syntax: in particular, the functional and cognitive aspects of human languages.

2-22-5-104, Ogawanishi-machi
Kodaira, Tokyo, 187-0035
Japan

Notes

1. For parsing or language processing, see also Kluender's (1998) Capacity Constrained Comprehension Theory. Kluender takes parsing to be constrained by working memory, and claims that 'syntactic' phenomena can be fully dealt with from the parsing perspective. Like Kluender, Gibson (1998) is also trying to analyze several linguistic phenomena in a unified way, based on his proposed Syntactic Prediction Locality Theory, from the viewpoint of parsing. As one can see from the following discussion, I go along with Kluender (1998) and Gibson (1998). As for parsing mechanisms, see also Chomsky (1965:14), Hawkins (1999), Phillips (1995), Frazier and

Clifton (1996), and references therein. Especially refer to Hawkins's (1994:77) Early Immediate Constituent hypothesis, which contends that node-counting is crucial to parsing.

2. Theme and rheme (or subject and predicate) correspond to topic and comment, respectively. Given this, the following quotation from Shi (2000:387) lends some support to my analysis of sentence parsing: 'It has often been observed that Chinese speakers use topic-comment constructions because they tend to first present the main thing they want to talk about and then organize their thoughts into a sentence to elaborate on the issue (Chen 1982, Lü 1986, cf. Chao 1968)'. This is in accordance with my intuition of Japanese.

3. But Phillips (1996, 1998), whose works build on the Minimalist Program, proposes that phrase structures are built from left to right on the basis of his Merge Right theory. His Merge Right theory is defined as follows (1996:27):

New items must be attached at the right edge of a structure.

Weinberg (1999), as well as Philips (1996, 1998), studies parsing in the framework of the Minimalist Program, and proposes, as Philips does, that a phrase structure is constructed from left to right. Her (1999:290) parsing algorithm is as follows:

A derivation proceeds left to right. At each point in the derivation, merge using the fewest operations needed to check a feature on the category about to be attached. If merger is not possible, try to insert a trace bound to some element within the current command path. If neither merger nor movement is licensed, spell out the command path. Repeat until all terminals are incorporated into the derivation.

She argues against Philips' approach to parsing by presenting several drawbacks of his analysis. As for her refutation, see Weinberg (1999:sec.11.5). Refer also to Richards (1999). The latter, as do Philips and myself, proposes that a phrase structure is constructed in a top-down manner.

4. That a phrase structure is actually not a tree but a root is often pointed out by several researchers. See, for example, Stokoe (1997:372).

5. I assume here that lexical insertion actually takes place immediately after one node branches into two nodes, that is, immediately after slots for lexical items are created; to put it another way, I do not consider here that lexical insertion occurs after all the frames of phrase structures have been constructed.

6. (11) is a restatement of Clark and Clark's (1977:548) generalization: 'Given information should appear before new information'. Gundel (1988:229) likewise formulates the information structure of natural languages as follows: 'State what is given before what is new in relation to it'. She calls this the Given Before New Principle.

7. Their Principle of Relevance I would be most related to the present discussion. The following is adapted from Sperber and Wilson (1995:260):

   Principle of Relevance I:

   Human cognition tends to be geared to the maximization of relevance.

8. Though the subject in (15a) is indefinite, its contents are clearly known to the speaker and the hearer; if the event that a picture is on sale were not present in the preceding context, we could not produce a question sentence such as (15a).

9. The following contrasts in (i)-(ii), which are borrowed from Hasegawa et al. (2000:198), indicate that extraction of NP from within subject NP is disallowed, but extraction of PP from within subject NP is allowed:

   (i) a.   *Which car$_i$ was [$_{NP}$ the hood of t$_i$] damaged?

   b.   Of which car$_i$ was [$_{NP}$ the hood t$_i$] damaged?

   (ii) a.   *the sonata which$_i$ [$_{IP}$ [$_{NP}$ [$_{IP}$ playing t$_i$]] is most difficult]

   b.   This is the kind of knife with which$_i$ [$_{IP}$ [$_{NP}$ [$_{IP}$ slicing soft breadt$_i$]] should be easiest].

   Based on data such as (i)-(ii), Hasegawa et al. argue that 'Subject Condition' is a constraint imposed on parsing, but not on syntax. See Hasegawa et al. (2000:198-200) for their analysis of the contrasts in question. I will leave for future research how my analysis of 'Subject Condition' can explain the contrast in (i)-(ii). But, since the present analysis of 'islands' is based on human cognition, Hasegawa et al.'s parsing-based analysis of the contrasts at issue is compatible with the present grammar system.

10. Persian, like Japanese, has no over *wh*-movement and behaves in the same way as does Japanese, as the following contrast indicates ((22a,b) are adapted from Karimi 1999:126):

    (22)   a. Kimea diruz [NP ye she'r az ki]   xund
             Kimea yesterday a   poem by who read
             'Who did Kimea read a poem by?'

    b. *Kimea [NP in   she'r az ki]   ro xund
             Kimea   this poem by who râ read
             '*Who did Kimea read this poem by?'

    râ in (22b) stands for the specificity marker. The data in (21)-(22) support the present functional analysis of the Definiteness effect.

11. But see Miyamoto (2000) for some problems with Erteschik-Shir's f-structure, in particular her (2000:204) Subject Constraint and her (2000:206) Principle A, both of which play an important role in her f-structure.

12. In Chomsky (1995a:332) it is suggested that adjuncts can appear in the complement of V. But in this case, too, the following question immediately arises: why doesn't an object NP in the complement of V constitute an island,

as an adjunct in that position does? Chomsky (1995a:sec. 4.7.5) further argues that both internal arguments and adjuncts, as well as external arguments, can occur in the specifier of V. But in this case, too, the following question will arise: why doesn't an internal argument in the specifier of V form an island, as an adjunct in that position does?

13. As the following examples show, time and location adjuncts can be preposed, whereas means and manner ones cannot. (i)-(iv) are all adapted from Kamio and Takami (1998:145-147).

    (i)   a. A man with blue eyes came yesterday.
          b. Yesterday, a man with blue eyes came.
    (ii) a. John drove a car with a sunroof in New York.
          b. In New York, John drove a car with a sunroof.
    (iii) a. A man with blue eyes came by bicycle.
          b. ?*By bicycle, a man with blue eyes came.
    (iv) a. John drove a car with a sunroof very carefully.
          b. ?*Very carefully, John drove a car with a sunroof.

    This indicates that with respect to category or type, the adjuncts under discussion are classified with time and location adjuncts, but not with means and manner ones.

14. There are many approaches to the interpretation of a *wh*-in-situ. As is shown directly, I assume here with Pesetsky (1987), who builds on insights of Baker (1970) and Heim (1982), that a *wh*-in-situ is interpreted through Absorption (or unselective binding), that is, without moving to its c-commanding *wh*-phrase. For other approaches, see Barss (2000), Boskovic (2000) and references therein. See also Cinque (1990) and Reinhart (1995). They consider *wh*-in-situ interpretation from different viewpoints than syntax; that is, from semantic and pragmatic viewpoints.

15. See Birner and Ward (1998), Chafe (1994), Fukuchi (1985), Gundel (1988), Gundel et al. (1993), Prince (1992), Vallduví (1990), and others for the close correlation between the information status of each word and its alignment in a sentence and its influence on choice of proper construction.

16. Aoun et al. (1987:560) observe (i) that *why* in (ia) can be associated with *say* and *sank*, whereas *why* in (ib) can only be associated with *say*; (ii) that *why* in (ic) can be associated with *think*, *said*, and *sank*, but *why* in (id) can only be associated with *think*; and (iii) that *why* in (ie) can be associated with *think* and *said*, but not with *sank*.

    (i)   a. Why did Fay say the boat sank?
          b. Why did Fay say that the boat sank?
          c. Why does Ray think Fay said the boat sank?
          d. Why does Ray think that Fay said the boat sank?

e.   Why does Ray think Fay said that the boat sank?

Based on this observation, Aoun et al. (1987:560) say that '[i]n short, *why* cannot be interpreted as originating in or moving through a clause with an overt complementizer. In effect, *why* shows *that*-trace effects, as our model predicts'. I will leave for future research why a *wh*-phrase *why* can be associated with a predicate of the embedded *wh*-interrogative, as illustrated by (35), but not with one of the embedded *that*-clause, as shown in (i).

17. See Chomsky (1995a:195ff). Chomsky (2000a), however, assumes that Case is a formal feature not innately given, but to be assigned through a derivation. Thus, he takes NPs not to bear any Case feature when they are taken from Numeration (or Lexical Array). This means that Chomsky (2000a) virtually does not adopt the Lexicalist Hypothesis. See also Lasnik, Depiante and Stepanov (2000), who propose a hybrid theory of English verbal morphology that includes elements of both Chomsky (1957) and Chomsky (1995a).

18. In passing, strictly speaking, what I have so far referred to as the PF component is not what is usually called 'PF' (an abbreviation for Phonetic Form), but the 'Punctuational Form' component, where the punctuation marks '.' and '?' are processed.

19. My proposed $SP_{[Dec/Q]}$ might constitute one class with Chomsky's (1957:65) transformations TA (a rule introducing a formal feature of contrast to phrase structure and by which the relevant sentence is construed as an affirmative sentence) and TNOT (a rule by which a negative element *not* is introduced to phrase structure), or Pollock's (1989:421, fn. 51) AssP (Assertion Phrase) and NegP (AssP corresponding to Chomsky's TA and NegP to his TNOT), or Laka's (1990:88; 1994) ΣP, which incorporates AffP (Affirmation Phrase, corresponding to Pollock's AssP) and NegP.

20. According to Langacker (1987:28-30, 48-49), Generative Grammar's approach to the language faculty, if viewed from the perspective of Cognitive Linguistics' paradigm, falls into the following two fallacies: rule/list fallacy and absolute predictability fallacy. He argues as follows:

   (i) '[I]t is perfectly plausible that the two [i.e. rules (viz. general statements) and lists (viz. particular statements)] might sometimes coexist [in grammar]'. (1987:29)

   (ii) 'Expectations of absolute predictability are sometimes unreasonable for natural language and commonly lead to erroneous conclusions, dubious claims, or conceptual confusion. We must scale our expectations down to a level of predictability that is appropriate and realistic for the subject matter'. (1987:48)

21. The question of how human cognition-based functional approach to the language faculty has relevance to Generative Semantics will be left for future research.

22. Contrary to Chomsky, Ndayiragije (2000) argues that in the system of the Minimalist Program, feature strength needs to be taken into consideration, and further that the operation Attract F cannot be reduced to Agree, thus refuting Chomsky's proposal.

23. In addition to Nakai (1999), Lappin, Levine, and Johnson (2000a) have carefully scrutinized the methodology of Generative Grammar, and criticize it as follows:

    (43) 'What is altogether mysterious from a purely scientific point of view is the rapidity with which a substantial number of investigators, who had significant research commitments in the Government-Binding framework, have abandoned that framework and much of its conceptual inventory, virtually overnight. In its place they have adopted an approach which, as far as we can tell, is in no way superior with respect to either predictive capabilities nor explanatory power'. (2000a:667)

    (44) '[I]n real science, arguments from authority do not, in general, determine the direction in which a field develops'. (2000a:669)

    (45) '[I]t is particularly surprising that so many linguists have chosen to adopt its [= the Minimalist Program's] assumptions without demanding additional clarification or empirical motivation. We know of no serious scientific discipline where theoretical paradigms are granted large scale acceptance in such a cavalier and uncritical manner'. (2000a:670)

    And they conclude their paper as follows:

    (46) 'We seem forced to the conclusion that a not insignificant part of our field is labouring under the manufacture of consent'. (2000a:670)

    Counter-criticisms have been offered by Holmberg (2000), Reuland (2000), Roberts (2000), Piattelli-Palmarini (2000), and Uriagereka (2000). In response, Lappin, Levine, and Johnson (2000b) have argued for their original position.

References

Ackerman, Farrell. 2000. Some Observations on the Typology of Linguistic Theories. Journal of Linguistics 36.365-381.

Aoun, Joseph & Li Audrey Yen-Hui. 1993. *Wh*-elements in Situ: Syntax or LF? Linguistic Inquiry 24.199-238.

Aoun, Joseph, Norbert Hornstein, David Lightfoot & Amy Weinberg. 1987. Two Types of Locality. Linguistic Inquiry 18.537-577.

Ariel, Mira. 1990. Accessing Noun-Phrase Antecedents. London: Routledge.

Arnold, Jennifer E. 1998. Reference Form and Discourse Patterns. Doctoral dissertation, Stanford University, Stanford, Calif.

Baker, Charles. 1970. Notes on the Description of English Questions: The Roles of an Abstract Question Morpheme. Foundations of Language 6.197-219.

Barss, Andrew. 2000. Minimalism and Asymmetric *Wh*-interpretation. In: Roger Martin, David Michaels & Juan Uriagereka (eds.), Step by Step: Essays on Minimalism in Honor of Howard Lasnik. Cambridge, Mass.: MIT Press, 31-52.

Birner, Betty J. & Gregory Ward. 1998. Information Status and Noncanonical Word Order in English. Amsterdam & Philadelphia: John Benjamins.

Boskovic, Zeljko. 2000. Sometimes in [Spec, CP], Sometimes in Situ. In: Roger Martin, David Michaels & Juan Uriagereka (eds.), Step by Step: Essays on Minimalism in Honor of Howard Lasnik. Cambridge, Mass.: MIT Press, 53-87.

Botha, Rudolf P. 1978. External Evidence in the Validation of Mentalistic Theories: A Chomskyan Paradox. Lingua 48.299-328.

Botha, Rudolf P. 1981. The Conduct of Linguistic Inquiry: A Systematic Introduction to the Methodology of Generative Grammar. The Hague: Mouton Publishers.

Bowers, John. 1988. Extended X-bar Theory, the ECP, and the Left Branch Condition. Proceedings of the West Coast Conference on Formal Linguistics 7. 47-62.

Bresnan, Joan. 1972. Theory of Complementation in English Syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Carr, Philip. 1990. Linguistic Realities: An Autonomist Metatheory for the Generative Enterprise. Cambridge: Cambridge University Press.

Chafe, Wallace L. 1994. Discourse, Consciousness, and Time. Chicago: University of Chicago Press.

Chao, Yuen-ren. 1968. A Grammar of Spoken Chinese. Berkeley: University of California Press.

Chen, Jianmin. 1982. Hanyu kouyuli de zhuijia xianxiang [The Afterthought in Spoken Chinese]. Yufa yanjiu yu tansuo [Studies and Explorations in Grammar] 1.117-132.

Cheng, Lisa. 1991. On the Typology of *Wh*-questions. Doctoral dissertation, MIT, Cambridge, Mass.

Chomsky, Noam. 1957. Syntactic Structures. The Hague: Mouton.

Chomsky, Noam. 1965. Aspects of the Theory of Syntax. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1973. Conditions on Transformations. In: Steven Anderson & Paul Kiparsky (eds.), A Festschrift for Morris Halle. New York: Holt, Rinehart, and Winston, 232-286.

Chomsky, Noam. 1975. Reflections on Language. New York: Pantheon.

Chomsky, Noam. 1977. On *Wh*-movement. In: Peter Culicover, Tom Wasow & Adrian Akmajian (eds.), Formal Syntax. New York: Academic Press, 71-155.

Chomsky, Noam. 1986a. Knowledge of Language: Its Nature, Origin, and Use. New York: Praeger.

Chomsky, Noam. 1986b. Barriers. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1992a. Explaining Language Use. Philosophical Topics 20.205-231.

Chomsky, Noam. 1992b. Language and Interpretation: Philosophical Reflections and Empirical Inquiry. In: J. Earman (ed.), Inference, Explanation and Other Philosophical Frustrations. Berkeley: University of California Press.

Chomsky, Noam. 1995a. The Minimalist Program. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1995b. Language and Nature. Mind 104.1-61.

Chomsky, Noam. 1999. Derivation by Phase. MIT Occasional Papers in Linguistics 18. Cambridge, Mass.: MIT Working Papers in Linguistics.

Chomsky, Noam. 2000a. Minimalist Inquiries: The Framework. In: Roger Martin, David Michaels & Juan Uriagereka (eds.), Step by Step: Essays on Minimalism in Honor of Howard Lasnik. Cambridge, Mass.: MIT Press. 89-155.

Chomsky, Noam. 2000b. New Horizons in the Study of Language and Mind. Cambridge: Cambridge University Press.

Chomsky, Noam. 2002. On Nature and Language. Cambridge: Cambridge University Press.

Chomsky, Noam & Howard Lasnik. 1977. Filters and Control. Linguistic Inquiry 8. 425-504.

Cinque, Guglielmo. 1990. Types of A'-Dependencies. Cambridge, Mass.: MIT Press.

Clark, Herbert H. & Eve V. Clark. 1977. Psychology and Language: An Introduction to Psycholinguistics. New York: Harcourt Brace Jovanovich.

Clark, Herbert H. & Susan E. Haviland. 1977. Comprehension and the Given-New Contrast. In: Roy O. Freedle (ed.), Discourse Production and Comprehension. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1-40.

Clark, Herbert H. & Thomas Wasow. 1998. Repeating Words in Spontaneous Speech. Cognitive Psychology 27.201-242.

Cook, V.J. 1988. Chomsky's Universal Grammar: An Introduction. Oxford: Basil Blackwell.

Croft, William A. 1995. Autonomy and Functionalist Linguistics. Language 71.490-532.

Denham, Kristin. 2000. Optional *Wh*-movement in Babine-Witsuwit'en. Natural Language and Linguistic Theory 18.199-251.

Derwing, Bruce L. 1973. Transformational Grammar as a Theory of Language Acquisition: A Study in the Empirical, Conceptual and Methodological Foundations of Contemporary Linguistic Theory. London: Cambridge University Press.

Diesing, Molly. 1992. Indefinites. Cambridge, Mass.: MIT Press.

Enç, Murvet. 1991. The Semantics of Specificity. Linguistic Inquiry 22.1-26.

Endo, Yoshio. 1995. Extraction, Negation and Quantification. In: Haraguchi Shosuke & Michio Funaki (eds.), Minimalism and Linguistic Theory. Tokyo: Hitsuji Shobo, 53-66.

Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima & Hisatsugu Kitahara. 1998. A Derivational Approach to Syntactic Relations. New York and Oxford: Oxford University Press.

Erteschik-Shir, Nomi. 1973. On the Nature of Island Constraints. Doctoral dissertation, MIT, Cambridge, Mass.

Erteschik-Shir, Nomi. 1997. The Dynamics of Focus Structure. Cambridge: Cambridge University Press.

Fiengo, Robert & James Higginbotham. 1981. Opacity in NP. Linguistic Analysis 7. 395-422.

Frazier, Lynn & Charles Clifton. 1996. Construal. Cambridge, Mass.: MIT Press.

Frazier, Lynn & Rayner Keith. 1982. Making and Correcting Errors During Sentence Comprehension: Eye Movements in the Analysis of Structurally Ambiguous Sentences. Cognitive Psychology 14.178-210

Fukuchi, Hajime. 1985. Danwa-no Kouzou [Structure of Discourse]. Tokyo. Taishukan.

Fukui, Naoki. 1997. Attract and the A-over-A Principle. UCI Working Papers in Linguistics 3.51-67.

Gibson, Edward. 1998. Linguistic Complexity: Locality of Syntactic Dependencies. Cognition 68.1-76.

Givón, Talmy. 1983. Topic Continuity in Discourse. Quantitative Cross-language Studies, TSL 3. Amsterdam: Benjamins.

Greenberg, Joseph H. 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In: Joseph H. Greenberg (ed.), Universals of Language. Cambridge, Mass.: MIT Press, 58-90.

Grice, Paul H. 1975. Logic and Conversation. In: Peter Cole & Jerrold Morgan (eds.), Speech Acts. New York: Academic Press, 41-58.

Gross, Maurice. 1979. On the Failure of Generative Grammar. Language 55.859-885.

Gundel, Jeanette K. 1988. Universals of Topic-Comment Structure. In: Michael Hammond, Edith A. Moravcsik & Jessica R. Wirth (eds.), Studies in Syntactic Typology. Amsterdam: Benjamins, 209-239.

Gundel, Jeanette K., Nancy Hedberg & Roger Zacharski. 1993. Cognitive Status and the Form of Referring Expressions. Language 69.274-307.

Halliday, Michael A.K. 1967-1968. Notes on Transitivity and Theme in English. Journal of Linguistics 3(1).37-81 (Part 1); 3(2).199-244 (Part 2); 4(2).179-215 (Part 3).

Hasegawa, Kinsuke, Ryouji Kasai, Sachie Kajita, Hiroshi Hasegawa & Noriko Imanishi. 2000. Bun (I) [Sentence (I)]. Tokyo. Kenkyusha Shuppan.

Hawkins, John. 1994. A Performance Theory of Order and Constituency. Cambridge: Cambridge University Press.

Hawkins, John. 1999. Processing Complexity and Filler-gap Dependencies Across Grammars. Language 75.244-285.

Heim, Irene. 1982. The Semantics of Definite and Indefinite Noun Phrases. Doctoral dissertation, University of Massachusetts.

Higginbotham, James & Robert May. 1981. Questions, Quantifiers, and Crossing. The Linguistic Review 1.41-79.

Holmberg, Anders. 2000. Am I Unscientific? A Reply to Lappin, Levine, and Johnson. Natural Language and Linguistic Theory 18.837-842.

Hornstein, Norbert. 1995. Logical Form: From GB to Minimalism. Oxford: Blackwell.

Hosaka, Yasuhito. 1994. Gengo Ruikeiron Saitou [Language Typology Reconsidered]. Tokyo: Gengo 23(3).52-58.

Huang, James C.-T. 1982. Logical Relations in Chinese and the Theory of Grammar. Doctoral dissertation, MIT, Cambridge, Mass.

Ikeda, Kiyohiko. 1995. Kagakuwa Dokomade Ikunoka [Science, Quo Vadis?]. Tokyo: Chikuma Shobo.

Ishii, Toru. 1997. An Asymmetry in the Composition of Phrase Structure and Its Consequences. Doctoral dissertation, University of California, Irvine.

Jackendoff, Ray S. 1972. Semantic Interpretation in Generative Grammar. Cambridge, Mass.: MIT Press.

Jackendoff, Ray S. 1990. Semantic Structures. Cambridge, Mass.: MIT Press.

Jackendoff, Ray S. 1997. The Architecture of the Language Faculty. Cambridge, Mass.: MIT Press.

Just, Marcel Adam & Patricia A. Carpenter. 1980. A Theory of Reading: From Eye Fixations to Comprehension. Psychological Review 87.329-354.

Kamio, Akio & Ken-ichi Takami. 1998. Danwa-to Jouhoukouzou [Discourse and Information Structures]. Tokyo: Kenkyusha Shuppan.

Karimi, Simin. 1989. Aspects of Persian Syntax, Specificity, and the Theory of Grammar. Doctoral dissertation, University of Washington.

Katz, Jerrold & Paul Postal. 1964. An Integrated Theory of Linguistic Descriptions. Cambridge, Mass.: MIT Press.

Kimball, Jack. 1973. Seven Principles of Surface Structure Parsing in Natural Language. Cognition 2.15-47.

Kluender, Robert. 1998. On the Distinction between Strong and Weak Islands: A Processing Perspective. In: Peter Culicover & Louise McNally (eds.), The Limits of Syntax: Syntax and Semantics 29. New York: Academic Press. 241-279.

Koizumi, Masatoshi. 1993. A Split VP Hypothesis. Unpublished manuscript. MIT, Cambridge, Mass.

Koizumi, Masatoshi. 1995. Phrase Structure in Minimalist Syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Kuno, Susumu. 1978. Danwa no Bunpoo [Grammar of Discourse]. Tokyo: Taishukan.

Laka, Itziar. 1990. Negation in Syntax: On the Nature of Functional Categories and Projections. Doctoral dissertation, MIT, Cambridge, Mass.

Lakoff, George. 1987. Women, Fire, and Dangerous Things. Chicago: The University of Chicago Press.

Langacker, Ronald W. 1987. Foundations of Cognitive Grammar. Vol. 1. Stanford: Stanford University Press.

Langacker, Ronald W. 1990. Concept, Image, and Symbol. Berlin and New York: Walter de Gruyter.

Lappin, Shalom, Robert D. Levine & David E. Johnson. 2000a. The Structure of Unscientific Revolutions. Natural Language and Linguistic Theory 18.665-671.

Lappin, Shalom, Robert D. Levine & David E. Johnson. 2000b. The Revolution Confused: A Response to Our Critics. Natural Language and Linguistic Theory 18.873-890.

Lasnik, Howard & Mamoru Saito. 1993. Move-α. Cambridge, Mass.: MIT Press.

Lasnik, Howard & Juan Uriagereka. 1988. A Course in GB Syntax: Lectures on Binding and Empty Categories. Cambridge, Mass.: MIT Press.

Lasnik, Howard, Marcela Depiante & Arthur Stepanov. 2000. Syntactic Structures Revisited: Contemporary Lectures on Classic Transformational Theory. Cambridge, Mass.: MIT Press.

Law, Paul. 2000. Review of Hisatsugu Kitahara. 1997. Elementary Operations and Optimal Derivations. Journal of Linguistics 36.425-431. Cambridge, Mass.: MIT Press.

Linell, Per. 1979. Psychological Reality in Phonology: A Theoretical Study. London: Cambridge University Press.

Lü, Shuxiang. 1986. Zhuwei weiyuju juli [Instances of Subject-Predicate Phrase as Predicate]. Zhongguo Yuwen [Chinese Linguistics] 5.334-340.

MacDonald, Mary Ellen, Neil Pearlmutter & Mark Seidenberg. 1994. The Lexical Nature of Syntactic Ambiguity Resolution. Psychological Review 101.678-703.

Mahajan, Anoop Kumar. 1992. The Specificity Condition and the CED. Linguistic Inquiry 23.510-516.

Martin, Roger & Juan Uriagereka. 2000. Introduction: Some Possible Foundations of the Minimalist Program. In: Roger Martin, David Michaels & Juan Uriagereka (eds.), Step by Step: Essays on Minimalism in Honor of Howard Lasnik. Cambridge, Mass.: MIT Press, 1-29.

Miyamoto, Yoichi. 2000. On the Stage- and Individual-level Distinction of Predicates. English Linguistics 17.134-159.

Nagao, Makoto, Hiroshi Nakagawa, Yuji Matsumoto, Koichi Hasida & John Bateman. 1999. Gengono Kagaku 8 (Linguistic Sciences 8). Tokyo: Iwanami Shoten.

Nakai, Satoru. 1999. Gengogaku-wa Shizenkagaku ka [Is Linguistics a Natural Science?]. Kyoto: Showado.

Ndayiragije, Juvénal. 2000. Strengthening PF. Linguistic Inquiry 31.485-512.

Newmeyer, Frederick J. 1998. Language Form and Language Function. Cambridge, Mass.: MIT Press.

Ochs, Elinor. 1979. Planned and Unplanned Discourse. In: Talmy Givón (ed.), Syntax and Semantics 12: Discourse and Syntax. New York: Academic Press, Inc., 51-80.

Oehrle, Richard. 1974. Some Remarks on the Painting of Rembrandt. Chicago Linguistic Society 10.

Pesetsky, David. 1987. *Wh*-in-situ: Movement and Unselective Binding. In: E. Reuland & A. ter Meulen (eds.), The Representation of (In)definiteness. Cambridge, Mass.: MIT Press, 98-129.

Phillips, Colin. 1995. Right Association in Parsing and Grammar. MIT Working Papers in Linguistics 26.37-93.

Phillips, Colin. 1996. Order and Structure. Doctoral dissertation, MIT, Cambridge, Mass.

Phillips, Colin. 1998. Linear Order and Constituency. Unpublished manuscript, University of Delaware.

Piattelli-Palmarini, Massimo. 2000. The Metric of Open-Mindedness. Natural Language and Linguistic Theory 18.859-862.

Pollock, Jean-Yves. 1989. Verb Movement, UG and the Structure of IP. Linguistic Inquiry 20.365-424.

Postal, Paul M. 1998. Three Investigations of Extraction. Cambridge, Mass.: MIT Press.

Prince, Ellen F. 1992. The ZPG Letter: Subjects, Definiteness, and Information Status. In: Sandra Thompson & William Mann (eds.), Discourse Description: Diverse Analyses of a Fundraising Text. Amsterdam: John Benjamins, 295-325.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. A Comprehensive Grammar of the English Language. London: Longman.

Radford, Andrew. 1988. Transformational Grammar. Cambridge: Cambridge University Press.

Reinhart, Tanya. 1995. Interface Strategies. (OTS Working Papers of Theoretical Linguistics 95-002.) Utrecht Institute of Linguistics OTS, Utrecht University.

Reuland, Eric. 2000. Revolution, Discovery, and an Elementary Principle of Logic. Natural Language and Linguistic Theory 18.843-848.

Richards III, Norvin W. 1999. Directionality of Tree Constructions. Unpublished manuscript. Kanda University and MIT.

Rizzi, Luigi. 1990. Relativized Minimality. Cambridge, Mass.: MIT Press.

Roberts, Ian. 2000. Caricaturing Dissent. Natural Language and Linguistic Theory 18.849-857.

Shi, Dingxu. 2000. Topic and Topic-comment Constructions in Mandarin Chinese. Language 76.383-408.

Sperber, Dan & Deirdre Wilson. 1995. Relevance: Communication and Cognition. Oxford: Blackwell.

Stemmer, Bernard. 1999. An On-line Interview with Noam Chomsky: On the Nature of Pragmatics and Related Issues. Brain and Language 68.393-401.

Stokoe, William C. 1997. Language Is Natural, but Not Automatic. Semiotica 113.369-383.

Takahashi, Daiko. 1994. Minimality of Movement. Doctoral dissertation, University of Connecticut.

Takami, Ken-ichi. 1995. Kinoutekikoubunron-ni yoru Nichieigohikaku [A Functional Syntactic Analysis of Japanese-English Comparison]. Tokyo: Kuroshio Shuppan.

Takami, Ken-ichi. 1997. Kinoutekitougoron [A Functional Syntax]. Tokyo: Kuroshio Shuppan.

Tomasello, Michael (ed.). 1998. The New Psychology of Language: Cognitive and Functional Approaches to Language Structure. Mahwah, N.J.: Lawrence Erlbaum Associates Publishers.

Torrego, Esther. 1998. The Dependencies of Objects. Cambridge, Mass.: MIT Press.

Toyoshima, Takashi. 1997. Derivational CED: A Consequence of the Bottom-up Parallel-Process of Merge and Attract. Proceedings of the Fifteenth West Coast Conference on Formal Linguistics: 505-519. Stanford, Calif.: CSLI Publications.

Tunstall, Susanne. 1996. The Processing of Definite Phrases with Extraposed Modifiers. University of Massachusetts Occasional Papers in Linguistics 19 (Linguistics in the Laboratory). 303-327.

Uriagereka, Juan. 1999a. Minimal Restrictions on Basque Movements. Natural Language and Linguistic Theory 17.403-444.

Uriagereka, Juan. 1999b. Multiple Spell-Out. In: Samuel Epstein & Norbert Hornstein (eds.), Working Minimalism. Cambridge, Mass.: MIT Press, 251-282.

Uriagereka, Juan. 2000. On the Emptiness of 'Design' Polemics. Natural Language and Linguistic Theory 18.863-871.

Vallduví, Enric. 1990. The Information Component. Doctoral dissertation, University of Pennsylvania, Philadelphia, Pa.

Weinberg, Amy. 1999. A Minimalist Theory of Human Sentence Processing. In: Samuel David Epstein & Norbert Hornstein (eds.), Working Minimalism. Cambridge, Mass.: MIT Press, 283-315.