Constraints on Discontinuity

Barbara Brunson
Dept of Linguistics
University of Toronto

1. Introduction.

I will accept without argument that a child is innately pre-equipped with a set of defining principles that constrains the possible class of human languages. These defining principles comprise the innate knowledge of language and its structures that a child possesses, and it is the primary task of a theory of language acquisition to determine the extent and nature of these principles and how data might be used as triggers on choices (where the grammar allows for parametric options). Such an approach is in keeping with the tradition of parametric Universal Grammar as in the Government-Binding (GB) framework, and, as is the practice in this framework, I will refer to those principles that the child brings to bear on the language acquisition task as Universal Grammar (UG).

A particular concern in the formulation of such a theory of UG is ensuring that the proposed theory is indeed universal and not founded on language specific biases. GB theory has been developed largely from analyses of widely studied European languages. As other languages have come under investigation, it has become clear that the early analyses proposed as universal in GB are inadequate for a general theory of linguistic universals (Saito, 1985; Hale, 1983; inter alia). Such inadequacies have been particularly clear when dealing with the varied ways that languages identify thematic arguments.

In this paper, I review claims made in the literature about thematic argument structure. This includes claims made by linguists in extending the theory to cover some of these rather varied language types, as well as claims made in
acquisition theory about the nature of argument structure. I then present a preliminary sketch of a theory of argument structure that is, I hope, somewhat more neutral with respect to language specific expectations. As a particular case in point, I consider a variety of languages with discontinuous arguments and propose a constraint that accounts for certain asymmetries found in the data. The consequences of this constraint for a theory of acquisition and learnability are discussed.

2. Thematic Argument Structure.

2.1 Word Order. In general, word order involves any statements required in the grammar about the relative positions of lexemes and phrasal constituents. These statements generally take the form of headedness statements (i.e. left or right headed), subcategorization, and directionality in case and theta role assignment. It has even been claimed that the phrase structure component is redundant in the grammar if we have an adequate theory of X-bar syntax together with appropriate statements from which word order can be derived (Stowell, 1981).

In the context of thematic argument structure, the primary concern of word order is the recoverability of grammatical relations since it is in terms of these, together with lexical specifications for predicates, that thematic roles are defined. The popular notion that languages can be characterized according to the relative position of verb, subject and object (e.g. VSO, SOV, SVO, etc.) is indicative of the importance of constituent order with respect to grammatical relations.

If we take word order as an independent principle of UG, then it is necessary that a child initially hypothesize a fixed order. The change to free order would be the result of positive evidence; that is, observation would be made of
the various permutations of a particular sentence type. If a child assumed free word order, only negative evidence (i.e. the absence of many of the possible permutations) could result in the modification of that assumption. However, if word order facts are derivative of other principles of UG, then this notion of a default fixed word order need not be the case.

2.2. Continuity of Arguments. It is widely assumed that in languages like English semantic units are continuous in the syntax. Many believe that there is no syntax, maintaining instead that syntactic phrases are just these semantic units (Hopper & Thompson, 1984, Schank & Abelson, 1977). Most parsing algorithms use syntactic contiguity as a necessary prerequisite for semantic composition. As an example, in the problem of determining the possible structural sites for attaching PPs, the candidate PPs for attaching into an NP are generally only the PPs that occur structurally adjacent (to the right in English) to the NP.¹

Wexler and Culicover (1980) maintain that regardless of constituent order, clauses and noun phrases must be continuous in the syntax (at least at deep structure). Any discontinuity must be the result of a transformation.

I maintain that discontinuity is not so unusual and continuity requirements are largely derivative of "uniqueness" properties of other syntactic principles, for example the uniqueness of a structurally cased position, the uniqueness of a phrasal head, the uniqueness of an adjacent position, etc. Syntactically discontinuous semantic expressions can be base-generated provided that no other principles of syntax are violated and provided that there be some way in the syntax to identify the joint semantic nature of the two expressions.

2.3. Configurationality. In recent years, linguists have focussed attention on

¹ Few parsing algorithms take into account such derived dislocations as those occurring from PP postposing.
languages that exhibit a great variety of argument structure properties. This has led to a typological characterization of configuration and non-configuration languages, where configurational languages typically do not allow great freedom of word order or discontinuity. The term non-configuration was originally used by Ken Hale in an attempt to characterize languages that allow free word order and discontinuity (Hale, 1981, inter alia). The term derives from an attempt to capture the differences through the base component — contrasting an X-bar type schema with a W base component. Clearly if the only phrase structure rule is of the form \( S \to W \), the language in question would have no hierarchical configuration intermediate between the sentence and the word — hence, non-configurational. Such a strong statement of configurationality is no longer maintained as putative non-configurational languages have been shown to exhibit configurational properties, albeit often with minimal configuration. Nevertheless, the general bipartite division of languages into two typological classes is still a common practice.

Under closer scrutiny it is evident that configurationality is a local property of structures and not a property of a language. There are many instances of typical "non-configurational" phenomena occurring in prototypical "configurational" languages. The examination of one such phenomenon will be central in later parts of this paper.

In studies of non-configurationality, Warlpiri has traditionally been pointed to as representative of the sorts of differences that languages of this type have compared to, say, English.

In Warlpiri virtually all permutations of the content words of a sentence count, semantically, as repetitions of one another. Noun phrases can be broken up into discontinuous pieces that may be arbitrarily far apart provided that
they all agree in case marking. Embedded clauses can also be discontinuous and interleaved with other clauses, including the matrix clause. In such cases a suffixed complementizer on each piece of the discontinuous subordinate clause indicates which pieces should, semantically, be put together.

The principle exception to the free word order of Warlpiri is the auxiliary which generally occurs as the second syntactic constituent of the sentence.

3. An Analysis of Thematic Identification.

In order that thematic interpretation be coherent and consistent with the sentence in question, the relevant thematic roles must be recoverable from the surface syntactic form of the sentence.

Within the syntax there are only three general types of information available:

(1) Features on nodes, or contents of leaf nodes (i.e. lexical types/tokens)
(2) Relative linear order
(3) Hierarchical constituency

In an endocentric, obligatory head syntax with an upper limit on $X^{\text{MAX}}$, (1) and (3), above, go hand in hand. The introduction of higher and higher levels of configuration and constituency generally involves the introduction of different phrasal projections, which in turn involves the introduction of heads for these phrases (i.e. lexically identified items).

In recent syntactic research, phrase structure has come to look more and more like exclusively binary branching tree structures. In such analyses, the distinction between linear order and hierarchical constituency collapses as cascading structures reduce template-type ordering statements to constituent subcategorization. For example, we can equivalently express the template ordered flat structure in A by a hierarchical structure as in B with appropriate
subcategorization statements.

A)

\[
\text{phrase} \quad \text{W} \quad \text{X} \quad \text{Y} \quad \text{Z}
\]

\text{template ordering: } [ \text{W} < \text{X} < \text{Y} < \text{Z} ]

B)

\[
\text{ZP} \quad \text{YP} \quad \text{XP} \quad \text{WP} \quad \text{X} \quad \text{Y} \quad \text{Z}
\]

where

i) \( X \) is subcategorized for a complement of type \( \text{WP} \),

ii) \( Y \) is subcategorized for a complement of type \( \text{XP} \),

and

iii) \( Z \) is subcategorized for a complement of type \( \text{YP} \).
The headedness in B is irrelevant, what is important is that a template type ordering can be equivalently expressed by more local orderings of heads and complements in a hierarchical structure.

It is entirely reasonable that, during the acquisition of syntax, hierarchical constituency is derivative of (1) and (2) together with an internalized constraint on possible phrase structure (X-bar schema). That is, hierarchical configuration may be acquired from the triggering evidence of linear order facts on lexical types and tokens.

At any rate, semantic and thematic interpretation requires non-redundant information to be encoded in at least one of these ways in the syntax.

What should be apparent is that the exceptional nature of languages like Warlpiri derives from the use of only lexical items, to the exclusion of linear order or hierarchical constituency, in the identification of the thematic roles of arguments in a sentence. Arguments are lexically identified according to what role they play and, in the case of subordinate clauses, with respect to which clause they bear that role in. There is little evidence that would lead one to propose a configurational hierarchy of arguments in a Warlpiri main clause and there appears to be no underlying or default word order.

That Warlpiri deserves an exceptional status is primarily derivative of the expectations from European languages that we bring to bear on the study of Warlpiri grammatical relations. However, European languages are not devoid of the sort of non-configurational structure found in Warlpiri. In English, for example, virtually all oblique arguments are identified by means of lexical prepositions. With such oblique arguments we find free ordering, as illustrated in (1).

(1) a) I saw the man at 2 o'clock in the park.

   b) I saw the man in the park at 2 o'clock.
We also find a single type of thematic relation being borne by several discontinuous PPs. This is shown by the two temporals in (2).

(2) a) I saw the man at 2 o'clock in the park on Tuesday.
    b) I saw the man on Tuesday in the park at 2 o'clock.

Generally it is assumed that these facts about English PPs derive from the fact that each PP comes with its own θ-identifier and case-assigner — the preposition. Analogously, in Warlpiri each case-marked nominal phrase can be seen to have its own lexical θ-identifier and case-assigner — the case-marker itself. What this means for Warlpiri is that the verb cannot be a case-assigner since each argument is inherently case-assigned by virtue of the case-marking suffix.

From the point of view of an acquisition model, each syntactic principle requires that the child attend to some triggering context that is identifiable in the s-structure (under the child's current hypothesis of its nature). Since we've identified three distinct types of syntactic information, one of the tasks of the child is to determine in which of these ways the trigger is identified if the means of identification is not prespecified by the principle itself.

Consider for example the sanctioning of NPs for case. The child is presumably attending to NPs² and must determine in what identifiable way these NPs are case-sanctioned in the syntax since there are universal principles requiring NPs to be case sanctioned. In Warlpiri the task appears to be relatively straightforward — NPs are case-marked and hence lexically identified as case-sanctioned. It is by no means a trivial task to determine that this is sufficient for case theory. Many European languages have case marking but also require

---

²I am excluding the NP attached reading [NP the man in the park].
³I ignore the problem of interdependencies involving the acquisition of syntactic categories.
specific argument orders. Taking an overly simplified view that fixed order is the initial parameter setting, Warlpiri children must also conclude that order is free in their language and hence order of arguments is irrelevant to case theory. For English, if we set aside the problem of the ordering and interactions of the acquisition of hierarchical structure with case theory, and presuppose that hierarchical structure is already acquired, there is positive evidence that all three types of syntactic information are used. Linear order and hierarchical constituency are required for subject and object positions, and lexically identified elements are employed in prepositional phrases. A complete analysis of the acquisition of the English case system would be rather complex, as is the theory itself. The main point to be made here is that if we assume that some version of the case filter⁴ is innate, the child must be able to establish how it is satisfied on NPs, given s-structure. Since, in English there is evidence for a variety of types of case-sanctioning — including the lexically identified option, then this option must be generally available to the child. That Warlpiri chooses just this means of case-sanctioning is, then, an allowed option and not necessarily degenerate. That it appears to be a marked option may be accounted for by the fact that, although there are at least 3 types of syntactically available methods of case-sanctioning, Warlpiri places the entire burden of the task upon one of these. This could be stated as a preference to distribute tasks among possible subsystems if these subsystems are independently required in the grammar.

3.1. Thematic Composition. Having made some preliminary comments regarding argument structures, I now return to Wexler and Culicover’s point, namely, that clauses must be contiguous and that each argument within the

---

⁴ The Case Filter is a constraint requiring that all lexical NPs have abstract case locally assigned to them by some case assigning syntactic entity.
clause must be contiguous. This constraint is derivative of their Binary Principle and is required to constrain the number of possible mappings between s-structure and LF with respect to grammatical relations.

To illustrate their claim, assume that the grammar is infinitely recursive. For each grammatical relation $R$, given $n$ levels of embedded clause structure $(S_1, ..., S_n)$, then $R$ potentially "belongs to" any of $S_1, ..., S_n$. As $n \to \infty$, the possible number of mappings for each such $R$ also approaches $\infty$, hence there is no constraint on the number of mappings of syntax under thematic identification to LF, unless the grammar somehow restricts and localizes the candidate $S$s for each $R$. Wexler and Culicover erroneously maintain that contiguity is the only possible means of appropriately limiting such mappings. In Warlpiri another method is used. Although arguments of separate clauses can be interleaved in the syntax, this requires that each argument — except matrix clause arguments — be suffixed with the complementizer of the clause to which they belong. In this way, all arguments of dependent clauses are identified with their respective clauses by means of these suffixes, even though they need not be contiguous. The number of possible mappings is restricted by a lexical rather than a structural means.

Extending this to the level of the arguments themselves, if we have $A_1, ..., A_n$, arguments of a given clause, and $D$ a discontinuous piece of an argument, $D$ potentially belongs to any of $A_1, ..., A_n$. By successive fragmenting of arguments, we again have an unconstrained mapping in associating each such $D$. Again, in Warlpiri the candidates are not structurally localized by such constraints as adjacency/contiguity, but, rather, are restricted by lexical identification — the case suffixes must be equivalent in order for two discontinuous nominal phrases to be interpreted together. It should be clear that Wexler
and Culicover's constraint is too narrow. Structural locality is not the only resource that language has to restrict mappings of s-structure and LF. Lexical items themselves can serve the functional role of indices, thus giving the appropriate mapping even when structural locality isn't maintained.

From this we could expect that, in discontinuous structures, such lexical indexing would allow for totally unconstrained linking and composition of thematically united, but syntactically disjoint, entities. This proves not to be the case. In the remainder of this section I consider some asymmetries in discontinuous structures and propose an analysis to deal with such data.

3.2. Discontinuity Constraints. We have already seen that in English it is possible for two or more PPs to bear the same thematic relation in a single sentence. This is evident in sentences (3) and (4), below, where the two PPs both provide temporal or locative modification (respectively) to the event in the sentence.

(3) a) John saw Mary at noon on Tuesday.
    b) John saw Mary on Tuesday at noon.

(4) a) John saw Mary in the park near the tree.
    b) John saw Mary near the tree in the park.

The order of the PPs in post-verbal position appears to be irrelevant. The PPs may even be discontinuous as in (5).

(5) John saw Mary at noon in the park on Tuesday.

It seems that with such multiple modification we have an intersection relationship defined for semantic interpretation. For example, in (5), the event of John seeing Mary in the park has temporal modification — it happened at that time which is both 'at noon' and 'on Tuesday'. Generally intersection, in set theoretic terms, is a symmetric relation. I will use the notation \( X_\theta \) to be the
extension set of one of the arguments bearing the $\theta_i$ $\theta$-role to some predicate. With symmetry we have:

$$A_{\theta_i} \cap B_{\theta_i} \equiv B_{\theta_i} \cap A_{\theta_i}$$

(*ie.* the order of occurrence is irrelevant, as is confirmed in the data)

Since such multiple modifiers take on a unified role with respect to the thematic structure of the sentence, we can define a process of thematic composition schematically as follows:

$\theta$-structure: [ ..... $\theta_i$ ..... ]

```
(A \cap B)_{\theta_i}
```

s-structure:

```
A_{\theta_i} \ldots B_{\theta_i}
```

What this means is that, however the thematic structure is realized at LF, there would be a "slot" for each theta role. The mapping from s-structure to this thematic structure would use the $\theta$-indices (prepositions and case markers) to determine which arguments go together via the intersection relation and it is this composed semantic entity that fills the appropriately indexed $\theta$-"slot".

In configurational languages with arguments that are identified by structural means, such pairing of arguments is normally blocked by the uniqueness of such structural positions as subject and object.

**3.2.1. PP Preposing.** The symmetric relation discussed above breaks down in cases of PP preposing. Consider the following sentences in which one of the PPs have been preposed:

(6) On Tuesday, John saw Mary in the park at noon.
(7) *At noon, John saw Mary in the park on Tuesday. According to Reinhart (1981), PP preposing is a characteristic of "sentential PPs" but not VP internal PPs. Such a structural approach proves unsatisfactory as a means of allowing (8) but blocking (9).

(8) At noon, John saw Mary in the park.

(9) *At noon, John saw Mary on Tuesday.

Without providing a very complex statement of what constitutes a VP internal PP, one must abandon a purely structural account of these asymmetries.

What is common to these examples is that in cases of more than one syntactic PP bearing the same thematic relation, only the more general of them can prepose. Semantically, if the extension set of $A_{i}$ is a subset of the extension set of $B_{i}$, then $B_{i}$ can prepose but $A_{i}$ cannot.

Taking this thematic/semantic description of the asymmetry, let’s now consider other thematically discontinuous elements to determine if they illustrate similar sorts of asymmetries.

3.2.2. English Passive By-Phrases. Following Baker (1985), I assume that English passive is formed by head movement of a passive morpheme into the verb. This passive morpheme is realized as verbal morphology but syntactically is an argument with very general semantics. So, for example, with sentence (10),

(10) The boy was bitten.

the passive morphology on the verb semantically contains the general agent argument, hence we can infer "by someone/thing" as agent of this event.

We do get overt by-phrases with passives, so this would be a case of multiple syntactic entities bearing a single thematic role. We find sentences like (11)

(11) The boy was bitten by the dog.
where Baker’s analysis would require that the thematic role of agent be instantiated by the semantic intersection of the general agent (i.e. the passive morpheme) and the extension set of ‘the dog’. Since the passive morpheme provides no restriction, except possibly for features such as animacy that are redundantly associated with agents, the intersection set here would be identical with the extension set of ‘the dog’ as we intuitively would have predicted.

Since the passive by-phrase is necessarily more restrictive and specific than the general passive agent, given our observation about preposing asymmetry, we would expect that preposing it is not allowed with passive by-phrases. This prediction is confirmed.

(12) *By the dog, the boy was bitten.

(13) *By the committee, the proposal was rejected.

Notice that in the least marked order, by-phrases occur last after other PPs.

(14) The boy was bitten in the park by the dog.

Such evidence as (14) has led Baker to propose that the passive by-phrase is VP external; but then, according to Reinhart, we should expect preposing to be grammatical.

If we accept Baker’s treatment of the passive, then we find that the restriction against preposing the passive by-phrase patterns with the preposing asymmetries that we have already seen.

3.2.3. Noun Incorporation. If we accept that noun incorporation is syntactic movement (as in Baker, 1985), we find that asymmetries with doubling of the incorporated argument pattern exactly like the PP preposing asymmetries5.

Consider the following data:

West Greenlandic Eskimo (Sadock, 1980)

5I am currently following a suggestion from Diane Massam to consider evidence from clitic doubling.
(15) Tuttu-p neq-itor-punga
    reindeer-erg meat-eat-indic/1sS
    'I ate reindeer meat.'

Tuscarora (Williams, 1976 (from Baker, 1985))
(16) ne-hra-taskw-ahk-hwa? ha? tsi:r
du-3M-domestic animal-pickup-asp prt dog
    'He is a dog-catcher.'

Onondaga (Woodbury, 1975 (from Baker, 1985))
(17) hati-hnek-sets o-e:ta:k-i?
    3Mpl-liquid-gather pre-syrup-suf
    'They gather maple syrup.'

In these examples we see that the incorporated noun is the more general, while the more specific argument remains unincorporated and counts as the doubled argument. Sentences that result from the incorporation of the more specific argument are ungrammatical if the general argument is also present.

3.2.4. Warlpiri Discontinuity. I have argued elsewhere (Brunson, 1986) that the Warlpiri first position constituent (which occurs immediately before the auxiliary) is located as the result of syntactic movement from a post-auxiliary position. This position may be occupied by a single part of a discontinuous constituent and when it is, it is the more general part that occurs in this moved position. The more specific part, on the other hand, cannot move into the pre-auxiliary position (Mary Laughren, personal communication).

3.3. Extensional Subsetting. The observation that has been repeatedly evidenced in the facts above is that the semantically more specific part of a discontinuous entity is distributionally asymmetric with the semantically more general
part. Using the metaphor of syntactic movement, we find that the semantically more specific part of a discontinuous expression cannot be moved\(^6\). Under application of move-α, the moved element will be in a position to command and/or precede its trace and thematically more specific part.

In lexical phonology it is accepted to allow phonological interpretation and the application of phonological rules to apply cyclically on morphologically defined structures. Cyclicity here is with respect to the hierarchical configuration of the morphology. Suppose, as a preliminary account of the asymmetry we have been investigating, we allow thematic interpretation to take place cyclically defined on the levels of structural hierarchy of the syntactic structure. I propose that one piece of a thematically discontinuous expression is sufficient to instantiate that θ-role during the interpretation at any given level. If we do thematic interpretation cyclically on the syntactic structure we find that the asymmetry occurs when the thematically discontinuous pieces occur on separate levels of syntactic structure, as occurs when one piece is moved by move-α. Specifically, we have found that if a particular θ-role is instantiated at one level, any subsequent specification of this θ-role on another level must have the semantic function of providing more and more specific restrictions.

I will now discuss why this should be the case. In set theory, set building involves the specification of some general set by a general property and then forming subclasses of it by applying restrictions to this set. It is a considerably more difficult problem to take a condition that applies to a more restricted set and constructs a superset of it consistent with that condition. In constructing a set by subsetting, the relevant universe is just the set that you are taking a

---

\(^6\) This condition may need to be revised to refer to specific rules since wh-movement seems to behave somewhat differently. We find grammatical sentences such as *By whom was the child bitten*. The difference here may have to do with the non-referential nature of wh-phrases, but clearly further study is required before a rule specific constraint can
subset of. To construct a superset from a given set, the entire universal set must be taken into consideration. The latter is clearly a more formidable exercise. Of course, this is only a speculative proposal and subject to empirical testing as to whether the analogy of relative ease of set construction applies to such cognitive capacities as the language faculty.

An immediate result of the previous proposal is that thematic structures are instantiated and filled cyclically on the levels of syntactic structure. Across levels we can compose discontinuous arguments via extensional subsetting — within levels we can compose discontinuous arguments via extensional intersection. This proposal has far reaching consequences for theories of language processing since it provides a cyclic procedure for semantic and thematic interpretation. It also provides independent support for the type of mechanism proposed in phonology whereby phonological rules operate cyclically defined on morphological structure.

These results also collaborate well with theories of hierarchical knowledge representation where entities are divided into successive subclasses along various dimensions. Such a representation would readily facilitate the decision procedure required for extensional subsetting.

If a child only has access to positive evidence, then these mechanisms of cyclicity and the division of extensional subsetting from extensional intersection must be innate parts of the child’s knowledge, possibly belonging specifically to the interpretive component of language. Since there is other evidence for cyclicity on structure at the interface of linguistic components, cyclicity between syntax and thematic structure involves no extra overhead to the language learner. Extensional intersection is independently required to interpret certain
types of conjunctions, and extensional subsetting is involved in many types of restrictive modification. The particular configuration of the grammar with respect to when subsetting applies and when intersection applies is the extent to which the proposed account forces a new mechanism in UG.


In this paper I have addressed issues in thematic argument structure. I have considered the role of lexical identification in the syntax as both a trigger for acquisition of case theory, and as a restriction on the cardinality of possible mappings from s-structure to thematic structure given discontinuity in s-structure.

A constraint on discontinuity that is evident in numerous languages and constructs has been examined and a preliminary analysis proposed so as to minimizes the need to introduce new mechanisms into UG. This analysis accounts for asymmetries in the distribution of semantically general and specific parts of discontinuous expressions.
Bibliography


