MARK STEEDMAN

LFG AND PSYCHOLOGICAL EXPLANATION*††

The publication of *The Mental Representation of Grammatical Relations* (hereafter, MRGR)[1] is an important event for linguists and all concerned with the psychology of language. It provides the most definitive statement to date of Lexical-Functional Grammar (LFG), a theory of natural language grammar which is claimed to generalise across radically different language types, to be recursive, learnable, realistically parsable and producible, and to provide "a stronger basis than ever before for a psychologically realistic theory of grammar". The present paper begins by briefly reviewing the nature of the psychological problem, before outlining LFG as a theory of grammar in section 1.[2] Section 2 then considers the broader psychological claim.

Grammars of natural languages are psychologically problematic in at least two ways. The first is a major concern among linguists: natural language syntax is complicated by the inclusion of various startling discontinuous and fragmentary constituents, which occur for example in relative clauses and "reduced" coordinate sentences, Other less surprising constructions still tend to be characterised by a puzzling complexity in the relation between surface grammatical functions like subject and object and underlying semantic or thematic categories like agent and patient.

Natural languages are also characterised by a surprising degree of syntactic ambiguity. In particular, points of "local" syntactic ambiguity, where it is locally unclear which of two or more analyses should be followed, are widespread, a fact which has led psychologists and computational linguists to postulate processing regimes embodied in various algorithms and heuristic "strategies". The psychological implications of these two types of untoward complexity are serious, both for learning and processing.

The two problems of grammar and local ambiguity resolution are logically quite separate. The former is a (partial) specification of WHAT is computed in natural language comprehension, whereas the latter relates to HOW it is computed. Since it is obvious that the space of algorithms for computing the relevant class of functions is very much larger than that class itself, there is an equally obvious methodological priority across the two problems: whether we count ourselves as psychologists or linguists, until the problem of grammar has been solved, we are unlikely to make much headway with the other problems of processing, such as local

Linguistics and Philosophy 8 (1985) 359–385. 0165–0157/85.10

ambiguity resolution. It is therefore entirely reasonable that MRGR, which makes very strong claims to psychological explanation, should consist for the most part of linguistic studies, and devote only three of its thirteen chapters to explicitly psychological questions.

The logical and methodological priority of the study of grammar has not prevented a certain impatience on the part of the psychologists at the fact that the formal theory of language which has held sway for the last twenty-five years has made it exceptionally hard to see how its rules might be processed or learned. So great has the gulf between such grammars and psychological mechanisms appeared that one very influential school of psycholinguists has actually proposed that the fundamental distinction between grammar and processor and the priority of the former should be ignored in favour of a monolithic mechanism of heuristic "strategies", whose relation to the concept of grammar has been notably obscure (Fodor et al. 1974).

It is only fair to point out that, since this darkest hour, computer scientists with quite explicit allegiance to transformational grammar (TG) have shown that there are simple and elegant computational equivalents of certain transformational operations like Wh-movement (as in the ATN HOLD mechanism of Woods 1970, 1973), and that local ambiguity resolution techniques exist which have been claimed to allow respectable fragments of English to be parsed in linear space and time according to grammars closely related to the "extended" standard theory (Marcus, 1980). However, it has for some time seemed likely that alternative theories of grammar with much nicer properties from the psychologists' point of view are attainable. The main reason for thinking this has been the proliferation in the various versions of TG of constraints on rules.

No theory is fully satisfactory if it requires ad hoc constraints. That is, if a theory exploits a degree of freedom, say to move constituents, in order to capature a generalisation about a particular language, then the exploitation of that degree of freedom in that way constitutes a prediction that other constructions and other languages may exploit that same degree of freedom in all the other ways that it allows. To the extent that constraints on rules are required to express the fact that this freedom is *not* exploited by the languages of the world, the prediction fails. The theory can only be redeemed in one of two ways: either it must be shown that the constraints arise in a principled fashion – for example, from the semantics which the syntax reflects, or from the requirements of processing or learning the grammar – or the theory must be modified to make it less general. The constraints on movement that TG required did not appear to be readily explainable in terms of the demands of the semantics or any other

functional component. And despite early hopes that they might be explained in terms of processing (Kuno and Robinson, 1972; Woods, 1973; Bach, 1977; Fodor, 1978; Marcus, 1980) or learnability (Wexler and Culicover, 1980), many of them did not seem to be any more readily susceptible to this sort of explanation.

The most important single constraint was the observation that the majority of standard transformations mapped onto configurations that were arguably identical to configurations already generated directly in the base, an observation which was embodied in the Structure Preserving Constraint of Emonds (1970, 1976). This observation gave rise to the Base Generation Hypothesis pioneered by Brame (1976, 1978) and Bresnan (1976, 1978), according to which the transformations in question could be avoided entirely, constructions like passive and dative-shift being generated directly. Bresnan, among others, placed the responsibility for defining these constructions in the lexicon, a characteristic which LFG inherits. A passive participle like *eaten* was to be directly represented there, together with the information that the surface subject of its clause was to act as the patient of the eating at the thematic or meaning-representational level. A similar approach to these constructions was also implicit in several of the computational approaches (Thorne *et al.*, 1968; Woods, 1970, 1973; Winograd, 1972), as Bresnan (1976) pointed out.[3] One advantage of base generation is that it immediately explains the fact that a great many of the structure preserving movements are clause-bounded in scope. The base generation hypothesis also permits the further hypothesis that the operations of the processor correspond directly to rules of the grammar. Bresnan and Kaplan, in the introductory essay, "Grammars as Mental Representations of Language" term this the "Strong Competence Hypothesis". It stands in contrast to the hypothesis that while the grammar defines the function that is computed, it corresponds in no simple way to the operations of the algorithm that computes that function, a possibility which Berwick and Weinberg (1982, 1983) have argued at some length exists for natural language processors, and which is in fact characteristic of certain compilers, for reasons of efficiency.

The strong competence hypothesis is attractive because it implies fewest additional assumptions. However, it is important to remember that it is only a working hypothesis. There is probably no conceivable data that could be used to test it. It can therefore never do more for us than any other instance of Occam's razor – that is, than allow us to choose between theories that are otherwise linguistically adequate. Much of the controversy between transformation-based and LFG-based processors that has been aired by Berwick and Weinberg, and is continued by Bresnan and

Kaplan here, is therefore of rather secondary importance. While it is true that all versions of TG to date are more or less incompatible with the strong competence hypothesis, and true that its abandonment appears to make for a more complex theory, it is equally true that such arguments alone cannot prove the psychological invalidity of TG or any other grammar in advance of the standard criteria of observational, descriptive and explanatory adequacy, to which we now turn.

1. LFG AS A THEORY OF GRAMMAR

1.1. *Overview*

LFG has two principal levels of analysis, one of which provides the input to semantic interpretation, and the other of which characterises constituent structure. The meaning-related level is called functional structure, (or f-structure). Structures at this level represent three kinds of information about strings.

(a) A hierarchical structure, which differs from an ordinary constituent structure in a number of ways. It is typically much "flatter" than the orthodox constituent structure of highly configurational languages like English, and the linear order of elements is not represented. Elements may be included which are not represented in the surface string.

(b) "Grammatical functions" of those elements, which resemble surface case roles. Again, the elements in question may be absent in surface structure.

(c) "Predicate argument structure", which defines the mapping of those grammatical functions onto arguments of the head of the constituent.

The mixture of information about both surface grammatical functions and meaning-related predicate argument relations in a single level of representation is distinctive.

Like a government-binding (GB) S-structure, an f-structure associates extracted items such as Wh-pronouns with a trace at the level of the extraction site, via functional variables or pointers, of which more later. Unlike S-structures, but like the corresponding entities in certain other non-transformational theories such as generalised phrase structure grammar (GPSG), the use of empty categories is confined to the analogue of Wh-movement: constructions which would in transformational terms involve NP movement are all generated directly.

The more surface level of analysis is called constituent- or c-structure and is defined by a context free grammar of a fairly familiar kind. The mapping from c-structure to f-structure is accomplished by a number of distinct components of grammar. A fairly straightforward combination of the context free syntax and the lexicon produces an object which can be thought of as an annotated c-structure, in which grammatical functions of constituents are identified via "metavariables" defining functional dependencies.[4] There are just two types of metavariable. The "immediate domination" type of metavariable is written ↑ ("up") or ↓ ("down"). "Up" stands for (the functional structure of) the node's parent node, while "down" stands for the node's own functional structure. (It should be noted that "up" and "down" are thus rather different sorts of thing. The latter just identifies an f-(sub)structure, while the former includes a tree-defining primitive. The other type of metavariable, of which discussion is deferred, is used to relate extracted elements and traces.) The distribution of the work done at this stage by syntax and lexicon may be different in different languages: in configurational ones, like English, the CF syntax defines grammatical functions via annotations on the constituents of a PS rule, using the first kind of metavariable. So for example, the expansion for S is the following:

(1) $S \rightarrow \underset{(\uparrow \text{SUBJ}) = \downarrow}{NP} \quad VP$

– which means that the NP's parent's f-structure's subject is the NP's f-structure. In a language with freer constituent order than English, the CF syntax may play a much lesser role in defining grammatical functions, with case inflection taking more of the responsibility via the lexicon. For example, nominative case might in a language like Latin define a nominative noun-phrase such as *ego* as being the SUBJ of ↑, its mother's f-structure.

A clear distinction should be drawn at this point between two contentions of LFG concerning grammatical functions in this technical sense. The first is that surface grammar must include the capacity to define grammatical functions of constituents both configurationally (via CF rules) and non-configurationally (via the lexicon). The freedom to accomodate both configurational and non-configurational elements of surface grammar is a feature of LFG which is shared by a number of other theories, such as Relational Grammar (RG) and certain versions of categorial grammar, and which is rapidly being taken up by others, such as GPSG (Gazdar and Pullum, 1982). The second contention is that these surface grammatical functions are preserved at the same level as inter-

pretable meaning-representation – that is, in f-structure – rather than being replaced by the thematic roles upon which interpretation itself depends. This suggestion is less appealing, because it implies a dissocation between syntax and semantics for which there is no apparent reason, and which therefore threatens the explanatory value of the theory.

Surface grammatical functions like subject have no intrinsic semantic role in f-structure. They merely provide a place-holder for the arguments of the predicate or head of a phrase. The responsibility for mapping these constituents onto predicate argument structure, and hence onto meaningful thematic relations such as agent and patient, rests with lexical entries for verbal forms. For example, a transitive verb like *kiss* and its past participle *kissed* have lexical entries which map different grammatical functions onto the thematic roles that go with the arguments of the predicate KISS, thus

(2)a. kiss: KISS⟨(SUBJ)(OBJ)⟩.
b. kissed: KISS⟨OBL_{AG})(SUBJ)⟩.

A further component of rules "instantiates" the metavariables with functional variables identifying the functional structures of the particular constituents that the individual instances of the two types of metavariable relate. At this point, the annotations on nodes such as the subject mentioned above can be entirely divorced from the c-structure, and considered as an unstructured unordered set of equalities defining the functional relations of the sentence. This set of equations or propositions is called an "f-description".

F-descriptions differ from f-structures in two important respects. First, although they contain all the information necessary for the construction of the semantic dependencies, including long range dependencies, they do not themselves embody them. In fact, an f-description merely defines an infinite class of f-structure. A further set of rules interprets the equalities of the f-description and maps it onto a unique f-structure, which *does* embody semantic dependencies, and can be interpreted. These rules play a central role in the grammar. They are responsible for basic phenomena like number agreement. They are crucial in more abstruse areas, like control (discussed below). And they also embody considerable power. For example, it is only by the operation of these rules, with their power to unify isomorphic subtrees (in a manner analogous to the discussion of 'WW' languages on p. 270 of MRGR), that the correct cross-serial dependencies in Dutch infinitival constructions can be established in the analysis of Bresnan *et al.* (1982).[5]

Another way in which the procedures which map f-descriptions onto

f-structures are grammatically crucial stems from the fact that LFG grammars for natural languages typically overgeneralise at the level of c-structure and f-description, allowing illegal strings to map onto grammatical structures at these levels. For example, the English relative clause is base generated by a PS rule much like the following:

(3) $S' \rightarrow XP\ S.$

The extracted phrase is linked to an empty category via the second type of metavariable, discussed below. However, such a rule will overgeneralise to illegal relative clauses lacking an appropriate empty category, such as:

(4) (a man) *who Harry likes Fred.

These overgeneralisations have to be filtered by the failure of the procedures in question to map the corresponding f-description onto an f-structure, in a manner reminiscent of Brame's (1978) semantic filtering of an overgenerating base grammar for relative clauses. The filtering effects of these further rules are enshrined in a succession of "conditions on grammaticality" in the form of "functional well-formedness conditions". For example, f-descriptions have to be *consistent* and *determinate*. That is, a given entity in a description (such as the number of the finite verb) must not be assigned more than one value, and all the entities must be assigned some value (p. 203, ex. 57). Similarly (p. 212, ex. 77) an f-description must define an f-structure which is *complete* and *coherent*. That is, all the governable functions that are demanded by a governer (such as a verb) must be present, and only those governable functions may be present.

The instantiation rules and functional well-formedness conditions amount to axioms of an *interpretation* of the LFG grammar, of a kind that would be required for a complete formalisation of even the simplest CF grammar. However, it is important to be clear *what* grammar they are an interpretation for. It is a grammar defined by the procedures that map between f-description and f-structure, not the grammar defined by the PS rules and the lexicon. Like the instantiation rules, the conditions can on occasion do substantial grammatical work, over and above what would be implicit in the related interpretation of a CF grammar. For example, it is only by the grace of completeness and coherence that the analysis of Dutch infinitivals mentioned above excludes strings with inappropriate numbers of verbs and arguments. Since the two arise from recursive expansions of quite independent PS rules, such strings are legal at c-structure.[6] It follows that these rules and conditions are a part of the syntax itself. They constitute an augmentation of the basic grammatical apparatus that maps sentences onto c-structure and f-description.

It seems natural to group these elements under three headings, and to consider them in turn. Under the first heading come components that are concerned with surface grammar (dually embodied in the configurational CF rules and the non-configurational lexicon), and the particular combination of grammatical-function and predicate-argument relational information which is embodied in f-structure. The second concerns those parts of this apparatus which handle Wh-movement. The third concerns the instantiation/unification procedures which deliver an interpretable f-structure and filter out overgeneralizations.

1.2. *Surface grammar and functional representations*

Virtually all of the evidence offered in MRGR is devoted to supporting the first of these components. Almost all of this evidence is confined to proving the superiority of lexicalism and functional representations for clause-bounded constructions, such as Passive, There-insertion and Raising, (Bresnan, ch. 1, "The Passive in Lexical Theory"), the related matters of ergativisation by reflexive clitic pronouns in Romance languages (Grimshaw, ch. 2, "On the Lexical Representation of Romance Reflexive Clitic"), other systematic relations between verbal constructions (Bresnan, ch. 3, "Polyadicity"), and case and control in English, Russian and Icelandic (Bresnan, ch. 5, "Control and Complementation"; Neidle, ch. 6, "Case Agreement in Russian"; Andrews, ch. 7, "The Representation of Case in Modern Icelandic"). Mohanon (ch. 8, "Grammatical Relations and Clause Structure in Malayalam") exploits the scope that LFG offers for surface grammatical roles to be defined in association with morphological case in the lexicon rather than configurationally, to provide an account of the very free clause constituent order in the Dravidian language Malayalam. By contrast, although a substantial section of the theoretical chapter 4, "Lexical Functional Grammar: A Formal System for Grammatical Representation" by Kaplan and Bresnan is also devoted to the mechanism for Wh-movement, (to which we turn below), very little of this discussion makes any serious claim for the linguistic superiority or originality of the theory in this respect. Similarly, while there is considerable discussion of the power of LFG in the theoretical chapter, the linguistic case studies tend not to exploit its full power in any syntactically significant way, as the study of Dutch already referred to does. An exception is the interesting ch. 9, "Sluicing: A Lexical Interpretation Procedure" by Levin, who attempts to show that "sluicing", a construction which is notoriously troublesome in transformational terms. can be accounted for in terms of copying operations at the level of functional

structure itself, exploiting the tree-unification powers of LFG. Where this power is exploited in some more constrained way, (as in the studies of control discussed below), there is very little discussion of potential explanations for the constraints.

Despite a certain narrowness of range in this sense, the papers provide elegant and convincing support for the lexicalist and functionalist bases of the theory. In particular, Bresnan's important and central chapter on control and complementation (which, like her ch. 3, has appeared elsewhere, as Bresnan, 1982) undertakes a searching comparison of the LFG analysis with GB and with Bach's 1979 Montague-grammatical treatment.

In a theory like GB, in which the notion subject is defined indirectly in terms of configuration, it appears to be predicted that languages which lack a VP should exhibit different control properties, either lacking control verbs or allowing categories other than the subject to be controlled. However, it is argued, VSO languages like Irish and SOV ones like Malayalam show strikingly similar control phenomena. LFG copes with this fact by using a combination of c-structure syntax and lexical case to build f-structures identical to those of English defining control at f-structure level. LFG is therefore argued to start with a considerable advantage in an area where GB appears at its murkiest. Similar claims are advanced against the accounts within Montague Grammar (MG) of Dowty (1978, 1982) and Bach (1979, 1980).

The claim is a strong one. There is no dispute that LFG provides an interesting and revealing account of this constellation of constructions across many languages. The inclusion of surface grammatical functions as primitives lends it some clear advantages over analyses that have been produced to date within GB, GPSG and MG. But the stronger implication that *no* analysis within these frameworks could be adequate is more open to question. The remainder of this section briefly examines the relation between LFG and these other theories, and proposes that the relation to GPSG and MG is closer than LFGs exotic and somewhat cumbersome notation might suggest. It is stressed that the intention is to expose this relation, rather than to prove the superiority of any particular alternative analysis.

The argument outlined above, to the effect that control phenomena are not susceptible to a configurational treatment because of the wide cross-linguistic variety of configurational realisations of the relevant constructions, is questionable. This variation is not really as great as the discussion on p. 319 of MRGR suggests. Most of the variation arises from the contrast between languages that have a flat structure for subject, verb and complement, and those that have a two-level structure involving a VP

node, as illustrated in MRGR, Fig. 5.7. But in a Montague analysis of these elements, where the verb is a functor and the subject and complement are its arguments, this is simply the difference between the configurations that ensue from the application of a two-argument function $\lambda x, y[\ldots x \ldots y \ldots]$ to its arguments, versus the application to an argument of a unary function $\lambda x[\lambda y[\ldots x \ldots y \ldots]]$ to yield another function, which is then applied to the other argument. The equivalence between n-ary functions and unary higher-order functions of this kind is well-known since the work of Schönfinkel (1924; see Dowty, 1982 for brief discussion). It follows that such variety is not necessarily problematic for theories which preserve a closer relation between configuration and function than LFG does, particularly those which exploit order-free rules.[7]

The fine detail of the theory of control also suggests a close relation between LFG and other non-transformational theories. Consider the fragments of syntax and lexicon that are involved in Bresnan's treatment of subject and object control verbs like *promise* and *persuade*. The lexical entries are roughly as follows (cf. MRGR, p. 220–223. Certain irrelevant details, including tense and the fact that the complement must be a *to* infinitive, have been suppressed):

(5) promise:

V, (↑ PRED) = PROMISE ⟨(↑ SUBJ)(↑ OBJ)(↑ VCOMP)⟩
(↑ VCOMP SUBJ) = (↑ SUBJ).

(6) persuade:

V, (↑ PRED) = PERSUADE ⟨(↑ SUBJ)(↑ OBJ)(↑ VCOMP)⟩
(↑ VCOMP SUBJ) = (↑ OBJ).

The general PS rule for expanding the VP can for present purposes be taken as (cf. MRGR, p. 206):

(7) VP → V (NP)(NP)
(↑ OBJ) = ↓ (↑ OBJ2) = ↓
PP* (VP′)
(↑ (↓ PCASE)) = ↓ (↑ VCOMP) = ↓.

The functional equations in the above examples potentially allow the encoding of a wide variety of tree properties and discontinuous relations, but only over a finite number of steps. (As previously noted, they therefore cannot be used to encode unbounded dependencies: the second type of metavariable, to be discussed in the next section, takes care of them.) For example, the subject control equation "my mother's f-structure's VCOMP's subject is my mother's f-structure's subject" can be encoded as

(↑ VCOMP SUBJ) = (↑ SUBJ) in example (5), by "composing" the grammatical functions VCOMP and SUBJ.

As is pointed out on p. 271 of MRGR, the inclusion of the functional equations and functional composition in this sense is one source of generative power in LFG.[8] This power is very narrowly constrained in a number of ways. First, the "up/down" metavariable is so restricted in its interpretation as to be almost entirely notationally redundant. There is no question of composing the tree relations implicit in two "ups" in a similar fashion, to encode relations like "my mother's mother's f-structure's subject is my mother's subject" as something which might be written (↑ ↑ SUBJ) = (↑ SUBJ). Even if the interpretation of "ups" would allow this equation, no analysis anywhere in the book demands such a use of the up arrow. While on occasion constituent f-structures can in effect be passes over longer distances (as in the analysis of Dutch infinitivals of Bresnan *et al.*, and in the curious Dutch examples considered here in Fig. 5.8), this possibility is the responsibility of the tree-unifying instantiation process, rather than the lexicon or the PS rates. As Grimshaw points out in her elegant analysis of Romance reflexives (p. 95), in lexical entries, information is only passed up to the immediately dominating node, and in PS rules, the grammatical function rules always use a "down" to refer to their own f-structures, and pass it to their parent, again via "up". (The passing of the feature PCASE from the PP to the parent in the PS rule (7), which accomplishes a a schematisation of the rule over PP types, is essentially the same, though it is a remarkably opaque notation.) For most purposes the up/down metavariable can therefore by suppressed by convention.

The operation of composing grammatical functions in functional equations is also restricted in LFG, this time by an explicit constraint, or "functional locality principle", which limits the number of grammatical functions which can be composed on one side of a control equation like those in (5) and (6) to two – the possibility which these examples in fact exploit. That is to say that equations like (a) are permitted in f-descriptions, while those like (b) are not:

(8)a. ↑ VCOMP SUBJ = ↑ SUBJ.
b. ↑ SCOMP VCOMP SUBJ = ↑ SUBJ.

This principle is conjectured to constitute a universal constraint.

Why should such restrictions hold? One possibility is that they hold because the equations encode "functional" relations in a sense that is closer to that used by Bach and Dowty than it is to the sense in which the term is central to LFG. In these terms, the verb *promise* would be a functor

taking a subject and a VP′ as arguments, and the control equations would encode relations among those arguments. (More precisely, the VP′ argument could itself be a functor taking an argument, so that the relation that is encoded could be as simple as the application of that functor to the same subject.) Such an interpretation is thoroughly compatible with a reasonable account of the semantics of the entities in question.

Under this interpretation, LFG would also bear a close relation to GPSG. The fact that LFG allows grammatical functions to be defined either lexically or syntactically means that there is a degree of freedom in the theory to analyse constructions either way. For example, in accounting for object control verbs like *persuade*, we could have as one of a whole set of more specific PS rules, the rule:

(9) $$VP \rightarrow V \quad \underset{\substack{(\uparrow \text{OBJ}) = \downarrow \\ (\uparrow \text{VCOMP SUBJ}) = \downarrow}}{NP} \quad \underset{(\uparrow \text{VCOMP}) = \downarrow}{VP'}$$

together with a lexical specification limited to predicate argument structure for the verb *persuade*, thus.[9]

(10) persuade: PERSUADE ⟨(↑ SUBJ)(↑ OBJ)(↑ VCOMP)⟩.

In effect, this is what some early versions of GPSG do, although like the grammars of Bach and Dowty, they do not use grammatical functions to mediate between surface and semantic interpretation. Instead, they put essentially the same information into the translation (cf. Gazdar 1982, p. 150). Of course, subject control cannot be handled quite the same way, because of the controller does not appear in the surface VP expansion. But GPSG handles this problem by making the semantic translation of a VP like *promise Harry to take a bath* into exactly the kind of function over subject NP intensions that is described earlier. Since all VPs are such functions, the semantics of the basic S expansion rule can be responsible for applying the function to its argument.

These similarities suggest that LFG's node-dominance interpretation of the metavariable notation may obscure the true state of affairs, and that certain rival approaches to these constructions (including GB) may be much closer than it presently implies. It remains to be seen whether the restricted nature of "up" and the functional locality principle can be explained in these terms, but there seems to be a possibility of a considerable notational simplification and reconciliation between the theories.[10] In particular, the necessity for explicitly represented grammatical functions in f-structure to mediate between surface grammar and interpretable predicate argument structure is called into question by the

Montague-based approach of Dowty (1978, 1982) and Bach (1979, 1980), who represent function-argument structure directly, and only implicitly define grammatical functions. (For example, the subject is defined as "an argument that combines with an intransitive verb IV to produce a sentence".)

Levin, in one of the most adventurous chapters in the book, addresses this question directly. She argues that "sluicing" – a construction that has given rise to some of the most terrifying deletion transformations ever conceived of – can be explained in terms of the particular mixture of lexical, surface-grammatical and thematic information that is embodied in f-structure. The following examples are commonly used to argue that the construction is on the one hand syntactic, but must on the other hand be defined at some level other than surface structure:

(11)a. *I don't know what.
b. Harry ate, but *I don't know what.*
c. Harry mentioned his plans to go on holiday, but *he didn't say who with.*

Hankamer and Sag (1976) and Sag and Hankamer (1984) have argued from the fact that isolated sentences like (a) cannot be used "deictically", i.e., under pragmatic control of a context, that the construction lies in the domain of sentence grammar.[11] Levin argues from examples like (b) and (c) that sluicing is an operation at the level of f-structure.[12] The crucial evidence for her analysis is that while sluicing may alter the number of arguments that the verb combines with, as in (b) above, such alterations must preserve both surface grammatical function and underlying thematic roles. Thus (a) below is ruled out by the former condition and (b) by the latter:

(12)a. *The door opened, but I don't know who (opened the door).
b. *John moved, but I don't know what (John moved).

This combination of information is exactly what is represented in f-structure.

Levin's proposal is that the Wh-phrase in a sluiced sentence in a discourse such as:

(13) Janet saw something. I wonder what.

is base generated as one expansion of an S′ complement. This *S′* is then fleshed out by the following process:

(1) The referent of the Wh-phrase in the f-structure of the preceding clause is identified and coindexed with it.

(2) A copy of the entire preceding clause's f-structure, with the co-indexed item deleted, is then merged (by the tree-unifying procedures already provided in LFG) with the f-structure of the S′ dominating the Wh-phrase, to give an f-structure which is exactly like that which would result from *I wonder what Janet saw*.

In order to accommodate "optional arguments" of the sluiced verb, Levin allows certain lexical rules to modify the antecedent f-structure and add arguments. (For example, (the inverse of) Existential Quantification (EQ) operates on the f-structure of *Janet ate* to yield an f-structure corresponding to *Janet ate something*, which allows the above procedure to product the sluice *I wonder what*.) However, not every lexical rule can be allowed to operate in this way. (If causativisation were to apply, the ungrammatical (12b) would result.) So lexical rules which change the mapping between existing grammatical functions and thematic rules are excluded by a "principle of lexical compatibility", or "PLC" (MRGR p. 618).

One could object at this point that the obvious explanation for the PLC is that the existentially quantifying lexical rule has made the antecedent intransitive sentence *Janet ate* include a dummy object *already*. On this assumption (which is already implicit to the analysis of certain optional "additional arguments", such as adjuncts), the empty argument could be coindexed with the Wh-expression as usual, and the copying procedure could operate without any lexical modification stage and without requiring the PLC.

Such a proposal is still lexically driven, and it is trivial to modify the theory in this respect. However, it suggests that Levin's analysis is less uniquely dependent on LFG itself than is claimed here, and by Kaplan and Bresnan elsewhere in the book [p. 231]. The claim is that it is the unique mixture of surface grammatical functions and thematic role-defining predicate-argument structure that is embodied in f-structure which is required for the account of sluicing. However, the PLC and the exclusion of examples like (12) means that sluicing must preserve the mapping between the two. It appears to follow that any theory which preserves the mapping between thematic roles and surface representations – even a configurational theory like GB with its Projection Principle – could accomodate a similar account.[13]

1.3. *Unbounded Dependencies*

It has already been noted that the "single up" metavariable is not capable of expressing unbounded dependencies. Instead, LFG captures Wh-

movement with the following extra device. Relative clauses and related constructions are base generated by rules like (3) above. A "bounded domination metavariable", written "$\Downarrow$" and known colloquially as "double down", is associated with the proposed phrase, along with other features defining its categorial nature, such as NP and +wh. The c-structure grammar includes the option of generating certain maximal phrases like NPs as empty categories. These also are associated with categorial information, and with the other bounded domination metavariable "$\Uparrow$" ("double up"). In the instantiation phase, these metavariables are replaced by functional variables, which are freely equated in the functional description, subject to a number of constraints which filter out certain f-descriptions as "improperly instantiated" [p. 246]. The most obvious of these constraints is that extracted items and their traces must bear the same major and minor features. There is also a definition of "constituent control domain" in terms of domination in c-structure and the absence of certain (language-specific) "bounding nodes" on the path between controller and controllee. This definition is reminiscent of Subjacency, and captures language specific island constraints. Because S′ and NP are bounding nodes in English, it is necessary in order to allow extractions from *that*-complements to include in the grammar a device which is implemented using bounded domination metavariables on the expansion of S′, thus (some details are suppressed):

(14) $S' \rightarrow (\textit{that})$ S
$\uparrow = \downarrow$
$\Uparrow = \Downarrow$.

The bounded domination metavariable equation "$\Downarrow = \Uparrow$" has the effect of equating the associated functional variables across the bounding node from a lower control domain to a higher one, and essentially constitutes a base-generative version of the COMP-to-COMP "escape hatch" of Chomsky (1977).[14] The inclusion of this mechanism for Wh-movement seriously compromises the earlier account of clause-bounded constructions. Since the above devices can be resorted to in order to accomplish unbounded dependencies between instances of this metavariable, they *could* have been used to allow unbounded versions of passive, raising and the rest. The fact that such constructions are bounded is therefore not explained, but only stipulated implicitly, by excluding the use of the double-up metavariable in the lexicon.

It is worth mentioning one more of the conditions listed on p. 246. The "nearly-nested condition", which is discussed further below, filters the results of crossing the metavariable correspondences in the ill formed (a)

below, whilst allowing the nested dependencies in (b):

(15)a. *I wonder [which sonata]$_1$[the violin]$_2$ is tough for her to play$_{-1}$ on $_{-2}$.

b. I wonder [which violin]$_1$[the sonata]$_2$ is tough for her to play $_{-2}$ on $_{-1}$.

(Some languages appear to allow a small number of non-nested dependencies, and the constraint allows this possibility.)

Whatever the fine differences of this account from those developed in mainstream TG, the most interesting comparison is with GPSG and the account of Wh-movement constructions first presented by Gazdar (1981). Despite some quite legitimate criticisms of the GPSG account, derived from work by Maling and Zaenen (1982), GPSG appears to retain a number of advantages. The most notable one is that it offers an explanation of the Coordinate Structure Constraint of Ross (1967) and of the bizarre "across the board" (ATB) exceptions to it discussed by Williams (1978), simply as an inevitable consequence of the particular mechanism that it uses for wh extraction. LFG preusumably still requires an ad-hoc constraint, qualified by an ATB escape clause. The second advantage is that it accomplishes all this with a mechanism which is essentially local, which more recent work in GPSG has shown to be equivalent to a feature passing convention on PS rules. Since it has been shown above that such a mechanism is closely related to LFG's "single up/down" metavariable, GPSG can lay claim to a considerable advantage in explanatory power and parsimony, compared to which the well-founded criticisms of Maling and Zaenen concerning the problem of multiple dependencies seem less important than is implied here. (If GPSG were allowed a nearly nesting condition on induction of multiple slashes, then it might do just as well as LFG on these constructions.) Whether or not the deficiencies in question in GPSG can be overcome, it seems clear that LFG is much less successful in explaining unbounded dependency than it is with the bounded variety. This impression is borne out by the almost complete neglect of conditions on extraction and coordination in the LFG literature. (MRGR has no reference to coordination as such in the index, and the only references in the text appear to be to unpublished and forthcoming papers.)

1.4. *From f-description to f-structure*

Almost the only purely linguistic arguments that can be brought to bear on the procedures that produce f-structures and filter out overgeneralisations

are arguments from power. These are necessarily inconclusive, since nobody really knows what power is implicit in natural language syntax, beyond a common (but not universal) belief that natural languages are likely to be contained within the recursive languages, that they probably cannot all be contained within the context-free languages, and that they may well be included in some (probably proper) subset of the context sensitive languages. It follows that mere power is not in itself necessarily a bad feature in a theory. There certainly seems to be plenty of it in LFG. It is true that the basic apparatus that accrues from the tree-property defining metavariables and the instantiation algorithm is shown to be decidable, and therefore to induce no more than recursive power. (The proof depends upon a condition upon "valid derivation" of c-structures which among other things prohibits indefinitely long non-branching subtrees (p. 266). Such a constraint could be expressed very simply as a constraint on familiar processing techniques, such as chart-parsing.) However, the claim that the system has less than that still very considerable power rests on the inclusion of a further constraint upon bounded domination metavariables, the "nearly nested restriction" referred to earlier. With this constraint, which it is pointed out embodies a well-known generalisation about long range dependencies in the languages of the world, it is claimed that LFG only has the power to generate some subset of the context sensitive languages.

The nearly nested constraint governs the linear ordering in the string of controlling and controlled metavariables (cf. p. 260, 262). Footnote 33 to chapter 4 points out that a number of other constraints on metavariable correspondencies would achieve the same effect, including a "nearly crossed restriction". Such constraints are therefore upon permissible *languages*, and look rather as though they might amount to a "don't" generate any non context-sensitive (or non-indexed or whatever) languages" constraint. If so, then the claim on p. 259 that "the nested pattern of acceptable dependencies is an empirical consequence of the requirement . . . that corresponding metavariables be nearly nested", and that this requirement offers "not only a description of the observed facts, but a formal basis for explaining them", is empty. It is clear from recent work in GPSG and in LFG itself that natural language phenomena requiring more than context free power, though they seem to exist, are comparatively rare. It follows that a truly explanatory theory must ultimately show why any power that it deploys in order to capture these phenomena is so little used. The task of explanation is incomplete until the restrictions in question are shown to stem from principled factors such as the requirements of semantics, the processor, or some other factor in linguistic

communication. LFG may be no worse than any other theory in this respect. But it is no better.

In sum, LFG provides elegant analyses of case and control in both configurational and non configurational languages, by exploiting grammatical functions in the lexicon and the rules of surface grammar. However, it contains a worrying profusion of devices for acommodating discontinuous constituents of various kinds, including two types of metavariables, the apparatus of composition of grammatical functions, and the instantiation procedures. Along with this profusion of devices comes a profusion of constraints, which appear to compromise the explanatory value of the theory for the reasons set out in the introduction above.

2. LFG AS PSYCHOLOGICAL EXPLANATION

2.1. *Psychological Grammar*

For a theory of grammar to be psychologically explanatory, it must, under the Competence Hypothesis, allow for a very direct relation of grammatical rules to the operations of a processor. Ideally, it should require nothing more to turn it into a complete theory of processing than the addition of a mechanism for local ambiguity resolution, to tell it which rule of the grammar to apply at a given point in analysis or generation.

Just about as little is known about the operations of the human sentence processing algorithm as is known about the power of its grammar, and what little there is is almost equally banal. The function that it computes must, obviously, be computable using the finite resources of a physical brain. It clearly operates to some substantial extent from left to right. It appears to be capable of carrying out semantic interpretation incrementally and in parallel with syntactic processing. That is, we appear to comprehend sentences in the fullest sense, more or less word by word as we hear them, and we can also interpret isolated fragments of sentences like ... *not told that* ... and ... *too difficult to attempt to* ....

The space of theories of grammar that are compatible with those very broad limits is very large. (For example, it includes standard-theoretic TG.) However, the Competence Hypothesis cuts this space down a lot. It demands a theory in which there is a rule-to-rule mapping between syntactic rules, rules of semantic interpretation, and operations of the processor. Such a theory must be base generative in a very strong sense indeed.[15] In particular, in order to exhibit the ability to produce analyses for fragments of the kind just mentioned, it must have a property which

Bresnan and Kaplan in the introduction to MRGR term "order-free composition".[16] The proposal to reflect in the grammar the order-free compositional property is one of the most interesting ideas in the whole book, for, under the competence hypothesis, it is a sine qua non for the suggestion by Marslen-Wilson, Tyler and others that human sentence processors carry out incremental semantic interpretation in the most intimate relation possible to syntactic processing, more or lese word by word.

LFG has this property, because of its base-generative grammar and lexicon, and the order-free nature of the instantiation process of producing interpretable f-structures. (It is odd that this important property does not appear to be explicitly proved anywhere in the book, but it is fairly obvious that it follows from the "direct encoding principle", which forbids rules from changing functional assignments.) Despite a number of points in the text at which it is implied that the algorithm provided in ch. 4 is one in which c-structure and f-structure are produced incrementally and in parallel (e.g., p. 787), footnote 10 to that chapter makes it clear that the algorithm that is described there is not in fact of this kind. However, the extence of such an algorithm appears to be essential for efficient parsing, because of the overgeneralisation-filtering effect of the instantiation stage. The more the grammar overgeneralises, the more a processor will tend to develop false analyses, which will occupy computational resources until they can be eliminated. It follows that LFG makes a very strong prediction that natural languages do not separate syntactic processing and f-structure building.[17] It is surprising that this necessary consequence of the theory, and its potential implications for local ambiguity resolution, are almost entirely ignored when it comes to the proposal of such a mechanism in chap. 11, which is considered next.

2.2. *Local Syntactic Ambiguity Resolution, Production, and Acquisition*

A grammar that is adequate under the Competence Hypothesis requires nothing more to make a processor than the addition of a mechanism to resolve local ambiguities and decide *which* rules to apply and when. Ford, Bresnan and Kaplan's ch. 11; "A Competence: Based Theory of Syntactic Closure", proposes such a mechanism for LFG. A number of proposals concerning the nature of the psychological ambiguity resolver have been made, using evidence from the biases in resolving ambiguities that are made obvious by "garden path sentences", such as the following well-known example from Bever (1970):

(16) The horse raced past the barn fell.

The interest of these sentences arises from the fact that the ambiguity in question (between a clause and a complex NP analysis for the substring *the horse raced past the barn*) is misresolved in favour of the clause reading. In this case, the misresolution generally leads naive readers to fail to produce any grammatical analysis at all. The mechanisms that have been proposed to explain this phenomenon have generally hinged upon some structural criterion (such as sheer number of nodes) for distinguishing analyses (Kimball, 1973; Frazier, 1979) or on rule ordering (Wanner and Maratsos, 1976).

Ford *et al.* put forward a proposal combining within an LFG framework a number of the tactics that have been proposed in earlier work. Their starting point is the observation that the resolution of such ambiguities of attachment as are exhibited by the PP in the following two examples is affected by the verb, contrary to purely structural criteria like the "minimal attachment" strategy of Kimball, Frazier and others:

(17)a. The woman wanted the dress on the rack.
b. The woman positioned the dress on the rack.

Accordingly, the most basic principle of their local ambiguity resolving mechanism is that the lexical entries for verbs include an ordering according to preference on the possible expansions of the verb phrase of which they are head. In processing, these alternatives are tried serially, from most preferred to least. A given construction, such as the expansion of a VP as V NP PP, may be differently ordered with respect to some other construction, such as the V NP expansion, in the lexical entries of different verbs. For example, (17) above suggests that it is less preferred by the verb *want*, but more preferred by the verb *position*.

Such a strategy is reminiscent of the rule ordering in ATNs proposed by Wanner and others. A principle of "Invoked Attachment" biasses the processor to direct attachment of NPs etc. to the VP node, and resembles Frazier's (1979) "Minimal Attachment", although its operation is obscure, and algorithm-dependent. It is augmented by a further strategy related to Kimball's (1973) "Right Association" and Frazier's (1979) "Late Closure", which is here called "Final Arguments". Under this principle, attachment of the final argument in a constituent (such as the NP *the dress* in example 17a (but not in 17b)) will be delayed, so that further constituents (like *on the rack* in 17a) will, grammar permitting, be attached to that argument rather than to the matrix. Wanner's 1980 discussion of Frazier's related principle of late closure suggests a similarly close relation between this strategy and rule ordering.

Ford *et al.* sketch an algorithm implementing these proposals, and present a simple experiment in its support. Subjects were presented with booklets of ambiguous sentences, each accompanied by two unambiguous paraphrases which they were asked to choose between. The majority of the results showed consistent preferences across subjects for given constructions for given verbs, irrespective of the particular arguments to which they were applied.

It is a virtue in the theory that the main proposals – lexical preferences, and the principles of invoked attachment and final arguments – are independent of LFG and of the particular algorithm proposed here. Even under the strong competence hypothesis, they would be compatible with almost any base generative grammar and processor, as their kinship with ATN rule ordering suggests. It follows that these results (like the results presented in Ford's ch. 12 on production, entitled "Sentence Planning Units") do not distinguish LFG from other such theories.

Conversely, these results do not exclude a quite different kind of local ambiguity resolving mechanism, in which lexical ordering plays no part, and the ambiguities are entirely resolved by comparing and evaluating interpretations of the rival analyses. Crain (1980, reported in Crain and Steedman, in press), and Altmann (forthcoming), have shown that garden path effects of the kind illustrated in (16) are to a very considerable extend under the control of referential context, and that the so-called "null" context that is typically used in processing experiments is misleading. Ford *et al.* acknowledge that contextual effects can over-ride lexical preferences.[18] But although they combine their verbs with a variety of arguments, they do not attempt to control the variety of these materials or the effects of context. The fact that there are broad consistencies across verbs is therefore hard to interpret. Indeed, their own appendix A.2 shows that in a few cases where factors relevant to reference *were* accidentially included in the material, the predictions of the lexical theory were not confirmed. For example, their (98)

(18) They signaled to someone that they couldn't hear

showed no bias between the two possible analyses, whereas, their (25b) and (28b) showed a massive bias toward complement analyses:

(19)a. They signaled to the guide that they couldn't hear.
b. They signaled to everyone that they couldn't hear.

The appendix puts this down to the "vagueness" of *someone*. However, it is arguable that the real difference is in *definiteness*, a referential category which Crain showed to affect relative clause attachment in the null context, and whose effect he explained in terms of the use of referential

context in the resolution of local ambiguities by the human sentence processor. Since Ford *et al.* admit the need to take into account entailment relations between their sentence pairs, and since, as Frazier (1983) has pointed out, such entailment relations can only be constructed on the assumption that LFG makes these interpretations available, it is unclear in the absence of an experimental measure more directly related to on-line processing whether there is any role at all for lexical ordering in the resolution of local ambiguities.

Pinker's chapter ("A Theory of the Acquisition of Lexical Interpretive Grammar," on the acquisition of LFG grammars by children adds two rather different ingredients to the arguments for psychological reality. It attempts to prove certain formal results relating to learnability. It also attempts to informally specify psychologically plausible heuristics that might actually get the job of acquisition done. The first of these enterprises (which has been criticised by Pullum, 1983) is less relevant to the present purpose than the second. Pinker provides an interesting overview of the problem, and discusses a wide range of learning strategies. Central to all of these is the assumption that the child starts with a complete though partly incorrect f-structure for the construction that he or she is to learn, consisting of the contextually given, correct, predicate-argument structure, plus the possibly incorrect grammatical functions assigned by the current c-structure rules and lexicon. Grammatical principles are offered for acquiring c-structure rules, lexical entries and control equations, and rules of long-distance binding, on the basis of inconsistencies in the f-structure. Some of these operations imply very complex tree analysing operations, notably those associated with the acquisition of some of the complexities discussed above, such as the control equations and the apparatus of Wh-extraction. Those who are already convinced by LFG as a theory of competence grammar will find this an interesting and elegant set of speculations. Those who are not so convinced will, in view of the logical primacy of the theory of grammar over performance mechanisms, probably not be converted.

3. Conclusion

It has been the fate of theories of grammar up till now to be replaced with some rapidity. The tendency of LFG to proliferate grammatical and notational devices, as well as principles and constraints, make it seem unlikely that it will be the exception. The most worrying problem of this kind lies in the sheer number of components which are responsible for discontinuous constituents of various kinds. The two types of metavariable, the composition or grammatical functions in lexical entries, and the instantiation algorithm are all in principle capable of encoding a wide range of tree properties, although the "single" metavariable and the composition property are narrowly restricted in this respect by constraints. But the success of LFG in accounting for bounded constructions, and the close relation that it bears to GPSG and Montague/Catorial Grammar, suggests that any successor will come from among the broad church of non-transformational theories which Bresnan has pioneered. (The particular nature of the constraints that LFG requires make it seem likely that the successor will be one which offers a more unified treatment of long range dependencies, perhaps preserving a closer relation between function-argument structure and surface syntax.) If so, the claims for the psychologically explanatory virtues of LFG will necessarily appear to have been somewhat premature. But the theory will hold the honourable position of having been one of the first to even attempt to take psychological questions seriously, and it is certain that the grammatical studies in diverse languages that are presented here will continue to be read and to testify to the importance of Lexical-Functional Grammar in the development of the theory of grammar.

NOTES

* I would like to thank Jean-Mark Gawron for many discussion on LFG. He, Ewan Klein and Henry Thompson read and commented upon a draft of this paper.

†† *Editor's note*: This is the first of two review articles about this work to be published in *Linguistics and Philosophy*. The second, by Paul Schachter, will appear in Vol. 8, No. 4.

[1] Joan Bresnan (ed.), *The Mental Representation of Grammatical Relations*, Cambridge, Mass., MIT Press, 1982, 874 pages.

[2] The book itself is a collection of individual essays rather than a unified text, and leaves something to be desired in presentation of such preliminaries. In particular, it is unclear why the theoretical overview in ch. 4 is positioned after chs. 1 to 3. Some quite basic points in these chapters are not fully comprehensible in advance of the theoretical chapter, (cf. ch. 1, p. 11–12, for example).

[3] Important differences between the ways in which the structure preserving constraint is interpreted in devices like the ATN and the present theory are discussed in note 13, MRGR p. 277. LFG does not allow the equivalent of register-changing SETR operations.

[4] The fact that these elements of the grammar are being described in a serial fashion is not to be taken as implying that processes would deal with them separately in series, of course. See end of section 2.1 below for some further remarks on this score.

[5] A grammar for this construction which does not exploit tree unification is discussed by Steedman (1984, 1985).

[6] Such overgeneralisations present potentially serious problems for a theory with claims to psychological explanation. They exacerbate the problem of local ambiguity resolution by introducing extra false paths. This feature of LFG and its implications are discussed in the second part of this paper, at the end of section 2.1. Some other conditions are also grammatically crucial. For example, many of the conditions on "double-up" metavariables (MRGR, p. 246, and section 1.3 below), which are required to exclude overgeneralisations of Wh-movement like (4) above, amount to special cases of the earlier conditions. But some

make reference to a notion of "control domain" which is defined over c-structures, rather than in f-description terms. Still others, such as the requirement for "nearly nesting" (also discussed below) are conditions on metavariable correspondences in f-descriptions which make reference to linear order.

[7] Steedman 1985 shows that other phenomena of long-range government of the kind illustrated in the (curiously special-case) example in MRGR, Fig. 5.8 are susceptible to a related categorial analysis.

[8] See Ades and Steedman (1982) and Steedman (1985) for proposals to include functional composition operations in categorial grammars for various natural languages, and Moortgat (1983) for related proposals concerning the lexicon. In these theories, functional composition bears the sole responsibility for accommodating discontinuous constituency.

[9] Such an alternative would require the relaxation of a constraint on LFG which currently prohibits double functional assignments in the PS rules, and would require the significant addition of some apparatus equivalent to GPSG rule numbering. To repeat, the concern here is to establish the relation between these theories, not to support any particular alternative analyses.

[10] There are other related constraints which suggest that LFG may be encoding function-argument relations in this sense – for example, the "first sister principle" (MRGR p. 37).

[11] It is not entirely clear that this argument is correct. A view which regarded the sluiced Wh-expression as pronominal would not be overthrown by the impossibility of pragmatic control if thematic relations could only be introduced into the context linguistically. It is at least arguable that, unlike physical entities and events, thematic roles are simply *not there* in the world, rather having the status of interpretations laid upon things and events. The suspicion that the phenomenon is not grammatical is lent support by the unique properties of the deletion operation involved in Sluicing, which pays scant regard to clause structure, unlike, say, gapping. However, for present purposes we will ignore this possibility.

[12] It might appear from example (11c) that sluicing cannot occur at surface structure, but only at some deeper level at which nominals are clause nuclei (cf. MRGR, p. 647). However, the existence of related sentences like

(i) Harry mentioned a holiday but he didn't say who with

suggests that this assumption is not necessary.

[13] Levin claims in her footnote 9 that current versions of REST cannot accommodate the construction. However, this argument depends upon the assumption that sluices like (a) have meanings corresponding to (b).

(i)a. I heard the claim that he bit someone, but I don't remember who.
b. *I don't remember who I heard the claim that he bit.

The point here is that (b) cannot be realised as a surface string, but the LFG distinction between c-structure and f-structure, and the way that the island constraints are imposed (see below) means that a properly instantiated f-structure corresponding to (b) *can* be constructed. However, the sluiced clause in (i,a) clearly *doesn't* mean (i,b), any more than the sluiced clause in (ii,a) means the perfectly well formed (ii,b):

(ii)a. I know that Harry kissed someone but I don't know who.
b. I don't know who I know that Harry kissed.
c. I don't know who Harry kissed.

what (ii,a) actually means is more like (ii,c). The possibility that a theory like GB can accommodate sluicing is therefore still open. What is more, it is clear that a substantial extension to the present theory is required. Mere copying of the entire f-structure must be replaced by a process of search for and copying of, relevant sub-structures of f-structures. At this point one again begins to wonder whether this phenomenon is related to sentence grammar at all (cf. note 11). The suspicion that it isn't is strengthened by the observation that

all the information that is required for an interpretation is, (given the above proposal to replace the PLC), made available in the first step of Levin's procedure, in which the Wh-expression and its antecedent are identified. The copying process itself seems entirely redundant.

[14] Some differences are discussed in note 27 to ch. 4.

[15] For example, Bresnan and Kaplan point out that the ATN and other instantiations of transformational theories as left-to-right processors would not qualify, since they allow previously determined grammatical relations to be modified by subsequent steps in the analysis, thus rendering order free composition impossible. Order-free combination and the direct encoding principles are also on the side of efficient processing, because such changes also have potentially disastrous effects on standard techniques for local ambiguity resolution, such as chart-driven parsing. (cf. Winograd, 1983).

[16] The earlier examples are from Bresnan and Kaplan. Naturally, such fragments will typically be more ambiguous than they are in context, and any given interpretation may be more or less preferred out of context. Nevertheless, interpretations are available, and they include the one(s) selected by any grammatical context.

[17] That is, LFG appears both to permit and to require for efficient processing that c-structure and f-structure be built in parallel, rather than in sequence. Note however that both types of structure must be maintained if bounded domination metavariables are to be properly instantiated.

[18] They do in fact make a proposal for how such an interaction might work in a processor like theirs. They propose to allow the context to "dynamically change the strength of lexical form as a sentence is being processed." It is interesting that this is what Crain and Steedman argue against as the "strong" interaction of syntactic and semantic processing.

References

Ades, Anthony E. and Mark J. Steedman: 1982, 'On the order of words', *Linguistics and Philosophy* **4**, 517–558.

Altmann, Gerry: 1985, 'Reference and the resolution of local ambiguity: interaction in human sentence processing', forthcoming Doctoral thesis, University of Edinburgh.

Bach, Emmon: 1977, 'Comments on a paper by Chomsky', in Peter Culicover, Tom Wasow and Adrian Akmajian (eds.), *Formal Syntax*, Academic Press, New York, pp. 133–155.

Bach, Emmon: 1979, 'Control in montague grammar', *Linguistic Inquiry* **10**, 515–531.

Bach, Emmon: 1980, 'In defence of passive', *Linguistics and Philosophy* **3**, 297–341.

Berwick, Robert C. and Amy S. Weinberg: 1982, 'Parsing efficiency, computational complexity, and the evaluation of grammatical theories', *Linguistic Inquiry* **13**, 165–192.

Berwick, Robert C. and Amy S. Weinberg: 1983, 'The role of grammars in models of language use', *Cognition* **13**, 1–62.

Bever, Tom: 1970, 'The cognitive basis for linguistic structures', in J. R. Hayes (ed.), *Cognition and the growth of language*, Wiley, New York, pp. 279–362.

Brame, Michael K.: 1976, *Conjectures and refutations in syntax*, Elsevier, North Holland, New York.

Brame, Michael K.: 1978, *Base generated syntax*, Noit Amrofer, Seattle.

Bresnan, Joan W.: 1976, 'Toward a realistic model of transformational grammar', unpublished paper presented to the MIT AT&T Convocation on Communications, "New approaches to a realistic model of language", Massachusetts Institute of Technology.

Bresnan, Joan W.: 1978, 'A realistic transformational grammar, In Morris Halle, Joan W. Bresnan, and George A. Miller (eds.), *Linguistic sructure and psychological reality*, MIT Press, Cambridge Mass., pp. 1–59.

Bresnan, Joan W.: 1982, 'Control and complementation', *Linguistic Inquiry* **13**, 343–434.

Bresnan, Joan W., Ronald M. Kaplan, Stanley Peters, and Annie Zaenen: 1982, 'Cross-serial dependencies in Dutch', *Linguistic Inquiry* **13**, 613–636.
Carlson, Greg N.: 1983, 'Review of MRGR', *Natural Language and Linguistic Theory* **1**, 261–280.
Chomsky, Noam: 1977, 'On Wh-movement', in Peter W. Culicover, Tom Wasow, and Adrian Akmajian (eds.), *Formal Syntax*, Academic Press, New York, pp. 71–132.
Chomsky, Noam: 1981, *Lectures on government and binding*, Foris, Dordrecht.
Crain, Stephen: 1980, 'Pragmatic constraints on sentence comprehension', unpublished Ph.D. dissertation, University of California, Irvine.
Crain, Stephen and Mark J. Steedman: 1982, 'On not being led up the garden path: the use of context by the psychological parser', in David Dowty, Lauri Kartunnen and Arnold Zwicky (eds.), *Natural language parsing: psychological, computational and theoretical perspectives, ACL Studies in Natural Language Processing*, CUP, Cambridge, in press.
Dowty, David: 1978, 'Lexically governed transformations as lexical rules in a Montague grammar', *Linguistic Inquiry* **9**, 393–426.
Dowty, David: 1982, 'Grammatical relations and Montague grammar', in Pauline Jacobson and Geoffrey K. Pullum (eds.), *The nature of syntactic representation*, Reidel, Dordrecht, pp. 79–130.
Emonds, Joseph: 1970, 'Root and structure preserving transformations', unpublished Doctoral dissertation, Massachusetts Institute of Technology.
Emonds, Joseph: 1976, *A transformational approach to English syntax*, Academic Press, New York.
Fodor, Jerry A., Tom B. Bever, and Merrill Garrett: 1974, *The Psychology of Language*, McGraw-Hill, New York.
Fodor, Janet Dean: 'Parsing strategies and constraints on transformations', *Linguistic Inquiry* **9**, 427–473.
Frazier, Lyn: 1979, 'On comprehending sentences: syntactic parsing strategies', Doctoral dissertation, University of Connecticut, Indiana University Linguistics Club.
Frazier, Lyn: 'Review of MRGR', *Natural Language and Linguistic Theory* **1**, 281–310.
Gazdar, Gerald: 1981, 'Unbounded dependencies and coordinate structure', *Linguistic Inquiry* **12**, 155–184.
Gazdar, Gerald: 1982, 'Phrase structure grammar', in Pauline Jacobson and Geoffrey K. Pullum (eds.), *The nature of syntactic representation*, Reidel, Dordrecht, pp. 131–186.
Gazdar, Gerald and Geoffrey K. Pullum: 1982, in Michael Moortgat, H. v.d. Hulst and Teun Hoekstra (eds.), *The scope of lexical rules*, Foris, Dordrecht, pp. 109–123.
Hankamer, Jorge and Ivan Sag: 1976, 'Deep and surface anaphora', *Linguistic Inquiry* **7**, 391–428.
Heny, Frank: 1979, 'Review article: Chomsky's "Logical structure of linguistic theory"', *Synthese* **40**, 317–352.
Kay, Martin: 1979, 'Functional grammar', *Proceedings of the fifth annual meeting of the Berkeley Linguistic Society*, University of California at Berkeley.
Kimball, John: 1973, 'Seven principles of surface structure parsing in natural language', *Cognition* **2**, 15–47.
Kuno, Susumo and Jane Robinson: 1972, 'Multiple Wh-questions', *Linguistic Inquiry* **3**, 463–488.
Maling, Joan M. and Annie Zaenen: 1982, 'A phrase structure account of Scandinavian extraction phenomena', in Pauline Jacobson and Geoffrey K. Pullum (eds.), *The nature of syntactic representation*, Reidel, Dordrecht, pp. 229–282.
Marcus, Mitch: 1980, *A theory of syntactic recognition for natural language*, MIT Press, Cambridge, Mass.
Moortgat, Michael: 1983, 'A Fregean restriction on metarules', *Proceedings of the fourteenth annual meeting of the Northeastern Linguistic Society*.

Pullum, Geoffrey K.: 1983, 'How many possible human languages are there?', *Linguistic Inquiry* **14**, 447–468.
Ross, John Robert: 1967, 'Constraints on Variables in Syntax', unpublished Doctoral dissertation, MIT.
Sag, Ivan and Jorge Hankamer: 1984, 'Toward a theory of anaphoric processing', *Linguistics and Philosophy* **7**(3), 325–345.
Schönfinkel, M.: 1924, Über die Bausteine der mathematischen Logic', *Mathematische Annalen* **92**, 305–316.
Steedman, Mark J: 1984, 'A categorial theory of intersecting dependencies in Dutch infinitival complements', in Wim de Geest and Yvan Putseys (eds.), *Sentential Complementation*, Foris, Dordrecht, pp. 215–226.
Steedman, Mark J.: 1985, 'Dependency and Coordination in the Grammar of Dutch and English', *Language* **61**(3), (in press).
Thorne, James Peter, Paul Bratley, and Hamish Dewar: 1968, 'The syntactic analysis of English by machine', in Donald Michie, *Machine Intelligence* 3, American Elsevier, New York, pp. 281–309.
Wanner, Eric: 1980, 'The ATN and the Sausage Machine: which one is baloney?', *Cognition* **8**, 209–225.
Wanner, Eric and Michael Maratsos: 1978, 'An ATN approach to comprehension', in Joan W. Bresnan, George A. Miller, and Morris Hallé (eds.), *Linguistic theory and psychological reality*, MIT Press, Cambridge, pp. 119–161.
Wexler, Kenneth, and Peter W. Culicover: 1980, *Formal principles of language acquisition*, MIT Press, Cambridge, Mass.
Williams, Edwin S.: 1978, 'Across-the-board rule application', *Linguistic Inquiry* **9**, 31–43.
Winograd, Terry: 1972, *Understanding natural language*, Edinburgh University Press, Edinburgh.
Winograd, Terry: 1983, *Language as a cognitive process: I*, Addison-Wesley, Reading, Mass.
Woods, William A.: 1970, 'Transition Network Grammars for natural language analysis', *Communications of the ACM* **13**, 591–606.
Woods, William A.: 1973, 'An experimental parsing system for transition network grammars', in Randall Rustin (ed.), *Natural language processing*, *Courant Computer Science Symposium* 8, Algorithm Press, New York, pp. 111–154.

School of Epistemics and Department of Artificial Intelligence
University of Edinburgh