MORPHOSYNTAX AND ITS GENERATION IN FUNCTIONAL DISCOURSE GRAMMAR: WHAT CAN BE LEARNED FROM ROLE AND REFERENCE GRAMMAR?

Christopher S. BUTLER

ABSTRACT: The morphosyntactic level in Functional Discourse Grammar (FDG) is currently underdeveloped. The aim of the present paper is to make proposals relevant to the structure of the morphosyntactic level in this theory, and to how the interpersonal (pragmatic) and representational (semantic) levels of the grammatical component map on to it. These proposals rely heavily on a second structural-functional theory, Role and Reference Grammar (RRG). It is suggested that morphosyntactic structures in FDG need to be much more like the layered structures of RRG, and that given appropriate additions to the representational level which have been shown to be needed cross-linguistically, FDG could then adopt (or adapt) the semantics-to-syntax linking algorithms of RRG in order to provide a generative interfacing mechanism between the interpersonal and representational levels, on the one hand, and the morphosyntactic level on the other.

KEYWORDS: Functional Discourse Grammar; Functional Grammar; Role and Reference Grammar; morphosyntax; linking algorithms.

1 Introduction

Functional Discourse Grammar (henceforth FDG; HENGEVELD, 2004a, 2004b, 2005; HENGEVELD; MACKENZIE, 2006; 2008, Forthcoming; MACKENZIE; GÓMEZ-GONZÁLEZ, 2004, 2005) offers a top-down model of grammar which is meant to reflect, in its general architecture, the ideas put forward by Levelt (1989, 1999) in relation to an account of language production. As shown in Figure 1, adapted from Hengeveld and Mackenzie (2006, p.669, 670),

1 Honorary Professor – Swansea University – Department of Applied Linguistics – Swansea SA2 8PP – UK. E.mail address: cbutler@telefonica.net

2 It should be noted that this version of the model outline differs in one important respect from that given in Hengeveld (2005, p61), which does not have an arrow feeding information from the interpersonal to the representational level, but rather formulates the interpersonal and representational structures independently. However, in the earlier article Hengeveld (2005, p.73) states that information at the interpersonal level can have consequences for the representational level (e.g. selection of an imperative frame forces the choice of a controlled event). I have therefore followed the later version of the model here.
Figure 1 – Components, operations and levels of representation in FDG
the central GRAMMATICAL COMPONENT of the overall model is driven by a CONCEPTUAL COMPONENT, interacts with a CONTEXTUAL COMPONENT, and passes its resulting structures to an OUTPUT COMPONENT concerned with the final articulation of the utterance in sound, writing or gestural sign. The grammatical component consists of four levels: the INTERPERSONAL, at which the utterance is planned in terms of discourse pragmatics as a move consisting of one or more acts, which in turn consist of subacts; the REPRESENTATIONAL, at which the semantics (predicate-argument structure, additional modifiers, etc.) is dealt with; the MORPHOSYNTACTIC, which takes the output of the interpersonal and representational levels and converts it to an ordered syntactic structure with appropriate morphology; and the PHONOLOGICAL, which converts the output of the morphosyntactic level into a pre-phonetic phonological representation. Each level is fed by a set of PRIMITIVES, which includes a subset with structuring function, a subset in phonemic form, and a subset of grammatically-realised operators. In order to respect the incrementality of production proposed in Levelt’s model, it is proposed that as soon as one level produces enough output for a lower level to act on, that part of the output is passed on without necessarily waiting for the rest of the production from a particular level. In the current state of development of FDG, we have simply four different, parallel specifications of the utterance, discourse-pragmatic, semantic, morphosyntactic and phonological. Clearly, however, if the theory is to develop into a fully generative model which is capable of specifying all the stages in going from conceptualisation to articulation, it is essential to describe how the levels interface. In other words, we must show how structures from the interpersonal level are mapped on to those at the representational level, how output from both of these levels is mapped on to the morphosyntactic level, and how what is produced by this level is mapped on to a phonological representation. Apart from some hints in the work of Bakker and Siewierska (2004) on how interpersonal and representational structures could be fused into a single underlying structure for input to dynamic expression rules based on earlier FG, little has been done on the FDG mapping rules. Furthermore, since the requirements of language production and comprehension are in some ways very different, it is likely that the mappings from one level to another will not simply be reversible, if they are to attain a high degree of cognitive adequacy. The purpose of the present paper is to make some suggestions for how the morphosyntactic level of FDG might be represented, and the way in which interpersonal and representational elements might be mapped on to it in the productive direction. These proposals will rest on a comparison between FDG and a second structural-functional grammar, Role and Reference Grammar (RRG).
2 The morphosyntactic level

2.1 The current model

The most recently published summary of FDG states that “[t]he morphosyntactic level accounts for all the linear properties of a linguistic unit, both with respect to the structure of sentences, clauses, and phrases, and with respect to the internal structure of complex words” (HENGEVELD; MACKENZIE, 2006, p.674). The morphosyntactic structure proposed is of an extremely orthodox kind, as shown in the linear structure for example (1) given in (2), or alternatively the tree diagram in (3):

(1) The clock ticked merrily, … (BNC^3 ASE 1923)

(2) [[[the_{Art} clock_{N-SG}]_{NP} [tick_{V-PAST} [merrily_{AdvP}]_{AdvP}]_{VP}]_{ClI}]_{Si}

(3)

In other words, the whole sentence is seen as consisting of a single clause, which in turn consists of a NP and a VP, the latter containing a verb and an AdvP; the NP consists of an article and a singular noun, and the AdvP of a single adverb. The structuring of the morphosyntactic level is achieved through the operation of templates for words, phrases, clauses and sentences, provided by the set of primitives which feed this level (HENGEVELD, 2005, p.68). I assume that it is intended that the templates for clauses will be broadly of the type proposed in Dik’s account of FG (DIK, 1997 p.408-414, p.424-427), consisting minimally of slots for Subject, Verb, Object^4 and pragmatically-significant positions such as initial and final position in the clause. The important question which arises is

---

^3 Examples marked BNC are from the British National Corpus (World Edition).

^4 Subject and Object are assigned as syntactic functions at the morphosyntactic level (HENGEVELD, 2004b, p.373).
whether the structural means available in FDG are sufficient, in view of the
evidence about morphosyntactic patterning available across a typologically-varied
range of languages. In order to make progress on this question it is instructive to
look at what is proposed in RRG.

2.2 The layered structure of the clause in Role and Reference Grammar

Role and Reference Grammar (RRG: VAN VALIN; LAPOLLA, 1997; VAN VALIN,
2005), like FDG, postulates separate levels of semantics and (morpho)syntax. In
what follows, I shall first summarise briefly the nature of the morphosyntax in
RRG and outline the evidence for the kind of syntax proposed, and then examine
what parallels already exist between the syntax of RRG and that of FDG.

The syntactic level of RRG is organised as a layered structure of units which,
although motivated in semantic terms, are argued for on the basis of syntactic
evidence. The clause is divided into a core, which contains the nucleus and core
arguments, and a periphery in which non-arguments occur. The nucleus houses
the semantic predicate, and the core arguments are also arguments in the
semantic representation for the predicate. These divisions, which are claimed to
be universal, are illustrated in Figure 2.

![Figure 2 – The layered syntactic structure of the clause in RRG](image)

There are also elements which are not claimed to be universal. Some
languages have a pre-core slot (PrCS), in which question words appear in
languages, such as English, which do not place them in their ordinary pattern
position in the clause (see example (4)); this is also the position for material which
would be regarded as ‘fronted’ in theories that admit transformations, and which
is integrated into the clause in English (example (5)). Some verb-final languages
have a corresponding post-core slot (PoCS). Some languages, again including
English, have a left-detached position (LDP) in which sentence-initial elements
can occur when separated from the clause by a pause, intonation break or comma (example (6)). Again there is a similar possibility after the clause, the right-detached position (RDP: example (6)).

(4) **What** did you leave behind? (BNC ASN 1841)

(5) **This one** I’m sure you’ll recognise. (BNC JK2 326)

(6) **As for him**, his heart was still set on finding the sea-king’s palace … (BNC FUB 43)

(7) She’s cunning, **that girl**. (BNC HTX 2808)

The structural elements of the clause are displayed in the form of a **CONSTITUENT PROJECTION**. There is also an **OPERATOR PROJECTION** showing the operators, for tense, illocutionary force, aspect, modality, etc, which attach to the various layers of the structure, and so show scopal relations. Finally, a **FOCUS PROJECTION** indicates the potential focus domain (PDF), within which focused elements must occur for a given language, and the actual focus domain (AFD) for a particular example. In Figure 3, all three projections are shown for the simple clause in Figure 2: ‘IU’ stands for ‘basic information unit’ (VAN VALIN, 2005, p.78).

The basis of the layered structure of the clause goes back to the first full presentation of RRG by Foley and Van Valin (1984). Three kinds of evidence for layering are adduced: the restricted scope of particular operators; the coding and behavioural properties of the units proposed; and the important role of the layered structure in explaining type of linkage between clauses in sentences. I shall examine each briefly in turn: for a rather fuller account, see Butler and Taverniers (2008, Forthcoming).

Operators for aspect (e.g. progressive in English) affect only the predicate itself, and so do those directional operators, in languages such as Kewa, which indicate the direction of the process itself: these operators constitute evidence for the nucleus. On the other hand, other types of directional, such as the prefixes *her-* and *hin-* in German, indicate the direction of movement of a participant in relation to the process. Similarly, root modalities (ability, obligation, intention) indicate some property of a participant in relation to whatever is signalled by the predicate. Such operators involve the whole core, without affecting non-core elements. Note, however, that this type of evidence assumes that operators are syntactic rather than semantic or pragmatic, a position which I shall argue against in a later section.
As far as coding properties are concerned, arguments in the core (e.g. what in most theories are labelled Subject and Object in English) are, cross-linguistically, normally coded by unmarked forms and those in the periphery (i.e. adjuncts) by marked forms, often adpositional. Further evidence comes from agreement and cross-referencing phenomena. English, for example, has verb agreement with the Subject but not with adjuncts, and in the Australian aboriginal language Gooniyandi arguments, but not adjuncts, are cross-referenced on the verb, as shown in examples (8) and (9), taken from McGregor’s (1990) corpus, in which the arguments girili and boojabij, but not the adjuncts ngilanggoowa and ngamoo nganyi marlami, show cross-referencing.5

5 Glosses for Gooniyandi examples have been provided with only the degree of detail necessary for our present purposes. The conventions are those of the Leipzig Glossing Rules, available at http://www.eva.mpg.de/lingua/files/morpheme.html
(8) (= McGregor’s 3-41, 1990, p.158)
girili wara-ari ngilanggoowa
tree stand:PRES:3SG:CLF eastern:end
‘The tree stands on the eastern end (of a row of trees).’

(9) (= McGregor’s 3-37, 1990, p.156)
gamoo nganyi marlami-ya ngarag-bidda\(^6\) boojabij
before 1SG nothing-LOC make:PST:3PL.NOM:3SG.ACC:CLF post:office
‘Before my time they built the old post office.’

The distinction between core and clause is motivated by, for example, the
behaviour of some Germanic languages with regard to verb position. In, for
instance, German, Dutch and Icelandic, the finite verb must appear in second
position in the clause, except in polar questions. Examples from German and
Dutch are given in (10) - (13):

(10) Gestern hab-e ich ein Bild
gestern have:PRES-1SG 1SG.NOM DET.INDF picture
ge-mal-t von der Mama, von mir
PTCP-paint-PTCP of DEF.F.SG.DAT Mama, of 1SG.DAT
und vom Clemens.
and of.DEF.M.SG.DAT Clemens
‘Yesterday I painted a picture of Mama, of me and of Clemens.’
(Peter Härtling …und das ist die ganze Familie: Tagesabläufe mit Kindern, cited as part of
the LIMAS corpus, Source no. 408, line 23, http://www.ikp.uni-bonn.de/Limas/)

(11) *Gestern ich habe ein Bild gemalt von der Mama, von mir und vom Clemens.

(12) Gisteren sprak-en Perez de Cuellar en Aziz, al
gisteren speak.PST-3PL Perez de Cuellar and Aziz already
vijf uur met elkaar.
five hour with each other
‘Yesterday Perez de Cuellar and Aziz spoke to each other for five hours.’ (Leiden Corpus
of Dutch, taken from ECI Corpus on CD-ROM, file DUT02A01)

(13) *Gisteren Perez de Cuellar en Aziz spraken al vijf uur met elkaar.

RRG explains these facts by positing that (except in yes-no questions) the finite
verb must be in second position in the clause, but that the Subject\(^7\) must be the first

\(^6\) The more modern spelling would be –birra rather than –bidda (Bill McGregor, pers. comm.).

\(^7\) In fact, as we shall see later, RRG does not use the traditional category of Subject, but rather that of the Privileged
Syntactic Argument (PSA) for a construction. For our purposes at present, however, we can continue to use the
more familiar label.
non-verbal element in the **core**. Thus when there is a constituent in the PrCS, inversion of Subject and verb must occur. Crucially, constituents in the LDP do not cause such inversion, as shown in examples (14) and (15) from German and Dutch respectively:

(14) Der Hans Sachs, der war
det.def.m.sg.nom hans sachs dem.m.sg.nom be.pst.3sg
ein Schuhmacher und Poet dazu …
det.indf.m.sg.nom shoemaker and poet with.it

‘Hans Sachs, he was a shoemaker and a poet too.’
(German Wikipedia corpus, id = 45776, available at http://corp.hum.sdu.dk/)

(15) Mijn zoon, ik heb hem het zestallig
1sg.poss son 1sg have.prs.1sg 3sg.m.obl det.def.n.sg six.number
stelsel uit<u>ge</u>leg-t, hij was pas vier
system <ptcp>explain-<ptcp> 3sg.m be.pst.3sg only four

‘My son, I explained the base six number system to him, he was only four.’

A second argument for the distinction between core and clause comes from the behaviour of head-marking languages (Nichols, 1984), in which dependency is marked on the head of the construction rather than on the dependent element. An example of such a language, as we have seen, is Gooniyandi. Consider an example such as (16):

(16) (= McGregor’s 5-241, 1990: 379)
gardiya gard-looni nganyi-ngga
white:man hit-pst:1sg.nom:3sg.acc:clf 1sg-erg

‘It was the white man that I hit.’

Note that both arguments are cross-referenced on the verb. Furthermore, the free NP arguments can be omitted, to leave simply *gardlooni* (McGregor, 1990, p.200). In RRG this is taken as evidence that the affixes on the predicate are the true core arguments, and that the omissible NPs are outside the core, though inside the clause.

A final important source of evidence for the layering of syntactic structure proposed in RRG is the syntax of clause combination. This is seen in terms of two cross-cutting dimensions, those of **juncture** and **nexus**. Juncture refers to the layer involved in the linkage (nucleus, core, clause), the default situation being that units of the same layer are linked; nexus refers to the type of linkage, coordinate, subordinate or cosubordinate, this last category referring to the
situation in which one clause is dependent on the other (in terms of the sharing of at least one operator), though not embedded within it. To the nine possible combinations generated by the three values on each dimension, all of which are said to occur in Korean (YANG, 1994), Van Valin and LaPolla (1997, p.469) add coordination of whole sentence structures containing LDP elements; in the latest version of the theory (VAN VALIN, 2005) there are also some refinements to subordination which need not concern us here. The combination of the theoretical constructs of unit layering and nexus type allows an explanatory account of the syntax of complex sentences which has been substantiated through analysis of a range of typologically diverse languages.

As the distinction between clause and core will be crucial to the arguments presented later, it is of interest here to look a little more closely at the difference between clausal and core junctures. For purposes of illustration I shall discuss only the coordination nexus type, but similar arguments apply to subordination and cosubordination types. First, it is important to note that coordination in RRG is not to be equated with conjunction of two units, this being just one of the ways in which the formal syntactic relation of coordination can be realised. Coordination is characterised by the lack of dependence between the units, while subordination involves structural dependence and cosubordination dependence between operators over the units. An example of clausal coordination is shown in (17): note the structural and operator independence, and the joining of two whole clauses.  

---

8 CLM stands for 'clause linkage marker'.

Examples (18) and (19) demonstrate that where whole clauses are the units involved in coordination, each can have its own independent illocutionary force:

(18) I hate myself for asking, but why did you leave? (BNC C9N 719)

(19) Forgive my asking, Sister, but why is he using a mask? (BNC CK0 370)

By way of contrast, (20) shows core coordination: again there is structural and operator independence, but here two cores are linked.
We have so far been concerned with evidence for the layering of syntactic units themselves. RRG grammarians also cite evidence for the attachment of operators to particular layers in the structure. The scope relations among operators lead to predictions, borne out in relation to typologically diverse languages, that operators will be expressed in an order, radiating out from the predicate itself, corresponding to the scheme nuclear < core < clausal. Note, however, that these facts could equally be taken to indicate semantic rather than syntactic scope: see later discussion.

2.3 Morphosyntactic structure in RRG and FDG: some parallels and differences

There is a considerable amount in the foregoing discussion which will be familiar to those with a knowledge of Functional Grammar. In particular, some of
the elements of the RRG layered structure are reminiscent of the special positions, labelled as P0-P3, proposed in FG in order to account for the location of elements with particular pragmatic functions. For instance, the LDP of RRG corresponds exactly to the P2 extra-clausal constituent proposed by Dik (1978, p.21), housing elements with the pragmatic function Theme, which “specifies the universe of discourse with respect to which the subsequent predication is presented as relevant” (DIK, 1978, p.19). Similarly, the RDP is equivalent to P3, housing a constituent with Tail function, which “presents, as an ‘afterthought’ to the predication, information meant to clarify or modify it” (DIK, 1978, p.19).

Within the clause itself, however, the mappings between FG and RRG are by no means exact. FG postulates a P1 position at the beginning of the clause, which is “used for special purposes, including the placement of constituents with Topic or Focus function” (DIK, 1997, p.408). Dik (1997, p.409) claims that this principle interacts with the typologically determined choice between ordering dependents before or after heads, and with the universal tendency for Subject to precede Object, to yield two basic word order patterns, P1 S O V and P1 V S O. Languages may have types of constituent which must be placed in P1: for English, Q-word constituents\(^9\) (interrogative pronouns and NPs with interrogative determiners), relative pronouns and subordinators fall into this category. Where P1 is not occupied by such a constituent, it may house elements with Given Topic or Focus function. An example of a constituent with contrastive Focus in P1 is given in (21):

\[(21) \text{This one I swapped with Christopher. (BNC KCT 7428)}\]

Since an element with Given Topic function is often the Subject, it follows that the Subject often occurs in P1.

Mackenzie and Keizer (1991), in a detailed examination of Topic and Focus in an English text, conclude that Dik’s analysis is untenable as far as Topic is concerned. They demonstrate that topical elements do not have any special treatment in English, and so, according to the principles of FG, the pragmatic function Topic cannot be assigned to them. This leads to the question of what use is made of P1 if no constituent specialised for that function is present, and Mackenzie and Keizer’s answer is to agree with Dik’s own rule, that in such circumstances this position is filled by the Subject.

The important point, then, in relation to English at least, is that there is incomplete correspondence between the FG P1 position and the PrCS of RRG, because of the stipulation that the Subject can go into P1 if there is nothing else that takes precedence over it for that position. The only circumstance in which

\(^9\) Except in echo questions, where they are in the normal position for their particular syntactic function.
the PrCS houses the Subject of an English clause is when that Subject is also a wh-constituent; otherwise, the default position is the first position within the core itself.

It should also be mentioned that a special position for Focus constituents, P0, has also been postulated to occur at the end of clauses in Polish, Czech and Bulgarian (DIK, 1997, p.426). Hannay and Martínez Caro (2008, Forthcoming) also present evidence for such a special position to account for certain focus phenomena in Spanish and English.

There are two problems for the F(D)G account in terms of syntactic templates with specified positions. Firstly, the templates proposed by Dik give us no way of recognising a syntactic unit which corresponds to the distinction between core and clause in RRG, which as we have seen is strongly motivated in terms of a number of types of evidence. Secondly, FDG inherits from FG a picture of constituent ordering in which structural templates make crucial use of the syntactic functions Subject and Object. Note, for instance, that de Groot (2005), in an article on morphosyntactic templates in FDG, still proposes clause templates containing Subject and Object. Likewise, Bakker’s dynamic model of the speaker (see e.g. BAKKER, 2001, 2005; BAKKER; SIEWIERSKA, 2004), which is the most detailed and cognitively-adequate account of FG expression rules available to date, still operates with templates involving the Subject and the P1 position. As we shall see, this is problematic in view of convincing evidence from RRG that Subject and Object functions are not well motivated cross-linguistically. Let us consider each of these points in turn.

2.4 The lack of a clause-core distinction in F(D)G

As shown earlier, the clause-core distinction in RRG is motivated by four types of evidence: constituent ordering in verb-second languages, the behaviour of core arguments in head-marking languages, the syntax of clause combination, and the fact that operators take a particular unit as the scope of their action. This last point will be discussed in a separate section later; here, I shall review briefly the other three types of evidence in relation to existing F(D)G accounts.

Even the early FG literature tackles the issue of constituent ordering in verb-second languages. For instance, Dik (1978, p.178-179) proposes a syntactic template for Dutch main clauses consisting of the sequence P1 Vf S O Vi, where Vf is the finite verb and Vi an infinitive, and P1 the initial position in the clause which must house certain types of constituent (including wh- and fronted items) if these are present. However, Dik goes on to comment that if there is no element which obligatorily goes into P1, then constituents with Topic or Focus function
may go there, and that since the Subject will often have one of these pragmatic functions, it is the main candidate for P1 position in such circumstances. There is thus still a blurring of the distinction between a core consisting of the predicate and its arguments, on the one hand, and a position before the core where wh- and fronted items must occur, on the other.

The account of constituent ordering by Connolly (1991, 2005), according to which no element is obligatorily filled, is likewise able to get constituents into the correct positions for particular languages, and represents a considerable improvement on the Dik scheme. For English (CONNOLLY, 1991, p.60-70, 2005, p.44), it recognises not only P1, but also a sequence of up to seven further ‘nuclear constituents’ (N1-N7). P1 houses only constituents which must go into this position (e.g. wh-items, subordinators) or, in the absence of such elements, a constituent with Focus pragmatic function. The unmarked position for the Subject in a declarative clause is N2, N1 being the position for the finite verb in constructions with Subject-finite inversion. This dissociation of P1 from Subject function potentially allows the recognition of a core grouping of elements in the clause, but Connolly does not go on to make this proposal, so that the model misses an opportunity to provide a unified account for the whole range of phenomena which motivate the recognition of the RRG core element.

As far as the syntax of head-marking languages is concerned, I am not aware of any F(D)G account which distinguishes between (i) a (‘core’) unit consisting of the predicate plus attached pronominal arguments, and (ii) the whole clause, in which there may be additional independent NPs which are coreferential with the pronominal elements.

Similarly, extant accounts of clause combination in F(D)G, partly because of their failure to recognise the clause-core distinction but also because the nexus category of cosubordination has no place in the theory, are unable to match the richness and elegance of the RRG account, where the matrix of juncture and nexus types gives rise to a principled set of clause relation types.

2.5 Syntactic templates and syntactic functions in RRG

In RRG as in F(D)G, patterns at the syntactic level are stored as templates (originally known as ‘constructional templates’, and now simply as ‘syntactic templates’), in what is known as the ‘syntactic inventory’ for a language (VAN VALIN; LAPOLLA, 1997, p.73, VAN VALIN, 2005, p.13). Templates are postulated for the various layers in the structure of the sentence. It is recognised that there is considerable variability in templates across languages: firstly, as we have seen, some languages do not have pre- or post-core slots or left/right detached
positions; secondly, some languages (e.g. English) impose particular orders on core elements, while others (e.g. Dyirbal) are very much more flexible. For English, Van Valin (2005, p.15) proposes templates for the PrCS and LDP, and six different templates for the core. In the construction of a sentence, appropriate templates are fitted together to form the final structure. Consider example (4), repeated for convenience as (22) below:

(22) What did you leave behind? (BNC ASN 1841)

This sentence consists of a single clause, with no element in the LDP. The clause requires the PrCS template, which itself contains the core; the core template needed is Van Valin’s Core-4, as shown in Figure 4:

Figure 4 – Templates in the construction of the structure for example 4/22
Note that the templates make no reference to functional categories such as Subject and Object. This reflects two characteristics of RRG: firstly, as noted briefly earlier, the traditional functions of Subject and Object are not recognised, and indeed it is postulated that there are languages which do not have syntactic functions at all; secondly, the syntactic function which corresponds to the Subject in English and many other languages is assigned by a general set of rules. Let us look at each of these features in a little more detail.

RRG and FG (and presumably FDG) agree in claiming that not all languages require the postulation of syntactic functions. The justification for postulating such functions in a language is that there are some phenomena in that language in which different semantic roles are neutralised for the purposes of the syntax. For instance, in English there is neutralisation of semantic roles as far as agreement on the verb is concerned. Consider examples (23) and (24):

(23) … our hearts have been smitten by the courage of athletes who come from Great Britain.
(BNC CAS 1067)

(24) … the courage of athletes who come from Great Britain has smitten our hearts.

In FG, it is the presence of voice alternations such as this in a language which determines whether syntactic functions need to be proposed or not. In RRG, because of its rather different categories, it must be demonstrated that the grammatical restrictions shown by the language cannot be stated simply in terms of semantic categories, together with reference to core or non-core status in the syntax. In order to discuss this further, we need to introduce a concept which will be treated in further detail later, that of SEMANTIC MACROROLES. RRG recognises two such roles, actor and undergoer, which generalise over sets of thematic roles which are treated as the same for particular purposes in the grammar: the prototypical actor is an Agent, while the prototypical undergoer is a Patient.

The agreement facts demonstrated by the above examples of the English active/passive alternation cannot be explained in this way, since finite verb agreement is with the undergoer of a transitive verb in (23), but with the actor in (24), so that this semantic opposition is neutralised for the syntactic purpose under discussion. There are languages, such as the Austronesian language Acehnese, in which all grammatical restrictions can be explained in terms of semantic categories, so that there is no need to postulate syntactic functions for this language.

Where syntactic function assignment is indeed needed, as in the case of the English active and passive alternation exemplified above, proponents of RRG argue that the traditional categories of Subject and Object run into considerable problems for some languages. Van Valin and LaPolla (1997, p.263-270), in a
discussion of two ergative Australian languages, Warlpiri and Dyirbal, demonstrate that although there is evidence for grammatical phenomena which require the assignment of a syntactic function in each language,\(^\text{10}\) the assignment of a blanket Subject function obscures differences in the alignment of syntactic function with the three different types of arguments which must be recognised in characterising ergativity, viz. the single argument of an intransitive verb, the actor of a transitive verb and the undergoer of a transitive verb. Furthermore, Van Valin (2005, p.94) points out that analysis in terms of the traditional Subject is problematic in Philippine languages.\(^\text{11}\) To avoid these difficulties, RRG postulates a different concept, that of the PRIVILEGED SYNTACTIC ARGUMENT OF A GRAMMATICAL CONSTRUCTION (PSA). PSAs come in two types, controllers and pivots (the latter normally relating to missing arguments in certain types of construction), and each of these can be either syntactic (as with the controller in the English passive, for example) or semantic (as in Acehnese). These details need not concern us here: the important point is that the PSA is defined relative to particular constructions in a language, and that “only syntactic pivots and controllers contribute to the generalized syntactic argument notion that is the traditional ‘subject’ in a language” (VAN VALIN, 2005, p.105).

RRG does not have any syntactic function which captures the notion of the traditional Object. Van Valin and LaPolla (1997, p.270-273) refer to work by Dryer (1986) which argues that in many accusative languages, the accusative NP does not have the syntactic and morphological properties associated with the traditional direct Object.

We are left with the question of why syntactic function, in the form of the PSA, does not figure in the syntactic templates proposed in RRG. The answer is that the PSA is assigned by general principles which make reference not only to the core status of arguments, but also to the semantic structure of sentences. It is, then, time for us to look briefly at how RRG describes the semantics.

3 Semantics and its mapping on to morphosyntax

3.1 The logical structures of sentences in RRG

The semantic structures proposed in RRG are based on the system developed by Dowty (1979) on the basis of earlier work by Vendler (1967), in which complex

---

\(^{10}\) Warlpiri, however, does not have a voice opposition (VAN VALIN; LAPOLLA, 1997, p.270), so that FG would be forced to conclude that it had no syntactic functions.

\(^{11}\) For discussion of voice in Philippine languages in relation to Subject assignment in FG, see Siewierska (1991, p.82-86, p.91-93).
predicates are decomposed into simpler elements, using abstract operators such as CAUSE and BECOME. At the heart of these structures are the LOGICAL STRUCTURES (LS) which form the main part of the lexical entries for predicates. Predicates are divided into six main classes (states, activities, achievements, semelfactives, accomplishments, active accomplishments) by means of a set of binary features as shown in Table 1. Static predicates represent states of affairs which do not involve ‘happening’; dynamic predicates express actions; telic predicates represent states of affairs which involve an inherent end-point; punctual predicates encode events which effectively occur instantaneously. These features can, in turn, be allocated by means of a battery of tests on a given sentence. In addition to the six main classes of predicate, there are causative variants of each.

<table>
<thead>
<tr>
<th>Predicate class</th>
<th>static</th>
<th>dynamic</th>
<th>telic</th>
<th>punctual</th>
</tr>
</thead>
<tbody>
<tr>
<td>State</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Activity</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Achievement</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Semelfactive</td>
<td>–</td>
<td>±</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Accomplishment</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Active accomplishment</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
</tbody>
</table>

Each major class has a distinctive pattern in its LS, as illustrated in (25) - (30) below:

(25) Bernice was tired. (BNC HTY 2826)
    tired´ (Bernice) [state]

(26) Miss Cress watched her, … (BNC AD1 1860)
    do´ (Miss Cress [watch´ (Miss Cress, her)]) [activity]

(27) The glass shattered. (BNC CKB 654)
    INGR shattered´ (glass) [achievement]

(28) Blue lights flashed. (BNC HTH 1766)
    SEML do´ (blue lights, [flash´ (blue lights)]) [semelfactive]

(29) Alfred Oliver had died. (BNC ANK 391)
    BECOME dead´ (Alfred Oliver) [accomplishment]

(30) They walked to the end of the road … (BNC BN1 1083)
    do´ (they, [walk´ (they)]) & INGR be-at´ (end of the road, they) [active accomplishment]
Achievements and semelfactives are both punctual (as shown by the operator INGR, for ‘ingressive’), the difference between them being that the former, but not the latter, result in a change of state. Elements such as dead’ or walk’ are not predicates of English, but rather are intended as placeholders for an eventual decomposition into cross-linguistically applicable semantic elements. A number of proposals have been made for more detailed semantic decomposition (VAN VALIN; WILKINS, 1993; VAN VALIN; LAPOLLA, 1997, p.116-118; MAIRAL USÓN; FABER, 2002, 2007; RUIZ DE MENDOZA IBÁÑEZ; MAIRAL, 2006).

RRG does not include thematic roles, such as Agent and Patient, as primitives of the theory, since such roles may be deduced from the positions of arguments in the LS. For example, the sole argument of a 1-place stative predicate is the PATIENT, the first argument of a pure location predicate is a LOCATION and the second a THEME, the first argument of a cognition predicate is a COGNIZER (itself a type of EXPERIENCER) and the second the CONTENT, and so on. However, as we saw briefly earlier, RRG does postulate generalised semantic roles, the macroroles of actor and undergoer. These roles are assigned to arguments in a LS by reference to the Actor-Undergoer Hierarchy shown in Figure 5 (VAN VALIN; 2005, p.61).

For any given sentence with a transitive verb, the argument which is furthest to the left in the hierarchy will be assigned actor status and the one which is furthest to the right will, in the default situation, be the undergoer. Thus in (26), the first argument of the do’ predicate, Miss Cress, is the actor, and the second argument of the activity predicate watch’ (her) the undergoer. In an intransitive construction, the sole argument may be either actor (as in (30)) or an undergoer (as in (29)).

3.2 Mapping from semantics to syntax in RRG: the role of the Privileged Syntactic Argument

While, as pointed out in the first section of this paper, FDG so far has no comprehensive account of how representations at the interpersonal (pragmatic)
and representational (semantic) levels are mapped on to the morphosyntax, RRG has a detailed algorithm for linking semantics and certain aspects of discourse pragmatics to syntax (and indeed, another algorithm for the reverse mapping, which will not concern us here). This algorithm has been developed using evidence from a wide range of typologically diverse languages, and the latest version is presented in Van Valin (2005, p.136-149), on which the following brief account is based.

The first stage in semantics-to-syntax linking is to construct the semantic representation for the sentence which, as we have seen, is based on the LS of the predating element. Actor and/or undergoer are then assigned according to the hierarchy in Figure 5. The next stage is to determine the morphosyntactic coding of the arguments. This involves, as a first step, the determination of the PSA, which is based on two sets of constraints. Firstly, PSA selection responds to a hierarchy which is identical to that for actor and undergoer assignment. Secondly, it obeys a set of principles which depend on the type of construction involved and/or the (type of) language. For accusative constructions, the default for PSA is the highest ranking core argument on the hierarchy in Figure 5, whereas for ergative constructions it is the lowest ranking argument. Also, there are languages (e.g. German, Italian, Dyirbal) in which only arguments with macrorole status can be PSA, and others (e.g. Icelandic, Japanese, Korean) in which the PSA can be assigned to non-macrorole direct core arguments. Finally, there are restrictions on the PSA in terms of coding: some languages (e.g. English, German) have case-sensitive PSAs, while others (e.g. Belhare, Tibetan) have PSAs which are case-insensitive.

Once the PSA has been assigned correctly the appropriate syntactic template can be selected, according to the following principle: “The number of syntactic slots for arguments and argument-adjuncts\textsuperscript{12} within the core is equal to the number of distinct specified argument positions in the semantic representation of the core” (VAN VALIN, 2005, p.130).

Note that only distinct specified argument positions count: if an argument position is specified more than once in the LS, it counts only once. This principle may be qualified by language-specific constraints: for instance, in English, all cores have a minimum syntactic valence of 1, certain constructions known as argument-modulation constructions (VAN VALIN, 2005, p.115-116) reduce the number of core slots by 1, and this last condition may override the first.

Finally, arguments are assigned to positions in the syntactic representation so far planned. Non-wh arguments are assigned to appropriate positions in the

\textsuperscript{12} Argument adjuncts occur in clauses such as those with put and a locative PP, where the PP is required by the verb and the preposition contributes an independent element of meaning to the clause, unlike the case with, for example, to after give, which simply marks the Recipient argument.
clause, and if there is a *wh* argument in the LS, it is assigned, according to the language concerned, to the normal position which a non- *wh* argument with the same function would occupy (‘pattern position’), or to the PrCS (as in English) or PoCS, or to some position with the potential focus domain of the clause, the unmarked position for focus in the language being the default. Optionally, a non- *wh* argument may be assigned to the PrCS or PoCS, provided that the structure conforms to the focus structure restrictions of the language. Finally, any non- *wh* arguments of the LS other than that of the main predicator in the nucleus are assigned to the periphery (the default), to the PrCS or PoCS, or to the LDP/RDP. All of these steps are subject to some degree of cross-linguistic variation.

An example of semantic-to-syntax linking is shown in Figure 6 for the example in (31):

(31) What did you give her? (BNC BNG 612)

Note that this sentence illustrates one of the language-specific rules for the assignment of the undergoer in English: with ditransitive verbs such as *give*, English allows undergoer assignment to either the Recipient (*her* in our example), or the Patient (as with *what* in (32)):

(32) What did you give to her?

The steps in the linking algorithm for a very similar example are discussed in more detail in Butler (2007).
Figure 6 – Semantics-to-syntax linking for example 31
3.3 The status of operators in FDG and RRG

FDG postulates that operators are required at all four levels of the grammar. At the interpersonal (pragmatic) level there are operators which effect grammatically-realised modifications of discourse moves, their component acts and the smaller constituents of which these acts are composed; at the representational (semantic) level, operators modify propositional contents, states of affairs, properties, individuals, etc.; at the morphosyntactic level we have secondary operators involved in morphological means of expression; finally, at the phonological level, we have phonological secondary operators such as those involved when the phonological structure is sensitive to syntactic organisation of a linguistic unit (Hengeveld; Mackenzie, 2006).

RRG, on the other hand, has a much smaller range of operator types, concentrating on those which are concerned with the grammatical expression of categories such as illocutionary force, tense, aspect, directionality, modality, evidentia, definiteness in the NP, and the like. Furthermore, as already noted, RRG treats all its operators as syntactic, though it is recognised (Van Valin, 2005, p.50) that they require semantic interpretation and that this is a complex matter.

FDG thus has the more sophisticated account of operators, and it seems reasonable to treat operators for illocution, evidentia, and the like as interpersonal, and so pragmatic in nature, while those for tense, aspect, etc. are representational, and so semantic, reserving morphosyntactic operator status to those grammatical modifications which occur at the morphosyntactic level itself. Note, for instance, that the restrictions imposed by operators at one level on those at a lower level are often predictable from the meanings involved: e.g. the fact that imperative illocutions require a controllable action. Note that the proposed shift from syntactic to semantic units for operator attachment is made possible by the fact that the universal units of the layered structure of the clause in RRG correspond to underlying semantic units: the syntactic nucleus contains the semantic predicate, core arguments represent arguments in the semantic representation of the predicate, the core contains the predicate plus its arguments, the periphery houses non-arguments, and the clause consists of the predicate plus arguments and non-arguments (see Van Valin; Lapolla, 1997, p.27; Van Valin 2005, p.5).

13 Nuyts (1992, 2001) has even proposed, within his Functional Procedural Grammar model, that distinctions of modality, evidentia, temporal and spatial modification, etc. should be seen as operating at an extralinguistic conceptual level, though with repercussions on the grammar.
4 The final reckoning: implications of RRG for FDG

FDG and RRG are extremely similar in their aims and many of their assumptions: both are concerned to show the relationships between form and function in language, and regard the syntax as (partially) semantically motivated; both place great emphasis on typological adequacy and also subscribe to principles of psychological/cognitive adequacy. It is therefore to be expected that ideas from one theory may well prove to be compatible with those from the other.

Although the discourse pragmatic and semantic levels of patterning are in some ways more highly developed in FDG than in RRG, FDG so far lacks a detailed account of lexical structure, and this is a particularly grave disadvantage for any theory with pretensions to cognitive adequacy, since it would seem that the syntactic component of language production is largely lexically driven (LEVELT, 1989, 1999). RRG provides such an account, through its postulation of logical structures in the lexical entries for predicates, and the semantic decomposition required is now being worked out in much more detail. Indeed, García Velasco and Hengeveld (2002), in an article which presages the development of FDG, actually suggest the adoption of abstract semantic decompositions of the type used in RRG. The discussion in the present paper suggests that such borrowing could profitably go even further, to embrace aspects of the morphosyntactic structure and semantics-to-syntax linking procedures.

First let us review the position with regard to the structural templates posited at the morphosyntactic level. We have seen that there are already equivalences between the LDP/RDP of RRG and the P2/P3 positions of F(D)G. Furthermore, the periphery of the clause in RRG is occupied by what were known in FG as satellites, and are referred to as lexical modifiers in FDG. The main differences are in the recognition of the core in RRG, and its separateness from the pre-core slot (or, in some languages, post-core slot). This distinction is strongly motivated by several kinds of evidence and, I have argued, needs to be reflected in FDG. We have seen that some elements which would go into the special clause-initial P1 position in F(D)G would be in the PrCS in RRG (e.g. wh-constituents and some ‘fronted’ elements with contrastive focus in English), but that although the Subject will often be placed in P1, it (or rather the PSA) does not go into the PrCS in RRG, except when it is a wh-item. This suggests that elements equivalent to the RRG core and PrCS/POCS should be recognised in FDG, and that in languages such as English the first NP in the core should be the default location for the Subject,14 rather than the P1 slot of the current model. We have seen that an adaptation of

14 In fact, it would seem advantageous for FDG to drop the label Subject in favour of something equivalent to the RRG PSA, given the arguments mentioned in §2.4.
Connolly’s (1991) proposals for English, which have the merit of dissociating P1 from the default position for Subject, could potentially achieve this aim. The syntactic templates which form part of the set of primitives available to the morphosyntactic level would then be reformulated in terms of the elements P2/LDP, P3/RDP, core and PrCS/PoCS, rather than in functional terms (i.e. containing S, O, as well as the positionally defined element P1). Furthermore, this proposal would require a realignment of the current position regarding the allocation of Focus and Topic constituents to positions in the clause: the focused subact of a discourse act could, for example, be placed in the PrCS-equivalent position (as in examples (21), (22), (31), (32)), while subacts with Topic function (in those languages which where topicalised elements have some overt reflex of their status in the grammar) would frequently end up mapping on to the core-initial position where this is the default for Subject/PSA.15

We have also seen that the adoption of the clause-core distinction would also allow FDG to give more detailed and insightful accounts of other aspects of syntax, such as predicate-argument relations in head-marking languages, and also clause combining. Crucially, we have here a set of fairly disparate phenomena, all of which can be elegantly accounted for in terms of a single, powerful proposal. If FDG were also to adopt the RRG concept of macroroles, which is again well motivated by cross-linguistic evidence, the scene would be set for adoption of some version of the semantics-to-syntax mapping algorithm already worked out for RRG. As we have seen, such a proposal has important advantages over current F(D)G work on morphosyntactic realisation, in that it makes use of the category of Privileged Syntactic Argument rather than Subject/Object, and has been thoroughly tested against a wide range of typologically diverse languages. Furthermore, this algorithm has been shown to be implementable in terms of the incremental processing which is widely agreed to occur in language production (BUTLER, 2007). On the other hand, RRG could benefit from closer attention to the stratification of operators proposed in FDG.

5 Conclusion

I hope to have shown in this paper that the morphosyntactic level in FDG is in need of considerable development, and that in the form presented so far in the FDG literature it fails to capture some important structural generalisations. On the other hand, proponents of RRG have succeeded rather better, a crucial element in their model being the distinction between the clause and the core which it contains. This distinction requires the recognition of a pre-core slot as well as a

---

15 For a comparison of focus in FG and RRG (and also Systemic Functional Grammar), see Butler (2005).
default position for the Subject in languages such as English, these being conflated into the P1 position in FG accounts of constituent order which have been taken over, for example, in the work of Bakker on expression rules. I conclude that proponents of FDG would do well to adopt some version of the distinctions made in RRG, a potentially promising avenue of exploration being an adaptation of the templates proposed by Connolly (1991, 2005), where, for English, P1 and the default Subject position are dissociated. Recognition of a core grouping of elements within the clause would also allow the adoption of the algorithms which map semantics on to syntax, and *vice versa*, in RRG, so providing a well worked out set of cross-linguistically validated rules for mapping meanings on to morphosyntax.

**Aknowledgments**

I am grateful to Lachlan Mackenzie and an anonymous reviewer for their comments on earlier drafts of this paper. Thanks also to Bill McGregor for supplying information on the Gooniyandi glosses.

References


Received May 2007
Approved October 2007