A new approach to clausal constituent order

Kees Hengeveld
University of Amsterdam

This chapter studies the consequences of the FDG approach to constituent ordering for the typology of constituent orders at the clausal level. After introducing the theoretical framework, it is argued that the fact that FDG in its dynamic approach to constituent ordering uses four absolute positions, which may be expanded with relative positions, leads to a situation in which each surface order in classical constituent order typology, such as VSO, SVO, etc. actually corresponds to 14 logically possible underlying orders. The chapter explores the consequences of this approach by comparing three V-medial and three V-initial languages, showing that each of these corresponds to a different type in the new approach advocated in this chapter.

1. Introduction[1]

In Functional Discourse Grammar (FDG, Hengeveld & Mackenzie 2008), constituent order is taken care of in a novel way in two different senses. First of all, the constituent ordering component of FDG makes use of templates that are dynamically constructed using at most four absolute positions: initial (P^I), second (P^2), middle (P^M) and final (P^F), which can be expanded by relative[2] positions when occupied by a constituent. Secondly, the placement of hierarchically higher constituents precedes the placement of constituents that are in a configurational relationship, which means that the place occupied by hierarchically higher constituents is crucial in deciding what the absolute positions relevant for a language are. In this chapter these two distinguishing features of the FDG approach to constituent ordering are used to set up a new typology of constituent orders at

1. I am greatly indebted to Lachlan Mackenzie and an external reviewer for comments on an earlier version of this chapter.

2. The notion 'relative position' was used earlier by Rijkhoff (2002) in a somewhat different sense, when he discusses the placement of constituents relative to a domain, relative to boundaries of a domain, and relative to constituents in domains. In this chapter relative means 'relative to a position that has already been filled in an earlier step during the ordering process'.

the clausal level. It will be shown that the availability of four absolute positions as well as the relative positions accompanying them leads to 84 logically possible constituent order patterns for transitive[3] clauses rather than the 6 possible patterns studied since Greenberg (1963). The chapter is organized as follows. Section 2 introduces the FDG approach to constituent order. Section 3 discusses classical constituent order typology and addresses the question whether the classification of clausal constituents in terms of S, O, and V is appropriate in a functional approach. Section 4 then presents the new classification of constituent ordering patterns that follows from the FDG approach. Section 5 illustrates the relevance of this classification by looking at three languages that are treated as V-medial and three languages that are treated as V-initial in the classical approach, but are of six different types in the new approach advocated here. The chapter is rounded off with conclusions in Section 6.

2. Constituent ordering in FDG[4]

2.1 Introduction

Hengeveld and Mackenzie (2008) propose a dynamic approach to morphosyntactic organization that consists of a number of steps taken by the Morphosyntactic Encoder, which together construct an underlying morphosyntactic representation at the Morphosyntactic Level. The main steps distinguished, not all necessarily relevant for every language, are the following:

i. Hierarchical ordering
ii. Non-hierarchical ordering
iii. Dummy-insertion
iv. Agreement

These steps apply at the Clause, Phrase, and Word Layers, in the order given here.

Only the first two steps are relevant for the present chapter, and both involve linearization. Both steps make use of dynamically constructed templates, which are introduced in Section 2.2. After that, Section 2.3 and Section 2.4 explain the processes of hierarchical and non-hierarchical ordering respectively.

---

3. Some languages have been claimed not to have transitive predicates at all (see e.g. Mosel 1991 on Samoan). For these a reduced classification of constituent order patterns would be in order.

4. This section is partly based on Hengeveld (2012).

2.2 Templates

The ordering process makes use of dynamically constructed templates. The construction process starts out from a number of absolute positions. Typological research has so far revealed that at least the initial (P^I), second (P^2), middle (P^M) and final (P^F) positions are potential starting points for the construction of templates.[5] These positions are crosslinguistically relevant, but are not all relevant for every language. The absolute positions relevant to a language have to be verified on a language-specific basis.

As soon as an absolute position is occupied, and only when it is occupied, the template is expanded with further relative positions. This is illustrated in (1):

(1)

P^I	P^{I+1}	P^{I+2}	etc.					
	P^2	P^{2+1}	P^{2+2}	etc.				
		etc.	P^{M-2}	P^{M-1}	P^M	P^{M+1}	P^{M+2}	etc.
					etc.	P^{F-2}	P^{F-1}	P^F

As (1) shows, P^I and P^2 may be expanded to the right, P^F to the left, and P^M to the right and the left. P^I and its expansions to the right may be called the *initial field*, P^2 and its expansions to the right the *secondary field*, P^M and its expansions to the left and to the right the *middle field* and P^F and its expansions to the left the *final field* (cf. Connolly 2012). A further important aspect shown by (1) is that the same superficial position may correspond to different absolute and or relative positions in the FDG framework. For instance, the superficial second position may correspond to P^{I+1}, P^2, and even to P^M or P^F in the appropriate circumstances. The superficial penultimate position may correspond to P^{F-1}, P^{M+n}, or even to P^I. Determining the relevant absolute or relative position corresponding to the surface position of a constituent is crucial to the approach taken here and will play an important role in what follows.

2.3 Hierarchical ordering

The process of hierarchical ordering involves the assignment of positions to elements (operators, modifiers) with higher scope before the assignment of positions to elements with lower scope. Operators capture grammatical elements, modifiers lexical elements. For instance, tense is an operator at the layer of the

5. It is interesting to note that while there is a wealth of evidence for the relevance of the absolute second position for constituent order typology, so far there is no evidence at all that points at the relevance of an absolute penultimate position, despite the relevance of the latter in phonology, for instance in stress assignment rules. Phonology might also be seen as working with templates.

episode, *frankly* is a modifier of the illocution. Modifiers are represented by Σ at the Interpersonal Level and by σ at the Representational Level, operators are represented by Π at the Interpersonal Level and by π at the Representational Level. The scope hierarchies are given in Figure 1. All the layers in Figure 1 may be provided with modifiers and/or operators.

Interpersonal Level	M	→	A	→	C	→	R	→	T
	↓								
Representational Level	p	→	ep	→	e	→	f^c	→	f^d

Figure 1. Hierarchical ordering

Three hierarchies are relevant and represented in Figure 1, in which the arrows point from elements with higher scope to elements with lower scope. Within each level (interpersonal and representational), a position is assigned to elements with higher scope before assigning a position to elements with lower scope. Across levels, a position is assigned to interpersonal units before assigning a position to representational units. The latter hierarchy is not argued for in Hengeveld and Mackenzie (2008) but is added here following Connolly (2012). I will illustrate the three different processes of hierarchical ordering with three different examples here.

First consider the following example, slightly adapted from a sentence found on the internet, which contains three different modifiers, of the communicated content (C), the propositional content (p), and the episode (ep), and two operators, of the episode (ep) and of the state-of-affairs (e).

(2) He reportedly (Σ^C) probably (σ^p) will (π^{ep}) not (π^e) play tomorrow (σ^{ep}).

The process of hierarchical ordering has to start with the highest operator or modifier. In this case this is the modifier *reportedly*. Since this constituent has to be assigned a place before the subject *he* is given a position, as this will only happen when non-hierarchical ordering starts, the modifier cannot be in P^{I+1}. It could only be in such a relative position if the absolute position had been filled at an earlier stage, which is impossible here, as *reportedly* itself is the first element to be assigned a position. As there is also no evidence for a P^2 position in English, as shown in Section 5.2.3, this highest modifier must be in P^M. The next modifier (*probably*) and the two operators (*will* and *not*) are then placed relative to the first one in P^{M+1}, P^{M+2}, and P^{M+3}, thus neatly reflecting their scopal order. Especially relevant is the placement of the modifier *tomorrow*, which occurs in the final position in the clause. This demonstrates the relevance of the P^F position in English, as the modifier *tomorrow* has to be assigned a position in the process

of hierarchical ordering; only after this is the verbal predicate *play* assigned a position in the process of non-hierarchical ordering. The modifier *tomorrow* can thus not be in a position relative to the verb, since at the point at which it has to be assigned a position, the verb is not there yet. As will be argued in Section 5.2.3, the predicate in English always occupies a position in the middle field. The subject is then in P^I. The final configuration is given in (3):

(3) P^I P^M P^{M+1} P^{M+2} P^{M+3}
He reportedly (Σ^C) probably (σ^p) will (π^{ep}) not (π^e)
P^{M+4} P^F
play tomorrow (σ^{ep}).

The next example, also adapted from a sentence on the internet, contains three different representational modifiers, of the episode (ep), the state-of-affairs (e) and the configurational property (f^c), and operators of the episode (ep) (expressed as the tense marker on the verb), the state-of-affairs (e), and the configurational property (f^c):

(4) The international gold price has (π^{ep}) been (π^e) rising (π^{fc}) continuously (σ^{fc}) again (σ^e) recently (σ^{ep}).

This example shows that for modifiers the hierarchical ordering at the representational level may start from P^F, which is the position occupied by *recently*, the modifier with the highest scope in (4). The next two modifiers in line are then assigned a position relative to the first one from right to left. They end up in the positions P^{F-1} and P^{F-2} respectively. The tense operator is assigned a position in the middle field. Since it is assigned a place before any non-hierarchical element it must be in the absolute P^M position. Further operators are then expressed relative to this position: the relative tense expression *been* is in P^{M+1}, the progressive *-ing* form is in P^{M+2}. Note that the tensed verb cannot be in P^{I+1} since it that case it would have been positioned after the subject had been located in P^I, which is impossible as the subject will receive its position later in the process of non-hierarchical ordering. Since the subject is in the initial field again, the full configuration is as in (5):

(5) P^I P^M P^{M+1} P^{M+2}
The international gold price has (π^{ep}) been (π^e) rising (π^{fc})
P^{F-2} P^{F-1} P^F
continuously (σ^{fc}) again (σ^e) recently (σ^{ep}).

For the hierarchical interaction across the interpersonal and representational levels, the vertical axis in Figure 1, consider the following examples:

(6) Perhaps (σ^p) I will (π^{ep}) accept that.

(7) That perhaps (σ^{p}) I will (π^{ep}) accept.

In (6) the order is as expected given the rules of hierarchical ordering at the representational level. The highest modifier, *perhaps*, goes to the absolute P^{I} position and therefore leaves the subject no choice but to go to P^{I+1}. The next highest operator, *will*, goes to P^{M}. It cannot be in P^{I+2} as the operator has to be assigned a position before any argument is placed. The predicate and object follow it in further relative positions, as indicated in (8):

(8)

P^{I}	P^{I+1}	P^{M}	P^{M+1}	P^{M+2}
Perhaps	I	will	accept	that

In (7), however, the object argument *that* occupies the absolute P^{I} position and has the highest representational modifier to its right. This means that a configurational element is assigned a position before a hierarchical element is placed. The reason for this is that *that* in (7) expresses contrast. Contrast is accounted for in FDG by means of a pragmatic function at the interpersonal level. In this case, therefore, the hierarchical relation across levels overrules the hierarchical relations within levels, which leads to the configuration in (9):

(9)

P^{I}	P^{I+1}	P^{I+2}	P^{M}	P^{M+1}
That	perhaps	I	will	accept

On the basis of these examples it may be noted that the process of hierarchical ordering leads to a situation in which the positions of modifiers and operators iconically reflect the scope relations between them. In this way it formalizes Rijkhoff's (2002) Principle of Scope.

2.4 Configurational ordering

In hierarchical ordering, which is based on considerations of scope, elements that are in a hierarchical relationship are assigned a position in a top-down fashion. In non-hierarchical ordering, which is based on alignment considerations, elements that are in a configurational relationship, such as a predicate-argument relation, are ordered on the basis of their pragmatic, semantic, and/or morphosyntactic properties, depending on the language under consideration.

In cases in which a language has a pragmatic alignment system, in the sense that pragmatic functions or other interpersonal factors determine order, these pragmatic considerations overrule all representational ones, even if hierarchical in nature, as shown in the previous section. Here hierarchical and configurational ordering thus interact strongly. A further good example of a language employing interpersonal configurational placement rules is Tzotzil (Aissen 1992, see also Hengeveld & Mackenzie 2008). Consider the following example:

(10) *A ti prove tzeb-e sovra ch'ak'bat.*
TOP DEF poor girl-TOP leftovers was.given
'It was leftovers that the poor girl was given.'

Tzotzil is generally considered to be a verb-initial language. However, if a Topic and a Focus constituent are present, the Topic occurs in the P^{I} position and the Focus in the P^{I+1} position. In the presence of a topic and a focus constituent the verb thus ends up in the P^{I+2} rather than the P^{I} position:

	P^{I}	P^{I+1}	P^{I+2}	P^{I+3}
(11)	A_ti_prove_tzebe^{TOP}	sovra^{FOC}	ch'ak'bat	---

Representational configurational placement rules are based on semantic factors such as semantic functions and animacy. In Movima (Haude 2006, see also Hengeveld & Mackenzie 2008), which has a hierarchical alignment system, the order in which the arguments are placed after the clause-initial predicate depends on their place on the animacy hierarchy (Silverstein 1976). The one highest on that hierarchy follows the predicate immediately, while the one lowest on the hierarchy follows afterwards, as illustrated in (12)–(13) (Haude 2006: 277):

(12) *Tikoy-na=sne os mimi:di.*
kill-DRCT=F.ABS ART.NML.PST snake
'She killed the/a snake.'

(13) *Tikoy-kay-a=sne os mimi:di.*
kill-INV-V=F.ABS ART.N.PST snake
'The/a snake killed her.'

Note that a shift in the semantic roles of the participants can only be indicated by using the direct (12)/inverse (13) marking on the verb, not by changing the word order.

The human participant in (12)–(13) outranks the non-human participant and thus immediately follows the verb in P^{I+1}.[6] The non-human participant goes to the next available position, P^{I+2}:

	P^{I}	P^{I+1}	P^{I+2}
(14)	Tikoy-na^{PRED}	sne^{HUM}	os_mimi:di^{ANIM}

When a language has morphosyntactic configurational placement rules the syntactic function or category of a constituent or its morphosyntactic complexity determines its position. For instance, in English the verb goes to the middle field just because it is the verb, the object has to follow it just because it is the object, and

6. The fact that *=sne* is a clitic is irrelevant here, as syntactically it is a constituent.

the subject goes to the initial field just because it is the subject, all of this, of course, unless overruled by hierarchical considerations. Thus, the simple English sentence in (15) fits into an ordering template as in (16):

(15) Paul read the book.

	P^{I}	P^{M}	P^{M+1}
(16)	$Paul^{SUBJ}$	$read^{V}$	the_book^{OBJ}

3. Classical constituent order typology

Before looking at the way in which the FDG approach to constituent ordering leads to a new classification of clausal word orders, I briefly go into some aspects of classical constituent order typology, mainly to straighten out some terminological issues.

Since Greenberg (1963) it has been generally assumed that the clausal constituent orders in Table 1 can be distinguished. The last three orders listed in this table are much less common than the first three, from which they differ in that the object precedes the subject rather than the other way around.

A basic problem with the approach visualized in Table 1 is that use is made of notions that are not crosslinguistically applicable. First of all, with many others (e.g. Falk 2006) FDG does not assume the universality of the grammatical relations of subject and object, and secondly, it does not assume the universality of verbs.

As for the non-universality of subjects and objects, it was shown in examples (12)–(13) from Movima that in this language order and other morphosyntactic properties have to make reference to animacy rather than to grammatical relations. The problem of the non-universality of grammatical relations in relation to constituent order was signalled in Rijkhoff (2002). As for the non-universality of verbs, I have shown in a number of publications (e.g. Hengeveld 2013) that several languages are better understood as having uncategorized lexemes, called 'contentives', that can be used in any function given semantic compatibility.

Given the non-universality of the notions used in Greenbergian constituent order typology I will replace the term 'verb', 'subject', and 'object' by 'predicate', 'actor', and 'undergoer', as illustrated for the six Greenbergian orders in Table 1. These are notions that are crosslinguistically valid and allow for an exhaustive classification of the languages of the world. This does not mean that verbhood, subjecthood or objecthood cannot be decisive factors in constituent ordering patterns, as

shown in 2.4. It just means that all languages can be captured by the terminology employed. Note that a predicate may be verbal or non-verbal in this approach, and that the terms actor and undergoer are used in the sense of Foley and Van Valin (1984). I will call the predicate, actor and undergoer 'nuclear constituents'.

Table 1. Greenbergian constituent orders

Constituent Order	Adapted terminology	Example language
V S O	Pred A U	Fijian
S V O	A Pred U	English
S O V	A U Pred	Turkish
V O S	Pred U A	Malagasy
O V S	U Pred A	Hixkaryana
O S V	U A Pred	Warao

4. A new approach to constituent order typology

The FDG approach to constituent ordering as summarized in Section 2 crucially affects the constituent order typology discussed in Section 3. The major question is to which positions Predicate, Actor, and Undergoer may be assigned. This question follows from the fact that FDG employs up to four absolute positions as well as many relative positions in its ordering templates, rather than the three positions that underlie classical constituent order typology. This gives a much larger range of potential orders. These will be introduced step by step.

Let me start with the situation in which A, U, and Pred all occupy an absolute position and in which the A precedes the U. This gives the logical possibilities listed in Table 2. As Table 2 shows, for every order in classical constituent order typology there are now three potential orders in this new typology. This has to do with the fact that there are distinct P^2 and P^M positions, rather than just one medial position. Note that P^2 only exists when P^I exists as well, such that P^2 can only contain a nuclear constituent when P^I contains a nuclear constituent as well. This excludes a number of orders that would otherwise be logically possible. Note that in this and the following tables I assume no hierarchical elements, which would have prior rights to absolute positions, are present. Furthermore, no distinction is made between situations in which a cell is empty because the relevant absolute position does not exist in the language in question, and those in which it is empty and is relevant for the language at hand but is not used for the

placement of non-hierarchical constituents. A further separation between these two situations would lead to an even richer classification.

Table 2. Absolute position only, A precedes U

P^I	P^2	P^M	P^F
Pred	A	U	
Pred	A		U
Pred		A	U
A	Pred	U	
A	Pred		U
A		Pred	U
A	U	Pred	
A	U		Pred
A		U	Pred

The same observation holds for orders in which A, U, and Pred occupy absolute positions and the U precedes the A, as shown in Table 3.

Table 3. Absolute positions only, U precedes A

P^I	P^2	P^M	P^F
Pred	U	A	
Pred	U		A
Pred		U	A
U	Pred	A	
U	Pred		A
U		Pred	A
U	A	Pred	
U	A		Pred
U		A	Pred

This does not exhaust the possibilities. As shown in the previous sections, A, U, and/or Pred may also occupy positions relative to one another. Let me start with situations in which two of the three nuclear constituents occupy a position within the same field, i.e. with one of these constituents being positioned relative to one of the other two. The possibilities are shown in Tables 4–9.

Table 4. Absolute and relative positions, Pred-A-U

P^I	P^2	P^M	P^F
Pred	A U		
Pred		A U	
Pred			A U
		Pred	A U
Pred A	U		
Pred A		U	
Pred A			U
		Pred A	U

Table 5. Absolute and relative positions, Pred-U-A

P^I	P^2	P^M	P^F
Pred	U A		
Pred		U A	
Pred			U A
		Pred	U A
Pred U	A		
Pred U		A	
Pred U			A
		Pred U	A

Table 6. Absolute and relative positions, A-Pred-U

P^I	P^2	P^M	P^F
A	Pred U		
A		Pred U	
A			Pred U
		A	Pred U
A Pred	U		
A Pred		U	
A Pred			U
		A Pred	U

Table 7. Absolute and relative positions, U-Pred-A

P^I	P^2	P^M	P^F
U	Pred A		
U		Pred A	
U			Pred A
		U	Pred A
U Pred	A		
U Pred		A	
U Pred			A
		U Pred	A

Table 8. Absolute and relative positions, A-U-Pred

P^I	P^2	P^M	P^F
A	U Pred		
A		U Pred	
A			U Pred
		A	U Pred
A U	Pred		
A U		Pred	
A U			Pred
		A U	Pred

Table 9. Absolute and relative positions, U-A-Pred

P^I	P^2	P^M	P^F
U	A Pred		
U		A Pred	
U			A Pred
		U	A Pred
U A	Pred		
U A		Pred	
U A			Pred
		U A	Pred

Finally, there is the possibility that A, U, and Pred all three occupy a position in the same field. This gives the further possibilities listed in Tables 10–15.

Table 10. A, U, and Pred in the same field, Pred-A-U

P^I	P^2	P^M	P^F
Pred A U			
		Pred A U	
			Pred A U

Table 11. A, U, and Pred in the same field, Pred-U-A

P^I	P^2	P^M	P^F
Pred U A			
		Pred U A	
			Pred U A

Table 12. A, U, and Pred in the same field, A-Pred-U

P^I	P^2	P^M	P^F
A Pred U			
		A Pred U	
			A Pred U

Table 13. A, U, and Pred in the same field, U-Pred-A

P^I	P^2	P^M	P^F
U Pred A			
		U Pred A	
			U Pred A

Table 14. A, U, and Pred in the same field, A-U-Pred

P^I	P^2	P^M	P^F
A U Pred			
		A U Pred	
			A U Pred

Table 15. A, U, and Pred in the same field, U-A-Pred

P^I	P^2	P^M	P^F
U A Pred			
		U A Pred	
			U A Pred

The overall result of the above is that, roughly equivalent to each of the six orders in classical constituent order typology, there are actually 14 logically possible underlying orders in the FDG approach. These are listed for convenience in Tables 16–21 in the classical order VSO, SVO, SOV, VOS, OVS, OSV.

Table 16. Pred A U (≈VSO)

P^I	P^2	P^M	P^F
Pred	A	U	
Pred	A		U
Pred		A	U
Pred	A U		
Pred		A U	
Pred			A U
		Pred	A U
Pred A	U		
Pred A		U	
Pred A			U
		Pred A	U
Pred A U			
		Pred A U	
			Pred A U

Table 17. A Pred U (≈SVO)

P^I	P^2	P^M	P^F
A	Pred	U	
A	Pred		U
A		Pred	U
A	Pred U		
A		Pred U	
A			Pred U
		A	Pred U
A Pred	U		
A Pred		U	
A Pred			U
		A Pred	U
A Pred U			
		A Pred U	
			A Pred U

Table 18. A U Pred (≈SOV)

P^I	P^2	P^M	P^F
A	U	Pred	
A	U		Pred
A		U	Pred
A	U Pred		
A		U Pred	
A			U Pred
		A	U Pred
A U	Pred		
A U		Pred	
A U			Pred
		A U	Pred
A U Pred			
		A U Pred	
			A U Pred

Table 19. Pred U A (≈VOS)

P^I	P^2	P^M	P^F
Pred	U	A	
Pred	U		A
Pred		U	A
Pred	U A		
Pred		U A	
Pred			U A
		Pred	U A
Pred U	A		
Pred U		A	
Pred U			A
		Pred U	A
Pred U A			
		Pred U A	
			Pred U A

Table 20. U Pred A (≈OVS)

P^I	P^2	P^M	P^F
U	Pred	A	
U	Pred		A
U		Pred	A
U	Pred A		
U		Pred A	
U			Pred A
		U	Pred A
U Pred	A		
U Pred		A	
U Pred			A
		U Pred	A
U Pred A			
		U Pred A	
			U Pred A

Table 21. U A Pred (≈OSV)

P^I	P^2	P^M	P^F
U	A	Pred	
U	A		Pred
U		A	Pred
U	A Pred		
U		A Pred	
U			A Pred
		U	A Pred
U A	Pred		
U A		Pred	
U A			Pred
		U A	Pred
U A Pred			
		U A Pred	
			U A Pred

5. An illustration

5.1 Introduction

It has taken several decades to discover instantiations of the six types of system Greenberg (1963) predicted in his six-way typology, especially due to the rarity of languages in which, using his terminology, the O precedes the S. Investigating the classification into 84 types presented in the previous section is therefore certainly not an option here. One important aspect of such an investigation would be to establish which of the logically possible orders are actually attested and which are not and to find explanations for possible restrictions. Since such an investigation is not within the scope of the current chapter, I will instead provide support for this fine-grained classification by presenting data from three different languages traditionally classified as V-medial and from three different languages traditionally classified as V-initial and by showing that in each case these actually represent three different types in the FDG classification.

5.2 Predicate-medial languages

5.2.1 *Introduction*

In this section four different Predicate-medial languages are presented and it is argued that they represent three different word order types in the FDG approach to word order typology.

5.2.2 *Dutch*

Example (17) illustrates the superficial SVO order of Dutch:

(17) *Peter las het boek.*
Peter read the book
'Peter read the book.'

By adding a hierarchical constituent in P^I, in this case the propositional modifier (σ^p) *waarschijnlijk* 'probably', it can be shown that the P^2 position is relevant for Dutch, as the finite verb has to stay in second position and the presence of the modifier forces to subject to go to a position after the verb:

(18) *Waarschijnlijk las Peter het boek.*
probably read Peter the book
'Probably Peter read the book.'

From this we may also derive that in (17) the actor is in P^I.

Modifiers may also appear after the verb in Dutch, as shown in (19), in which the hierarchical constituent *gisteren* 'yesterday' is a modifier of an episode (σ^{ep}) and *snel* 'quickly' a modifier of a state-of-affairs (σ^{e}).

(19) *Peter las gisteren snel het boek.*
Peter read yesterday quickly the book
'Peter read the book quickly yesterday.'

Since modifiers have to be placed first in the process of hierarchical ordering, they cannot be in a position relative to P^2, since the verb is placed later in non-hierarchical ordering. Therefore they have to be in the middle field, in the case of (19) in P^M and P^{M+1}. This also means that the undergoer cannot be in a position relative to P^2 but has to be in the middle or final field. Given that the modifiers may also follow the undergoer, the conclusion must be that the undergoer is in the middle field:

(20) *Peter las het boek gisteren snel.*
Peter read the book yesterday quickly
'Peter read the book quickly yesterday.'

In all, the conclusion must be that Dutch exhibits the system given in Table 22:

Table 22. Dutch

P^I	P^2	P^M	P^F
A	Pred	U	

5.2.3 *English*

The superficial SVO order of English is shown in (21):

(21) Peter read the book.

By adding a constituent in P^I it can be shown that English does not have a P^2 position, as all constituents shift to the right:

(22) Probably Peter read the book.

Modifiers may appear between the actor and the predicate in English, as shown in (23):

(23) Peter probably read the book.

This means that the verb is not in a position relative to P^I but in the middle field. This is so because the modifier *probably* cannot be in P^{I+1}, as the subject *Peter* will only be assigned a position later, during the operation of non-hierarchical ordering.

Unlike Dutch, it is impossible for a modifier to appear in between the predicate and the undergoer, as shown in (24):

(24) *Peter read probably the book.[7]

This means that the undergoer is in a position relative to the predicate, so in P^{M+N}.

The English constituent ordering facts may now be represented as in Table 23.

Table 23. English

P^I	P^2	P^M	P^F
A		Pred U	

5.2.4 *Leti*

The last predicate-medial language to be discussed here is Leti. Its superficial SVO order is illustrated in (25) (Van Engelenhoven 2004: 231):

(25) *Püata=e* *n-vèvla=e* *tani=la* *ròna=e.*
woman=EXCT 3.SG-forge=EXCT earth.EXCT=DIR vessel=EXCT
'The woman makes the vessel out of clay.'
"The woman forges the clay into a vessel."

Adding a constituent in P^I makes all constituents shift one position to the right, which means that Leti does not employ a P^2 position (Van Engelenhoven 2004: 208):

(26) *Apo* *püata-samtua=de* *n-vava* *upu* *Pui=o.*
SEQ woman-adult=DEM 3.SG-carry.name grandparent Pui=IND
'And the old lady was called Lady Pui.'

Modifiers and operators may not occur in between the actor and the predicate, nor in between the predicate and the undergoer. They have to be placed in P^I or P^F, as illustrated in (27) (Engelenhoven 2004: 221), with the locative modifier *lo Tutküèi* in P^I and the indicative clitic *o* in P^F:

(27) *Lo* *Tutküèi* *Solemaana* *n-vava* *Seli=o*
at Tutukei Salomon 3.SG-carry.name Seli=IND
'Salomon is called Seli in Tutukei.'

7. As pointed out to me by Lachlan Mackenzie it is possible to have a modifier between predicate and undergoer in English if the undergoer is lengthy, as in *He read quickly all the books that his uncle had recommended to him*. In this case the Undergoer can be said to be in P^F and is assigned that position on the basis of its weight. This also shows that languages can have multiple orders.

From these facts it may be concluded that the actor and undergoer are in positions relative to the predicate, and that the three of them are in the middle field, as indicated in Table 24:

Table 24. Leti

P^I	P^2	P^M	P^F
		A Pred U	

5.2.5 *Summary*

The three different systems discussed in the previous subsections compare as indicated in Table 25.

Table 25. Three predicate-medial languages

Language	P^I	P^2	P^M	P^F
Dutch	A	Pred	U	
English	A		Pred U	
Leti			A Pred U	

5.3 Predicate-initial languages

5.3.1 *Introduction*

In this section three different Predicate-initial (traditional V-initial) languages are presented and it is argued that they represent three different word order types in the FDG approach to word order typology.

5.3.2 *Scottish Gaelic*

The superficial VSO order of Scottish Gaelic is illustrated in (28) (Mackenzie 2009: 888):

(28) *Ghlac mi an cù an.dé.*
catch.PST.INDEP 1.SG DEF dog yesterday
'I caught the dog yesterday.'

Adding a constituent in P^I makes all constituents shift one position to the right, which means that Scottish Gaelic does not employ a P^2 position (Mackenzie 2009: 888):

(29) *Cha do=ghlac mi an cù an.dé.*
NEG.DECL DEP=catch.PST 1.SG DEF dog yesterday
'I did not catch the dog yesterday.'

Actors always follow the verb immediately, so that they must be in a position relative to P^I.

Operators may appear between the actor and the undergoer in Scottish Gaelic, as shown in (30), with the actor *e* and the undergoer *fìona*:

(30) *Tha* *e* *ag* *òl* *fìon-a.*
COP.PRS.INDEP 3.SG PROGR drink.NMLZ wine-GEN
'He is drinking wine.'

This means that there is a P^M position in Scottish Gaelic, as this hierarchically higher operator has to be assigned a position before Actor and Undergoer are placed. The Undergoer is in the middle field, as illustrated in (29) above.[8]

From these facts it may be concluded that the Actor is in a position relative to the predicate in the initial field, while the Undergoer is in the middle field, as indicated in Table 26:

Table 26. Scottish Gaelic

P^I	P^2	P^M	P^F
Pred A		U	

5.3.3 *Tzotzil*

The superficial V-initial order of Tzotzil is illustrated in (31) (Aissen 1987: 1):

(31) *7i-s-pet* *lok'el* *7antz* *ti* *t'ul=e.*
PFV-3-carry away woman the rabbit=CL
'The rabbit carried away the woman.'

Note that the predicate is followed by the Undergoer, which in turn is followed by the Actor. As illustrated in (10) above, repeated here as (32), the predicate may be preceded by hierarchically higher constituents, in which case it occupies a position relative to P^I.

(32) *A* *ti* *prove* *tzeb-e* *sovra* *ch'ak'bat.*
TOP DEF poor girl-TOP leftovers was.given
'It was leftovers that the poor girl was given.'

8. In contrast, a weak pronominal undergoer of a finite verb goes to the absolute final position in Scottish Gaelic (Lachlan Mackenzie, personal communication). It is unclear how this should be handled in the current approach. One possible explanation would be that its referentiality (an interpersonal operator) triggers early placement in P^F.

Tzotzil has a large set of second position clitics, one of them (=*xa*) being illustrated in (33) (Aissen 1987: 9):

(33) *7i-bat=xa li Xun=e.*
PFV-go=already the Xun=CL
'Xun has already gone.'

This shows that a P^2 position is relevant in Tzotzil. Since Undergoer and Actor immediately follow the second-position clitic when present, we may assume that they are in the P^2-field as well.

The P^F-position is relevant in Tzotzil too. For one thing, because there are sentential clitics that necessarily occupy the clause-final position, as illustrated in (33) with the clitic =*e*; for another, because modifiers go to the P^F-position, as illustrated in (34) (Cowan 1987: 20):

(34) *Tixk'éltik yán sénya ʔók'om.*
we.will.look.at another sign tomorrow
'We will look at another sign tomorrow.'

This shows that the clause-final position is relevant in Tzotzil as well, though not as a host of nuclear clausal constitituents.

In all, it may be concluded that the predicate is in the initial field, while the Undergoer and Actor are in the P^2-field, as indicated in Table 27:

Table 27. Tzotzil

P^I	P^2	P^M	P^F
Pred	U A		

5.3.4 *Kokota*

The superficial V-initial order of Kokota is illustrated in (35) (Palmer 2009: 279):

(35) *N-o fa-lehe=ri ago kokorako are.*
REAL-2.SG CAUS-die=3.PL.OBJ 2.SG chicken REM
'You are killing those chickens.'

Arguments with topic function may be placed before the predicate, in which case all other elements shift to the right, as illustrated in (36) (Palmer 2009: 286):

(36) *Ago n–o fa-lehe=au ara.*
2.SG REAL-2.SG CAUS-die=1.SG.OBJ I
'You are killing me.'

This shows that there is no P^2 position in Kokota.

Adjuncts normally go to the initial or final position, but in some cases they may also intervene between the predicate and the following arguments, as in (37) (Palmer 2009: 303):

(37) *N-a-ke* *lao* *buala* *ara.*
REAL-1.EXCL-PFV go Buala I
'I went to Buala.'

As adjuncts are hierarchically higher than arguments and therefore have to be assigned a position before predicate and arguments are placed, the adjunct in (37) must be in P^M. As a result, the arguments must be in the medial field as well.

In all, it may be concluded that the predicate is in the initial field, while the Actor and Undergoer are in the P^M-field, as indicated in Table 28:

Table 28. Kokota

P^I	P^2	P^M	P^F
Pred		A U	

5.3.5 *Summary*

The three different systems discussed in the previous subsections compare as indicated in Table 29.

Table 29. Three predicate-initial languages

Language	P^I	P^2	P^M	P^F
Scottish Gaelic	Pred A		U	
Tzotzil	Pred	U A		
Kokota	Pred		A U	

6. Conclusion

This chapter argues that the FDG approach to constituent ordering leads to a classification of constituent orders that covers 84 different types in a fine-grained system making use of absolute and relative positions. An exploration of a number of languages traditionally classified as belonging to the V-medial and V-initial types shows that these display subtle differences in their ordering possibilities that can be fruitfully described in terms of the many subtypes distinguished within the FDG approach. A large-scale typological investigation of many more languages will have to reveal which of the 84 types distinguished actually exist, and what the

typological restrictions and preferences are. For instance, is it a coincidence that in none of the six languages studied is the final field used for nuclear constituents? Is it a coincidence that all three predicate-medial languages use the middle field and all three predicate-initial languages use the initial field? It would have been logically possible for a predicate-initial language to have all three nuclear constituents in the middle field, or for a predicate-medial language to have, say, the actor and the predicate in the initial field and the undergoer in the secondary field. The small sample investigated thus yields interesting hypotheses for a large-scale investigation of clausal word order from the perspective defended in this chapter.

References

Aissen, Judith. 1987. *Tzotzil Clause Structure*. Dordrecht: Reidel.

Aissen, Judith. 1992. Topic and focus in Mayan. *Language* 51(1): 43–80.

Connolly, John H. 2012. The constituent ordering process in Functional Discourse Grammar. *Language Sciences* 34(4): 455–467.

Cowan, Marion. 1987. *Tzotzil grammar*. Norman OK: Summer Institute of Linguistics.

Engelenhoven, Aone van. 2004. *Leti, a Language of Southwest Maluku* [Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 211]. Leiden: KITLV Press.

Falk, Yehuda N. 2006. *Subjects and Universal Grammar. An Explanation Theory.* Cambridge: CUP.

Foley, William A. & Van Valin Jr, Robert D. 1984. *Functional Syntax and Universal Grammar*. Cambridge: CUP.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Human Language*, Joseph H. Greenberg (ed.), 73–113. Boston MA: The MIT Press.

Haude, Katharina. 2006. A Grammar of Movima. Ph.D. dissertation, Radboud University Nijmegen.

Hengeveld, Kees. 2012. Referential markers and agreement markers in Functional Discourse Grammar. *Language Sciences* 34(4): 468–479.

Hengeveld, Kees. 2013. Parts-of-speech system as a basic typological determinant. In *Flexible Word Classes. Typological Studies of Underspecified Parts of Speech*, Jan Rijkhoff & Eva van Lier (eds). Oxford: OUP.

Hengeveld, Kees & Mackenzie, J. Lachlan. 2008. *Functional Discourse Grammar. A Typologically-Based Theory of Language Structure*. Oxford: OUP.

Mackenzie, J. Lachlan. 2009. Aspects of the interpersonal grammar of Gaelic. *Linguistics* 47(4): 885–911.

Mosel, Ulrike. 1991. Transitivity and reflexivity in Samoan. *Australian Journal of Linguistics* 11(2): 175–194.

Palmer, Bill. 2009. *Kokota Grammar*. Honolulu HI: University of Hawai'i Press.

Rijkhoff, Jan N.M. 2002. *The Noun Phrase*. Oxford: OUP.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In *Grammatical Categories in Australian Languages*, R.M.W. Dixon (ed.), 112–171. Canberra: Australian Institute of Aboriginal Studies.