

General principles of linearization in Functional Discourse Grammar

Kees Hengeveld & Evelien Keizer

Abstract

This chapter proposes a number of refinements of the approach to linearization in Functional Discourse Grammar. It argues for an additional sentential layer at the Morphosyntactic Level, and shows how the introduction of such a layer may provide a means to account for the placement of, among others, backgrounded elements. It furthermore introduces a number of additional precedence principles in linearization that take into account the hierarchical structure of FDG, such as the precedence of predicates over arguments and the precedence of functions over operators over modifiers, while also recognizing the importance of a pragmatic and a semantic function hierarchy, as well as other pragmatic and semantic factors. Subsequently, the various innovations are applied to examples from languages with different basic constituent orders: two predicate-medial languages (English and Dutch, the latter a V2 language), a predicate-final language (Turkish), and a predicate-initial language (Tagalog). In doing so, we show that the modified approach allows us to account for a range of frequently attested linear patterns in these languages, while excluding unacceptable sequences.

Keywords

linearization, Functional Discourse Grammar, sentence, information structure, hierarchical ordering

1 Introduction

Functional Discourse Grammar (FDG) is exceptional in that it is a functional theory with a very specific system of placement rules, inspired by the overall organization of the model in the sense that linearization is dealt with in a top-down and dynamic fashion (Hengeveld and Mackenzie 2008: Chapter 4; Hengeveld 2013; Keizer 2015: Chapter 5). The approach is top-down in the sense that elements that are hierarchically higher at the Interpersonal and Representational Levels are assigned a position before hierarchically lower elements; in addition, hierarchically related elements are placed before non-hierarchically related (configurational) elements, such as predicates and arguments. The approach is dynamic in that, starting with a language-specific template with absolute positions, further relative positions are created dynamically relative to existing positions as these become occupied.¹

This approach is both powerful and flexible, and has been able to explain many (general, as well as language-specific) ordering tendencies and patterns. However, as it turns out, it may actually be too flexible, resulting in a certain degree of overgeneration. Our aim in this paper is to introduce further constraints on the current system of placement rules so as to increase its

¹ Note that we do not claim that the placement rules presented here represent the way utterances are produced by speakers (online); rather, they are meant to capture the functionally inspired general tendencies and constraints that apply within a language.

predictive power and its usefulness in understanding linearization patterns. The need to formulate such additional constraints is also acknowledged by Cinque (2023: 1) when he states: “The bewildering variation in word order among the languages of the world should not detain us from researching what, if anything, determines which orders are possible (and attested or attestable) and which orders are impossible (and not attested/nonattestable”.

In pursuing this aim, we will add rules for the placement of hierarchically related units at the same level (functions, operators, modifiers), as well as for the placement of non-hierarchically related units; in addition, we will consider the question of at which stage in the ordering process pragmatic and semantic factors play a role. For reasons of space, we will concentrate on linearization within the Linguistic Expression and the (main) Clause, but the changes we propose will also have wider application.

In Section 2 we will present the existing FDG approach, after which we propose a number of modifications to this approach in Section 3. Section 4 is dedicated to the ranking of the ordering principles proposed in Section 3. We put the principles and their ranking to the test in Section 5, by analysing examples from two predicate-medial languages (English and Dutch, the latter a V2 language), a predicate-final language (Turkish), and a predicate-initial language (Tagalog). Our conclusions are presented in Section 6.

2 The FDG approach to linearization

2.1 The current approach²

FDG distinguishes a number of levels of analysis: the Interpersonal, the Representational, the Morphosyntactic, and the Phonological Level. Only the first three of these are of interest to us in this paper. The Interpersonal Level is actional in nature and captures pragmatic distinctions that have a formal reflex in languages. The Representational Level is denotational in nature and captures semantic distinctions that have a formal reflex in languages. The Morphosyntactic Level captures the morphological and syntactic properties of languages. The Interpersonal and Representational Levels are levels of Formulation, an operation that determines what constitute valid underlying pragmatic and semantic representations in a language. The Morphosyntactic Level is produced by Morphosyntactic Encoding, an operation that concerns the rules that convert pragmatic and semantic representations into morphosyntactic ones. Linearization is an important part of the operation of Morphosyntactic Encoding.

The overall structure of the levels that are of interest here is given in (1), (2), and (3). In (1) and (2), the symbols Π/π represent operators, Σ/σ modifiers, and Φ/ϕ functions, while \diamond indicates a slot for lexemes.

² This section is partly based on Hengeveld (2013).

<p>(1) Interpersonal Level (IL)³</p> <p>(πM_1: [(πA_1: [(πF_1: ILL (F_1): $\Sigma (F_1)$) (πP_1: ... (P_1): $\Sigma (P_1)$)_S (πP_2: ... (P_2): $\Sigma (P_2)$)_A (πC_1: [(T_1)/(R_1) (Cm_1: [(πT_1: [...] (T_1): $\Sigma (T_1)$)_{Φ} (πR_1: [...] (R_1): $\Sigma (R_1)$)_{Φ}] (Cm_1)_{Φ}] (C_1): $\Sigma (C_1)$)_{Φ}] (A_1): $\Sigma (A_1)$)_{Φ}] (M_1): $\Sigma (M_1)$)_{Φ}</p>	<p>Move Discourse Act Interpersonal Property⁴ Speaker Addressee Communicated Content Any Subact Comment Ascriptive Subact Referential Subact Comment Communicated Content Discourse Act Move</p>
<p>(2) Representational Level (RL)⁵</p> <p>(πp_1: (πep_1: (πe_1: (πs_1: [(πf_1: $\blacklozenge (f_1)$: $\sigma (f_1)$) (πx_1: ... (x_1): $\sigma (x_1)$)_{Φ} (πt_1: ... (t_1): $\sigma (t_1)$)_{Φ} (πl_1: ... (l_1): $\sigma (l_1)$)_{Φ} etc.] (s_1): $\sigma (s_1)$) (e_1): $\sigma (e_1)$) (ep_1): $\sigma (ep_1)$) (p_1): $\sigma (p_1)$)</p>	<p>Propositional Content Episode State-of-Affairs Situational Property Lexical Property Individual Time Location Situational Property State-of-Affairs Episode Propositional Content</p>

³ We follow Smit (2010) in distinguishing a Comment layer within the Communicated Content. Note, however, that we do not adopt his proposal to regard Topic as a layer.

⁴ When a lexical element is used at the Interpersonal Level, the variable F (for Interpersonal Property) is used (cf. Giomi's [2021] Lexical Deed). The Illocution, the main (lexical or abstract) predicate at the Interpersonal Level, is also represented by this variable; see e.g. example (6) below.

⁵ The Situational Property is equivalent to what was called the Configurational Property in earlier publications.

(3)	Morphosyntactic Level (ML)	
	(Le ₁ :	Linguistic Expression
	(Cl ₁ :	Clause
	(Xp ₁ :	Phrase
	(Xw ₁ :	Word
	[
	(Xs ₁)	Stem
	(Aff ₁)	Affix
] (Xw ₁)	Word
	(Xp ₁)	Phrase
	(Cl ₁)	Clause
	(Le ₁)	Linguistic Expression

During linearization the units at IL (1) and RL (2) have to be mapped onto ML (3). In our view, this entails two steps: first, morphosyntactic templates have to be selected and combined, after which they have to be filled with material from IL and RL and possibly by dummies. In this paper we only focus on the second step, in which use is made of dynamically constructed templates. The basic template consists of a number of absolute positions. Typological research has so far revealed that at least the initial (P^I), second (P²), middle (P^M) and final (P^F) positions are potential starting points for the construction of templates. These positions are cross-linguistically relevant, but are not all relevant for every language; which absolute positions are relevant for the various templates of a given language can only be determined on a language-specific basis. As soon as an absolute position is occupied, the template may be expanded with further relative positions. This is illustrated in (4), which shows that P^I and P² may be expanded to the right,⁶ P^F to the left, and P^M to the right and the left:

(4)	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

Templates constructed in this way may be applied to Words, Phrases, and Clauses. When Clauses are part of Linguistic Expressions, the template is expanded by extracausal positions (precausal, interpolated,⁷ and postcausal). The maximal template for a Linguistic Expression, in terms of absolute positions, is given in (5) (but see Section 2.2).

(5)	[P ^{PRE}		P ^{INT}		P ^{POST}]	LinguisticExpression
	[P ^I	P ²	P ^M	P ^F]	Clause	

In filling the various positions, hierarchical ordering precedes non-hierarchical (configurational) ordering. The process of hierarchical ordering involves the assignment of positions to elements from IL and RL expressing functions, operators and modifiers, starting with those with the widest scope and then proceeding to those with lower scopes, according to the hierarchical relations represented in (1) and (2). In configurational ordering, which is based on alignment considerations, elements that are in a configurational relationship, such as a predicate-

⁶ Later in the paper, we will show that this formulation is too restrictive.

⁷ For arguments for introducing the interpolated position, see e.g. Keizer (2020).

argument relation, are ordered on the basis of their pragmatic, semantic, and/or morphosyntactic properties, depending on the language in question. After the placement of hierarchical and configurational elements, dummies are added (in those languages that have them) to fill obligatory morphosyntactic positions for which no material is available from IL and RL.

Example (6) may serve to illustrate these various steps.

	p^I	p^M	p^{M+1}	p^{F-1}	p^F
(6)	<i>It</i>	<i>frankly</i>	<i>rained</i>	<i>constantly</i>	<i>yesterday.</i>
IL:	(A _i : [(F _i : DECL (F _i): (F _j : frankly (F _j)) (F _i) (P _i) _S (P _j) _A (C _i : [(C _m _i : [(T _i) (T _j) (R _i)] (C _m _i)) _{FOC}] (C _i))				
RL:	(p _i : (past ep _i : (e _i : (s _i : [(f _i : rain (f _i))] (s _i)) (e _i): (f _j : constantly (f _j)) (e _i)) (ep _i): (t _i : -yesterday- (t _i)) (ep _i)) (p _i))				
ML:	(Le _i : [(Cl _i : [(Np _i : it (Np _i)) (Advp _i : -frankly- (Advp _i)) (Vw _i : -rained- (Vw _i)) (Advp _j : -constantly- (Advp _j)) (Advp _k : -yesterday- (Advp _k)) (Cl _i))] (Le _i))				

Starting with the process of hierarchical ordering, the illocutionary adverb *frankly*, as the highest modifier in (6), is the first element to be assigned a position, going to the clause-medial position. The next elements to be placed are the past tense operator and the adverb *yesterday*, which both operate at the layer of the Episode. The adverb functions as the head of a t-variable, which is used for temporal entities. The past tense marker goes to P^{M+1} , where it will later be joined by the verb, and the adverb goes to P^F . Next in line in hierarchical ordering is the adverb *constantly*, which modifies the State of Affairs. Since the adverb *yesterday* has gone to P^F , a new relative position P^{F-1} is now available to host the adverb *constantly*. This completes the process of hierarchical ordering. In configurational ordering, the only element present in the underlying RL representation is the verb *rain*, which joins tense in P^M . In the resulting configuration, the P^I subject position, which is an obligatory slot in English, is still empty, and therefore has to be filled by the dummy *it*.

2.2 A modification

The current version of the FDG Morphosyntactic Level does not have a layer in between the Linguistic Expression and the Clause. We feel that this is problematic, as there are cases in which individual Clauses combine into larger units, which may exhibit syntactic behaviour that is different from that of the individual Clauses that constitute it. We therefore propose that a new morphosyntactic layer, that of the Sentence, be introduced. To show the need for this new layer, let us consider the second position that is relevant to two of the languages that we study in this paper: Dutch and Tagalog.

Unlike what has generally been assumed so far, we would like to argue that in Dutch main clauses the second position is the second sentential position rather than the second clausal position.⁸ The examples in (7)-(8) demonstrate this.

(7)	<i>Peter</i>	<i>was</i>	<i>ziek.</i>
	Peter	COP.PST.3.SG	ill
	'Peter was ill.'		

⁸ Note that we use the capitalized terms *Clause* and *Sentence* when referring to the layers of the Morphosyntactic Level in FDG, while we use the non-capitalized terms *clause/clausal* and *sentence/sentential* in a general descriptive sense.

- (8) *Gisteren was Peter ziek.*
 yesterday COP.PST.3.SG Peter ill
 'Yesterday Peter was ill.'

In (7) the subject precedes the verb, which has to be in second position. When a constituent other than the subject occurs in initial position, the subject occurs in postverbal position, such that the verb can stay in second position, as shown in (8). Subordinate clauses, on the other hand, do not have a second position, as shown in (9):

- (9) *Ik dacht dat Peter (gisteren) ziek was.*
 1.SG think.PST.1.SG CONJ Peter (yesterday) ill COP.PST.3.SG
 'I thought that Peter was ill (yesterday).'

In subordinate clauses the verb occupies the clause-final position, and the second position is irrelevant for the placement rules. This strongly suggests that in Dutch the absolute second position is a sentential rather than a clausal position.

In Tagalog the situation is different. Tagalog clause structure exhibits a second position that hosts a special set of clitics. These clitics occupy the second position in both main and subordinate clauses, the latter showing that they are not occupying a sentential, but a clausal position. This is shown in (10), in which both the main clause and the subordinate clause contain second position clitics.

Tagalog (Kroeger 1998: 3)

- (10) *Sinabihan=ako ni=Luz na [ibinigay=mo=na ang=pera kay=Charlie].*
 PRF.say.DV=1.SG.NOM GEN=Luz CONJ
 IV.PRF.give=2.SG.GEN=already NOM=money DAT=Charlie
 'I was told by Luz that you already gave the money to Charlie.'

The clausal nature of the clitics in Tagalog is furthermore reflected in the fact that, as we will demonstrate later, sentences such as (10) may be preceded by other sentential constituents, such that the second clausal position effectively corresponds to the third sentential position. This is not possible in Dutch.⁹

We therefore propose that the representation in (5) be enriched by an additional sentential layer with its own pre-, post- and interpolated positions (PRESNT, POSTSNT and INTSNT), as shown in (11). Within the sentence, we find the clause, as well as a preclausal (PRECL), postclausal (POSTCL) and interpolated clausal (INTCL) position.

- (11) [P^{PRESNT} [P^{PRECL} [P^{INTSNT} [P^{INTCL} [P^{POSTCL}]_{Sent}]LingExpr
 [P^I P² P^M P^F]_{Clause}

Given that languages make use of a subset of this maximal set of positional possibilities, we could hypothesize that in Tagalog clitics occupy the P² position in the Clause, whereas in Dutch, which does not have a clausal second position, the second position constituent, i.e. the finite verb in

⁹ With a few exceptions; see discussion of example (34) below.

main clauses, goes to the first Clausal position, which effectively corresponds to the second Sentential position. In other languages, e.g. English and Turkish, the first Clausal position is typically reserved for the subject.

The interpolated positions in (11) do not correspond to fixed positions within the Clausal and Sentential templates but may interrupt the Clause and Sentence at various points. We use the interpolated Sentential position for prosodically independent interpolated constituents, while the interpolated Clausal position is now used for prosodically integrated, backgrounded interpolated constituents. This is illustrated in examples (12) and (13): in (12) the illocutionary adverb *frankly* is interpolated at the layer of the Sentence (cf. Keizer 2018), in (13) at the layer of the Clause.

(12) *And he COVered, FRANKly, a LOT of ground when he was in that BRIEging room (COCA, spoken) (where frankly is prosodically detached and has its own pitch accent)*

(13) *JOHN frankly has NO iDEA (where frankly is prosodically integrated and deaccented)*

The modification proposed allows for a number of further refinements of the analyses within the FDG linearization approach that we will present below. Anticipating the results, we will show that the preclausal and postclausal positions within the Sentence may host a variety of pragmatically marked constituents, such as emphasized Settings in Dutch and English, Topics in Dutch and Tagalog, and Background constituents in English and Turkish.

Note that (11) also contains a representation for the layer of the Linguistic Expression, which itself exhibits presentential, postsentential, and interpolated positions (corresponding to the pre-clausal, post-clausal and interpolated positions so far distinguished in FDG). In what follows, however, we will concentrate on the newly conceptualized extra-clausal positions.

3 Restrictions on linearization

3.1 Introduction

As mentioned in Section 1, the approach outlined above, although very powerful, is in need of revision and refinement. In this section, we will suggest some ways in which the placement rules can be adapted and expanded to avoid them from overgenerating. After that, in Section 4, we will discuss how the different principles proposed should be ordered.

3.2 Hierarchical and configurational ordering

A first refinement we want to propose concerns the relationship between hierarchical and configurational ordering. In Hengeveld and Mackenzie (2008) (i) all hierarchical ordering precedes all configurational ordering, the latter being restricted to RL, and (ii) the expression of Illocution is treated as part of hierarchical ordering. These two steps are not consistent with the way Illocution is represented in FDG, namely as a predicate taking the Speech Participants and the Communicated Content as its arguments (see e.g. Hengeveld and Mackenzie 2008: 70). When the Illocution is taken as a predicate, configurational ordering is also relevant within the head of the Discourse Act, with the Communicated Content as well as the two Speech Participants functioning as arguments, just like any argument at RL. This means that the process of hierarchical ordering is in a sense interrupted at the Interpersonal Level. Compare this to the situation in which a predicate

at RL takes a complement clause as its argument. In that case, too, configurational ordering applies first between the predicate and its arguments, and only after that does hierarchical ordering within the complement clause take place. All this is shown in the following overview of hierarchical and configurational rules that we propose.

Hierarchical (IL before RL)

I. IL

1. Hierarchical (higher functions, operators, modifiers before lower ones)

$(\Pi M_i: (\Pi A_i: [\dots] (A_i): \Sigma (A_i))_\Phi (M_i): \Sigma (M_i))_\Phi$

2. Configurational (predicate before arguments)

$[(F_i: ILL (F_i)) (P_i)_A (P_j)_S (C_i)]_\Phi$

3. Hierarchical (higher functions, operators, modifiers before lower ones)

$(\Pi F_i: ILL (F_i): \Sigma (F_i))$

$(\Pi P_i: [\dots] (P_i): \Sigma (P_i))_S$

$(\Pi P_j: [\dots] (P_j): \Sigma (P_j))_A$

$(\Pi C_i: [\dots] (C_i): \Sigma (C_i))_\Phi$

II. RL

4. Hierarchical (higher functions, operators, modifiers before lower ones)

$(\pi p_i: (\pi ep_i: (\pi e_i: (\pi s_i: [\dots] (s_i): \sigma (s_i) (e_i): \sigma (e_i)) (ep_i): \sigma (ep_i)) (p_i): \sigma (p_i)))$

5. Configurational (predicate before arguments)

$[(f_i) (\alpha_i)_{\Phi^n}]$

6. Hierarchical (higher functions, operators, modifiers before lower ones)

$(\pi f_i: \blacklozenge (f_i): \sigma (f_i))$

$(x_i)_\Phi (t_i)_\Phi (l_i)_\Phi$, etc.

As shown here, hierarchical ordering first applies at the layers of the Move and the Discourse Act, and then comes to a halt when the configuration that fills the head position of the Discourse Act (consisting of the Illocution, the Participants, and the Communicated Content) is reached. At this point the configurational ordering of these units takes over. Then, within each of the four configurational units, hierarchical ordering takes place again. Once this process has been completed, hierarchical ordering at RL begins, until this comes to a halt again, as the units making up the Situational Property become the target of configurational ordering. Subsequently, within each of these units hierarchical ordering takes place again.

This recursive process can be captured by the following principles:

1. Interpersonal > Representational
2. Hierarchical > Configurational (recursive)
3. Hierarchically higher > Hierarchically lower

3.3 Predicates before arguments

It has been shown in the FDG literature (e.g. Fang and Hengeveld 2022; Hengeveld 2004; Keizer 2018) that grammatical expressions of Illocution and lexical expressions of Illocutionary modification are placed earlier than elements of the Communicated Content. Now that we have changed the treatment of Illocutions, we need a different way of explaining this preferential placement. The solution we propose is that, both at IL and at RL, predicates are placed before arguments. This makes sense, as predicates, despite not having scope over their arguments, do govern their arguments in other ways. Consider the following examples:

- (14) *He must be ill.*
(15) *He is certainly ill.*
(16) *I am sure he is ill.*

In these examples the Propositional Content ‘he is ill’ is specified in similar ways: in (14) through an operator, in (15) through a modifier, and in (16) through a higher predicate. It thus seems to make sense to have predicates receive a place before their arguments do, not because they have a higher scope in the technical sense, but because their arguments are dependent on them. So since the predicate can have a similar function to operators and modifiers, it should be placed earlier than its arguments, much like operators and modifiers are placed earlier than their heads. Note that this principle applies within the configurational part of the ordering process. The principle can be given as:

4. Predicate > Argument

3.4 Functions before operators before modifiers

In hierarchical ordering, functions are expressed before operators and modifiers, since they are external to the entire layer, thus scoping over everything within that layer (Hengeveld and Mackenzie 2008: 311). For the expression of operators and modifiers, however, no specific ordering has been proposed. We argue that these should also be ordered with respect to each other, such that functions are expressed first, then operators, and then modifiers.

The idea that functions should be expressed first directly follows from the way these are represented in FDG.

- (17) $(\pi \alpha_1: \text{head}(\alpha_1): \sigma(\alpha_1))_\phi$

The function is attached to the outer brackets, thus scoping over the entire unit, including the operators and the modifiers pertaining to (α_1) . Consider (18):

- (18) a. *from the high mountains*
b. $(m \ x_i: (f_i: \text{mountain}(f_i)(x_i)): (f_j: \text{high}(f_j)(x_i)))_{\text{Abl}}$

In (18b), we have a function (Ablative, expressed as *from*), an operator (for plurality), and a modifier (*high*). It would not make sense to say that the function pertains, for instance, just to

mountain and that identifiability, plurality and the modifier *high* are then added to that combination. Instead, the entire noun phrase has the function Ablative.

Operators should be expressed next, before modifiers of the same layer. First of all, because this makes sense semantically. For instance, in a phrase like *three blue balloons*, we do not say that we have a set of three balloons, which are then assigned the property of being blue: in that case 'blue' would no longer function as a restrictor, since we already know the size of the set. Instead, we have a set of objects whose members are subsequently restricted by assigning the properties 'balloon' and 'blue'. Finally, the size of the set is restricted to three members. Secondly, there is also grammatical evidence that operators have scope over a head and its modifier, as the head and modifier can be replaced by a pronoun (19):

(19) *this blue balloon and that one*

Taken together, these facts reflect a further principle, that may be formulated as:

5. Functions > Operators > Modifiers

3.5 Pragmatic factors

Pragmatic factors play an important role in the placement of constituents: as we have seen, the Interpersonal Level is the highest level in the top-down architecture of FDG, reflecting the idea that pragmatics governs semantics, and pragmatics and semantics together govern morphosyntax. Thus, unlike in generative approaches, where pragmatic factors are assumed to lead to changes in a basic order, in FDG, which does not allow for transformations, pragmatics factors are an integral part of the theory. We propose here that these factors, despite being specified within the Communicated Content at IL, come into play at the moment that the ordering of the relevant elements at RL starts out. The reason for this is that in most cases all the elements contained in the Communicated Content at IL are lexically instantiated at RL only.¹⁰ We suggest that these elements await placement until the process of representational placement starts. This means, for instance, that pragmatic functions of constituents will only be expressed once the lexical means to express those constituents have been selected.¹¹ The priority given to pragmatic information is then reflected in the fact that constituents with a pragmatic function are placed before those without a pragmatic function.

Within the group of pragmatic functions, we have reason to believe that Topic takes precedence over Focus. Consider the following example from Tzotzil (Aissen 1992; see also the discussion in Hengeveld and Mackenzie (2008: 334)). Here we see that if a Topic and a Focus constituent are present, the Topic occurs in the initial position and the Focus in the post-initial position. The language is otherwise predicate-initial, so the Topic and the Focus constituents push the verb to the right, such that it now ends up in P^{I+2}:

¹⁰ Proper names (Hengeveld and Mackenzie 2008: 117)) and placeholders such as *thingummy* (Hengeveld and Keizer 2011) are exceptions to this.

¹¹ Note that the same holds for the placement of information within the Subacts, since their operators and modifiers do not take clausal positions. In other words, all this information has to be put on hold until the representational unit corresponding to the Subact has been assigned a position (cf. Giomi, this volume). At that point the general rule again applies: functions are placed before operators, which are placed before modifiers.

Where factors like these play a role in languages, a principle like the following might be relevant (cf. Dik 1997: 37):

8. Definite (+id/+s) > Indefinite specific (-id/+s) > Non-specific (-id/-s)

Finally, we follow Keizer (2014) in regarding Subject function as being triggered by a Perspective operator on a Referential Subact at IL, constituting a pragmatic factor.¹³ In certain languages the presence of this operator may trigger prioritization of placement of an argument in the linearization process. For instance, in Kinyarwanda the Subject of a sentence is in initial position. This happens irrespective of the semantic function of the argument, the influence of which will be discussed in Section 3.6 below. Thus, in the following examples an Actor, Undergoer and Recipient occupy the first position in the sentence due to the fact that they carry the perspective operator.

Kinyarwanda (Keenan and Dryer 2007: 349)

- (24) a. *Umugabo y-a-haa-ye umugóre igitabo.*
 man 3.SG-PST-give-ASP woman book
 'The man gave the woman the book'
- b. *Umugóre y-a-haa-w-e igitabo n-ûmugabo.*
 woman 3.SG-PST-give-PASS-ASP book AG-man
 'The woman was given the book by the man'
- c. *Igitabo cy-a-haa-w-e umugóre n-ûmugabo.*
 book 3.SG-PST-give-PASS-ASP woman AG-man
 'The book was given to the woman by the man'

The following principle accounts for facts like these:

9. Perspectivized > Not perspectivized

3.6 Semantic factors

Further linearization rules within configurational ordering at RL may be determined by semantic factors. Placement may be sensitive to the semantic functions of arguments, as in the following example:

Turkish (Kornfilt 1997: 90, see also Hengeveld and Mackenzie 2008: 336)

- (25) *Hasan-Ø kitab-ı Ali-ye ver-di-Ø.*
 Hasan-NOM book-ACC Ali-DAT give-PST-3.SG
 'Hasan gave the book to Ali.'

The initial Subject of this sentence is placed according to the previous principle, but for the remaining constituents their semantic functions are relevant, in the sense that, barring pragmatic factors, the Undergoer (U) precedes the Recipient (which is a subtype of L(ocation)). In cases like these the following placement principle applies:

¹³ Hengeveld, Keizer and Giomi (in prep.) furthermore propose a secondary perspective operator.

10. A > U > L

In other cases, placement may be sensitive to person and/or animacy features of arguments. This is illustrated in the following example:

Plains Cree (Wolvengrey 2005: 423, see also Hengeveld and Mackenzie 2008: 321):

- (26) a. *Ni-wîcih-â-nân-ak.*
1-help-DRCT-1.PL-3.PL
'We help them.'
b. *Ni-wîcih-iko-nân-ak.*
1-help-INV-1.PL-3.PL
'They help us.'

In Plains Cree a first person argument always has to precede a third person argument in linear order. The semantic functions of these arguments can then be read off from the fact that the verb is marked for direct (first linear argument is Actor) or inverse (second linear argument is Actor).

A placement principle such as the following might capture facts like these (Dik 1997: 37):

11. 1st/2nd > 3rd human > 3 animate > 3 inanimate force > other inanimate

4 Ranking the linearization principles

The principles discussed in Section 3 are the following:

1. Interpersonal > Representational
2. Hierarchical > Configurational
3. Hierarchically higher > Hierarchically lower
4. Predicate > Argument
5. Functions > Operators > Modifiers
6. Topic > Focus > No pragmatic function
7. Contrast > No contrast
8. Definite (+id/+s) > Indefinite specific (-id/+s) > Non-specific (-id/-s)
9. Perspectivized > Not perspectivized
10. A > U > L
11. 1st/2nd > 3rd human > 3 animate > 3 inanimate force > other inanimate

The above rules would give different results when ranked in different ways. We propose the following overall ranking:

1. Interpersonal > Representational

- > Interpersonal
 - 2. Hierarchical > Configurational
 - > Hierarchical ordering at IL
 - 3. Hierarchically higher > Hierarchically lower
 - At each layer:
 - 5. Functions > Operators > Modifiers
 - > Configurational ordering at IL
 - 4. Predicate > Arguments
 - > Hierarchical ordering at IL (continued)
 - 3. Hierarchically higher > Hierarchically lower
 - At each layer:
 - 5. Functions > Operators > Modifiers
 - > Representational
 - 2. Hierarchical > Configurational
 - > Hierarchical ordering at RL
 - 7. Contrast > No contrast
 - 6. Topic > Focus > No pragmatic function
 - 3. Hierarchically higher > Hierarchically lower
 - At each layer:
 - 5. Functions > Operators > Modifiers
 - > Configurational ordering at RL
 - 4. Predicate > Arguments
 - > Predicate
 - > Arguments
 - 6. Contrast > No contrast
 - 7. Topic > Focus > No pragmatic function
 - OR: 8. Definite (+id/+s) > Indefinite specific (-id/+s) > Non-specific (-id/-s)
 - OR: 9. Perspectivized > Not perspectivized
 - 10. A > U > L
 - OR: 11. 1st/2nd > 3rd human > 3 animate > 3 inanimate force > other inanimate

We start with hierarchical ordering at IL, first with the Move, then with the Discourse Act, in each case placing functions before operators before modifiers. Since the head of the Act is configurational in nature, configurational ordering takes over and the Predicate > Argument rule becomes active, such that operators and modifiers of the Illocution (or its head when segmentally expressed) are placed before any information from the Participants and the components of the Communicated Content. At the layer of the Communicated Content hierarchical ordering is resumed, and at each layer functions are again placed before operators before modifiers. All the elements at IL that have a counterpart at RL await placement until the process of representational placement starts. At RL the first step in hierarchical ordering is to assign a place to those hierarchical elements that carry a pragmatic function, after which hierarchical placement is resumed for the remaining constituents that are in a hierarchical relationship, a process that operates down to the layer of the Situational Property. As usual, at each layer functions are placed before operators before modifiers, unless the latter carry a pragmatic function. The remaining constituents are in a configurational relationship. In ordering these, first the Predicate > Argument

rule is applied, after which pragmatic and semantic factors, depending on the language, guide the placement of the remaining arguments.

If there is a tie in the application of one principle, the next principle kicks in. For instance, if the Comment is focal, then the elements that constitute the Comment, all being alike in terms of focality, are ordered in terms of the next relevant principle, e.g. the semantic function hierarchy.

5 Applying the principles

5.1 Introduction

In this section we apply the principles proposed in Section 3 and their ranking in Section 4 to examples from a predicate-medial languages (English), a V2-language (Dutch), a predicate-final language (Turkish), and a predicate-initial language (Tagalog).

5.2 English

Let us begin our treatment of English data with the analysis of a categorical sentence containing two interpersonal modifiers:

- (27) P^{PRECL} P^I P^M P^{M+1} P^{M+2} P^{M+3} P^{M+4}
frankly we unfortunately will be avoiding this restaurant. (Internet, adapted)
 IL: (A_i : [$(F_i$: DECL (F_i): (emph F_j : frankly (F_j)) (F_i)) (P_i) $_A$ (P_j) $_S$ (C_i : [(persp R_i) (C_m : [(T_i) (R_j)] (C_m))] $_{FOC}$] (C_i): (F_k : unfortunately (F_k)) (C_i) $_{\phi}$] (A_i))
 RL: (p_i : (fut ep_i : (e_i : (prog si : [(fi : avoid (fi)) (m x_i) $_A$ (1 prox x_j : -restaurant- (x_j) $_U$] (si)) (e_i)) (ep_i)) (p_i))
 ML: (Le_i : (Sent $_i$: [(AdvP $_i$: -frankly- (AdvP $_i$)) (Cl_i : [(Np $_i$: we (Np $_i$)) [(AdvP $_j$: -unfortunately- (AdvP $_j$)) (Vw $_i$: will (Vw $_i$)) (Vw $_j$: be (Vw $_j$)) (Vw $_k$: -avoiding- (Vw $_k$)) (Np $_j$: -this restaurant- (Np $_j$))] (Cl_i))] (Sent $_i$)) (Le_i))

As can be seen from the IL representation, the two modifiers, *frankly* and *unfortunately*, scope over different interpersonal layers: the Illocution and the Communicated Content, respectively. Since there are no functions, operators or modifiers specifying the Move or Discourse Act, the first placement rule to apply concerns the units within the configurational head of the Discourse Act. As pointed out before, the Illocution is now regarded as an interpersonal predicate, taking the Speech Participants and the Communicated Content as its arguments. Given that elements specifying the predicate are placed before those specifying any of the arguments, *frankly* is placed first; due to the presence of the emphasis operator, this element is placed in the preclausal position P^{PRECL} . Subsequently, *unfortunately* is placed in P^M .¹⁴

Moving on to the Representational Level, we return to hierarchical ordering, starting with elements from the Propositional Content. Since there are no functions to be dealt with, the first element to be placed is the absolute tense operator ‘future’ at the layer of the Episode, expressed

¹⁴ Note that the predicate-argument analysis proposed here for the head of the Discourse Act, together with the newly stipulated rule that predicates are assigned a position before their arguments, explains why *frankly* is placed before *unfortunately*. This could not be accounted for before, because they are not in a hierarchical relation.

as *will* in the relative position P^{M+1} . This is followed by the progressive operator at the layer of the Situational Property, expressed by the auxiliary *be*, which takes the next relative position, P^{M+2} .

We now come to the head of the Situational Property. Since Focus is assigned to the Comment as a whole (wide Focus), we start by assigning positions to the elements that make up the Comment: first the predicate *avoid* (P^{M+3}), then the Undergoer *this restaurant* (P^{M+4}). The final element to be placed is the remaining argument, the Actor *we*, which, due to the fact that the corresponding Subact is specified by the operator ‘perspective’, goes to the default position for English subjects, P^1 . Note that although this Referential Subact evokes the referent that the Comment “is about”, it is not assigned Topic function, since in English this function is not (systematically) coded (Mackenzie and Keizer 1991) and as such not present in underlying representation.

Our next example comes from the spoken component of the ICE-GB Corpus, and is provided with a bit of context:

(28) Context: So the prospect of the United Kingdom which is [sic] only accounts for only about ten per cent of the combined gross domestic product of the European Community becoming economically marginalised is all too real.

P^1 P^M P^{INTCL} P^{M+1} P^{M+2}

Europe could frankly get along without us (ICE-GB: S2A-023_039)¹⁵

IL: (A_i: [(F_i: DECL (F_i): (F_j: frankly (F_j))_{BCKGR} (F_i)] (P_i)_A (P_j)_S (C_i: [(emph persp R_i) (Cm_i: [(T_i) (T_j)] (Cm_i))_{FOC}] (C_i))_Φ] (A_i))

RL: (p_i: (past ep_i: (e_i: (abil s_i: [(f_i: get along (f_i)) (x_i)_A] (s_i: (f_j: -without us- (f_j)) (s_i)) (e_i)) (ep_i)) (p_i))

ML: (Le_i: (Sent_i: (Cl_i: [(Np_i: -Europe- (Np_i)) (Vw_i: can-past (Vw_i)) (AdvP_i: -frankly- (AdvP_i)) (Vw_i: get along (Vw_i)) (Adpp_i: -without us- (Adpp_i))] (Cl_i)) (Sent_i)) (Le_i))

As in the previous example, the illocutionary adverb *frankly* is the first element to be placed. However, since in this case the adverb is assigned the pragmatic function Background, it now appears in P^{INTCL} (leading to a prosodically deaccented realization, following the pitch accented element *Europe* in the same Phonological Phrase).¹⁶ Since there is no further hierarchically ordered interpersonal information to be dealt with, we continue with hierarchical ordering at RL, starting again with tense (‘past’), expressed as a placeholder in P^M . This is followed by the ability operator at the layer of the Situational Property, triggering the modal auxiliary *can*, which joins tense in P^M . Next, we come to the modifier of the Situational Property, the additional participant *without us*, which is placed in P^F .

We then proceed with the configurational ordering of the units within the head of the Situational Property. The first element to be placed here is predicate *get along*, which goes to P^{M+1} (note that the interpolated *frankly* does not take a Clausal position). This leaves us with the Actor *Europe*, which, due to the operator ‘perspective’ on the corresponding Subact, goes to P^1 . Note that the corresponding Subact at IL contains the emphasis operator, resulting in prosodic prominence at the Phonological Level.

¹⁵ The entire utterance in the corpus is *Europe could frankly get along without us perfectly happily*. However, the adverb phrase *perfectly happily* forms a separate Intonational Phrase, and would as such be analysed as a separate Discourse Act at the Interpersonal Level (Afterthought), forming a new Linguistic Expression. As such it has not been included in the analysis here.

¹⁶ For a detailed prosodic analysis of this example, see Kojadinović (2022: 605-606, 609).

The next example differs from the previous examples in that narrow Focus is assigned to one Subact within the Comment:

- (29) P^I P^M P^{M+1} P^{M+2} P^{F-1} P^F
We frankly saw some evidence just recently with the U.S.S. Vincennes (COCA, spok)
 IL: (A_i: [(F_i: DECL (F_i): (F_j: frankly (F_j)) (F_i)) (P_i)_A (P_j)_S (C_i: [(persp R_i) (Cm_i: [(T_i) (R_j) (R_k) (R_L)_{FOC}] (Cm_i))] (C_i)_Φ] (A_i))
 RL: (p_i: (past ep_i: (e_i: (s_i: [(f_i: see (f_i)) (x_i)_A (some x_j: -evidence- (x_j))_U] (s_i): -with the U.S.S. Vincennes- (s_i)) (e_i)) (ep_i): (t_i: -just recently- (t_i)) (ep_i)) (p_i))
 ML: (Le_i: (Sent_i: (Cl_i: [(Np_i: -we- (Np_i)) (Advp_i: -frankly- (Advp_i)) (Vw_i: see-past (Vw_i)) (Np_j: -some evidence- (Np_j)) (Advp_j: -just recently- (Advp_j)) (Adpp_i: -with the U.S.S. Vincennes- (Adpp_i))] (Cl_i)) (Sent_i)) (Le_i))

Once again, we begin with the illocutionary adverb *frankly*, which, in this case, can be assumed to be both integrated and accented, leading to placement in the Clausal medial position. We then proceed with hierarchical ordering at RL, starting with the modifier *with the U.S.S. Vincennes*, which goes to P^F . As a modifier at the layer of the Situational Property, this is not, hierarchically speaking, the highest element at RL, but since the corresponding Subact at IL is assigned a pragmatic function (Focus), it takes precedence over any hierarchically higher elements. We then continue with the placement of the other two hierarchically related elements: past tense (in P^{M+1}) and *just recently* (in P^{F-1}). This example illustrates the need to place modifiers with a pragmatic function before other elements – without this rule, the element *just recently* would have had to be placed first, in which case it would have had to go to P^F , leaving no room for the hierarchically lower modifier *with the U.S.S. Vincennes*.

The next step is the configurational ordering of elements within the Situational Property, where we start with the predicate *see*, which joins tense in P^{M+1} . Since we are dealing with a case of narrow Focus, rather than Focus on the Comment as a whole, we simply follow the general ordering rules for the placement of the non-focused arguments, which means that we first place the Actor *we* (representing the perspective) in P^I , followed by the Undergoer *some evidence* in P^{M+2} .

Our next example again illustrates the relevance of the pragmatic function Background, here assigned to the element *to the post story* (R_k):

- (30) P^I P^M P^{M+1} P^F P^{POSTCL}
Thiessen unsurprisingly responded yesterday to the Post story, ... (COCA, web-newsp)
 IL: (A_i: [(F_i: DECL (F_i)) (P_i)_A (P_j)_S (C_i: [(persp R_i: -Thiessen- (R_i)) (Cm_i: [(T_i) (R_j) (R_k)_{BCKGR}] (Cm_i))] _{FOC}] (C_i): (F_j: unsurprisingly (F_j)) (C_i)_Φ] (A_i))
 RL: (p_i: (past ep_i: (e_i: (s_i: [(f_i: respond (f_i)) (x_i)_A (1 p_j: -Post story- (p_j))_{All}] (s_i)) (e_i)) (ep_i): (t_i: yesterday (t_i)) (ep_i)) (p_i))
 ML: (Le_i: (Sent_i: [(Cl_i: [(Np_i: -Thiessen- (Np_i)) (Advp_i: -unsurprisingly- (Advp_i)) (Vw_i: respond-past (Vw_i)) (Advp_j: -yesterday- (Advp_j))] (Cl_i)) (Adpp_i: -to the Post story- (Adpp_i))] (Sent_i)) (Le_i))

As the only hierarchically related element at IL, the attitudinal adverb *unsurprisingly* is placed first, in P^M . At RL, the tense operator first goes to P^{M+1} , after which the modifier of the Episode *yesterday* is placed in P^F . Proceeding to the placement of configurational elements, the predicate *respond* first joins tense in P^{M+1} . Given the fact that Focus function is assigned to the Comment as a whole,

we then proceed with the only argument within the Comment, *to the Post story*. Since the corresponding Subact at IL is assigned the pragmatic function Background, this element goes to P^{POSTCL} , where it will be realized as prosodically deaccented (following the (focal) nuclear pitch accent on *yesterday*). Finally, the only remaining element, the perspectivized Actor *Thiessen*, is placed in P^I .

5.3 Dutch

As mentioned in Section 2.2, in Dutch main clauses the finite verb occurs in the second Sentential position (our first Clausal position), and can only be preceded by a single constituent, typically the subject (in our initial Sentential, i.e. preclausal, position), as shown in example (31):

- (31) *Het ministerie van Binnenlandse Zaken en Koninkrijksrelaties heeft die boodschap gisteren gegeven aan de eilandbewoners die na orkaan Irma hun heil in het Europese deel van Nederland zochten.*
 The ministry of internal affairs and kingdom.relations have.PRS.3.SG that message yesterday give-PTCP to the island-inhabitants who after hurricane Irma their salvation in the European part of the.Netherlands seek.PST.PL
 ‘The ministry of internal affairs and kingdom relations gave this message yesterday to the inhabitants of the island who after Hurricane Irma have sought refuge in the European part of Netherlands.’ (*De Telegraaf*, 27-09-2017)

P^{PRECL}	P^I	P^{M-1}	P^M	P^{M+1}	P^F
<i>Het ministerie ...</i>	<i>heeft</i>	<i>die boodschap</i>	<i>gisteren</i>	<i>gegeven</i>	<i>aan de eilandbewoners</i>
IL: (A _i : (F _i : DECL (F _i)) (P _i) _A (P _j) _S (C _i : [(persp R _i) (Cm _i : [(T _i) (R _j) (R _k)(R _L)] (Cm _i) _{FOC}] (C _i) _Φ] (A _i))					
RL: (p _i : (pres ep _i : (ant e _i : (s _i : [(f _i : geven (f _i)) (1 x _i : -ministerie ...- (x _i) _A (1 dist p _j : -boodschap- (p _j) _U (m x _k : -de eilandbewoners ...- (x _k) _{REC}] (s _i)) (e _i)) (ep _i : (t _i : -gisteren- (t _i)) (ep _i)) (p _i))					
ML: (Le _i : (Sent _i : [(Np _i : -Het ministerie ...- (Np _i)) (Cl _i : [(Vw _i : hebben-pres (Vw _i)) (Np _j : -die boodschap- (Np _j)) (Advp _i : -gisteren- (Advp _i)) (Vw _j : gegeven (Vw _j)) (Adpp _i : -aan ... zochten- (Adpp _i))] (Cl _i))] (Sent _i)) (Le _i))					

Since there are no IL elements to be assigned a Clausal position during hierarchical ordering, we start with hierarchical ordering at RL. Here, we first get to the present tense operator at the layer of the Episode, which takes up the position for the finite verb in Clause-initial position, P^I , and then, at the same layer, to the temporal modifier *gisteren* ‘yesterday’, which goes to P^M . Moving on to the layer of the State of Affairs, we come to the anteriority operator, expressed by the auxiliary *hebben* ‘have’, which joins tense in P^I . After this, we turn to the units within the head of the Situational Property. First, the predicate goes to P^{M+1} , where it appears in its participial form (*gegeven* ‘given’). Since Focus is assigned to the Comment, we turn to the arguments that are part of the Comment: the Undergoer (*die boodschap*, ‘that message’), going to P^{M-1} , and the Recipient *aan de eilandbewoners* ‘to the inhabitants of the island ...’, which, due to its complexity, is placed in P^F . Finally, the remaining argument, the Actor *Het Ministerie ...* ‘The ministry ...’ (representing the perspective) goes to the default positions for subjects in Dutch, P^{PRECL} .

In the underlined part of example (32), however, the preclausal position is not filled by the subject, but by the object (*dat* ‘that’). This triggers subject-verb inversion, as the finite verb must be in P^I :

- (32) *zij zou hier nog op terug-komen. Dat heeft zij helaas niet gedaan.*¹⁷
 she would here still PART back-come that have.PRS.3.SG she
 unfortunately not do.PTCP
 ‘...she would come back to this. Unfortunately, she hasn’t done so.’

p^{PRECL} p^I p^{I+1} p^M p^{M+1} p^{M+2}
Dat heeft zij helaas niet gedaan
 IL: (A_i: (F_i: DECL (F_i)) (P_i)_A (P_J)_S (C_i: [(R_i)_{TOP} (C_{m_i}: [(T_i) (persp R_J)] (C_{m_i})_{FOC}] (C_i): (F_J: helaas (F_J)) (C_i)_φ] (A_i))
 RL: (p_i: (pres ep_i: (ant neg e_i: (s_i: [(f_i: doen (f_i)) (1 x_i)_A (1 dist x_j)_U] (s_i)) (e_i)) (ep_i)) (p_i))
 ML: (Le_i: (Sent_i: [(Np_i: dat (Np_i)) (Cl_i: [(Vw_i: hebben-pres (Vw_i)) (Np_j: zij (Np_j)) (Advp_i: -helaas- (Advp_i)) (Gw_i: niet (Gw_i)) (Vw_j: gedaan (Vw_j))] (Cl_i))] (Sent_i)) (Le_i))

The only IL element that takes a Clausal position here is the attitudinal adverb *helaas*, which modifies the Communicated Content; this element is placed in P^M . At RL, we place the tense operator in P^I , where it is subsequently joined by the anterior operator, expressed by the auxiliary *hebben*. Next, the negation operator, here scoping over the State of Affairs (Hengeveld and Mackenzie 2018: 26-27), goes to the relative P^{M+1} position. Then the predicate *doen* ‘do’ is placed in P^{M+2} . Next, we turn to the arguments, starting with the Undergoer *dat* ‘that’, since this argument corresponds to a Subact with the pragmatic function Topic. Note that it is the presence of this pragmatic function that causes the Undergoer argument to be placed in P^{PRECL} . This means that, unlike English, Dutch does have a specific way of coding Topic function (the placement of a non-contrastive, non-Actor and non-perspectivized argument in P^{PRECL}). Finally, the remaining argument, the subject *zij* ‘she’, goes to P^{I+1} .

In example (33) we find a passive sentence with a modifier in first position, triggering subject-verb inversion:

- (33) *Nu al worden trainingen gegeven aan ondermeer onderwijzers
 now already PASS.AUX.PRS.PL training.PL give-PTCP to among.others teachers
 en gevangenvaarders over 'hoe om te gaan met seksualiteit'.
 and prison-guards about how around to go with sexuality
 ‘Even now training sessions are being provided to amongst others teachers and prison guards
 on ‘how to deal with sexuality’.*

¹⁷ Where no source is given, the Dutch examples are taken from the *Corpus Hedendaags Nederlands* (Corpus of Contemporary Dutch).

p^{PRECL} p^I p^{I+1} p^M p^{M+1} p^F
Nu al worden trainingen gegeven aan onder meer... over 'hoe ...'
 IL: (A_i: (F_i: DECL (F_i)) (P_i)_A (P_j)_S (C_i: [(persp: R_i) (Cm_i: [(T_i) (R_j) (emph R_k)] (Cm_i)_{FOC}] (C_i)_Φ] (A_i))
 RL: (p_i: (pres ep_i: (ant ei: (s_i: [(f_i: geven (f_i)) (x_i)_A (m ej: (s_j: [(f_j: training (f_j)) (e_k: -over ...
 seksualiteit- (e_k)_{Ref} (s_j)) (e_j)_U (x_k: -onder meer ...gevangenbewaarders- (x_k)_{Rec}] (s_i)) (e_i))
 (ep_i): (t_i: -nu al- (t_i)) (ep_i)) (p_i))
 ML: (Le_i: (Sent_i: [(Advp_i: -nu al- (Advp_i)) (Cl_i: [(Vw_i: worden-pres (Vw_i)) (Np_i: -trainingen-
 (Np_i)) (Vw_j: gegeven (Vw_j)) (Adpp_i: -aan... gevangenbewaarders- (Adpp_i)) (Adpp_j: -
 over ... seksualiteit- (Adpp_j))] (Cl_i))] (Sent_i))] (Le_i))

In this example, the placement of elements only starts with hierarchical ordering at RL: first present tense is placed in P^I, then the emphasized temporal modifier *nu al* is placed in P^{PRECL}. Subsequently, the predicate, *gegeven* 'given', goes to P^M, after which the argument within the focal Comment, the Recipient *aan onder meer ...* 'to amongst other ...', takes up P^{M+1}. Turning to the remaining arguments, we see that the Actor, represented as a headless variable (x_i), does not correspond to any Subact at the Interpersonal Level; as such it is not expressed. The Undergoer consists of a head (*training*) and its argument (*over 'hoe ...'* 'about 'how...'). Due to the complexity of this argument, it is extraposed; thus, while the head (the predicate *training*) is placed in P^{I+1}, the embedded argument is placed in P^F. Since the perspective operator is assigned to the Subact corresponding to the Undergoer, the result is a passive. This means that as a final step the support element *worden* has to be added. Since in (33) this element functions as the finite verb, it is added to tense in P^I.

What is of particular interest in our next example (the underlined part in (34)) is the fact that two elements precede the finite verb *zullen*:

(34) *Gelukkig kunnen wij als kleine cinema snel schakelen,*
 fortunately can.PRS.PL we as small cinema quick switch
waardoor ik meteen opnieuw open kan.
 because-of-which I immediately anew open can.PRES.SG.
Grotere bioscopen daarentegen zullen meer voorbereidingen
 bigger cinemas however will.PL more preparations
moeten doen.
 must.INF do.INF

'Fortunately, as a small cinema we can change things quickly, which means that I can immediately open again. Bigger cinemas on the other hand will need to make more preparations.'

$p^{PRECL-1}$ p^{PRECL} p^I p^M p^F
Grotere bioscopen daarentegen zullen meer voorbereidingen moeten doen.
 IL: (contr A_i: (F_i: DECL (F_i)) (P_i)_A (P_j)_S (C_i: [(emph persp R_i) (Cm_i: [(T_i) (R_j)] (Cm_i)_{FOC}] (C_i)_Φ] (A_i))
 RL: (p_i: (fut ep_i: (ei: (obl s_i: [(f_i: doen (f_i)) (m x_i: -grotere bioscopen- (x_i)_A (m ej: -meer
 voorbereidingen- (e_j)_U] (s_i)) (e_i)) (ep_i)) (p_i))
 ML: (Le_i: (Sent_i: [(Np_i: -grotere bioscopen- (Np_i)) (Advp_i: -daarentegen- (Advp_i)) (Cl_i: [(Vw_i:
 zullen (Vw_i)) (Np_j: -meer voorbereidingen- (Np_j)) (Vp_i: [(Vw_i: moeten (Vw_i)) (Vw_i:
 doen (Vw_i)) (Vp_i))] (Cl_i))] (Sent_i))] (Le_i))

As mentioned before, in Dutch the finite verb is commonly assumed to occur in second position, and as such can only be preceded by one constituent. This would mean that the position of the contrast marker *daarentegen* in (34) can only be accounted for by assuming that it forms one constituent with the subject *grotere bioscopen* ‘bigger cinemas’; this, in turn, would mean that the contrast marker has to be regarded as having narrow scope over the preceding element. We do not find such an interpretation very convincing; instead, we regard *daarentegen* as having wide scope, indicating a contrast between the contents of the current Discourse Act and information present in the previous discourse. We therefore analyse *daarentegen* as an operator on the Discourse Act, which, in this case, is assigned a preclausal position. This leaves the initial Clausal position available for the finite verb (*zullen*), which now occurs in the third position of the sentence (cf. Van der Wouden 2015, 2020). Note, however, that this additional preclausal position can host only a small set of elements (wide scope contrast markers like *daarentegen*, *echter*, *evenwel*), which can only be preceded by an emphatic element (in this case the subject *grotere bioscopen*) in the Sentence-initial position (as evidenced by the fact that pronouns in this position can only occur in their strong form (*zij/*ze daarentegen ...* ‘they on the other hand ...’); see Van der Wouden 2020: 248).

Turning to the hierarchical ordering of RL elements, we come to the participant-oriented modality operator ‘obligation’, expressed as *moeten* ‘must’, which is placed in the verb phrase that occupies the P^F position. The predicate *doen* also ends up in this verb phrase; in other words, these two elements are regarded as constituting a phrase that takes up a single position within the Clause. This means that, generally speaking, the final position in Dutch Clauses can host more than one element, provided that the verbal elements filling this position are obligatorily placed adjacent to each other (as already proposed for Dutch subclauses in Hengeveld and Mackenzie (2008: 355)).¹⁸

This leaves us with the two arguments. Since Focus is assigned to the Comment, we start with the argument that is part of the Comment, i.e. the Undergoer argument *meer voorbereidingen* ‘more preparations’, which goes to P^M. Finally, as already mentioned, the Actor argument (both perspectivized and emphasized) goes to the Sentence-initial position (in this case p^{PRECL-1}).

5.4 Turkish

Turkish is a language in which in the basic constituent order the predicate occupies the final position in the clause, the subject comes first, and other arguments occur in between subject and predicate, as shown in (35).

- (35) *Hasan kitab-ı Ali-ye ver-di-ø.*
 Hasan book-ACC Ali-DAT give-PST-3.SG
 ‘Hasan gave the book to Ali.’
 (Kornfilt 1997: 90)

¹⁸ Note that, as in the case of verbal elements in the final position of Dutch subclauses, the two elements can also appear in the reversed order (*doen moeten*), although this is relatively rare.

Hierarchical material may be placed in clause-initial (36), clause-medial (37), and clause-final (38) position.

- (36) Birkaç gün önce *birisi istakoz-u Ali-ye ver-di-ø.*
 some day before someone lobster-ACC Ali-DAT give-PST-3.SG
 ‘Several days ago someone gave the lobster to Ali.’
 (Kılıçaslan 2004: 730)
- (37) *Istakoz-u birkaç gün önce Ali ye-di-ø.*
 lobster-ACC some day before Ali eat-PST-3.SG
 ‘Ali ate the lobster several days ago.’
 (Kılıçaslan 2004: 730)
- (38) *On-u Ali ye-di birkaç gün önce.*
 it-ACC Ali eat-PST some day before
 ‘Ali ate it several days ago.’
 (Akan and Hartmann 2019: 134)

Higher order modifiers may also occupy the initial and medial positions, but are less felicitous (Beyza Sümer, personal communication) in the final position.

- (39) Açıkçası muhtemelen *ben gel-me-yeceğ-im.*
 frankly possibly 1.SG come-NEG-FUT-1.SG
 ‘Frankly, I will probably not come.’
 (Beyza Sümer, speaker; adapted from Wilson and Saygin 2001: 4)
- (40) *Ben açıkçası muhtemelen gel-me-yeceğ-im.*
 1.SG frankly possibly come-NEG-FUT-1.SG
 ‘Frankly, I will probably not come.’
 (Beyza Sümer, speaker; adapted from Wilson and Saygin 2001: 4)
- (41) Çok_şükür *bu fareler bozuk peynir-i ye-di.*
 fortunately PROX mice spoiled cheese-ACC eat-PST
 ‘Fortunately these mice ate the spoiled cheese.’
 (Gürer 2014: 235)
- (42) *Bu fareler çok_şükür bozuk peynir-i ye-di.*
 PROX mice fortunately spoiled cheese-ACC eat-PST
 ‘Fortunately these mice ate the spoiled cheese.’
 (Gürer 2014: 235)

Topic constituents in Turkish take, when not preceded by hierarchically placed constituents, the initial position in the clause, as shown in (43):

- (43) p^I p^{I+1} p^{I+2} p^F
Istakoz-u Hasan Ali-ye ver-di-ø.
 lobster-ACC Hasan Ali-DAT give-PST-3.SG
 ‘(Speaking of) the lobster, Hasan gave (it) to Ali.’
 (Kornfilt 1997: 200)

Several authors (e.g. Kornfilt 1997) have claimed that Focus constituents in Turkish occur in the preverbal position.

- (44) *Kitab-ı Ali-ye Hasan ver-di-ø.*
 book-ACC Ali-DAT Hasan give-PST-3.SG
 'HASAN gave the book to Ali.'
 (Kornfilt 1997: 190)

However, Kılıçaslan (2004: 720) shows that focal constituents may, and sometimes have to be, separated from the verb, as in (45):

- (45) (Context: who saw a dog in the garden?)
 p^I p^M p^{F-1} p^F
Bahçe-de Oya bir köpek gör-dü-ø.
 garden-LOC Oya one dog see-PST-3.SG
 'OYA saw a dog in the garden.'

In Turkish non-specific Undergoers have to occur in pre-verbal position, such that the focus constituent cannot occur in that position. Since the non-specific Undergoer cannot be separated from the verb, it must be in p^{F-1} . And since focal arguments must be assigned a position before arguments without a pragmatic function, the focal constituent in (45) must be in p^M .

Contrastive foci may occur in situ and are then marked prosodically, not morphosyntactically, as in (46b). However, they may also occur in the clause initial field, as in (46c).

- (46) Context: What did the waiter bring for you?
 a. p^I p^{I+1} p^M p^F
Garson bana kızartma-yı getir-di-ø
 Waiter 1.SG.DAT fries-ACC bring-PST-3.SG
 'The waiter brought the fries for me.'
 b. p^I p^{I+1} p^M p^F
Hayır, garson sana meyve-ler-i getir-di-ø
 no waiter 2.SG.DAT fruit-PL-ACC bring-PST-3.SG
 'No, THE FRUITS the waiter brought for you (not the fries).'
 p^I p^{I+1} p^{I+2} p^F
 c. *Hayır, meyve-ler-i garson sana getir-di-ø*
 no fruit-PL-ACC waiter 2.SG.DAT bring-PST-3.SG
 'No, THE FRUITS, the waiter brought for you (not the fries).'
 (Grigoraş 2020: 50).

When both a contrastive focus and a contrastive topic are present in the initial field, the topic precedes the focus.

- (47) Context: Biden and Warren visited our campus. Who brought which presidential candidate to the campus?
 p^I p^{I+1} p^{I+2} p^M p^F
Biden-i^{TOPCONTR} Seren^{FOCCONTR} kampüs-e bu sabah getir-di-ø.
 Biden-ACC Seren campus-DAT this morning bring-PST-3.SG

‘Seren brought Biden to campus this morning.’
(Grigoraş 2020: 79)

The only situation in which a non-hierarchical constituent can end up after the verb is when it is backgrounded, as in (48):

- (48) Context: ‘Who married Kaya?’
 p^M p^F p^{post}
Oya^{FOC} *evlen-di-∅* *Kaya-yla*^{BACKGR}
 Oya marry-PST-3.SG Kaya-COM
 ‘Oya married Kaya.’
 (Kılıçaslan 2004: 727)

In this case, we consider these backgrounded constituents to be in the postclausal position P^{POSTCL} , as they are deaccented.

With these remarks on linearization in mind, we will consider the ordering process in a number of examples in more detail. We will start with example (37), repeated here as (49).

- (49) Context: ‘What about the lobster? Who ate it?’
 p^I p^M p^{M+1} p^F
Istakoz-u *birkaç gün önce* *Ali* *ye-di-∅*.
 lobster-ACC some day before Ali eat-PST-3.SG
 ‘Ali ate the lobster several days ago.’
 (Kılıçaslan 2004: 730)
 IL: (A_i: (F_i: DECL (F_i)) (P_i)_S (P_j)_A (C_i: [(R_i)_{TOP} (C_{m_i}: [(T_i) (persp R_j: -Ali- (R_j))_{FOC} (R_k)] (C_{m_i}))] (C_i)) (A_i))
 RL: (p_i: (past ep_i: (e_i: (s_i: [(f_i: ye- (f_i)) (x_i)_A (x_j: - istakoz- (x_j))_U] (s_i)) (e_i)) (ep_i): (t_i: -birkaç gün önce- (t_i)) (ep_i)) (p_i))
 ML: (Le_i: (Sent_i: (Cl_i: [(Np_i: istakoz (Np_i)) (Adpp_i: -birkaç gün önce- (Adpp_i)) (Np_j: -Ali- (Np_j)) (Vw_i: -yedi- (Vw_i))] (Cl_i)) (Sent_i)) (Le_i))

Since there are no hierarchical elements from IL present in this sentence, we start with hierarchical ordering at RL. The relevant elements are the past tense marker on the verb and the temporal modifier *birkaç gün önce* ‘several days ago’, both operating at the layer of the Episode. Since operators are placed before modifiers, the past tense goes to P^F first, after which the temporal modifier goes to P^M . In configurational ordering the predicate is then placed first and joins tense in P^F . Of the two arguments one is Topic and the other is Focus. As Topic precedes Focus in placement, *istakozu* is placed in P^I first, after which the Focus constituent *Ali* is placed in P^{M+1} .

The following example contains two backgrounded constituents.

- (50) Context: ‘When did Oya marry Kaya?’
 p^M p^F p^{POSTCL} $p^{POSTCL+1}$
İki yıl önce *evlen-di* *Oya* *Kaya-yla*
 two years before marry-PST-3.SG Oya Kaya-COM
 ‘Two years ago Oya married Kaya.’ (litt: “with Kayla”) (Kılıçaslan 2004: 728)
 IL: (A_i: (F_i: DECL (F_i)) (P_i)_S (P_j)_A (C_i: [(C_{m_i}: [(T_i) (persp R_j: -Oya- (R_j))_{BACKGR} (R_k: -Kaya- (R_k))_{BACKGR} (+id -s R_L)_{FOC}] (C_{m_i}))] (C_i))] (A_i))

RL: (p_i: (past ep_i: (e_i: (s_i: [(f_i: evlen (f_i)) (x_i)_A (x_j)_{Com}] (s_i)) (e_i)) (ep_i: (t_i: -iki yıl önce- (t_i)) (ep_i)) (p_i))
 ML: (Le_i: (Sent_i: [(Cl_i: [(Adpp_i: -iki yıl önce- (Adpp_i)) (Vw_i: -evlendi- (Vw_i))] (Cl_i)) (Np_i: -Oya- (Np_i)) (Np_j: -Kayayla- (Np_j))] (Sent_i)) (Le_i))

In this example hierarchical ordering starts at the Representational Level with the past tense operator and the temporal modifier, both at the Episode layer. Since the temporal modifier has a pragmatic function, it is placed before the past tense operator. The temporal modifier being in focus, it goes to P^M, after which the past tense operator goes to P^F. The next principle to be applied is that predicates are placed before arguments, which means that the verb goes to P^F, where it joins the tense operator. Finally, the arguments have to be placed. Both have Background function, hence the next principle has to be applied, which states that Actors are placed first. Therefore *Oya* is placed in P^{POSTCL}, after which the comitative argument can go to P^{POSTCL+1}.

The latter example shows that in P^{POSTCL} expansion is to the right. This can also be seen in the following example:

(51) *Gel-me-yeceğ-im* açıkçası muhtemelen.
 come-NEG-FUT-1.SG frankly possibly
 ‘Frankly, I will probably not come.’
 (Akan and Hartman 2019: 141)

In this example the adverbs are presented as afterthoughts. Since *açıkçası* ‘frankly’ has to be placed before *muhtemelen* ‘possibly’, the former must be in P^{POSTCL} and the latter in P^{POSTCL+1}.

Example (40), repeated here as (52), contains the same modifiers.

(52) p^I p^M p^{M+1} p^F
Ben açıkçası muhtemelen *gel-me-yeceğ-im*.
 1.SG frankly possibly come-NEG-IRR-1SG
 ‘I frankly may not come.’
 (Beyza Sümer, speaker; adapted from Wilson and Saygin 2001: 4)
 IL: (A_i: (F_i: DECL (F_i: (F_j: açıkçası (F_j)) (F_i)) (P_i)_S (P_j)_A (C_i: [(persp R_i) (Cm_i: [(T_i) (T_j)] (Cm_i)) (C_i))] (A_i))
 RL: (p_i: (irr ep_i: (neg e_i: (s_i: [(f_i: gel (f_i)) (x_i)_A] (s_i)) (e_i)) (ep_i)) (p_i: (f_j: muhtemelen (f_j)) (p_i))
 ML: (Le_i: (Sent_i: (Cl_i: [(Np_i: -ben- (Np_i)) (Advp_i: -açıkçası- (Advp_i)) (Advp_j: -muhtemelen- (Advp_j)) (Vw_i: -gelmeyeceğim- (Vw_i))] (Cl_i)) (Sent_i)) (Le_i))

In configurational ordering at IL, the illocutionary adverb has to be placed first and is assigned the P^M position. In hierarchical ordering at RL, the modal modifier has to be assigned a position, as well as the irrealis and negative operators. Since the modal modifier operates at the layer of the Propositional Content, it is assigned first the P^{M+1} position. The irrealis and modal operators then go to P^F, where they will later be joined by the verb. In configurational ordering the predicate is placed first in P^F, and finally the only argument, which carries the perspectivizing operator, goes to P^I.

The following example shows the interaction of various properties of the Turkish linearization system.

(53) p^{PRECL} p^I p^{I+1} p^M p^F

Duy-duğ-um-a göre, öğretmen öğrenci-ler-e kitap-lar-ı ver-miş.
 hear-NMLZ-POSS-DAT according teacher students-PL-DAT book-PL-ACC give-PST
 ‘As I hear, the teacher gave the students the books.’

(Akan and Hartmann 137)

IL: (A_i: (F_i: DECL (F_i)) (P_i)_S (P_j)_A (C_i: [(persp R_i)_{TOP} (Cm_i: [(T_i) (R_j)_{FOC} (R_k)] (Cm_i)) (C_i): -duyduğuma göre- (C_i))] (A_i))

RL: (p_i: (pst ep_i: (e_i: (s_i: [(f_i: ver (f_i)) (1 x_i: -öğretmen- (x_i))_A (m x_j: -kitap- (x_j))_U (m x_k: öğrenci (x_k))_L] (s_i)) (e_i)) (ep_i)) (p_i))

ML: (Le_i: (Sent_i: [(Adpp_i: -duyduğuma göre- (Adpp_i)) (Cl_i: [(Np_i: -öğretmen- (Np_i)) (Np_j: -öğrencilere- (Np_j)) (Np_k: -kitapları- (Np_k)) (Vw_i: -vermiş- (Vw_i))] (Cl_i)) (Sent_i)) (Le_i))

At the Interpersonal Level, the reportative modifier has to be placed first. It is assigned the preclausal position P^{PRECL}, which is consistent with the fact that it is separated from the remainder of the clause by means of an intonation break. This allows the Subject to stay in its default position P^I. At the Representational Level, the only hierarchical element is the past tense marker, which goes to P^F. Then the predicate has to be placed before any arguments, and it joins the tense marker in P^F. Among the arguments, the Undergoer argument carries the Focus function, and the Actor argument the Perspective operator. As functions rank higher than operators, the Undergoer is placed in P^M, after which the Actor argument is placed in P^I. Finally, the Recipient argument is placed in P^{I+1}.

5.5 Tagalog

In its basic constituent order, Tagalog is a predicate-initial language as demonstrated in (54). It combines this feature with having a second clausal position for enclitics, as shown in Section 2.2.

(54) p^I p²⁻¹ p² p^M p^{M+1} p^F
B<in>ili =ko =na ang bulaklak para kay Weng kahapon.
 PRF:bought =1SG.ERG=ALREADY TOP flower for DAT Weng yesterday
 ‘I have already bought the flower for Weng yesterday.’
 (Nagaya 2007: 346)

The effect of the second position is clearly seen in the following example, in which an adverb occupies the initial position, and the verb, as a result, is placed in a position after the second position clitics.

(55) p^I p²⁻¹ p² p^M p^{M+1}
Ngayon =ko =lang na-basa ang=email=mo.
 now =1.SG.ERG =just PRF-read TOP=email=your
 ‘I read your e-mail just now.’
 (Nagaya 2007: 350)

The first clausal position is important for the analysis of Focus. When the predicate or the comment is in focus, the order is as illustrated in (54), with the predicate in initial position. When the focus is on an adjunct, this adjunct may be placed in initial position, as in (55), and also in (56). For identificational Focus a cleft construction is used, in which case the focused non-verbal

Higher adverbs, on the other hand, do occur in P^{PRECL}.

- (62) P^{PRECL} p^I p²
Malamang ay nan-daya =sila.
 probably PART AV.PRF-cheat =3.PL.SBJ
 'They probably cheated.'
 (Kaufman 2006: 157)

Adverbs that occupy a middle position in the hierarchy allow both possibilities.

- (63) p^I p^M p^F
Um-u~ulan nang madalas dito.
 AV-PROG~rain LNK often here
 'It rains often here.'
 (Kaufman 2006: 161)

- (64) P^{PRECL} p^I p^F
Madalas ay um-u~ulan dito.
 often PART AV-PROG~rain here
 'It rains often here.'
 (Kaufman 2006: 161)

The clausal second position may be occupied by several clitics. In that case their hierarchical ordering reflects the scope relations between them, where elements with the highest scope occupy the rightmost position.

- (65) p^I p²⁻² p²⁻¹ p² p^M p^F
Nag-bitiw =na =raw =ba ang komisyoner kahapon?
 AV.PRF-quit =ALREADY=RPRT =Q SBJ commissioner yesterday
 'Did the commissioner reportedly quit yesterday?'
 (Kaufman 2006: 178)

This shows that the assumption in Hengeveld and Mackenzie (2008) that the absolute P² position may only expand to the right does not hold. The following variant of (65) shows that the positions of the clitics do not depend on the position of the predicate either:

- (66) p^I p²⁻² p²⁻¹ p² p^M p^{M+1}
Kahapon =na =raw =ba nag-bitiw ang komisyoner?
 Yesterday=ALREADY=RPRT =Q AV.PRF-quit TOP commissioner
 'Was it yesterday that the commissioner reportedly quit?'
 (Kaufman 2006: 178)

As shown in (66), the order of the clitics does not change when an adverbial phrase rather than the predicate is placed in initial position.

With these facts in mind, we now turn to the way the linearization process in Tagalog takes place. We will start with example (58), repeated here as (67).

- (67) p^1 p^2 p^{2+1} p^M p^F
kailan =ba =siya ba-balik dito?
 when =Q =3.SG.ABS AV-will.return here
 ‘When will she come here?’
 (Nagaya 2007: 361)
 IL: (A_i: (F_i: INT (F_i)) (P_i)_S (P_j)_A (C_i: [(R_i) (Cm_i: [(T_i) (+id -s R_j)_{FOC} (R_k)] (Cm_i)] (C_i))] (A_i))
 RL: (p_i: (fut ep_i: (e_i: (s_i: [(f_i: balik (f_i)) (x_i)_A] (s_i: (l_i)_L (s_i)) (e_i)) (ep_i: (t_i) (ep_i)) (p_i))
 ML: (Le_i: (Sent_i: [(Cl_i: [(Advp_i: -kailan- (Advp_i)) (Gw_j: ba (Gw_j)) (Gw_k: siya (Gw_k)) (Vw_i: -ba-balik- (Vw_i)) (Advp_j: -dito- (Advp_j))] (Cl_i))] (Sent_i)) (Le_i))

At the IL, the interrogative clitic has to be assigned a position first and goes to P². Hierarchical ordering at RL then involves the future tense marker on the verb, the questioned temporal modifier *kailan* ‘when’ and the directional modifier *dito* ‘here’. Since within hierarchical ordering at RL elements with a pragmatic function have priority, the question word is placed in P¹, after which the future tense marker, which applies at the Episode layer, has priority over the directional modifier, which applies at the layer of the Situational Property. It is placed in P^M, after which the directional modifier goes to P^F. The predicate is then placed in P^M, where it joins tense. Finally, the Topic is resumed within the clause. The pronoun *=siya* is placed in the P²⁺¹ slot.

In our next example there are three expressions of operators:

- (68) p^{PRECL} p^1 p^2 p^M p^{M+1}
Maaaring hindi =sila maaaring mag-aral.
 can.LNK NEG =3PL.SBJ can.LNK AV.INF-study
 ‘It is possible that they are unable to study.’
 (Kaufman 2006: 176)
 IL: (A_i: (F_i: DECL (F_i)) (P_i)_S (P_j)_A (C_i: [(Cm_i: [(persp R_i) (T_i)] (Cm_i)] (C_i))] (A_i))
 RL: (p_i: (poss ep_i: (neg e_i: (abil s_i: [(f_i: -aral- (f_i)) (x_i)_A] (s_i) (e_i)) (ep_i)) (p_i))
 ML: (Le_i: (Sent_i: [(Vw_i: -maaaring- (Vw_i)) (Cl_i: [(Gw_i: -hindi- (Gw_i)) (Gw_j: sila (Gw_j)) (Vw_j: -maaaring- (Vw_j)) (Vw_i: -magaral- (Vw_i))] (Cl_i))] (Sent_i)) (Le_i))

. The auxiliary *maaari* occurs twice, once with an epistemic reading and once with a dynamic reading. In the former reading it operates at the layer of the Episode, in the latter on the layer of the Situational Property. Furthermore, there is a grammatical word expressing negation, which operates at the layer of the State of Affairs. Since there are no hierarchical elements at IL, we have to start with the placement of these elements in hierarchical order. The epistemic auxiliary goes to P^{PRECL} first, a position that it has to be in, given that the second position clitic *sila* follows the next constituent. The negative particle then goes to P¹, and the facultative auxiliary to P^M. Then the predicate has to be assigned a position before the only argument and goes to P^{M+1}. Finally, the pronominally expressed argument goes to P².

The next example we analyze is (65), repeated here as (69).

- (69) p^1 p^{2-2} p^{2-1} p^2 p^M p^F
Nag-bitiw =na =raw =ba ang komisyoner kahapon?
 AV.PRF-quit =ALREADY=RPRT=Q SBJ commissioner yesterday
 ‘Did the commissioner reportedly quit yesterday?’
 (Kaufman 2006: 178)
 IL: (A_i: (F_i: INT (F_i)) (P_i)_S (P_j)_A (rp_{rpt} C_i: [(Cm_i: [(persp R_i) (T_i) (R_i)] (Cm_i)]_{FOC}] (C_i))] (A_i))

RL: (p_i: (ep_i: (e_i: (res s_i: [(f_i: -bitaw- (f_i)) (x_i: -komisyoner- (x_i))_A] (s_i) (e_i)) (ep_i: (t_i: -kahapon- (t_i)) (ep_i)) (p_i))
 ML: (Le_i: (Sent_i: (Cl_i: [(Vw_i: -nagbitaw- (Vw_i)) (Gw_i: -na- (Gw_i)) (Gw_j: -raw- (Gw_j)) (Gw_k: ba (Gw_k)) (Adpp_i: -ang komisyoner- (Adpp_i)) (Advp_i: kahapon (Advp_i))] (Cl_i)) (Sent_i)) (Le_i))

At the Interpersonal Level, the interrogative illocution has to be assigned a place first in configurational ordering. It goes to P². Then hierarchical ordering takes over at the layer of the Communicated Content, and the reportative marker is placed in P²⁻¹. At the Representational Level a perfect operator and a temporal modifier have to be expressed. Since the latter is part of the focused Comment, it has to be expressed before the perfect operator. It is placed in P^F, after which the resultative operator goes to P²⁻². Then the predicate has to be expressed before its single argument and is placed in P^I. Finally, the single argument is placed in P^M.

Our final example shows that there may be multiple elements in preclausal position in Tagalog, as in the following example, in which the two pre-Clausal constituents are marked by the postposed particle *ay*.

(70) p^{PRECL} p^{PRECL+1} p^M
Ngayon ay siya ay natátákot.
 now PART 3.SG PART STA.INGR.fear
 ‘Now he is scared.’ (Kaufman 2005)
 IL: (A_i: (F_i: DECL (F_i)) (P_i)_S (P_j)_A (C_i: [(R_i)_{Top} (persp R_j)_{Top} (Cm_i: [(T_i)] (Cm_i))_{Foc}] (C_i)) (A_i))
 RL: (p_i: (ep_i: (e_i: (s_i: [(f_i: -takot- (f_i)) (x_i)_U] (s_i) (e_i)) (ep_i: (t_i: -ngayon ay- (t_i)) (ep_i)) (p_i))
 ML: (Le_i: (Sent_i: [(Advp_i: -ngayon ay- (Advp_i)) (Np_i: -siya ay- (Np_i)) (Cl_i: [(Vw_i: -takot- (Vw_i))] (Cl_i)) (Sent_i)) (Le_i))

There are no elements from IL that have to be assigned a position, so we start with hierarchical ordering at RL, where the adverb *ngayon* ‘now’, which is furthermore a Topic, is the only element to be assigned a position. It is placed in the preclausal position P^{PRECL}. Next, the predicate has to be assigned a place, and it goes to the medial position. The topical Undergoer is next in line, and goes to P^{PRECL+1}. This example shows that the preclausal position in Tagalog expands to the right.

6 Conclusion

In this paper we have proposed a number of additions to the FDG linearization rules for clauses, introducing several ranked priority rules that take into account the hierarchical structure of FDG, the precedence of predicates over arguments, the precedence of functions over operators over modifiers, the importance of a pragmatic functions hierarchy and other pragmatic factors, and the importance of the semantic hierarchy and other semantic factors. In addition, we have argued for the introduction of a new morphosyntactic layer (the Sentence), which allows for the creation of a new set of extra-clausal positions for pragmatically marked elements. Applied to a variety of languages, these innovations have been shown to account for some frequently attested linear patterns in these languages, while excluding unacceptable sequences.

Acknowledgements

We are grateful to Beyza Sümer and Liberty Notarte Balanquit for their kind help in interpreting the Turkish and Tagalog data, respectively. We also thank Daniel García Velasco, Riccardo Giomi, Lachlan Mackenzie, Elnora ten Wolde and an anonymous reviewer for their insightful comments on earlier versions of this paper.

Abbreviations used

1 first person, 2 second person, 3 third person, ABS absolutive, ACC accusative, AG agent, ASP aspect, AUX auxiliary, AV actor voice, COM comitative, CONJ conjunction, COP copula, DAT dative, DECL declarative, DEF definite, DEM demonstrative, DRCT direct voice, DV dative/locative voice, ERG ergative, FUT future, GEN genitive, INGR ingressive, INDEF indefinite, INF infinitive, INV inverse voice, IV instrumental voice, LNK linker, LOC locative, NEG negation, NOM nominative, NR nominalizer, PART particle, PASS passive, PL plural, POSS possessive, PRES present, PRF perfect, PROGR progressive, PROX proximate, PST past, PTCP participle, Q question, REAL realis, RPRT reportative, SG singular, STA stative, SBJ subject, TOP topic, UV undergoer voice.

References

- Akan, Tamer and Hartmann, Katharina. 2019. SOV-X: Syntactic and pragmatic constraints of the postverbal domain in Turkish. In Josef Bayer and Yvonne Viesel (eds), *Proceedings of the Workshop clause typing and the syntax-to-discourse relation in head-final languages*, 123-144. Konstanz: Universität Konstanz.
- Aissen, Judith. 1992. Topic and Focus in Mayan. *Language* 68(1). 43–80.
- Cinque, Guglielmo. 2023. *On Linearization: Toward a Restrictive Theory*. Cambridge, Mass.: MIT Press.
- Dik, Simon C. 1997. *The Theory of Functional Grammar. Part I: The Structure of the Clause* (Functional Grammar Series 20), 2nd, revised edn. Edited by Kees Hengeveld. Berlin and New York NY: Mouton de Gruyter.
- Ernst, Thomas. 2002. *The Syntax of Adjuncts*. Cambridge: Cambridge University Press.
- Fang Hongmei and Hengeveld, Kees. 2022. Sentence-final particles in Mandarin. *Studia Linguistica*. Advance online publication. <https://doi.org/10.1111/stul.12198>.
- Giomi, Riccardo. 2021. The place the place of interpersonal lexemes in linguistic theory, with special reference to Functional Discourse Grammar. *Corpus Pragmatics* 5(2). 1-36.
- Giomi, Riccardo. this volume.
- Grigoraş, Sözen Özkan. 2020. *Word Order and Information Structure in Turkish*. PhD Diss., University of California, Los Angeles.
- Gürer, Asli. 2015. *Semantic, prosodic and syntactic marking of information structural units in Turkish*. PhD Diss., Boğaziçi University.
- Hasselgård, Hilde. 2010. *Adjunct adverbials in English*. Cambridge: Cambridge University Press.
- Hengeveld, Kees. 2004. Illocution, mood, and modality. In Geert Booij, Christian Lehmann and Joachim Mugdan (eds), *Morphology: A handbook on inflection and word formation*, Vol. 2, 1190-1202. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110172782.2.14.1190>.

- Hengeveld, Kees. 2013. A new approach to clausal constituent order. In J. Lachlan Mackenzie and Hella Olbertz (eds), *Casebook in Functional Discourse Grammar* (Studies in Language Companion Series 137), 15-38. Amsterdam: John Benjamins.
<https://doi.org/10.1075/slcs.137.02hen>.
- Hengeveld, Kees and Keizer, Evelien. 2011. Non-straightforward communication. *Journal of Pragmatics* 43(7). 1962-1976. <http://dx.doi.org/10.1016/j.pragma.2011.01.001>.
- Hengeveld, Kees and Keizer, Evelien and Giomi, Riccardo. in prep. *Layering in Functional Discourse Grammar: The hierarchical structure of the language system*. Oxford: Oxford University Press.
- Hengeveld, Kees and Mackenzie, J. Lachlan. 2008. *Functional Discourse Grammar: A typologically-based theory of language structure*. Oxford: Oxford University Press.
- Hengeveld, Kees and J. Lachlan Mackenzie. 2018. Negation in Functional Discourse Grammar. In Evelien Keizer and Hella Olbertz (eds.), *Recent developments in Functional Discourse Grammar*, 17-46. Amsterdam: Benjamins.
- Kaufman, Daniel. 2005. Aspects of pragmatic focus in Tagalog. In I Wayan Arka and Malcolm Ross (eds), *The many faces of Austronesian voice systems*, 175-196. Canberra, Pacific Linguistics.
- Kaufman, Daniel. 2006. Rigidity versus Relativity in Adverbial Syntax: Evidence from Tagalog. In Hans-Martin Gärtner and Paul Law and Joachim Sabel (eds), *Clause Structure and Adjuncts in Austronesian Languages*, 151-194. Berlin: Mouton de Gruyter.
- Keenan, Edward L. and Dryer, Matthew S. 2007. Passive in the world's languages. In Timothy Shopen (ed.), *Language typology and syntactic description. Volume 1: Clause structure*, 325-361. Cambridge: Cambridge University Press.
- Keizer, Evelien. 2014. Context and cognition in Functional Discourse Grammar: What, where and why? *Pragmatics* 24(2). 399-423.
- Keizer, Evelien. 2015. *A Functional Discourse Grammar for English*. Oxford: Oxford University Press.
- Keizer, Evelien. 2018. Interpersonal adverbs in FDG: the case of *frankly*. In Evelien Keizer and Hella Olbertz (eds), *Recent developments in Functional Discourse Grammar* (Studies in Language Companion Series 205), 48-88. Amsterdam: John Benjamins.
<https://doi.org/10.1075/slcs.205>.
- Keizer, Evelien. 2020. The placement of extra-clausal constituents in Functional Discourse Grammar. *Revista Canaria de Estudios Ingleses* 80. 89-121.
- Kılıçaslan, Yılmaz. 2004. Syntax of information structure in Turkish. *Linguistics* 42(4). 717-765.
- Kornfilt, Jaklin. 1997. *Turkish* (Descriptive Grammars). London: Routledge.
- Kroeger, Paul. 1998. Clitics and clause structure in Tagalog. In Lourdes S. Bautista (ed.), *Pagtanaw: Essays on language in honor of Teodoro A. Llamzon*, 53-72. Manila: Linguistic Society of the Philippines.
- Mackenzie, J. Lachlan and Evelien Keizer. 1991. On assigning pragmatic functions in FG, *Pragmatics* 1/2: 169-215.
- Nagaya, Naonori. 2007. Information structure and constituent order in Tagalog. *Language and Linguistics* 8(1). 343-372.
- Smit, Niels. 2010. *FYI: Theory and typology of information packaging*. PhD. Diss, University of Amsterdam.
- Van der Wouden, Ton. 2015. A marked construction to mark a marked phenomenon: how to shift topics in Dutch, or, the Barabbas construction. In Sander Lestrade, Peter de Swart and Lotte Hogeweg (eds.), *Addenda: artikelen voor Ad Foolen*, 539-572. Radboud Universiteit Nijmegen.

- Van der Wouden, Ton. 2020. Er is nog steeds geen partikelwoordenboek, hoewel echter het partikelonderzoek niet stilligt. *Nederlandse Taalkunde* 25(2-3). 241-253.
- Vermeulen, Reiko. 2010. The syntax of topic, contrast and contrastive topic in Japanese and Korean. Paper presented at *On Linguistics Interfaces, II (OnLI II)*, University of Ulster.
- Wilson, Stephen and Saygin, Ayşe Pinar. 2001. Adverbs and Functional Heads in Turkish: Linear Order and Scope. *Proceedings of WECOL*, 1-12.
- Wolvengrey, Arok. 2005. Inversion and the Absence of Grammatical Relations in Plains Cree. In Casper de Groot and Kees Hengeveld (eds), *Morphosyntactic expression in Functional Grammar*, 419-445. Berlin: Mouton de Gruyter.