



UvA-DARE (Digital Academic Repository)

Transparency in Functional Discourse Grammar

Hengeveld, K.

Publication date

2011

Document Version

Final published version

Published in

Linguistics in Amsterdam

[Link to publication](#)

Citation for published version (APA):

Hengeveld, K. (2011). Transparency in Functional Discourse Grammar. *Linguistics in Amsterdam*, 4(2), 1-21. <http://linguisticsinamsterdam.nl/aclc/04/nr02/a01>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Transparency in Functional Discourse Grammar

Kees Hengeveld

Amsterdam Center for Language and Communication

Languages differ widely from one another in the extent to which they are transparent, i.e. obey one-to-one relationships between meaning and form. This paper sets up a framework in which these cross-linguistic differences can be studied systematically, using the theory of Functional Discourse Grammar. Transparent and transparent relationships are defined using the multi-level architecture of this model of language, representing these relationships as mappings between and within levels.

1 Introduction

Languages being first and foremost means of communication, one would expect them to be maximally transparent, in the sense of displaying a one-to-one relationship between meaning and form, in order to be successful instruments of interaction. Yet most of the languages of the world are less than fully transparent, and many of them even exhibit a high degree of opacity. This observation leads to two interesting general research questions. The first is if there is any systematicity in how languages may lose transparency, i.e. acquire opaque features. The second is how this systematicity can be explained. In order to answer the first question one has to compare language systems with higher and lower degrees of transparency with one another, in order to determine whether there are implicational relationships between transparent and opaque features across languages. In order to answer the second question one has to explore the implicational hierarchies following from the above in order to see how these can be related to other known communicative principles, such as pragmatic highlighting, economy, and processability.

In order to come to a systematic answer to these questions, transparency needs to be defined. Such a definition is bound to be more successful if it is implemented in a coherent framework that allows to formally define the units between which a one-to-one relation should exist. The framework adopted here is Functional Discourse Grammar (FDG, Hengeveld & Mackenzie 2008). This model is introduced in §2. FDG offers four different levels of analysis: an interpersonal, representational, morphosyntactic, and phonological level. In §3

transparency relations are defined as one-to-one relations between and within these four levels. This way a list of very precise relations between units can be defined for which languages may adopt a transparent or opaque solution. This list is provided in the concluding §4.

2 Functional Discourse Grammar

2.1 *FDG and verbal interaction*

As shown in Figure 1, FDG is conceived of as the Grammatical Component of an overall model of verbal interaction in which it is linked to a Conceptual Component, an Output Component and a Contextual Component. These three non-grammatical components interact in various ways with the Grammatical Component, more specifically with the operations of Formulation and Encoding. Formulation concerns the rules that determine what constitute valid underlying pragmatic and semantic representations in a language. Encoding concerns the rules that convert these pragmatic and semantic representations into morphosyntactic and phonological ones. FDG assumes that both Formulation and Encoding are language-specific, i.e. no universal pragmatic, semantic, morphosyntactic or phonological categories are postulated until their universality has been demonstrated through empirical research.

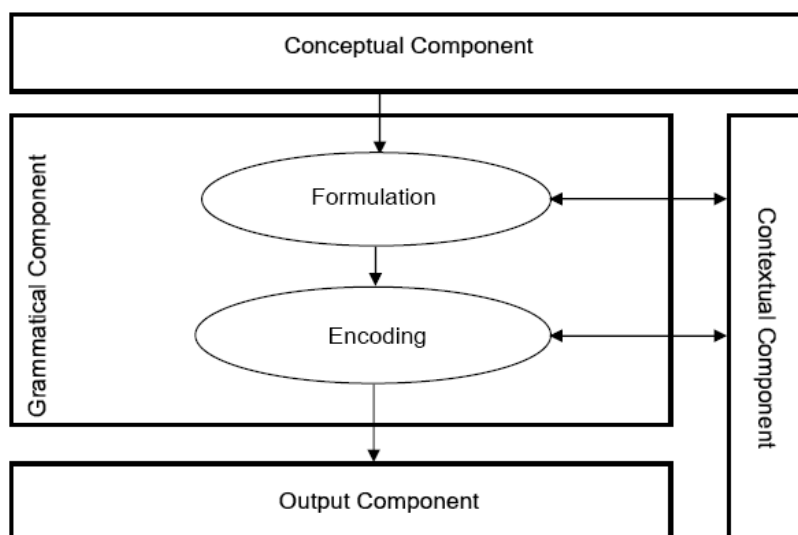


Figure 1. FDG as part of a wider theory of verbal interaction

The Conceptual Component is responsible for the development of both a communicative intention relevant for the current speech event and the associated conceptualizations with respect to relevant extra-linguistic events, and is thus

the driving force behind the Grammatical Component as a whole. The Output Component generates acoustic or signed expressions on the basis of information provided by the Grammatical Component. Its function may be seen as translating the digital (i.e. categorical, opposition-based) information in the grammar into analogue (i.e. continuously variable) form. The Contextual Component contains a description of the content and form of the preceding discourse, of the actual perceivable setting in which the speech event takes place, and of the social relationships between Participants. This type of information is relevant to many grammatical processes, such as narrative chaining, reflexives, and passives.

2.2. The architecture of FDG

The general architecture of FDG itself, in relation to the components that flank it, may now be represented as in Figure 2, in which the Grammatical Component is presented in the centre, the Conceptual Component at the top, the Output Component at the bottom, and the Contextual Component to the right.

A distinguishing feature of FDG shown in Figure 2 is its rigorous top-down architecture: FDG starts with the speaker's intention and works down to articulation. This is motivated by the assumption that a model of grammar will be more effective the more its organization resembles language processing in the individual. Psycholinguistic studies (e.g. Levelt 1989) clearly show that language production is indeed a top-down process. The implementation of FDG reflects this process and is organized accordingly.

In Figure 2 ovals contain operations, boxes contain the primitives used in operations, and rectangles contain the levels of representation produced by operations. I will describe the the general top-down process on the basis of a simple example, given in (1), produced in a context in which the Addressee wants to enter a field that hosts a bull:

- (1) There's a bull in the field!

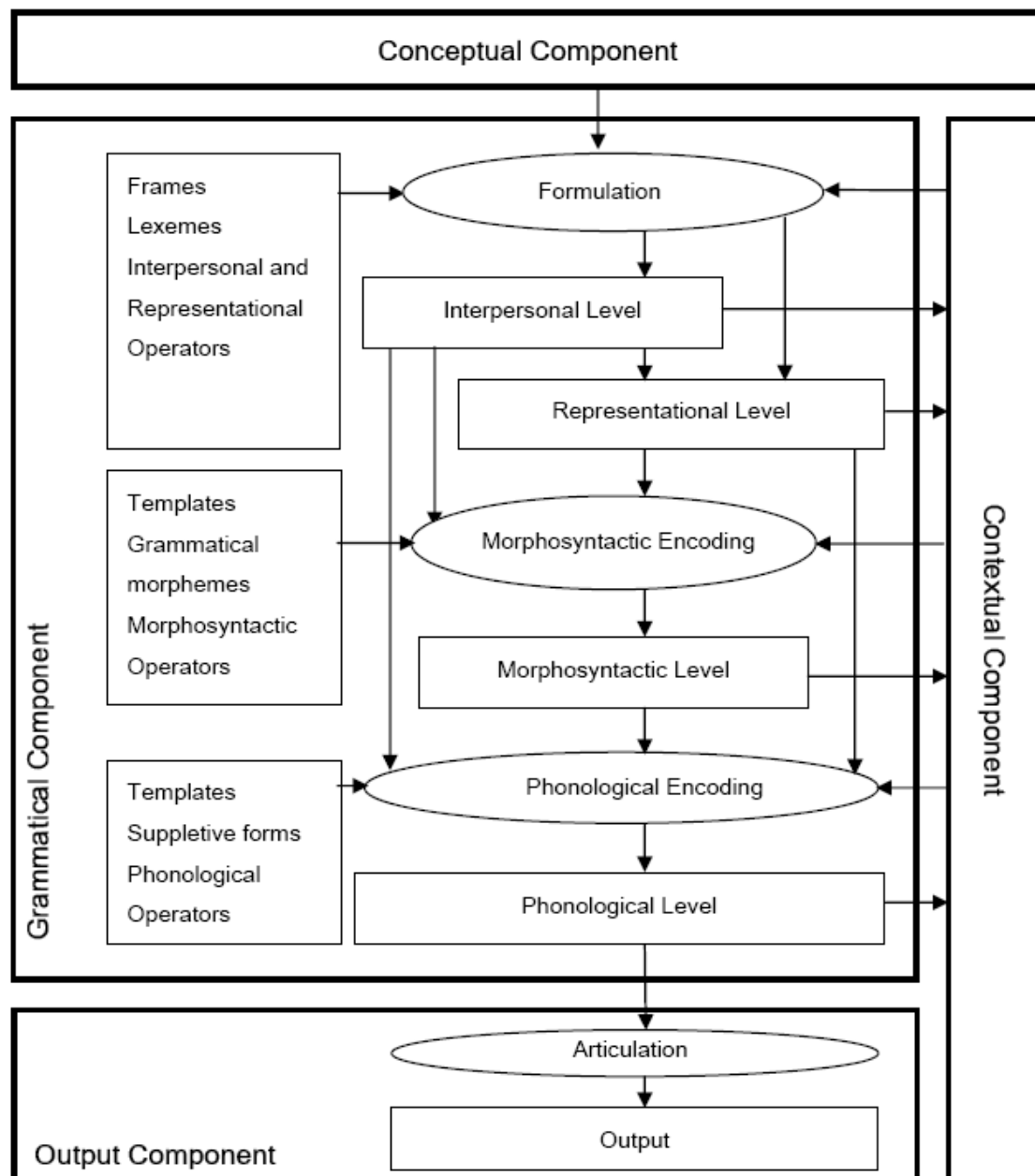


Figure 2. General layout of FDG

In the prelinguistic Conceptual Component a communicative intention (issuing a warning) and the corresponding mental representation (of the event causing danger) are relevant. The operation of Formulation translates these conceptual representations into pragmatic and semantic representations at the Interpersonal and the Representational Levels respectively. Warnings are not a separate illocutionary category in English, but the Speaker solves this problem by selecting a Declarative Illocution combined with an Emphatic operator at the Interpersonal Level. At the Representational Level the Speaker chooses to designate the entity causing danger as part of a locative predication frame. The configurations at the Interpersonal and the Representational Levels are

translated into a morphosyntactic structure at the Morphosyntactic Level through the operation of Morphosyntactic Encoding. In (1) this involves, for instance, the word order characteristic of existentials, the insertion of dummy *there*, etc. Similarly, the structures at the Interpersonal, Representational and Morphosyntactic Levels are translated into a phonological structure at the Phonological Level. In this case, for instance, the selection of the declarative illocution combined with an emphatic operator is responsible for the overall intonation contour with a high fall on the Focal Topic *bull*. By organizing the Grammatical Component in the way illustrated here, FDG takes the functional approach to language to its logical extreme: within the top-down organization of the grammar, pragmatics governs semantics, pragmatics and semantics govern morphosyntax, and pragmatics, semantics and morphosyntax govern phonology.

The Phonological Level of representation is the input to the operation of Articulation, which contains the phonetic rules necessary for an adequate utterance. Articulation takes place in the Output Component, outside the grammar proper.

The various levels of representation within the grammar feed into the Contextual Component, thus enabling subsequent reference to the various kinds of entity relevant at each of these levels once they are introduced into the discourse. The Contextual Component feeds into the operations of Formulation and Encoding, so that, for instance, the availability of antecedents may influence the composition of (subsequent) Discourse Acts.

2.3 *Levels and Layers*

2.3.1 *The Interpersonal Level*

The Interpersonal Level captures all distinctions of Formulation that pertain to the interaction between Speaker and Addressee. These cover, at the higher layers, rhetorical notions of the overall structuring of discourse, to the extent that they are reflected in linguistic form, and at the lower layers, the pragmatic distinctions that reflect how Speakers mould their messages in view of their expectations of the Addressee's state of mind, again only to the extent that these are grammatically relevant. The hierarchical structure arises through the application of an appropriate set of frames from those available to the Speaker. The following shows the hierarchical relationships that apply at the Interpersonal Level:

(2)	(M ₁ :	[Move
		(A ₁ :	[Discourse Act
			(F ₁)	Illocution
			(P ₁) _S	Speaker
			(P ₂) _A	Addressee
		(C ₁ :	[Communicated Content
			(T ₁)	Ascriptive Subact
			(R ₁)	Referential Subact
](C ₁))	Communicated Content
](A ₁))		Discourse Act
](M ₁))			Move

The Move (M₁) is the largest unit of interaction relevant to grammatical analysis. It may be defined as an autonomous contribution to the ongoing interaction: it either calls for a reaction, or is itself a reaction. A Discourse Act (A₁) is the smallest unit of communicative behaviour. It consists of an Illocution (F₁), the Speaker (P₁)_S, the Addressee (P₂)_A and a Communicated Content (C₁). The latter contains the totality of what the Speaker wishes to evoke in his/her communication with the Addressee. There are two types of Subact within the Communicated Content: an Ascriptive Subact (T₁) is an attempt by the Speaker to evoke a property, while a Referential Subact (R₁) is an attempt by the Speaker to evoke a referent.

2.3.2 *The Representational Level*

The Representational Level deals with the semantic aspects of a linguistic unit. Whereas the Interpersonal Level takes care of evocation, the Representational Level is responsible for designation. The use of the term 'semantics' is thus restricted to the ways in which language relates to the possible worlds it describes. The layers relevant at the Representational Level are defined in terms of the semantic categories they designate. Semantic categories are the language-specific linguistically relevant manifestations of ontological categories. They are hierarchically organized as indicated in (3):

(3)	(p ₁ : [Propositional Content
	(ep ₁ : [Episode
	(e ₁ : [State-of-Affairs
	(f ₁ : [Configurational Property
	(f ₁)	Lexical Property
	(x ₁)	Individual
	...	
] (f ₁))	Configurational Property
] (e ₁))	State-of-Affairs
] (ep ₁))	Episode
] (p ₁))	Propositional Content

Propositional Contents (p), the highest units at the Representational Level considered here, are mental constructs, such as pieces of knowledge, beliefs, and hopes. Propositional Contents contain Episodes (ep), which are sets of States-of-Affairs that are thematically coherent, in the sense that they show unity or continuity of Time (t), Location (l), and Individuals (x). States-of-Affairs (e) include events and states and are characterized by the fact that they can be located in time and can be evaluated in terms of their reality status. States-of-Affairs can thus be said to '(not) occur', '(not) happen', or '(not) be the case' at some point or interval in time. A State-of-Affairs is characterized by a Configurational Property (f), which is compositional in nature and contains a combination of semantic units that are not in a hierarchical relationship with respect to each other. Configurational Properties constitute the inventory of predication frames relevant to a language. Configurational Properties are built up using semantic categories that are in a non-hierarchical relationship with one another. These semantic categories may be of various types, and include Individuals (x), i.e. concrete objects that can be located in space, and Lexical Properties (f), which have no independent existence and can only be evaluated in terms of their applicability to other types of entity. Further semantic categories may be relevant to the grammar of an individual language and enter into the constitution of a Configurational Property, such as Location (l), Time (t), Manner (m), Reason (r), and Quantity (q). In all cases, only those semantic categories are postulated for a language that trigger formal processes within the grammar of that language.

2.3.4 The Morphosyntactic Level

The Morphosyntactic Level deals with the structural aspects of a linguistic unit. Together with the Phonological Level, it takes care of the encoding of interpersonal and representational distinctions. In view of this function, much of

what happens at the Morphosyntactic Level is functionally motivated: ordering principles are motivated by iconicity, domain integrity, and the preservation of scope relations. At the same time, morphosyntax has its own principles of organization, as for instance in the arbitrary imposition of a basic constituent order pattern, which in itself cannot be argued to be functionally motivated. FDG does not make a distinction between a syntactic and a morphological level of analysis, as the principles used in the formation of words are the same as those used in the formation of phrases and clauses.

The layers relevant at the Morphosyntactic Level are listed in (4):

(4)	(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

A Linguistic Expression (Le₁) is any set of at least one morphosyntactic unit; where there is more than one unit within a Linguistic Expression, these will demonstrably belong together in their morphosyntactic properties. The units combining into a Linguistic Expression may be Clauses, Phrases, or Words. By introducing Linguistic Expressions as the highest category in its morphosyntax, FDG creates a possibility of dealing straightforwardly with holophrases and non-sentential utterances. A simple Clause (Cl₁) is a grouping of one or more Phrases and possibly (grammatical) Words and is characterized, to a greater or lesser extent, by a template for the ordering of those Phrases and, also to a greater or lesser extent, by morphological expressions of connectedness (notably government and agreement). A Phrase (Xp₁) is headed by a lexical item that is passed on from the Interpersonal Level or the Representational Level. The Word (Xw₁), especially in incorporating languages, can be highly complex. Apart from the fact that it may consist of Stems (Xs) and Affixes (Aff), in some languages it may, just like any other layer of morphosyntactic analysis, embed higher layers such as phrases and clauses, obeying full recursivity.

2.3.5. The Phonological Level

The Phonological Level is responsible for every aspect of Encoding not covered by the Morphosyntactic Level. It receives input from all three other levels and provides input to the Output Component. Whereas the latter deals with such ‘analogue’ matters as formant frequency, intensity, duration and spectral characteristics, the Phonological Level – being within the grammar – is ‘digital’, containing representations in phonemes that are ultimately based in binary phonological oppositions. Just like the other levels, phonological representations are hierarchical in nature (as in the tradition of Prosodic Phonology initiated by Nespor & Vogel 1986). Here too, FDG makes the assumption that not all layers are active in every Utterance or indeed are relevant to every language system. And as at the Morphosyntactic Level, FDG does not exclude the possibility of recursion at certain layers. The maximum layering of the Phonological Level is as follows:

(5)	(U ₁ : [Utterance
	(IP ₁ : [Intonational Phrase
	(PP ₁ : [Phonological Phrase
	(PW ₁ : [Phonological Word
	(F ₁ : [Foot
	(S ₁) ⁿ	Syllable
] (F ₁)	Foot
] (PW ₁)	Phonological Word
] (PP ₁)	Phonological Phrase
] (IP ₁)	Intonational Phrase
] (U ₁)	Utterance

The Utterance (U₁) is the largest stretch of speech covered by the Phonological Level. A Speaker will tend to use more substantial pauses to separate Utterances than Intonational Phrases. An Utterance may in addition display pitch distinctions called paratones which help to mark it off as a self-contained group of Intonational Phrases. The Intonational Phrase (IP₁) is characterized by a nucleus, i.e. a pitch movement localized on one or more Syllables which is essential to the interpretation of the Intonational Phrase as a whole. The Phonological Phrase (PP₁) in stress languages contains one Syllable that is more strongly stressed than the others; this Nuclear Syllable is typically also the primary location for the global fall or rise within the Intonational Phrase. In tone languages, in which pitch movement is used for lexical distinctions, Phonological Phrases have a different *raison d’être*, namely as the domain of tone sandhi. The Phonological Word (PW₁), for those languages in which such a

category needs to be recognized, is a slice of phonological structure which displays at least one criterial characteristic, which may relate to the number of segments, to prosodic features or to the domain of phonological rules. Phonological Words are divided into Syllables (S_1), which in stress languages (i.e. those with stressed and unstressed Syllables) group into Feet (F_1).

3 Transparent and opaque mappings

3.1 Introduction

Transparency or the lack thereof may obtain between all the levels shown within the Grammatical Component in Figure 2. It may also obtain within each of the encoding levels: the Morphosyntactic Level the Phonological Level. I will discuss relations between levels in §3.2 and within levels in §3.3.

3.2 Transparent relations between levels

3.2.1 Introduction

In principle, relations may obtain between any pair of the four levels in Figure 2. These are (i) Interpersonal-Representational, (ii) Interpersonal-Morphosyntactic, (iii) Interpersonal-Phonological, (iv) Representational-Morphosyntactic, (v) Representational-Phonological, and (v) Morphosyntactic-Phonological. These will be grouped here in terms of the endpoint of the relation, where the endpoint is determined on the basis of FDG's top down perspective. Such groupings are warranted by the fact that what is passed on to the Morphosyntactic Level is the cumulative result of the Interpersonal and Representational Levels, and that what is passed on to the Phonological Level is the cumulative result of the Interpersonal, Representational, and Morphosyntactic Levels. This leads to the following groupings:

- Interpersonal-Representational (§3.2.2),
- Interpersonal/Representational-Morphosyntactic (§3.2.3),
- Interpersonal/Representational/Morphosyntactic-Phonological (§3.2.4).

3.2.2 Interpersonal-Representational

- No apposition

A transparent relation between the Interpersonal and Representational Level obtains when a single Subact at the Interpersonal Level corresponds to a single semantic category at the Representational Level. Such a relation is absent in

cases of apposition, as in the following example from Sri Lankan Malay (Nordhoff this volume):

- (6) Mr Sebastian aada, se aada kitham duuva arà-oomong.
 Mr Sebastian exist 1.SG exist 1.PL two NON.PAST-speak
 ‘You are here, I am here, the two of us are talking.’

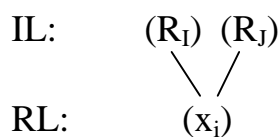
Here both *kitham* ‘1.PL’ and *duuva* ‘two’ refer to the same entity, and might do so independently of one another. Thus there are two Referential Subact that correspond to a single semantic category.

A special case of apposition is the one in which a referential pronominal marker on a verb cooccurs with a verb-external noun phrase, as in the following example from Chickasaw (Munro & Gordon 1982: 110):

- (7) Aboha anõ’k-akõ Dan ib-aa-binni’li-li-tok.
 house in-CONTR.NONSUBJ Dan COM-LOC-sit-1.SG.A-PST
 ‘I sat with Dan in the house.’

Within the verb form the prefixes *ib-* ‘COM’ and *aa-* ‘LOC’ refer independently to company and location respectively, just as the suffix *-li* ‘1.SG.A’ is capable of independently referring to a first singular actor. The comitative and locative prefixes are, however, accompanied by external constituents lexically specifying company and location, thus providing a second reference to these semantic categories.

Apposition, including crossreference as illustrated in (7), may be represented as:



- No limitations on which semantic units can be chosen as predicates

Predication is implemented in FDG through Ascriptive Subacts. One would expect these to be allowed to pick any semantic category and turn it into a predicate. However, most languages do not allow just any category to be used predicatively. Kharia (Peterson 2006: 60; Leufkens this volume) is an exception:

- (8) Lebu $\text{d}\text{el}=\text{ki}$.
 man come=M.PST
 ‘The man came.’
- (9) Bhagwan lebu=ki ro $\text{d}\text{el}=\text{ki}$.
 God man=M.PST and come=M.PST
 ‘God became man [=Jesus] and came [to earth].’

The lexeme *lebu* ‘man(hood)’ is used referentially in (8) and predicatively in (9). In contrast, only the first use would be allowed in the Dutch equivalents of (8) and (9):

- (10) De man kwam.
 DEF man come.PST.SG
 ‘The man came.’
- (11) *God man-de.
 God man-PST.SG
 ‘God became man.’

This restriction may be represented as:

$$\begin{array}{l} \text{IL:} \quad (\text{T}_I) \\ \quad \quad \downarrow \\ \text{RL:} \quad (\alpha_i) \end{array}$$

in which α is a variable ranging over semantic categories. This indicates that certain semantic categories cannot be used in a certain interpersonal function, thus blocking a potential one-to-one relationship.

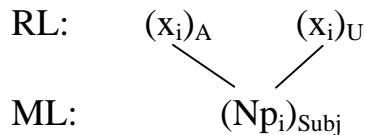
3.2.3 *Interpersonal/Representational-Morphosyntactic*

- No grammatical relations

A grammatical alignment system lines up pragmatic/semantic and syntactic units in a non-transparent way, in the sense that a semantic/pragmatic unit may receive different expression depending on the syntactic configuration. A transparent alignment expresses pragmatic/semantic units always in the same way, independently of the syntactic configuration. The latter is the case in Acehnese (Durie 1985: 55-58), as illustrated in (12)-(13):

- (12) Lôn teungöh=lôn=jak.
 1 M=1.A=go
 'I am going.'
- (13) Gopnyan galak=geuh that.
 3.POL happy=3.POL.U very
 'He is very happy.'

In Acehnese Actors are always expressed by a proclitic, Undergoers by an enclitic, the forms of the markers being identical for every person. Thus, the two intransitive constructions in (12) and (13) are realized differently based on semantic considerations. The translations show that in English the distinction between Actors and Undergoers is neutralized through the application of the grammatical relation Subject. This non-transparent situation may be represented as:

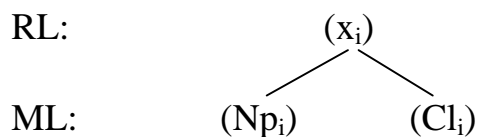


- No discontinuity

One would expect that in a transparent language that which belongs together is expressed together, in accordance with Behaghel's (1932) first law. A discontinuous configuration is illustrated in (14) (Hengeveld & Mackenzie 2008):

- (14) The guy has arrived who is going to fix my lock.

Here the semantic constituent *the guy who is going to fix my lock* is interrupted by the verbal complex *has arrived*. Such an opaque situation may be represented as:



- Function marking not sensitive to nature of input

A language is transparent if a certain function is always marked in the same way, independently of the nature of the constituent on which this function is to be expressed. A clear example of this can be found in Nama, where the

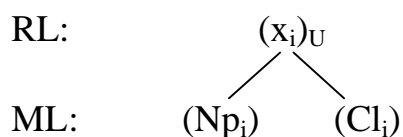
Undergoer marker =à attaches to noun phrases as well as to clauses (Hagman 1977: 54, 138):

- (15) ||'iĩp ke 'áop=à kè †áí.
 3.SG.M DECL man=ACC REM.PAST call
 'He called the man.'
- (16) Siíx̄m ke kè ||nàú ||'iĩp kò !úú !xáís=à .
 1.PL.M.DU DECL REM.PAST hear 3.SG.M REC.PAST go COMP=ACC
 'We heard that he had just left.'

The fact that the same marker attaches to a noun in (15) and a clause in (16) leads to a completely transparent marking of the function of the constituent, irrespective of its internal constitution. Transparent languages thus make use of phrase marking rather than of head marking. The latter is illustrated for Finnish (Sulkala & Karjalainen 1992: 212,) in (17)-(18):

- (17) Varas ott-i-∅ lompako-n.
 thief take-IMPF-3.SG wallet-ACC
 'The thief took the wallet.'
- (18) Sairas valitt-i-∅ että kurkku on kipeä.
 patient complain-IMPF-3.SG COMP throat COP.3.SG sore
 'A patient complained that his throat was sore.'

The head of the noun phrase in (17) is marked by an accusative suffix, while the complement clause in (18) is not marked at all for its function. This may be represented as in (where U stands for the function of Undergoer):



3.2.4 *Interpersonal/Representational/Morphosyntactic-Phonological*

- Phonological phrasing and morphosyntactic phrasing run parallel

A language is transparent if there is a one-to-one mapping between morphosyntactic and prosodic units. Such is for instance the case in Achenese (Durie 1985: 29-30), in which every Lexeme corresponds to a Phonological Word and every Morphosyntactic Phrase corresponds to a Phonological Phrase. Phonological Words in Achenese are characterized by the fact that they have a

word final word stress position which is realized when the word is the stressed word in a phrase. Phonological Phrases are characterized by the fact that they contain a single stressed word and can be separated from other phrases by a pause. The following example illustrates this:

- (19) Ureueng='nyan ka=geu=jak='woe ba'roe.
 person=DEM INCH=3=go=return yesterday
 'That person returned yesterday.'

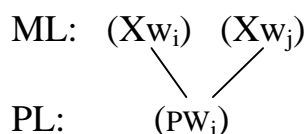
In (19) there is a one to one mapping between the NP, the VP and the AdvP on the one hand, and the three phonological phrases on the other hand. Quite the opposite is the case in Dutch, as shown in (20):

- (20) Ik wou dat hij kwam.
 I want.PST COMP he come.PST
 'I wish he would come.'

The phonological phrasing of this example is as in (21):

- (21) (PP_i: (PW_i: -kʋɑu- (PW_i)) (PP_i)) (PP_j: [(PW_j: -dati- (PW_j)) (PW_k: -kʋɑm- (PW_k))] (PP_j))

which shows, for instance, that the Subject and the verb form one Phonological Word, just as the conjunction and the subordinate subject, thus leading to a serious mismatch between morphosyntactic and phonological phrasing. The latter situation may by way of illustration be represented as:



- Phonological weight does not influence morphosyntactic placement

It is a rather common phenomenon that phonological weight influences syntactic placement. For instance, in Spanish an Undergoer that is expressed by a clitic precedes the verb, while an Undergoer expressed by a lexical noun phrase follows it, as shown in (22) and (23):

- (22) Lo=ví.
 3.SG.ACC=see.PRF.PST.IND.3.SG
 'I saw him.'

- (24) Marami-ng pera.
 lot-LNK money
 ‘There is a lot of money.’
 ‘A lot of money’

An existential constructions in Tagalog consists in the naming of the existing object, without the use of expletive elements. English, as can be seen in the translation, uses two dummy elements, *there* indicates the absence of a predicate, *be* carries the tense specification. The latter, opaque, situation may be represented as:

ML: $\emptyset \rightarrow$ Expletive

- no tense copying

A transparent indirect speech report is one in which the tense of the embedded verb is the one that was used by the original speaker. Such a situation obtains in Amele (Roberts 1987: 48):

- (25) Naus uqa ege [qila bele-q-an fo=ec] sisil-t-en.
 Naus he I today go-1.PL-FUT Q=NMLZ ask-1.SG/3.SG-REM.PST
 ‘Naus asked me whether we would go today.’

The future tense expressed by the original speaker is here retained in the embedded clause. Some languages, however, apply a rule of sequence of tenses or tense copying, leading to situations as the one represented by the English translation in English translation in (25). This may be represented as:

ML: Tense sub \rightarrow Tense main + Tense sub

- no raising

Raising of a constituent of an embedded clause to a main clause position leads to discontinuity and therefore to lack of transparency. This is the case in the following example from Spanish:

- (26) Los profesor-es parec-en ser inteligentes.
 DEF.PL.M teacher-PL seem-PRES.IND.3.PL COP.INF intelligent-PL
 ‘The teachers seem to be intelligent.’

A more transparent counterpart of this example would be:

- (27) Parec-e que los profesor-es son
 seem-PRES.IND.3.SG COMP DEF.PL.M teacher-PL COP.PRES.IND.3.PL
 inteligent-es.
 intelligent-PL
 ‘It seems that the teachers are intelligent.’

The non-transparent situation illustrated in (26) may be represented as:

ML: Clause → NP + Clause

- no grammatical gender, declination, conjugation

The organization of a word class into grammatical genders and declination and conjugation classes leads to an arbitrary subdivision of words that does not correspond to their meaning or function. Thus in Spanish the word *casa* ‘house’ is arbitrarily assigned to the class of feminine nouns, and the word *árbol* ‘tree’ to the class of masculine nouns. This arbitrary classification leads to a non-transparent situation that may be represented as:

ML: word class → word subclasses

- no agreement

Languages exhibiting agreement copy a feature of one constituent to another one, thus creating double marking and multiple forms for the same meaning. The two Spanish words just used to illustrate arbitrary gender assignment trigger different forms of agreement:

- (28) la-∅ casa-∅ viej-a-∅
 DEF.F-SG house(F)-SG old-F-SG
 ‘the old house’

- (29) el árbol-∅ viej-o-∅
 DEF.M-SG tree(M)-SG old-M-SG
 ‘the old tree’

The article and the adjective thus exhibit two different forms that do not correlate with a difference in meaning. Furthermore, a single feature is expressed more than once.

- no fusional morphology

In fusional morphology the boundaries between forms are not respected. This leads to a situation in which two or more aspects of meaning are expressed in a single, fused, form. This is a non-transparent of languages. Fusion may occur in two different forms. In stem alternation the stem of a lexeme has a different form in different morphological circumstances. This is for instance the case in the Wambon example (30) (de Vries 1989: 23):

- (30) en- ande- na-
 eat(basic stem) eat(PAST/FUT/IMP.PL stem) eat(IMP.SG stem)

A second type of fusion is *cumulation*, which concerns the simultaneous expression of multiple grammatical meanings in a single morpheme. A case in point is the Spanish morpheme *-é* in (31), which simultaneously expresses the notions indicative, past, perfective, first person, and singular.

Spanish

- (31) compr-é.
 buy-IND.PAST.PF.1.SG
 ‘(I) bought.’

3.3.3 The Phonological Level

At the Phonological Level too a number of operations take place that add elements or features to the structures that have been built up on the basis of material handed over from the higher components. This is the case when phonological rules apply that adapt an underlying phoneme to its phonological environment. A number of examples from a wide array of possibilities are Quechua nasal assimilation (32, Grández Ávila, this volume), Spanish diphthongization (33), Dutch degemination (34), and Turkish vowel harmony (35):

- (32) *tayta-n=paq* ‘father-3.POSS=PURP’ ‘for his father’ → *taytampaq*
 (33) *dormir* ‘sleep’ *duerme* ‘sleeps’
 (34) *pakkans* ‘chance to be caught’ → *pakans*
 (35) *gel-miş* ‘come-RES’, *gör-müş* ‘see-RES’

In all these cases the a one-to-one relation between meaning and form is obscured.

4 Summary and outlook

In the preceding sections a long list of transparent properties of languages was given, and it was shown how these properties may be violated. The complete list of properties is given in Table 1.

Table 1: Transparent mappings between and within level(s)

Interpersonal - Representational
- No apposition
- No limitations on which semantic units can be chosen as predicates
Interpersonal/Representational - Morphosyntactic
- No grammatical relations
- No discontinuity
- Function marking not sensitive to nature of input
Interpersonal/Representational/Morphosyntactic - Phonological
- Phonological phrasing and morphosyntactic phrasing run parallel
- Phonological weight does not influence morphosyntactic placement
Morphosyntactic
- no expletive elements
- no tense copying
- no raising
- no grammatical gender, declination, conjugation
- no agreement
- no fusional morphology
<i>Phonological</i>
- no phonological adaptation

In the following papers four languages will be subjected to the transparency test using the criteria listed in Table 1. These languages are Esperanto, Kharia, Quechua, and Sri Lankan Malay. The ensuing epilogue will try to give a tentative answer to the question whether an implicational hierarchy can be established that expresses the extent to which languages resist opaqueness as regards these features. The resulting tentative hierarchy will help us understand both the degree of opaqueness associated with each feature and the degree of opaqueness of each of the languages studied.

References

- Behaghel, Otto (1932), *Deutsche Syntax: Eine geschichtliche Darstellung. Band IV: Wortstellung-Periodenbau*. Heidelberg: Carl Winter.
- Durie, Mark (1985), *A Grammar of Acehnese: On the Basis of a Dialect of North Aceh* (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 112). Dordrecht: Foris.
- Hagman, Roy Stephen (1973), *Nama Hottentot Grammar*. PhD dissertation, Columbia University.
- Hengeveld, Kees & Mackenzie, J.Lachlan (2008), *Functional Discourse Grammar: A typologically-based approach to language structure*. Oxford: Oxford University Press.
- Levelt, Willem J.M. (1989), *Speaking*. Cambridge MA: MIT Press.
- Munro, Pamela & Gordon, Lynn (1982), Syntactic relations in Western Muskogean. *Language* 58: 81–115.
- Nespor, Marina & Vogel, Irene (1986), *Prosodic Phonology*. Dordrecht: Foris.
- Peterson, John (2006), *Kharia. A South Munda Language*. Habilitationsschrift, Universität Osnabrück.
- Roberts, John R. (1987), *Amele* (Croom Helm Descriptive Grammars Series). London: Croom Helm.
- Schachter, Paul & Otnes, Fe T. (1972), *Tagalog Reference Grammar*. Berkeley CA: University of California Press.
- Sulkala, Helena & Karjalainen, Merja (1992), *Finnish* (Descriptive Grammars). London & New York NY: Routledge.
- Vries, Lourens de (1989), *Studies in Wambon and Kombai*. Dissertation, University of Amsterdam.

Contact information:

Kees Hengeveld

Amsterdam Center for Language and Communication, University of Amsterdam

Spuistraat 210, NL-1012 VT Amsterdam

p.c.hengeveld@uva.nl