Transparency in complementation

Hengeveld, K.; Luberti, M.

Publication date
2020

Document Version
Final published version

Published in
Linguistics in Amsterdam

License
Unspecified

Link to publication

Transparency in complementation[*]

Kees Hengeveld
ACLC, University of Amsterdam

Merel Luberti
ACLC, University of Amsterdam

In this paper, we investigate to what extent the degree of transparency of a language (as defined in Hengeveld & Leufkens 2018) determines its system of complement clauses (CCs), both as regards the number of CCs a language has, and as regards the types of CCs it displays. The investigation of a sample of nine languages with different degrees of transparency reveals that the more transparent a language is, the fewer CC types it displays, and the more their CCs are like main clauses.

1 Introduction

Language tends "to strive to maintain a one-to-one mapping between underlying semantic structures and surface forms, with the goal of making messages easily retrievable for listeners" (Slobin 1977: 186). In other words, in order for a speaker to get their message across as effectively as possible, a language must be transparent. It seems, however, that in natural languages a one-to-one mapping between meaning and form rarely exists. Rather, languages often allow for ambiguity, mismatches, and other features that make language less transparent. The lack of transparency in language has become a topic of interest for many theoretical linguists as first language acquisition studies report that transparent features of a language are easier to acquire than features that are opaque (Slobin 1977).

In a recent study on the transparency of languages, Hengeveld & Leufkens (2018, see also Leufkens 2015) use the theory Functional Discourse Grammar (FDG, Hengeveld & Mackenzie 2008) to examine the notion of transparency allowing them to define transparency in terms of the interaction between levels within FDG: the Interpersonal (pragmatic), Representational (semantic),

Morphosyntactic and Phonological Levels. The theory is used to systematically establish transparent and non-transparent features of languages. This framework is then applied to a sample of 30 languages resulting in the establishment of an implicational transparency hierarchy, and a ranking of the 30 languages in terms of their degree of transparency.

The present study aims to extend the research of Hengeveld & Leufkens (2018), by studying the transparency of embedded clauses, more specifically complement clauses: Are these structurally the same as or different from main clauses? Are there different types of complement clauses? If the latter is the case, the language may be considered less transparent than a language exhibiting a single complement clause type, especially when this clause type is identical in form to the main clause. This lack of transparency may take different forms. First of all, the fact that a language has multiple manifestations of a single morphosyntactic unit within the Morphosyntactic Level may be considered opaque by itself. Opacity increases furthermore when multiple clause types are used to express the same meaning unit, as in (1):

(1)   a.   I believe that John is ill.
       b.   I believe John to be ill.

We will limit ourselves here to the first type of opacity, as the data provided by reference grammars do not allow us to focus on the second type.

A language exhibiting multiple types of complement clauses may be said to exhibit *syntactic suppletion*[1] (Hengeveld & Mackenzie in prep.). Just as in the case of morphological suppletion, in which there is morphologically conditioned alternation of a stem, there are different variants of a syntactic construction, the use of which is required in certain circumstances. A language with syntactic suppletion can therefore be considered less transparent than a language exhibiting a single complement clause type, especially when this clause type is identical in form to the main clause. In order to address these general questions, a typological study is carried out.

The present paper is structured as follows. First, we introduce the notion of transparency in Section 2. In Section 3, we then define embedded clauses and more specifically elaborate on complement clauses. The method of this study will be discussed in Section 4, in which we list the languages to be examined, and note the selection criteria for the inclusion of these languages. The research question and hypothesis will also be stated in this section. The analysis of the data is then presented in Section 5, and the results in Section 6. In Section 7, we situate the

---

[1] Alternatively, one could call these constructions 'syntactically conditioned alloclauses' (Sterre Leufkens, p.c.).

transparency feature investigated in this paper in Hengeveld & Leufkens' (2018) transparency hierarchy, and in Section 8, we present our conclusions.

2 Transparency

2.1 Simplicity vs transparency

As mentioned by Slobin (1977), language should strive to be as transparent as possible, as transparent features of a language are the features most easily acquired during first language acquisition. Transparency, according to Slobin (1977), denotes the one-to-one relation between form and meaning. In contrast, features that prove to be more difficult to acquire are those that are opaque and thus do not exhibit a one-to-one relation between form and meaning. In order to elucidate the notion of transparency, it is important to distinguish the notion of transparency from the notion of simplicity. Transparency concerns the mapping between aspects of form and aspects of meaning, while simplicity is measured in terms of the number of distinctions or categories present in a language in a certain domain (Leufkens 2015: 39). Slobin (1977) illustrates this by comparing the L1 acquisition of the Turkish inflectional system to the Serbo-Croatian inflectional system. The Turkish language is said to be morphologically complex. This is due to the large number of suffixes present in the language that can be attached to a verbal word. However, this language is also said to be highly transparent as each suffix is clearly separable and has a single meaning. Despite the language being morphologically complex, Turkish children are able to acquire and command the inflectional morphemes "well before the age of 2" (Slobin 1977: 190). In contrast to the Turkish inflectional system, the Serbo-Croation one is more opaque. For instance, the inflectional system is sensitive to phonology, and grammatical gender also influences the choice of inflection. These features make the language more opaque, resulting in children not fully acquiring the system "until about age 5 – 3 years or so later than the Turkish child" (Slobin 1977: 191).

On the other hand, Slobin (1977) notes that embedded clauses in Serbo-Croatian are very transparent, in the sense that they generally take the same form. In Turkish, however, there is a wide variety of verb forms used in embedded clauses, resulting in a much less transparent system. This opacity evidently affects the rate in which these embedded structures are acquired. Turkish children tend to master these constructions around the age of 5 (Slobin 1977), while the Serbo-Croatian children acquire embedded clauses much earlier, at around the age of two (Slobin 1977) due to their transparent structure.

3   Complement clauses

3.1   Introduction

When a clause is contained within another clause, the former clause is said to be embedded within the latter. This process is called embedding or subordination, and the clause that is embedded is called the embedded clause or the subordinate clause. There are several types of embedded clauses: complement clauses, relative clauses, and adverbial clauses. [2] Noonan defines complementation as "the syntactic situation that arises when a notional sentence or predication is an argument of a predicate" (Noonan 2007: 52). In other words, the function of the complement clause (CC) is to be an argument of the complement taking predicate (CTP) (Yule 2016). For example, complement taking predicates like *wish* and *promise* are those that need a clausal argument to complete it (as *somebody* wishes/promises *something*). By way of illustration, the following example contains a transitive predicate, the second argument of which is expressed by a CC:

(2)   Jane thought [she should leave].

The CTP of (1) is *think*. The CC, in brackets, functions as the object of the predicate.

        Adverbial (3) and relative (4) clauses differ from CCs in not being required to complete the predicate:

(3)   Olaf bought the bananas [when he went to the grocery store].
(4)   The girls [that didn't go to the party] stayed home.

3.2   Finite and non-finite verb forms

The verb form in an embedded clause can be finite or non-finite. Finiteness is usually simply defined as the verb carrying tense (Aarts 2018). An alternative definition is for the verb to carry tense, aspect, mood, and person agreement. The verb that does not carry these distinctions is then said to be non-finite. However, this definition seems to be too narrow to define the notion of finiteness cross-linguistically, as languages differ widely as regards the marking of tense, mood, and aspect (TMA) distinctions and agreement on verbs. To solve this problem, Hengeveld (1998) offers a functional classification of verb forms and proposes that a verb form can be either independent or dependent. A verb form can be said to be independent when it can be used in a main clause. A dependent verb form is

---

[2] A further type of embedded clause is the *predicate clause*, a situation that occurs when the predicate of a clause is itself a clause, as in *Driving too fast is risking your life.*

then one that may only occur in a subordinate clause. In this paper, the definition of independent and dependent verb forms is adopted to replace the distinction between finite and non-finite verb forms.

3.3   *Argument and adjunct realization*

Just as verbs can be realized dependently and independently, so can arguments. In an embedded clause, arguments and adjuncts can be realized as they would be in main clauses, or they can be realized differently. The latter situation arises in certain types of nominalizations in English, as illustrated in (5), as compared to the corresponding main clause in (6):

(5)   [Peter's unexpected leaving of the meeting] was unfortunate.
(6)   Peter left the meeting unexpectedly.

In (5) the Subject and Undergoer arguments are realized with genitive marking as opposed to zero nominative/accusative marking, and the adjunct is realized as an adjective rather than as an adverb. As the verb in (5) takes a dependent form as well, one can say that every single constituent of the embedded clause in (5) is realized differently from its main clause equivalents.

This discussion shows that, parallel to the distinction between dependent and independent verb forms, it is useful to distinguish between dependent and independent argument/adjunct forms.

3.4   *Complement clause types*

CCs come in various forms that differ from language to language. As Noonan (2007) notes, the CCs in English are of four types, given in (7):

(7)   a.   It is a pity [that Mary ate the cake].
            (that-clause)
       b.   I want [Mary to eat the cake].
            (infinitival clause)
       c.   I saw [Mary eating the cake].
            (participial clause)
       d.   [Mary's eating of the cake] was a mistake.
            (nominalized clause)

The CCs in (7) differ from one another in significant ways. The CC introduced by *that* in (7a) differs from the other three in containing an independent verb form, i.e. one that could be used in a main clause as well. Furthermore, the arguments are realized as they would be in a main clause, hence they have independent argument/adjunct forms.

The infinitival CC in (7b) partly resembles the clause in (7a), as the arguments of the infinitive are realized in the same way as they would be in a main clause, i.e. they have independent forms. The verb, however, takes a dependent form.

The participial CC in (7c) is one that contains a participle, which is an adjectival or an adverbial form of a verb (Noonan 2007). In certain circumstances, participles may develop into verb forms used in CCs as well, especially with verbs of perception (see Dik & Hengeveld 1991: 250-255). The participial construction resembles the infinitival one in (7b), as the verb form is dependent, while arguments and adjuncts are realized independently, i.e. in the same form as they would be in main clauses.

The nominalized complement in (7d) is a CC the structure of which resembles that of a noun phrase. The predicate takes the form of a verbal noun and hence is a dependent verb form. The arguments have a dependent realization too, as they are realized as genitival phrases.

Another type of CC according to Noonan (2007) is called parataxis. The construction "involves the juxtaposition of two clauses within the same sentence" (Noonan & Bavin 1981: 45). The construction may occur with or without a (coordinating) conjunction. The definition of the construction does not seem to be constant across linguists, as Aarts (2018: 45) claims that the paratactic construction functions to link "units of equal syntactic status". Similarly, Lehmann (1988: 3) defines parataxis as "the coordination of clauses". It is true that the juxtaposition of two clauses can fulfill a function similar to complementation, as shown in (8) and (9):

(8)   I left. He forced me.
(9)   He forced me to leave.

However, the construction in (8) does not contain an embedded clause but juxtaposes two clauses of the same rank. We therefore exclude paratactic constructions from our analysis.

3.5   *Complementizers*

It is often said that a CC may be best identified through complementizers, which may take the form of clitics, particles or affixes (Noonan 2007; Yule 2016). They help speakers identify a complement within a sentence. An example is the complementizer *that* which introduces a *that*-clause. Using a complementizer to introduce a CC may be optional, or may come with certain constraints. The complementizer *to* for example, used in infinitival CCs exhibits this optionality. Noonan (2007) notes that the use of this complementizer is not obligatory, except when it occurs in object position with certain complement taking predicates. CTPs

like *want*, *force*, and *allow* need the complementizer in order for the argument to be grammatical:

(10) a. Evelle forced Jerry to change his plans.
      b. *Evelle forced Jerry change his plans.   (Noonan 2007: 56)

Some complementizers may be specific to the type of clause they occur with, while others may occur with several complement types. Lango, for example, has only one complementizer (*nî*) that is used with both subjunctive and indicative complement clauses:

Subjunctive:
(11) *Rwòt òdìò   lócà   nî   òryêt.*
      king 3SG.PRES.PERF   man   COMP   3SG.winnow.SUBJ
      'The king pressed the man to winnow.'   (Noonan 1992: 222)

Indicative:
(12) *Rwòt   bínô   tàmmò   nî   lócà   òdɔk*
      king   3SG.COME.HAB   think.INF COMP   man   3.SG.go.back.PERF
      yesterday
      àwó'ró.
      'The king will think that the man went back yesterday'
                                                      (Noonan 1992: 220)

Since complementizers are not constituents of the complement clause, which itself is a complement of the complementizer, we will not take complementizers into consideration in classifying our data.

3.6   *The definition and classification of CCs used in this paper*

On the basis of the considerations above, we define complement clauses (CCs) as those clauses that are embedded in a main clause, and that are required by, that is, an argument of, the complement taking predicate (CTP). This means that clauses that are not needed to complete the CTP, such as paratactic clauses, are excluded from our analysis.

For the classification of CCs, we make use of two parameters. One concerns the distinction between dependent and independent verb forms, the other the one between dependent and independent argument/adjunct forms. The combination of these two parameters leads to the following classification:

Independent verb form
      Independent argument/adjunct encoding
      Dependent argument/adjunct encoding
Dependent verb form
      Independent argument/adjunct encoding
      Dependent argument/adjunct encoding

One would expect all these four logically possible combinations to exist. However, in our sample, we have not come across situations in which an independent verb form combines with dependent argument/adjunct encoding. This may have to do with the fact that dependent argument encoding manifests itself in the context of nominalization. It thus seems that independent verb form encoding implies independent argument/adjunct encoding. The actual classification to be used in the rest of this paper is therefore the following:

Independent verb form
Dependent verb form
      Independent argument/adjunct encoding
      Dependent argument/adjunct encoding

4   Methodology

The sample to be studied in this paper consists of nine languages. The languages are selected based on the sample of Hengeveld & Leufkens (2018). Their language sample consists of 30 languages, which they chose based on the selection criteria of Rijkhoff et al. (1993). The sample of Hengeveld & Leufkens (2018) thus consists of languages that were selected on the basis of their genetic and geographical independency, as well as their typological characteristics, including "fusional, agglutinative, and isolating languages, as well as polysynthetic ones" (p. 159). Moreover, languages that are known to have a high degree of transparency, like Sri Lanka Malay and Japanese, and languages that are known to exhibit a high degree of opacity, like French and Georgian, were included.

      The study by Hengeveld & Leufkens (2018) leads to a classification of languages in nine different classes as regards their degree of transparency. The sample to be used for the present study takes one language from each of these nine classes. The choice for a specific language from each class was determined by the availability of the data required. The languages selected are given in Table 1, in which the languages listed in the second column belong to the language families given in the first column. The data gathered for the nine languages were

obtained by analyzing grammars. The third column lists the grammar(s) consulted.

Table 1: Sample languages

Language Family	Sample Language	Grammar(s) consulted
Afro-Asiatic	Arabic, Egyptian	Gary & Gamal-Eldin (1981)
Basque	Basque	Hualde & De Urbina (2003)
Uto-Aztecan	Cupeño	Hill (2005)
Indo-European	French	Batchelor & Chebli-Saadi (2011)
Quechuan	Quechua, Huallaga	Weber (1983, 1989)
Nilo-Saharan	Lango	Noonan (1992)
Mapudungun	Mapudungun	Smeets (2008)
Trans-New Guinea	Teiwa	Klamer (2003)
West Papuan	Tidore	Van Staden (2000)

The data were categorized according to the classification provided in Section 3.6. This classification is based on Hengeveld (1991), a questionnaire originally designed to collect data on various types of adverbial clauses. The questionnaire aims to collect and structure data through researching the general subordinating system of a language. For the present study, the questionnaire was adapted, replacing the distinction between finite and non-finite verb forms with a distinction between independent and dependent verb forms, as explained above. Working with a questionnaire allows for a systematic way to categorize, analyze and compare the data of different languages. For every language, the different types of CC types were listed. These CC types were then classified in terms of the classification given in Section 3.6. We did not count CCs with different complementizers but otherwise identical properties as different strategies, for the reason given in Section 3.5.

Based on the preceding explanations, we put forward two hypotheses. The first is a quantitative one. We hypothesize that the more transparent a language is, the fewer CC types it will exhibit. A transparent language avoids *syntactic suppletion* (Hengeveld & Mackenzie in prep.). As the languages selected for this study have already been classified in terms of their degree of transparency in Hengeveld & Leufkens (2018), we predict that, as regards their number of CC strategies, the languages from our sample will follow the ordering in terms of

transparency in Hengeveld & Leufkens (2018). This ordering is given in Table 2, which starts from the most opaque language (French) at the top, and moves to the most transparent one (Mapudungun) at the bottom.

Table 2: Sample languages ordered in terms their degrees of transparency in Hengeveld & Leufkens (2018) from least transparent to most transparent

French
Arabic, Egyptian
Basque
Lango
Cupeño
Quechua, Huallaga
Tidore
Teiwa
Mapudungun

Thus, it is predicted that for the languages in our sample, the number of CC types will be highest for French, the least transparent language in the sample of Hengeveld & Leufkens (2018), and lowest for Mapudungun, the most transparent language in Hengeveld & Leufkens (2018).

Our second hypothesis is a qualitative one. As shown above, some CC types are more similar to main clauses than others. Specifically, a CC type with an independent verb form and independent argument marking shares many properties with a main clause, while a CC type with a dependent verb form and dependent argument marking shares fewer properties with a main clause. We thus predict that transparent languages will show CC types with independent verb forms and independent argument marking, while opaque ones will allow CC types with dependent verb forms and dependent argument marking. Given the dependency of dependent argument marking on dependent verb marking (see Section 3.6), we expect dependent marking to show up in transparent and opaque languages in the following way:

(13) Independent verb   ⊂   Dependent verb ⊂   Dependent argument
       forms   forms   marking

       Transparent   Opaque

That is, the more transparent a language is, the more its CCs will be like main clauses.

5 Data

The data for the nine sample languages are presented in Table 3. This table should be read as follows: the languages are presented horizontally in alphabetical order, while the complement strategies (independent or dependent verb form) and the sub-strategies (dependent or independent argument encoding) of every language are presented vertically.

Table 3: The data (unsorted)

	Basque	*Cupeño*	*Arabic, Egyptian*	*French*	*Lango*	*Mapudungun*	*Quechua, Huallaga*	*Teiwa*	*Tidore*
Independent verb form	1	1	1	1	1	1	1	1	1
Dependent verb form	2	1	2	3	2	1	1	0	0
Independent argument encoding	2	1	1	2	2	0	1	0	0
Dependent argument encoding	0	0	1	1	0	0	0	0	0

In Table 3, the number of strategies present within a language are indicated by a numeral. Whereas a zero indicates that a certain type of CC does not exist in a language, a 1, 2 or more indicates that a language has one or multiple CC strategies belonging to that subtype. The examples on which the figures in Table 3 are based, are given in Appendix 1.

6 Results

To test our hypotheses, the data presented in Table 3 are rearranged in Table 4 along the horizontal axis in accordance with the degree of transparency of the languages as established in Hengeveld & Leufkens (2018). Vertically, the verb forms and argument/adjunct encoding of the complementation strategies are listed.

Table 4: The data (sorted)

	French	*Arabic, Egyptian*	*Basque*	*Lango*	*Cupeño*	*Quechua, Huallaga*	*Tidore*	*Teiwa*	*Mapu-dungun*
Independent verb form	1	1	1	1	1	1	1	1	1
Dependent verb form	3	2	2	2	1	1	0	0	0
Independent argument encoding	2	1	2	2	1	1	0	0	0
Dependent argument encoding	1	1	0	0	0	0	0	0	0
Total number of CC types	4	3	3	3	2	2	1	1	1

As regards our quantitative hypothesis, Table 4 shows that it is confirmed. The highest number of CC types is found in French, which has four types. The next group of languages (Egyptian Arabic, Basque, and Lango) then has three types, followed by one language (Cupeño) with two. The remaining languages (Huallaga Quechua, Tidore, Teiwa, and Mapudungun) then have just one type. So indeed, the number of CC types decreases systematically moving from less transparent to more transparent languages.

Note that we find a scale confirming our prediction, but this does not mean that we find different patterns for the nine different classes of languages in Hengeveld & Leufkens (2018). For instance, in contrast to the Hengeveld & Leufkens (2018) study, where Basque and Lango show different degrees of transparency, in complementation strategies, Basque and Lango exhibit the same number of CC types as well as sub-types; both languages have one independent verb form and two complementation strategies with dependent verb forms and independent argument encoding. Similarly, Tidore and Teiwa both have only one strategy for complementation, each with an independent verb form. In Hengeveld & Leufkens (2018), these languages show different degrees of transparency.

As regards our qualitative prediction, we find that this is confirmed as well. First of all, it is interesting to note that all languages possess CC types with an independent verb form while not all have CC types with a dependent verb form. This shows that there is an implicational hierarchy as in (14):

(14)   Independent verb forms   ⊂   Dependent verb forms

Moving to dependent argument/adjunct encoding, there is a similar pattern. Again, there is a systematic decrease from left to right, as French and Egyptian Arabic show one CC type with dependent argument encoding, and the other languages do not. Our qualitative hypothesis is thus also confirmed: not only the number of CC types, but also the nature of the CC type plays an important role. The closer the CC type resembles a main clause, the more transparent it is.

When we relate these data to the implicational hierarchy in (13), proposed in Section 4, and expand the hierarchy in (14), the following result is obtained:

(15)

Independent verb forms	⊂	Dependent verb forms	⊂	Dependent argument marking
Teiwa		Basque		French
Tidore		Lango		Arabic, Egyptian
Mapudungun		Cupeño		
		Quechua, Huallaga		

The representation in (15) should be read cumulatively: Teiwa, Tidore, and Mapudungun have independent verb forms only; Basque, Lango, Cupeño, and Huallaga Quechua have independent verb forms and dependent verb forms, but no dependent argument marking; and French and Egyptian Arabic have independent verb forms and dependent verb forms, the latter with both independent and dependent argument marking.

7   Syntactic suppletion and the transparency hierarchy

As mentioned in the introduction, this paper studies a transparency feature that was not included in Hengeveld & Leufkens' (2018) study, in which they propose a transparency hierarchy relating a whole range of transparency features to one another. We may now address the question how this new feature fits into the hierarchy they propose, though with some caution, as the number of languages studied in this paper is limited.

Table 5: Syntactic suppletion in the Transparency Hierarchy

Property \ Language	French	Arabic, Egyptian	Basque	Lango	Cupeño	Quechua, Huallaga	Tidore	Teiwa	Mapudungun
Nominal expletives	+	–	–	–	–	–	–	–	–
Grammatical agreement (clausal)	+	–	–	–	–	–	–	–	–
Grammatical gender assignment	+	+	–	–	–	–	–	–	–
Tense copying	+	–	+	–	–	–	–	–	–
Grammatical agreement (phrasal)	+	+	–	+	–	–	–	–	–
Morphologically based stem or affix alternation	+	+	+	+	+	–	–	–	–
Syntactic suppletion	+	+	+	+	+	+	–	–	–
Discontinuity	+	+	+	+	+	+	–	–	+
Morphophonologically based stem or affix alternation	+	+	+	+	+	+	+	–	–
Grammatical relations	+	+	+	+	+	+	+	+	–
Crossreference	Na	+	+	+	+	+	+	+	+
Apposition	+	+	+		+	+	+	+	+
Phonologically based stem or affix alternation	+	+	+	+	+	+	+	+	+

Table 5 shows the transparency features of the languages studied in this article as presented in Hengeveld & Leufkens (2018) and our findings concerning syntactic suppletion, which are presented in such a way that the hierarchical relations are respected. It then turns out that syntactic suppletion occupies a place in between Morphologically based stem or affix alternation and Discontinuity. As a result, the feature occupies a position between a feature that is internal to morphosyntax (Morphologically based stem or affix alternation) and one that concerns the interface between semantics and morphosyntax (Discontinuity). As argued in the introduction, the occurrence of multiple types of complement clauses can be interpreted as an opaque feature in two different ways, but we applied the purely morphosyntactic interpretation in this paper and classified the data as such. The provisional placement of this feature on the Transparency Hierarchy is consistent with this interpretation, as it groups with other morphosyntactic features higher up in the hierarchy.

8   Conclusion

The present study set out to research the transparency of complement clauses, the main hypothesis being that the more transparent a language is, the less types of CCs it will have, and the more these CCs will resemble main clauses. A representative sample of languages was selected based on previous research and it was hypothesized that the languages of this sample would exhibit the same pattern of transparency in the domain of complement clauses as that established in previous research for other grammatical domains. The results of this study support both our quantitative and qualitative hypotheses, and confirm that the languages analyzed follow the same transparency ranking as in previous research.

9   References

Aarts, B. 2018. *English Syntax and Argumentation.* London: Palgrave.

Batchelor, R.E. & M. Chebli-Saadi. 2011. *A Reference Grammar of French*. Cambridge: Cambridge University Press.

Dik, S.C. & K. Hengeveld. 1991. The hierarchical structure of the clause and the typology of perception verb complements. *Linguistics* 29(2): 231–259.

Gary, J.O. & S. Gamal-Eldin. 1981. *Cairene Egyptian Colloquial Arabic*. Amsterdam: North-Holland.

Hengeveld, K. 1991. *Questionnaire: The Internal Structure of Adverbial Clauses*. https://www.eva.mpg.de/lingua/tools-at-lingboard/questionnaire/adverbial-clauses_description.php.

Hengeveld, K. 1998. Adverbial clauses in the languages of Europe. In J. Van der Auwera & P. Dónall Ó. Baoill (eds.), *Adverbial Constructions in the Languages of Europe*. Berlin: Mouton de Gruyter, 335–419.

Hengeveld, K. & S.C. Leufkens. 2018. Transparent and non-transparent languages. *Folia Linguistica* 52(1): 139–175.

Hengeveld, K. & J.L. Mackenzie. 2008. *Functional Discourse Grammar: A Typologically-based Theory of Language Structure.* Oxford: Oxford University Press.

Hengeveld, K. & J.L. Mackenzie. in prep. Interfaces in Functional Discourse Grammar.

Hill, J.H. 2005. *A Grammar of Cupeño.* Berkeley, CA: University of California Press.

Hualde, J.I. & J.O. de Urbina. 2003. *A Grammar of Basque.* Berlin: Mouton de Gruyter.

Klamer, M. 2003. *A Grammar of Teiwa.* Berlin: Mouton de Gruyter.

Lehmann, C. 1988. Towards a typology of clause linkage. In J. Haiman & S.A. Thompson (eds.), *Clause Combining in Grammar and Discourse.* Amsterdam: John Benjamins, 181–225.

Leufkens, S.C. 2015. *Transparency in Language: A Typological Approach.* PhD dissertation, University of Amsterdam.

Noonan, M. & E. Bavin. 1981. Parataxis in Lango. *Studies in African Linguistics* 12(1): 45–69.

Noonan, M. 1992. *A Grammar of Lango*. Berlin: Mouton de Gruyter.

Noonan, M. 2007. Complementation. In T. Shopen (ed.), *Language Typology and Syntactic Description. Vol. II: Complex Constructions* (2nd ed.). Cambridge: Cambridge University Press, 52–150.

Rijkhoff, J.N.M., D. Bakker, K. Hengeveld & P. Kahrel. 1993. A method of language sampling. *Studies in Language* 17(1): 169–203.

Slobin, D.I. 1977. Language change in childhood and in history. In J. Macnamara (ed.), *Language Learning and Thought.* New York: Academic Press, 185–214.

Smeets, I. 2008. *A Grammar of Mapuche.* Berlin: Mouton de Gruyter.

Staden, M. van. 2000. *Tidore: A Linguistic Description of a Language of the North Moluccas.* PhD dissertation, Leiden University.

Weber, D.J. 1983. *Relativization and Nominalized Clauses in Huallaga (Huanuco) Quechua*. Berkeley, CA: University of California Press.

Weber, D.J. 1989. *A Grammar of Huallaga (Huallaga Quechua)*. Berkeley, CA: University of California Press.

Yule, G. 2016. *The Study of Language*. Cambridge: Cambridge University Press.

Kees Hengeveld
Amsterdam Center for Language and Communication
Spuistraat 134
1012 VB Amsterdam
p.c.hengeveld@uva.nl

Merel Luberti
Amsterdam Center for Language and Communication
Spuistraat 134
1012 VB Amsterdam
merel.luberti@student.uva.nl

Appendix: CC types in the languages of the sample

1. Arabic, Egyptian

(i) Independent verb form, independent argument marking

This is a finite complement introduced by a complementizer.

(1)   *Huwwa   ʔaal inn=u   ha-y-ruuh̄   il   kuweet.*
       he   said COMP=3.SG   FUT-3.SG-go   DEF   Kuwait
       'He said that he was going to Kuwait.'   (Gary & Gamal-Eldin 1982: 16)

(ii) Dependent verb form, independent argument marking

In cases in which the predicate is non-verbal, a non-finite form of the verb *kaan* 'be' replaces the complementizer.

(2)   *Koon=u   ʕajjaan   h̄aaga   miʃ   kuwajjisa.*
       be.NR=3.SG   sick   thing   not   good
       'That he is sick isn't a good thing.'   (Gary & Gamal-Eldin 1982: 16)

(iii) Dependent verb form, dependent argument marking

Nominalized clauses take possessive subjects.

(3)   *Ruguuʕ-u   mitʔaχχar   ʔalaʔ   giraan-u.*
       return.back.NR-3SG.POSS   late   disturbed   neighbors-3.SG.POSS
       'His returning late disturbed his neighbors'   (Gary & Gamal-Eldin 1982: 17)

2. Basque

(i) Independent verb form, independent argument marking

This is a finite complement with the complementizer in final position.

(4)   *Udaltzainek   ukatu   dute Rubioren   bizkartzain zir=ela.*
       policemen.ERG   deny   AUX   Rubio.GEN bodyguard were=COMP
       'The two policemen have denied that they were Rubio's bodyguards.'
                                              (Hualde & De Urbina 2003: 635)

(iia) Dependent verb form, independent argument marking

The subjunctive, though finite by most criteria, is a dependent verb form, as it may not be used in main clauses.

(5)   *Euskara   Euskal   Herri   osoan   ofiziala izan*
      Basque   Basque   Country   entire.LOC official be
      dad-in   nahi/espero   dute euskaldun   askok.
      AUX.SUBJ-COMP   want/hope   AUX   Basque.speaker many.ERG
      'Many Basque speakers want/hope that the Basque language be official
      in the entire Basque Country.'   (Hualde & De Urbina 2003: 640)

(iib) Dependent verb form, independent argument marking

Different non-finite verb forms that govern arguments marked as in main clauses are used.

(6)   *Bilbon   etxe   bat   erosi nahi dut.*
      Bilbao.LOC   house   one   buy   want AUX
      'I want to buy a house in Bilbao.'   (Hualde & De Urbina 2003: 694)

3. Cupeño

(i) Independent verb form, independent argument marking

Complement clauses with the realis subordinator *-ve* or the irrealis subordinator *-pi* belong to this group.

(7)   *Tew-qa'=ne   yaq-pem-yax-weni-ve-y.*
      see-PRS=1.SG.ERG   absent-3PL-CL-PST.IMPF.PL-SUB.REAL-OBJ
      I see that they are all gone.   (Hill 2014: 412)

(ii) Dependent verb form, independent argument marking

The infinitive construction belongs in this group.

(8)   Atax-m-i   mi=chixne-t   elel'i-sh.
      person-PL-O   3PLO=kill-NPN   bad-NPN
      'Killing people is bad.'   (Hill 2014: 299)

4. French

(i) Independent verb form, independent argument marking

Indicative complements belong to this group.

(9)　　*Je　leur　ai　　dit　　que　　je　l'　avais　　déjà　　écrit.*
　　　　I　　them have　told　COMP　I　it　have.IND　already　written
　　　　'I told them that I had already written it.'　　　(Batchelor & Chebli-Saadi 2011: 180)

(iia) Dependent verb form, independent argument marking

The subjunctive is a finite form limited to subordinate contexts.

(10)　　*Je　ne　　crois　　pas　que　　Cédric　vienne.*
　　　　I　　NEG　believe　NEG　COMP　Cédric　comes.SUBJ
　　　　'I don't believe Cédric will come.'　　　　　(Batchelor & Chebli-Saadi 2011: 384)

(iib) Dependent verb form, independent argument marking

The arguments of the infinitive are expressed as they would be in main clauses.

(11)　　*Elle　entrevit　　sortir　　l'　animal　de　la　　brume.*
　　　　she　　glimpsed　come.out　DEF　animal　from　DEF　mist
　　　　'She glimpsed the animal coming out of the mist.'
　　　　　　　　　　　　　　　　　　　　　　(Batchelor & Chebli-Saadi 2011: 384)

(iii) Dependent verb form, dependent argument marking

Nominalized complements take genitive arguments.

(12)　　*On attend　l'　　arrivée　　du　　　premier ministre　　dans la soirée.*
　　　　IMP expect　DEF　arrive.NR　of.DEF　prime　minister　in　　DEF　evening
　　　　'We are expecting the prime minister's arrival this evening.'
　　　　　　　　　　　　　　　　　　　　　　(Batchelor & Chebli-Saadi 2011: 393)

5. Lango

(i) Independent verb form, independent argument marking

Indicative complements belong to this group.

(13) *ɲákô   òkòbbì   dákô   nî   dyèl   òcàmò*
      girl   3.SG.say.BEN.PERF   woman   COMP   goat   3.SG.EAT.PERF
      'The girl told the woman that the goat ate it.'   (Noonan 1992: 220)

(iia) Dependent verb form, independent argument marking

The subjunctive is a finite form limited to subordinate contexts.

(14) *rwòt   òdìò   lócɔ̀   nî   òryêt*
      king   3.SG.press.PERF   man   COMP   3.SG.winnow.SUBJ
      'The king pressed the man to winnow.'   (Noonan 1992: 222)

(iib) Dependent verb form, independent argument marking

The arguments of the infinitive are expressed as they would be in main clauses.

(15) *ògwàŋ   òcàkò   òbùccò   òpíô*
      O.   3.SG-start-PERF   3SG.yell.INF   O.
      'Ogwang started to yell at Opio.'   (Noonan 1992: 225)

6. Mapudungun

(i) Independent verb form, independent argument marking

Mapuche does not display indirect speech, only direct speech in indicative complements.

(16) *"ruka-w-pu-a-y-u"*   *pi-y*   *ta-ñi*   *d̄omo*
      house-REFL-LOC-IRR-IND-1.NON.SG-DU   say-IND.3   DEF-POSS.1.SG   woman
      ka   ta-ñi   llalla   ka.
      and   DEF-POSS.1.SG   mother.in.law   also
      '"We will build a house and settle down" said my wife and my mother-in-law.'
                                                                       (Smeets 2008: 362)

At first sight, there seem to be many nominalized complements as well, as illustrated in (17):

(17)   *ta-mi*   *wülel-ün*   *ayü-la-fi-n*
       the-POSS.2.SG   beat-NR   love-NEG-OBJ-IND.1.SG
       'I don't like you to beat.'   (Smeets 2008: 193)

However, Mapudungun has subject and object markers on the verb that are referential by themselves, such that in (17) the actual meaning is: 'I don't like it, your beating.'. This means that the nominalization is actually not an argument, but an apposition. The real argument is the object marker (*fi* in 17).

7. Quechua, Huallaga

(i) Independent verb form, independent argument marking

In Quechua, there are no indirect speech complements, only direct ones.

(18)   *Chawra   buurru   ni-n-shi: "Watqa-rku-shaq-chu?"*
       then   donkey   say-3-IND   spy-up-1.FUT-Y/N
       'Then the donkey said: "Shall I peek?".'   (Weber 1989: 125)

(ii) Dependent verb form, independent argument marking

Huallaga Quechua furthermore has several non-finite verb forms, all of which govern arguments realized in the same way as in main clauses.

(19)   *noqa   aywa-naa   risiisan*
       I   go-SUB   it.is.certain
       'It is certain that I will go.'   (Weber 1983: 81)

8. Teiwa

(i) Independent verb form, independent argument marking

This is the only type of complement clause in Teiwa.

(20)   *Na   wan   kruanyi   na'   aria-n   maan.*
       1.SG   know   2.P   possibly   arrive-REAL   NEG
       'I know you will not come.'   (Klamer 2003: 373)

9. Tidore

(i) Independent verb form, independent argument marking

This is the only type of complement clause in Tidore.

(21)   *nene=ge*   *nane*   *waje*   *mina*   *hoda*   *fayaa*   *jang*
       grandmother=there   dream   COMP   3.SG.F   see   woman   beautiful
       'Grandmother dreamed that she saw a beautiful woman.'   (van Staden 2000: 290)