Genesis: Data collection between 2008 and 2012 (in the context of a larger project on subordination systems) – Draft of paper and acceptance (for a special issue of *Linguistics*) in 2013 – Accepted for regular submission in *Linguistics* in 2015 – publication in 2017

Cross-linguistic patterns in the structure, function and position of (object) complement clauses

Karsten Schmidtke-Bode and Holger Diessel

Abstract: The present contribution examines object complement clauses from the perspective of constituent-order typology. In particular, it provides the first principled empirical investigation of the position of object clauses relative to the matrix verb. Based on a stratified sample of 100 languages, we establish that there is an overall cross-linguistic preference for postverbal complements, due largely to the heterogeneous ordering patterns in OV-languages. Importantly, however, we also show that the position of complement clauses correlates with aspects of their structural organisation: Preverbal complement clauses are significantly more likely to be coded by morphosyntactically 'downgraded' structures than postverbal complements. Given that previous research has found a parallel correlation between structural downgrading and the semantics of the complement-taking predicate (Givón 1980, Cristofaro 2003), one needs to analyse how positional, structural and semantic factors interact with one another. Our data suggest that the correlation between clause order and morphosyntactic structure holds independently of semantic considerations: All predicate classes distinguished in the present study increase their likelihood of taking downgraded complements if they are preceded by the complement clause. We thus propose that, in addition to the well-known 'binding hierarchy', a second correlation needs to be recognised in the typology of complementation: the co-variation of linear order and morphosyntactic structure.

Key words: Complementation, complement clause, linearisation, word order typology, binding hierarchy.

1 Introduction

Cross-linguistic research on complement clauses has been thriving since the 1970s, and has produced a number of important insights into the grammatical organisation of languages (cf. Horie 2001 for an overview). Perhaps the most intensively studied aspect in the typology of complementation is the relationship between the form and meaning of complement clauses. Ransom's (1986) monograph, for example, is devoted to the various 'modalities', i.e. attitudes towards propositional content, that complement clauses can instantiate, and shows how each of them is realised by typical formal means, such as specific complementizers, modal markers, word-order restrictions, constraints on argument sharing, etc. In a similar vein, Givón (1980) established a systematic cross-linguistic correlation between certain semantic aspects of complement-taking predicates and the degree of syntactic 'downgrading' of the complement clause. Specifically, he proposed that "the stronger the influence exerted over the agent of the complement

clause by the agent of the main-clause verb", the less will the complement clause "tend to be syntactically coded as an independent/main clause" (Givón 1980: 335–337). This correlation, dubbed the "binding hierarchy" of complementation, has found ample support ever since (e.g. Foley and VanValin 1984, Cristofaro 2003), and Givón himself considers it "one of the best, and cross-linguistically most reliable, examples of iconicity in syntax" (Givón 2001: 39).

Our aim in the present paper is to investigate how these well-established form-function relationships in object complementation interact with a syntactic variable that has received less attention in the typological literature: the position of complementation constructions relative to the matrix verb. Clause and constituent order have been important topics in linguistic typology, but in contrast to other types of subordinate clauses, notably relative clauses (e.g. Lehmann 1984, Andrews 2007, Dryer 2011a) and some kinds of adverbial clauses (e.g. Diessel 2001, Schmidtke-Bode 2009, Diessel and Hetterle 2011), the positional patterns of complement clauses have not yet been examined in comparable detail. All previous studies are limited in their choice of languages, analytical parameters and overall scope. Dryer (1980), for example, provides qualitative evidence from about 20 languages for certain positional tendencies of complement clauses. It is broader and more powerful in its generalisations than an earlier study by Grosu and Thompson (1977), yet it remains limited to a non-sampled set of languages and qualitative observations. Similar remarks apply to a more recent study by Ogihara (2009). Since her analysis of complement clauses is situated in a larger investigation of "verb-final typology", it is exclusively concerned with complement clauses in OV-languages. While this is arguably the more interesting type of language in this context (as we shall see below), her discussion of complements in OV-languages is still not exhaustive in many respects.

The current study is thus the first to examine the position of complement clauses in a larger and more balanced typological sample. It shows, first, that there is an overall cross-linguistic preference for postverbal complements and, second, that the position of complement clauses correlates in principled ways with the morphosyntactic structure of the subordinate clause: Complement clauses that are placed before the matrix verb are significantly less likely to be coded by main-clause-like structures than complements that follow their matrix verb. This applies to several aspects of morphosyntactic organisation and constitutes a robust statistical trend within and across languages. However, given that, according to Givón and many others, the morphosyntax of the complement also co-varies in similar ways with the semantics of the matrix verb, one may ask how precisely the three parameters of position, structure and meaning interact with one another. Our study demonstrates that the structure of complement clauses correlates with both the semantics of the matrix predicate and the position of the complement, but that there is no such correlation between clause order and meaning: Each of the commonly distinguished predicate classes occurs with pre- and postverbal complements, and for all of them, the likelihood of structural downgrading of the complement increases in preverbal position. This finding thus suggests that the morphosyntax of complementation is not only determined by semantic factors, but also by considerations of linear order.

The structure of the paper is as follows: Section 2 defines complement clauses as a "comparative concept" (Haspelmath 2010) and addresses methodological questions relating to our language sample and the collection and analysis of the data. Section 3 establishes the basic typological patterns in the positioning of complement clauses, while Section 4 goes on to elaborate the correlation between these positional patterns and the morphosyntactic structure of the complement. In Section 5, the interplay of structure, function and position will be investigated, followed by a brief conclusion in Section 6.

2 Conceptual and methodological preliminaries

2.1 Complementation constructions

In the typological literature, complementation is commonly defined as a grammatical phenomenon that arises when "a predication is an argument of a predicate" (Noonan 2007: 52). Such a notional or semantic definition ensures cross-linguistic comparability (since this situation is arguably found in all languages) and does not impose any a priori restrictions on the form of complements. In fact, the constructions that are subsumed by a purely semantic definition can, in principle, range from deverbal nouns as in (1) to fully developed clauses as in (2):

(1) English

[*His prompt refusal (of the offer)*] *was surprising.*

(2) Jamsay (Niger-Congo/Dogon: Mali)

[*Á*	*ùrò-bɔ̀rɔ́*	*dìŋ-â:-ø*]	*jùgɔ́-jɛ̀-w.*
2SG.POSS	house.L-rear	sit.down-PFV-3SG.SBJ	know-RECPF-2SG.SBJ

'You have known that your house foundation has been settled.'
(Heath 2008: 602)

Moreover, Cristofaro (2003: 95–98) notes that a semantic definition is sufficiently open to accommodate different degrees of syntactic integration of the complement and the main verb: While many complement clauses are syntactic arguments of the matrix predicate and hence embedded as constituents of the main clause, others are 'adjoined' to a syntactically complete main clause. Consider, for instance, example (3) from To'aba'ita, in which the matrix verb *ade* 'do' has a causative interpretation and takes a clause as its notional complement. Syntactically, however, the clause in brackets is adjoined to a structurally saturated matrix clause that could stand alone as an independent sentence, since *ade* takes the NP *wane* 'man' as its direct object:

(3) To'aba'ita (Austronesian/Eastern Malayo-Polynesian: Solomon Islands)

Meresina	*qeri*	*qe*	*ade-a*	*wane*	[*ka*	*qiiqida*	*qasia*	*naqa*].
medicine	that	3SG.NFUT	do-3OBJ	man	3SG.SEQ	sweat	INTS	INTS

'The medicine made the man sweat a lot.' (Lit.: 'The medicine did the man, he sweated a lot.')
(Lichtenberk 2008: 984)

As Lichtenberk (2008: 984–985) comments, "it is the causee phrase and not the complement clause that is the object of the causative verb"; therefore, the clause in brackets is not embedded as a syntactic argument of the matrix predicate.

In view of this formal variation, Dixon (1995, 2006) has proposed to distinguish between a core domain, which he calls '**complement clauses**' proper, and functionally equivalent devices called '**complementation strategies**'. The most important defining features of a complement clause in this narrower sense are (i) that it has (most of) the internal structure of a clause, and (ii) that it functions as a genuine syntactic argument of a predicate. Example (2) above fulfils both criteria: It is a fully clause-like construction that functions as the direct object of the transitive verb 'watch'. Examples (1) and (3), by contrast, are complementation strategies. The nominalisation in (1) has all the vestiges of an NP rather than of a clause (determiner, genitive subjects and objects, internal modification by an adjective), and the adjoined clause in (3) is not embedded as an argument in the main clause.

In accordance with Dixon, we also consider it useful to distinguish between complement clauses proper and complementation strategies. In practice, however, the line can be very hard to draw, chiefly due to the elusive nature of many of the parameters involved in the distinction (e.g. *degrees* of nominalisation and syntactic argumenthood, cf. Schmidtke-Bode 2014: Ch.2 for discussion). For the purposes of the present paper, a nominalisation will be considered a complementation strategy if its internal object is coded differently from that of an independent clause and/or the modification of the complement predicate is adjectival rather than adverbal, as in (1) above (cf. Dixon 1995 for the same criteria[1]). And the so-called adjoined clauses from (3) above are considered complement strategies if this is their syntactic status with all predicates they complement. In other words, if there is no evidence that a complementation pattern in question can function as a syntactic argument of any of the predicate classes relevant to the present study (cf. Section 5 below), it is considered a complementation strategy.[2]

The focus of our study is on complement clauses in the narrow sense. Complementation strategies are only taken into account if no genuine complement clauses are available. For example, in Yimas (Lower Sepik-Ramu: Papua New Guinea), complementation is rendered exclusively by strongly nominalised constructions that are

[1] Note that the form of the internal subject is not criterial: Nominalised complement clauses may have all internal characteristics of an independent clause except for the omission or the possessive coding of their subject (e.g. English *John's playing the national anthem competently*, as discussed in Dixon [2006: 16]). We follow Dixon in considering such structures complement clauses proper rather than complementation strategies.

[2] Ultimately, this distinction would have to be made for every single predicate that a given complementation pattern co-occurs with. For example, a fully sentential type of complement could be embedded as a syntactic argument of some predicates, but adjoined as a non-argument to others (e.g. verbs of speech, cf. Munro 1982, Güldemann 2008). For the present paper, however, this predicate-specific level of coding was infeasible. Instead, then, every complementation pattern that can function as a genuine syntactic argument of at least a *subset* of predicates it complements will be considered a complement clause in the narrow sense. This practice is also adopted in Dixon's (2006) framework (in conjunction with other criteria, of course), i.e. he applies the distinction to complementation structures as a whole rather than to their individual uses with specific matrix predicates.

"clearly noun phrases" (Foley 1991: 394). In the absence of proper complement clauses, those constructions were taken into account. This scenario, in which an entire language lacks 'core complements' in Dixon's sense, is rather rare in our data. More commonly, however, we find that complementation strategies fill certain functional roles in complementation that complement clauses proper 'leave open'. As observed by Givón (1980), many languages have a productive complement clause for a wide array of matrix predicates, but resort to strongly nominalised constructions in certain semantic domains (e.g. phasal or same-subject desiderative predicates). Similarly, the adjoined construction in To'aba'ita (cf. [3] above) is the standard complementation pattern of perception, jussive and causative predicates. In such cases, complementation strategies were taken on board.

2.2 Sampling and collection of the data

The data for the present study are drawn from a world-wide sample of 100 languages, which was compiled for a larger project on subordination systems in the world's languages. Representative sampling in the domain of complex sentences (and clause order, in particular) is compromised by a severe "bibliographic bias" (Bakker 2011: 106) in language documentation: About 40% of the stocks selected by formalised sampling algorithms (e.g. the Diversity Value algorithm, kindly provided by Dik Bakker) needed to be discarded due to insufficient information on the complex array of variables of interest in our work. This is why an initially envisaged sample of 150 languages had to be reduced and modified in accordance with the available information, until we finally arrived at a sample of 100 languages. This final selection represents what Bakker (2011: 121) calls a "pseudo-probability sample": Like a variety sample, it is based on family-internal genetic diversity calculations (cf. Rijkhoff et al. 1993) and is large enough in scope to capture the typological variability in clause-combining systems; at the same time, it is sufficiently controlled to statistically test for universal coding trends in this domain, and at least in this respect resembles a probability sample.

Note that the inherent difficulties of the sampling procedure are counterbalanced to some extent by choosing appropriate statistical tests: Following a suggestion by Janssen et al. (2006), we will apply non-parametric, distribution-free tests that do not, strictly speaking, draw a statistical inference from the sample to the underlying population. This inference, according to Janssen et al., is rather a logical one: If the sample is sufficiently representative of the world's languages due to principled sampling procedures (cf. above), a statistically significant signal in the analysis of the sample is likely to reflect a generalisable typological trend. A list of the sample languages can be found in the Appendix.

For each of the languages in our sample, we collected the relevant information from a wide variety of sources, primarily grammatical descriptions and language-specific articles. In addition, we are extremely grateful to many experts of individual languages, who kindly acted as consultants on particularly challenging questions concerning aspects of complementation in their respective language (cf. also Appendix). The analytical procedure was then as follows: We analysed the complementation system of each language and extracted the major complementation patterns recognised in our sources, again with primary attention to complement clauses proper. For all languages

in the sample, we neglected complementation patterns that were described as minor constructions (e.g. in terms of type frequency), unless they complement predicate classes that are not covered by the major constructions. Each construction in the database was then classified as either a complement clause or a complementation strategy, according to the criteria discussed above. This yielded a construction-specific database of 205 complementation patterns in total (i.e. roughly 2 data points per language on average), comprising 172 genuine complement clauses and 33 complementation strategies. The general practice for the rest of the paper is such that each analysis will be performed on the entire data set, and, where appropriate, a second analysis will be run on the more restrictive 'control sample' of complement clauses proper. In this way, typological generalisations can be tested for complement clauses in a narrower and in a wider sense.

3 Positional patterns of complement clauses

As was stated in the introduction to this paper, the positional tendencies of complement clauses have been the concern of mostly qualitative typological studies with a limited scope of languages. Two such studies, i.e. Grosu and Thompson (1977) and Dryer (1980), argue that complement clauses tend to avoid centre-embedding by being placed at the sentence margins (e.g. Dryer's [1980: 125] "final-over-internal" and "initial-over-internal" hypotheses). In addition, they also propose a "final-over-initial" hypothesis, according to which the most common position of complement clauses is to occur postverbally. However, especially the data on this latter constraint conflate subject and object complement clauses, so that we still need to establish the empirical picture for object clauses as such. In the current study, we will leave specific questions of centre-embedding and extraposition aside and concentrate on the positioning types that are most relevant from the perspective of constituent-order typology and the specific goal of our study.

In this context, the major parameter of interest is the position of each construction in the database vis-à-vis its associated head element, i.e. the matrix verb. Examples (4) and (5) below illustrate a typical representative of pre- and postverbal complement clauses, respectively.

(4) preverbal complement clause from Awa Pit (Barbacoan: Ecuador, Colombia)

Gregorio=na [*ashaŋpa=na* *az-tu*] *kizh-t̶i̶-zi.*
Gregorio=TOP woman=TOP cry-IPFV.PTCP say-PST-NONLOCUT
'Gregorio said that his wife was crying.'
(Curnow 1997: 260)

(5) postverbal complement from (Chalcatongo) Mixtec (Oto-Manguean: Mexico)

X^{w}ã *nì-kunì=Ø* [*xa=nà-šuk^{w}ĩ́=rí́*].
Juan COMPL-want=3 COMP=REP-turn=1
'Juan wanted me to go back home.'
(Macaulay 1996: 153)

The complement clauses in (4) and (5) are conventionalised in their respective position, i.e. it is not normally possible to simply swap the relative order of complement clause and matrix verb. In languages like English, such preposing is possible (*That he really intended to cheat us*] *I still can't believe*. [cf. Huddleston and Pullum 2002: 952]), but even there it constitutes a highly marked option, so that the postverbal position of finite complements is still a very strong tendency. For our purposes, the situations in Awa Pit, Chalcatongo Mixtec and English will be grouped into complements with a strong positional preference. For lack of a better term, we will refer to such constructions as 'rigidly preverbal' and 'rigidly postverbal' complements, respectively, bearing in mind that an alternative order may not always be categorically excluded.

In other languages, there is still a discernible preference for a particular order, but the complement can be found in alternative positions more freely or commonly. This holds, for example, for finite complements in Hungarian, which "are most naturally placed finally in main clauses" but where "preverbal positions are also possible and frequent" (Kenesei et al. 1998: 29). Similarly, participial complement clauses in Tümpisa Shoshone "more commonly occur before the superordinate verb, [but] they may also follow it" (Dayley 1989: 381) without being a mere afterthought. Complements of this type will be referred to as 'non-rigidly postverbal' and 'non-rigidly preverbal', respectively. Where both orders of complement clause and matrix verb are attested and no significant positional tendency could be identified by the authors of the sources or by us, we will speak of a 'flexible' type (e.g. 'dependent embedded clauses' in Motuna [Onishi 1994: 461]). Needless to say, the data available at present make it notoriously hard to compare such performance patterns across a wide range of languages; our assessments of 'rigidity' and 'flexibility' must, therefore, be seen as preliminary and awaiting further research. For the present purposes, the positional patterns of complement clauses can be summarised as follows:

- **Rigidly preverbal**: The complement clause normally precedes the matrix verb.[3]
- **Non-rigidly preverbal**: The complement clause typically precedes the matrix verb but is not uncommon in postverbal position.
- **Rigidly postverbal**: The complement clause normally follows the matrix verb.
- **Non-rigidly postverbal**: The complement clause typically follows the matrix verb but is not uncommon in preverbal position.
- **Flexible**: The complement clause is common in both pre- and postverbal position, without a discernible marked preference.

These positional patterns can now be compared to that of NP complements. Since we are exclusively dealing with complements in object function here, it is useful to take the

[3] For complementation patterns that are left-adjoined rather than embedded (and are hence complementation strategies), the correct formulation here would be that they precede the matrix *clause* rather than just the matrix verb. This proviso holds for all of the categories in the current list (i.e. right-adjoined complements are subsumed under postverbal here, but since these are not syntactic arguments of the matrix verb, their attachment site is actually the entire matrix clause.) For comparative purposes, however, what counts is that these complements are distributionally still 'pre' or 'post' relative to the matrix verb, regardless of their syntactic status.

order of phrasal objects relative to the verb as a reference point against which the position of clausal objects is investigated. Along this dimension, we shall draw a three-way distinction between OV-languages, VO-languages, and languages without a dominant order of verb and object (cf. also Dryer 2011b). In our database, the latter group comprises languages in which the order of object and verb follows discourse-pragmatic rather than grammatical principles and no general preference is discernible (e.g. Yuracaré, Kayardild), as well as languages in which both VO and OV patterns exist due to grammatical principles (e.g. German, Ma'di, Trumai); in the following, we shall collectively abbreviate this third group as OV/VO-languages. Table 1 outlines how the different positional types of complement clauses pattern with regard to the OV/VO distinction. Note that, in this and all following analyses, syntactically adjoined constructions (cf. example [3] above) are subsumed under pre- or postverbal complements, depending on whether they are left- or right-adjoined to the main clause.

Table 1. Positioning types of the complement clauses in OV- and VO-languages

	Pre:rigid	Pre:nonrig	Flexible	Post:nonrigid	Post:rig	Totals
OV	61	13	12	11	12	109
VO	0	1	0	3	77	81
OV/VO	1	0	3	4	7	15
Totals	62	14	15	18	96	205

There are several observations to be gleaned from Table 1. To begin with, the postverbal position is the cross-linguistically 'dominant' type in Greenbergian terminology (cf. Greenberg 1966), accounting for about 56% of all constructions. Preverbal and flexible complements are both 'recessive' types by comparison, accounting for roughly 37% and 7%, respectively.[4] This overall distribution remains constant if all complementation strategies are removed from the analysis: Complement clauses proper are dominant in postverbal position (55%), less common in preverbal position (38%) and least common with flexible ordering (7%).

With regard to constituent-order correlations, Table 1 shows that VO-languages clearly conform to the typological null hypothesis that complement clauses occupy the same postverbal position as nominal objects. The only aberrant construction here comes from Mapudungun, where the VO trend for nominal objects is rather weak to begin with and a quotative complementation strategy preferably precedes the utterance predicate (Smeets 2008: 361). If this is taken out, we have a uniform distribution. By contrast, the behaviour of OV-languages is more heterogeneous, which aligns the positional patterns of complement clauses with that found for relative and adverbial clauses in OV-languages (cf. Dryer 2011a, Diessel 2001). As can be seen in Table 1, 74 of the 109 (= 67.9%) constructions in OV-languages are preverbal and hence occur in the same relative position as nominal objects. In fact, the dominant construction type overall is a rigidly (i.e. strongly) preverbal complement (61/109 = 56%). However, there is also a notable amount of postposing in OV-languages. For one thing, 13 of the 74

[4] It should be borne in mind at this point that we concentrated on the *major* complementation patterns in each language; if additional complementation strategies had all been taken into account, the number of postverbal complements would have been likely to increase further.

preverbal constructions (= 17.6%) can be right-shifted, i.e. they are non-rigid constructions. And there is, of course, also a sizeable number of complements that are postverbal to begin with: Table 1 shows that 23 of the 109 constructions in OV-languages (= 21.1%) are usually positioned after the matrix verb. It is precisely these constructions that lead to the overall cross-linguistic preference for postverbal complements noted above. Interestingly, this skewing is also observable in languages that have both OV and VO patterns, i.e. in the third row of Table 1. As can be seen, the postverbal type is clearly dominant here (11/15 = 73.3%).

Apart from the general distributions, a further aspect of Table 1 is noteworthy. If we compare the proportion of 'non-rigid' as opposed to 'rigid' orderings in OV- and VO-languages, it appears that OV-languages are more 'liberal' here, in at least two respects: First, they immediately contrast with VO-languages on the postverbal constructions: Postverbal complements in OV-languages have a significantly higher likelihood of being 'non-rigid' than in VO-languages (11/23 versus 3/80 cases, Fisher exact test, $p < 0.0001$). Second, 'flexible' complements, i.e. those without a marked preference for a particular order, occur only in OV-languages (or languages with a significant portion of OV patterns), and never in a VO-language, at least in our data (cf. last column of the table). Taken together, the positional rigidity of complement clauses appears to interact in interesting ways with the order of O and V, and each of them underlines the overall dominance of postverbal complementation patterns.

Thus far, we have considered the positional preferences of complement clauses at the level of individual constructions. To wrap up our discussion of clause order, let us finally examine the positional patterns at the level of languages. The 100 languages in our sample comprise 40 VO-languages, 53 OV-languages and 7 languages with mixed VO/OV-patterns. As was stated above, VO-languages are homogeneous as far as the position of their complements is concerned. By contrast, the OV- and partially OV-languages in our sample (N = 53 + 7 = 60) fall into several different types. The largest group, comprising 31 out of the 60 languages (= 51.7%), has exclusively preverbal complements. In this group, we find languages like Amele, Barasano, Jarawara, Mekens, Urarina and Wolaytta. In the traditional terms of word-order typology, such languages are thus consistently 'left-branching' in their primary complementation patterns. The second, much smaller, group of languages has postverbal complements only, either of the embedded or the right-adjoined type (14/60 = 23.3%). Languages in this category include, for example, Epena Pedee, German, Gooniyandi, Skou and Somali. Finally, a third group of languages (15/60 = 25%) has either flexible complements only (Kwazá, Yuracaré) or a mixture of preverbal complements with flexible or postverbal constructions (e.g. Awa Pit, Santali, Persian, Tümpisa Shoshone).[5]

Overall, then, the distribution of complements in OV-languages leads to a situation in which slightly more than 50% of *all* languages investigated in this study have exclusively postverbal object complements, despite the fact that only 40% of the sample languages are VO and would thus be expected to occur with the postverbal type. The

[5] Note that we also examined the geographical distribution of the three OV-types, which turned out to be inconspicuous: All positioning types can in principle be found wherever OV-languages are attested, and there is no straightforward areal bias of any positioning type in our data.

deviations from expected positional patterns (i.e. OV|pre and VO|post) in our sample achieve statistical significance in a binomial test (p = 0.001). Our results thus confirm, on a more robust empirical basis, earlier suggestions (e.g. Dryer 1980) that complement clauses have a greater tendency than nominal objects to occur in sentence-final position.

4 The co-variation of positional and structural choices

Having established the positional patterns of complement clauses, we are now going to take a closer look at their morphosyntactic structure. Although this has never been examined systematically, typological research has suggested that positional patterns often go hand in hand with certain structural choices: In the domain of relative clauses, for example, Lehmann (1984) finds that prenominal constructions tend to be coded by participles and similarly reduced structures, while postnominal relatives are more likely to resemble full-fledged 'finite' clauses. In this part of the paper, we show that there is a similar morphosyntactic asymmetry in the domain of complementation. This asymmetry is observable along several dimensions of the internal structure of complement clauses. Most prominently, however, it affects the core element of the complement, i.e. the subordinate verb. In the following, therefore, two aspects of the verb will be examined in more detail: Section 4.1 is devoted to the derivational status of the verb, while Section 4.2 is concerned with its inflectional properties. In Section 4.3, we report selected findings on additional aspects of grammatical organisation.

4.1 The derivational status of the complement verb

Given that prototypical complement clauses function as constituents of main clauses, it is not surprising that their dependent status and their inherently nominal function is often overtly marked by employing verb forms that are different in status from those in independent clauses. Such dependent verb forms have sometimes been referred to as 'deranked' structures (Stassen 1985, Cristofaro 2003), and they can also be found in our sample. More specifically, 118 of the 205 complementation patterns (58%) contain the same basic form of the verb as independent clauses, while the remaining 87 constructions employ dependent verb forms. As would perhaps be expected, there is considerable cross-linguistic diversity in the nature of such dependent verb forms. The majority (71/87 = 82%) is constituted by derivational morphology of various kinds. This includes, first and foremost, nominalising morphemes that create a deverbal-noun *head*, while typically leaving at least some other elements of the complement verbal in nature (cf. our above discussion of complement clauses versus complementation strategies again). A nominalised complement clause is illustrated for Matses in (6):

(6) Matses (Panoan: Brazil, Peru)

[*Kuesban-ø kues-te*] *bun-e-bi.*
bat-ABS kill-NMLZ want-NPST-1S
'I want to shoot at/the bats.'
(Fleck 2006: 232)

Note that nominalisations also subsume 'infinitival' morphemes: It is well known that such infinitives are often nothing but erstwhile purposive action nominalisations that have come to be reanalysed as complement clauses and are synchronically more widespread as complements than as purpose clauses (Haspelmath 1989). As far as we can discern, most of the so-called 'infinitives' in our data follow this pattern.

Two less common derived verb forms in complement clauses are participles and converbs.[6] They are similar from a diachronic point of view since both participles and converbs originate as forms that are specifically associated with other types of subordinate clauses (i.e. relative and adverbial clauses, respectively), but come to be reanalysed as complement clauses of certain matrix verbs (cf. also Ylikoski 2003). Participial complement clauses, for instance, are typically found with perception verbs first and may then spread to selected other predicate classes (cf. Noonan 2007: 72–73 on participial complements). In contrast to infinitives, their primary synchronic function is still in the source domain, i.e. relative or adverbial clauses. Example (7) below illustrates a participial complementation strategy from Martuthunira, while (8) shows a converbal complement clause from Evenki:[7]

(7) Martuthunira (Pama-Nyungan: Australia)

Nhuwana nhuura nganaju [*yilangu karri-nyila-a purnumpuru*].
2PL knowing 1SG.ACC here stand-PR.REL-ACC quiet
'You know that I'm standing here quietly.'
(Dench 1995: 255)

(8) Evenki (Altaic/Tungusic: Russia)

Asatkan omngo-ro-n [*dukuvun-ma ung-de:-vi*].
girl forget-NFUT-3SG letter-ACC send-CVB-POSS.REFL
'The girl forgot to send the letter.'
(Nedjalkov 1997: 28)

Apart from such derivations, dependent verb forms comprise a wide range of other phonological and morphological processes that are not associated with word-class transpositions or specific functional subtypes of subordinate clauses. We will collectively refer to such cases as 'other dependent forms', as they can take on a variety of different

[6] Just like nominalisations, participles and converbs are "derived verb forms" in Haspelmath's (1995: 3) sense of creating 'deverbal adjectives' and 'deverbal adverbs', respectively. However, in contrast to (strong) nominalisations, these are derivational only to the extent that they change the word-class assignment of the head of the subordinate clause, i.e. the verb, while the other elements of the clause may remain completely or dominantly verbal in nature (cf. also footnote 7 below).

[7] As was pointed out in footnote 6 above, the internal structure of participles and converbs is typically more verbal than that of (strong) nominalisations, i.e. their internal modifiers and object arguments are often coded as in an independent clause. Therefore, when participles and converbs are classified as complementation strategies here, this is usually because of their syntactic status (e.g. right-adjoined rather than embedded in example [7]) and not because of their internal structure.

shapes. In Wappo, for example, verbs of all subordinate clauses lack a final glottal stop that is characteristic of the verbs of independent clauses (Thompson et al. 2006: 109). In complement clauses in Jarawara, the verb-final vowel /a/ is replaced by /i/ (Dixon 2004: 91). In yet other languages, dependent verb forms are deprived of an otherwise obligatory sentential-mood marker (e.g. the declarative marker in Santali, which changes the status of the verb to a subordinate form but still leaves room for tense-aspect-mood inflections to vary independently, cf. Neukom 2001: 181). And they may even be reduced to a bare verb stem without any inflectional or derivational marking (e.g. Barasano, cf. Jones and Jones 1991: 137). These examples will suffice to give a flavour of the morphosyntactic variation in the verbs of complement clauses.

What is now most interesting to observe is that the occurrence of dependent verb forms co-varies with the position of the complement clause in a non-accidental fashion. In Table 2, we cross-classify the basic distinction between 'independent' and 'dependent' verb forms with the three major positioning types of complement clauses[8]:

Table 2. Verb form and the basic position of complement clauses

	Postverbal	Preverbal	Flexible	Totals
Independent verb form	78	35	5	118
Dependent verb form	36	41	10	87
Totals	**114**	**76**	**15**	**205**

Table 2 shows that the proportion of dependent verb forms is very different in the three positioning types. Most importantly, it is significantly different for pre- and postverbal complement clauses (53.9% versus 31.6%, Fisher exact test, $p = 0.002$). It thus appears that pre- and postverbal complements tend to strive into opposite directions as far as the derivational status of their verb is concerned. Although it is not the case that dependent verb forms are overwhelmingly preferred in preverbal position, the conclusion is still that preverbal complements have a significantly higher likelihood of appearing in such special forms. Flexibly ordered complements seem to pattern with (and even outperform) the preverbal clauses in this regard, as 66.7% of the flexible constructions include dependent verb forms. Apparently, then, flexibility of constituent ordering is associated with the loss of the morphological independence of the complement clause. Vice versa, one may say that complements with independent verb forms cannot be positioned as freely as nominalised and other dependent constructions.

The reader may wonder if the proportions change if the data are reduced to complement clauses 'proper', i.e. if all complementation strategies are removed from Table 2. We thus conducted a separate analysis, which yielded even slightly stronger results: The amounts of dependent verb forms change to 27% (postverbal) and 57% (preverbal), respectively. This is most likely due to the removal of certain left-branching direct-speech complements, but note that some of the nominalisations (i.e. the more strongly nominal ones) are also taken out. Therefore, our results do not seem to be

[8] The analyses in the remainder of this paper relate to variables with increasing complexity as far as their number of levels is concerned (from two-way to nine-way categorical contrasts later on). For this reason, we decided to reduce the positional patterns of complements to the three basic types ('preverbal', 'postverbal' and 'flexible') and to disregard the rigidity of the ordering pattern.

dependent on the inclusion or exclusion of less typical instances of complementation constructions, and instead reveal a fairly robust asymmetry in the morphosyntax of pre- and postverbal clauses.

This can be observed directly in languages where structurally different types of complementation patterns obey different rules of linearisation. In the Nilo-Saharan language Ma'di, for instance, a certain class of verbs (the so-called 'inflected verbs') takes object NPs in preverbal (SOV) position. When these verbs license complement clauses with independent verb forms, the construction is always right-adjoined to the matrix clause; it cannot occupy the preverbal object position directly but is represented there by a pronominal object clitic (cf. [9a] below). Crucially, Ma'di also employs nominalisation as a complementation strategy (9b), and "unlike any other kind of complement clause", such nominalised complements "can also precede an inflected verb" like "a true object" (Blackings and Fabb 2003: 208):

(9) Ma'di (Nilo-Saharan/Central Sudanic: Uganda, Sudan)

a. right-adjoined "full clausal complement"

Nɨ́ ˋfō ʔā [k-ē-mú ɔ̀bɔ́].
2SG N-say OBJ 3DIR-VE-go tomorrow
'You will tell her to come tomorrow.'
(Blackings and Fabb 2003: 410)

b. centre-embedded "suffixed clausal complement"

Àmá [bás̀ɨ ʔà ŋgā-kā] ˋ nɨ̄ tè áō.
1PL.EXCL bus POSS (N)-depart-NMLZ SPEC PR (N)-wait FOC
'It is the bus' departure that we await.'
(Blackings and Fabb 2003: 410)

Similar patterns can be found, for instance, in Awa Pit, Basque, Khwe, Motuna, Persian, Santali and Tümpisa Shoshone. In other languages, complement clauses generally precede the matrix verb, but the constructions in question either all employ dependent verb forms (e.g. Lavukaleve, Sanuma, Urarina, Wappo, Warao, Wolaytta), or there are more constructions with dependent than independent verb forms (e.g. Barasano, Lezgian, Malayalam, Matses, Turkish). All of these language-specific patterns thus contribute to the overall skewing of dependent verb forms towards preverbal complement clauses.

4.2 Inflectional properties of the complement verb

A further dimension of morphosyntactic dependency relates to the inflectional properties of the verb in the complement clause. While these properties tend to correlate with the derivational status of the verb, they are logically independent parameters of cross-linguistic variation. For example, some of the nominalised constructions in our database can retain a considerable amount of TAM marking (e.g. in Krongo, Malayalam, Mapudungun, Musqueam) and/or preserve the canonical subject-

agreement pattern of independent clauses (e.g. in Kwazá, Lavukaleve, Tariana). For this reason, the present section briefly considers how these two inflectional categories interact with the position of the complement clause.

With regard to TAM, we cannot possibly outline the diverse individual patterns of temporal and modal marking that occur in the data. What counts primarily in the present context is whether the expression of TAM categories largely follows that of independent clauses or is changed in a conspicuous way. In the following analysis, we thus distinguish three types of scenario:

- **Non-reduced**: TAM marking is the same as, or only minimally different from, that in independent clauses. (Note that, as for 'mood', our focus is with distinctions like realis versus irrealis, potential, conditional etc., and not with 'sentence mood' [i.e. illocutionary-force] distinctions like declarative, interrogative, assertive, etc.)
- **Reduced**: TAM marking needs to be suppressed entirely, fixed to a specific TAM value or changed into a subordinate TAM paradigm (e.g. different kinds of tense contrasts in Quechuan nominalised clauses [Muysken 1994: 2814] or the relative-tense distinctions expressed by different kinds of participial markers [e.g. Evenki, Tümpisa Shoshone]).
- **Neutral**: The language in question does not realise TAM expressions inflectionally to begin with (e.g. Mandarin Chinese).

Based on these classifications, we can now proceed in a similar way as with the dependent verb forms above. Table 3 cross-classifies the basic contrast in TAM inflection with the three positioning types of complement constructions.

Table 3. Basic TAM contrasts and the position of complement clauses

	Postverbal	Preverbal	Flexible	Totals
Non-reduced TAM	66	35	5	106
Reduced TAM	40	41	10	91
Neutral TAM	8	0	0	8
Totals	**114**	**76**	**15**	**205**

Table 3 yields a very similar distribution to our earlier one on dependent verb forms: The reduction of TAM categories is again most likely in flexible complements (10/15 = 66.7%), intermediate in preverbal complements (41/76 = 53.9%), and lowest for postverbal complements (40/106 = 37.7% [the 'neutral' cases were subtracted from the totals]). If we submit the most relevant contrast, i.e. the figures framed in Table 3, to a Fisher exact test, the difference between pre- and postverbal constructions is statistically significant ($p = 0.03$). Again, the results improve if all complementation strategies are removed from the analysis. In this case, the proportion of TAM-reduced preverbal complements rises to 56%, while that for postverbal complements decreases to 33.3% (Fisher exact test, $p = 0.005$).

Our second inflectional category is that of person indexation, and for each language in our sample, we determined how subject indexes are coded in complement clauses. Again, we distinguished three basic levels:

- **Canonical**: The subject index is retained and appears in its canonical form.

– **Non-canonical**: The subject index is obligatorily deleted or else takes a special form (e.g. possessive rather than personal indexes, or a special subordinate agreement paradigm [e.g. Urarina 'D-forms', cf. Olawsky 2006: 489]).
– **Neutral**: The language in question lacks subject indexes as a conjugation category.

Where agreement properties differ depending on whether or not the subject of the complement clause is coreferential with the subject of the matrix clause, indexation under different-subject conditions was taken as criterial; this is because same-subject contexts are more prone to equi-deletion in the first place, and this would have concealed a substantial amount of overt agreement that is present in all other contexts.

The distribution of the three agreement categories over the basic positional types of complement clauses is displayed in Table 4.

Table 4. Basic subject indexation contrasts and the position of complement clauses

	Postverbal	Preverbal	Flexible	Totals
Canonical index	54	23	6	83
Non-canonical index	25	31	6	62
Neutral indexation	35	22	3	60
Totals	**114**	**76**	**15**	**205**

As one can see, subject indexation, too, is distributed unevenly across the three position types. This time, the highest amount of non-canonical marking is found in preverbal complements (31/54 = 57.4%, again discarding the 'neutral' cases). This is followed by flexible complements (6/12 = 50%) and, finally, by the postverbal constructions (25/79 = 31.6%). The two-way contrasts framed in Table 4 are highly significant in a Fisher exact test ($p = 0.004$), and again the results become more pronounced if complementation strategies are removed: Non-canonical agreement rises to 60% in preverbal complements and slightly reduces to 29.2% in postverbal complements. Therefore, subject indexation shows a similar contrast as TAM marking, which in turn demonstrates that the inflectional categories of the complement verb co-vary with the position of the subordinate clause in the same way that their derivational status does.

4.3 Further morphosyntactic properties of the complement clause

The general picture that emerges from the previous results is that preverbal complementation constructions, whether they are complement clauses proper or include complementation strategies, are relatively more susceptible to being 'desententialised' (Lehmann 1988) or structurally constrained. This hypothesis is further supported by a range of qualitative observations in our data. A thorough discussion is beyond the scope of this paper, but in most general terms, these observations fall into two different classes.

First, there are some constructions in our database whose overall syntactic complexity is either categorically or preferably reduced to a small number of 'core' elements, e.g. the dependent verb only, the verb and its object (but no adjuncts), etc. Often, this yields the impression of "impoverished clauses" (McGregor [1988: 45] on one construction in Gooniyandi), and we find them, for instance, with certain complementation patterns in Barasano, Imonda, Jarawara, Warao, Wari' or Wambaya.

In Barasano, for example, one complementation construction involves a bare verb stem deprived of any morphological marking; in addition, argument NPs are either obligatorily deleted (subject) or preferably left implicit (object), and typically there is no other material present in the subordinate clause:

(10) Barasano (Tucanoan: Colombia)

B~~ü~~a-re [*sãha*] *roti-be-a-ha* *y~~ü~~.*
3SG-OBJ enter order-NEG-PRS-3 1SG
'I am ordering him/her not to enter.'
(Jones and Jones 1991: 163)

The syntactic complexity of the subordinate clause can also be restricted by constraints on the transitivity of the complement: In some of the sample languages, one of the complement clauses needs to be syntactically intransitive (or detransitivised by valency-changing operations if a second participant is to be expressed, e.g. Mosetén, Tümpisa Shoshone, West Greenlandic) or cannot appear in the canonical transitive pattern (e.g. Yuchi). Although the occurrences of such restrictions on the complexity or argument structure of complements are limited and hence defy rigorous quantification at the present stage, we can discern a certain bias for the relevant constructions to be preverbal complements. It remains to be seen whether future research can confirm this in a larger dataset.

The second class of qualitative observations involves cases where different positions of the same type of complement clause are possible, but the preverbal position requires a morphosyntactic reduction of the complement in comparison to the postverbal position. To give an example of this pattern, in some VO-languages a postverbal complement clause can be fronted for discourse-pragmatic purposes, and this marked option is concomitant with a certain pressure to keep the complement structurally 'compact'. Thus for Mosetén (Mosetenan: Bolivia), Sakel (2004: 433) reports that "this fronting of an object complement is found when the object complement is rather short". The mirror image of this sort of constraint is found when certain OV-languages with preverbal complement clauses restrict the possibility of right-extraposition to the structurally more elaborate types of complement. In Choctaw, for instance, extraposition is "possible and relatively frequent" (Broadwell 2006: 46) with the most productive, relatively 'finite' type of complement, but it appears to be strongly dispreferred with so-called 'equi constructions' that "are somewhat comparable to English infinitivals" (Broadwell 2006: 281). The following examples thus reflect common distribution patterns in Choctaw:

(11) Choctaw (Muskogean: USA)

a. right-extraposed 'neutral' complement with complementizer

John-at *anokfilli-h* [*alikchi-it* *Bill* *lhakoffichi-tokã*].
John-NOM think-TNS doctor-NOM Bill cure-PST.COMP.DS
'John thinks that the doctor cured Bill.'
(Broadwell 2006: 47)

b. in-situ 'equi' complement without person marking and 'default' tense only

[*Tamaaha'* *iya-h*] *sa-banna-h.*
town go-TNS 1SG.II-want-TNS
'I want to go to town.'
(Broadwell 2006: 282)

On the whole, what we hope to have established in this part of the paper is that there is an important pattern of co-variation in the typology of complement clauses that biases preverbal complements more strongly towards morphosyntactic reduction than postverbal ones, which typically share much of the morphosyntax of independent clauses.

5 The interaction of structure, function and position in complement clauses

The preceding sections of this paper have argued that complement clauses are overall more frequent after the matrix predicate than before it, and that preverbal complements are more strongly desententialised or downgraded than postverbal complements. As was noted in the introduction to this paper, previous research has shown that the degree of downgrading also correlates with the semantics of the complement-taking predicates, henceforth CTPs (Givón 1980). Therefore, the final step in this paper is now to investigate how these two correlations interact with one another: Does the correlation between clause order and syntactic structure arise because position, just like structure, correlates with the semantics of the CTP, or does it hold independently of the CTP? In order to shed light on this issue, we first analyse the relationship between the form and the function (i.e. CTP semantics) of the complement clause, so as to see whether our data are in keeping with Givón's 'binding hierarchy' (Section 5.1). Secondly, we examine whether different CTP classes are associated with significant preferences for pre- or postposing of the complement (Section 5.2).

5.1 Evidence for the binding hierarchy in complementation systems

Building on previous classifications by Noonan (2007: 120–145) and others, we distinguished nine contexts for the occurrence of object complements, each defining a relatively coherent semantic class of CTP that also plays a role in Givón's original study:

- **Phasal** (sometimes also called 'aspectual') predicates include CTPs profiling the inception ('begin', 'start'), continuation ('continue') or termination ('finish') of an event.
- **Causative** predicates include CTPs profiling the (physical) coercion of an affected participant into (bringing about) a resultant situation ('make', 'cause', 'force', etc.).
- **Jussive** predicates are similar to causative ones (and often accommodated together with them under 'manipulative CTPs'), but they profile the verbal coercion of an affected participant into (bringing about) a resultant situation ('command', 'order', 'ask'). They thus include an illocutionary act and in this respect resemble utterance predicates (cf. below).

- **Desiderative** predicates are defined here narrowly as those corresponding to English *want*.[9] The class of desideratives was, however, divided into two subgroups. A pilot study and previous research (e.g. Givón 1980, Haspelmath 2013, Khanina 2009) had indicated that **same-subject** and **different-subject** 'want'-constructions can exhibit very different selection patterns as far as their complement clauses are concerned, and hence we coded the two scenarios as different CTP classes.
- **Perception** predicates are defined here as those encoding immediate perception of the complement event by an experiencing participant of the matrix clause. The sensory mode is typically visual ('see', 'watch') or auditory ('hear'); evidential uses of the same predicates are very different and were not be considered in this study (cf. Boye 2010 for an overview of these different uses of perception verbs).
- **Knowledge** predicates prototypically comprise the equivalents of English *know* and *realise/discover*, respectively, and are restricted in the present paper to knowledge of a declarative proposition (e.g. *I know **that** Hannah quit her job.*). Procedural-knowledge complements (e.g. *I know **how to** fix a car.*) are again very different semantically (they rather code the modal notion of 'ability') and are hence not taken into account.
- **Propositional-attitude** predicates profile "an attitude regarding the truth of the proposition expressed as their complement" (Noonan 2007: 124), and are limited here to 'positive' attitudes of this kind (i.e. 'think', 'believe' but not 'doubt' or 'deny').
- **Utterance** predicates profile "a simple transfer of information initiated by an agentive subject" (Noonan 2007: 121) and correspond to English *say* or *tell*. The complementation of utterance predicates often involves a certain amount of deictic adjustments of the quoted material, resulting in 'indirect speech', but this is not a necessary requirement across languages (cf. also Güldemann 2008, Jäger 2007, Spronck 2012 for recent treatments of quotative constructions in typological perspective[10]). We included direct-speech clauses as complementation strategies where no indirect speech is available, and only if the direct-speech clause functions as a proper syntactic argument in the main clause. Finally, a restriction was made to declarative complements; 'indirect questions' were not considered.

We then coded whether a given complementation pattern in our data can co-occur with these predicate classes or not. Determining the precise CTP distribution of each complement clause was one of the most time-consuming parts of the analysis, and we are particularly grateful for the help of some informants here. For the present purposes, we recorded whether a given construction is or is not attested with matrix verbs of the above predicate classes.[11] The functional profile thus obtained supplements the

[9] This is because other desiderative notions, like 'wish' and 'hope' (cf. Noonan 2007: 132, Khanina 2008), are not covered systematically across all sources considered (cf. also Cristofaro 2003 for the same procedure due to the same predicament). Where a language has lexicalised 'want', 'wish' and 'hope' separately, only 'want' was taken into account.

[10] Note that the term 'quotative' is used in a variety of different senses in the literature (cf. Wiemer and Kampf 2013 for discussion). For the purposes of the present paper, 'quotative' is to be understood in reference to any biclausal reported-speech construction involving an explicit referential source to which the quote is attributed (this is sometimes contrasted with 'reportative constructions', cf. Spronck 2012). We would like to thank Björn Wiemer for insightful discussions of the terminology.

[11] This binary contrast between 'attested' and 'not attested' does, of course, reduce the complexity of the actual co-occurrence patterns. In reality, there is a much more fine-grained gradation of the type and token frequency with which a given CTP class attracts the complement in question, but this will have to

morphosyntactic information on each complement construction in the database, and it now becomes possible to take a different analytical perspective on the data: For each of the nine predicate classes above, we can determine whether the constructions that co-occur with it tend to have a full-blown or a rather reduced structural make-up. The relatively small number of total observations in our sample (N = 205) militates against including several dimensions of syntactic structure simultaneously; therefore, we concentrated once more on the variable that is most suggestive of the morphosyntactic nature of the complement, i.e. the form of the verb. With regard to this parameter, we shall continue to work with the basic contrast between 'dependent' and 'independent' verb forms introduced in Section 4.1 above. Figure 1 displays the relative amount of dependent verb forms for our nine CTP classes.

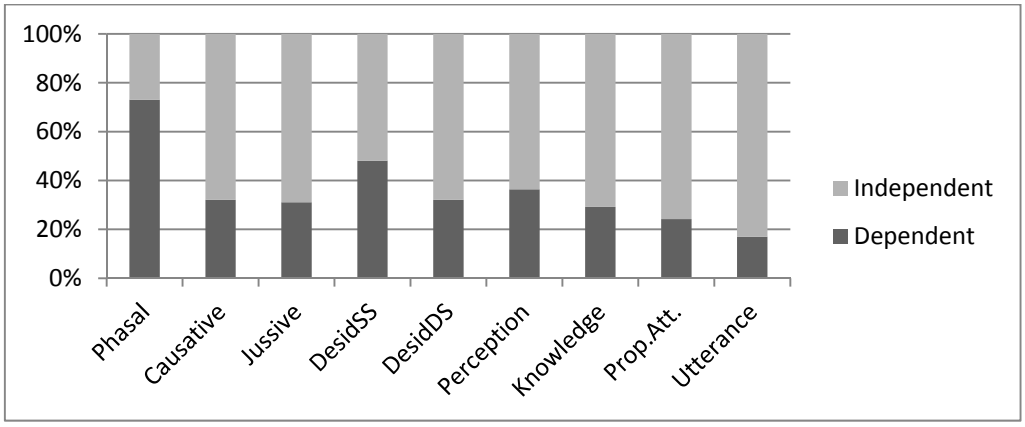

Figure 1. Proportions of dependent verb forms in complementation constructions of different CTP classes

A first glance at Figure 1 reveals that dependent verb forms are distributed unevenly across the predicate classes; not surprisingly, therefore, the overall distribution is highly significant in a randomised Chi-squared analysis ($\chi^2 = 60.75$, $p < 0.0001$, B = 100,000, Cramer's $V = 0.276$). It is easy to see that phasal and utterance predicates behave in opposite ways and thus induce the strongest skewing in the data. If we think of the predicate classes as forming a continuum, defined by decreasing preferences for dependent verb forms, phasal and utterance verbs can be seen as the end points of this continuum on either side. The part in the middle, by contrast, is much harder to evaluate based on visual inspection of Figure 1 alone. A more revealing picture can be obtained if we calculate the relative (dis)similarity of the predicate classes to each other, based on their proportions of dependent verb forms. Using multidimensional scaling as a visualisation technique[12], we can then arrange the predicate classes in a one-

be ignored for the moment (cf. Schmidtke-Bode 2014: Ch.6 for an analysis that incorporates such additional dimensions of co-occurrence).

[12] The dissimilarity calculation was performed via the function `dist` in *R* 2.13.0 (*R* Development Core Team 2011). The dissimilarity matrix was then subjected to metric multidimensional scaling (cf. Kruskal and Wish 1978) and the resulting MDS values (applying `cmdscale` with k=1) were plotted as a one-dimensional vertical scale.

dimensional space, so that the dissimilarities are reflected by relative Euclidian distances on a scale. The first column of Table 5 presents the result of this procedure.

Table 5. Implicational hierarchies of form-function mappings in complementation

our data	Givón 1980	Cristofaro 2003
Phasal	Phasals	Phasals
	Causative (DesiderativeSS)	Causative
DesidSS	Jussive	Jussive
	Desiderative	Desiderative
Perception	Knowledge	Perception
DesidDS	Propositional Attitude	Knowledge
Jussive		Propositional Attitude
Knowledge	Utterance	Utterance
Prop.Att.		
Utterance		

What we superimposed on our scale in Table 5 (indicated by dotted lines) is the statistical 'breakpoints', as it were, i.e. the cut-off points on the hierarchy at which a significant change in the preference for dependent verb forms can be noted. That is, same-subject desiderative verbs already differ significantly from phasal verbs in the amount of dependent verb forms. The next breakpoint of this sort is found with jussive verbs (which are only very slightly lower on the scale than causative verbs and hence overlap in the graph); the classes in between (i.e. perception, different-subject desiderative and causative) do not differ significantly from one another, although the graph also shows that they do not occur at equal intervals. Essentially the same applies to the gradation of jussive, knowledge, propositional-attitude and utterance verbs: the difference between the end poles is significant, but the internal dissimilarities are too small to yield a statistical signal.

Let us briefly compare our results to the hierarchies proposed by Givón (1980) and Cristofaro (2003), as displayed in the other columns of Table 5. All three studies are unanimous as far as the very top of the hierarchy, i.e. phasal predicates, is concerned, and basically the same holds for the lower part of the scale: In Givón (1980), knowledge, propositional-attitude and utterance predicates are subsumed under a larger rubric of 'cognition-utterance verbs'. These are shown to be internally graded according to certain semantic 'binding' principles (roughly: the degree of emotional commitment of the matrix agent to the truth or realisation of the complement event), which in turn yield the cline 'hope/remember/forget' > 'think/believe/know' > 'say'. Our data can confirm this, although, in keeping with Cristofaro, we find that the group-internal differences between the relevant notions (e.g. 'think/believe', 'know', 'say') are not

statistically significant, so that a genuine ranking is infeasible here. This empirically reflects Givón's suspicion that "languages do not always exhibit many coding points along [the] cognition-utterance verb portion" (Givón 1980: 363). Interestingly, Givón (1980) does not discuss perception verbs, but it is likely that they qualify as 'cognition verbs' in his sense of the term, and our data are in keeping with Cristofaro's in that perception verbs select markedly different complements from the other cognition-utterance verbs. They thus shade into the upper portions of the hierarchy, and it is in these portions of Givón's and Cristofaro's studies that we find differences to our own results.

Most notably, the relative position of causative and jussive CTPs is radically different in our data. For jussives, this is likely due to the fact that Givón's notion of jussive does not appear to include the cases in which a quotative construction is used for (indirect) commands. As was stated above, however, it is precisely the presence of an illocutionary act that sets jussive complements apart from other manipulative ones and induces certain similarities with utterance predicates; therefore, it is not entirely surprising that the jussive predicates in our data oscillate between the two poles of morphosyntactic coding. However, following this logic, the relatively low ranking of causative CTPs is somewhat unexpected; in fact, they hardly differ from jussives CTPs in our data. This finding is likely due to the following things: First, Givón's paper emphasises the fact that, cross-linguistically, causative predicates are particularly prone to clause union, lexicalisation ('lexical causatives') or grammaticalisation (causative affixes), and they share this property with phasal predicates. In our data, too, the overall number of constructions that can cover phasal or causative CTPs is much lower than that of the other CTPs, indicating their propensity for being coded by monoclausal alternatives to complementation. Where causative CTPs *are* coded by complement clauses, however, they do not generally object to independent verb forms. And here the difference to Cristofaro's finding may be grounded in the fact that she explicitly neglected predicates of more indirect causation, such as 'cause' itself, while we adopted a more embracing approach.

Finally, our results demonstrate the empirical reality of two types of 'want'-scenarios: While these are not systematically distinguished in either Givón's or Cristofaro's study, Givón (1980: 345) does suggest that 'want' often partakes in several predicate classes, and this typically entails a difference in participant sharing: same-subject 'want' predicates belong to Givón's 'modality verbs', and hence are expected to pattern at least to some degree with phasals and other high-ranking verbs, while different-subject 'want' predicates constitute the lower end of Givón's 'manipulative' class, which often shades seamlessly into the cognition-utterance group. This is, by and large, what we find in our data: same-subject desiderative contexts claim second rank in the preference for dependent verb forms, while different-subject desideratives pattern statistically with perception, causative and jussive verbs.

Overall, then, our data confirm that there is a systematic correlation between the form of the complement and the semantic class of the CTP. The data are best conceived of as a continuum of increasing morphosyntactic independence from a 'phasal' to an 'utterance' pole, along which some significant cut-off points may be identifiable (depending on the precise definition of the predicate classes under scrutiny).

5.2 Clause order and the binding hierarchy

Given our previous findings that certain predicate classes tend towards morphosyntactic independence (notably '[cognition-]utterance verbs' in Givón's sense) and that such morphosyntactic independence is more characteristic of postverbal than of preverbal complement clauses, a logical conclusion could be that utterance verbs are particularly closely associated with postverbal position, while phasal predicates prefer preverbal position. In other words, the CTP classes may show, in their preferred positional choices, a hierarchical pattern similar to the 'binding hierarchy'. In fact, if this is true, then the correlation between form and position of the complement clause would be an epiphenomenon, i.e. a side-effect of the different predicate classes being coded *and ordered* in particular ways. In order to examine this possibility, we can perform the same kind of analysis as in Section 5.1 above, but with regard to positional patterns.

Specifically, we can determine the proportion of pre- and postverbal complements for each predicate class; since we saw in Section 4.1 above that complements with a flexible ordering pattern structurally with preverbal clauses, we can conflate the two types into a single class that, as a whole, represents a deviation from the cross-linguistically dominant (i.e. postverbal) positioning type of complement clauses. The results of the analysis are displayed in Figure 2.

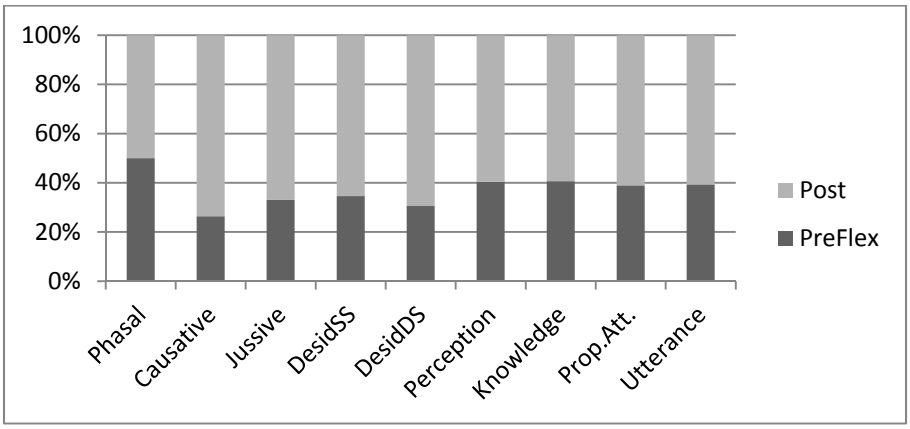

Figure 2. Proportions of preverbal and flexible complementation constructions in different CTP classes

The overall distribution looks much more homogeneous than the one in Figure 1 above: In contrast to the binding hierarchy, there are no glaring oppositions here; the only predicate class with a marked preference in our data is that of causative predicates, and when this is compared to the most dissimilar data points (notably phasal predicates), individually significant results can be obtained, but these are not strong enough to yield a statistical signal for the distribution as a whole (randomised $\chi^2 = 9.91$, $p = 0.27$, $B = 100{,}000$), and most of the other contrasts are non-significant, anyway. The statistics remain constant if the flexible complements are removed from the analysis, and also if all complementation strategies are discarded. Apparently, then, there is no particular skewing of individual predicate classes towards a certain positioning type (except for causatives), and no straightforward implicational hierarchy suggests itself. In fact, if the

data in Figure 2 are coerced onto a scale of relative dissimilarity again, the result looks strikingly different from the one for the binding hierarchy (Figure 3):

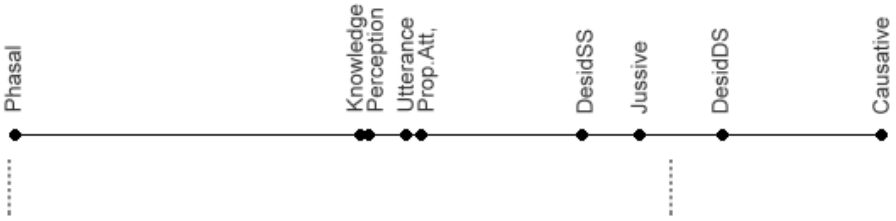

Figure 3. Scale of relative dissimilarity of predicate classes according to preferred position of their complements

Figure 3 indicates, firstly, the absence of significant differences between the predicate classes: Looking at the cline from the 'phasal' end, there is only one 'breakpoint' in the data, between jussive and different-subject desiderative CTPs (the significance level of $\alpha = .05$ lies between them, jussives scoring slightly above and DS-'want' slightly below this value). Secondly, while phasals persist at the top of the scale, just as on the binding hierarchy, the remainder is organised very differently. The 'shift' of causatives to the right end presumably reflects iconicity of sequence, the causative CTP preceding the effect coded by the complement. What is remarkable from the perspective of the binding hierarchy is that Givón's 'cognition-utterance' verbs have now moved up the scale; utterance verbs, in particular, do no longer enter into a significant contrast with the phasal class at the top, which was the most pronounced opposition on the form-function scale.

In fact, a qualitative analysis of our data suggests that it is not uncommon for quotative clauses to *precede* their CTP. There are three recurring qualitative observations in our data that are interesting in this connection. First, there are languages in which utterance complements have a stronger tendency for preverbal position than other CTPs. In Tümpisa Shoshone (SOV), for instance, an indirect-speech complement appears to be less prone to right-extraposition than other complementation constructions (Dayley 1989: 374–383). Similarly, for clausal objects in Choctaw (SOV), we learn that "both right and left extraposition are possible and relatively frequent. Left extraposition appears to be more frequent with verbs of saying, while right extraposition is more common with other verbs" (Broadwell 2006: 46). Second, there are languages in which direct speech is (or can be) right-branching but indirect speech is coded by a preverbal complement clause, thus contributing to the picture in Figure 3; such languages include Awa Pit (cf. example [4] above), Turkish and Lezgian. Finally, we also find the opposite constellations, i.e. indirect speech being drawn into the postverbal pattern of complement clauses and direct speech being expressed by preverbal (or at least more flexible) clauses. While these cases do not immediately contribute to the distribution above, they are noteworthy because they underline the fact that it is not uncommon for reported discourse to precede the utterance verb. Examples of this pattern can be found (to varying degrees) in Karo Batak, Epena Pedee, Mayogo, Lao, Tepehua, Tzutujil, Yagua and others. In sum, while it is true, as the binding hierarchy suggests, that utterance verbs typically co-occur with fully clausal structures from a

morphosyntactic point of view, there is no reason to assume that these constructions generally appear in postverbal position (whether they code direct or indirect discourse). Similar remarks apply to the complements of propositional-attitude and knowledge predicates.

Apparently, then, complement-taking predicates are relatively more selective with regard to the *structure* of their complements than with regard to their *positional* preferences. If this is true, the combined results of Section 4 and Section 5 of this paper would suggest the following: Given that (i) preverbal complements are relatively more prone to morphosyntactic reduction than postverbal complements, and that (ii) all predicate classes occur with both pre- and postverbal complements with few statistically significant differences, the complements of all predicate classes should reflect the correlation between position and structure. This prediction is tested in the last analysis of this paper. Since we are now investigating the association structure of three categorical variables (position, structure and function of the complement), an appropriate multivariate method is called for. We employed loglinear analysis (cf. Agresti and Finlay 2009: 504), one important component of which is measuring the mutual association of two variables (here: clause order and verb form) at each category of a third variable (here: the CTP class). Therefore, the method is precisely geared towards testing the above prediction. We retain the variables and levels from the previous sections:

- **Position** of the complement: postverbal versus preverbal/flexible
- **Structure** of the complement: independent versus dependent verb form
- **CTP classes** that the complement can co-occur with: the 9 classes from above

The loglinear analysis reveals that the best-fitting statistical model of our data[13] contains

i. a significant association of verb form and position of the complement (in keeping with Table 2),
ii. a significant association of verb form and CTP class (in keeping with Figure 1),
iii. no significant association of CTP class and position (in keeping with Figure 2),
iv. and no significant three-way interaction between all variables.

The latter finding is the most important one in the present context because it essentially means that the CTP class does not change the association pattern between the verb form and the position of the complement. Instead, this association follows the same direction across all CTP classes. This can be shown by pairwise post-hoc tests to the loglinear analysis, and the results are summarised in Table 6 below.

[13] The loglinear analysis was performed by using the function `loglm` in *R*. Applying hierarchical backward elimination of terms from a saturated model, we arrived at a final model that neither contains the three-way interaction between verb form, position and CTP class, nor the two-way interaction between position and CTP class, but all other terms. The final model has a likelihood ratio of $\chi^2\,(16) = 17.18$, $p = 0.374$, indicating that this model is not significantly different from (and hence provides a reasonably good fit to) our actually observed frequencies in the data.

Table 6. Proportions of dependent verb forms in pre- and postverbal complements of each predicate class

	PreFlex	Post	Total	*p* (FET)
Phasal	0.88	0.58	0.76	0.027*
Causative	0.50	0.26	0.32	0.11
Jussive	0.41	0.28	0.32	0.183
DesidSS	0.71	0.36	0.48	0.002**
DesidDS	0.58	0.22	0.33	0.003**
Perception	0.63	0.20	0.37	4.09E-05***
Knowledge	0.54	0.14	0.30	4.19E-05***
Prop.Att.	0.41	0.16	0.25	0.008**
Utterance	0.32	0.09	0.18	0.005**

Significance levels in a Fisher exact test (FET): *p < 0.05; **p < 0.01; ***p < 0.001.[14]

Table 6 shows that all of the CTP classes distinguished in the present paper feature a larger proportion of dependent verb forms if they are coded by preverbal or flexible complements (the results remain stable even if flexible complements are removed). Crucially, apart from causative and jussive predicates, all CTP classes yield a statistically significant contrast. That is, even utterance predicates, while showing the lowest degree of structural reduction overall, are significantly more dependent in preverbal position. Thus indirect-speech complements in, say, Korafe, Dolakha Newar, Turkish or Wolaytta attract the dependent morphology characteristic of preverbal complements. Conversely, some of the predicate classes higher up on the binding hierarchy (e.g. same-subject 'want' and perception verbs) are *not* preferably reduced when they are coded by postverbal complements, but *are* often reduced as preverbal clauses. Therefore, we come to conclude that the morphosyntactic structure of complement clauses is not only determined by semantic considerations, but also, to a considerable degree, by the position of the complement vis-à-vis the CTP.

6 Summary and conclusion

This paper has been an endeavour to approach the typology of complement clauses from the perspective of linear order, and to shed new empirical light on how these ordering patterns interact with structural and semantic properties of the complement. Based on a genetically controlled sample of 100 languages, we first established the major positioning types of complement clauses. We then went on to argue that preverbal complements are typologically 'marked', not only with regard to their position, but also with respect to their morphosyntactic structure: In contrast to postverbal complements, which typically have an internal structure similar to independent main clauses, preverbal complements are significantly more likely to contain dependent, often derived verb forms, and to show a corresponding or additional loss of inflectional categories and morphosyntactic complexity overall. A similar contrast had previously been noted to

[14] The fact that the nine individual calculations draw on partially overlapping samples can be accommodated by adjusting the significance level: $\alpha' = \alpha/n = 0.05/9 = 0.006$. On this more rigorous (but also very conservative) calculation, most of the results remain highly significant. Only phasal and propositional-attitude predicates can no longer hold up to the lower α-level.

exist at a paradigmatic level and hence orthogonally to our syntagmatic dimension: Givón (1980) and others have argued that different classes of complement-taking predicates tend to go radically different ways as far as the structure of their complements is concerned, and that the predicate classes can be arranged accordingly on an implicational scale. In keeping with this proposed 'binding hierarchy', our data attest that some predicates are relatively more prone to select morphosyntactically reduced complements than others. However, we added the insight that *all* predicate classes increase the likelihood of choosing such reduced complements when these appear in preverbal position. This holds for high-ranking and low-ranking classes alike, and some of them may even change their overall structural preferences depending on the position of the complement: Desiderative, perception and knowledge predicates, for example, preferably co-occur with independent clauses if these are postverbal constructions, but acquire a preference for reduced complements if these precede the CTP. Our findings suggest, therefore, that the relationship between form and position is of a more general nature that can apply independently of semantic considerations. We thus propose that, in addition to the binding hierarchy, a second type of associative relationship should be recognised in the typology of complementation: the co-variation of linear order and morphosyntactic structure.

Correspondence address

Department of English, Project "Grammatical Universals", Leipzig University (IPF 141199), Nikolaistr. 6-10, 04109 Leipzig, Germany. e-mail: kschmidtkebode@gmail.com

Acknowledgements

We would like to express our particular gratitude to our informants, who are listed alongside the respective language in the Appendix, to the reviewers of the first draft version of this paper, and to two anonymous reviewers for *Linguistics* for their valuable comments. Part of the research for the present study was supported by a grant from the *Deutsche Forschungsgemeinschaft* (DFG), which is hereby gratefully acknowledged.

Abbreviations

The paper abides by the Leipzig Glossing Rules. Additional glosses adopted from the original sources include:

DIR	directive pronoun	DS	different subject	INTS	intensifier
L, N	low-tone prefix	NFUT	non-future	NONLOCUT	non-locutor
NPST	non-past	PR	pronominal	PR.REL	present relative form
RECPF	recent perfect	REP	repetition	SEQ	sequential
SPEC	specific determiner	TNS	(default) tense	VE	ventive

References

Agresti, Alan & Barbara Finlay. 2009. *Statistical methods for the social sciences,* 4th edn. Upper Saddle River, NJ: Pearson Prentice Hall.

Andrews, Avery D. 2007. Relative clauses. In Timothy Shopen (ed.), *Language typology and syntactic description, Vol. II: Complex constructions,* 2nd ed., 206–236. Cambridge: Cambridge University Press.

Bakker, Dik. 2011. Language sampling. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*, 100–129. Oxford: Oxford University Press.

Blackings, Mairi & Nigel Fabb. 2003. *A grammar of Ma'di.* Berlin & New York: Mouton de Gruyter.

Boye, Kasper. 2010. Reference and clausal perception-verb complements. *Linguistics* 48(2). 391–430.

Broadwell, George Aaron. 2006. *A Choctaw reference grammar*. Lincoln & London: University of Nebraska Press.

Cristofaro, Sonia. 2003. *Subordination.* Oxford: Oxford University Press.

Curnow, Timothy J. 1997. *A grammar of Awa Pit (Cuaiquer)*. Canberra: Australian National University dissertation.

Dayley, Jon P. 1989. *Tümpisa (Panamint) Shoshone grammar*. Berkeley: UC Press.

Dench, Alan C. 1995. *Martuthunira.* Canberra: Pacific Linguistics.

Diessel, Holger. 2001. The ordering distribution of main and adverbial clauses: A typological study. *Language* 77(2). 433–455.

Diessel, Holger & Katja Hetterle. 2011. Causal clauses: A cross-linguistic investigation of their structure, meaning, and use. In Peter Siemund (ed.), *Linguistic universals and language variation,* 21–52. Berlin & New York: Mouton de Gruyter.

Dixon, Robert M.W. 1995. Complement clauses and complementation strategies. In Frank R. Palmer (ed.), *Grammar and meaning. Essays in honour of Sir John Lyons,* 175–220. Cambridge: Cambridge University Press.

Dixon, Robert M.W. 2004. *The Jarawara language of Southern Amazonia*. Oxford: Oxford University Press.

Dixon, Robert M.W. 2006. Complement clauses and complementation strategies in typological perspective. In Robert M.W. Dixon & Alexandra Y. Aikhenvald (eds.), *Complementation: A cross-linguistic typology,* 1–48. Oxford: Oxford University Press.

Dryer, Matthew S. 1980. The positional tendencies of sentential noun phrases in universal grammar. *Canadian Journal of Linguistics* 25(2). 123–195.

Dryer, Matthew S. 2011a. Order of relative clause and noun. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures online.* Max Planck Digital Library, chapter 90A. http://wals.info/chapter/90A (accessed 31 January 2013).

Dryer, Matthew S. 2011b. Order of object and verb. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures online.* Max Planck Digital Library, chapter 83A. http://wals.info/chapter/83A (accessed 25 February 2013).

Farr, Cynthia J.M. 1999. *The interface between syntax and discourse in Korafe*. Canberra: Pacific Linguistics.

Fleck, David W. 2006. Complement clause type and complementation strategies in Matses. In Robert M.W. Dixon & Alexandra Y. Aikhenvald (eds.), *Complementation: A cross-linguistic typology,* 224–244. Oxford: Oxford University Press.

Foley, William A. 1991. *The Yimas language of New Guinea.* Stanford: Stanford University Press.

Foley, William A. & Robert D. VanValin Jr. 1984. *Functional syntax and universal grammar.* Cambridge: Cambridge University Press.

Givón, Talmy. 1980. The binding hierarchy and the typology of complements. *Studies in Language* 4. 333–377.

Givón, Talmy. 2001. *Syntax: An introduction, Vol. II.* Amsterdam & Philadelphia: John Benjamins.

Greenberg, Joseph H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of language,* 73–113. Cambridge, MA: MIT Press.

Grosu, Alexander & Sandra A. Thompson. 1977. Constraints on the distribution of NP clauses. *Language* 53. 104–151.

Güldemann, Tom. 2008. *Quotative indexes in African languages: A synchronic and diachronic survey*. Berlin & New York: Mouton de Gruyter.

Haspelmath, Martin. 1989. From purposive to infinitive – a universal path of grammaticization. *Folia Linguistica Historica* 10(1–2). 287–310.

Haspelmath, Martin. 1995. The converb as a cross-linguistically valid category. In Martin Haspelmath & Ekkehard König (eds.), *Converbs in cross-linguistic perspective*, 1–55. Berlin & New York: Mouton de Gruyter.

Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86(3). 663–687.

Haspelmath, Martin. 2013. On the cross-linguistic distribution of same-subject and different-subject 'want' complements: Economic vs. iconic motivation. *SKY Journal of Linguistics* 26. 41–69.

Heath, Jeffrey. 2008. *A grammar of Jamsay.* Berlin & New York: Mouton de Gruyter.

Horie, Kaoru. 2001. Complement clauses. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, & Wolfgang Raible (eds.), *Handbook of language typology and language universals.* 979–993. Berlin & New York: Walter de Gruyter.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language.* Cambridge: Cambridge University Press.

Jäger, Andreas. 2007. Coding strategies of indirect reported speech in typological perspective. *University of Queensland Working Papers in Linguistics* 1. http://www.library.uq.edu.au/ojs/index.php/uqwpl/article/view/16/17 (accessed 14 November 2011).

Janssen, Dirk, Balthasar Bickel & Fernando Zúñiga. 2006. Randomization tests in language typology. *Linguistic Typology* 10. 419–440.

Jones, Wendell & Paula Jones. 1991. *Barasano syntax.* Arlington: The Summer Institute of Linguistics and The University of Texas at Arlington.

Kenesei, István, Robert M. Vago & Anna Fenyvesi (1998). *Hungarian*. London & New York: Routledge.

Khanina, Olesya. 2008. How universal is *wanting*? *Studies in Language* 32(4). 818–865.

Khanina, Olesya. 2009. The symbiosis of descriptive linguistics and typology: A case study of desideratives. In Patience Epps & Alexandre Arkhipov (eds.), *New challenges in typology: Transcending the borders and redefining the distinctions,* 199–220. Berlin & New York: Mouton de Gruyter.

Kruskal, Joseph B. & Myron Wish. 1978. *Multidimensional scaling*. Newbury Park etc.: Sage Publications.

Lehmann, Christian. 1984. *Der Relativsatz: Typologie seiner Strukturen. Theorie seiner Funktionen. Kompendium seiner Grammatik*. Tübingen: Gunter Narr.

Lehmann, Christian. 1988. Towards a typology of clause linkage. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse,* 181–226. Amsterdam & Philadelphia: John Benjamins.

Lichtenberk, Frantisek. 2008. *A grammar of Toqabaqita.* Berlin & New York: Mouton de Gruyter.

Macaulay, Monica. 1996. *A grammar of Chalcatongo Mixtec*. Berkeley etc.: University of California Press.

Maslova, Elena. 2003. *A grammar of Kolyma Yukaghir*. Berlin: Mouton de Gruyter.

McGregor, William. 1988. Mood and subordination in Kuniyanti. In Peter Austin (ed.), *Complex sentence constructions in Australian languages,* 37–67. Amsterdam & Philadelphia: John Benjamins.

Munro, Pamela. 1982. On the transitivity of 'say' verbs. In Paul J. Hopper & Sandra A. Thompson (eds.), *Studies in transitivity,* 301–318. New York: Academic Press.

Muysken, Pieter. 1994. Nominalizations. In Ronald E. Asher & James M.Y. Simpson (eds.), *The encyclopedia of language and linguistics,* 2811–2815. Oxford: Pergamon Press.

Nedjalkov; Igor. 1997. *Evenki*. London & New York: Routledge.

Neukom, Lukas. 2001. *Santali.* München: Lincom Europa.

Noonan, Michael. 2007. Complementation. In Timothy Shopen (ed.), *Language typology and syntactic description, Vol. II: Complex constructions,* 2nd edn, 52–150. Cambridge: Cambridge University Press.

Ogihara, Saeko. 2009. *Verb-final typology*. Boulder: University of Colorado dissertation.

Olawsky, Knut J. 2006. *A grammar of Urarina.* Berlin & New York: Mouton de Gruyter.

Onishi, Masayuki (1994). *A grammar of Motuna*. Canberra: Australian National University dissertation.

R Development Core Team. 2011. *R: A Language and Environment for Statistical Computing*. Version 2.13.0. Vienna: R Foundation for Statistical Computing. http://www.r-project.org. (accessed 15 July 2012)

Ransom, Evelyn R. 1986. *Complementation: Its meanings and forms.* Amsterdam: John Benjamins.

Rijkhoff, Jan, Dik Bakker, Kees Hengeveld & Peter Kahrel. 1993. A method of language sampling. *Studies in Language* 17. 169–203.

Sakel, Jeanette. 2004. *A grammar of Mosetén.* Berlin & New York: Mouton de Gruyter.

Schmidtke-Bode, Karsten. 2009. *A typology of purpose clauses.* Amsterdam & Philadelphia: John Benjamins.

Schmidtke-Bode, Karsten. 2014. *Complement clauses and complementation systems: A cross-linguistic study of grammatical organization.* Jena: Friedrich Schiller University Jena dissertation.

Smeets, Ineke. 2008. *A grammar of Mapuche.* Berlin & New York: Mouton de Gruyter.

Spronck, Stef. 2012. Minds divided: Speaker attitudes in quotatives. In Isabelle Buchstaller & Ingrid van Alphen (eds.), *Quotatives: Cross-linguistic and cross-disciplinary perspectives,* 71–116. Amsterdam & Philadelphia: John Benjamins.

Stassen, Leon. 1985. *Comparison and universal grammar.* Oxford: Blackwell.

Thompson, Sandra A., Joseph S. Park & Charles N. Li. 2006. *A reference grammar of Wappo*. Berkeley etc.: University of California Press.

Wiemer, Björn & Veronika Kampf. 2013. Gesten und Mimik als semiotische Substitute für Sprechakte oder umgekehrt? Zur speziellen Verwendung einiger Reformulierungsmarker im Bulgarischen. *Съпоставително езикознание XXXVIII* 2/3. 153–191.

Ylikoski, Yussi. 2003. Defining non-finites: Action nominals, converbs and infinitives. *SKY Journal of Linguistics* 16. 185–237.

Appendix: Sample languages

The sample languages are organised by genetic affiliation (according to *WALS* online) in alphabetical order, and informants (on whatever aspect of the respective complementation system, i.e. not necessarily ordering issues) are listed in brackets.

Afro-Asiatic
- Berber: **Tamashek** (Jeff Heath)
- Chadic: **Hausa** (Mahamane Abdoulaye)
- Cushitic: **Somali**
- Omotic: **Wolaytta**
- Semitic: **Gulf Arabic** (Clive Holes)

Altaic
- Turkic: **Turkish**
- Tungusic: **Evenki** (Igor Nedjalkov)

Arauan: **Jarawara**
Araucanian: **Mapudungun** (Fernando Zúñiga)
Arawakan: **Tariana**
Australian
- Bunuban: **Gooniyandi** (Bill McGregor)
- Mangarrayi: **Mangarayi**
- Pama-Nyungan: **Martuthunira**
- Tangkic: **Kayardild**
- West Barkly: **Wambaya** (Rachel Nordlinger)
- Wororan: **Ungarinjin**

Austro-Asiatic
- Mon-Khmer/Aslian: **Semelai**
- Mon-Khmer/Viet-Muong: **Vietnamese**
- Munda: **Santali** (Lukas Neukom)

Austronesian
- Central Malayo-Polynesian: **Tetun**
- Eastern Malayo-Polynesian/Oceanic: **To'aba'ita**
- Eastern Malayo-Polynesian/ South Halmahera-West New Guinea: **Taba**
- Western Malayo-Polynesian/Sulawesi: **Tukang Besi** (Mark Donohue)
- Western Malayo-Polynesian/Borneo: **Begak Ida'an**
- Western Malayo-Polynesian/Sundic: **Karo Batak** (Geoffrey Woollams)
- Western Malayo-Polynesian/Sama-Bajaw: **Yakan**

Barbacoan: **Awa Pit**
Chapacura-Wanhan: **Wari'**
Chibchan: **Rama**
Chocó: **Epena Pedee**
Chumash: **Barbareño Chumash**
Dravidian: **Malayalam**
East Bougainville: **Motuna**
Eskimo-Aleut: **West Greenlandic** (Michael Fortescue)
Hmong-Mien: **Hmong Njua** (Bettina Harriehausen-Mühlbauer)
Hokan/Yuman: **Jamul Tiipay**
Indo-European
- Germanic: **German**
- Iranian: **Persian** (Peter Öhl)
- Slavic: **Serbo-Croatian** (Wayles Browne, Margita Soldo)

Kartvelian: **Georgian** (Merab Geguchadze)
Khoe-Kwadi: **(Modern) Khwe** (Christa Kilian-Hatz)
Kiowa-Tanoan: **Kiowa**
Lower Mamberamo: **Warembori**
Lower Sepik-Ramu/Lower Sepik: **Yimas**
Makú/Vaupés-Japurá: **Hup** (Pattie Epps)
Mayan/Quichean: **Tzutujil**
Mixe-Zoque: (Chimalapas) **Zoque** (Heidi Johnson, Terje Faarlund, Roberto Zavala Maldonado)
Mosetenan: **Mosetén**
Muskogean: **Choctaw**
Na-Dene/Athapaskan: **Slave** (Keren Rice)
Nakh-Dagestanian: **Lezgian**
Niger-Congo
- Northern Atlantic: **Noon**
- Benue-Congo/Cross-River: **Kana**
- Benue-Congo/Bantoid: **Nkore-Kiga**
- Dogon: **Jamsay** (Jeff Heath)
- Kwa: **Fongbe** (Clair Lefebvre)
- Adamawa-Ubangian: **Mayogo** (Kenneth Sawka)
- Gur: **Supyire**

Nilo-Saharan
- Kadugli: **Krongo**
- Moru-Ma'di: **Ma'di**
- Nilotic: **Lango**
- Songhay: **Koyra Chiini**

Oto-Manguean/Mixtecan: **Chalcatongo Mixtec** (Monica Macaulay)
Panoan: **Matses**
Peba-Yaguan: **Yagua**
Quechua: **Huallaga Quechua**
Salish/Central Salish: **Musqueam**
Sino-Tibetan
- Tibeto-Burman/Bodic: **Dolakha Newar** (Carol Genetti)
- Tibeto-Burman/Burmeso-Lolo: **Burmese**
- Chinese: **Mandarin Chinese** (Martin Schäfer)

Siouan: **Lakota**
Sko/Western Sko: **Skou**
Solomons-East Papuan: **Lavukaleve** (Angela Terrill)
Tai-Kadai: **Lao**
Tarascan: **Purépecha**
Totonacan: **Tepehua** (Susan Kung)
Trans-New Guinea
- Angan: **Menya**
- Binanderean: **Korafe**
- Border: **Imonda**
- Engan: **Kewa**
- Madang: **Amele** (John Roberts)

Tucanoan: **Barasano**
Tupian/Tupí-Guaraní: **Mekens**
Uralic: **Hungarian**
Uto-Aztecan/Numic: **Tümpisa Shoshone**
Wappo-Yukian: **Wappo**
West Papuan/North-Central Bird's Head: **Abun**
Yanomam: **Sanumá**
Yukaghir: **Kolyma Yukaghir**
Isolates
- **Ainu** (Anna Bugaeva)
- **Basque**
- **Japanese** (Kyoko Maezono, Toshio Ohori)
- **Korean**
- **Kwazá**
- **Urarina** (Knut Olawsky)
- **Trumai**
- **Warao**
- **Yuchi**
- **Yuracaré** (Rik van Gijn)

Pidgins and Creoles: **Ndyuka**