Abstract: This typological study based on data from a parallel text corpus is a two-step investigation of selectives (“topic markers”). First, a set of selectives in 81 languages from all continents is compiled on the basis of their occurrence with emphatic personal pronouns in contrast constructions. In a second step, it is explored how this set of markers is used across 19 subordinate clause domains. The results indicate that, despite much crosslinguistic diversity, the distribution of selectives across subordinate clauses is strongly constrained. Selectives in subordinate clauses are distributed following a tendency scale (no strict hierarchy, but no blatant exceptions): general relative clauses rank highest followed by conditional and temporal clauses with concessive and purpose clauses ranking lowest. No postposed subordinate clauses attract selectives. It is further found that selectives tend to occur at the end of the constituent which they have scope over where there is minimal risk of scope ambiguity. Despite the frequent occurrence of selectives on conditional clauses, selectives tend not to be conditional connectives unless this happens to be their grammaticalization source.

Keywords: concessive clauses; conditional clauses; contrast; headless relative clauses; information structure; parallel texts; subordination; temporal clauses; topic; topic marker

1 Introduction

This article is the first large-scale crosslinguistic investigation of a diverse set of grammatical elements that are often referred to as “topic markers” in the descriptive literature and I will investigate to what extent such items are used across various kinds of subordinate clause domains. The markers in question have a broad range of use as illustrated with the element *kuu* in Yonggom (Ok, New Guinea) in (1) and (2) (described as a “topic marker” by Christensen 2010). They can
occur on nominal, pronominal (1) and clausal (2) constituents and can have both contrastive (1) and non-contrastive (2) use.

(1) Yonggom: “topic marker” on pronominal constituents in a contrast construction

\[ Yiib \text{ kuu} \ kiri-dan, \quad ne \text{ kuu} \ ari-mbed \ man-aan. \]
\[ \begin{array}{llllll}
2\text{pl} & \text{top} & \downarrow \text{there-people} & 1\text{sg} & \text{top} & \text{on-LOC} & \text{come-pst.1sg} \\
\end{array} \]

‘You are from below; I am from above.’
(yon-x-bible, 43008023)

(2) Yonggom: “topic marker” on clausal constituent (conditional) in non-contrastive context

\[ \ldots \text{sibi} \ \text{wedm-an-een} \ \text{kuu} \ \text{kubi} \ \text{darewoob} \ \text{ker-an-een} \ldots \]
\[ \begin{array}{llllll}
\text{sheep} & \text{see-fut-3sg.m} & \text{top} & \text{joy} & \text{big} & \text{become/do-fut-3sg.m} \\
\end{array} \]

‘...if he happens to find the sheep, he will rejoice...’
(yon-x-bible, 40018013)

Readers of earlier versions of this work have dissuaded me from using the term “topic marker”, which is why I am launching the new term “selective” here (somewhat reluctantly: I will retain the conventional gloss \text{top}). The underlying idea is that a crucial – but admittedly vague – function of selectives is to indicate that the main predicate only applies partially – only with respect to the selective-marked constituent – not unrestrictedly.2

Selectivity includes contrast, but is more comprehensive, it also embraces conditional clauses, as in (2). In natural languages, almost all illocutions (assertions, questions and commands) have restricted universes of discourse (Dahl 1975: 50) and selective constituents are one possibility to make restrictions of illocutions explicit. Thus in (1) “below” and “above” only hold for “you” and “me”, respectively, and in (2), “he will rejoice” holds under the condition that “he finds the sheep”. However, selectives are at least loosely associated with the

---

1 Numbers in examples refer to verses in the N.T. “43” is John, the 43rd book of the Bible, and “008023” is Book 8 and Verse 23.

2 This is well in line with Chafe’s (1976: 50) “restricted domain”, Krifka’s (2008) and Krifka and Musan’s (2012: 32) notion of delimitation unifying contrastive topics and frame setting and Constant’s (2014: 48) notion of “partiality” characterizing contrastive topics (see also Büring 2016: 68). If I am coining a new term “selective” here, this is because “restrictive” is preempted for restrictive focus particles such as only, “delimitation” will be used in this paper in a syntactic rather than semantic sense and \text{partiality} mainly means ‘preference, bias’, which is a misleading connotation. Constant (2014: 48) derives “partiality” from the expression “partial answer”. In question-answer sequences, a contrastive topic in the answer signals that the answer only accounts for a selected part of what the question addresses.
realm of what the theoretical literature treats under the headings of “topic” within “information structure”, since they often occur on some sort of “point of departure” (Dooley and Levinsohn 1999: 35) within sentences – if there is such a unit as sentence in spoken language – and they tend to have in common that they “limit the applicability of the main predication to a certain restricted domain” (Chafe 1976: 50). Selectives are closely associated with contrast and contrastive topics, even though selectivity is broader (conditionals are usually not treated in the literature on contrastive topics).

This study will define selectives in terms of an arbitrary conceptual framework for crosslinguistic comparison in the sense of Lazard (2002: 147) rather than Haspelmath’s (2010) debated “comparative concept” (see Dahl 2016; Spike 2020). It is important to emphasize that selectives are not a pre-established formal crosslinguistic category in the sense of Newmeyer (2007). Arbitrary conceptual frameworks are a tool for crosslinguistic investigation and are (i) “the product of the linguist’s free choice”, but (ii) “not haphazardly chosen”, are (iii) “best limited to particular semantic fields” and are (iv) “provisional […] Their value rests in their fruitfulness” (Lazard 2002: 147–148). A difference to Lazard’s specific example – transitivity – is that it is not possible to start with a semantic framework (such as, in Lazard’s study of transitivity, prototypical action), since the prototypical semantics of the elements often called “topic markers” is not sufficiently known and there is reason to assume that these elements also have a number of crucial formal properties that cannot easily be captured in an abstract semantic framework. However, selectives lend themselves to a distributional definitional approach.

I will choose selectives on pronominal constituents in contrast construction as in (1) as definitional context (for the details see Section 3). A first prediction – inspired by Kuno’s (1973: 44) finding that Japanese wa is both “thematic” and contrastive (see also Kim 2015 for Korean -(n)un) – is that the set of markers thus obtained will not be restricted to contrast. A second prediction is that the markers found in the definitional context will also occur on many subordinate clauses, notably on conditionals, as in (2), which assumption is inspired by Haiman (1978), who has argued that conditional clauses, as in (2), are intimately associated with topics. It may be expected, in view of the findings of Haiman (1978), Marchese (1977), Bickel (1993) and others, that selectives are particularly common in conditional clauses as in (2), but selectives can also be found, for instance, on temporal clauses (Matić et al. 2014: 12), and Bittner (2001) has also pointed out the relevance of individual-centered correlative clauses (which, more generally, suggests that headless relative clauses are a relevant domain).

The reason why personal pronouns and not nouns are chosen for definitional purposes is that selectives in many languages have the same form as definite articles, demonstrative pronouns or third person personal pronouns (the close link
to definiteness is emphasized by Haiman 1978). Since these elements do not usually occur on speech-act-participant personal pronouns, first and second person pronouns are important as diagnostic contexts for selectives, as illustrated with (3) from Ewe, where la is both definite article and – as Ameka (1991) puts it – a terminal particle that marks scene-setting constituents.

(3) Ewe

\[
\begin{align*}
\text{Miawo} & \quad \text{la, anyigba sia dzi-e mie-tso,} \\
2\text{pl} & \quad \text{DEF/TOP earth DEM top-EMPH 2pl-come.from} \\
\text{ke} & \quad \text{nye la dzifo-e me-tso} \\
\text{but} & \quad 1\text{sg} \quad \text{DEF/TOP sky-EMPH 1sg-come.from}
\end{align*}
\]

‘You are from below; I am from above.’

(ewe-x-bible, 43008023)

In almost all languages where they occur, selectives are extremely frequent; if selectives are words and not affixes, they are often the most frequent word in a language (Yonggom kuu rank 1, 8.3% of all tokens; Ewe la rank 1, 4.5% of all tokens in the New Testament [N.T.]). If words and morphemes on average have the more meanings the more frequent they are (as already shown by Zipf 1945 for lexical items), it cannot be expected that any marker qualifying as selective only has a single meaning. This is a problem for typological studies and for language-specific descriptive studies alike. Strangely, many language-particular studies take for granted that such markers as Japanese wa or Korean -(n)un might have only a single meaning. If a marker qualifies as selective for the purpose of this paper, this never means that this is the only function of that marker.

Unlike Lazard (2002: 145) and Haspelmath (2010), who emphasize the difference between description and crosslinguistic comparison – because of using reference grammars as major data source and due to their structuralist stance – I do not believe that language description and crosslinguistic comparison should be kept strictly apart. For the purpose of the present study, reference material cannot be relied upon for the simple reason that there is considerable disagreement about the notion of “topic markers” in the literature. This paper takes a “corpus-typological” approach. The major source of data is a massively parallel text – translations of the New Testament (Mayer and Cysouw 2012). A special feature of this paper is that many specific properties from different languages of the sample are discussed and that the data for all glossed examples come from the corpus. The kind of data source used assures us that languages can be compared on the level of specific language use and that largely the same amount of data can be considered per language irrespective of the degree and kind of previous descriptions of a

---

3 For instance, for Yonggom kuu Christensen (2010: 25) suggests a deictic origin.
language (which prevents solipsism, see Newmeyer 2007: 146). The work underlying this article has thus at least some descriptive component for all of the 81 languages sampled (Appendix A). Reference materials, such as grammars and dictionaries, are used as secondary sources only wherever available (Appendix B). The choice of written material also has obvious disadvantages, notably the impossibility to study exclusively prosodic marking strategies that are not distinguished in orthography, the lack of spontaneous interaction, and the impossibility to include sign languages. This entails a restriction to segmental selectives. It is thus important to keep in mind that this study is biased toward written language and segmental markers, biases it shares with many typological studies of grammatical phenomena.

The rest of this paper is structured as follows. Section 2 provides further background about the basic notions of this study: selectives (2.1), contrast (2.2) and subordinate clauses (2.3)–(2.4). Section 3 deals with the method of how a sample of selectives is compiled and introduces the sample. Section 4 presents the results of how selectives in the 81 languages of the sample are distributed across subordinate clause domains. Section 5 discusses aspects of the results from a more general perspective and Section 6 concludes the paper.

2 Background

2.1 Toward a set of hypotheses about the functions of selectives

Authors of grammars and dictionaries make highly different choices as to whether or not calling the elements considered in this article “topic markers”, “topicalizers” or the like, which is due to entirely different traditions in the literature about what topics are considered or not considered to be.

Topic and focus (formerly also termed “theme” and “rheme” or in Czech základ ‘basis’ and ohnisko ‘focal point’; Mathesius 1939) are traditional notions of linguistic information structure, also called functional sentence perspective. Non-topic parts of sentences are called “comment” and non-focus parts “background”. A classical definition of topics is the aboutness-definition: “the speaker announces a topic and then says something about it” (Hockett 1958: 201). Due to the close association of the understanding of subject and topic going back to Aristotle’s hypokeimenon, translated to Latin as subiectum, aboutness is often understood in such a way that the topic is a sort of semantic or psychological subject and occurs as a variable in the valence of the semantic predicate (Jacobs 2001: 647). This makes it difficult to
accommodate adjunct topics, often called “hanging” or “dangling” or “Chinese-style” topics (Chafe 1976; Li and Thompson 1976), which is why Haiman (1978: 585) entirely rejects the aboutness-definition. Haiman (1978) has been very influential in the descriptive tradition but has largely been ignored in theoretical approaches to topics, where various versions of aboutness continue to thrive. Jacobs (2001: 650) interprets aboutness as addresstion (Jacobs 2001: 650) and Reinhart (1982) has suggested that topics function like file cards as units under which information is stored (see also Krifka 2008: 265).

Given that selectives tend to be extremely frequent, we must expect them to have a considerable range of different functions in all or at least most languages where they occur but notice that “different” does not mean “entirely different” here, but rather “similar”. Massive crosslinguistic comparison is useful for investigating which functions are more generally relevant and which are entirely parochial.

I will now introduce some likely candidates for functions of selectives in form of hypotheses that the empirical part of this study can relate to. The hypotheses are summarized in Table 1.

Selectives are likely to have something to do with the flow of information in speech and since the flow of information on the most superficial level of language structure is identical with linear order, a straightforward assumption is that selectives will strongly correlate with word order. Reasonable hypotheses are that selectives tend to occur early in the sentence (Hypothesis 2), but late within the constituents they are associated with (H3) (recall Ameka’s [1991] term “terminal marker” for Ewe).

Table 1: Hypotheses about selectives to be considered in this paper.

| H1   | Selectives tend to be units of the sentence (in the sense of unit with one main clause). |
| H2   | Selectives tend to occur early in the sentence. |
| H3   | Selectives tend to occur after or at the very end of their constituent. |
| H4   | Selectives are associated with (have scope over) a syntactic constituent. |
| H5   | Selectives mark constituents of very different length. |
| H6   | Grammaticalized selectives have a high degree of freedom of host selection. |
| H7   | Freedom of host selection increases when selectives grammaticalize. |
| H8   | Selectives often have the function of grouping (one way of singling out the constituent they are associated with). |
| H9   | Selectives often have the function of delimiting (another way of singling out the constituent they are associated with). |
| H10  | Selectives often have the function of linking their constituent to other parts of the sentence. |
| H11  | Selectives tend not to mark clausal units carrying illocutionary force. |
| H12  | Selectivity is an important semantic function of selectives. |
H2 actually contains an even more basic hypothesis H1 which is that the domain of selectives is not the clause and not discourse units larger than the sentence, but rather the sentence (or, more precisely, a unit containing one main clause with its associated subordinate clauses). This entails that selectives will be largely indifferent to units of the clause such as the grammatical relation subject. Thus, I do not mean the same thing as Hockett (1958: 201) when he speaks of “the topic of the sentence” (this is rather topic of the clause when he discusses subject) – or Maslova and Bernini (2006) when they speak of “sentence topics” in a broader sense than what is meant here. The specific domain sentence is also one reason why “European-style” topic markers such as English as to, concerning, when it comes to, and French en ce qui concerne, quant à are excluded here, because these markers and constructions often have wider discourse scope than just the sentence that contains them.

H3 ("Selectives tend to occur after or at the very end of their constituent") presupposes a connection to some kind of syntactic constituent. Selectives can be expected to be usually associated with a constituent of some kind (H4). If H4 is correct, this means that selectives are of high interest for empirical syntactic studies, since they are a kind of contrast agent for constituency. It follows from H4 and H2 that if selectives do not come particularly early, this can be because they mark long initial constituents, which is most efficiently achieved if the marker is positioned at the end of that constituent (H3). Selectives can have scope over more than just initial words. This is illustrated in (4) from Nalca with two selectives following two different nominalizers that trigger two different genders, the first one at the end of a headless relative clause (double underline), the second one at the end of the complement clause (underline).

(4) Nalca

\[\begin{align*}
\text{Nimi bleb-si-na’} & \quad \text{dub-nya,} \\
\text{men save-0.3PL-VN above-NMLZ1} & \quad \text{si-lim-ak-nya} & \quad \text{be-ra,} \\
\text{call.name-IPV-PRS.3PL-NMLZ2} & \quad \text{M-TOP} \\
\text{ella u-lum-na.} & \\
\text{si-lim-ak-nya} & \quad \text{be-ra,} \\
\text{call.name-IPV-PRS.3PL-NMLZ2} & \quad \text{M-TOP} \\
\text{ella u-lum-na.}
\end{align*}\]

(I know that Messiah is coming (the one called Christ).’ lit. “That the men savior chief, the one who is called Christ, is coming knowing I am.”

(name-x-bible, 43004025; Sap5)

It can further be expected that selectives mark constituents of very different length (H5) – compare Examples (1) and (4) – which will entail a high degree of psycholinguistic complexity.
If selectives, as this study intends to show, occur on constituents or “hosts” of very different kinds, ranging from noun or determiner phrases including pronouns to subordinate clauses and including arguments and adjuncts, it can be expected that selectives will have a high degree of freedom of host selection (H6), which is a typical property of clitics. Selectives can hence be expected to exhibit clitic-like behavior (but not necessarily phonologically). Since high degree of freedom of host selection is rather rare, it can be expected that the increase of freedom of host selection will characterize the grammaticalization processes of selectives whatever their sources (H7).

Marking constituency entails that selectives will have such effects as grouping (H8) (indicating which words form a constituent together), delimiting (H9) (indicating where a constituent ends or starts) and linking (H10) (indicating that something else is expected to follow). Linking is related to the notion of projection in interactional linguistics (see Schegloff 1996, and Tanaka 2015 for Japanese wa), but – due to H1 – largely restricted to the realm of the sentence. Delimiting and linking may seem to be contradictory at first glance but are in fact not. If there is nothing that follows there is no need for delimiting what precedes.

The hypotheses above are very general and highly underspecifying. They apply to many more environments than those where selectives are actually attested. It must therefore be expected that the use of selectives is constrained semantically or pragmatically. This article has no ambition to flesh out a comprehensive semantics of selectives, which is actually impossible given its focus on subordinate clause domains. The expected outcome of the empirical study is that different types of subordinate clauses differ in their propensity to take selectives will result in some sort of hierarchy that can be interpreted semantically (scales are simple semantic maps, see Croft 2003, Ch. 5).

Selectives are usually absent from units that bear illocutionary force (assertion, interrogation, command) (H11). This restriction is most important for clauses. Many complement clauses, for instance, have illocutionary force of their own, so complement clauses may be expected to be less prone to display selectives than other kinds of subordinate clauses. (However, as will be discussed in Section 3, selectives can anticipate the illocution associated with the main predicate – question or command – which testifies to the linking function of selectives.)

As already suggested in the introduction, selectives are expected to express selectivity (H12). As far as subordinate clauses are concerned, selectivity is akin to condition rather than concession. The apodosis of a conditional constructions applies only partially: under the condition that the protasis is true. The protasis has thus a selective function. Concessive clauses, however, do not restrict the occurrence of events reported in main clauses. Hence, if “selectives” express selectivity (H12) they can be expected to occur more often with conditional than with concessive clauses.
2.2 Contrast

In the vast literature on contrast, a frequent distinction made is the one between (i) semantic opposition (John is tall, but Bill is short) and (ii) denial of expectation (John is tall, but he’s no good at basketball) (Lakoff 1971: 133). The examples also differ between “double contrast” (two pairs of alternatives) and “single contrast” (one set of alternatives; Umbach 2005: 218). Unified approaches to (i) and (ii) include those of Umbach (2005), who suggests the confirm+deny condition (one part of the contrast is explicitly or implicitly affirmed and one part is explicitly or implicitly denied) and Winterstein (2012: 1883), who argues that two elements are contrastive if they are counter-orientated toward an argumentative goal. For our purposes, the distinction between semantic opposition and denial of expectation is relevant, because pairs of selective-marked constituents in contrast sequences can be semantically opposed, as in (1) “You are from below; I am from above”, but need not, as in (5), where it is not “my house” and “you” that are contrasted.

(5) Dimasa

\[ A-ni \text{ noh de} \text{ bi-ma-ni noh se thi-hi reb du.} \]
\[ Thikhabo, nisi la bo khe makhao-rao ni thinkhoh slam kha. \]

“It is written, ‘And my house will be a house of prayer,’ but you have made it a cave of robbers!”

(dis-x-bible, 42019046)

In (5), the selective constituents (underlined) are semantically independent of each other. The first one introduces the topic of the whole contrast sequence, the second one narrows down the range of application of the counter-orientated statement.5

Mann and Thompson’s (1988) Rhetorical Structure Theory operates with the two units “nucleus”6 (roughly corresponds to a semantically non-hypotactic clause) and “satellite” (a semantically hypotactic clause). According to Mann and Thompson (1988: 248), the contrast relation has two nuclei; it is thus not a hypotactic relation. For our purpose it is important to notice that the two nuclei can have different properties, which is why it is useful to label them distinctively. I will call

---

4 Denial of expectation is important for Verhagen’s (2007: 44) intersubjective approach, according to which the function of contrastive connectives, such as English but, is to cancel inferences.

5 It is important to emphasize that selective constituents are never implicit, unlike Sæbø’s (2003: 269) implicit topics adduced in a unified approach to contrast in terms of information structure.

6 Note that “nucleus” in Rhetorical Structure Theory is entirely different from “nucleus” in Role and Reference Grammar (Van Valin 2005).
the first contrast nucleus “non-adversative” and the second one “adversative”; thus, “contrastive” [both contrast nuclei] and “adversative” [the second contrast nucleus] do not mean the same thing. The distinction allows us to discuss such questions as to whether Dimasa distinguishes non-adversative (de) and adversative (la) selectives, as Example (5) suggests (Jacquesson 2008 speaks of non-adversative and adversative topic markers). Example (6) illustrates that the notions non-adversative and adversative apply to clausal points of departure in the same way as to nominal ones.

(6) Dimasa

```
Jodi ang thi-ba gajai kha thi-kha-de,
if 1SG speak-VN sin PRF say-PRF-TOP
bo gajai ni basao ha thi-phin ri.
that sin GEN body/above LOC/ALL say-put.on give
Thikhabo, ang gibi khe se jo ba thi-kha-la,
but 1SG real/true ACC one speak PST say-PRF-ADS.TOP
snadi ja-ba-ni ning ang khe do ba?
what become-VN-GEN 2SG 1SG ACC beat PST
```

‘Jesus replied to him, “If I have spoken wrongly, testify about what is wrong! But if I have spoken correctly, why do you strike me?”’

(dis-x-bible, 43018023)

A further frequently posited use of but and other adversative connectives is concessive. According to Hermodsson (1978: 59–60) and König (1991), concessives are negatively causal or “incausal”. Umbach (2005: 227) argues that “there is no ‘concessive but’ just as there is no ‘causal and.’ The meaning of but does not include incausality, just as the meaning of and does not include causality.” This is a problem for Rhetorical Structure Theory where a discourse relation cannot be contrast (two nuclei) and concessive (satellite and nucleus) at the same time. As we will see, it is very common across the languages of the sample to use adversative connectives as translation equivalent for European-style concessive subordinators. The concessive use of adversative connectives, however it should be interpreted, is thus highly relevant for the study of interaction of selectives and contrast across subordinate clause domains.

Many languages have several different adversative connectives, but according to Winterstein (2012: 1866) this is not an argument against a uniform analysis of English but; rather other languages “sport connectives with meanings that are more specified”. Contrast is so far rather poorly investigated crosslinguistically,

---

7 It might be more appropriate to gloss thikha as a grammaticalized connective stem, it also recurs in thikabo ‘but’. 
hence general statements seem to be premature. However, there does not seem to be any clear borderline between contrast and lack of contrast. For instance, in the N.T. but is often used in an inflationary manner at the beginning of new episodes (But when Herod was dead… 40002019 eng-x-bible-kingjames vs. Now after Herod had died eng-x-bible-lexham) which is probably due to translation effects from Koine Greek de and Latin autem (Defuncto autem Herode…). In fact, Classical Greek de has similar functions as selectives. However, as will be argued in Section 3, what makes it differ from selectives is that it does not occur in the non-adversative part of contrast sequences. It is similar to the adversative selective la in Dimasa in (5) and (6), but there is no non-adversative selective in Koine Greek.

2.3 Subordination

There are many conceivable types of subordinate clauses. Kortmann (1997: 81) distinguishes 32 interclausal relations only for adverbial clauses. No typological investigation of subordinate clauses is comprehensive, and this study is no exception. Since we know at least since Haiman’s (1978) seminal article that selectives are expected in conditionals, this study focuses first of all on various kinds of conditional clauses. Table 2 lists the types of subordinate clauses considered in this paper – the order anticipates the one of Figure 1 in Section 4. The full list of clauses is given in Appendix E.

By subordinate clauses of a certain type I mean constructions that display translation-equivalence with such clause types irrespective of their syntactic status. As far as adverbial clauses are concerned, this is well in line with Matthiessen and Thompson’s (1988: 301) hypothesis that “enhancing hypotactic clause combining” [which is their term for adverbial subordination] “has evolved as a grammaticalization of rhetorical relations in text of the enhancing Nucleus-Satellite kind”. In many languages, constructions are used that exhibit the hallmarks of what is called “cosubordination” in Role and Reference Grammar: shared operators and switch reference (Van Valin 2005: 187). Coordination cannot be excluded either. In Cucatec and Sierra de Juárez Zapotec, the selectives have the same form as the coordinator ‘and’.

My approach to subordination is well in line with the one adopted by Cristofaro (2003: 34–48), who has launched the “Asymmetry Assumption” according to which the state of affairs expressed by the subordinate clause is pragmatically not-asserted irrespective of the structural properties of any clause type. Non-assertion of subordinate clauses also fits H11 “Selectives tend not to mark clausal units carrying illocutionary force”. Note also that according to Maslova and Bernini (2006: 68) topics are kept outside the scope of assertion.
Table 2: Subordinate clause types considered in this paper.

<table>
<thead>
<tr>
<th>Label</th>
<th>Clause type</th>
<th>Kortmann’s (1997) interclausal relations (only adverbial clauses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GenRel</td>
<td>General relative clauses</td>
<td>[Relative clause]</td>
</tr>
<tr>
<td>CondAny</td>
<td>Conditional clauses with an indefinite pronoun subject (“if anybody...”)</td>
<td>Condition ‘if’ (COND)</td>
</tr>
<tr>
<td>CondHyp</td>
<td>Hypothetical conditional clauses</td>
<td>Condition ‘if’ (COND)</td>
</tr>
<tr>
<td>Cond</td>
<td>(Standard) conditional clauses</td>
<td>Condition ‘if’ (COND)</td>
</tr>
<tr>
<td>CondFin</td>
<td>Postposed conditional clauses</td>
<td>Condition ‘if’ (COND)</td>
</tr>
<tr>
<td>CondNeg</td>
<td>Negative conditional (“unless”) clauses</td>
<td>Negative Condition ‘unless’ (N_COND)</td>
</tr>
<tr>
<td>GenTime</td>
<td>Generalizing temporal clauses (“whenever...”)</td>
<td>Contingency ‘whenever’ (CONTIN)</td>
</tr>
<tr>
<td>Anterior</td>
<td>Anterior clauses (“after...”)</td>
<td>Anteriority ‘after’ (ANTE)</td>
</tr>
<tr>
<td>Sequent</td>
<td>Sequential clauses (“when...”)</td>
<td>Simultaneity Overlap ‘when’ (SIOVER)</td>
</tr>
<tr>
<td>Simult</td>
<td>Simultaneous clauses (“while...”)</td>
<td>Simultaneity Duration ‘while’ (SIDUR) and Immediate Anteriority ‘as soon as’ (IMANTE)</td>
</tr>
<tr>
<td>Termin</td>
<td>Terminal clauses (“until...”)</td>
<td>Terminus ad quem ‘until’ (TAQUEM)</td>
</tr>
<tr>
<td>Post</td>
<td>Posterior clauses (“before...”)</td>
<td>Posteriority ‘before’ (POST)</td>
</tr>
<tr>
<td>ConCond</td>
<td>(Scalar) concessive conditional clauses (“even if...”)</td>
<td>Concessive Condition ‘even if’ (COCOND)</td>
</tr>
<tr>
<td>Concess</td>
<td>Concessive clauses</td>
<td>Concession ‘although’ (CONC)</td>
</tr>
<tr>
<td>Sap</td>
<td>Complement of verbs of knowledge clauses</td>
<td>[Complement clause]</td>
</tr>
<tr>
<td>Perc</td>
<td>Complement of perception verb clauses</td>
<td>[Complement clause]</td>
</tr>
<tr>
<td>Dic</td>
<td>Complement of verbs of saying clauses</td>
<td>[Complement clause]</td>
</tr>
<tr>
<td>Cause</td>
<td>Causal clauses</td>
<td>Cause/Reason ‘because’ (CAUSE)</td>
</tr>
<tr>
<td>Purpose</td>
<td>Purpose (“because...”) and result (“so that...”) clauses</td>
<td>Purpose ‘in order that’ (PURPOSE) and Result ‘so that’ (RESULT)</td>
</tr>
</tbody>
</table>

Functional equivalence of subordinate clauses across languages makes Cristofaro (2003) rely on translation, which is well in line with the parallel text approach used here. However, Reesink (2014) has argued against Cristofaro that idiomatic translation cannot be relied upon. I agree with Reesink (2014: 259) that “different translation possibilities for various examples” is a relevant issue, but it can be addressed by using the cumulative evidence of multiple examples rather than just isolated ones. Hetterle (2015: 36) argues that Cristofaro underestimates the continuous character of the subordination–coordination distinction. For practical reasons, I use “subordination” as a convenient cover term for the clause types that are traditionally subsumed under subordination. This does not necessarily imply syntactic subordination or a cognitive construal of non-assertion.
Finally, I fully endorse Hetterle’s (2015: 146) point that linguistic variation in clause-linkage phenomena should be investigated in a bottom-up way, this is exactly what the parallel text approach is good for.

2.4 Conditionals and other types of subordinate clauses

Conditionals (consequence relations) differ from many other interclausal relations in that they are not veridical (Asher and Lascarides 2003: 157). Veridical means ‘and’-like in the sense that the relation is true if and only if both members of the relation are true. Concessives, most temporal clause constructions, purpose and result and cause (explanation) are all veridical, which makes them more amenable to coordination-like construals than conditionals. Non-hypotactic construal of conditionals is possible, however, if the protasis is framed as a question (which also entails non-veridicality).

2.4.1 Conditional and concessive as different reasoning connectives

A conditional is a reasoning bifurcation. In standard conditionals (reflecting material implication \( p \rightarrow q \) with truth table TFTT), the alternative path of reasoning not pursued (the “else”-path) usually remains unexpressed and undetermined. In this sense conditionals entail selectivity of reasoning.\(^8\) An if-and-only-if conditional (biconditional \( p \leftrightarrow q \) with truth table TFFT) is also a reasoning bifurcation, but with the alternative path being determined (albeit not expressed explicitly) in form of the opposite of the reasoning path chosen. Concessives and conditional concessives are also reasoning connections, but they are pseudo-bifurcations: whatever path of reasoning pursued, the outcome is the same. The state of affairs in the concessive clause is irrelevant for the outcome, hence there is a lack of selectivity in concessives. If “selectives” express selectivity, it is therefore expected

\(^8\) Note also Lehmann’s (1974) approach to conditionals, reported in Comrie (1986: 86) “Lehmann notes that in any discourse it is necessary for the participants to gain common ground step by step. In this process of establishing common ground, a conditional protasis represents progress in its establishment in a disjunctive situation: there are two possibilities (namely, \( p \) and \( \sim p \)), and before communication can progress, it is necessary for the speaker to establish which of the disjuncts is to be considered; only then can the argumentation proceed. From this perspective, the linear order of the clauses is iconic to the sequence of steps in the argumentation.” Postposed conditionals, however, often have a character of hedges, such as And this we will do, if God permits (CondNeg9) and do not contribute to gaining common ground step-by-step in the same way.
that concessives lack selectives, perhaps with concessive conditionals being intermediate, since they, as demonstrated by Haspelmath and König (1998), have both properties of concessives and conditionals.

If-and-only-if conditionals are unfortunately poorly represented in the parallel text considered in this study. What comes closest to them is negative conditionals (“unless”-clauses) which can be viewed as biconditionals by implicature. Only if \( p \), \( q \) is a natural paraphrase of unless \( p \), not \( q \), as illustrated in (7) from Yonggom. [A natural paraphrase is a paraphrase that occurs as translation-equivalent in at least some languages of the world.]

\[(7)\]
Yonggom

\[
\begin{array}{llllllllllllllll}
\text{Karub} & \text{kuu} & \text{Ewen} & \text{ari-mbed} & \text{komo} & \text{kon-oon} \\
\text{man} & \text{TOP} & \text{heaven} & \text{up/top-LOC} & \text{what} & \text{give.to.him-PST.3M} \\
\text{ku-mbed-mo} & \text{awin-e} & \text{kand-an-een} \\
\text{TOP-LOC/ERG-ONLY} & \text{grab-SS} & \text{get-FUT-3M} \\
\end{array}
\]

‘A man can receive not one thing unless it is granted to him from heaven!’

lit. “Man will receive only if something is given to him from heaven.”

(yon-x-bible, 43003027; CondNeg9)

For information structure it may be relevant that ‘only’ (in biconditionals) and ‘even’ (in scalar conditional concessives) are known as restrictive and additive focus particles (König 2000) and, to the extent that focus is incompatible with topic, this may in some languages (which does not actually hold for Yonggom in (7)) be a reason for not using selectives in biconditionals and in conditional concessives. In Imbabura Quichua, for instance, the selective -ka -ca is in complementary distribution with the additive suffix -pash ‘also, even’ (typically following concessive and conditional concessive clauses) and with the limitative suffix -lla ‘only’. Similarly, in Awa-Cuaiquer, the selective -na cannot be combined with the additive marker -kas (used in concessive and conditional concessive clauses) or with the restrictive marker -miŋ (Curnow 1997: 358, 373).

Mann and Thompson (1988) show that the concessive relation is often expressed with adversative coordination, such as English \( S[\text{atellite}], \text{but } N[\text{ucleus}] \). \( S, \text{but } N \) is a natural paraphrase of although \( S, N \) (but only if the concessive clause is preposed, which is the preferred order for concessives according to Verhagen 2007: 170). In this study, only concessive examples where English uses subordination are sampled, but, as we will see, many languages use adversative coordination more widely than Standard Average European languages in concessive relations. Concessive relations are closely related to contrast, since in concessives, two propositions are asserted “against the background of an assumption” (König 1991; Verhagen 2007: 167).
2.4.2 Temporal and conditional clauses

Many languages use the same connective for conditional ‘if’ and temporal ‘when’ (such as German wenn; Auer 2000: 175). However, most temporal clauses are factual, i.e., the relation is veridical. Non-hypothetical conditional clauses are semantically closer to temporal clauses than hypothetical and especially counterfactual ones. Generalizing temporal clauses (‘whenever’, contingency, also called “Indefinite Time”; Kortmann 1997: 85) are intermediate between conditional and temporal clauses, they both express habituality (time) and a general condition (see also Nau 2018: 72). Temporal clauses are similar in function to temporal adverbs, which are often followed by selectives. In posterior clauses, the relationship is inverted, but many languages use negative construals in ‘before’ clauses, *time not yet p, q* is a natural paraphrase of *before p, q* (Hetterle 2015: 136).

2.4.3 General relative clauses and conditionals

In many languages, translation-equivalents of headless relative clauses have light head nouns, such as ‘person’, ‘people’, ‘man’, which is why I use the term “general relative clause” as a cover term for headless relative clauses and relative clauses with light head nouns, such as (8):

(8) Una

\[
\begin{array}{llllll}
\text{Sun} & \text{nun-ti ‘Nun li,’} & \text{tene-si-n} & \text{kun nang ara} \\
3\text{PL}/2\text{PL} & 1\text{PL-for} & 1\text{PL} & \text{not.like/not.want think-O.1\text{PL-VN} NEG people top} \\
\text{nun} & \text{kwit-yabwe} & \text{kun-ung…} \\
1\text{PL} & \text{friend-PL} & \text{be-CONT.3\text{PL}} \\
\end{array}
\]

‘For whoever is not against us is for us.’, lit. “People who do not think about us ‘We don’t like (them)’ are our friends.”

(mtg-x-bible, 41009040; GenRel4)

To the extent that such constructions have nominal heads, it is expected that selectives are more likely to occur in general relative clauses in languages where verbs are less inclined to bear selectives, which is, for instance, the case in Una. General relative clauses are akin to correlative clauses, which are often topical (Lipták 2009: 12). Bittner (2001) has pointed out a parallelism between correlative constructions and conditionals. Conditional clauses with indefinite pronoun subjects (“if anybody”) are intermediate between conditional clauses and general relative clauses and can adopt a general relative construal as in (9) from Yuracaré, where general relative clauses are formed with compounds with the noun *(j)bëshë ‘thing, entity’ (van Gijn 2006: 68):
Conversely, general relative clauses often have conditional construal “the one who” = “if anybody”.

Even though relative clauses are often embedded (Mathiessen and Thompson 1988: 282), general relative clauses tend to behave rather like enhancing hypotactic clauses. They share with restrictive relative clauses the property of some sort of co-reference across the clauses, which is usually individual-centered. However, most kinds of adverbial clauses are also amenable to correlative-like constructions with a resumptive element in the main clause picking up the subordinate clause event anaphorically.

2.4.4 Complement clauses and conditional clauses

On the one hand, complement clauses are quite different from adverbial clauses and this might be an argument to entirely disregard complement clauses here. According to Verhagen (2007: 96–97), complement clauses contribute to the main line of discourse whereas their matrix clauses mainly deal with intersubjective conceptualization, which is not the case for adverbial clauses. However, there are also clear similarities between at least some kinds of adverbial clauses and some kinds of complement clauses. For instance, it is well-known that many languages encode ‘if’ and ‘whether’ in similar ways, and conditionals are particularly important in this study. This is why complement clauses are included here.

Complement clauses can be subclassified according to the type of main clause verb, as opted for in Table 2 – complements of verbs of saying, perception or cognition (but differences can be gradual, see Horie 1993) – or according to the illocution of the complement (assertion or question). The translation-based approach used here implies that complement clauses with speech verbs will not be strictly separated from constructions with direct speech or construals intermediate between direct and indirect speech.

Complementizers often grammaticalize from demonstrative pronouns (as in English that) which are erstwhile arguments of main clauses. Selectives can then occur either on the complement clause or, as in (10) from Hills Karbi, on the anaphoric demonstrative that picks up the complement clause. While (10)
is an isolated example in Hills Karbi and does not reflect the most common construction to form indirect questions in that language, it illustrates how selectives occasionally can be associated with anaphors picking up complement clauses.

(10) Hills Karbi

\[
\text{A}l\text{a}n\text{-}l\text{i ke } i\text{-}\text{n}u\text{t } a\text{-papi } l\text{o-ne, } m\text{ate kali } l\text{o-ne,}
\]

3-HON TOP one-CL.HUM.SG POSS-impiety FOC-INDF or NEG.EQU.COP FOC-INDF

\[
l\text{a-k}\text{e } n\text{-l}i \quad c\text{hini-}n\text{e.}
\]

this-TOP 1EXCL-HON know-NEG

‘Whether he is a sinner I do not know.’ (literally “...that I do not know”)

(mjw-x-bible, 43009025; Sap1)

Since complement clauses can have illocutionary force, they are in conflict with the lack of illocutionary force hypothesis of selectives (H11). However, anaphors picking up complement clauses lack illocutionary force.

Purpose/result and causal clauses do not entertain close relationships with conditional clauses and are mostly included as control groups where selectives are not expected. Purpose and result clauses also have a tendency to be postposed in many languages, which is expected to be unfavorable for selectives.

3 Compiling a sample of selectives and assessing the range of diversity of the set

I define selectives here on the basis of how they are used in particular contexts. Thus, being used as a selective and being a selective is the same thing for the purpose of this study. Notably, being a selective does not mean being a selective only. Elements sampled can also be demonstratives or determiners (such as Ewe la DEF), temporal adverbs (such as Sougb kaba ‘then’) or case forms (Borong -noŋ LOC) at the same time. It follows from this that the class of elements sampled exhibit a considerable amount of heterogeneity.

For the reasons mentioned in Section 1, the starting point are contrastive uses of personal pronouns as in examples such as (1) “You are from below; I am from above” (43008023). More specifically, I have concentrated on occurrences of the second person plural. The process of identifying second person plural forms plus selective (listed in Appendix C) is iterative in a similar way as the procedure applied in Dahl and Wälchli (2016) for the identification of perfect and iamitive grams – the major difference being that the selection of markers had to be done manually because selectives cannot be identified by simply searching by
automated means for parallel markers with similar distributions in other languages. Appendix D lists the top ten verses with selectives on second person plural pronouns. A selective occurs in at least one of the two top examples in all 81 languages of the sample and in all languages of the samples there is a selective in at least two of the top 10 contexts. As can be seen from the examples in Appendix D, not all of them are equally contrastive.

Selectives are difficult to distinguish from a whole range of other phenomena types (Table 3). The most characteristic context of use of selectives are copular sentences in contrast sequences, where they can be confounded, among other things, with definite articles or demonstratives, copulas, subject markers, inversion markers and adversative connectives.

In languages where selectives have specific allomorphs when associated with personal pronouns, other allomorphs were also included. For instance, Botolan Sambal has hi(-) for pronouns and person names and hay_ for common nouns. Konai has a pronominal -me and a non-pronominal -be variant of the selective.

Verb-initial languages sometimes have inversion markers indicating that some other constituent precedes the verb. Inversion markers, such as Tagalog ay, are thus functionally very close to selectives. However, if more than one constituent precedes the verb, as in (11), the inversion marker goes together with the verb. (12) illustrates two preverbal constituents in Sierra de Juárez Zapotec, where both have a following selective of their own.

(11) Tagalog

\[\text{Nguni’it kayo, mga kapatid, ay wala sa kadiliman…}\]

but \(2\text{PL.PIV PL brother inv NEX OBL}\) darkness

‘But you, brothers, are not in the darkness…’

(tgl-x-bible-1905, 52005004)

(12) Sierra de Juárez Zapotec

\[\text{Attu te cq nna, para guni bá cq ne prueba nna,}\]

\[\text{other PL 3PL.OBJ TOP for FUT.make still 3PL 3SG.OBJ test TOP}\]

\[\text{gunàba cq qui’ni guni-e ttu milagro de yiabara’.}\]

\[\text{PST.demand 3PL that/because FUT.make-3SG.RESP one miracle of sky}\]

‘And others, in order to test him, were demanding from him a sign from heaven.’

(zaa-x-bible, 42011016; Purpose2)

As pointed out in 2.2, contrast sequences have a non-adversative and an adversative nucleus, and languages can have different non-adversative and adversative selectives (in 2.2, Dimasa was discussed as a potential case in point). It is important to notice here that “adversative selective” is a secondary (or conditioned) arbitrary
conceptual framework: a language can have an adversative selective only if there is also another (more general in the sense of more frequent) selective that can occur in non-adversative contexts. If a language has a frequent marker in adversative nuclei, such as Koine Greek *de*, and occasionally a less frequent different parallel marker in non-adversative nuclei, such as Koine Greek *men* (as in *as in ho men...ho de* ‘the one...the other’), these do not qualify as selectives. For the same reason, Guaraní *katu* is excluded (against Tonhauser 2012 and Büring 2016: 69).

However, delimitation is not always straightforward since a few languages have selectives with a preference for adversative contexts, such as Sochiapan Chinantec. In the two Mongolian languages of the sample, Khalkha *bol* and Kalmyk *bolxla* are mostly restricted to adversative contexts. Khalkha and Kalmyk are included here in the sample mostly because they share a specific grammaticalization path \textit{CONDITIONAL} \textgreater \textit{(ADVERSATIVE) SELECTIVE} (see, 4.3.2), which is highly relevant for this study. Khalkha and Kalmyk also have another selective \textit{n'}, deriving

\begin{table}[h]
\centering
\begin{tabular}{|l|l|l|}
\hline
\textbf{Type of phenomenon} & \textbf{Reason for exclusion} & \textbf{Diagnostics for selective excluding that type} \\
\hline
Definite article only & \textit{TOP} has higher degree of freedom of host selection & \textit{TOP} also occurs with personal pronouns of 1st/2nd persons \\
Emphasis on personal pronouns & \textit{TOP} has higher degree of freedom of host selection & \textit{TOP} does not occur only with personal pronouns \\
Subject or other relational role (S, A, P, or combinations thereof) & Selectives are not strictly limited to one or some grammatical relations & \textit{TOP} can usually also occur with adjuncts \\
Copula & \textit{TOP} is not part of the predicate and is not dedicated to a particular type of predicate & \\
Inversion marker in verb-initial languages & \textit{TOP} is independent of finite verbs & If two selective constituents precede the verb, both can have \textit{TOP}. \\
European-style “topic markers” such as \textit{As far as X is concerned} & European-style “topic markers” have a different distribution in texts and are much less frequent. & European-style “topic markers” do not occur in the definitional context. \\
Adversative connectives “but” & \textit{TOP} is not restricted to contrast (even though contrast is a favorable context) & \textit{TOP} can occur in the non-adversative part of the contrast sequence and if there is an adversative selective, there is also a more frequent non-adversative one. \\
\hline
\end{tabular}
\caption{Excluded phenomena (no selectives).}
\end{table}
from the third person possessive, which, however, is not attested with personal pronouns in the N.T.

So far I have identified selectives in 81 languages represented in the N.T. corpus. Since the procedure applied is manual, I will certainly have missed a considerable number of languages with selectives in the Bible corpus. With the exception of the many Quechua languages, most of which seem to have some sort of selectives, no restrictive sampling procedure was applied, but only four Quechuan languages were included. When using electronically available parallel texts, there is no need to strive for genealogically and areally unbiased samples, since structural similarities in genealogically and areally related languages will be visible in the results. Generally, the sample is a heuristic sample, which has a strong areal bias, reflecting a tendency of selectives to exhibit areal patterns (for Eurasia, see Yurayong 2018).

The sample is listed in Appendix A and displayed on four maps in Appendix A2. The markers detected are listed in Appendix B. Appendix C lists the strings of second person plural form with selectives in the corpus and Appendix D lists the 10 top contexts for 2PL (a selective occurs in either or both of the two top examples in all languages sampled). There is no single context in the N.T. where all languages of the sample have selectives. The top context, illustrated from Yonggom (1) above, has selectives in 71 of 81 languages (88%). Despite their high text frequency, selectives tend to be communicatively optional in the sense that there are few contexts where they always can be expected (and this is a difference to grammatical gram types such as past or plural). Cheung (2007: 35) notes for Jingpo that “the presence of the topic marker go is always optional.” One reason is that selectives are redundant in their most characteristic environment: contrast. Contrast can efficiently be expressed without selectives with an adversative connective or just by two opposed predicates.

Many languages in the sample have more than one selective. In Motu, the three particles na, be and ese in Example (13), all meet the definitional criteria for selectives.

(13) Motu

Umui ese tama-mui ena kara o kara-mu. Idia eto
2PL top father-POSS.2PL POSS custom/habit 2 do-PRS 3PL be.about.to.3
Ai na dia ariara natu-dia;
1PL.EXCL top NEG street child-POSS.3PL
ai Tama-mai be tamona, Dirava bunai.
1PL.EXCL father-POSS.1PL.EXCL top one God there
“You are doing the deeds of your father!” Then they said to him, “We were not born from sexual immorality! We have one father, God!”
(meu-x-bible; 43008041)
Ese is also an optional ergative marker (Yam 2020: 54). According to Yam (2020), na and be differ in engagement (Evans et al. 2018a, 2018b), the relative accessibility of a state of affairs to speaker and/or addressee. In roughly half of the languages of the sample (see Appendix B), selectives also express some other grammatical feature. The markers found in the N.T. are not always the same as in the reference sources. Only one of three markers described as topic marker in Farr (1999) for Korafe-Yegha meets the selective definitional criteria. According to Matisoff (2006: 265), Lahu has as many as six topicals, one of them being qo <k’o> ‘if; when’. However, according to the definition adopted here, only lē <leh_2> meets the definition of selective. According to Rigden (1985), topic markers in Karkar-Yuri include the expression of local position and direction, but only one combination, t-i TOP-POSITION.HERE.AND.ELSEWHERE, occurs in the N.T. For Zaiwa, Lustig (2010: 253) discusses three variants <ge_i ~ ga_i1 ~ ge_i11> reflecting different kinds of attitude not distinguished in the orthography of the N.T. However, the Zaiwa N.T. reveals different interrogative and adversative selectives, which I cannot find in the grammar. Here, using parallel texts as source of data has the advantage that exactly the same criteria can be applied to all languages considered. However, taking into account what the reference materials report, it must be concluded that this study based on written texts will underrate the amount of possible distinctions.

The bewildering variety of distinctions can be roughly sorted into three types. The additional distinction or distinctions can reflect (i) a feature of the selective constituent (such as gender, number, case, deixis or part of speech of the selective constituent), (ii) a feature of the nucleus (such as illocution or counter-expectation) or (iii) a feature of the larger construction (notably non-adversative vs. adversative selective).

(i) **Features of the selective constituent**: Features of selective constituents reflected in selectives can be highly language specific. Amanab -ba TOP has an allomorph -bi if the clausal constituent preceding it is transitive and has a plural subject (i.e., triggered by ergative case role and plural number), which reflects a kind of morphological assimilation. Number of subjects in transitive stems in Amanab is expressed by vowel raising (a/e;i; Minch 1992: 107), which spreads to the suffix to the right. In (14), there is in addition the selective eba, which is also a demonstrative.

(14) Amanab

\[
\text{Afa ne ka-m numwehe anwena-fi-g.bi eba ka-na then 2 1SG-DAT ??? know-do.PL-PST-TOP.PL DEM/TOP 1-GEN}
\]
\[
\text{Alagi-m gafa ne ehe-m anwena-fi-fi father-DAT also 2 3-DAT know-do.PL-FUT.PL}
\]

‘If you have known me, you will know my Father also. And from now on you know him and have seen him.’
(amm-n-x-bible, 43014007; CondHyp8)
(ii) *Features of the nucleus*: Selectives do not express illocution, but they can anticipate illocutionary distinctions in the nucleus (question or command), which testifies to the linking function of selectives. In the sample, Ewe, Parkwa, Zaiwa and Ama, languages from four different families and three continents, have interrogative selectives (the sentence that contains the selective is a question), and one language, Fon, uses a particle /hùn/ ‘therefore, thus, certainly’, as an imperative selective (Lefebvre and Brousseau 2002: 177). As illustrated in (15), it is also used for deontic modality.

(15) Fon

\[
\begin{align*}
\text{É(-)nyì méde} & \quad \text{dqò mi me, bò dqò awu we} \\
\text{hùn,} & \quad \text{ni ná dqò-kpó me e ma dqò dqè} \\
\text{therefore} & \quad \text{SBJV give be.at-together person 3SG,OBJ NEG have one} \\
\text{NEG 3SG} & \quad \text{á é.}
\end{align*}
\]

‘The one who has two tunics must share with the one who does not have one…’

(fon-x-bible, 42003011; GenRel16)

Selectives can also anticipate engagement. Sochiapan Chinantec has a counter-expective selective má’nà² indicating “that the comment will contain information that is in some way counter to the known or assumed expectation of the addressee” (Foris 2000: 389). Note that engagement can be both a feature of the selective constituent (Haka Chin⁹) and a feature of the nucleus (Sochiapan Chinantec).

In Lambayeque Quechua and in Cajamarca Quechua, but not in Huallaga Quechua (Weber 1989: 396) and Imbabura Quichua, evidential markers can be suffixed directly to the selective. In (16) from Cajamarca Quechua, ch’i ‘hypothetical’ is one of the sets of evidential markers, which are usually connected with assertion.

(16) Cajamarca Quechua

\[
\begin{align*}
\text{Dyus-manda mana shamù-shpa-ga-chi, mana nima milagru-ta} \\
\text{God-ABL NEG come-SUB-TOP-HYP NEG nothing miracle-OBJ} \\
\text{atinanda-chu rura-y-ta,…} \\
\text{be.able-Q do-INF-OBJ}
\end{align*}
\]

‘If (this man) were not from God, he would not be able to do anything!’

(qvc-x-bible, 43009033; CondHyp2)

---

⁹ Selectives in Haka Chin are postposed demonstratives. The demonstrative kha (obl. khan), as opposed to cu (obl. cun), is used when the listener has personal knowledge about the referent (Barnes 1998: 60).
Features of the larger construction: Among all distinctions in selectives, adversative selectives tend to be most underdescribed. As exemplified in (17), Huallaga Quechua uses the attitudinal adverb ichan ‘certainly, probably, perhaps’ plus selective /-qa/ (Weber 1989: 66) as an adversative selective. (According to Weber 1989: 380, “The syntax of ichan-qa ‘perhaps’ […] is somewhat of a mystery” in an example that clearly reflects an adversative selective context.)

(17) Huallaga Huánuco Quechua

\begin{verbatim}
Mana aywa-r-ga Yanapăcoj Espíritu Santu-ta mana-mi
\end{verbatim}

\begin{verbatim}
\end{verbatim}

\begin{verbatim}
send-here-FUT.1-Q. go-REFL-ADV.SS-DIR.EV probably-TOP send-here-FUT.1
\end{verbatim}

‘For if I do not go away, the Advocate will not come to you; but if I go, I will send him to you.’

(qub-x-bible, 43016007; Cond6)

Incidentally, as in Wik Mungkan, several selectives can be stacked. The form Niiy-an-iy-a, [2PL-DEF-TOP-SEQ] has as many as three markers qualifying as selectives, a definite suffix, what Kilham et al. (1986: 415) call the topic suffix -iy and the sequential clitic -a, (with comma, as opposed to -a “indicative”, without comma) (Kilham et al. 1986: 401). In subordinate clauses, it is the third of these that is most relevant.

If we now reconsider the set of hypotheses from 2.1 in Table 1, it has to be admitted that the definitional procedure interferes partly, but only partly, with two of them. Speech-act personal pronouns are some sort of constituent already, which may have an effect on biasing selectives to syntactic constituents (H4). Opting for personal pronouns but excluding markers that only occur with personal pronouns already entails a moderately high degree of freedom of host selection but does not imply clauses as possible hosts (H6).

In this section I have compiled a sample of selectives. They all occur with explicit personal pronouns in contrastive uses. Let us now consider in Section 4 how this set of markers from different languages behave across various subordinate clause domains in the same parallel text corpus.

4 Selectives across subordinate clause domains

4.1 Introduction

This section will introduce the selectives-across-subordinate-domains tendency scale and discuss the relationship of selectives and preposed versus postposed
order of subordinate clauses (4.1) before addressing some minor exceptions of the
tendency scale in Section 4.2. Section 4.3 deals with the preferred position of
selectives after the constituent they have scope over and proposes a functional
explanation in terms of ambiguity avoidance. Section 4.4 addresses the question
as to why selectives, which are frequent on conditional clauses, tend not to be
conditional connectives. While 4.1–4.5 emphasize general trends that can be
observed for selectives across subordinate clause domains, the summarizing remark
4.6 also highlights their crosslinguistic heterogeneity.

4.2 The selectives-across-subordinate-domains tendency
scale emerging from the visualization of the database

A major result of surveying selectives in 81 languages across subordinate clause
domains is that there is a tendency scale (no strict hierarchy, has minor, but no
blatant exceptions):

(18) The selectives-across-subordinate-domains tendency scale

```
GENERAL.RELATIVE>CONDITIONAL>TEMPORAL>CONCESSIVE>COMPLEMENT/CAUSE/PURPOSE
```

The scale emerges from Figure 1, which visualizes the database underlying this
study. The underlying set of codings is available at https://doi.org/10.5281/
zenodo.5825998. The 81 translations of the N.T. reflecting 81 doculects are given as
rows (start by three-letter ISO 639-3 codes) ordered according to their frequency of
selectives in the 200 clauses contexts surveyed (x-axis). You can see the tendency
scale as a triangle in the left upper corner when looking at the figure from a
distance and when disregarding different colors. Colors indicate the linguistic area
(see legend). For instance, purple rows reflect the five Oto-Manguean languages all
from the Cuicatlán area and it can directly be seen that, for instance, Chiquihuitlán
Mazatec (maq) and Teutila Cuicatec (cut) exhibit very similar patterns. Each col-
umn stands for a clause and columns are grouped according to rough semantic
domains as explained earlier in Table 2 in 2.3. Domains (separated by black vertical
lines) are ordered according to a compromise of frequency of selectives and se-
metric similarity of domain. Within each domain, the clauses with most selectives
across doculects are placed to the left.

The generalization in (18) formulates the tendency scale on a rather abstract
level. As one zooms in to more fine-grained domains and exemplars, there are
many special circumstances which complicate matters. For instance, concessive
conditionals (ConcCond) are intermediate between conditionals and concessives,
generalized temporal clauses (GenTime) are intermediate between conditional and temporal, and ‘until’ clauses (Termin) behave in many ways like purpose clauses rather than other temporal clauses.

Small black crosses indicate postposed subordinate clauses (the main verb corresponding to the verb in the English main clause precedes the main verb corresponding to the verb in the English subordinate clause). Postposed clauses exhibit a largely complementary distribution to selectives (almost all small crosses have white background). In no language of the sample do postposed clauses have a preference for selectives. This finding supports H2 “Selectives tend to occur early in the sentence.”

Figure 2 displays the percentage of postposed subordinate clauses (x-axis) as compared to percentage of selectives across the doculects of the sample, which reflects an implicational relationship as the space on one side of the diagonal is empty. However, word order determines occurrence of selectives only partly. Concessive clauses, for instance, are almost always preposed in the dataset, but lack selectives almost completely. Put differently, no postposed clauses attract selectives, but not all preposed subordinate clauses attract selectives.

Turning back to Figure 1, “Symbols” in the legend distinguish different kinds of marking. Notice that there is a difference between “None” (small dot: no
selective) and “Not coded” (empty: the passage is lacking in the translation, the translation is too literal, or it was not possible to make a decision). There are special symbols for selectives following after a comma explicitly marking a break at the end of the subordinate clause (“TOP after comma”, see 4.4 vi) and selectives following an anaphor picking up the subordinate clause in the main clause (“TOP on anaphor”, see 4.4 vii). These two kinds of special cases have been counted as selectives. Other symbols refer to selectives occurring with main rather than subordinate clauses, lexicalized selectives with scope only over connectives, and “Similar markers” that are not identical with selectives. All these special cases will be illustrated and discussed below.

4.3 Apparent counterexamples to the scale (but no blatant violations of it)

4.3.1 Apparent counterexamples to conditional > temporal

There are four kinds of phenomena suggesting that selectives can occasionally be more extended in temporal than in conditional clauses: (i) preposed ‘until’ clauses,
(ii) relative clauses in the temporal domain in a range of different languages, (iii) contrast, and (iv) the special case of the non-future time demonstrative in Dii. These are now treated one-by-one.

(i) Preposed ‘until’ clauses
Selectives in ‘until’ (Termin) in the three Sino-Tibetan languages Dimasa (dim), Amri (ajz) and Hills Karbi (mjw) seem at first glance to be a violation of the selectives-across-subordinate-domains scale. However, these clauses only violate the tendency scale when (i) preposed and (ii) when constructed with expletive negation. The Sino-Tibetan languages in the sample all have preposed ‘until’ clauses, which goes against the general trend for postposed position. Moreover, in Dimasa (Dundas 1908: 17), Amri Karbi and Hills Karbi, ‘until’ is constructed with expletive negation, which renders the construction non-veridical (see 2.4) and in this respect similar to conditional clause constructions. The Hills Karbi connective tik expresses ‘as long as’ in affirmative construction and ‘until’ in (19) only together with negation (Grüßner 1978: 192; for Dimasa, see Dundas 1908: 17).

(19) Hills Karbi

“Nang-tum nohok a-dim-le dam settame halabangso
2-PL everywhere POSS-place-FOC.IRR go nevertheless, that
a-dim pen bar∼be tik ke nang-tum
POSS-place from exit∼NEG as.long.as/until TOP 2-PL
ke-lut-dam a-hem-le do-non
NMLZ-enter-go POSS-house-FOC.IRR stay-COND.IMP.
‘Whenever you enter into a house, stay there until you depart from there’,
lit. “Wherever you go, as long as you do not exit from that place, stay at the house you entered.”
(mjw-x-bible, 41006010; Termin5)

Amri Karbi has the connective leng (always combined with negation, often following a nominalized clause, perhaps from (h)leng ‘reach’, N. Philippova, p.c.), which is often followed by a selective in negative conditionals and in ‘until’. In Amri Karbi, the selectives in ‘until’ can thus be viewed as a language-specific close connection to negative conditional clauses.

Also, Japanese wa <ha> can occur on terminal clauses. This is because made ‘until’ is part of the set of case particles that can be followed by the topic marker wa

---

10 Unlike ‘before’ clauses, where expanded negation is crosslinguistically common, expletive negation in ‘until’ is mostly restricted to South Asia and Eastern Europe (Hetterle 2015: 136; Wälchli 2018b).
(Wlodarczyk 1998). ‘Until’-clauses are thus in a sense more nominal than other subordinate clauses that are not followed by *wa*.

(ii) Relative clauses headed by a temporal noun or headless with such a noun implied

Temporal clauses can be relative clauses headed by temporal nouns, such as ‘time’ or headless relative clauses with such a temporal noun implied (for Abau, see Lock 2011: 219). Since they have properties of definite noun phrases, such clauses are often inclined to bear selectives. In Abau (aau), as in (20), this contributes especially to a higher incidence of selectives in simultaneous clauses.

(20) Abau

\[ \text{Abau} \]

\[ \text{Ama, enekwei, ha ki mon ma lwak o-ho-kwe, ...} \]

\[ \text{ADDR.PL time 1SG ground LOC REL be this-GL.M-TOP} \]

‘While I am in the world, [I am the light of the world.]’

(aau-x-bible, 43009005; Simult1)

In Zarma-Kaado (dje, see (32) in 4.4), relative clauses cause an apparent disruption of the selectives-across-subordinate-domains scale. In Umanakaina (gdn), temporal clauses with selectives have *make-ya* ‘time/day-LOC’. The Ewe (ewe) connective *esi /èsì/ ‘when’ is the nominalized (=headless) relative pronoun *si*. In Gaa (gaa), postposed relative (*ni*) and temporal clauses (*be ni*) can bear the selective *lɛ* (same marker as determiner), whereas with conditionals, selectives are restricted to preposed clauses. In Yawa (yva), which has few subordinate clauses with selectives, selectives occasionally occur in ‘before’ clauses with the head noun *arono* ‘time’, as in (21):

(21) Yawa

\[ \text{Yawa} \]

\[ \text{arono Abraham avaki rainy-amo, Risy-amro ri-no to} \]

\[ \text{time.period/when Abraham be.born never-top 1SG-TOP 1SG-be-at go} \]

‘Before Abraham was, I am!’

(yva-x-bible, 43008058; Post1)

(iii) Contrast and other orthogonal distinctions

Figure 1 suggests that Sochiapan Chinantec (cso) is more inclined to have selectives in temporal than in conditional clauses. However, what actually triggers selectives is adversative context in contrast. This is illustrated in (22) where the counter-expective selective *máñá* /máñá/ occurs, in an adversative context.
Sochiapan Chinantec

\( Ti³la³ \quad m³-ca³-jenh³ \quad jná³ \quad jáun² \quad m³’ná³, \ldots \)

but \( \text{when.FUT PERF-PST-raise 1SG then CTOP.} \)

‘But after I am raised, [I will go ahead of you into Galilee.]’

Foris (2000: 388) points out that “topic”-marked objects tend to be contrastive. In the N.T., both general \( n³ \) and counter-expective selective \( m³’ná³ \) (Foris: \( m³’ná² \)) occur in adversative contexts.

Toura (neb) has two items functioning as selectives: \(-le \ TOP\) and adversative selective \(-le \ laá \ [TOP \ TOP.FOC].\)\(^{11}\) Only the adversative selective \(-le \ laá\) occurs with subordinate clauses and only in adversative contexts, but – unlike Sochiapan Chinantec – only in general relative, conditional, generalizing temporal and concessive conditional clauses – an example is given in (23) – which is fully compatible with the tendency scale.

(23) Toura

\begin{align*}
\text{Atoosiwo } & \text{móò } \text{ lò’, ke } \text{ ka } \text{ zvugbaan=mèè láà àà nu-ii } \text{ ka } \text{ pé} \\
& \text{ for 1SG.NEG go CONJ 2PL comforter(?) that NEG come-PROG 2PL side.} \\
\text{Kè } & \text{ an } \text{ lò’ } \text{ laá, } \text{ le } \text{ an } \text{ à } \text{ bo } \text{ ka } \text{ nè } \text{ le.} \\
& \text{ but 1SG go TOP.FOC CONJ 1SG him send 2SG to TERMIN} \\
\end{align*}

‘…For if I do not go away, the Advocate will not come to you; but if I go, I will send him to you.’

(neb-x-bible, 43016007; Cond6)

In Yuracaré (yuz), \(-ja\) “topic” (after NPs) functions as same subject marker following subordinate clauses. Interestingly, however, the switch-reference system does not work alike in all types of subordinate clauses. Many more subordinate clauses can take \(-ti\) ‘different subject’ than \(-ja\) ss. According to van Gijn (2006: 328, 316), \(-ja\) ss only occurs in time, simultaneous, conditional and relative clauses, which is consonant with the tendency scale (but in the data visualized in Figure 1 there is also one purpose and one causal clause). We can thus formulate a local hierarchy for Yuracaré: general relative clauses (very often same subject) > temporal (no obstacles for \(-ja\) ss to occur) > conditional (sometimes other markers – potential \(-ta\) or non-veridical \(-ya\) – instead of \(-ti\) DS and \(-ja\) ss) > cause/purpose (almost only \(-ti\) DS) > complement and concessive \(-ti\)-jsha DS-ABL, where \(-ja\) ss never

\(^{11}\) Bearth (1992) calls \(-le F(OCUS)1\) and \(-le F(OCUS)2\), but Matić and Wedgwood (2013: 142) point out that \(-le r²\) rather resembles a contrastive topic and the examples given by Bearth and Matić and Wedgwood for the combination of \(le\)-’ > \(lāā\) \(r²+F1\) are compatible with an adversative selective analysis. To add further to the confusion, the conjunction \(le\) and the terminal marker \(le\) at the beginning and end of the main clause in (22) (Bearth 1971: 368–369) are homonymous with \(le\) \(r²\).
occurs). This is only a slight deviation from the tendency scale since conditionals have more possibilities to deviate than temporal clauses, but no blatant violation of the scale.

(iv) The non-future time demonstrative in Dii
Dii (dur) has two clause-final demonstratives máa ‘non-future time’ and tée ‘future time’ (Bohnhoff 2010: 112). Only máa ‘non-future time’ is used as an adversative selective following personal pronouns. Now, while máa ‘non-future time’ occurs in temporal clauses, conditional clauses have tée ‘future time’ (in Figure 1 indicated as “Similar marker”). Whether Dii is an exception to the selectives-across-subordinate-domains scale boils thus down to the question as to whether the tight paradigmatic relationship between the two clause-final demonstratives is sufficient for arguing that even the marker tée ‘future time’ is a selective. Example (24) illustrates that their functions are very similar, they can both be used following general relative clauses:

(24) Dii

\[
\begin{align*}
Nán \ & ka \ & dì \ & kan \ & sqqm \ & idù \ & tée, \ & a \ & púm\text{an} \\
\ & \ & \ & \ & \ & \ & \ & \text{be.there with} \\
\ & \ & \ & \ & \ & \ & \ & \text{cloth} \\
\ & \ & \ & \ & \ & \ & \ & \text{two} \\
\ & \ & \ & \ & \ & \ & \ & \text{TOP.FUT} \\
nán \ & í- \ & ka \ & 'yé-n \ & né \ & máa \\
\ & \ & \ & \ & \ & \ & \text{NEG} \\
\ & \ & \ & \ & \ & \ & \text{TOP.N.FUT} \\
\ & \ & \ & \ & \ & \ & \text{one} \\
\ & \ & \ & \ & \ & \ & \text{put-CONEG} \\
\ & \ & \ & \ & \ & \ & \text{NEG} \\
\ & \ & \ & \ & \ & \ & \text{TOP.N.FUT} \\
\ & \ & \ & \ & \ & \ & \text{one} \\
\ & \ & \ & \ & \ & \ & \text{have one.}'
\end{align*}
\]

(dur-x-bible, 42003011; GenRel16)

However, Dii is no blatant violation of the scale. Interestingly, Dii máa is a marker only secondarily expanded to noun phrase selectives. This brings us to the discussion of selectives originating in subordinate clauses.

4.3.2 From conditionals to selectives, and apparent counterexamples to GenRel > Cond

Including Dii, six languages of the sample provide evidence for a subordinate clause origin of selectives.

Longchuan Achang, Khalkha Mongolian and Kalmyk suggest a development CONDITIONAL > SELECTIVE. Spoken Indonesian kalau ‘if’ (not in the N.T. sample; Bruno Olsson, p.c.) also sorts here. In Khalkha Mongolian and Kalmyk, it is more precisely the conditional converb of the verb ‘to be’ that turns into an (adversative) selective. Khalkha bol is a shortened form of bol-vol [be-COND.CVB] with a conditional converb suffix (Poppe 1951: 87, 99). In Kalmyk bol-xla [be-COND.CVB], the conditional converb suffix originates from the future participle -x and the comitative -la (Benzing 1985: 184).
Since Khalkha -vAl and Kalmyk -x1A as verbal suffixes cannot be considered selectives, they are indicated with the symbol for “Similar marker” in Figure 1. Longchuan Achang (acn) gas 'verbal subjunctive particle; if, in case that’ (Sampu et al. 2005: 33) is the only selective in the sample that is more extended in use in conditional clauses than in general relative clauses.

Gulay and Soughb provide evidence for a more extended grammaticalization path: TEMPORAL > CONDITIONAL > SELECTIVE. Gulay (gvl) baa /bâ:/ derives from bâ /bâ:/ ‘already’ plus an older selective -â (Djarangar 1989: 786) and must have originated in anterior contexts such as (25), where bâ ‘already’ is now followed by baa.

(25)  
Gulay  
\[\text{Loo ke sei inge-i bá baa}\ldots\]  
place REL 2PL find-O.3SG already TOP  
‘…and when you have found him [report to me so that I also may come and worship him]’

(gvl-x-bible, 40016002)

The current use of baa suggests a development ANTERIOR > SIMULTANEOUS > CONDITIONAL > GENREL > NP SELECTIVE. A similar development must have occurred in Soughb (mnx), where kaba ‘then’ is used as a selective also following general relative clauses, as in (26), and NPs:

(26)  
Soughb  
\[\text{G-areg dau yen m-en sansun hwai, kaba, en eic}\]  
NMLZ-REL from 2PL POSS.3SG-POSS cloth two then/TOP 3SG take  
homa dou me-si gedig m-en sansun ingma-ro.  
one to POSS.3SG-friend REL POSS.3SG-POSS cloth some-NEG  
‘The one who has two tunics must share with the one who does not have one…’

(mnx-x-bible, 42003011; GenRel16)

From conditional clauses the use of kaba ‘then’ is even extended to fronted indirect questions. Kaba also occurs on temporal adverbs, such as iteitogin, kaba [now TOP] (see Example (47)) which shows that it is not just a temporal adverb. Occasionally, kaba is further expanded by another element originating from a temporal adverb kaba, cum, [then/TOP in.a.while].

4.3.3 The lowest part of the scale

Complement, causal and purpose clauses all tend to lack selectives. Figure 1 only reveals one systematic exception: verb of knowing clauses in Nalca (Illustrated
earlier in Example (4)). However, complement clauses are not entirely without relevance for selectives, as will be discussed in 4.4 (iv).

Purpose clauses have selectives very rarely. Brown (1990: 106) reports that in Waris (wrs) – a language with very extended use of selectives – purpose clause can have selectives even in sentence-final position, but Figure 1 shows that no language of the sample, not even Waris, generally has selectives in purpose or result clauses.

4.4 The position of selectives with respect to subordinate clauses

There is an astonishing uniformity in the positioning of selectives. Usually, selectives simply follow the subordinate clause: $S_{\text{TOP}}$ or $S=\text{TOP}$, which supports H3: selectives tend to occur after or at the very end of their constituent. This section discusses the few deviations from this general trend, summarized in Table 4.

(i) **Clause-final selectives**: In most languages of the sample, selectives occur at the very end of subordinate clauses – often as clause-final clitics (see, for instance, Curnow 1997: 387 for Awa-Cuaiquer). This preferred order best averts ambiguity between clausal (wide) and phrasal (narrow) scope of the selective. According to H5 “Selectives mark constituents of very different length”, which may cause scope ambiguities. H4 “Selectives are associated with a syntactic constituent” presupposes that a selective can be unambiguously attributed to a constituent it scopes over. Ambiguity-avoidance thus testifies to the relevance of scope of selectives. Since clause-final constituents are usually not topic-like, $S=\text{TOP}$ usually prevents all risk of ambiguity between selective clause-internal constituents and

<table>
<thead>
<tr>
<th>Table 4: Kinds of positioning of selectives.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) Clause-final</td>
</tr>
<tr>
<td>(ii) Clause-initial</td>
</tr>
<tr>
<td>(iii) Clause-internal</td>
</tr>
<tr>
<td>Following clitic</td>
</tr>
<tr>
<td>Following verb</td>
</tr>
<tr>
<td>Following head noun</td>
</tr>
<tr>
<td>(iv) Following connective</td>
</tr>
<tr>
<td>(v) Following anaphor</td>
</tr>
<tr>
<td>(vi) Following comma</td>
</tr>
<tr>
<td>Most languages of the sample</td>
</tr>
<tr>
<td>Gude, Bolinao, Botolan Sambal</td>
</tr>
<tr>
<td>Parecis</td>
</tr>
<tr>
<td>Waris</td>
</tr>
<tr>
<td>Zarma</td>
</tr>
<tr>
<td>Iyo, Muyang (“TOP on connective”)</td>
</tr>
<tr>
<td>Borong, Warlpiri and few others (“TOP on anaphor”)</td>
</tr>
<tr>
<td>Suena, Zia, Waskia, Doromu-Koki, Cuicatec (“TOP after comma”)</td>
</tr>
</tbody>
</table>
selective-marked clauses. Yanesha’ (ame) (27) – where subjects are sometimes final in subordinate clauses – can also disambiguate by the choice between simple -pa’[TOP] and a complex -ña-pa’[SEQUENCE.MARKER-TOP]. Only -pa’ can occur at the end of subordinate clauses (Duff-Tripp 1997: 190).

(27) Yanesha’
\[ Allempo-ña aña\text{'}ña cha\text{’}tarat̃ Herodes e\text{’}man atet̃ ñoñets-pa’ ſña-ña-pa’ \]
then-seq governor/king Herodes hear thus word-TOP 3SG-SEQ-TOP
\[ ñeñt̃a na’ u-ña p-ueyočhr-o ama eñmat-o p-ueyočhr-o. \]
that.no more POSS.3SG-heart-LOC NEG be.happy-CONEG POSS.3SG-heart-LOC
\[ ‘And when King Herod heard it, he was troubled...’ \]
(ame-x-bible, 40002003, Sequent1)

(ii) Initial selectives: Only three languages of the sample – Gude (gde), Botolan Sambal (sbl) and Bolinao (smk) – have initial selectives. In the two Sambalic languages, selectives are restricted to NPs and headless relative clauses and the Botolan Sambal inversion marker ay preceding the verb often follows the selective constituent, hereby indicating where it ends. In Gude, nominal selective constituents have an initial particle ma and can be followed by prepause -a (Hoskison 1983: 127). According to Hoskison (1983: 129), ma also introduces conditional clauses. However, the N.T. has in fact mada and counterfactual maci (Hoskison 1983: 130), which contain the initial selective ma, but cannot mark NPs. Thus, in (28) ma-ci is the clausal selective and the nominal selective constituent ‘this person’ is clearly singled out by final -a.

(28) Gude
\[ Ma-ci ma ñanda-ts-a. ñentaa da-cii ñantafó shi ca. \]
TOP-CNTRFC TOP person-that.close-TOP NEG.FOC with-hand God come he
\[ ka-da pooshi uushi cii kuya mbee ka da-na. \]
POT-do not thing hand 3_M.POSS be.able POT do-NMLZ
\[ ‘If this man were not from God, he would not be able to do anything!’ \]
(gde-x-bible, 43009033; CondHyp2)

(iii) Clause-internal selectives: In Parecis (pab), the selective atyo often follows the first element – a position with potential scope conflict. However, in subordinate clauses, it typically follows the conditional clitic -(i)ya, which in turn follows the first or the “thematic” element according to Rowan and Burgess (2008: 26), which prevents scope ambiguity, as shown in (29). [In hypothetical contexts -(i)ya also occurs in main clauses.]
(29) Parecis

_Only when first tying the strong man, he can then plunder his house_

(pab-x-bible, 40012029; CondNeg11)

In contrast to the closely related language Amanab, Waris (wrs) exhibits such a considerable inflation of selectives that it is perhaps easiest to describe which words do not have selectives. They are mostly lacking with finite verbs in main clauses, some function words and closely integrated attributes in continuous NPs. Selectives on subordinate clauses follow the verb, which in conditional clauses bears an irrealsis markers (-ta ‘unanticipated/undesired’ or -ra ‘anticipated/desired’; Brown 1990: 64). The verb need not be clause-final, but often is, as in (30).

(30) Waris

_‘…if you had been here, my brother would not have died.’_

(wrs-x-bible, 43011032; CondHyp6)

The Parecis and Waris examples also show that there is no fundamental difference between clause-final and clause-internal order. In (29) and (30), the selectives are actually also final. Also positioning of selectives following the verb is not exclusive to Waris, in many languages with clause-final selectives the verb is always clause-final.

(iv) Selectives on connectives: Clause-final positioning is not the only scope-disambiguating position. If the selective immediately follows the connective, it cannot have NP-scope. In Iyo (nca), the combination of relative _qu_ and selective _-ko_ is lexicalized as adversative connective _quko_ ‘but’ (Minter 2009: 68). This is a case of a selective in a lexicalized unit indicated as “TOP on connective” in Figure 1. As in many other languages, adversative connectives in Iyo are used in concessive constructions (although _p, q = p, but q_), illustrated in (31).
A parallel can be found in Muyang ay ti [thus TOP ‘but’, where the source for lexicalization, is a dangling adverb.

It has been pointed out that selectives often develop from definite articles or demonstratives. However, interestingly, most languages where such a grammaticalization occurs, place determiners in absolutely final position or the demonstrative or article shifts to final position during the grammaticalization process. This suggests that the grammaticalization of determiners to selectives is conditioned by word order properties (implying absolutely phrase-final position as attractor position). The question thus arises as to whether there can be any non-final selective determiners. In the sample, there is only one candidate. Zarma (dje) has the demonstrative /wô/ ‘this’ that follows selective personal pronouns (araŋ wo [2PL this/TOP]), hereby meeting the criteria for selectives. If we consider the definite article -o to be related to this demonstrative, it is instructive to look at its word order, which is N-o S. The unusual position N-MARKER S for a selective goes hand-in-hand with an unusual distribution of markers in Zarma on the selectives-across-subordinate-domains scale, not including conditional clauses (see Figure 1: dje). Many temporal clauses consist of relative clauses with a temporal head noun such as waat ‘time’. Example (32) shows that occasionally even generalizing (“indefinite”) temporal clauses can have such definite marking.

(v) The selective on anaphor strategy: An alternative to selectives within or on the clause is to place the marker after an anaphor following the subordinate clause: S,
ANAPH-TOP M. The anaphoric strategy – to the extent that it does not involve clausal anaphors – is best suited for general relative clauses and for conditional clauses with the same subject as the main clause. Example (33) from Burum-Mindik (bm) indicates that the anaphoric strategy may occur side-by-side with clause-final selectives, here with a nominalized subordinate clause. Burum-Mindik exhibits a further unexpected parallelism in that the anaphoric pronoun with the selective bears the same nominalizing marker as the selective-marked clause (underlined in (33)). [Borong (ksr), closely related to Burum-Mindik, only has anaphoric selectives.]

(33) Burum-Mindik

\[
\begin{align*}
\text{An} & \text{t} - \text{u} & \text{t} - \text{ö} & \text{kö} - \text{l} - \text{ö} & \text{m} - \text{a} - \text{k} - \text{z} - \text{a} - \text{w} - \text{i}, & \text{y} - \text{a} - \text{j} - \text{ö} - \text{n} & \text{mönö} \\
\text{one/a} & \text{nother-AD} & \text{j} & \text{God} & \text{ask-HAB-PRS-3SG-DET} & \text{3PL-NMLZ-TOP} & \text{just} \\
\text{like} & \text{pə} & \text{ŋ} & \text{buni} & \text{ŋ} & \text{qem} & \\
\text{on.the} & \text{other.side} & \text{mess} & \text{age/promise} & \text{hit.ss (>possess)} \\
\text{an} & \text{γu-ma} & . & \text{Yuai-go} & \\
\text{carry.on.shoul} & \text{der/hold.fast-CUST} & \text{thing-POSS/for} \\
\text{jar} - \text{u} - \text{m} - \text{k} - \text{z} - \text{a} - \text{w} - \text{an} - \text{j} - \text{o} - \text{n} & \text{mönö} & \text{m} & \text{mi} & \text{mi} & \text{wi} & \text{k}_{\text{ŋ} - \text{ai-ma}} \\
\text{seek-} & \text{HAB-PRS-3SG-NMLZ-TOP} & \text{just} & \text{that} & \text{find/create-CUST}. & \\
\text{‘For every} & \text{one who asks receives, and the one who seeks finds…’} & \\
\end{align*}
\]

(bmu-x-bible, 40007008; GenRel14)

Among the languages of the sample, the anaphoric strategy is most pervasive in Warlpiri (wbp), which has a clausal anaphor ngula and where the anaphoric strategy is the only possibility to mark clauses with selectives (see also Bittner 2001). Warlpiri is well-known for its many second position clitics and the selective is one of them. Restricting selectives of clauses to the anaphoric strategy has the effect of minimizing the potential for scope ambiguity. Whereas the selective with general relative and conditional clauses usually follows the bare anaphor ngula, temporal clauses can have more extended forms such as ngula=jangka-ju [then=after.that-TOP] or ngula=kurra-ju [then=ALL-TOP]. Posterior clauses such as (34) exhibit the most baroque forms with both negative (“privative”) and temporal adverbial suffixes.

(34) Warlpiri

\[
\begin{align*}
\text{…} & \text{Kuja-lpa Yipuruyamu nyina-ja kirri-ngka yirdi-ngka Yarrana-rla,} \\
\text{REL-IPV} & \text{Abraham be-PST village-LOC name-LOC Haran-LOC,} \\
\text{ngula=} & \text{wangu-rla=wiyi} & \text{ngula-ju-lpa} & \text{nyina-ja nguru-ngka} \\
\text{then=} & \text{PRIV-LOC/SEQ=first/before then-TOP-IPV} & \text{be-PST country-LOC} \\
\end{align*}
\]
(vi) The selective “after comma” strategy: A more laconic, but similar, strategy is to skip the anaphor as a separate element, but to place the selective still outside of the subordinate clause together with the main clause following a clause break indicated by comma, as illustrated in (35) from Suena (sue).

(35) Suena

Ge pot-inowe-na awa dim-amuno-wa,

awl talk put-PST.TODAY.CONT-1SG 1SG-OBJ TOP that-object carry-FUT.PUNCT-2/3PL

awl TOP/1SG friend PL.of.friend 1SG-OBJ/POSS do-FUT.PUNCT-2/3PL

‘You are my friends if you do what I command you.’

(sue-x-bible, 43015014; CondFin1)

As indicated in glossing, Suena awa is not only the selective but also the object form of the distal demonstrative pronoun, and (35) illustrates that it is not only used as a selective following conditional clauses, but also as complementizer in object complement clauses. It is not unlikely that the selective use after conditional clauses has originated from object complement clauses. The “subject” or perhaps rather ergative forms of the distal demonstrative Suena ami and Zia amimene in general relative clauses are coded as “similar markers” in Figure 1.

In Waskia (wsk), the selective mu – identical with the determiner mu ‘the, that’ – can occur on either side of the comma in conditional constructions. The occurrence of agi ‘or’ in conditional clauses (36) testifies to the close relationship with complement questions, which also can bear the selective (37):

(36) Waskia

Nina ani ninguru aga iki-man agi mu, nina

2PL 1SG truly 1SG.OBJ/POSS hear/know-PST.1/2PL or TOP 2PL

aga Ait ago ko iki-man.

1SG.OBJ/POSS my.father(?) with 3SG.OBJ hear/know-PST.1/2PL

‘If you have known me, you will know my Father also.’

(wsks-x-bible, 43014007; CondHyp8)
Waskia

...se nu Jesus kuer-am agi mu ko ninguru isu-am.

and 3SG Jesus die-PST.3SG or TOP 3SG.OBJ truly ask-PST.3SG

‘...(and Pilate) asked him whether (Jesus) had died already.’

(wsk-x-bible, 41015044; Dic5)

Given the definition of selective in Section 3 as forming a unit together with a constituent (the personal pronoun), it may be surprising that there is such a thing at all as a selective “after comma” strategy but note that the definition only applies to selectives following a noun-phrase-like constituent. Selectives cling “to the wrong side” only with subordinate clauses, not with NPs. However, the question arises as to whether such selectives following commas are in fact spurious selectives at least diachronically, as the case of Suena and the closely related language Zia (zia) suggests, where it is likely that the markers are an instance of demonstrative-to-complementizer grammaticalization, which is well-attested even in languages where no selectives are involved, as in English that.

While selectives in Suena, Zia and Waskia have pronominal origin, I do not know what their origin in Doromu-Koki (kqc) and Cuicatec (cux, cut) is, where the selective “after comma” strategy is attested as well, but in all these languages, forms used as selectives also occur sentence-initially, which suggests that they all have some sort of anaphoric or conjunctional properties. According to Bradshaw (2012: 50), the Doromu-Koki selective bi can follow a zero anaphor. Cuicatec na is also a coordinator ‘and’ (Anderson and Roque 1983: 630).

(vii) Selectives on (and in) main clauses: Given that selectives are not prototypical subordinators, it is not surprising that they occasionally follow preposed main clauses (M=TOP S), which is common in some Chadic languages, as illustrated in (38) from Matal (mfh).

(38) Matal

Mo-səl kà masla kona gami

1PL-know TOP 3SG.ANAPH son 1PL.POSS

‘We know that this man is our son.’

(mfh-x-bible, 43009020)

In Matal, the selective kà also introduces postposed purpose clauses and, in collocation with the relativizer, kà uwana [TOP REL] ‘because’ introduces causal clauses (Verdizade 2018: 34). This appears in Figure 1 with the symbol “TOP on main clause”.

A further phenomenon I have so far only encountered in Chadic languages is illustrated here from Merey (meq), a language with rather extended use of selectives. In (39), the selective na follows the background (single underline) with the
question word in focus sentence—finally (see also Gravina 2007: 6). It is interesting to notice here that the selective-marked background (underlined) is not a straightforward simple phrase (NP, adverb or clause) as in all other examples, even though it probably does not violate H4 “Selectives are associated with (have scope over) a syntactic constituent”.

(39) Merey

\[
\text{Na tsik-akum de} \quad \text{dedek na, ka dzal-um ha}
\]

\[
\text{faya bay na, hərwi mey?}
\]

‘If I am telling the truth, why do you not believe me?’

(meq-x-bible, 43008046, Cond14)

To summarize, despite much diversity in how selectives are used in complex sentences, selectives almost always occur after the selective constituent, largely confirming H3 “Selectives tend to occur after or at the very end of their constituent”. This is a position with maximally little potential for scope ambiguities, and in the few cases where the position is not absolutely terminal, there are usually other, more language-specific, strategies to avert ambiguity.

4.5 The relationship between conditionals and selectives

According to Comrie (1986: 88), conditionals differ in degree of hypotheticality, without there being any neat divisions between real, hypothetical and counterfactual. Degree of hypotheticality plays a role for selectives only in a few languages of the sample, and there is no uniform trend. In Abau (aau), non-hypothetical conditionals lack selectives, because they take the same connective ankın as temporal clauses not followed by the selective (Lock 2011: 367). Non-hypothetical conditionals are more akin to temporal clauses than hypothetical ones, and Abau testifies to the fact that temporal clauses can occasionally trigger the lack of selectives in non-hypothetical conditionals. Siane (snp) shows the opposite trend. The hypothetical-conditional connective neko (perhaps from ne-ko [exist-since/only]) is compatible with the demonstrative and selective ya ‘there’ (Potts and James 1983: 12) only if ya follows after the comma. In Ga (gaa), the counterfactual is marked initially by àjí and finally by kúllé (or kúllé), with the definite article, which serves as selective, as a second component (Campbell 2017: 619).

If word order is disregarded, conditionals generally behave rather uniformly when it comes to selectives, which suggests that there is considerable potential for lexicalization. Indeed, several languages have markers in conditional clauses that
look suspiciously similar to selectives, even though they are not selectives strictly speaking (in Figure 1 indicated with the symbol for “Similar marker”): Meyah (bera//beda), Moskona (era//eda), Yawa (-amo//weamo), Una (-ra//ura), Mian (-le//mole), Iyo (-ko//-qo), Korean (-nun//-myen) and Japanese (wa//ba). This is hardly a coincidence. In some of these, selectives and the markers attested in conditional clause will ultimately derive from the same source. For Korean (kor), there is actually diachronic evidence that this is the case (Koo 1999). In Japanese, ba and wa share some morphophonological properties (Shibatani 1990: 164). Also, in casual speech, the conditional and topic particles drop their consonants and become homophonous (Shibatani 1990: 176). In Moskona (mtj), TOP erá and the apodosis marker edá ‘then’ are minimal pairs (Gravelle 2010: 24). The same holds for the closely related language Meyah (mej) with TOP berá and the apodosis marker bedá ‘then’ (Gravelle 2004: 33). However, in generalized relative clauses both markers occur side-by-side, as in (40):

(40) Meyah

\[\text{Jeska} \quad \text{r-usnok} \quad \text{nomnaga} \quad \text{ongga} \quad \text{ri-rejgei} \quad \text{gu} \quad \text{Ofa} \quad \text{rot}\]

\[\text{because} \quad \text{3PL-person} \quad \text{all} \quad \text{REL} \quad \text{3PL-request} \quad \text{to/at} \quad \text{3SG} \quad \text{concerning}\]

\[\text{mar} \quad \text{bera} \quad \text{ri-m-esma} \quad \text{mar} \quad \text{insa} \quad \text{koma} \quad \text{si}. \quad \text{Noba}\]

\[\text{thing} \quad \text{TOP} \quad \text{3PL-IRR-receive} \quad \text{thing} \quad \text{ANAPHOR that} \quad \text{probably} \quad \text{and}\]

\[\text{r-usnok} \quad \text{nomnaga} \quad \text{ongga} \quad \text{rua} \quad \text{ri-m-ek mar}\]

\[\text{3PL-person} \quad \text{all} \quad \text{REL} \quad \text{3PL-search} \quad \text{thing} \quad \text{then} \quad \text{3PL} \quad \text{3PL-IRR-find}\]

\[\text{insa} \quad \text{koma} \quad \text{si}. \quad \text{thing} \quad \text{ANAPHOR that probably}\]

‘For everyone who asks receives, and the one who seeks finds…’

(mej-x-bible, 40007008; GenRel14)

In Yawa (yva), Una (mtg), and Mian (mpt), the only problem is that the initial parts of the markers used in conditional constructions have no etymology, while the final parts are identical with selectives.

There is also evidence for on-going lexicalization. Chipaya (cap) nïki ‘if’ contains a fossilized masculine demonstrative form *ni(i)-ki [that/the.M-TOP] (Cerrón-Palomino 2006: 264). The Chipaya form may have originated in general relative clauses, where it occurs side-by-side with selective-marked relative forms, as shown in (41):

(41) Chipaya

\[\text{Nižaša} \quad \text{ni} \quad \text{tanš-i-ki}, \quad (...)\]

\[\text{thus} \quad \text{that.M} \quad \text{catch-REL-TOP} \quad (...)\]

\[\text{xaqši-lla-žlax} \quad \text{ana} \quad \text{tanš-i} \quad \text{nï(-)ki}, \quad (...)\]

\[\text{which-PL-???} \quad \text{NEG} \quad \text{catch-REL} \quad \text{(that.M(-)TOP)if}\]

‘For whoever has, (more will be given to him,) and whoever does not have…’

(cap-x-bible-2004, 41004025; GenRel17)
Ama (amm) has both optional and non-optional uses of selectives in conditional clauses. There is a lexicalized future conditional V-aki-PERSON-mo [TOP], with the future conditional morpheme -aki- occurring only in this construction, where selectives are obligatory (Årsjö 1993: 14). A token of lexicalization is that the construction with -aki- can occur in postposed conditional clauses as illustrated in (42):

(42) Ama

Moti mo sai, Musi-ni ami imo mo itouniyaimo,
1DU.INCL TOP know, Moses-GEN strong word TOP good
aiyolo-ko itouniyaim-aki-koni-mo.
teach-3N.SG.O/S good-FUT.COND-1N.SG.INCL.O/S-TOP

‘But we know that the law is good, if anyone makes use of it lawfully.’
(amm-x-bible, 54001008; CondFin11)

Also, Chipaya nīki ‘if’ occurs in postposed conditional clauses which further testifies to its lexicalization.

Comrie (1986: 86) argues that, if conditionals are topics, there must be some degree of grammaticalization (which is the same thing as lexicalization of the conditional marker), since conditionals can occur in non-topical uses, notably in responses to questions, such as Under what circumstances will you leave? with answers with conditional clauses in focus, such as I will leave, if you pay me. Comrie’s argument is similar to the one adduced by Chappell and Creissels (2019) against topic possessives as a possible basic type in the typology of predicative possession (Stassen 2009). However, clear cases of focal conditional clauses in the sense of answers to information questions are rare, in the N.T. only attested in some cases of free translation. An example is (43) from Huallaga Quechua (qub), where conditionals are adverbial clauses with adverbial markers distinguishing switch reference (-r ss in (43)) and usually followed by the selective <-ga> /-qa/ (Weber 1989: 415). However, in (43), the conditional clause, which is an answer to a question, takes the direct evidential marker that also indicates focus.

(43) Huallaga Huánuco Quechua

¿Imanōpataj musya-nchi Tayta God-OBJ rejsi-shanchī-ta?
how know-1PL.INCL father Dios-ta recognize-SUB.1INCL>3-OBJ
Mandamintu-n-cuna-ta wiya-cu-r-mi musya-nchi
commandment-POSS.3-PL-OBJ hear-REFL-ADV.SS-DIR.EV know-1PL.INCL

‘And by this we know that we have come to know him, if we keep his commandments.’

Literally: “How do we know to recognize our father God? We know him if we hear his commandments.”
(qub-x-bible, 62002003; CondFin3)
The example suggests that -qa is no lexicalized expression of conditional in Huallaga Quechua.

The question thus arises as to what extent selectives in conditional clauses can be considered conditional connectives. Markers with similar meanings have similar distributions in parallel texts and this can be easily measured with collocation measures (see, e.g., Dahl and Wälchli 2016 and Wälchli 2019). When searching by automated means for candidates for conditional markers in the 81 N.T.s of the sample with English *if* as a seed gram, in most languages not selectives, but entirely different conditional markers, are extracted. This is because selectives generally tend to have very high text frequency. Put differently, even though selectives often occur on conditional clauses, they are not dedicated markers of conditional clauses. The selectives that turn out to match best in distribution with English *if* (technically speaking “collocate best”) fall in three groups:

(i) selectives originating in subordinate clauses (see 4.3.2): Gulay ɓaa;
(ii) a less frequent selective or a less frequent allomorph: Fon hünk ‘TOP.IMP/thus’, Kenga jaay, Korafe a-mo THAT-TOP, Aymara -xa; and
(iii) The extracted marker is a combination of the selective with some other marker: Gude mada and maci (see 4.3), Toura laad TOP.FOC, Haka Chin ah-cun LOC-TOP.OBL, Waris -ta-va IRR-TOP, Konai de-ba-be AUX/DO-PFV.IRR-TOP and ba-be PFV.IRR-TOP, Chipaya ni-ki DEM.M-TOP.

This suggests that selectives do not tend to grammaticalize (or lexicalize) to conditional connectives, but occasionally combinations of selectives with another morpheme can lexicalize to conditional markers, as in Haka Chin locative plus the oblique case form of the selective (see Barnes 1998: 68).

Further evidence that selectives are not conditional connectives comes from word order. In a considerable number of languages of the sample (Chadic, Kwa, East Bird’s Head and Oto-Manguean languages), selectives and conditional connectives end up at the two different edges of the conditional clause, which results in a kind of bracketing marking of conditional clauses. Table 5 lists some examples for verse 43014007 ‘If you have known me, [you will know my Father also.]’ (CondHyp8).

| Matal mfh | *Bañα kasαlaw kà,…* |
| Gen gej | *Ne mì jēsi mû-a,…* |
| Soughb mnx | *Sug Yesus ens enin dag gini dou Tomasohob, kabo,…* |
| Sierra de Juárez Zapotec zaa | *Canchu hualigani nabìa’ni le inte’ nna,…* |

Table 5: “Bracketing” marking of conditional clauses with selectives.
Almost all of these languages are VO languages, where adverbial subordinators are expected to be initial (Dryer 1992: 108).

### 4.6 Summarizing remark

In Section 4 I have argued that the range of possible variation of selectives across languages in subordinate languages is strongly constrained. The distribution of selectives across subordinate clause domains follows a scale without blatant exceptions (4.2 and 4.3). Word order is strongly biased toward initial, rather than final, selective-marked clauses (4.2), which supports H2 “Selectives tend to occur early in the sentence” and toward occurring at the end or after the clause they scope over, which supports H3 “Selectives tend to occur after or at the very end of their constituent” (4.4). Despite the frequent occurrence of selectives in conditional clauses, selectives tend not to be conditional subordinators (4.5), if they do not originate from subordinators (4.3.2).

However, the regularities go hand-in-hand with a very high degree of language-specific heterogeneity. Here is a just one further example. In Folopa (ppo), the selective -ta almost exclusively follows nouns and pronouns and there happens to be a lexical preference to express events related to speech with relative clauses with the head noun fo ‘talk’: yq-lo yale fo [1SG-ERG do.REL.PST talk] ‘what I said’, lit. “talk I did” (Anderson 2010: 50), which triggers the use of the selective, as in (44):

(44) Folopa

\[
\begin{array}{llllll}
\text{Take} & \text{yq-lo} & \text{díg-paae} & \text{asqe-mó} & \text{yale} & \text{fo-ta}, \\
\text{earlier/later} & \text{1SG-ERG} & \text{2PL-to write/draw-LOC/while} & \text{do.REL.PST} & \text{talk-TOP} \\
\text{ala} & \text{dowi} \\
\text{bad} & \text{custom} \\
\text{dere} & \text{whí-mó} & \text{kisipa} & \text{mutu, ó ai dowi ala} \\
\text{do.REL.PRS} & \text{man-LOC/BEN} & \text{knowledge lay.INFL} & \text{or that bad custom/way} \\
\text{era-tere} & \text{whí-mó} & \text{kisipa} & \text{mutu inipó.} \\
\text{do-thing.doing.action} & \text{man-LOC/BEN} & \text{knowledge lay.INFL do.NEG} \\
\end{array}
\]

‘Consequently, even if I wrote to you, it was not because of the one who did wrong or because of the one who had been wronged …’

(ppo-x-bible, 47007012; ConcCond1)

This idiosyncrasy accounts for the specific pattern of Folopa (ppo), different from all other languages in Figure 1 with sporadic occurrences of the selective -ta across different subordinate clause domains.
5 Discussion

5.1 Are selectives a fruitful conceptual framework?

If – as argued in Sections 1 and 3 – selectives are an arbitrary conceptual framework for crosslinguistic comparison, how fruitful has the concept been so far?

First, it can be pointed out that the notion is easily applicable from a cross-linguistic point of view and is especially well-suited for a massive parallel text corpus where even languages that so far are poorly described can be considered.

Second, it can be observed that there are micro- and macro-areal patterns that testify to the usefulness of the concept, both such that already have been associated with “topic markers” (e.g., Ga-Ewe-Gen-Fon; New Guinea) and such that do not yet figure prominently in the literature (notably Cuicatlán).

Third, the concept has proven to be relevant beyond its definitional domain. Markers that were identified on the basis of their use with personal pronouns in contrastive use also occur in non-contrastive uses and in many languages also across various subordinate clause domains (see Section 4), very much as has been suggested already by Haiman (1978).

Fourth, and most importantly, selectives have turned out to be sensitive to a number of scalar phenomena such as summarized in Table 6. Among those, (iii) Contrast is most trivial, since it at least partly follows from the definitional domain, but (i) and (ii) testify to very strong correlations with word order (irrespective of other word order typology of the languages involved), confirming or at least partly confirming Hypotheses 1 and 2 in Table 1 (see Section 2.1) and the selectives-across-subordinate-domains tendency scale (iv) suggests that there is a hierarchy of more and less prototypical contexts for selectives across subordinate clauses.

Table 6: Scales and tendency scales relevant for selectives.

<table>
<thead>
<tr>
<th>Scale name</th>
<th>Scale ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) Word order of selective constituent (4.2):</td>
<td>sentence-initial &gt;&gt;&gt; sentence-final</td>
</tr>
<tr>
<td>(ii) Word order of selective in relation to constituent (4.4):</td>
<td>absolutely final &gt;&gt;&gt; all other positions</td>
</tr>
<tr>
<td>(iii) Contrast (3)</td>
<td>more contrast &gt; less contrast</td>
</tr>
<tr>
<td>(iv) Selectives-across-subordinate-domains (4.2):</td>
<td>general relative &gt; conditional &gt; temporal &gt; concessive &gt; purpose</td>
</tr>
</tbody>
</table>

However, it cannot be denied that there are still very many open questions connected to the suggested concept, some of which will be addressed below.
5.2 Are selectives a gram type?

Dahl (2016: 435) argues that clusters of language-specific category types that he calls crosslinguistic gram types are not only comparative concepts constructed by the researcher, but “objects that are ‘out there’ to be discovered”. Members of a gram type are not the same, but just “similar enough to be put in the same bucket”. They are expected to “share a set of prototypical uses, where the probability of their use comes close to 1” (Dahl 2016: 435). The underlying idea is that prototypical uses define the semantic core of the gram type which is crosslinguistically uniform whereas grams belonging to a gram type can be rather diverse in their peripheries. Gram types can be “polycentric”, which can mean that members of a gram type can have different preferred sources of grammaticalization or that there can subclusters within the more general larger cluster.

Selectives are a challenge to the gram approach for various reasons. As pointed out in Section 3, selectives tend to be communicatively optional, despite their high text frequency, except sometimes when they fuse with other markers for specific functions such as conditional as discussed in 4.5. This suggests that selectives are perhaps more of a pragmatic rather than a strictly grammatical category type. While this study has identified more prototypical contexts of use for selectives with personal pronouns, on the one hand, and with subordinate clauses, on the other hand, it is not yet clear whether there is a semantic prototype for the set of markers irrespective of particular uses. As pointed out, a major difficulty is that selectives used with nouns are difficult to distinguish from other phenomena (all associated with definiteness). Another difficulty is that selectives tend to have an extremely wide range of use in many languages where they occur, which makes it rather difficult to make out the most prototypical contexts of use. However, it can be taken for granted that the crosslinguistic set of selectives have prototype-like (scalar) behavior at least when associated with personal pronouns and with subordinate clauses, which makes them at least similar in this respect to a gram type.

One thing that is certain is that selectives are not a monocentric gram type. In 4.3.2, we have seen that a number of selectives in different languages originate in conditional or in temporal clauses. At least in Dii and Gulay, the clausal uses clearly prevail in frequency. Thus, there is at least a distinct subtype with clausal origin (including Dii māa, Gulay baa, Khalkha Mongolian bol, Kalmyk bol-xla and Sougb kaba).

5.3 Further dimensions of variability

Throughout this paper I have emphasized the crosslinguistic heterogeneity of selectives, but this does not entail that a selective in one language is totally different from selectives in all other languages. If this were the case, the concept
would obviously not be fruitful. Rather there are a number of dimensions of variability, whose identification actually may contribute to the fruitfulness of the concept of selectives. Dimensions of variability are also useful for understanding how selectives relate to similar categories that do not meet the definitional criteria for selectives. Let us consider here a few important dimensions.

(i) *Some selectives can be sensitive to adversative meaning*: In 4.3.1, we have seen that Sochiapan Chinantec and Toura selectives prefer adversative contexts in subordinate clauses. This groups them, on the one hand, with adversative selectives, such as Dimasa *la*, as opposed to non-adversative selectives (see Sections 2.2 and 3), but also with adversative markers that have been excluded from consideration (such as Koine Greek *de*, see Section 3).

(ii) *Some selectives can be separated from subordinate clauses*: These selectives follow anaphora (see 4.4v), have anaphoric properties themselves (4.4vi) or are at the same times conjunctions (4.4vi). Among the phenomena excluded, these are similar to apodosis-introducers, such as English *then* or clausal conjunctions meaning ‘and then’.

(iii) *Some selectives are wholly or mostly restricted to NPs*: This holds for most markers at the bottom of Figure 1, which do not combine with subordinate clauses or only to subordinate clauses with nominal heads, such as generalized relative clauses. There is reason to believe that a smaller proportion of markers would have been attested in subordinate clauses if nouns rather than personal pronouns had been chosen as a definitional domain. This means that the fact that this study seems to confirm H7 (high degree of freedom of host selection of selective) is largely a consequence of how selectives have been defined.

In many Central Eurasian languages, possessive suffixes have developed topic-like uses (Yurayong 2018; for Colloquial Persian, see; Etebari et al. 2020) and these elements seem to have difficulties to spread both to personal pronouns and to subordinate clauses (i.e., they do not easily increase in freedom of host selection). Eastern Mari is a language where third person singular possessive affixes can occur as selective on personal pronouns of speech act participants. However, not in all styles and registers to the same extent. Forms, such as *me-že* [1PL-TOP/POSS.3SG], are common especially in folk poetry (see, for instance, Saarinen 1994: 220, 390). In the Eastern Mari N.T. translation, there is just a single occurrence, Example (45) with second person plural:

(45) Eastern Mari

\[\begin{array}{llll}
\text{te-že} & cylan-at & isas-šol’-ak & gaje \\
\text{2PL-TOP/POSS.3SG} & \text{all-ADD} & \text{older.brother-PART=younger.brother-PART} & \text{like} \\
\text{ul-ya} & \text{be-PRES.3SG} & \text{ul-ya} & \text{ul-ya} \\
\end{array}\]

‘(because one is your teacher,) and you are all brothers.’

(mhr-x-bible, 40023008)
The Mari example reveals that the test for selectives suggested in Section 3 is not always equally easy to apply. What to do if the definitional criterion holds, but is extremely rare? In this case I opted for not including Mari, because the use with personal pronouns was too marginally attested. It follows from this discussion that markers which do not have a high degree of freedom of host selection simply did not make it into the sample.

(iv) **Multiple selectives within a sentence**: Especially some languages of New Guinea allow for multiple selectives within the same sentence, a phenomenon that Heeschen (1994: 57, 1998: 162) and de Vries (2006: 814) call “thematizing” (a discourse preference), here illustrated in (46) from Abau with the beginning of the Parable of the Prodigal Son where selectives follow two NPs containing new (off-stage) information.

(46) Abau

\[
\begin{array}{llllll}
Uwr & \text{prueyn} & \text{hiy-kwe} & \text{ney} & \text{ho-kwe} & \text{prueys} & \text{non.} \\
\text{man} & \text{one.HUM} & \text{3.SG.M.SUBJ-TOP} & \text{child} & \text{GL.M-TOP} & \text{two.CL1 with/DU} & \\
\end{array}
\]

‘A certain man had two sons.’ Lit. “Man one, child(ren), two with”

(aau-x-bible, 42015011)

The most extreme case in the sample is Waris (see Section 4.3). Indeed, the Waris selective has a very different overall distribution, for instance, when compared to Gulay ɓaa that mainly follows subordinate clauses and has clausal origin. However, both Waris and Gulay selectives have similar uses in person pronoun contrast and on subordinate clauses.

These are just some examples of how various kinds of selectives may differ from one another, but this does not mean that the markers differ in completely unrestricted ways, otherwise there would not be any scalar effects.

### 5.4 Are there any semantic invariants?

The multiplicity of scales (Table 6) clearly testifies to the semantic and functional complexity of selectives. Let us now consider some possible semantic explanations of the selectives-across-subordinate-domains tendency scale.

Why are concessives so much less prone to have selectives than conditionals? There is no major difference in word order. In 2.4, I have suggested that selectivity (H12) is part of the explanation. Furthermore, concessives are veridical, which allows for non-hypotactic construals. In many languages, concessives are expressed mostly by adversative coordination. Adversative coordination is a natural paraphrase of concessives: although \( p, q = p, but q \). Put differently, concessives are often expressed with adversative connectives that also occur in contrast constructions. Now, contrast constructions actually favor selectives, but not in the position where the concessive relationship is expressed (double underline), as illustrated in (47) from Sougb.
In Section 4 it has also been shown that complement clauses are not a characteristic environment for selectives. This may seem strange since there is some evidence that indirect question complements may have contributed to the origin of selectives in conditional clauses in such languages as Suena and Waskia (see 4.4). However, given that complement clauses have illocutionary force, it is no surprise that selectives hardly ever become fully entrenched in complement clauses, which is consistent with H11 “Selectives tend not to mark clausal units carrying illocutionary force”. Such marginal contexts as ‘Whether he is a sinner I do not know’ (see (10) from Hills Karbi in 2.4) are most inclined to bear selectives. These are examples where the question in the complement is actually not to be answered and rather a point of departure than an illocutionary act (indicated in English by the marked initial word order). This renders such complement clauses similar to alternative concessive conditionals, such as *Whether we get any financial support or not, we will go ahead with our project* (Hauspelmahr and König 1998: 653).

As far as temporal clauses are concerned, veridicality may also be a factor why conditionals outperform temporal clauses. Another important factor is the higher degree of delimiting of conditional clauses. Temporal sequence is one of the most characteristic contexts for what Foley and Van Valin (1984) and Van Valin (2005) call co-subordination, where operators, such as tense, are shared. In Figure 1, it can be seen that especially many languages of New Guinea, where co-subordination is widespread, have selectives in conditionals, but not usually in temporal clauses. Foster (1981: 15) points out for Timbe that one of the functions of *ámâ* is that it limits the scope of tense/mood/aspect and negation to that clause. In (48), where the subordinate clause has its own TAM and polarity values, here in the only example from the concessive domain where *ámâ/amâ* is used:

(48) Timbe

<table>
<thead>
<tr>
<th>Ye</th>
<th>ámâ</th>
<th>ek-mâ</th>
<th>amâ</th>
<th>bo</th>
<th>ek-mâ</th>
<th>nângâ-mai.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2PL</td>
<td>TOP</td>
<td>see-PRS.HAB.3PL</td>
<td>TOP</td>
<td>NEG</td>
<td>see-SS</td>
<td>hear-PRS.HAB.3PL</td>
</tr>
</tbody>
</table>

‘Although you have eyes, do you not see?’

(tim-x-bible, 41008018; Concess6)
Similarly, Årsjö (2016: 255) reports that the selective in Konai may “be used as a barrier, to certain features spreading”; see also Reesink (2014: 256) for Usan. All this supports H9 “Selectives often have the function of delimiting”. However, note that delimiting is not an automatic consequence of selectives; Bickel (1993: 33, 38) discusses instances of ambiguity involving illocution and negation in Belhare. Delimiting is not restricted to clauses. Several grammar writers report that selectives preferably follow zero-marked subjects (Poppe 1951: 99 for Khalkha bol, van Gijn 2006: 114 for Yuracaré, and Cerrón-Palomino 2006: 123 for Chipaya).

However, selectives do not only delimit constituents from the rest of the sentence or clause, they also do what appears to be the opposite: linking (H[ypothesis]10) the selective constituent with the predicate or nucleus that follows. Linking is most apparent in instances where selectives anticipate properties of the nucleus predicate. In Section 3, we have seen that some languages have interrogative and imperative selectives. There are also examples for less grammaticalized instances of the same tendency. Usan uses the form eng originating from a demonstrative (Reesink 2014) as a selective. In many instances with imperative-like main predicates, this marker is preceded by /gâb/, the same-subject-form of the verb ‘see’, as in (49). The range of use is broader than the imperative selective in Fon (see Section 3), most important is that the nucleus predicate somehow engages an addressee.

(49) Usan

\[
\text{Munon ger min-goan ombur igam-a } \text{gab eng munon ger}
\]

\[
\text{man one be(?)-skin/bark two be-3SG.DS see.ss top man one}
\]

\[
\text{ue gab eng ger wo u-t-iner.}
\]

\[
\text{NEX see.ss top one 3SG him-give-UNCERTAIN.FUT.3SG}
\]

‘The one who has two tunics must share with the one who does not have one’, lit. “Seeing that a person has two shirts, seeing that a person has no shirt, give that person one.”

(wnu-x-bible, 42003011; GenRel16)

This construction suggests that conditioned commands are primarily persuasion by directing perception to crucial circumstances.

However, there are also uses of selectives where there is nothing to link to, which are apparent counterexamples to linking (H10) and to H11 “Selectives tend not to mark clausal units carrying illocutionary force”. In many languages selectives can occur sentence-finally, often with some modal connotation. Sentence-final selectives are akin to insubordination, “the conventionalized main clause use of what, on prima facie grounds, appear to be formally subordinate clauses” (Evans 2007: 367). Insubordination has the functions of “indirectionalizing, modalizing, and presuppositionalizing” (Evans 2007: 423). Object-topicalized clauses in Kayardild are actually one of Evans’ (2007: 425) examples. In Sochiapan Chinantec, sentence-final nê³ is used as a request for further clarification (Foris
In Aymara, sentence-final -xa is an attenuator “I guess, OK?” (Hardman 2001: 171). In Usan, “[c]onstructions with sentence-final e(ng) convey the speaker’s attitude of assertiveness, indignation or surprise” (Reesink 2014: 249). Also, at least in some languages, there are referent-plus-selective-questions, such as Japanese Jun kun wa? ‘What about Jun-kun?’ (Tanaka 2015) or Tena Lowland Quichua Jenny=ga? ‘What about Jenny? Where is Jenny?’ (Grzech 2016: 40), whose cross-linguistic distribution unfortunately cannot be assessed in the N.T., where there are no such examples. These are an interesting type of turns without overt predicates. They provide counterevidence to H1 that selectives are units of the sentence. However, it might be argued that in such “what-about”-questions the predicate is simply omitted, and that if it is not, as in (50) from Nalca, the illocution follows the selective constituent, as expected.

(50) Nalca

...ban be-nye-k a-na·ra, hek-a’ da-nam-la do?

this M-DEM-DAT DP-DEM-TOP which-thing become-FUT-3SG Q

‘[Lord], but what about this one?’, lit. “this one, what will it become?”

The questions as to how selectives relate to the domain of the sentence or any related unit and the extent to which selective constituents can bear illocutionary force needs further investigation.

6 Conclusions

This study demonstrates that the elements often referred to as “topic markers” in the descriptive literature can be subject to a typological investigation that can take into account their similarities and differences at the same time if the items considered are defined in a narrow context of use in form of a Lazardian arbitrary conceptual framework and if parallel text examples, which allow for filling descriptive gaps in an “inconvenience sample”, are used.

This investigation has focused on grammaticalized selectives with a high degree of freedom of host selection, which mark constituents of different kinds (nouns, pronouns, adverbs, and clauses): the definitional domain is pronominal (Section 3), and the domains surveyed are clausal (Section 4). It has been found that selectives are distributed across subordinate clause domains following a tendency scale where semantically more selective and non-veridical clause types – notably general relative clauses and conditional clauses – have the highest propensity of being combined with selectives.
However, there are also very strong correlations with word order, which suggests that selectives are highly sensitive to the linear flow of information in discourse. Selective constituents tend to occur early in the sentence but tend to follow the constituent they have scope over, which is motivated by ease of scope disambiguation, given that selectives can mark constituents of very different length.

Selectives can combine with a large range of different types of grammatical features both of the selective constituent, the nucleus (notably the illocution of the nucleus) and the larger construction (Section 3 and Appendix B), which suggests that selectives are not elements of a detached “information structure” layer of language structure, but rather multifunctional markers reflecting many different facets of grammar and discourse at the same time and in language-particular ways. However, the scalar effects identified clearly demonstrate that selectives also have highly similar functions across languages.

Abbreviations in glosses

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>=</td>
<td>hyphen(-) in orthography</td>
</tr>
<tr>
<td>1</td>
<td>first person</td>
</tr>
<tr>
<td>2</td>
<td>second person</td>
</tr>
<tr>
<td>3</td>
<td>third person</td>
</tr>
<tr>
<td>ABL</td>
<td>ablative</td>
</tr>
<tr>
<td>ACC</td>
<td>accusative</td>
</tr>
<tr>
<td>ADD</td>
<td>additive</td>
</tr>
<tr>
<td>ADDR</td>
<td>address</td>
</tr>
<tr>
<td>ADJ</td>
<td>adjective</td>
</tr>
<tr>
<td>ADS</td>
<td>adversative</td>
</tr>
<tr>
<td>ADV</td>
<td>adverbial</td>
</tr>
<tr>
<td>AFF</td>
<td>affirmative</td>
</tr>
<tr>
<td>AGT</td>
<td>agentive</td>
</tr>
<tr>
<td>AGT.N</td>
<td>agentive noun</td>
</tr>
<tr>
<td>ALL</td>
<td>allative</td>
</tr>
<tr>
<td>AN</td>
<td>animate</td>
</tr>
<tr>
<td>ANAPH</td>
<td>anaphoric</td>
</tr>
<tr>
<td>BEN</td>
<td>benefactive</td>
</tr>
<tr>
<td>CL</td>
<td>classifier</td>
</tr>
<tr>
<td>CNTRFC</td>
<td>counterfactual</td>
</tr>
<tr>
<td>COMP</td>
<td>complementizer</td>
</tr>
<tr>
<td>COND</td>
<td>conditional</td>
</tr>
<tr>
<td>CONEG</td>
<td>conegative</td>
</tr>
<tr>
<td>Conj</td>
<td>conjunction</td>
</tr>
<tr>
<td>CONT</td>
<td>continuous aspect</td>
</tr>
<tr>
<td>COP</td>
<td>copula</td>
</tr>
<tr>
<td>COREF</td>
<td>coreferential</td>
</tr>
<tr>
<td>CTOP</td>
<td>counter-expective selective</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Meaning</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
</tr>
<tr>
<td>CVB</td>
<td>converb</td>
</tr>
<tr>
<td>DAT</td>
<td>dative</td>
</tr>
<tr>
<td>DEF</td>
<td>definiteness, definite article</td>
</tr>
<tr>
<td>DEM</td>
<td>demonstrative</td>
</tr>
<tr>
<td>DES</td>
<td>desiderative</td>
</tr>
<tr>
<td>DIR</td>
<td>direct</td>
</tr>
<tr>
<td>DP</td>
<td>default phrase gender</td>
</tr>
<tr>
<td>DS</td>
<td>different subject</td>
</tr>
<tr>
<td>DU</td>
<td>dual</td>
</tr>
<tr>
<td>ELA</td>
<td>elative</td>
</tr>
<tr>
<td>EMPH</td>
<td>emphatic</td>
</tr>
<tr>
<td>EQU</td>
<td>equational</td>
</tr>
<tr>
<td>ERG</td>
<td>ergative</td>
</tr>
<tr>
<td>EX</td>
<td>existence</td>
</tr>
<tr>
<td>EXCL</td>
<td>exclusive</td>
</tr>
<tr>
<td>EV</td>
<td>evidence</td>
</tr>
<tr>
<td>EVID</td>
<td>evidential</td>
</tr>
<tr>
<td>F</td>
<td>feminine</td>
</tr>
<tr>
<td>FOC</td>
<td>focus</td>
</tr>
<tr>
<td>FUT</td>
<td>future tense</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive</td>
</tr>
<tr>
<td>GL</td>
<td>general</td>
</tr>
<tr>
<td>HAB</td>
<td>habitual</td>
</tr>
<tr>
<td>HON</td>
<td>honorific</td>
</tr>
<tr>
<td>HUM</td>
<td>human</td>
</tr>
<tr>
<td>HYP</td>
<td>hypothetical</td>
</tr>
<tr>
<td>ILL</td>
<td>illative</td>
</tr>
<tr>
<td>IMP</td>
<td>imperative</td>
</tr>
<tr>
<td>INCL</td>
<td>inclusive</td>
</tr>
<tr>
<td>INDF</td>
<td>indefinite</td>
</tr>
<tr>
<td>INF</td>
<td>infinitive</td>
</tr>
<tr>
<td>INFL</td>
<td>inflection</td>
</tr>
<tr>
<td>INS</td>
<td>instrumental</td>
</tr>
<tr>
<td>INT</td>
<td>interrogative</td>
</tr>
<tr>
<td>INV</td>
<td>inversion marker</td>
</tr>
<tr>
<td>IPV</td>
<td>imperfective</td>
</tr>
<tr>
<td>IRR</td>
<td>irrealis</td>
</tr>
<tr>
<td>LOC</td>
<td>locative</td>
</tr>
<tr>
<td>M</td>
<td>masculine</td>
</tr>
<tr>
<td>MED</td>
<td>medial</td>
</tr>
<tr>
<td>N.</td>
<td>non-</td>
</tr>
<tr>
<td>NEG</td>
<td>negation</td>
</tr>
<tr>
<td>NEX</td>
<td>negative existence</td>
</tr>
<tr>
<td>NMLZ</td>
<td>nominalizer</td>
</tr>
<tr>
<td>Q</td>
<td>question particle/ marker</td>
</tr>
<tr>
<td>PART</td>
<td>particle</td>
</tr>
<tr>
<td>PERF</td>
<td>perfect</td>
</tr>
</tbody>
</table>
Acknowledgments: I am extremely grateful to two very careful anonymous reviewers, who must have spent a lot of time to help me improve this paper, to Carl Börstell, Östen Dahl, Karolina Grzech, John Haiman, Dejan Matić, Dmitry Nikolaev and Bruno Olsson for many highly useful comments on earlier versions of this work and to Volker Gast, Ann Kelly, Dominique Knuchel, Maria Koptjevskaja-Tamm, Christian Lehmann, Thomas Mayer and Robert Östling for the discussion of particular matters. Particular language data have been discussed with Sebastian Fedden (Mian), Nailya Philippova (Amri Karbi) and Stephanie Yam (Motu). Many thanks to Carl Börstell for the maps.
Appendix A: Sample of languages with selectives considered in this study
(81 languages)

ATLANTIC-CONGO, CAMEROUN-UBANGIAN: Dii dur; GA-DANGME: Ga gaa; GBE: Ewe ewe, Gen gej, Fon fon
MANDE: MANO-DAN: Toura (Côte d'Ivoire) neb
CENTRAL SUDANIC, BAGIRMIC: Kenga kyi, SARAIC: Gulay gvl
SONGHAY: Zarma-Kaado dje
AFRO-ASIATIC, CHADIC: Gude gde, Parkwa pbi, Matal mfh, Zulgo-Gemzek gnd, Merey meq, North Mofu mfk, Muyang muy
MONGOLIAN: Khalkha Mongolian khk, Oirad-Kalmyk xal
KOREAN: Korean kor
JAPONIC: Japanese jpn
AUSTRO-NEUROPEAN, CENTRAL LUZON: Botolan Sambal sbl, Bolinao smk; PAPUAN TIP: Iduna viv, Motu meu
EAST BIRD’S HEAD: Meyah mej, Moskona mtj, Sougb mnx
YAWA-SAWERU: Yawa yva
BORDER: Waris wrs, Amanab amn
SEPIK: Abau aau
PAUWI: Karkar-Yuri yuj
LEFT MAY: Ama (Papua New Guinea) amm
NUCLEAR TRANS NEW GUINEA, MEK: Nalca nlc, Una mtg; OK: Yonggom yon, Mian mpt;
MADANG: Waskia wsk, Usan wnu; FINISTERRE-SARUWAGED: Rawa rwo, Iyo nca;
KAINANTU-GOROKA: Siane snp; WESTERN HUON: Timbe tim, Borong ksr, Burum-Mindik bmw; BINANDEREAN: Suena sue, Zia zia, Korafe-Yegha kpr
EAST STRICKLAND: Konai kxw
TEBERAN: Folopa: ppo
MANUBARAN: Doromu-Koki kqc
DAGAN: Umanakaina gdn
PAMA-NYUNGAN, DESERT NYUNGIC: WIK: Wik Mungkan wim; Warlpiri wbp;
OTO-MANGUEAN, CHINANTECAN: Sochiapan Chinantec cso; MAZATECAN: Chiquihuitlán Mazatec maq; MIXTEC-CUICATEC: Teutila Cuicatec cut, Tepeuxila Cuicatec cux;
ZAPOTECAN: Sierra de Juárez Zapotec zaa
ARAWAKAN, CENTRAL MAIPURAN: Parecis pab, Western Maipuran: Yanesha’ ame,
ARAUAN: Culina cul
PÁEZ: Páez pbb
BARBACOAN: Awa-Cuaiquer kwi
QUECHUAN, QUECHUA I: Huallaga Huánuco Quechua qub; QUECHUA II: Lambayeque Quechua quf Quechuan, Cajamarca Quechua qvc, Imbabura Highland Quechua qvi
AYMARAN: Central Aymara ayr
YURACARÉ: Yuracaré yuz
URU-CHIPAYA: Chipaya cap
The classification follows Hammarström et al. (2020)

Appendix A2: Maps

Map 1: Languages of the sample with selectives in Africa

Map 2: Languages of the sample with selectives in Eurasia
Map 3: Languages of the sample with selectives in New Guinea and Australia

Map 4: Languages of the sample with selectives in Meso- and South America
Appendix B: Languages and selectives sampled

*: selectives not attested in the search domain with contrastive use of second person plural personal pronouns

<table>
<thead>
<tr>
<th>Language</th>
<th>ISO</th>
<th>Markers</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dii</td>
<td>dur</td>
<td>TIME PST mâa, {TIME FUT tèe*}</td>
<td>Bohnhoff (2010)</td>
</tr>
<tr>
<td>Ga</td>
<td>gaa</td>
<td>le</td>
<td>Campbell (2017)</td>
</tr>
<tr>
<td>Gen</td>
<td>gej</td>
<td>-a</td>
<td>Jondoh (1980)</td>
</tr>
<tr>
<td>Fon</td>
<td>fon</td>
<td>ṭ, IMP hān* ‘thus, certainly’</td>
<td>Lefebvre and Brousseau (2002), Höftmann and Ahohounkpanzon (2003), Hofmann (1977), Delafosse (1894)</td>
</tr>
<tr>
<td>Kenga</td>
<td>kyg</td>
<td>se; ADS(?) le; jaay</td>
<td>Vandame (1968), Palayer (2004)</td>
</tr>
<tr>
<td>Gulay</td>
<td>gvl</td>
<td>baa</td>
<td>Djarangar (1989)</td>
</tr>
<tr>
<td>Zarma-Kaado</td>
<td>dje</td>
<td>wo, -o</td>
<td>Bernard and White-Kaba (1994)</td>
</tr>
<tr>
<td>Gude</td>
<td>gde</td>
<td>ma(-a)</td>
<td>Hoskison (1983)</td>
</tr>
<tr>
<td>Parkwa</td>
<td>pbi</td>
<td>laki, INT mâ</td>
<td>Lagona and Jarvis (1992)</td>
</tr>
<tr>
<td>Matal</td>
<td>mfh</td>
<td>kà</td>
<td>Verdizade (2018)</td>
</tr>
<tr>
<td>Zulgo-Gemzek</td>
<td>gnd</td>
<td>ka</td>
<td>Haller (1986)</td>
</tr>
<tr>
<td>Meray</td>
<td>meq</td>
<td>na</td>
<td>Gravina (2007), Gravina et al. (2003)</td>
</tr>
<tr>
<td>Muyang</td>
<td>muf</td>
<td>ti, ADS nahama</td>
<td>Smith (2003a, 2003b)</td>
</tr>
<tr>
<td>Khalkha</td>
<td>khk</td>
<td>bol &lt; bol-vol [be-COND.CVB]; (-)n*n [POSS.3]</td>
<td>Poppe (1951)</td>
</tr>
<tr>
<td>Mongolian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Korean</td>
<td>kor</td>
<td>-nun</td>
<td>Koo (1999)</td>
</tr>
<tr>
<td>Japanese</td>
<td>jpn</td>
<td>/wa/ &lt;ha&gt;</td>
<td>Shibatani (1990), Heycock (2008)</td>
</tr>
<tr>
<td>Dimasa</td>
<td>dis</td>
<td>de, ADS la</td>
<td>Jacquessson (2008), Dundas (1908)</td>
</tr>
<tr>
<td>Language</td>
<td>ISO</td>
<td>Markers</td>
<td>Source</td>
</tr>
<tr>
<td>---------------</td>
<td>------</td>
<td>--------------------------------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>Southern</td>
<td>kac</td>
<td><strong>gaw</strong>, <strong>ADS chyawm gaw</strong>, <strong>ADS mahtang</strong></td>
<td>Cheung (2007)</td>
</tr>
<tr>
<td>Jinghpaw</td>
<td>cnh</td>
<td><strong>cu</strong> <strong>OBL cun</strong>; <strong>KNOWN kha</strong>, <strong>OBL khan</strong></td>
<td>Peterson (2003), Barnes (1998)</td>
</tr>
<tr>
<td>Haka Chin</td>
<td>cnw</td>
<td><strong>cu</strong> <strong>OBL cun</strong></td>
<td></td>
</tr>
<tr>
<td>Ngawn Chin</td>
<td>csy</td>
<td><strong>sia</strong></td>
<td>Stern (1963)</td>
</tr>
<tr>
<td>Sizang Chin</td>
<td>acn</td>
<td><strong>gas</strong></td>
<td>Sampu et al. (2005)</td>
</tr>
<tr>
<td>Longchuan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Achang</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lahu</td>
<td>lhu</td>
<td><strong>leh</strong></td>
<td>Matisoff (1973, 1988, 2006)</td>
</tr>
<tr>
<td>Hills Karbi</td>
<td>mjuw</td>
<td>(-)ke</td>
<td>Konnerth (2014), Grüßner (1978)</td>
</tr>
<tr>
<td>Amri Karbi</td>
<td>ajz</td>
<td>(-)ke</td>
<td>N. Philippova (p.c.)</td>
</tr>
<tr>
<td>Botolan Sambal</td>
<td>sbl</td>
<td><strong>hay</strong></td>
<td>Antworth (1979)</td>
</tr>
<tr>
<td>Bolinao</td>
<td>smk</td>
<td><strong>PL say</strong>, <strong>SG si</strong>, <strong>PRO si</strong></td>
<td>Viray and Scheerer (1929)</td>
</tr>
<tr>
<td>Iduna</td>
<td>viv</td>
<td>(-)ma</td>
<td>Loving (1976), Huckett (1976)</td>
</tr>
<tr>
<td>Motu</td>
<td>meu</td>
<td><strong>na</strong>, <strong>ADDRESS.(A)SYMMETRY be</strong>, <strong>ERG ese</strong></td>
<td>Lawes (1896), Yam (2020)</td>
</tr>
<tr>
<td>Meyah</td>
<td>mej</td>
<td><strong>bera</strong></td>
<td>Gravelle (2004)</td>
</tr>
<tr>
<td>Moskona</td>
<td>mtj</td>
<td><strong>era</strong></td>
<td>Gravelle (2010)</td>
</tr>
<tr>
<td>Sougb</td>
<td>mnx</td>
<td><strong>kaba</strong>, [then]; <strong>gin(i)</strong>, <strong>kaba</strong>, [this then]; <strong>gin(i)</strong> [this]</td>
<td>Reesink (2002)</td>
</tr>
<tr>
<td>Yawa</td>
<td>yva</td>
<td><strong>-amo</strong></td>
<td>Jones (2003)</td>
</tr>
<tr>
<td>Waris</td>
<td>wrs</td>
<td><strong>-mba</strong>, <strong>-ba</strong>*, <strong>-va</strong>*, <strong>-pa</strong>*, <strong>RESUMPTIVE -oa</strong></td>
<td>Brown (1990)</td>
</tr>
<tr>
<td>Amanab</td>
<td>amn</td>
<td><strong>-ba</strong>, <strong>PL.AGT -bi</strong>, <strong>eba</strong>, <strong>-ba/bi eba</strong></td>
<td>Mlinch (1992)</td>
</tr>
<tr>
<td>Abau</td>
<td>aau</td>
<td><strong>-kwe</strong></td>
<td>Lock (2011)</td>
</tr>
<tr>
<td>Karkar-Yuri</td>
<td>yuj</td>
<td><strong>l-e</strong> <strong>TOP-HERE.WITHIN</strong></td>
<td>Rigden (1985)</td>
</tr>
<tr>
<td>Ama (Papua New</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guinea)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nalca</td>
<td>nlc</td>
<td><strong>dara</strong>, <strong>PRON also -da</strong>, <em><em>NP bera</em>/ara</em> <strong>+GENDER</strong></td>
<td>Wälchli (2018a)</td>
</tr>
<tr>
<td>Una</td>
<td>mtg</td>
<td><strong>PRON -da</strong>, <strong>NP ara</strong>*, <strong>FN bira</strong></td>
<td>Louwse (1988)</td>
</tr>
<tr>
<td>Yonggom</td>
<td>yon</td>
<td><strong>kuu</strong>, <strong>LOC/ERG kumbed</strong></td>
<td>Christensen (2010)</td>
</tr>
<tr>
<td>Mian</td>
<td>mpt</td>
<td><strong>le</strong></td>
<td>Fedden (2011)</td>
</tr>
<tr>
<td>Waskia</td>
<td>wsk</td>
<td><strong>mu</strong></td>
<td>Ross and Paol (1978), Lee and Barker (1985)</td>
</tr>
<tr>
<td>Usan</td>
<td>wnu</td>
<td><strong>e-ng</strong> [this-GIVEN] <em>that</em>, <strong>eng-u</strong> [that-HESITATIVE]</td>
<td>Reesink (1984, 2014)</td>
</tr>
<tr>
<td>Rawa</td>
<td>rwo</td>
<td><strong>ngu</strong></td>
<td>Toland and Toland (1991)</td>
</tr>
<tr>
<td>Iyo</td>
<td>nca</td>
<td><strong>-ko</strong></td>
<td>Minter (2009)</td>
</tr>
<tr>
<td>Siane</td>
<td>snp</td>
<td><strong>ya</strong> <strong>DEM ya-te</strong> [<strong>DEM-ERG.PL</strong>] <strong>ya-kafo</strong> [<strong>DEM-ERG.SG</strong>]</td>
<td>Potts and James (1983)</td>
</tr>
<tr>
<td>Language</td>
<td>ISO</td>
<td>Markers</td>
<td>Source</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----</td>
<td>-----------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Timbe</td>
<td>tim</td>
<td>ámbí/amá</td>
<td>Foster (1981, n.d.)</td>
</tr>
<tr>
<td>Borong</td>
<td>ksr</td>
<td>-noŋ</td>
<td>Olkkonen and Olkkonen (2000)</td>
</tr>
<tr>
<td>Burum Mindik</td>
<td>bmu</td>
<td>-nön*/-an*/-ōn</td>
<td>Olkkonen and Olkkonen (2007), Olkkonen (1990)</td>
</tr>
<tr>
<td>Korafe-Yegha</td>
<td>kpr</td>
<td>-mo, amo* [that.top], prox emo* [this.top]</td>
<td>Farr (1999); Farr and Farr (1975)</td>
</tr>
<tr>
<td>Konai</td>
<td>kxw</td>
<td>pron -me, -be*</td>
<td>Årsjö (2016)</td>
</tr>
<tr>
<td>Doromu-Koki</td>
<td>kqc</td>
<td>bi</td>
<td>Bradshaw (2012)</td>
</tr>
<tr>
<td>Umanakaina</td>
<td>gdn</td>
<td>ka</td>
<td>Evensen (1993)</td>
</tr>
<tr>
<td>Wik Mungkan</td>
<td>wim</td>
<td>def -an, “top” -iy, seq =a,</td>
<td>Kilham et al. (1986)</td>
</tr>
<tr>
<td>Warlpiri</td>
<td>wbp</td>
<td>-ju/-ji*</td>
<td>Nash (1980), Hall (1976)</td>
</tr>
<tr>
<td>Sochiapan</td>
<td>cso</td>
<td>nē³, ctōp mā¹ nā¹; jáun² nē³*</td>
<td>Foris (2000)</td>
</tr>
<tr>
<td>Chinantec</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chiquihuitán</td>
<td>maq</td>
<td>ne</td>
<td>Jamieson Capen (1996)</td>
</tr>
<tr>
<td>Mazatec</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teutila Cuicatec</td>
<td>cut</td>
<td>ne</td>
<td></td>
</tr>
<tr>
<td>Cuicatec</td>
<td>cux</td>
<td>ní, tuu’mi ni* [then top]</td>
<td>Anderson and Roque (1983), Bradley (1991)</td>
</tr>
<tr>
<td>Sierra de Juárez</td>
<td>zaa</td>
<td>nna</td>
<td>Nellis and Nellis (1983)</td>
</tr>
<tr>
<td>Zapotec</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parecis</td>
<td>pab</td>
<td>_ atyo _</td>
<td>Rowan and Burgess 2008 (2008)</td>
</tr>
<tr>
<td>Culina</td>
<td>cul</td>
<td>m -pa, r -pi</td>
<td>Tiss (2004)</td>
</tr>
<tr>
<td>Awa-Cuaiquer</td>
<td>kwi</td>
<td>-na &lt;-ne&gt;</td>
<td>Curnow (1997)</td>
</tr>
<tr>
<td>Quechua</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lambayeque</td>
<td>quf</td>
<td>-qa, top+dir.ev -qami, ads =shuypqaqami, =shuypaqami</td>
<td>Shaver (1996)</td>
</tr>
<tr>
<td>Imbabura High-land Quichua</td>
<td>qvi</td>
<td>-ca</td>
<td>Cole (1985)</td>
</tr>
<tr>
<td>Central Aymara</td>
<td>ayr</td>
<td>-x, -xa*</td>
<td>Hardman (2001)</td>
</tr>
<tr>
<td>Yuracaré</td>
<td>yuz</td>
<td>-ja</td>
<td>van Gijn (2006)</td>
</tr>
<tr>
<td>Chipaya</td>
<td>cap</td>
<td>-ki</td>
<td>Cerrón-Palomino (2006)</td>
</tr>
</tbody>
</table>
Appendix C: Search strings for second person plural personal pronoun with selective

<table>
<thead>
<tr>
<th>Language</th>
<th>String</th>
</tr>
</thead>
<tbody>
<tr>
<td>dur-x-bible</td>
<td>í ví míáa</td>
</tr>
<tr>
<td>gaa-x-bible-newworld</td>
<td>nyc le</td>
</tr>
<tr>
<td>ewe-x-bible</td>
<td>miawo la</td>
</tr>
<tr>
<td>gej-x-bible</td>
<td>kə-a</td>
</tr>
<tr>
<td>fon-x-bible</td>
<td>midɛ lɛɛ ñ</td>
</tr>
<tr>
<td>neb-x-bible</td>
<td>kale laá</td>
</tr>
<tr>
<td>kyuq-x-bible</td>
<td>naase se</td>
</tr>
<tr>
<td>gvl-x-bible</td>
<td>sei baa</td>
</tr>
<tr>
<td>dje-x-bible</td>
<td>araŋ wo</td>
</tr>
<tr>
<td>gde-x-bible</td>
<td>ma una</td>
</tr>
<tr>
<td>pbi-x-bible</td>
<td>waku laki</td>
</tr>
<tr>
<td>mfh-x-bible</td>
<td>akul kà</td>
</tr>
<tr>
<td>gnd-x-bible</td>
<td>kurum ka</td>
</tr>
<tr>
<td>meq-x-bible</td>
<td>nakurom na</td>
</tr>
<tr>
<td>mfk-x-bible</td>
<td>kuray na</td>
</tr>
<tr>
<td>muy-x-bible</td>
<td>lekulum ti</td>
</tr>
<tr>
<td>khk-x-bible</td>
<td>ta nar bol</td>
</tr>
<tr>
<td>xal-x-bible</td>
<td>тəдн болхла</td>
</tr>
<tr>
<td>kor-x-bible-latinscript</td>
<td>nehuyunun</td>
</tr>
<tr>
<td>jpn-rom</td>
<td>anatatagataha</td>
</tr>
<tr>
<td>dis-x-bible</td>
<td>nisi de</td>
</tr>
<tr>
<td>kac-x-bible-common</td>
<td>nanhte gaw</td>
</tr>
<tr>
<td>cnh-x-bible</td>
<td>nannih cu</td>
</tr>
<tr>
<td>cnw-x-bible</td>
<td>nohte cun</td>
</tr>
<tr>
<td>csw-x-bible</td>
<td>note sia</td>
</tr>
<tr>
<td>acn-x-bible</td>
<td>namoq gas</td>
</tr>
<tr>
<td>atb-x-bible</td>
<td>nungmoq gi</td>
</tr>
<tr>
<td>lhu-x-bible</td>
<td>naw hu leh</td>
</tr>
<tr>
<td>mjw-x-bible</td>
<td>nangtum ke</td>
</tr>
<tr>
<td>ajz-x-bible</td>
<td>nali ke</td>
</tr>
<tr>
<td>sbl-x-bible</td>
<td>hikawo</td>
</tr>
<tr>
<td>smk-x-bible</td>
<td>si’kamo</td>
</tr>
<tr>
<td>viv-x-bible</td>
<td>omi’iyao ma</td>
</tr>
<tr>
<td>meu-x-bible</td>
<td>umui na</td>
</tr>
<tr>
<td>mej-x-bible</td>
<td>iwa bera</td>
</tr>
<tr>
<td>mtj-x-bible</td>
<td>yuwa era</td>
</tr>
<tr>
<td>mnx-x-bible</td>
<td>yen, kaba</td>
</tr>
<tr>
<td>yva-x-bible</td>
<td>weapamo</td>
</tr>
<tr>
<td>Code</td>
<td>Meaning</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>wrs-x-bible</td>
<td>yimba</td>
</tr>
<tr>
<td>amn-x-bible</td>
<td>nengel eba</td>
</tr>
<tr>
<td>aau-x-bible</td>
<td>homkwe</td>
</tr>
<tr>
<td>yuj-x-bible</td>
<td>yumo te</td>
</tr>
<tr>
<td>amm-x-bible</td>
<td>moisu mo</td>
</tr>
<tr>
<td>nlc-x-bible</td>
<td>ugunuk dara</td>
</tr>
<tr>
<td>mtg-x-bible</td>
<td>sunda</td>
</tr>
<tr>
<td>yon-x-bible</td>
<td>yiib kuu</td>
</tr>
<tr>
<td>mpt-x-bible</td>
<td>ibole</td>
</tr>
<tr>
<td>wsk-x-bible</td>
<td>nina mu</td>
</tr>
<tr>
<td>wnu-x-bible</td>
<td>an eng</td>
</tr>
<tr>
<td>rwo-x-bible-karo</td>
<td>ye ngu</td>
</tr>
<tr>
<td>nca-x-bible</td>
<td>yeko</td>
</tr>
<tr>
<td>snp-x-bible-komongu</td>
<td>linate ya</td>
</tr>
<tr>
<td>tim-x-bible</td>
<td>ye amâ</td>
</tr>
<tr>
<td>ksr-x-bible</td>
<td>ononoŋ</td>
</tr>
<tr>
<td>bmw-x-bible</td>
<td>enön</td>
</tr>
<tr>
<td>sue-x-bible</td>
<td>nikare awa</td>
</tr>
<tr>
<td>zia-x-bible</td>
<td>niye awiya</td>
</tr>
<tr>
<td>kpr-x-bible</td>
<td>nemo</td>
</tr>
<tr>
<td>kww-x-bible</td>
<td>njme</td>
</tr>
<tr>
<td>ppo-x-bible</td>
<td>diŋta</td>
</tr>
<tr>
<td>kqc-x-bible</td>
<td>ya bi</td>
</tr>
<tr>
<td>gdn-x-bible</td>
<td>wi ka</td>
</tr>
<tr>
<td>wim-x-bible</td>
<td>niiyan-i-a</td>
</tr>
<tr>
<td>wbp-x-bible</td>
<td>nyurrurlakuju</td>
</tr>
<tr>
<td>csn-x-bible</td>
<td>hnoh⁻² ne³</td>
</tr>
<tr>
<td>maq-x-bible</td>
<td>ngayun ne</td>
</tr>
<tr>
<td>cut-x-bible</td>
<td>nchuñh ne</td>
</tr>
<tr>
<td>cux-x-bible</td>
<td>ndiš'ii ní</td>
</tr>
<tr>
<td>zaa-x-bible</td>
<td>lebi'i nna</td>
</tr>
<tr>
<td>pab-x-bible</td>
<td>xiso atyo</td>
</tr>
<tr>
<td>ame-x-bible</td>
<td>sañapa’</td>
</tr>
<tr>
<td>cul-x-bible</td>
<td>tiadenidsapi</td>
</tr>
<tr>
<td>pbb-x-bible</td>
<td>i’cue’sha’</td>
</tr>
<tr>
<td>kwi-x-bible</td>
<td>une</td>
</tr>
<tr>
<td>qub-x-bible</td>
<td>gamcunaga</td>
</tr>
<tr>
<td>quf-x-bible</td>
<td>qamkunaqami</td>
</tr>
<tr>
<td>qvc-x-bible</td>
<td>qamkunaqam</td>
</tr>
<tr>
<td>qvi-x-bible</td>
<td>cangunaca</td>
</tr>
<tr>
<td>ayr-x-bible-2011</td>
<td>jumanakax</td>
</tr>
<tr>
<td>yuz-x-bible</td>
<td>paaja</td>
</tr>
<tr>
<td>cap-x-bible-2004</td>
<td>ançhukki</td>
</tr>
</tbody>
</table>
Appendix D: Ten top verses for second person plural plus selective

<table>
<thead>
<tr>
<th>Verse No</th>
<th>Text (English, Lexham Bible)</th>
<th>No of lgs</th>
<th>Perc</th>
</tr>
</thead>
<tbody>
<tr>
<td>43008023</td>
<td>and he said to them, “you are from below; i am from above. you are from this world; i am not from this world.</td>
<td>71</td>
<td>88%</td>
</tr>
<tr>
<td>60002009</td>
<td>but you are a chosen race, a royal priesthood, a holy nation, a people for god’s possession, so that you may proclaim the virtues of the one who called you out of darkness into his marvelous light,</td>
<td>69</td>
<td>85%</td>
</tr>
<tr>
<td>46004010</td>
<td>we are fools for the sake of christ, but you are prudent in christ! we are weak, but you are strong! you are honored, but we are dishonored!</td>
<td>63</td>
<td>78%</td>
</tr>
<tr>
<td>62004004</td>
<td>you are from god, little children, and have conquered them, because the one who is in you is greater than the one who is in the world.</td>
<td>57</td>
<td>70%</td>
</tr>
<tr>
<td>43014019</td>
<td>yet a little time and the world will see me no longer, but you will see me; because i live, you also will live.</td>
<td>56</td>
<td>69%</td>
</tr>
<tr>
<td>43015005</td>
<td>“i am the vine; you are the branches. the one who remains in me and i in him — this one bears much fruit, for apart from me you are not able to do anything.</td>
<td>55</td>
<td>68%</td>
</tr>
<tr>
<td>46003009</td>
<td>for we are god’s fellow workers; you are god’s field, god’s building.</td>
<td>54</td>
<td>67%</td>
</tr>
<tr>
<td>40023008</td>
<td>but you are not to be called ‘rabi,’ because one is your teacher, and you are all brothers,</td>
<td>53</td>
<td>66%</td>
</tr>
<tr>
<td>52005004</td>
<td>but you, brothers, are not in the darkness, so that the day should catch you like a thief,</td>
<td>52</td>
<td>64%</td>
</tr>
<tr>
<td>62002020</td>
<td>and you have an anointing from the holy one, and you all know.</td>
<td>51</td>
<td>63%</td>
</tr>
</tbody>
</table>

Appendix E: List of domains and clauses

The initial number indicates the number of languages in the sample that this clause has been coded for (maximum of 81). The second number is the verse number in the N.T.

GenRel
81 43007018 The one who speaks from himself seeks his own glory. But the one who seeks the glory of the one who sent him — this one is true, and there is no unrighteousness in him.
81 40023011 And the greatest among you will be your servant.
73 43010002 But the one who enters through the door is the shepherd of the sheep.
81 41009040 For whoever is not against us is for us.
80 46006017 But the one who joins himself to the Lord is one spirit with him.
81 40016025 For whoever wants to save his life will lose it, but whoever loses his life on account of me will find it.
Now those who ate were about five thousand men, in addition to women and children.

 Whoever seeks to preserve his life will lose it, but whoever loses it will keep it. It will not be like this among you! But whoever wants to become great among you must be your servant, The one who believes in the Son has eternal life, but the one who disobeys the Son will not see life — but the wrath of God remains on him.

The one who believes and is baptized will be saved, but the one who refuses to believe will be condemned.

And when he heard it, he said, “Those who are healthy do not have need of a physician, but those who are sick.” For everyone who asks receives, and the one who seeks finds, and to the one who knocks it will be opened. And whoever is not offended by me is blessed.”

The one who has two tunics must share with the one who does not have one, and the one who has food must do likewise. For whoever has, more will be given to him, and whoever does not have, even what he has will be taken away from him.

And if anyone says anything to you, ‘Why are you untying it?’ you will say this: ‘The Lord has need of it.’"
Then Mary, when she came where Jesus was and saw him, fell at his feet, saying to him, “Lord, if you had been here, my brother would not have died.”

and you say, ‘If we had lived in the days of our fathers, we would not have been partners with them in the blood of the prophets!’

If you have known me, you will know my Father also. And from now on you know him and have seen him.

If you were of the world, the world would love its own. But because you are not of the world, but I chose you out of the world, for this reason the world hates you.

Jesus said to them, “If you were blind, you would not have sin. But now you say, ‘We see,’ your sin remains.

But if it is from God, you will not be able to overthrow them, lest you even be found fighting against God.” So they were persuaded by him.

And saying, “If you are the king of the Jews, save yourself!”

And if Satan expels Satan, he is divided against himself. How then will his kingdom stand?

So the demons implored him, saying, “If you are going to expel us, send us into the herd of pigs.”

Truly, truly I say to you, unless a grain of wheat falls into the earth and dies, it remains by itself alone. But if it dies, it bears much fruit.

But I tell you the truth, it is better for you that I go away. For if I do not go away, the Advocate will not come to you; but if I go, I will send him to you.

Jesus replied to him, “If I have spoken wrongly, testify about what is wrong! But if I have spoken correctly, why do you strike me?”

And if he happens to find it, truly I say to you that he rejoices over it more than over the ninety-nine that did not wander away.

for she said to herself, “If only I touch his cloak I will be healed.”

But if Timothy comes, see that he is with you without cause to fear, for he is carrying out the Lord’s work, as I also am.

But if we say, ‘From men,’ all the people will stone us to death, because they are convinced that John was a prophet.”

Or also if he will ask for a fish, will give him a snake?

Who among you convicts me concerning sin? If I am telling the truth, why do you not believe me?

Then Jesus said to them, “Truly, truly I say to you, unless you eat the flesh of the Son of Man and drink his blood, you do not have life in yourselves!

No, I tell you, but unless you repent you will all perish as well!

and said, “Truly I say to you, unless you turn around and become like young children, you will never enter into the kingdom of heaven!

So Jesus said to him, “Unless you people see signs and wonders, you will never believe!”
Jesus answered, “Truly, truly I say to you, unless someone is born of water and spirit, he is not able to enter into the kingdom of God.

And unless those days had been shortened, no human being would be saved. But for the sake of the elect, those days will be shortened.

So also my heavenly Father will do to you, unless each of you forgives his brother from your hearts!”

This man came to him at night and said to him, “Rabbi, we know that you are a teacher who has come from God, for no one is able to perform these signs that you are performing unless God were with him.”

John answered and said, “A man can receive not one thing unless it is granted to him from heaven!

And he said, “So how could I, unless someone will guide me? And he invited Philip to come up and sit with him.

Or how can someone enter into the house of a strong man and steal his property, unless he first ties up the strong man? And then he can thoroughly plunder his house.

And when they come from the marketplace, they do not eat unless they wash. And there are many other traditions which they have received and hold fast to—for example, the washing of cups and pitchers and bronze kettles and dining couches.)

You are my friends if you do what I command you. Know that our brother Timothy has been released, with whom I will see you, if he comes quickly enough.

And by this we know that we have come to know him, if we keep his commandments.

For I do not want to see you now in passing, for I hope to remain some time with you, if the Lord allows it.

And let us not grow weary in doing good, for at the proper time we will reap, if we do not give up.

And he said to him, “I will give to you all these things, if you will fall down and worship me.”

For we have become partners of Christ, if indeed we hold fast the beginning of our commitment steadfast until the end,

And do not marvel, brothers, if the world hates you.

And this we will do, if God permits.

And they asked him and said to him, “Why then are you baptizing, if you are not the Christ, nor Elijah, nor the Prophet?”

But we know that the law is good, if anyone makes use of it lawfully,

How are you able to believe, if you accept glory from one another, and do not seek the glory which is from the only God?

But he said to him, “If they do not listen to Moses and the prophets, neither will they be convinced if someone rises from the dead.”

For you put up with it if someone enslaves you, if someone devours you, if someone takes advantage of you, if someone is presumptious toward you, if someone strikes you in the face.
but whenever one turns to the Lord, the veil is removed.

whenever the perfect comes, the partial will pass away.

whenever they persecute you in this town, flee to another, for truly I say to you, you will never finish going through the towns of Israel until the Son of Man comes.

whenever Moses is read aloud, a veil lies upon their heart.

whenever you give a banquet, invite the poor, the crippled, the lame, the blind.

whenever you pray, enter into your inner room and shut your door and pray to your Father who is in secret, and your Father who sees in secret will reward you.

Now learn the parable from the fig tree: Whenever its branch has already become tender and puts forth its leaves, you know that summer is near.

whenever an unclean spirit has gone out of a person, it travels through waterless places searching for rest, and does not find it.

And the unclean spirits, whenever they saw him, were falling down before him and crying out, saying, “You are the Son of God!”

Whenever you enter into a house, stay there until you depart from there.

whenever they hear it, immediately Satan comes and takes away the word that was sown in them.

And no one after drinking old wine wants new, because he says, ‘The old is just fine!’

And after he had said farewell to them, he went away to the mountain to pray.

And the unclean spirits, whenever they saw him, were falling down before him and crying out, saying, “You are the Son of God!”

And after he had entered into the house, his disciples asked him privately, “Why were we not able to expel it?”

And very early in the morning on the first day of the week they came to the tomb after the sun had risen.

And after he set out from the synagogue, he went into Simon’s house. And Simon’s mother-in-law was afflicted with a high fever, and they asked him on behalf of her.
And when King Herod heard it, he was troubled, and all Jerusalem with him,

Now when he saw the crowds, he went up the mountain and after he sat down, his disciples approached him.

And when they got into the boat, the wind abated.

And when they came to the crowd, a man approached him, kneeling down before him.

And when the men of that place recognized him, they sent word into that whole surrounding region, and they brought to him all those who were sick.

"Why do your disciples break the tradition of the elders?? For they do not wash their hands when they eat a meal."

While I am in the world, I am the light of the world.

But while his people were sleeping, his enemy came and sowed darnel in the midst of the wheat and went away.

And while he was still approaching, the demon threw him down and convulsed him. But Jesus rebuked the unclean spirit and healed the boy, and gave him back to his father.

Now it happened that when all the people were baptized, Jesus also was baptized, and while he was praying, heaven was opened,

And while they were eating he said, "Truly I say to you, that one of you will betray me."

But Mary stood outside at the tomb, weeping. Then, while she was weeping, she bent over to look into the tomb,

And while traveling toward Jerusalem, he was passing through the region between Samaria and Galilee.

And while he was still speaking, behold, Judas — one of the twelve — arrived, and with him a large crowd with swords and clubs, from the chief priests and elders of the people.

Now while the Pharisees were assembled, Jesus asked them,

And while all the people were listening, he said to his disciples,

Telling them, "Say 'His disciples came during the night and stole him while we were sleeping.'"

Then Jesus went with them to a place called Gethsemane, and he said to the disciples, "Sit here while I go over there and pray."

and did not have sexual relations with her until she gave birth to a son. And he called his name Jesus.

But he did not want to, but rather he went and threw him into prison until he would repay what was owed.

My children, for whom I am having birth pains again, until Christ is formed in you!

And they did not know anything until the deluge came and swept them all away. So also the coming of the Son of Man will be.

And he said to them, "Whenever you enter into a house, stay there until you depart from there.

Nevertheless, hold fast to what you have until I come.
Jesus said to them, “Truly, truly I say to you, before Abraham was, I am!”

Now the birth of Jesus Christ occurred in this way. His mother Mary had been betrothed to Joseph, but before they came together, she was found to be pregnant by the Holy Spirit.

And Jesus said to him, “Truly I say to you that today — this night — before the rooster crows twice, you will deny me three times!”

Nathanael said to him, “From where do you know me?” Jesus answered and said to him, “Before Philip called you, when you were under the fig tree, I saw you.”

Therefore do not be like them, for your Father knows what you need before you ask him.

So he said, “Men — brothers and fathers — listen: The God of glory appeared to our father Abraham while he was in Mesopotamia, before he settled in Haran,

Consequently, even if I wrote to you, it was not because of the one who did wrong or because of the one who had been wronged, but in order that your diligence on our behalf might be revealed to you before God.

Even if he should come back in the second or in the third watch of the night and find them like this, blessed are they!

But even if we or an angel from heaven should proclaim a gospel to you contrary to what we proclaimed to you, let him be accursed!

Jesus said to her, “I am the resurrection and the life. The one who believes in me, even if he dies, will live,

But even if they all fall away, certainly I will not!

for even if I am absent in the flesh, yet I am with you in spirit, rejoicing and seeing your good order and the steadfastness of your faith in Christ.

And Isaiah cries out concerning Israel, “Even if the number of the sons of Israel is like the sand of the sea, the remnant will be saved,

Therefore we do not lose heart, but even if our outer person is being destroyed, yet our inner person is being renewed day after day.

But even if I am unskilled in speech, yet I am not in knowledge; certainly in everything we have made this clear to you in every way.

So Saul got up from the ground, but although his eyes were open he could see nothing. And leading him by the hand, they brought him into Damascus.

Therefore if you, although you are evil, know how to give good gifts to your children, how much more will the Father from heaven give the Holy Spirit to those who ask him??”

The Jews answered him, “We are not going to stone you concerning a good deed, but concerning blasphemy, and because you, although you are a man, make yourself to be God!”

Wälchli
And they did not find it, although many false witnesses came forward. And finally two came forward. 

For although they knew God, they did not honor him as God or give thanks, but they became futile in their reasoning, and their senseless hearts were darkened.

Although you have eyes, do you not see?? And although you have ears, do you not hear?? And do you not remember?? 

then that man replied, “Whether he is a sinner I do not know. One thing I know — that although I was blind, now I see!”

And although they found no charge worthy of death, they asked Pilate that he be executed.

And Simon answered and said, “Master, although we worked hard through the whole night, we caught nothing. But at your word I will let down the nets.”

And after eight days his disciples were again inside, and Thomas with them. Although the doors had been shut, Jesus came and stood in their midst and said, “Peace to you.”

Righteous Father, although the world does not know you, yet I have known you, and these men have come to know that you sent me.

And although he wanted to kill him, he feared the crowd, because they looked upon him as a prophet.

(Now I also baptized the household of Stephanus. Beyond that I do not know if I baptized anyone else.)

“You know that after two days the Passover takes place, and the Son of Man will be handed over in order to be crucified.”

How do you not understand that I did not speak to you about bread? But beware of the leaven of the Pharisees and Sadducees!”

The woman said to him, “I know that Messiah is coming” (the one called Christ); “whenever that one comes, he will proclaim all things to us.”

There is another who testifies about me, and I know that the testimony which he testifies about me is true.

But the angel answered and said to the women, “Do not be afraid, for I know that you are looking for Jesus, who was crucified.

And they were watching him closely to see if he would heal him on the Sabbath, in order that they could accuse him.

“Have you heard that it was said, ‘Do not commit adultery.’

But the others said, “Leave him alone! let us see if Elijah is coming to save him.”

For which of you, wanting to build a tower, does not first sit down and calculate the cost to see if he has enough to complete it?

And when they looked up, they saw that the stone had been rolled away (for it was very large).

Now when Jesus saw that a crowd was running together, he rebuked the unclean spirit, saying to it, “Mute and deaf spirit, I command you, come out of him, and enter into him no more!”
And they saw that some of his disciples were eating their bread with unclean — that is, unwashed — hands.

The woman said to him, “Sir, I see that you are a prophet.

And when he saw from a distance a fig tree that had leaves, he went to see if perhaps he would find anything on it. And when he came up to it he found nothing except leaves, because it was not the season for figs.

And when he came out he was not able to speak to them, and they realized that he had seen a vision in the temple. And he kept making signs to them, and remained unable to speak.

But Peter and John answered and said to them, “Whether it is right in the sight of God to listen to you rather than God, you decide!

And Peter said to her, “Tell me whether you both were paid this much for the piece of land.” And she said, “Yes, this much.”

Everything that the Father has is mine. For this reason I said that he takes from what is mine and will proclaim it to you.

And Pharisees came up to him in order to test him, and asked if it was permitted for a man to divorce his wife for any cause.

And Pilate was surprised that he was already dead, and summoning the centurion, asked him whether he had died already.

Now when Pilate heard this, he asked if the man was a Galilean.

So he answered and said to them, “When evening comes you say, ‘It will be fair weather because the sky is red,’

And because he did not have enough to repay it, the master ordered him to be sold, and his wife and his children and everything that he had, and to be repaid.

But because he knew their maliciousness, Jesus said, “Hypocrites! Why are you testing me?

So the master of that slave, because he had compassion, released him and forgave him the loan.

And because he was angry, his master handed him over to the merciless jailers until he would repay everything that was owed.

But the chief priests took the silver coins and said, “It is not permitted to put them into the temple treasury, because it is blood money.”

And do not be called teachers, because one is your teacher, the Christ.

But when the sun rose it was scorched, and because it did not have enough root, it withered.

Therefore be on the alert, because you do not know what day your Lord is coming!

And the rain came down and the rivers came and the winds blew and beat against that house, and it did not collapse, because its foundation was laid on the rock.

But he did not answer her a word. And his disciples came up and asked him, saying, “Send her away, because she is crying out after us!”

he said, “Go away, because the girl is not dead, but is sleeping.” And they ridiculed him.
(continued)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>81</td>
<td>40019016</td>
</tr>
<tr>
<td>80</td>
<td>42011016</td>
</tr>
<tr>
<td>80</td>
<td>40019003</td>
</tr>
<tr>
<td>79</td>
<td>40026012</td>
</tr>
<tr>
<td>81</td>
<td>40019013</td>
</tr>
<tr>
<td>81</td>
<td>40027026</td>
</tr>
<tr>
<td>81</td>
<td>40027031</td>
</tr>
<tr>
<td>80</td>
<td>40014015</td>
</tr>
<tr>
<td>80</td>
<td>40002008</td>
</tr>
<tr>
<td>81</td>
<td>40012014</td>
</tr>
<tr>
<td>80</td>
<td>40003013</td>
</tr>
</tbody>
</table>

References


Duff-Tripp, Martha. 1998. Diccionario yanesha’ (amuesha)–castellano. Lima: SIL.


Foster, F. M. n.d. *Indicating prominence in Timbe texts*. SIL Ms.


Heeschen, Volker. 1994. How long are clauses and sentences in a Papuan language like Eipo? In Ger P. Reesink (ed.), *Topics in descriptive Papuan linguistics* (Semaian 10), 50–74. Leiden: Department of Languages and Cultures of South-East Asia and Oceania, Leiden University.


