

# The typology of non-argument clauses

*Oxford Handbook of Clausal Embedding*

Karsten Schmidtke-Bode and Holger Diessel

Friedrich-Schiller-Universität Jena

**Draft version, November 2019**

This is an **extended version** of the published paper. The full version is more elaborate with regard to the data and references that support the claims made in the main text – thus containing substantially more footnotes to point the reader to the relevant work. Additionally, §2.3 on accessibility in relativization includes more phenomena and examples than the printed version and, most importantly, this extended version features a section on supplementary relative clauses (§2.4) that cannot be found in the publication at all.

## ABSTRACT

The present contribution surveys prominent patterns in the typology of relative and adverbial clauses, with particular reference to the notion of embedding. To this end, we unfold the notion of embedding into a three-dimensional space consisting of a functional, a distributional and a formal axis along which a clause may be argued to be embedded into a (constituent of a) main clause. We show how these dimensions of embedding interact with each other to yield cross-linguistically recurrent subtypes of relative and adverbial clauses. In doing so, we also point to correlations between these dimensions and other grammatical properties of the non-argument clause, such as accessibility to relativization or the omissibility of overt markers of the embedding relationship.

## STRUCTURE OF THE PAPER

1. Non-argument clauses and the notion of embedding
    - 1.1 Goal and structure of the paper
    - 1.2 Dimensions and reflexes of embedding
      - 1.2.1 The functional dimension of clausal embedding
      - 1.2.2 The distributional dimension of clausal embedding
      - 1.2.3 The formal dimensions of clausal embedding
  2. Relative clauses in typological perspective
    - 2.1 Headedness
    - 2.2 Integration with the main clause
      - 2.2.1 Internally-headed RRCs (IHRRCs)
      - 2.2.2 Externally-headed RRCs (EHRRCs)
    - 2.3 Orientation of the RRC
    - 2.4 Outlook: Supplementary relative clauses (SRCs)
  3. Adverbial clauses in typological perspective
    - 3.1 Coding and position of ACs
    - 3.2 Desententialization of ACs
    - 3.3 Properties of adverbial subordinators
  4. Conclusion: The dynamics of clausal embedding
- Abbreviations  
References

## 1. Non-argument clauses and the notion of embedding

### 1.1 Goal and structure of the paper

The present chapter is concerned with clauses that function as adjuncts to a particular element of a main clause in a complex sentence<sup>1</sup>. As such, they normally provide circumstantial information on the element in question, rather than being an argument of a predicator, and hence they are collectively referred to as **NON-ARGUMENT CLAUSES** here. The traditional semantic classification of non-argument clauses rests on whether they relate to a nominal of the main clause (= prototypical relative clauses) or the predicate or the entire proposition expressed by the matrix clause (= prototypical adverbial clauses). Accordingly, our goal in the present chapter is to provide an overview of prominent cross-linguistic patterns of relative-clause and adverbial-clause formation.

As it proves impossible for a single article to do justice to the enormous amount of typological research on these clause types that has accumulated over the last decades, we will pursue a narrower, more focussed agenda here. Specifically, we will attempt to systematize the typology of adjunct clauses from the central perspective of the present handbook, that of embedding. To this end, we begin the paper (in §1.1) by unfolding the notion of embedding into a three-dimensional space consisting of a functional, a distributional and a formal axis along which a clause may be argued to be embedded into a (constituent of a) main clause. In doing so, we also intend to lay the terminological groundwork for all subsequent sections. The ensuing §2 and §3 will then be concerned with the typology of relative and adverbial clauses, respectively, before §4 concludes the paper by situating commonalities between the two domains in a dynamic view of clausal embedding and grammatical structure more generally.

### 1.2 Dimensions and reflexes of embedding

In Lehmann's (1988) typology of clause linkage, embedding is the endpoint of a continuum of "hierarchical downgrading", a situation where a clause comes to behave as a "well-defined constituent" (ibid.: 184) of another unit, i.e. a clause or even a phrase. In the present paper, we understand this to mean that a clause is *functionally*, *distributionally* and *formally* integrated with that other unit. In the following subsections, we briefly introduce each of these three dimensions and begin to carve out their relevance for the typology of non-argument clauses.

#### 1.2.1 The functional dimension of clausal embedding

A clause is fully integrated from a *functional* point of view if it has a specific semantic and syntactic function inside another unit. This is arguably the case when a clause "elaborates a site" (Langacker 1987: 304) or fills a gap (Lehmann 1983, Diessel 2019b: Ch. 9) projected by a relational element of the matrix clause, i.e. by a verbal, nominal

---

<sup>1</sup> We follow Quirk et al. (1985: 991) in labelling the superordinate clause in a complex sentence the **MAIN CLAUSE**; it is the entire unit in which a subordinate clause may have a syntactic function. The term **MATRIX** or **MATRIX CLAUSE**, by contrast, designates that part of the main clause that remains when the subordinate clause is removed. Hence MAIN[MATRIX[I like the book] SUB[that you've written]]. The distinction is not particularly crucial in the discussion of adverbial clauses, but it is vital for typologizing relative clauses.

or adjectival predicate in the matrix. Non-argument clauses, by contrast, are not projected in the same way. This holds for all of the following clauses, taken from typologically very diverse languages:

- (1) relative clause from Chiapas Zoque (Mixe-Zoque: Mexico; Faarlund 2012: 162)
- $_{NP}$ [*Te'* [*jo'nchi*  $_{REL}$ [*y-ku't-u=pü='is* *te' tüm*]]] *ø-kek-u.*
- DET bird.ABS  $\exists$ ERG-eat-COMPL=REL=ERG DET fruit.ABS  $\exists$ ABS-fall-COMPL
- 'The bird which ate the fruit fell down.'
- (2) adverbial clause from Noon (Atlantic-Congo, North-Central Atlantic, Cangin: Senegal; Soukka 2000: 278)
- [*Balaa mi le' kaad-aa*], *mi híd-oh-ha na húnismun.*
- before 1SG arrive house-IRR 1SG meet-RECP-NARR with friend
- 'Before I arrived home, I met with a friend.'
- (3) clause chaining from Mauwake (Nuclear Trans New Guinea, Madang: Papua New Guinea; Berghäll 2015: 335)
- [*Nainiw ekap-ep*] *maa me sesenar-e-mik.*
- again come-SS.SEQ food not sell-PST-1/3PL
- 'They did not come back and sell food.'

In the literature, it is thus said that the presence of the dependent clauses in (1)–(3) is not specifically licensed or governed by a predicator. It has been pointed out, however, that the difference between arguments and non-arguments is gradient rather than categorical (see, e.g., Deutscher 2000 on Akkadian), and we will see later on that some relative and adverbial clauses are more closely associated with their head element than others, with interesting typological consequences (see §3.2–3.3).

Despite not being selected by a main-clause element, non-argument clauses can still fulfil a semantic and syntactic role in another unit and thus be functionally embedded in that unit. This is what happens most clearly in (1) and (2) above: In (1), the dependent clause functions as an attributive modifier of a nominal, and hence has a dedicated function inside an NP; in (2), the dependent clause spells out part of the setting of the matrix-clause situation and thus functions just like a phrasal constituent with similar semantics. In both cases, then, the dependent clause acts as a modifier of an element in the matrix. This contrasts with the situation in (3), where we see a succession of events in which the formally dependent clause cannot be said to function as a modifier of the final finite clause. The grammatical construction in (3) has aptly been called **CLAUSE CHAINING** in the typological literature, best-known from (but not limited to) languages of Papua New Guinea (see, e.g., Roberts 1988 for an overview), and the dependent clauses in such chains have been argued not to have a syntactic function in a larger clause and hence not to be embedded within it (see Van Valin 1984, 2005: Ch. 6, who uses the term **COSUBORDINATION** for this phenomenon).<sup>2</sup>

<sup>2</sup> The term 'cosubordination' is supposed to capture the fact that, in the languages in question, clause-chaining constructions are grammatically distinct from both coordinate and subordinate clauses. Unlike subordinate (and like coordinate) clauses, they do not fulfil a semantic or syntactic role in another clause and rather encode sequential (and often non-backgrounded) events in a narrative; unlike coordinate clauses (and like subordinate clauses in these languages), they are grammatically dependent on another

Accordingly, while the typological domain of ‘non-argument clauses’ also includes dependent clauses like the one in (3), the latter are not functionally embedded like the dependent clauses in (1)–(2), which may be referred to more narrowly as **ADJUNCT CLAUSES**.

It has been widely recognized that adjuncts are extremely heterogeneous themselves (see, e.g., Maienborn and Schäfer 2011 for a survey). A coarse-grained but useful distinction is that between adjuncts that modify the propositional meaning of an element of the matrix clause (= **MODIFIERS**) and adjuncts that provide additional comment on the element in question (= **SUPPLEMENTS**). Modifiers of a nominal element typically restrict the referential potential of the NP headed by that nominal and hence go by the name of **RESTRICTIVE RELATIVE CLAUSES**; modifiers that restrict a situation to a specific setting (time, place and manner) or specific contingent circumstances (condition, cause, purpose, result, etc.) are traditionally known as **ADVERBIAL CLAUSES**. They are often said to spell out “peripheral” information on the “core” of the matrix proposition (e.g. in Van Valin and LaPolla’s (1997: 26) terms). Adjuncts that provide additional comment are called supplements here (extending a descriptive term used by Huddleston and Pullum (2002: 59) for English). Relative clauses can be used to express supplementary information on both an NP in the matrix clause (4) and a whole clause (5); they are also known as **NON-RESTRICTIVE, SUPPLEMENTARY** and – in the case of (5) – **SENTENTIAL** relative clauses. In many languages, the clausal constructions used as adverbial modifiers can also be employed as adverbial supplements (see Tsunoda 2018 for examples). In this function, they often provide the speaker’s attitude towards the matrix proposition or relate to the speech act (rather than the propositional content) expressed by the matrix clause (6–8):

- (4) supplementary relative clause (with a nominal antecedent) from Basque (isolate: Spain, France; Hualde and Ortiz de Urbina 2003: 804)

*Gure Jainkoak, [beti xuxen dabilana], ongi gidatzen baitu.*  
 our god.ERG always right acts.SUB.DET well guide.IMP SUB.AUX  
 ‘Because our God, who always acts rightly, guides well.’

- (5) supplementary relative clause (with a sentential antecedent) from Avatime (Atlantic-Congo, Volta-Congo, Kwa Volta-Congo: Ghana; van Putten 2014: 85)

*Àà-gu kpe [gì lí-tá-bùbò].*  
 C<sub>1</sub>:SG.POT-speak put.in REL C<sub>3</sub>:SG.NEG-INT-easy  
 ‘He will be speaking into it, which will not be easy.’

- (6) supplementary (conditional) adverbial clauses from English (Huddleston and Pullum 2002: 774)

- a. *Dick is coming to the party, [in case you’re interested].*  
 b. *[If you must know], I wasn’t even shortlisted.*

---

clause, i.e. the final one in the chain, notably for the interpretation of tense, aspect and mood operators. The dependent clauses in such chains are also called ‘medial clauses’.

- (7) supplementary (purposive) adverbial clause from Yakkha (Sino-Tibetan, Himalayish, Kiranti: India, Nepal; Schakow 2015: 450)

[*Yeppa cok-ma bhoŋ*]      *i-ha-ca*                      *im-ma*    *por-a*  
 true    do-INF COND/PURP    what-NMLZ.NSG-ADD    buy-INF    must-NATIV

*n-joŋ-me-ŋa-n.*

NEG-do-NPST-1SG-NEG

‘To be honest/If I’m honest, I do not have to buy anything.’

- (8) supplementary (concessive) adverbial clause from Georgian (Kartvelian: Georgia; Kojima 2018: 436)

*Ak sač'mel-i-a,*                      [*tumca*    *es*                      *šen*                      *i-c-i*].  
 here food-NOM=be.PRS.3SG    although this.NOM 2SG.ERG    i-know-THEM

‘There is food here, although you know this.’

### 1.2.2 The distributional dimension of clausal embedding

There are several ways in which a clause can be more or less *distributionally* integrated into another unit. First, the more internally coherent the unit in question, the more likely it is to be considered a constituent of a higher unit. Coherence is achieved by keeping the constituting elements of the unit in linear adjacency (Behaghel 1932); the opposite pattern, sometimes known as non-configurationality (Hale 1983), is to separate these elements from one another. Cross-linguistically, adverbial clauses seem averse to being split up by their matrix clause, even in languages that otherwise license discontinuous constituents (see, e.g., Reinholtz 1999 on Swampy Cree). If adverbial clauses are interrupted at all, it is usually by further adverbial (or other dependent) clauses, as in (9):

- (9) Koyra Chiini (Songhay: Mali; Heath 1999: 281)

[*jaa* [*nda baana kar*], *ganji-ije*    *di*    *yo*    *o*    *ñin*]  
 since if    rain    hit    forest-child DEF PL IMPF drink  
 ‘because, if rain falls, the wild animals will drink’

Discontinuities are more common in relativization, when the relative clause is separated from its nominal head. This so-called **EXTRAPOSITION** from NP is illustrated in (10):

- (10) Slave (Athabaskan-Eyak-Tlingit, Northern Athabaskan: Canada; Rice 1989: 1327)

*John*    *łi*    *wehk'é*    [*?eyi*    *ts'ó'dani*    *kayihk'a*    *yil'é*    *i*].  
 John dog 3.shot    the child 3.bit    PST REL/NMLZ  
 ‘John shot the dog that bit the child.’

This leads us to the position of non-argument clauses more generally, as a second indication of their distributional integration with the matrix clause: the more closely a dependent clause can be interwoven with the matrix clause, the clearer its status as a constituent of the main clause and hence its degree of embedding. Along this continuum, we find dependent clauses that (i) are able to disrupt the core predicate-argument structure of the matrix clause (**CENTRE-EMBEDDING**, as in (11)), (ii) attach to the left or right of the this core (12), (iii) are relegated to more marginal positions of the matrix clause, such as strictly before or after other adverbial modifiers of the core

((10) and (13)), or (iv) are removed from the matrix clause into a left- or right-detached position (14). The latter pattern typically goes hand in hand with the intonational separation of the two clauses (see also Van Valin 2005: §6.2 for discussion of these positional types):

- (11) Urarina (isolate: Peru; Olawsky 2006: 690)

*Nii hetau=te [ahe-urra-a=ne hana] tura-a ku-e.*  
 that HRS=FOC get.drunk-PL-3.DF=SUB when arrive-NTR go-3.EF  
 ‘That [man] arrived when they [the people] were getting drunk.’

- (12) English

*Anna was angry [after she left the party] because of Mike’s sexist jokes.*

- (13) Martuthunira (Pama-Nyungan, Pilbara: Australia; Dench 1995: 252)

*Ngayu kartungu parla-marta purra-rninyji pal.ya-a, [pungka-waa-rru].*  
 1SG.NOM 2SG.ACC stone-PROP hit-FUT temple-ACC fall-PURP.S=P-now  
 ‘I’ll hit you in the temple with a stone, so that you fall down.’

- (14) Supyire (Atlantic-Congo, North Volta-Congo, Senufo: Burkina Faso; Carlson 1994: 492)

*[Myàhíí u a cèè gé], ci náhá mii fúnjí í.*  
 song.DEF.3PL she PRF sing REL 3PL be.here my inside in  
 ‘The songs which she sang, they are here inside me.’  
 (‘I remember the songs which she sang.’)

In (13), we see a typical representative of what has been called an “adjoined clause” (Hale 1975) in Australian languages, a dependent clause that precedes or follows the matrix clause, may be “bound intonationally” to it (e.g. Nordlinger (1998: 217) on Wambaya), but is not normally found inside it. In (14), finally, the relative clause is found in a left-detached position. As Carlson (1994: 488) states, “relative clauses in Supyire are unembedded,” and they are “typically followed by a short pause before the main clause.” In contrast to the relative clause in (10) above, constructions like the one in (14) are not due to an option of extraposition; they standardly occur in a left- or right-adjoined position and are not normally inserted between elements of the matrix clause (for further discussion, see §2.2.2 below). Examples (11) to (14), then, illustrate a continuum from clearly integrated to clearly non-integrated adjunct clauses. As we shall see, this parameter is a crucial typological variable that distinguishes different construction types of relative and adverbial clauses cross-linguistically.

### 1.2.3 The formal dimension of clausal embedding

When clauses are functionally and distributionally integrated into another unit, it is common for them to also approximate the form of the typical constituent, namely that of a phrase. Haiman and Thompson (1984) and Lehmann (1988) introduce a wide range of dimensions that reflect the “deategorization” (Malchukov 2004) of a clause and its “recategorization” as a phrasal unit, a phenomenon also known as **DESENTENTIALIZATION** (Lehmann 1988). Care must be taken not to interpret this as a unidirectional diachronic process, but as a descriptive device for comparing the extent to which language-specific constructions exhibit formal properties associated with clauses and phrases, respectively. These can emerge in very different ways, notably

also by “expanding” a phrase into a (more) clausal unit (Heine 2009) rather than by reducing a full-fledged clause. For example, when the NP governed by a benefactive adposition (‘for my wife’) is replaced by a deverbal noun with its own arguments and modifiers, the result can be a fully desententialized adverbial clause of purpose (‘for (the) feeding (of) the children’). It has been documented how such constructions can gradually rid themselves of nominal properties and acquire a more clause-like treatment of their arguments and modifiers (e.g. Disterheft and Viti 2010 for infinitives in Indo-European, Fanego’s (2004) detailed study of English gerunds or Trask’s (1997: 215) similar suggestion for Basque nominalizations in *-t(z)e*).

Among the hallmarks of desententialization are the restriction or lack of clausal and sentential operators (illocutionary-force marking (see Verstraete 2007, Bickel 2010 for specific parameters), tense-aspect-mood marking, negation), the use of dependent verb forms, and the reduction and/or special coding of the arguments and modifiers. Dependent verb forms that are specifically associated with relativization have been called **PARTICIPLES** (see Shagal 2019 for a monographic typological study), while dedicated adverbial verb forms have been labelled **CONVERBS** in the typological literature (Haspelmath 1995, Nedjalkov 1998), although it must be stated that there is considerable diachronic interaction and hence synchronic overlap between participles, converbs and nominalizations (Ylikoski 2003). In (15)–(20) below, we illustrate different degrees and facets of desententialization in non-argument clauses.

- (15) weakly desententialized concessive clause: internal structure as in independent clauses, except for a clause-initial subordinator

Tz’utujil (Mayan: Guatemala; Dayley 1985: 371)

[*Maanaan xtipeeti ja Aa Lu?*], *majun nuub’an.*  
 although 3ABS.might.come the youth Pedro nothing 3ABS.3ERG.do  
 ‘Even though Pedro might come, he won’t do anything.’

- (16) weakly desententialized relative clause: no subordinator, internal structure as in independent clauses, except for reduced person indexation

Musqueam (Salishan: Canada; Suttles 2004: 76)

*K<sup>w</sup>θə swə̀yqe?* [*céw-ət-Samx-∅ ce?*]

ART man help-TR-me-∅ FUT

‘the man who will help me’

(compare *céw-ət-S(amx)-əs ce?*. = help-TR-me-3TR FUT = ‘He will help me.’)

- (17) more strongly desententialized temporal/causal clause: subordinating conjunction, non-canonical subject coding (zero Accusative instead of overt Nominative case), dependent verb form (phonologically reduced), but retention of tense-aspect inflections

Wappo (Yuki-Wappo: USA; Thompson et al. 2006: 155)

*K’anihtuč’m-i nale?-iš-khi?* [*k’ešu-∅ pulu:mek-ta wen*].

chief-NOM angry-INCH-STAT deer-ACC run:away-PST.DEP **when**

‘The chief got angry when/because the deer ran away.’

- (18) strongly desententialized temporal clause: semantically specific subordinating suffix -*ka* attached to a bare verb stem (= converb), reduction of TAM and person inflection

Awa Pit (Barbacoan: Colombia; Curnow 1997: 272)

[*Santos a-ka*]=*na*,                      *kula-ta-w*.  
 Santos    come-**when**=TOP            hide-PST-LOCUT:SBJ  
 ‘When Santos came, I hid.’

- (19) strongly desententialized purpose clause: nominalized verb form, reduction of TAM and person expression, dative case marker (= “purposive phrase”)

Konso (Afro-Asiatic, Cushitic: Ethiopia; Mous and Oda 2009: 349)

*Innaannó*                      [*golpaytá pidd-a-é*]                      *urmala-pá*                      *aan-é*.  
 brother:1PL.POSS    goat                      buy-NMLZ-DAT    market-DEST    go-PRF  
 ‘Our brother went to the market in order to buy a goat.’

- (20) strongly desententialized relative clause: nominalized verb form, reduction of tense, mood, person-number inflection and lack of separate negative suffixes

Dolakha Newar (Sino-Tibetan, Himalayan: Nepal; Genetti 2007: 390)

[*Am pipāna ye-u*]                      *mi=pen*                      *gun?*  
 that veranda come-NMLZ/REL    person=PL                      who  
 ‘Who are the people who came to the veranda?’

In recent typological studies, the gradient nature of Lehmann’s desententialization has been captured empirically by developing fine-grained structural variables along which language-specific clause-linking constructions may differ (e.g. Bickel 2010, Schmidtke-Bode 2014, Hetterle 2015, Shagal 2019). This, in turn, allows a more rigorous comparison of the degree to which different types of adjunct clauses (e.g. different adverbial relations) are structurally downgraded and thus approximate the formal make-up of phrases rather than clauses (as in (19) above).

In the following sections, we will show how the three dimensions of embedding combine with one another to yield typologically prominent patterns of relative-clause and adverbial-clause formation. In the terms of the present section, the remainder of this article is thus concerned with clauses that are *functionally* embedded as adjuncts of another syntactic unit (phrase, clause or sentence), to which they relate as modifiers or supplements. This removes coordinate, purely sequential chaining constructions and all kinds of argument clauses from the investigation, in full acknowledgement of the fact that we do not establish clear dividing lines between argument–adjunct, phrase–clause, modifier–supplement and so on.



## 2. Relative clauses in typological perspective<sup>5</sup>

A relative clause is a clausal construction that is semantically “oriented towards one of its NP positions” (Lehmann 2014: 4), in order to specify an entity of the main clause in such a way that the designatum of that entity occupies a participant position in the situation encoded by the relative clause (see also Bickel 2005 for a similar definition).<sup>4</sup> The NP towards which the relative clause is oriented is known as the **RELATIVIZED POSITION**.

The main-clause entity that is specified by the relative clause is most commonly a **NOMINAL** (e.g. *old man*) centred around a **NOMINAL HEAD** (*man*); in these situations, relative clauses are typically used to restrict the referential potential of the NP headed by the nominal head, as in *the old man I saw at the pub last night*, or our earlier example (1). These relative clauses thus act as modifiers of a nominal and are known as **RESTRICTIVE RELATIVE CLAUSES (RRCs)**. By contrast, **SUPPLEMENTARY RELATIVE CLAUSES (SRCs)** specify an NP, the matrix predicate or the matrix clause by providing elaborative information on it without restricting its referential potential, as in (4)–(5) above. In the following overview, we will concern ourselves primarily with the typology of restrictive relative clauses (§2.1–§2.3), as most cross-linguistic work has been done on this type of relativization.<sup>5</sup> However, we will conclude our exposition by a brief outlook on supplementary relative clauses in §2.4.

### 2.1 Headedness

The first parameter of cross-linguistic variation for RRCs is whether the nominal head to be modified is syntactically part of the RRC, appears outside of it, or is left implicit. Accordingly, many typologies distinguish between **INTERNALLY-HEADED**, **EXTERNALLY-HEADED** and **HEADLESS RRCs**. When we compare (1) above to (21) below, we see that *úsa* in (21) is (zero-)marked for its syntactic function inside the RRC, namely its absolutive S-role. It is the RRC as a whole that bears an ergative flag for its syntactic function in the main clause. This is, therefore, an internally-headed RRC. In (1), by contrast, *jo'nchi* heads an NP that is case-marked for its syntactic function in the main clause, namely that of an absolutive S-argument, while its function in the RRC is that of an ergative A-argument. Consequently, the RRC in (1) is externally-headed.

---

<sup>5</sup> The most comprehensive typological analysis of relative clauses is Lehmann (1984), crucial aspects of which are succinctly summarized in various follow-up publications (Lehmann 1986, 2008, 2014). Non-monographic systematizations of the typology of relativization can be found in work by Bernard Comrie (e.g. Comrie 1981, 1998, 2006), in Downing (1978), Keenan (1985), De Vries (2001, 2005), Andrews (2007), and a number of volumes on (functional) syntax in cross-linguistic perspective (e.g. Dik 1997: Chs. 3–4, Givón 2001: Ch. 14, Dixon 2010: Ch. 17). Finally, Hendery (2012) provides a diachronic typology of relative clauses.

<sup>4</sup> The distinction between main and matrix clause (see fn. 1 above) becomes crucial now: our definition does not require the element specified by the (restrictive) relative clause to be in the matrix clause (as this would cover only particular structural types of relative constructions); but the element in question is, of course, part of the main clause.

<sup>5</sup> In fact, many typological studies take restrictiveness as a crucial part of their comparative concept for relative clauses, such as Comrie and Kuteva (2013) or Andrews (2007).

- (21) Epena Pedee (Chocoan: Colombia, Ecuador; Harms 1996: 167)

[Úsa-ø t<sup>h</sup>ot<sup>h</sup>óo bɛ̃]-pa peróra pee-pa-čí.  
 dog-ABS white be-ERG spotted.cavy kill-HAB-PST  
 ‘The dog that is white used to kill spotted cavies.’

A headless RC is illustrated in (22):

- (22) Ute (Uto-Aztecan, Numic: USA; Givón 2011: 288)

[kani-vaa-tɬ wúuka-qha-tɬ]  
 house-LOC-DIR work-ANT-NMLZ  
 ‘the one who worked at the house’

In this example, the proposition of the subordinate clause is oriented towards one of its arguments by virtue of a participant nominalizer, the suffix *-tɬ*, yielding the English translation ‘the one who ...’ (rather than the unoriented proposition ‘someone worked at the house’). Recent work on different areas of the world has brought out the importance of (participant) nominalizations for the typology of relative clauses. In particular, it has been pointed out that when such nominalizations are placed in apposition to an NP in the main clause, the resulting structure and initial interpretation is that of a supplementary relative clause – a looser, often phonetically disintegrated juxtaposition of two NPs with the same referent:

- (23) Ute (Uto-Aztecan, Numic: USA; Givón 2011: 291)

Mamachi<sub>i</sub> [(‘ú) kani-vaa-tɬ wúuka-tɬ], tɬgɬvɬ-n ‘ura-’ay.  
 woman.SBJ that.SBJ house-LOC-DIR work-NMLZ friend-1SG be-IMM  
 ‘The woman, the one who works at home, is my friend.’

Given the right circumstances, however, such appositions can come to be reanalyzed as restrictive, by the processes summarized succinctly in Lehmann (2008: 219ff., 2014: 3–4). This may or may not be accompanied by the formal integration of the participant nominalization with the nominal head, yielding a difference in the degree to which the two nominals form a single NP. The following example illustrates the integrated externally-headed RRC in Ute<sup>6</sup>:

- (24) Ute (Uto-Aztecan, Numic: USA; Givón 2011: 288)

mamachi ‘u [kani-vaa-tɬ wúuka-qha-tɬ]  
 woman.SBJ 3SG.SBJ house-LOC-DIR work-ANT-NMLZ  
 ‘the woman who worked at home’

## 2.2 Integration with the main clause

In the context of the present volume, perhaps the most crucial parameter of cross-linguistic variation is that RRCs of all headedness types above can be either **EMBEDDED** in the main clause, where they form a syntactic constituent with the nominal head (if

<sup>6</sup> In Ute, non-restrictive RRCs are rather rare (Givón 2012: 291), as are prosodically integrated restrictive relatives (ibid.: 285). The most common type in unelicited oral discourse thus appears to be the headless construction illustrated in (22), either on its own or with an apposed NP which it modifies restrictively (ibid.: 287). This nicely illustrates how a looser (= appositive) syntactic structure is associated with a tighter (= restrictive) semantic interpretation; see also Lehmann (2014: 3) on this point.

there is one), or be **ADJOINED** to a syntactically saturated matrix clause. We will discuss this difference separately for internally- and externally-headed RRCs.

### 2.2.1 Internally-headed RRCs (IHRRCs)

IHRRCs are embedded in the main clause if they directly fulfil a syntactic function in it, as is the case in (21) above.<sup>7</sup> Lehmann (1984, 1986) reserves the term **CIRCUMNOMINAL RC** for this embedded variant of an IHRRC. In much of the typological literature, the label IHRRC is often used more narrowly to designate only this type of RRC (see also Cole 1987, Culy 1990 and Basilico 1996 for relevant work). In Dryer’s (2013c) sample of more than 800 languages, 24 exhibit the circumnominal type as their primary RRC construction; these languages cluster conspicuously in North America and, moreover, they are overwhelmingly OV-languages (all 24 except for Kutenai and Oneida).<sup>8</sup> A further 39 languages have the circumnominal type as one of several RRC constructions. Since the hallmark of circumnominal RRCs is that they function as an NP in the main clause, they typically bear some nominal morphology at their right boundary, such as a case marker in (21) or a case-marked determiner in (25):

(25) Choctaw (Muskogean: USA; Broadwell 2006: 50)

[*Hattak-mat ofi’ chōpa-tok-mā*] *p’isa-li-tok.*

man-DEM.NOM dog buy-PST-DEM.ACC see:NGR-1SG.I-PST

‘I saw the dog that the man bought.’ (or: ‘I saw the man that bought the dog.’)

Circumnominal RRCs contrast with those internally-headed constructions whose nominal head is also represented in the matrix clause, in different degrees of explicitness. In all cases, this leads to the RRC and the matrix clause containing a **CORRELATIVE** element, and typologists thus often use the term **CORRELATIVE RRC**. The most explicit but arguably least economical way of creating this structure is by having a full noun in each clause, either a direct “copy” of the head nominal or a corresponding light noun like ‘thing’ or ‘person’. Only a single language in Dryer’s (2013c) database, Kombai, exhibits such **DOUBLY-HEADED RRCs** as its primary RC type, and only four others (Jamsay, Mina (both Africa), Kobon and Yagarian (both Nuclear Trans-New Guinea)) as a secondary type. Two examples beyond Dryer’s database come from Gooniyandi (26), where doubling occurs “occasionally” (McGregor 1990: 438), and from Santali (27), where doubling is part of a distinct RRC construction dubbed a **CORRELATIVE DIPTYCH** (Haudry 1973):

(26) Gooniyandi (Bunaban, Australia; McGregor 1990: 438)

[*Thangarndi garndiwangooddoo gooddoomba-ya yoodjidi*]

word many paper-LOC we.put.it

*thangarndi binaddigmiloona.*

word I.taught.them

<sup>7</sup> Strictly speaking, therefore, such RC constructions are actually argument clauses in terms of their syntactic status, but they are semantically oriented towards one of their NP positions and hence to be interpreted as a modifier of a nominal.

<sup>8</sup> VO languages with circumnominal RCs beyond Dryer’s (2013e) sample that have been mentioned in the literature include Mooré (Culy 1990: 214) and Dagbani (Lehmann 1984: 118).

‘I taught them (some of) the many words we had put on paper.’

- (27) Santali (Austroasiatic, Mundaic: Bangladesh; Neukom 2001: 199)

[*Oka disom-re onko gadel hər-ko jarwa-akan-tahēkan*],  
**which country-LOC** those.ANIM.PL crowd person-3PL.SBJ gather-PRF.MID-COP.PST

*ona disom-ren raj-də tis-re cə-e gəc'-akan-tahēkan.*  
**that.INAN country-GEN.ANIM** king-TOP when-LOC ever-3SG.SBJ die-PRF.MID-COP.PST

‘The king of the country where these crowds of people had come together had died some time previously.’

(27) represents a correlative construction that is familiar from Latin or Hittite and which is a well-known areal feature of RRC formation in South Asia as well as in Mande languages and their neighbours (Dryer 2013c); in this ‘question-and-answer’ construction, the IHRRC contains a relative/interrogative determiner accompanying the nominal head, while the main clause either exhibits a demonstrative determiner with a copy of the head nominal (as in (27)) or else simply a (demonstrative or personal) pronoun. In fact, the latter pattern, with only a pronominal correlate in the matrix clause, is also found in Santali and even preferred to double heads in Gooniyandi. We encountered this type of correlative clause earlier, in (14) from Supyire. It was mentioned there that the IHRRC in Supyire is relegated to a left-detached (and less commonly, a right-attached) position, usually with a slight intonation break, and obligatorily correlated with a personal pronoun in the matrix clause. Another representative of this type is Bambara:

- (28) Bambara (Mande, Manding-Vai: Guinea; Givón 2001: 183)

[*Cε min ye muru san*], *n ye o ye.*  
 man REL PST knife buy I PST him see  
 ‘The man who bought the knife, I saw him.’

The pattern that unites all correlative constructions is their (preferred) occurrence at the left sentence boundary (“preposed”, “left-dislocated”), in a topical or expository discourse function (Lehmann 1984, 2008); this also holds for the RC illustrated in (26) from Gooniyandi (McGregor 1990: 438)). What distinguishes the constructions in (26) and (27–28) is thus chiefly their relative degree of desententialization: the RRC in Gooniyandi shows all vestiges of an independent clause, while the ones in (27–28) bear a relative marker that would not be present if these were independent clauses. Some authors have thus described (26) as a **PARATACTIC** relative clause (e.g. Comrie and Kuteva 2013). Overall, however, correlative RRCs show a low degree of desententialization, which reflects their diachronic origin as a discourse strategy in which a (typically indefinite-specific) referent is “introduced as something to be resumed” in the next clause (Lehmann 2008: 217).

In the present context, it is interesting to point out how these adjoined clauses may develop into embedded structures. In fact, depending on the specific type of correlative RC, only small adjustments might be needed to turn it into circumnominal ones, notably the combination of the two clauses without a prosodic break and the omission of the anaphoric demonstrative in the second clause. If the demonstrative carried case markers, these may come to cliticize to the RC as a whole (which would account, for example, for a pattern like (21) above). Alternatively, the anaphoric

demonstrative may be reanalyzed as the final element of the first clause, also yielding a clausal nominalization (which would account for structures such as (25) above). Also, as observed by Lehmann (1984: 388), some languages allow their correlative diptychs to change the order of relative determiner and noun, making the head noun topical (e.g. [*book which you gave me*], *I lost it*). This structure may then be reanalyzed as postnominal, and if the correlative element in the matrix clause is dropped, the result is an embedded subordinate clause. Finally, it is possible for preposed correlative clauses, as a whole, to be attracted to other positions in the main clause. Firstly, a correlative clause may be postposed (i.e. right-adjoined) to the main clause, as an “inverted diptych” that codes presupposed information, from which it can develop into a postnominal RC (Lehmann 2008: 218; see also §2.2.2 below). Secondly, the normally preposed RC may be inserted directly into the main clause right after the nominal to be modified; this latter option is illustrated by Comrie (2006: 139) for Hittite, and he argues that these constellations gave rise to postnominal RRCs with relative pronouns (§2.2.2), which subsequently spread across Europe through language contact. Thirdly, correlative constructions may show up, perhaps rather unexpectedly, in centre-embedded positions, as in Wappo (Thompson et al. 2006: 115) or in Cabecar:

(29) Cabecar (Chibchan: Costa Rica; Gonzáles Campos and Lehmann 2019: 10)

Bá te [jǰí tk-á=ju yiki] jé s-á?  
 2SG ERG earthquake cross-PRFV=AM yesterday DEM feel-PRFV  
 ‘Did you feel the earthquake that happened yesterday?’

Structurally, RRCs in Cabecar are of the paratactic type, i.e. there is no trace of desententialization, and they are often (but not obligatorily) resumed by the medial demonstrative *jé* in the main clause. Gonzáles Campos and Lehmann argue that this construction evolved from a sequence of independent main clauses, functionally equivalent to the preposed correlative clauses of the present section. By intonational integration and omission of the demonstrative, one can easily obtain a ‘standard’ circumnominal RC which directly occupies a syntactic slot in the main clause. But as (29) shows, it is also possible to centre-embed the RRC while retaining the co-referential demonstrative. As a result, the RRC is no longer left-adjoined, unlike a typical correlative clause, which yields an interesting surface combination of an embedded clause that seems adjoined (or rather juxtaposed) to an anaphoric demonstrative. However, Gonzáles Campos and Lehmann (2019: 25) argue that “the sole function of the demonstrative” in these cases “is to mark the final boundary of the relative clause,” which effectively makes (29) exactly parallel to the circumnominal RRC in (25) and thus marks the “last step” (ibid.) in the development from a paratactic to an embedded structure.

### 2.2.2 Externally-headed RRCs (EHRRCs)

EHRRCs can, in principle, occur before or after the nominal they modify. The latter option was exemplified in (1), (4), (5) and (10) above, while the former option is illustrated in (30):

(30) Lezgian (Nakh-Daghestanian, Lezxic: Azerbaijan; Haspelmath 1993: 340)

[*Mu'minat-a ktab ga-ji*]                      *ruš-a q<sup>h</sup>fe-na*.  
 Muminat-ERG book give-AOR.PTCP girl-ERG go.away-AOR  
 'The girl to whom Mu'minat gave the book went away.'

Cutting across this basic ordering distinction is, once more, the parameter of embedding: specifically, the question here is whether the EHRRRC occurs **ADNOMINALLY**, i.e. adjacent to the nominal it modifies (as in (30)) or **EXTRAPOSED** from it (as in (10)), and whether it is syntactically (and not just semantically) an attributive modifier of the head.

We begin by inspecting the **ADNOMINAL** type of EHRRCs, which is the typologically dominant relative construction by a clear margin. However, there is a strong asymmetry in the distribution of the **PRENOMINAL** and the **POSTNOMINAL** subtype: whereas postnominal RRCs are extremely widespread (in geographical, genealogical and typological respects), prenominal RRCs are much more confined. First, barring a few exceptions in mainland China and Taiwan (e.g. Bai (Tibeto-Burman), Chinese languages, Amis (East Formosan)), their occurrence is restricted to OV languages (Dryer 2013e), while postnominal RCs are common in both VO and OV ordering types.<sup>9</sup> Second, among the OV languages, prenominal RRCs are dominant mainly in Asia, while OV languages in the other macro areas appear to prefer different kinds of RC constructions (*ibid.*).<sup>10</sup> Finally, while languages with prenominal RCs sometimes have a postnominal alternative construction (29/170 languages = 17% in Dryer 2013c), the vast majority of languages with postnominal RCs do not have a prenominal alternative (579/610 = 95% in Dryer 2013c).<sup>11</sup>

Several factors are responsible for this asymmetry. To begin with, it appears that attributive modification has a certain postnominal bias, as attributive adjectives – the structurally 'smaller' equivalents of RRCs – are cross-linguistically also preferred after the noun they modify (Dryer 1992).<sup>12</sup> Furthermore, in contrast to postnominal RCs (and the proposed internally-headed types we surveyed above), prenominal RCs are

---

<sup>9</sup> According to Greenberg's (1963: 71) Universal 24, prenominal RCs occur in languages with either postpositional phrases or prenominal adjectives (or both). The latter restriction thus accounts for prenominal RCs in languages like Amis and Mandarin (VO&Prep but Adj-N) as well as for Tigré (OV&Prep but Adj-N in Dryer 2013c,d,e). On prenominal RCs in VO languages, specifically in Amis, see Comrie (2008a).

<sup>10</sup> Local pockets of OV and RelN are found in "(i) New Guinea; (ii) Ethiopia and Eritrea; and (iii) southern Colombia and the adjacent area of Brazil" (Dryer 2013e). However, all of these regions are parts of larger macro areas used for typological studies, and the OV languages of these macro areas do not generally favour prenominal RRCs.

<sup>11</sup> These figures are not controlled for geographic or genealogical dependencies, and neither do they imply that the languages in question use the *same* RC construction in pre- and postnominal order. All that matters is that there is *some* prenominal counterpart to the dominant postnominal construction or vice versa.

<sup>12</sup> Thus, in Dryer's (2013b) massive but uncontrolled sample, languages with a clear order of noun and attributive adjective tend to have postnominal adjectives (878/1251 languages = 70.1%). In a more controlled sample, Dryer (1992: 95) finds that genera with prenominal adjectives outnumber those with postnominal adjectives only in Eurasia, so that postnominal adjectives come out as the preferred type across most macro areas as well.

oriented towards a certain participant before that entity is actually fully mentioned. This order of “gap before filler” (Hawkins 2004: 175) goes against the grain of the usual management of discourse referents, which is primarily anaphoric, i.e. backward-looking. Prenominal RCs have hence sometimes been argued to be less efficient for online processing than their postnominal counterparts (Antinucci et al. 1979, Hawkins 1994, 2004, 2014). All other types of RRC avoid this problem by introducing the nominal first and resuming it, implicitly or explicitly, in the ensuing clause.<sup>13</sup> The most explicit anaphoric devices are ...

- (i) **RELATIVE PRONOUNS**, which typically occur at the beginning of the postnominal RC, combining the function of a subordinator and an anaphoric pronoun. As argued by Comrie (2006) and documented in Comrie and Kuteva (2013), relative pronouns are rather common in Eurasia but rarely used elsewhere.<sup>14</sup> (31) below provides an example from Hungarian, where the clause-initial interrogative pronoun inflects for the case role of the relativized position inside the RC:

- (31) Hungarian (Uralic, Hungarian: Hungary; Kenesei et al. 1998: 38)

A könyv, [*amely-et* Anna olvas-ott], érdekes volt.  
the book **which-ACC** Anna read-PST interesting was  
‘The book that Anna was reading was interesting.’

- (ii) **RESUMPTIVE PRONOUNS**, i.e. in-situ pronominal representations of the antecedent inside the RC, as in (32):

- (32) Krongo (Kadugli-Krongo: Sudan; Reh 1984: 253)

Òdínkò bílyáatà [ŋ-òkúrò-n-tú i'íŋ].  
fall.M.PRFV child CONN.M-push.PRFV-TR-2SG **him**  
‘The boy you pushed has fallen down.’

Indexing-prominent languages, i.e. those with verbal agreement markers for more roles than the subject, may show an index on the RC predicate instead, as in (33):

- (iii) **RESUMPTIVE VERBAL INDEXES**

- (33) Modern Standard Arabic (Afro-Asiatic, Semitic; Holes 2004: 283)

?al-qiššatu [llati: qaraza-ha:]  
ART-story.F REL read.PST.3SG.M.SBJ-3SG.F.OBJ  
‘the story that he read (it)’<sup>15</sup>

Even where the most economical and hence most widespread technique of co-reference, namely a so-called **GAP** in the relativized position, is chosen, postnominal

<sup>13</sup> For further functional advantages of postnominal RCs over the other RC types, see Lehmann (1984: 405f.).

<sup>14</sup> In fact, the only language outside of Europe to use relative pronouns in Comrie and Kuteva’s (2013) 112-language sample is Acoma (Keresan: USA). Kuteva and Comrie (2006) argue that what looks like relative pronouns in some African languages often turns out to be a different phenomenon.

<sup>15</sup> As discussed in Comrie (1981: 220), resumptive pronouns and indexes are person forms whose presence is not required in the corresponding simple sentence. This is the case in both (32) and (33).

RCs are still often marked by a clause-initial subordinating morpheme that has its diachronic roots in a topical pronoun, notably a demonstrative, such as *inde* ‘this’ in Mayogo:

- (34) Mayogo (Atlantic-Congo, Volta-Congo, Ubangi: The Democratic Republic of the Congo; Sawka 2001: 170)

*Ma dje engú [inde mɛ nga-pa-e ma \_\_\_\_] de.*

1SG PST.hear affair REL 2SG PROG-say-REF 1SG NEG

‘I don’t understand (the) problem that you are talking to me about.’

Thus in Hendery’s (2012: 267ff.) sample, demonstrative and interrogative pronouns constitute the most common sources of subordinators of postnominal RCs (see also Kuteva et al. 2019: 144ff. and Diessel 1999: §6.3.2 for further examples). Therefore, while invariant markers like Mayogo *inde* are not synchronically pronominal and merely serve to mark the RC as such rather than a particular referent inside the RC, they are diachronically precisely the kinds of anaphoric elements we would expect in postnominal RCs.

The situation in prenominal RCs is conspicuously different, as has been pointed out by numerous authors. Wu’s (2011) recent survey of the syntax of prenominal RCs confirms the absence of relative pronouns, the prevalence of other types of subordinators (which we will turn to below), and a less systematic use of “resumptive” pronouns. As for the latter, it is telling, firstly, that in Comrie and Kuteva’s (2013) sample, all languages with resumptive pronouns have postnominal relative clauses. Wu (2011) makes out only three clusters of prenominal RC-languages in which resumption is used more regularly for certain relativized positions (notably oblique positions), namely Semitic, Chinese and “some Causasian languages” (Wu 2011: 594). Secondly, as noted by Lehmann (1984: 230), the resumptive elements of prenominal RCs can be markedly different from the anaphoric pronouns in postnominal RCs; in particular, they can take the form of reflexive pronouns (e.g. Turkish (Kornfilt 1997: 61), Korean (Sohn 1994: 66) or Tanti Dargwa (Sumbatowa and Lander 2014: 192), whose antecedent is unlikely to be found in the preceding sentence and thus expected in the upcoming matrix clause.

What these observations and the general filler-gap preference suggest is that prenominal RCs are the most difficult type of RC to grammaticalize from a sequence of referentially coherent (independent) clauses.<sup>16</sup> In keeping with this, we find that the internal structure of prenominal RCs resembles that of independent clauses much less commonly than that of postnominal RCs. On the contrary, they characteristically show dependent verb forms with reduced verbal and clausal operators, as in the nominalized RC from Dolakha Newar in (20) above and the participial RC from Lezgian in (30). In

<sup>16</sup> In her diachronic typology of RCs, Hendery (2012: 169) is very pessimistic with regard to paratactic structures being the origins of relative clauses more generally, as often promoted, for example, by Givón (e.g. Givón 2009). Hendery thus concludes that “in the case of relative clauses, an origin in paratactic structures may be the exception rather than the rule. In general, relative clauses seem to be far more commonly modeled on pre-existing complex sentences, either non-relative patterns in the same language, or relative clause patterns borrowed from other languages.” Whatever further investigations will bring to light with regard to parataxis as the source of RCs, it seems safe to assume that it yields postnominal (or adjoined) rather than prenominal RCs.



fact, of all RC types, the prenominal one is associated with the highest degree of desententialization and thus formal embedding into the matrix clause. In this respect, prenominal RCs also differ crucially from IHRCs: circumnominal RCs are typically ‘externally’ nominalized, i.e. they are sentential structures flagged by appropriate case markers and demonstratives at their right boundary; and correlative RCs are not normally nominalized at all, i.e. they, too, are basically sentential structures (Keenan 1985: 161).

Therefore, it has been proposed that prenominal RCs chiefly develop by processes of expansion of and substitution within nominal constructions rather than the integration of successive sentences in discourse. The expansion scenario accounts for RCs which carry morphological traits of other noun modifiers, such as attributive adjectives and genitives. Lehmann (1984: 376f.) mentions the Turkish *-An*, the Dravidian *-a* and the Dyrirbal *-ɲu* participles as representative examples (see also Aristar 1991 for diachronic relationships between adjectives, genitives and relative clauses). Along the same lines, Wu (2011: 581) points out that what looks like a general “complementizer” in some relatively sentential prenominal RCs is actually a genitive marker (e.g. Mandarin *de* or Alambalak *-ho*; see also Trask (1997: 247) on a possible genitive origin of the Basque relativizer *-en*). The general thrust of these examples is thus that a phrasal attributive construction becomes expanded into a clausal one, either by adding arguments and adjuncts to a deverbal adjective or by inserting a clause into the slot of a genitival or other modifying construction.

The second prominent diachronic source of prenominal RCs is that of participant nominalizations, such as agent-, patient- or various oblique-role nominalizations. A widely cited example comes from Lhasa Tibetan, where the suffixes used in relativization go back to lexical nouns in compounds:

- (35) Lhasa Tibetan (Sino-Tibetan, Bodic: Bhutan, Nepal; DeLancey 1986)
- a. *mkhan* ‘teacher, expert’
  - b. *s’i’n-mkhan* lit. wood-expert = ‘carpenter’
  - c. [*stag gsod-mkhan*] *mi*  
     tiger kill-NMLZ man  
     ‘the man who killed the tiger’

While *-mkhan* is, quite predictably, a relativizer for the subject position, locative and instrument relativizers harken back to the nouns for ‘place’ and ‘tool’, respectively. Similar uses of participant nominalizers for relative-clause formation have been reported for many languages, particularly in Asia (Yap et al. 2011) and the Americas (Comrie and Estrada-Fernández 2012). In fact, many of the “oriented participles” in Shagal’s (2019) world-wide study of relativization are participant nominalizers. Importantly, while the diachronic pathway from (35a)–(35c) involves successively larger structures, and in this sense an expansion, it still often proceeds from looser to tighter syntactic configurations at the final stages: as we saw in (22)–(24) from Ute, the RC interpretation arises when an expanded participant nominalization is placed in apposition to another nominal element, i.e. ‘man’ in (35c) and ‘woman’ in (23). At least at the beginning, this is a relationship of sociation rather than dependency (Lehmann 1983), with the nominalization constituting a separate NP rather than a modifier embedded into an NP (see also Comrie and Thompson 2007: 378). This is

nicely illustrated by Awa Pit, where the nominalization can be placed in either order with regard to the noun it modifies semantically (and can be completely separated from it), independently marked as a clausal topic and for its syntactic function in the main clause:

(36) Awa Pit (Barbacoan: Colombia; Curnow 1997: 287–288)

- a. [A-*mtu=mika*]=*na*      *ashaypa=na*    *wan*    *pyan-ti-zi*.  
 come-PTCP=NMLZ.SG=TOP    woman=TOP    all    hit-PST-NLOCUT  
 ‘The woman who was coming hit everyone.’
- b. [*Santos=ta*    *pyanta-ta=mika*]=*ta*      *pyan-ta-w*,      *ashaypa=ta*.  
 Santos=ACC    kill-PTCP=NMLZ.SG=ACC    hit-PST-LOCUT.SBJ    woman=ACC  
 ‘I hit the woman who killed Santos.’  
 (lit. ‘I hit the one who killed Santos, the woman.’)

In (36b), the RC and the ‘head noun’ are non-adjacent and hence they cannot form a constituent. In (36a), they are adjacent and hence the RC could be said to be ‘adnominal’, but it is still not embedded within a single NP. Thus while Awa Pit has retained the syntactic independence of the semantic modifier, Ute has kept this pattern (23) alongside a more integrated one where the nominalization is embedded as a RRC (24).

Example (36) is also revealing for another reason: while the stages in (35) suggest a diachronic expansion, it has recently been argued that nominalizers in prenominal RCs can also come from semantically light and hence frequent head nouns in earlier RC constructions, such as ‘thing’, ‘place’, etc. as well as nominal classifiers. Following the general “suffixing preference” (Bybee et al. 1990), these head nouns may come to cliticize to the RC to yield exactly the participant nominalizers that are characteristic of prenominal RCs. This would account for =*mika* in Awa Pit, which synchronically cliticizes to a participial suffix. Curnow (1997: 290) speculates that this may be a remnant of an earlier RC construction which is no longer grammatical by itself, and Epps (2009) argues that these developments may generalize to various languages in South America (see also Genetti et al. 2009 for similar developments in Tibeto-Burman languages).

Whatever the precise direction by which these nominalizations emerge, our claim in the present section is *not* that this pathway, just like the expansion of deverbal adjectives and genitive modifiers, is restricted to prenominal RCs. As Ute demonstrates, it can also underlie the genesis of postnominal RCs (see also Lehmann 1984 on German, Deutscher 2009a on Akkadian and Cristofaro 2014 on Ewe). But the current typological picture suggests that these developments of expansion, “insertion” (Kibrik 1992) and apposition are *relatively* more important, and perhaps even the only ones, for the emergence of prenominal RCs.

We round off our discussion of EHRRCs with a type of construction that is externally headed but attaches to the right of the matrix clause rather than the nominal head; in other words, this type of RC is not adnominal but **POSTPOSED** or **RIGHT-ADJOINED**. It can, however, appear in this position for two rather different reasons. On the one hand, a postposed RC can result from a performance option of extraposition, as introduced in (10) above from Slave. Very similarly, “it is common”

for the IHRRC from Choctaw in (25) above “to be extraposed to the end of the [sentence], leaving the head of the relative clause in situ” (Broadwell 2006: 301):

- (37) Choctaw (Muskogean: USA; Broadwell 2006: 301)
- |               |                 |                  |                                  |
|---------------|-----------------|------------------|----------------------------------|
| <i>Ofi-it</i> | <i>balii-it</i> | <i>kaniiya-h</i> | <i>[ipiita-li-k-aash-m-at]</i> . |
| dog-NOM       | run-PTCP        | go.away-TNS      | feed-1SG.I-TNS-PREV-NOM          |
- ‘The dog that I fed ran away (lit. went away running).’<sup>17</sup>

Where extraposition affects nominalized RCs (as in Slave and Choctaw), the result looks like the mirror image of the NP apposition pattern we saw in (36b) above, i.e. ‘The dog ran away, the one that I fed.’ Sometimes the pattern of extraposition is such that the whole relative construction, i.e. the RRC along with its head noun, is shifted to the right sentence boundary. In Evenki, for example, nominalized (or participial) RRCs are typically prenominal constructions, as in (38a), but the order can be changed in such a way that the modified nominal and the RRC appear sentence-finally:

- (38) Evenki (Tungusic: China, Mongolia; Nedjalkov 1997: 137)
- a. [*himat agi-tki tuksa-d’ari-va*] *asi-va*.  
 quickly forest-ALL run-PTCP-ACC.DEF woman-ACC.DEF  
 ‘a/the woman who quickly ran into the forest’
- b. *Bi archa-ø-m asi-va* [*himat agi-tki tuksa-d’ari-va*].  
 1SG meet-NFUT-1SG woman-ACC.DEF quickly forest-ALL run-PTCP-ACC.DEF  
 ‘I met a/the woman who quickly ran into the forest.’

While each language has its own specific set of motivating factors for RRC extraposition (see, e.g., Nefedov 2012 on Ket, Strunk 2014 on German, Rasekh-Mahand et al. 2016 on Persian), a factor that appears to play a role in all of the languages for which corpus data have been scrutinized is the relative weight of clausal as opposed to single-word or phrasal modifiers. By shifting the RRC, one can “minimize domains” (Hawkins 2004: 31) for processing the fundamental predicate-argument structure of the matrix clause. As this applies to languages of all basic word-order types, patterns of RRC extraposition are productive in both OV- and VO-languages.<sup>18</sup>

The second type of postposed RC is not due to a performance option of extraposition, but one that is canonically adjoined to the matrix clause in this position. In (13) above, we introduced this as a characteristic type of subordinate clause in Australian languages. And indeed, in Dryer’s (2013c) database, there are only two languages outside of Australia in which adjoined RRCs are the dominant RC construction, namely Mekens (Tupian) and Kuikúro (Cariban), both spoken in Brazil. Another example of an adjoined RRC from Australia is given in (39):

<sup>17</sup> Interestingly, this pattern of extraposition can be accompanied by the same kind of “case floating” we will see for Huallaga Quechua later on, with the nominal head receiving the case marker of the relativized position, e.g. ‘dog-ACC ran away [I fed]-NOM’.

<sup>18</sup> Some (randomly chosen) examples of such VO languages licensing extraposition from NP are Hausa, Arabic, Mekens, Karo Batak, French, Koyra Chiini, Mandarin Chinese, Mosestén, Ndyuka and To’aba’ita; corresponding OV languages (beyond the ones we have seen) include Georgian, Kalaallisut, Kiowa and Lakota.

- (39) Kayardild (Tangkic: Australia; Evans 1995: 512, 517)
- a. *Ngada birrmurdami [ngijin-inja kajakaja-ntha bukawa-thurrk].*  
 1SG.NOM sad.NOM my-COBL daddy-COBL die-IMM.COBL  
 ‘I can feel that my daddy has just died.’
- b. *Jina-a maku, [ngjuwa kurri-jurrk].*  
 where-NOM woman.NOM 1SG.SBJ.COBL see-IMM.COBL  
 ‘Where’s this woman who I’m seeing?’

The construction illustrated in (39) is characterized by special (“complementizing oblique”) case marking on all elements of the dependent clause. And as in other Australian languages, this subordinate clause has a number of different grammatical functions and hence interpretations in Kayardild, as reflected by (39a–b). But interestingly, when it is supposed to function as a relative clause, the free constituent order of Kayardild is usually adapted in such a way that the modified nominal appears as the last element of the matrix clause, effectively achieving an adnominal structure without centre-embedding it in the main clause (just as in (38b) from Evenki). According to Hale (1975), this is a signal of the “attraction” of the subordinate clause to a nominal element in the matrix: Hale argues that such adnominal environments at the right sentence boundary enable a reanalysis of adjoined subordinate clauses as postnominal modifiers, and when this pattern is analogically extended to non-final NPs in the main clause, the result is a (centre-)embedded RC. This, then, constitutes another way in which embedding can emerge from distributional adjunction. In fact, Hale’s “attraction”-and-extension scenario is also involved when postnominal RRCs develop from independent clauses with topicalized anaphoric pronouns (as discussed, for example, by Lehmann 1984, 2008 for Homeric Greek, Haader 1992 for Hungarian and McConvell 2006 for Ngumpin-Yapa). In both cases, the overall process is one of combining and successively integrating two clauses.

### 2.3 Orientation of the RRC

A defining property of RCs is that they are semantically oriented towards one of their NP positions. Therefore, a major parameter of cross-linguistic variation is how explicitly this orientation is coded morphosyntactically, so as to enable the unambiguous and efficient processing of the RC and the complex sentence of which it is part. The different types of RRC formation surveyed above present different challenges in this regard.

For IHRRCs, the major processing challenge is to discern the fact that a nominal element inside the RRC is meant to be modified by the clause as a whole; in other words, one needs to discern the head. This problem is illustrated by our earlier example (25) from Choctaw: as the translation shows, there are two possible ways of interpreting the relative clause, which results from the fact that there is no explicit means of orientation. It is thus not surprising that languages with IHRRCs often develop grammatical strategies of what Lehmann (1984: 207) calls **NUCLEUS FORMATION**. The most common strategy is to move the relativized NP to the beginning of the RRC; this is illustrated by our earlier example (14) from Supyire again, and found in a similar fashion in, for instance, Slave and Menya. Alternatively, speakers of Supyire can also leave the relativized NP in situ, but then it needs to be marked by a relative determiner (Carlson 1994: 491ff.), just as we saw for Bambara in (28) and in

the typical correlative diptych in (27). Further means of nucleus formation include special agreement markers on the RRC verb (Lavukaleve), an accompanying indefinite article (Lakota) or tonal adjustments (Jamsay).

For EHRRCs, the processing challenges are, in a way, the opposite of those in IHRRCs: the modified noun is given externally, but now the crucial task is to figure out the relativized position inside the RRC. This is sometimes called the “CASE RECOVERABILITY PROBLEM” (Givón 2001: 180), and it becomes apparent whenever the RC does not contain an explicit representation of the relativized NP, i.e. in what we introduced as gapping above. The syntactic function associated with such gaps can be recovered comfortably in languages with (i) high “referential density” (Bickel 2003), i.e. overt coding even of contextually salient participants, and/or (ii) fixed constituent order and/or (iii) the possibility of stranded case markers. In Abun, for example, constituent order is rather rigidly SVO, and adpositions can be left in situ without an overt complement:

(40) Abun (West Papuan: Indonesia; Berry and Berry 1999: 153–156)

- a. *ii bi buku [gato an gre \_\_\_\_\_ mo bot]*  
 1SG POSS book REL 3SG burn LOC fire  
 ‘my book that she burned in the fire’
- b. *Ji mu kenyak nu [gato Isak sem mo \_\_\_\_\_ ne].*  
 1SG go locate house REL Isak sleep at DET  
 ‘I went to find the house that Isak slept at.’

In Cibak, gapped core arguments may leave their case marker behind, but instead of being stranded in its usual position, the marker gravitates towards the subordinator introducing the RC:<sup>19</sup>

(41) Cibak (Afro-Asiatic, Chadic: Nigeria; Fraijzyngier 1987: 436)

- mwala [nam \_\_\_ tə zər ni tara tsar-ba]*  
 woman REL OBJ boy DEF SBJ choose-VENT  
 ‘the woman whom the young man has chosen’

And for Huallaga Quechua, Weber (1989: 41) describes a phenomenon of “case floating,” whereby an oblique case affix can be stranded at the nominal head of the entire RC construction. This applies to the comitative suffix *-wan* in (40):<sup>20</sup>

<sup>19</sup> The resulting adjacency of a relativizer and a case marker comes close to typical resumptive pronouns, but Comrie (2006) argues that they constitute a phenomenon in its own right (chiefly because the relativizer is not a pronoun of any kind).

<sup>20</sup> At first sight, this marking pattern appears inefficient since an internal syntactic function is marked externally. However, Weber (1983: 45) argues that this is still more transparent for reconstructing the RC situation than gapping the comitative phrase altogether. Furthermore, the semantic and syntactic relations between ‘clothes’ and ‘drying’ in the matrix clause in (42) are entirely clear and leave no room for the comitative marker to ‘interfere’ with their interpretation. Following this train of thought, one might hypothesize that an oblique marker on the head nominal would work best (i.e. interfere with processing the least) if the NP has the *same* oblique role in both the main and the relative clause. And indeed, there are languages which use gapping for oblique roles only under such role parallelism (see, e.g., Saltarelli (1988: 38) and Aldai 2003 on Basque) and require alternative strategies (e.g. resumptive pronouns) if no such parallelism is given.

(42) Huallaga Quechua (Quechuan: Peru; Weber 1983: 42)

[ $\emptyset$  *yaku-man* *yayku-shan*] *roopa-wan* *chakikuykan*  
 $\emptyset$  water-GOAL enter-NMLZ clothes-COM 3PL.are.drying  
 ‘the clothes with which he entered the water are drying’

In these examples, perhaps most clearly in (38), it makes sense to speak of a gap *strategy*, as is often done, but where constituent order is more flexible and adposition stranding is ungrammatical, the gap itself cannot be called a strategy for role recovery (incidentally, this applies to most RCs in Huallaga Quechua as well). In such cases, the relativized position needs to be identified on the basis of other devices.

One such tool can emerge directly as a by-product of grammaticalization, namely when different participant nominalizers (or contrasting participles) come to be associated with different relativization sites. This is what we saw in (35) from Tibetan above, and a similar system is found, for example, in Mapudungun. (43) shows the agent nominalizer *-t* yielding relativization on a subject and the “objective verbal noun” in *-el* yielding relativization on an object, but there are several other nominalizers in addition to these.

(43) Mapudungun (Araucanian: Argentina, Chile; Smeets 2008: 216, 200)

- a. *Kim-nie-fi-y-m-i* *ti* *wentru* [*mi* *leli-nie-e-t-ew*]?  
 know-PROG-OBJ-IND-2-SG the man 2SG.POSS watch-PROG-OBJ-AVN-SBJ:DAT  
 ‘Do you know that man who is watching you?’
- b. *Fey-chi* *chanchu* [*eymi* *mi* *ngilla-el*] *trongli-le-y- $\emptyset$* .  
 that-ADJZ pig you 2SG.POSS buy-OVN lean-STAT-IND-3  
 ‘That pig you bought is lean.’

In a similar fashion, many languages distinguish (at least) between a subject-extracting affix (often an agent nominalizer or an ‘active participle’) and a non-subject-extracting affix (e.g. from a patient or an action nominalization or a ‘passive participle’), e.g. Quechua, Wolaytta, Dolakha Newar, Finnish, Aguaruna, Kolyma Yukaghir and Maricopa. Some languages, like Yuracaré, exploit their switch-reference markers to make exactly the same contrast, i.e. here, too, a grammatical system that evolved for the purposes of reference tracking can function as a cue to serve the role recovery problem when such constructions are used for relativization. All of these cases have thus been referred to collectively as demonstrating a “verbal affix” strategy (e.g. Comrie 2003).

Considering the potential processing difficulties and the nominalizing/adjectival origin of prenominal RRCs we discussed above, one would expect that such role-distinguishing dependent verb forms are particularly common in prenominal RRCs; however, Shagal (2019) finds exactly the opposite: in her study, languages which *only* have “contextually oriented participles” (i.e. verb forms that can be flexibly oriented towards several core roles) strongly tend to have them in prenominal RCs (28:6 languages), while languages which *only* have role-specific verb forms are more likely to have them in postnominal RCs (23:12 languages).

A similar observation can be made with regard to resumptive pronouns, the second device that can be used when a gap is judged not to be sufficient for role recovery. In their seminal study, Keenan and Comrie (1977) established an implicational hierarchy

for the distribution of gaps and resumptive pronouns in RRCs. A modified version of this hierarchy, adapted from Lehmann (1984: 211), is provided in (44):

- (44) The Accessibility Hierarchy (AH) of relativization<sup>21</sup>  
 SBJ/ABS > DO/ERG > IO/‘time’, ‘place’ > oblique complements > adjuncts

According to (44), whenever a language uses resumptive pronouns instead of a gap for a certain position on the AH, it will also use resumption on all lower positions. This is a statistical generalization that applies to all types of RRC. However, it is noteworthy that pronominal RRCs do not only make less frequent and less systematic use of resumptive pronouns (as we mentioned above), but that their resumptive elements do not extend as highly on the AH than those of postnominal RRCs: it is very uncommon to find pronominal RRCs with resumptive elements for core syntactic functions (S, A, P), whereas resumptive pronouns for P are often at least licensed, if not required, in postnominal RRCs (e.g. East Ambae, Arabic, Kera, Persian, Tongan).

It appears, therefore, that pronominal RRCs are overall less prone than postnominal RRCs to code their orientation explicitly, given the lack of relative pronouns, less systematic use of resumption, use of unoriented verb forms, infrequent case stranding and flexibility of constituent order in the OV languages in which they typically occur.<sup>22</sup>

In the absence of orienting devices, the relativized position of EHRCs can be “infer[red] by subtraction” (Givón 2001: 184), notably from the semantic and syntactic requirements of the RRC predicate and the arguments that are already supplied in the RRC. In this connection, it is noteworthy that some languages restrict gaps to those arguments that can be indexed on the dependent verb (e.g. Ungarinjin, Basque) or restrict relativization to the core syntactic functions more generally (e.g. Jarawara, Mangarrayi). Furthermore, where resumption is applied across argument and adjunct positions (e.g. Hebrew), corpus evidence shows that it is significantly more common for adjuncts (Ariel 1999). All of this suggests, in accordance with (44) above, that RRCs are most frequently oriented towards core syntactic positions, and hence core participants, whose involvement in the RRC event is easiest to retrieve and in this sense most “accessible”. Deviations from these most accessible roles correlate typologically with special marking strategies, either more explicit ways of indicating the relativized position (i.e. various forms of resumption) or by deriving relativizable core roles through valency-changing operations. The latter strategy is illustrated in the following example from Yakan, where only S and P are relativized on directly while A and lower roles require appropriate means of promotion to a higher argument position:<sup>23</sup>

---

<sup>21</sup> Keenan and Comrie’s (1977) original hierarchy included possessors and objects of comparisons as the two lowest roles. But as pointed out by Lehmann (1984: 211), one needs to distinguish at least between an *ad-verbal* AH (where the relativized position is a syntactic function in a *clause*) and an *ad-nominal* AH (where the relativized position is a syntactic function in a *phrase*, like possessor or standard of comparison). In (44) above, we limit our attention to the *ad-verbal* AH.

<sup>22</sup> This also ties with the observation that pronominal RCs are often more general “noun-modifying constructions” (Matsumoto et al. 2017) whose specific semantic link to the nominal head can be highly diverse and needs to be inferred from the context.

<sup>23</sup> Since such restrictions on the relativized position target sets of arguments, i.e. {SA} or {SP}, relative clauses are also relevant to the typology of grammatical relations and alignment types (see Bickel 2011 for

(45) Yakan (Austronesian, Malayo-Polynesian, Sama-Bajaw: Philippines; Brainard and Behrens 2002: 166–167)

- a. {A} becomes {derived-S} by means of antipassivization and then takes the nominalized relative clause appropriate to {S}

*Iyan nakanak [ma-N-polong buwa'-buwa'-in].*  
 that child NMLZ-ANTIPASS-break toy-DEF  
 'That is the child who broke the toy.'

- b. {OBL} recipient ('to the dog') becomes {derived-P} by means of applicativization and then takes the non-nominalized relative clause appropriate to {P}

*Iyan asu [-in-teppad-an-ne tolang-in].*  
 that dog TR-throw-APPL-3SG.ERG bone-DEF  
 'That is the dog that she threw a bone to.'

There has been considerable debate in the literature as to how exactly constraints on the relativized position can be explained. On the one hand, there are discourse-pragmatic preferences that have been argued to lead to different degrees of accessibility even within the core roles (S, P > A, see Fox 1987). On the other hand, there is a wealth of psycholinguistic research attempting to motivate accessibility differences between the core roles (subject > object) and between these and the non-core roles (see Norcliffe et al. 2015 for an overview of recent work). But as summarized succinctly in Hawkins (2014: 22–28), even the processing motivations are multifactorial in nature, involving several different ways in which the lower roles on the AH create more processing efforts and hence profit from more explicit coding strategies. While these different facets of processing are beyond the scope of the present overview, we can conclude our discussion of RRCs with a more general observation related to language processing. It has been argued, most prominently by Hawkins (2004, 2014) again, that OV-languages face conflicting pressures in the grammatical expression of relativization. On the one hand, these languages favour structures in which dependent elements like NP complements of verbs and adpositions as well as subordinate clauses appear before their respective heads; this would favour the grammaticalization of prenominal RRCs. On the other hand, the nature of reference management in discourse favours anaphoricity (“filler before gap”) and hence actually disfavours prenominal RRCs. The result of these conflicting pressures is what we have seen throughout the present section, namely a greater diversity of RRC types in OV-languages: the internally-headed types deliver a dependent clause early in

---

a recent overview). As “argument selectors” (Witzlack-Makarevich and Bickel 2019), they can be systematically compared to other grammatical constructions like case markers, indexes, control constructions, etc. Kazenin (1994), for example, proposed the following implicational hierarchy for ergative alignment (see also Croft 2001: 155):

case > indexation > relativization > purpose clauses > coordination

According to this hierarchy, if a language exhibits gaps in relative clauses that define an ergative argument set, i.e. which apply to S- and P-arguments to the exclusion of A, the language also exhibits ergative alignment in verb indexation and case marking (but not necessarily in purpose clauses or coordinating constructions). This is true for Yakan above. However, it remains to be explored whether this generalization also applies to other relativization strategies. Urarina (isolate: Peru), for example, has distinct relative affixes for {SP} and {A}, but it does not exhibit ergative alignment in either case marking or indexation (Olawsky 2006: 325ff.); see also Shagal (2019) for similar findings.



a sentence while adhering to the filler-gap principle; the postposed RRC also ensures that fillers come first while still allowing efficient processing of all main-clause dependencies; and the two major externally-headed RRC types resolve one of these pressures each (see also Lehmann 1984 and Dryer 2013e). In VO-languages, by contrast, no conflicting pressures arise in the first place, and different pathways of grammaticalization all converge on the postnominal type(s) most straightforwardly.

#### 2.4 Outlook: Supplementary relative clauses (SRCs)

Supplementary (or “elaborative” (Lichtenberk 2008: 1064)) relative clauses do not serve to restrict the reference of an NP, but “have the communicative intent of parenthetical assertions” (Givón 2001: 179). Put differently, SRCs typically encode backgrounded but non-presupposed information on an element in the main clause. We have seen above that such constructions may actually pave the way for the emergence of RRCs, and this would seem to require that, apart from intonation, there is little structural difference between the two RC types. Indeed, it is easy to come by descriptions from typologically and geographically widely diverse languages which emphasize precisely the lack of formal differentiation between restrictive and supplementary interpretations of their primary RRC construction.<sup>24</sup> What is noteworthy here is that, contrary to earlier assumptions in the literature, this formal identity applies to RCs of all major types surveyed above: while it would seem natural for parenthetical assertions to be made after the NP in question, and thus by postnominal RC constructions, it is not impossible for prenominal and internally-headed RCs to fulfil such a supplementary function. We can illustrate this by the non-restrictive counterparts to restrictive examples we have encountered before (see (29) and (30) above):

- (46) Lezgian (Nakh-Daghestanian, Lezgitic: Azerbaijan; Haspelmath 1993: 343)

*Ada-z [bulax.di-n pataw weq' ne-zwa-j] balk'an aku-na.*  
 he-DAT well-GEN beside grass eat-IPFV-PTCP horse see-AOR  
 ‘He saw the horse, which was grazing beside a well.’

- (47) Cabecar (Chibchan: Costa Rica; Gonzáles Campos and Lehmann 2019: 17)

*[i dawá dul kal jula nā] ká jék dalí-n-é.*  
 3 brother-in-law POS.stand tree hand/arm INESS NEG REFL MOVE-DYN.MID-NEG.PRFV  
 ‘His brother-in-law, who was standing on the branch, did not move.’

Having said this, there are languages (i) in which the RC construction is always restrictive (e.g. Lavukaleve (Terrill 2003: 398)); (ii) in which RRC and SRCs do show some formal differentiation. The SRC in (4) above from Basque is flagged by a clause-final determiner which is not attached to the predicate of a RRC (which is otherwise grammatically the same) and it does need to follow the NP it relates to (in contrast to the RRC, which is normally prenominal). This effectively turns the SRC into a nominalized but headless RC (‘the (one) that’). In English and Hausa, similarly, SRCs do not normally use the invariant relativizer *that/dà* but a relative pronoun (‘who’,

<sup>24</sup> For some examples spanning the major macro areas, see Berghäll (2015: 360) on Mauwake, Epps (2008: 691) on Hup, Gonzales Campos and Lehmann (2019: 18) on Cabecar, Haspelmath (1993: 343) on Lezgian, Nedjalkov (1997: 35) on Evenki and Timberlake (2004: 209) on Russian.

‘which’), yielding a similarly headless (or “free”) relative (Newman 2000: 542). And the final stage along this continuum of adding head-like elements is what we see in Rapa Nui, where RCs cannot be used in supplementary function directly; instead, a semantically appropriate nominal like ‘person’, ‘place’ or ‘thing’ is placed in apposition to the NP to be elaborated, and it is this nominal that can take the usual RRC:

- (48) Rapa Nui (Austronesian, Central-Eastern Malayo-Polynesian, Oceanic: Chile; Kieviet 2017: 273)

*He turu a Rovi, he tajata [hāpa’o i te poki].*  
 NEUT.ASP go.down PROP Rovi NEUT.ASP person care.for ACC ART child  
 ‘Rovi, the man who took care of the child, came down.’

Some other structural differences will be mentioned exemplarily: In Abun, illustrated in (40) above, the relativizer *gato* shortens to *to* in SCRs and the clause-final determiner becomes obligatory, so as to signal the right boundary of the parenthetical insertion. In Western Balochi (and its close relative Persian), the SCR as such is structurally identical to an RRC, but the NP it elaborates loses its indefiniteness suffix which normally identifies the nominal head of a RRC construction (Axenov 2006: 253). In Japanese, it is the position of demonstrative determiners like *sono* that marks a subtle difference between RRCs and SRCs: if the linear order is DEM [REL] N, only a restrictive interpretation is possible, but if the order is [REL] DEM N, both readings are allowed (Ishizuka 2008: 3). And perhaps more generally, SRCs seem averse to extraposition, typically having to occur directly adjacent to the NP they elaborate (e.g. English and Hungarian (Kenesei et al. 1998: 38)).

A rather rare and hence also understudied type of SRC is the one from (5) above with a sentential antecedent. Its cross-linguistic rarity is certainly connected to the fact that it functions like an afterthought and, as such, can easily be rendered by non-embedded structures, such as paratactic clause combinations or successive sentences in discourse. The essential questions for typological research are, therefore, whether a language can extend its RC construction(s) to different kinds of non-restrictive environments, up to a sentential antecedent, and whether these different antecedents require structural modifications. In Hebrew, for example, the canonical RRC type can also attach to referentially unique heads like personal pronouns, yielding a SRC interpretation (e.g. ‘I, that don’t have any experience’); sentential antecedents, by contrast, need an overt nominal anchor in the matrix clause (Glinert 1989: 370), just like SRCs with nominal antecedents in Rapa Nui above:

- (49) Modern Hebrew (Afro-Asiatic, Semitic: Israel; Glinert 1989: 370)

*Hu niftar pitom, \*(ma/davar) [she-ziazia otam].*  
 he died suddenly, what/thing REL-shocked them  
 ‘He died suddenly, which shocked them.’

Again, it has been claimed that sentential SRCs cannot occur in SOV languages (de Rijk 1972: 69), but more research is needed to investigate this issue properly. In Basque, for example, the SRC from (4) elicits mixed grammaticality judgements when used with a sentential antecedent (Hualde and Ortiz de Urbina 2003: 807), so it remains to be seen what the exact typological generalizations will be.

### 3. Adverbial clauses in typological perspective<sup>25</sup>

For the purpose of cross-linguistic comparison, adverbial clauses (ACs) are commonly defined extensionally, as clauses which encode (more or less explicitly) one or more of the following semantic relations, henceforth adverbial relations (ARs):

- temporal simultaneity (WHEN/WHILE), anteriority (AFTER), posteriority (BEFORE), end points (UNTIL) and beginnings (SINCE)
- realis, irrealis and counterfactual CONDITIONALITY
- CAUSE and REASON<sup>26</sup>
- PURPOSE
- RESULT
- CONCESSION
- MANNER

For reasons of space, we will deal exclusively with ACs in a modifying function, leaving aside their uses as supplementary adjuncts (as in (6)–(8) above; see, e.g., the papers in Tsunoda 2018 for relevant work).

#### 3.1 Coding and position of ACs

Just as some languages have been reported to make very little use of relative clauses in discourse (e.g. Donohue 1999 on Warembori), some have been said to make little use of grammaticalized adverbial clauses, even if the latter are available in the language system (see Deutscher 2009b on Akkadian and similar assessments for Lango (Noonan 1992), Warao (Romero-Figeroa 1997) and again Warembori). In Akkadian, for example, ARs are commonly inferred from an underspecified and hence highly versatile coordinate clause combination (*-ma* ‘and’). And even where ARs are coded more explicitly, some of them may behave formally like coordinate constructions (see also Verstraete 2007 and Cristofaro 2014 on this issue) or simply be rendered by coordinate alternatives (Mithun 1984). This is particularly common with result clauses and concessive clauses: for example, the logical relation expressed by concessive clause linkage can equally well be expressed by adversative (BUT) coordination (see Dixon (2009: 28–30)). It is not surprising, therefore, that distinct ACs of result and concession are less frequently grammaticalized than those encoding the other ARs from above.

ARs are most commonly encoded by a specific marker at the boundary of the AC, such as an adverbial conjunction (as in (2), (8), (9) and (11) above) or a bound marker

---

<sup>25</sup> Monographic treatments of the typology of adverbial clauses include Kortmann (1997) on European languages and Hetterle (2015) with a world-wide scope. For non-monographic typological surveys of the adverbial domain, the reader is referred to Thompson et al. (2007) and Dixon (2009). In contrast to these latter overviews, we will not survey the different semantic relations and their coding properties individually, but focus instead on the comparison of the major adverbial relations from the perspective of the different facets of embedding introduced in §1.2 (see also Diessel 2013 for a like-minded overview).

<sup>26</sup> The terms CAUSE and REASON are not always distinguished systematically, but where they are, the term CAUSE tends to be reserved for external causality (*She couldn't go skiing because she had broken her leg*), while the term REASON captures internal causality, i.e. motivation (*She left the party because it was boring*). We will have to gloss over this distinction here, but Verstraete (2008) shows that it has empirical typological consequences.

that attaches to the AC verb stem (18), to a nominalized or other dependent form of the verb (19) or to the whole clause:

(50) Trumai (isolate: Brazil; Guirardello 1999: 391)

[Kumaru-*k*      *tete-kma-n-es*]                                  *hen hi wa-pata.*  
 Kumaru-ERG    body.paint-PRFV-3ABS-**when**    then 2    *wa-arrive*  
 ‘When Kumaru finished body-painting her, you arrived.’

In keeping with the general suffixing preference for grammatical material (Bybee et al. 1990), bound adverbial markers tend to be enclitics or suffixes (Dryer 2013a), but proclitics and adverbial prefixes are also occasionally attested (Grossman et al. 2018).

As laid out in detail by Hetterle (2015: Ch. 5), there are recurrent patterns of co-expression among adverbial subordinators. Particularly prominent overlaps (and hence diachronic processes of reanalysis and analogical extension) are found between (i) WHEN and conditional relations, (ii) purpose, cause and result relations (sometimes including temporal UNTIL), and (iii) manner/instrument and WHEN/WHILE relations (see also Kortmann 1997, Martowicz 2011). In languages with highly polyfunctional (i.e. semantically underspecified) dependent clauses, such as the “generalized subordinate clauses” common in Australian languages, the specific AR may be cued by other grammatical devices (see Hetterle 2015: 39ff. for discussion). Chief among those are specific TAM markers and the position of the dependent clause. In Yuracaré (isolate: Bolivia), for example, a versatile same-subject irrealis clause can receive a temporal WHEN and a conditional interpretation in preposed position and is interpreted purposively in postposed position; and the temporal interpretation can change from WHEN to BEFORE when the AC is additionally marked for future tense (van Gijn 2006: 290ff.). In Haruai (Piawi: Papua New Guinea), a cosubordinate construction normally occurs in typical clause chains with a loosely successive interpretation (‘and then’), but it can be centre-embedded in the matrix clause if it is meant to signal a more specific AR, such as a particular temporal interpretation (Comrie 2008b: 15). This shows, once more, an interesting correlation between the functional and the distributional dimension of clausal embedding: the more a dependent clause is conceived of as an adverbial modifier of a matrix clause, the more it can be distributionally integrated with that clause.

This observation leads us to the more general point that the semantic nature of the AR often engenders typical discourse-pragmatic functions of the AC in question, and both of these factors give rise to characteristic ordering patterns of ACs cross-linguistically. In Table 1, we juxtapose corpus data on the position of different types of ACs in Barbareño Chumash with the corresponding cross-linguistic data from a recent study by Diessel (2019a):

Table 1. Positioning patterns of ACs in corpus data and a typological sample

AR	Barbareño Chumash corpus (Wash 2001: 345)			Diessel's (2019a: 101) sample of 100 languages (numbers count constructions)			
	PREPOSED	POSTPOSED	TOTAL	PREPOSED	FLEXIBLE	POSTPOSED	TOTAL
TEMPORAL	142 (66%)	72 (34%)	214	119 (60%)	68 (34%)	12 (6%)	199
CONDITION	76 (86%)	12 (14%)	88	94 (91%)	9 (9%)	0 (0%)	103
CAUSE	3 (6%)	47 (94%)	50	40 (35%)	24 (21%)	49 (43%)	113
PURPOSE	8 (9%)	83 (91%)	91	33 (29%)	19 (17%)	63 (54%)	115
CONCESSION	3 (27%)	7 (73%)	10				
MANNER	2 (13%)	13 (87%)	15				

Both types of data show that temporal and particularly conditional relations are biased towards preposed position, while causal and purposive clauses gravitate towards postposed position. The postposing preference of concessive and manner clauses in Chumash, a VO-language, may be due to an overall proclivity of dependent clauses towards the postposed position; from a typological perspective, concessive and manner relations are more typically coded by preposed ACs (Hetterle 2015: 124). And result clauses, which are missing in Table 1, show the strongest cross-linguistic bias towards the postposed position (*ibid.*). In Ojibwe (Algonquian: USA, Canada), for instance, ACs show very flexible orders, but result clauses strictly follow the matrix clause (Valentine 2009: 207). Importantly, when the category of time clauses is broken up into more specific temporal relations, the picture usually changes such that UNTIL- and BEFORE-clauses deviate from the general preposed trend (in Chumash, for example, 93% of all UNTIL-clauses are placed after the matrix clause). This is all the more striking when it happens in verb-final languages that generally keep dependent clauses in preverbal position. Thus in Akkadian, it is only UNTIL- and BEFORE-clauses that can appear in postposed position while all other ACs precede the matrix clause (Deutscher 2009b: 64).

The primary semantic motivation for these ordering preferences is that of iconicity of sequence. As already observed by Greenberg (1963), AC order typically mirrors the underlying cause-effect or antecedent-consequence order of the AR in question, and this would account for the positioning preferences of conditional, SINCE/AFTER/BEFORE/UNTIL as well as purpose and result clauses (see also Diessel 2008). In addition to semantic considerations, however, there are important aspects of discourse organization that affect the placement of ACs (e.g. Chafe 1984, Thompson 1985, Ramsay 1987, Verstraete 2004). As Wash (2001: 551) remarks on her corpus of Chumash, “initial adverbial clause tokens often played a global role in structuring the discourse, while the final adverbial clauses tended to be limited to the more localized role of qualifying the [matrix] clause.” The discourse-structuring or grounding function is particularly prominent with WHEN- and conditional clauses (Haiman 1978), and when other ARs assume such expository roles, they are sometimes explicitly marked with topic morphology (e.g. preposed purpose clauses in Awa Pit (Curnow 1997: 280)). Conversely, the fact that causal clauses are often used to provide a more “local” justification for an immediately preceding statement (Diessel and Hetterle 2011) may explain their anti-iconic ordering preference across languages: As cause-effect relationships, we would expect causal clauses to precede their matrix, but as Table 1 shows, this is strongly dispreferred in Chumash and mildly dispreferred from a

typological perspective. In contrast to temporal, conditional and concessive relations, then, causal relations are less frequently used to provide an interpretative framework for the ensuing matrix clause (see also Ford 1993, Thompson et al. 2007: 280ff.). Similarly to purpose and result clauses, they rather tend to encode new information. In fact, when languages have several types of causal clauses that differ in their discourse-pragmatic functions, it is usually the case that preposed causal clauses encode presupposed information ('since') while postposed causal clauses are reserved for new or focal information ('because'), as in Iquito (Zaparoan: Peru; Michael 2009: 157f.).

In sum, we have seen that the semantic and discourse-pragmatic functions of the different ARs naturally lead to mixed ordering patterns in the adverbial domain. The only factor militating against such heterogeneity is a consistent ordering preference for all ACs, and this appears to happen virtually only in strong verb-final languages (like Korean) that keep all ACs in preverbal position (see Diessel 2019a: 100). This makes ACs crucially different from relative and complement clauses, which have an overall cross-linguistic bias towards postnominal and postverbal position, respectively, and where languages exhibit mixed ordering patterns of RCs or CCs, they will be OV-languages. In other words, homogeneity of clause order is typical of RCs and CCs in VO-languages, but for ACs, it can only be found in OV-languages (if at all). VO-languages always allow at least some types of AC to precede the matrix clause.

### 3.2 Desententialization of ACs

Another important aspect of the typology of ACs is the systematic correlation between specific ARs and their relative degree of desententialization. Cristofaro (2003) and Hetterle (2015) provide ample empirical evidence for a downgrading hierarchy of ARs, i.e. a systematic implicational scale which allows us to predict that the degree of desententialization of any given AR is at least as pronounced as that of all ARs higher on the hierarchy. Fig. 1 is the empirical result of Hetterle's (2015) study; the different positions on this hierarchy show statistically significant differences in their mean degree of desententialization across languages, while the ARs that share a position do not differ in a significant way:

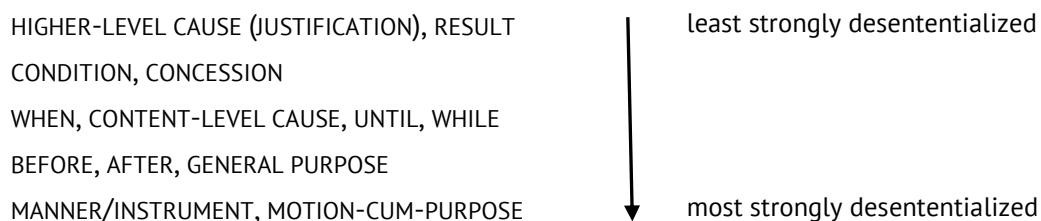


Figure 1. The downgrading hierarchy of ACs (adapted from Hetterle 2015: 179)

There are (at least) two strands of functional explanation that can account for these patterns of grammatical coding, and they are convergent rather than competing motivations.

Firstly, building on a proposal by O'Dowd (1992), both Cristofaro (2003) and Hetterle (2015) invoke the idea that some ARs are open to reification, i.e. a construal as a THING in Langacker's (1987) sense. In keeping with this hypothesis, Hetterle (2015: 84) finds that the different ARs are unevenly associated with nominalized verb forms:

AFTER, BEFORE, purpose, cause and WHEN relations show a significantly higher degree of nominalization than, for example, condition, result and WHILE relations. Cristofaro uses this event-as-object metaphor to explain why conditional relations outrank all other ARs in her study in terms of the downgrading of the verb form, as there are indeed languages (like Mapudungun) in which conditional clauses are more sentential than any of the other ACs. Conversely, it is easy to see how some of the strongly desententialized ARs can emerge as expansions of NPs and PPs (see also Genetti 1991). In (51), we have a nominalized causal clause with ablative case morphology that treats the event of the AC as the metaphorical source for the event in the matrix clause:

- (51) Kolyma Yukaghir (Yukaghir: Russian Federation; Maslova 2003: 432)
- [*Tine tet ibil'ōl-**get***] *ediŋ nodo-pe āj kel-ŋi.*  
 recently you cry-NMLZ-ABL this bird-PL CONN come-3PL:INTR  
 'Because you cried before, these birds came again.'

In a similar fashion, purposes can be seen as metaphorical goals (Lakoff and Johnson 1980), and events can also easily mark temporal reference points (AFTER, BEFORE, SINCE, UNTIL). This is again reflected in their diachronic origins: Purpose clauses commonly bear allative, dative or benefactive marking (Thompson et al. 2007, Rice and Kabata 2007, Schmidtke-Bode 2009), and AFTER-, BEFORE-, SINCE- and UNTIL-clauses are transparently based on the corresponding adpositions or case affixes (e.g. Harris and Campbell 1995: 291f.). This line of explanation is also consistent with the fact that converbs (i.e. special adverbial affixes rather than case-marked nominalizations) are preferably found with ARs that are not as easily conceptualized in nominal terms, notably manner and WHILE-clauses (Hetterle 2015: 92f.).

Secondly, the different ARs exhibit different degrees of what we called functional embedding in §1.2 above: Whilst ACs are generally not seen as argument clauses (because they are not selected by a matrix predicator), some ACs do contract a closer semantic relationship with that predicator than others. In the same way that, for example, aspectual marking affects verb meaning more than tense marking and is hence more “relevant” and in closer proximity to the verb stem (Bybee 1985), some ARs elaborate more crucial aspects of the matrix situation than others. Manner clauses, to begin with, directly specify in which the verbal action of the matrix is carried out and are hence particularly relevant, in Bybee’s sense, to the matrix verb. Time clauses are sometimes explicitly referred to as “temporal arguments” (see, e.g., Van Valin 1984, who uses this criterion to say that temporal ACs are embedded in the main clause while chaining constructions are not). Incidentally, this is reflected in the ordering pattern in our earlier example (12) above, where the temporal clause attaches to the core of the matrix predication while the causal clause appears only afterwards (see Van Valin’s (2005: Ch. 6) difference between ‘ad-core’ and ‘ad-clausal’ relations). Finally, a particular subtype of purpose clause, the so-called “motion-cum-purpose” construction (Aissen 1984), could likewise be seen as an argument of a matrix verb of directed motion: Wherever we go deliberately, it is for the sake of carrying out an intended action, and hence it is not surprising that many languages have a distinct motion-cum-purpose construction which formally behaves like an argument clause (Schmidtke-Bode 2009: 94ff.). In line with Fig. 1, we thus find that motion-cum-

purpose clauses can be more desententialized than general purpose clauses, as in the following example:

(52) Huallaga Quechua (Quechuan: Peru; Weber 1989: 117)

- a. [Miku-*na-nchi:-paq*]            *yanu-ku-shka-:*  
 eat-NMLZ-1INCL.POSS-PURP cook-REFL-PRF-1  
 ‘I cooked it so that we could eat it.’
- b. [Rika-*q*]    *aywa-shka-:*  
 see-NMLZ see-PRF-1  
 ‘I cooked it so that we could eat it.’

As can be seen, the construction in (52b) employs a different (and shorter) nominalizer than the one in (52a) and the purposive marker as such is optional and often missing; we will return to this observation in §3.3. Note, however, that even beyond motion-cum-purpose clauses, purposive relations are quite different from the remaining ARs in that they involve an intentional component (Verstraete 2008), fixed time reference (relative to the matrix clause) and a high probability of subject sharing (Schmidtke-Bode 2009); therefore, “the semantic difference between purpose relations and the other ARs is greater than the difference between any of the other ARs” (Cristofaro 2003: 179), and the features involved all favour the omission of grammatical information (TAM, person marking) in purpose clauses and hence a higher degree of desententialization.<sup>27</sup> As a result, purpose clauses are prone to syntactic reduction (e.g. control constructions) even in languages without explicit desententializing morphology. Overall, we thus repeatedly find systems of ACs in which purpose clauses stand out as the only structurally deranked, or a significantly more deranked, construction type (e.g. in Lango, Mekens, Ungarinjin, to name but three typologically and geographically diverse languages).

The two parameters of nominal construal and elaboration of the matrix predicator are logically independent, but they converge on the same result, namely ACs that behave formally and functionally like typical constituents of the main clause. As such, they also stand a higher chance to be distributionally embedded in the main clause, as illustrated by the centre-embedded purpose clause in (19) above. What we thus observe is that the more downgraded types of AC can occur in centre-embedded position while the less downgraded ACs in the same language need to occur at the periphery of the main clause (see, e.g., Warao and Wambaya<sup>28</sup>). Once again, there is a correlation between the formal and the distributional dimension of embedding, both of which may themselves be determined by the degree of functional embedding of an AR into the main clause.

<sup>27</sup> Note that Cristofaro’s (2003) study does not include manner clauses, and these are, of course, not unlike purpose clauses in many of the above respects (fixed TAM constellation, same-subject constraint) and thus show a predictably high degree of desententialization, as we see in Fig. 1.

<sup>28</sup> We encountered the so-called adjoined subordinate clause in Wambaya earlier in this paper (§1.2.2). Importantly, this clause type contrasts with more desententialized clauses also in that the latter can be found inside the matrix clause rather than strictly preceding or following it (Nordlinger 1998: 217). In Warao (isolate: Guyana), converbal and infinitival ACs can be found in centre-embedded position, while more sentential types of AC occur ad-clausally, notably in preposed position (Romero-Figeroa 1997: 18ff).



The opposite patterns are found with the ARs at the top of Fig. 1: As we saw above, conditional and concessive clauses provide an interpretative frame for the proposition in the matrix clause and are hence more peripheral – both functionally and distributionally – to the matrix clause than temporal, manner or goal specifications of the matrix predicator. Causal clauses, especially in their discourse function of providing post-hoc justification of the matrix proposition, are equally peripheral and, like result clauses, resemble an independent (i.e. coordinated) discourse move (Diessel and Hetterle 2011). And since, additionally, none of these relations predetermines any particular TAM or person constellation, they tend to be full-fledged clauses, or least significantly less desententialized, than the ARs discussed above. Interestingly, Hetterle’s (2015) study also shows that when causal and concessive clauses appear in preposed position (and hence with topical/presupposed rather than focal function), they are significantly more downgraded than in postposed position (see also Hengeveld 1998 on this correlation). For purpose, manner and some temporal relations, by contrast, downgrading is equally prominent in pre- and postposed position. This suggests that causal and concessive clauses, in keeping with their different pre- and post-matrix functions, may also be associated with different diachronic origins in the two positions: an expansion scenario in preverbal position and an integration scenario in postverbal position. This is exactly parallel to the developmental trajectories of pre- and postnominal RRCs from above.

It is clear, then, that there is a systematic triangular relationship between the formal, functional and distributional degree of embedding of ACs across the world’s languages, just like in relative and complement clauses.

### 3.3 Properties of adverbial subordinators

The final correlation we wish to point out is that between the order of the AC and the primary marker of the AR in question. Based on what is known from other types of complex sentence, the typologically expected position of adverbial subordinators would be at the clause boundary that links the AC to the matrix clause, such that postposed ACs are introduced and preposed ACs are terminated by an adverbial marker. This expectation is nicely borne out, for example, in Abun, as illustrated in (53):

(53) Abun (West Papuan: Indonesia; Berry and Berry 1999: 195, 209)

- a. [Ji ma sa] an yo ma nde tó.  
 1SG come **when.RL** 3SG NEG come NEG INCOMPL  
 ‘When I came, he had not yet arrived.’
- b. Ji yo ma nde [we ji bi nggon i].  
 1SG NEG come NEG **because** 1SG POSS woman sick  
 ‘I did not come because my wife was sick.’

It turns out, however, that preposed ACs are actually often found with initial markers. In particular, as shown by Diessel (2001, 2019a), preposed conditional and temporal clauses (but not causal and purpose clauses) are often marked by free initial subordinators. This is in stark contrast again with relative and complement clauses, which only rarely have clause-initial markers in prenominal or preverbal position, respectively (Lehmann 1984, Hawkins 1994, Dryer 2009).

These patterns can arise, of course, when an AR is grammaticalized in postposed position but then commonly shifted to sentence-initial position for discourse-pragmatic reasons. However, a further explanation lies in the diachronic origins of a sizeable number of preposed ACs in postnominal RRCs on “adverbial” head nouns like ‘time’. This is illustrated by (54) from Fongbe:

- (54) Fongbe (Atlantic-Congo, Volta-Congo, Kwa Volta-Congo, Gbe: Benin, Togo; Lefebvre and Brousseau 2002: 171)

[*Hwè̀nù* *dé-è* à *xá* *àtín jí* *é*], *ùn* *mò* *wè*.

**time** OP-REL 2SG climb tree on DEF 1SG see 2SG

‘When you climbed up the tree, I saw you.’

In (44) above, we saw that ‘time’ ranks relatively high on the AH of relativization, thus making it an easily available source for grammaticalization into a temporal AC construction. Crucially, it is this source construction itself that is commonly placed sentence-initially because of its expository function, even in VO languages, and the temporal AC resulting from it simply retains this positional preference. In this scenario, the combination of sentence-initial ACs and clause-initial subordinators is thus a “persistence effect” (Hopper 1991) from a source construction that was subject to similar discourse-pragmatic pressures. In keeping with this hypothesis, Hetterle (2015) finds that the adverbial linkers in her study that had a clearly identifiable origin in nouns (= 10% of all linkers) are distributed unevenly over the different ARs: they are common for WHILE, BEFORE and WHEN relations, but rather uncommon, for example, for PURPOSE and RESULT.<sup>29</sup>

For conditional clauses, too, such persistence effects can be found. On the one hand, many sources have pointed out that conditional markers are often indistinguishable (or develop) from temporal WHEN-markers, thus occupying the same position in the clause. On the other hand, conditional connectors also derive from copulas (‘is (the case) that ...’, ‘it being that ...’; see Hopper and Traugott (2003: 186) on Swahili, Japanese and Chicasaw), modal markers of various kinds (‘perhaps’, ‘suppose’) and markers of polar interrogatives (ibid., Kuteva et al. 2019: 479). Where these sources are clause-initial structures, they can give rise to preposed conditional clauses with initial markers. And as also shown by Hopper and Traugott (2003: 187f.), temporal and conditional clauses are themselves sources for other ARs, thereby bequeathing their (initial) markers to causal (‘since’) and concessive (‘(even) if’) clauses. Overall, then, it appears that the position of adverbial connectives and the resulting AC as a whole falls out directly from their origins. As argued by Diessel (2019a), this motivates the cases in which we find a neat correlation between the position of the AC and the placement of the connector, as in (53) above, as well as (at least some of) the non-correlating cases of preposed ACs with initial markers and postposed ACs with final markers.

A final aspect of subordinators worth pointing out concerns the reduction of their substance depending on contextual factors. We saw above that ACs may contain multiple markers that cue the specific AR in question, and if the remaining markers

<sup>29</sup> Thompson et al. (2007: 244f.) propose that the pathway from relative to adverbial clauses is particularly common for those ARs that have single-word (i.e. adverb) equivalents, notably temporal, locative and manner clauses.

are judged sufficiently informative in this regard, the primary adverbial connector may be dropped (see Hetterle 2015: 108). In the following example from Turkish, the ablative marking on the nominalized verb licenses the optionality of the causal marker at the end of the AC:

(55) Turkish (Turkic: Turkey; Kornfilt 1997: 69)

*Hasan [kitab-ı san-a ver-diğ-im-den (dolayı)] çok kız-dı.*  
 Hasan book-ACC you-DAT give-NMLZ-1SG-ABL (because) very angry-PST  
 ‘Hasan got very angry because I gave the book to you.’

In the recent typological and psycholinguistic literature, such patterns have attracted increasing attention under the label of redundancy management in grammar (e.g. Hawkins 2014, Jaeger and Buz 2018). As such, they join many other grammatical devices whose optional or differential occurrence is said to be motivated by efficiency: whenever a particular grammatical meaning is predictable (from contextual or long-term frequency effects), it tends to receive shorter expression (relative to the less predictable contexts, see Haspelmath 2018 for an overview).

Following this line of argumentation, it has recently been shown that a strong predictor of complementizer and relativizer omission is the relatively likelihood that a complement or relative clause will appear given a certain matrix verb or noun phrase, respectively (see Jaeger 2010, Wasow et al. 2011 for relevant corpus-linguistic work, and Norcliffe and Jaeger 2016 for experimental evidence). In this light, it is certainly no accident that ARs which are more strongly projected by the matrix predicator show the same trend towards shortening and omission. First and foremost, this applies to the motion-cum-purpose clauses we encountered in (52) above; as we saw there, the purposive connector as such is optional in purpose clauses with a matrix verb of directed motion, and the nominalizer used in these constructions is also shorter than the one in general purpose clauses. Schmidtke-Bode (2009) shows that this is a more general typological pattern. In Korean, for example, the purposive connector commonly used with motion verbs *-(u)le* is shorter than the general-purpose connector *-ki wiha-ye* (Hwang 1997), and in Amele and Hausa (to name but two relevant languages), the purposive marker can be omitted particularly with a matrix verb of directed motion.

Moreover, the fact that purpose clauses (in general) have a strong usage bias towards subject co-reference across matrix and subordinate clause leads, in some languages, to exactly the same reduction effects: Where same-subject and different-subject purpose clauses are distinct constructions, the different-subject marker will usually be at least as long as the same-subject marker (e.g. Awa Pit); and a purpose connector may become optional only in same-subject constellations (e.g. Lango). These effects are thus not based on immediate contextual predictability but long-term effects of frequency-sensitive linguistic memories (see also Haspelmath 2013 for the same effects in desiderative complementation).

#### 4. Conclusion: The dynamics of clausal embedding

In this chapter, we have laid out some prominent typological patterns in the structure, function and use of relative and adverbial clauses. Apart from presenting the major

cross-linguistic types of RCs and ACs as such, we have attempted to situate them in the more general context of clausal embedding. Specifically, we have argued that, for both RCs and ACs, there are systematic correlations between *functional*, *formal* and *distributional* dimensions of embedding: prototypically, the more a clause can be conceptualized as being functionally embedded into another unit (i.e. a phrase or a clause), the more it will assume the formal and distributional properties of typical constituents of that unit. It needs to be emphasized, however, that there can also be mismatches between form and function (e.g. syntactically appositive nominalizations used with restrictive meaning), and that our three dimensions of embedding are not to be understood as discrete entities but as continua: as we saw throughout the chapter, clauses can be more or less embedded along all of the three dimensions, and the manifold combinations that can result from this reflect the enormous diversity of clause-combining constructions across the world's languages (see also Lehmann 1988 on this point).

Therefore, the view of embedding espoused in this paper is a synchronically gradient and diachronically dynamic one, in keeping with the more general understanding of grammar as a usage-based dynamic system (Beckner et al. 2009, Diessel 2019b). On this view, clausal embedding is an emergent phenomenon: for both RCs and ACs, we observed that sociative syntactic configurations – such as juxtaposed clauses with anaphoric links, afterthoughts or appositions of NPs and clauses – can come to be integrated into formally more asymmetrical structures over time. On this account, many RRCs follow similar diachronic trajectories as other constituents of NPs, such as determiners, which have also been argued to emerge from appositive demonstrative pronouns (Himmelmann 1997). Alternatively, clausal embedding may come about by successively expanding the scope of phrasal constituents: RRCs can thus be derived from the expansion of other noun modifiers, such as genitival or adjectival slots, and ACs can be expansions of flagged NPs and adverbial PPs. And once grammaticalized, embedded clauses continue to be dynamic entities, being reanalyzed into semantically and discourse-pragmatically related construction types (e.g. relative > temporal clauses, temporal > conditional clauses, purpose > result clauses, etc.). They may also be analogically extended from one domain to the next (e.g. from purposive to complement environments (Haspelmath 1989), from quotative complement to various adverbial relations (Heine et al. 1991, Saxena 1995) or from relative to both adverbial and complement relations (e.g. Givón 1991)). These developments yield dynamic networks of polyfunctionality (see van Lier 2009 and Hetterle 2015 for detailed studies), and as we have argued for ACs in particular, they can often directly account for larger typological regularities (such as word-order correlations between adpositional phrases, dependent clauses and subordinators; see Diessel 2019a).

In surveying different types of RRCs and ACs, we have also seen that both domains of clause linkage are subject to very similar communicative and processing-related pressures. Concerning the former, we saw for both domains that preposed and postposed constructions are often associated with distinct information-structural profiles and hence different types of RRCs and ACs, respectively. As far as processing is concerned, we saw for both domains that the more predictable the dependent clause is from its associated matrix element (i.e. the head nominal or the matrix predicator) or from other aspects of the construction, the less overtly it needs to be marked itself.

This principle extends from the subordinator to other aspects of economical coding, such as the gaps in relativization: the more recoverable the relativized position is from other cues in the RC construction (or the context), the more easily it can be gapped. Communicative pressures motivate which semantic roles are most commonly chosen as relativization sites and these, in turn, are again the ones with the least overt coding.

Overall, then, both RCs and ACs can not only be embedded in a structural sense, but they are always embedded (more metaphorically) into a language-specific ecosystem that is shaped by the cognitive, communicative and social demands of the respective language users.

## Abbreviations

The paper follows the *Leipzig Glossing Rules* (Comrie et al. 2008). Additional abbreviations are as follows:

I, II, III	agreement classes (e.g. Choctaw)	NARR	narrative (tense)
ADD	additive focus (Yakkha)	NATIV	nativizer (verbal affix in Yakkha attached to certain loan words from Nepali)
ADJZ	adjectivalizer		
AM	autonomous motion (Cabecar)		
ANIM	animate	NGR	n-grade (aspectual marker for duration and incompleteness in Choctaw)
ANT	anterior (aspect)		
AOR	aorist (perfective past tense)	NLOCUT	non-locutor (3 <sup>rd</sup> -person index in Awa Pit declarative clauses)
AVN	action verbal noun (Mapudungun)		
C <sub>NUMBER</sub>	noun class	NTR	neutral form (specific inflectional form of Urarina)
COBL	complementizing oblique case (Kayardild)	OP	operator
CONN	connector	OVN	object verbal noun (Mapudungun)
DEP	dependent form	POT	potential
DEST	destination	PREV	previous mention (Choctaw)
DIR	direction	PROP	proprietary
DF	D-form (specific inflectional form of Urarina)	REF	reference marker (Mayogo)
DYN	dynamic	RL	realis
EF	E-form (specific inflectional form of Urarina)	SEQ	sequential
HRS	hearsay (evidential)	SS	same subject
INAN	inanimate	STAT	stative
INESS	inessive (case)	SUB	subordinator
INT	intentional	THEM	thematic suffix (Georgian)
LOCUT	locutor (1 <sup>st</sup> -person index in Awa Pit declarative clauses)	TNS	tense marker
		TR	transitivizer
		VENT	ventive

## References

- Aldai, Gontzal (2003). The prenominal [-case] relativization strategy of Basque: Conventionalization, processing and frame semantics. Ms., Departments of Linguistics, University of Southern California and University of California Los Angeles.
- Aissen, Judith L. (1984). Control and command in Tzotzil purpose clauses. *Berkeley Linguistics Society* 10: 559–571.
- Andrews, Avery D. (2007). Relative clauses. In Timothy Shopen (ed.), *Language Typology and Syntactic Description, Vol. II: Complex Constructions*, 2<sup>nd</sup> ed, 206–236. Cambridge: Cambridge University Press.

- Antinucci, Francesco, Alessandro Duranti and Lucyna Gebert (1979). Relative clause structure, relative clause perception, and the change from SOV to SVO. *Cognition* 7(2): 145–176.
- Ariel, Mira (1999). Cognitive universals and linguistics conventions: The case of resumptive pronouns. *Studies in Language* 23: 217–269.
- Aristar, Anthony Rodrigues (1991). On diachronic sources and synchronic patterns: An investigation into the origin of linguistic universals. *Language* 67(1): 1–33.
- Auer, Peter (2009). On-line syntax: Thoughts on the temporality of spoken language. *Language Sciences* 31: 1–13. DOI:10.1016/j.langsci.2007.10.004
- Axenov, Serge (2006). *The Balochi Language of Turkmenistan: A Corpus-Based Grammatical Description*. PhD dissertation, Acta Universitatis Upsaliensis.
- Basilico, David (1996). Head position and internally headed relative clauses. *Language* 72(3): 498–532.
- Beckner, Clay, Richard Blythe, Joan L. Bybee, Morten H. Christiansen, William Croft, Nick C. Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman & Tom Schoenemann (2009). Language is a complex adaptive system: Position paper. *Language Learning* 59(s1). 1–26.
- Behaghel, Otto (1932). *Deutsche Syntax: Eine geschichtliche Darstellung, Vol. 4: Worstellung, Periodenbau*. Heidelberg: Winter.
- Berghäll, Liisa (2015). *A Grammar of Mauwake*. Berlin: Language Science Press.
- Berry, Keith and Christine Berry (1999). *A Description of Abun: a West Papuan Language of Irian Jaya*. Canberra: Pacific Linguistics.
- Bickel, Balthasar (2003). Referential density in discourse and in syntactic typology. *Language* 79(4): 708–736.
- Bickel, Balthasar (2005). On the typological variables of relativization. Paper presented at the *Workshop on the Typology, Acquisition, and Processing of Relative Clauses*, Max Planck Institute for Evolutionary Anthropology Leipzig. Available online at <[https://www.comparativelinguistics.uzh.ch/dam/jcr:00000000-774a-e877-0000-00005667596c/rc\\_mpi-eva\\_june2004bb.pdf](https://www.comparativelinguistics.uzh.ch/dam/jcr:00000000-774a-e877-0000-00005667596c/rc_mpi-eva_june2004bb.pdf)> (May 2019).
- Bickel, Balthasar (2010). Capturing particulars and universals in clause linkage: A multivariate analysis. In Isabelle Bril (ed.), *Clause Linking and Clause Hierarchy: Syntax and Pragmatics*, 51–102. Amsterdam, Philadelphia: John Benjamins.
- Bickel, Balthasar (2011). Grammatical relations typology. In Jae Jung Song (ed.), *The Oxford Handbook of Linguistic Typology*, 399–444. Oxford: Oxford University Press.
- Brainard, Sherry and Dietlinde Behrens (2002). *A Grammar of Yakan*. Manila: Linguistic Society of the Philippines.
- Broadwell, George Aaron (2006). *A Choctaw Reference Grammar*. Lincoln, London: University of Nebraska Press.
- Bybee, Joan L. (1985). *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam, Philadelphia: John Benjamins.
- Bybee, Joan L., William Pagliuca and Revere D. Perkins (1990). On the asymmetry in the affixation of grammatical material. In William Croft, Keith Denning and Suzanne Kemmer (eds.), *Studies in Typology and Diachrony: Papers Presented to Joseph H. Greenberg on his 75th Birthday*, 1–42. Amsterdam, Philadelphia: John Benjamins.
- Carlson, Robert (1994). *A Grammar of Supyire*. Berlin, New York: Mouton de Gruyter.
- Chafe, Wallace (1984). How people use adverbial clauses. In *Berkeley Linguistics Society* 10: 437–449.
- Cole, Peter (1987). The structure of internally headed relative clauses. *Natural Language & Linguistic Theory* 5(2): 277–302.
- Comrie, Bernard (1981). *Language Universals and Linguistic Typology*. Chicago: The University of Chicago Press.
- Comrie, Bernard (1998). Rethinking the typology of relative clauses. *Language Design* 1: 59–86.

- Comrie, Bernard (2003). The verb-marking relative clause strategy, with special reference to Austronesian languages. *Linguistik Indonesia* 21: 1–18.
- Comrie, Bernard (2006). Syntactic typology: just how exotic ARE European-type relative clauses? In Ricardo Mairal and Juana Gil (eds.), *Linguistic Universals*, 130–154. Cambridge: Cambridge University Press.
- Comrie, Bernard (2008a). Prenominal relative clauses in verb-object languages. *Language and Linguistics* 9(4): 723–733.
- Comrie, Bernard (2008b). Subordination, coordination: Form, semantics, pragmatics – setting the scene. In Edward J. Vajda (ed.), *Subordination and Coordination Strategies in North Asian Languages*, 1–16. Amsterdam, Philadelphia: John Benjamins.
- Comrie, Bernard and Zarina Estrada-Fernández (eds.) (2012). *Relative Clauses in the Languages of the Americas*. Amsterdam, Philadelphia: John Benjamins.
- Comrie, Bernard and Tania Kuteva (2013). Relativization on obliques. In Matthew S. Dryer and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*, Chapter 123. Munich: Max Planck Digital Library. Available online at <<http://wals.info/chapter/123>> (May 2019).
- Comrie, Bernard and Sandra A. Thompson (2007) [1985]. Lexical nominalization. In Timothy Shopen (ed.), *Language Typology and Syntactic Description. Vol. III: Grammatical Categories and the Lexicon*, 334–381. Cambridge: Cambridge University Press.
- Comrie, Bernard, Martin Haspelmath and Balthasar Bickel (2008). The *Leipzig Glossing Rules*: Conventions for interlinear morpheme-by-morpheme glosses. Department of Linguistics of the Max Planck Institute for Evolutionary Anthropology and the Department of Linguistics of the University of Leipzig. Available online at <<https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf>> (July 2016).
- Cristofaro, Sonia (2003). *Subordination*. Oxford: Oxford University Press.
- Cristofaro, Sonia (2014). Is there really a syntactic category of subordination? In Laura Visapää, Jyrki Kalliokoski and Helena Sorva (eds.), *Contexts of Subordination: Cognitive, Typological and Discourse Perspectives*, 73–91. Amsterdam, Philadelphia: John Benjamins.
- Croft, William (2001). *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- Culy, Christopher D. (1990). *The Syntax and Semantics of Internally Headed Relative Clauses*. PhD dissertation, Stanford University.
- Curnow, Timothy J. (1997). A Grammar of Awa Pit (Cuaiquer). PhD dissertation, Australian National University.
- Dayley, Jon P. (1985). *Tzutujil Grammar*. Berkeley: University of California Press.
- de Rijk, Rudolf P.G. (1972) Relative clauses in Basque: A guided tour. *Anuario del Seminario de Filología Vasca 'Julio de Urquijo' VI. (Papers from the Basque Linguistics Seminar, University of Nevada, Summer 1972)*: 55–69.
- de Vries, Mark (2001). Patterns of relative clauses. *Linguistics in the Netherlands* 18(1): 231–243.
- de Vries, Mark (2005). The fall and rise of universals of relativization. *Journal of Universal Language* 6: 1–33.
- DeLancey, Scott (1986). Relativization as nominalization in Tibetan and Newari. Paper presented at the 19<sup>th</sup> International Conference on Sino-Tibetan Languages and Linguistics, Ohio State University.
- Dench, Alan C. (1995). *Martuthunira*. Canberra: Pacific Linguistics.
- Deutscher, Guy (2000). *Syntactic Change in Akkadian: The Evolution of Sentential Complementation*. Oxford: Oxford University Press.
- Deutscher, Guy (2009a). Nominalization and the origin of subordination. Talmy Givón and Masayoshi Shibatani (eds.), *Syntactic Complexity: Diachrony, Acquisition, Neuro-Cognition, Evolution*, 199–214. Amsterdam, Philadelphia: John Benjamins.

- Deutscher, Guy (2009b). The semantics of clause linking in Akkadian. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *The Semantics of Clause Linkage: A Cross-linguistic Typology*, 56–73. Oxford: Oxford University Press.
- Diessel, Holger (1999). *Demonstratives: Form, Function, and Grammaticalization*. Amsterdam, Philadelphia: John Benjamins.
- Diessel, Holger (2001). The ordering distribution of main and adverbial clauses: A typological study. *Language* 77(3): 433–455. DOI:10.1353/lan.2001.0152
- Diessel, Holger (2008). Iconicity of sequence: A corpus-based analysis of the positioning of temporal adverbial clauses in English. *Cognitive linguistics* 19(3): 465–490. DOI:10.1515/COGL.2008.018
- Diessel, Holger (2013). Adverbial subordination. In Silvia Luraghi and Claudia Parodi (eds.), *Bloomsbury Companion to Syntax*, 341–354. London: Continuum.
- Diessel, Holger (2019a). Preposed adverbial clauses: Functional adaptation and diachronic inheritance. In Karsten Schmidtke-Bode, Natalia Levshina, Susanne Maria Michaelis and Ilja A. Seržant (eds.), *Explanation in Typology: Diachronic Sources, Functional Motivations and the Nature of the Evidence*, 97–122. Berlin: Language Science Press.
- Diessel, Holger (2019b). *The Grammar Network: How Linguistic Structure is Shaped by Language Use*. Cambridge: Cambridge University Press.
- Diessel, Holger and Katja Hetterle (2011). Causal clauses: A cross-linguistic investigation of their structure, meaning, and use. In Peter Siemund (ed.), *Linguistic universals and language variation*, 23–54. Berlin, New York: Mouton de Gruyter. DOI:10.1515/9783110238068.23
- Dik, Simon C. (1997). *The Theory of Functional Grammar, Part 2: Complex and Derived Constructions*. Ed. by Kees Hengeveld. Berlin, New York: Mouton de Gruyter.
- Disterheft, Dorothy (1980). *The Syntactic Development of the Infinitive in Indo-European*. Columbus, Ohio: Slavica Press.
- Disterheft, Dorothy and Carlotta Viti (2010). Subordination. In Silvia Luraghi and Vit Bubenik (eds.), *Continuum Companion to Historical Linguistics*, 230–249. London: Continuum.
- Dixon, R. M. W. (2004). *The Jarawara Language of Southern Amazonia*. Oxford: Oxford University Press.
- Dixon, R. M. W. (2009). The semantics of clause linking in typological perspective. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *The Semantics of Clause Linkage: A Cross-linguistic Typology*, 1–55. Oxford: Oxford University Press.
- Dixon, R. M. W. (2010). *Basic Linguistic Theory, Vol. 2: Grammatical Topics*. Cambridge: Cambridge University Press.
- Donohue, Mark (1999). *Warembori*. München: LINCOM Europa.
- Downing, Bruce T. (1978). Some universals of relative clause structure. In Joseph H. Greenberg (ed.), *Universals of Human Language, Vol. 4*, 357–418. Stanford CA: Stanford University Press.
- Dryer, Matthew (1992). The Greenbergian word order correlations. *Language* 68(1): 81–138.
- Dryer, Matthew S. (2009). The branching direction theory of word order correlations revisited. In Sergio Scalise, Elisabetta Magni and Antonietta Bisetto (eds.), *Universals of Language Today*, 185–208. Dordrecht: Springer.
- Dryer, Matthew S. (2013a). Order of adverbial subordinator and clause. In Matthew S. Dryer and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*, Chapter 94. Munich: Max Planck Digital Library. Available online at <<http://wals.info/chapter/94>> (June 2019).
- Dryer, Matthew S. (2013b). Order of adjective and noun. In Matthew S. Dryer and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*, Chapter 87. Munich: Max Planck Digital Library. Available online at <<http://wals.info/chapter/87>> (May 2019).



- Dryer, Matthew S. (2013c). Order of relative clause and noun. In Matthew S. Dryer and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*, Chapter 90. Munich: Max Planck Digital Library. Available online at <<http://wals.info/chapter/90>> (May 2019).
- Dryer, Matthew S. (2013d). Relationship between the order of object and verb and the order of adposition and noun phrase. In Matthew S. Dryer and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*, Chapter 95. Munich: Max Planck Digital Library. Available online at <<http://wals.info/chapter/95>> (May 2019).
- Dryer, Matthew S. (2013e). Relationship between the order of object and verb and the order of relative clause and noun. In Matthew S. Dryer and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*, Chapter 96. Munich: Max Planck Digital Library. Available online at <<http://wals.info/chapter/96>> (May 2019).
- Epps, Patience (2008). *A Grammar of Hup*. Berlin, New York: Mouton de Gruyter.
- Epps, Patience (2009). Between headed and headless relatives: Evidence from Amazonia. Paper presented at the 8<sup>th</sup> Meeting of the Association for Linguistic Typology, Berkeley.
- Evans, Nicholas (1995). *A Grammar of Kayardild*. Berlin, New York: Mouton de Gruyter.
- Faarlund, Jan Terje (2012). *A Grammar of Chiapas Zoque*. Oxford: Oxford University Press.
- Fanego, Teresa (2004). On reanalysis and actualization in syntactic change: The rise and development of English verbal gerunds. *Diachronica* 21(1): 5–55.
- Fortescue, Micheal (1984). *West Greenlandic*. London: Croom Helm.
- Fox, Barbara A. (1987). The noun phrase accessibility hierarchy reinterpreted: Subject primacy or the absolutive hypothesis? *Language* 63(4): 856–870.
- Frajzyngier, Zygmunt (1987). The relative clause in Proto-Chadic. In Herrmann Jungrathmayr and Walter W. Müller (eds.), *Proceedings of the 4<sup>th</sup> International Hamito-Semitic Congress*, 425–450. Amsterdam, Philadelphia: John Benjamins.
- Genetti, Carol (1991). From postposition to subordination in Newari. In Elizabeth Closs Traugott and Bernd Heine (eds.), *Approaches to Grammaticalization. Vol. II*, 227–256. Amsterdam, Philadelphia: John Benjamins.
- Genetti, Carol (2007). *A Grammar of Dolakha Newar*. Berlin, New York: Mouton de Gruyter.
- Genetti, Carol, Ellen Batee, A.R. Coupe, Kristine Hildebrandt and You-Jing Lin (2009). Syntactic aspects of nominalization in five Tibeto-Burman languages of the Himalayan Area. *Linguistics of the Tibeto-Burman Area* 31: 97–143.
- Givón, Talmy (1991). The evolution of dependent clause morpho-syntax in Biblical Hebrew. In Elizabeth Traugott and Bernd Heine (eds.), *Approaches to Grammaticalization, Vol. II*, 257–310. Amsterdam, Philadelphia: John Benjamins.
- Givón, Talmy (2001). *Syntax: An Introduction, Vol. II*. Amsterdam, Philadelphia: John Benjamins.
- Givón, Talmy (2009). *The Genesis of Syntactic Complexity*. Amsterdam, Philadelphia: John Benjamins.
- Givón, Talmy (2011). *Ute Reference Grammar*. Amsterdam, Philadelphia: John Benjamins. DOI:10.1075/clu.3
- Glinert, Lewis (1989). *The Grammar of Modern Hebrew*. Cambridge: Cambridge University Press.
- González Campos, Guillermo and Christian Lehmann (2019). The Cabecar relative clause. Ms., available at <[http://www.christianlehmann.eu/publ/cabecar\\_rel\\_clause.pdf](http://www.christianlehmann.eu/publ/cabecar_rel_clause.pdf)> (accessed 03/09/2019).
- Greenberg, Joseph H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of Language*, 58–90. Cambridge, MA: MIT Press.
- Grossman, Eitan, Guillaume Jacques and Anton Antonov (2018). A cross-linguistic rarity in synchrony and diachrony: Adverbial subordinator prefixes exist. *STUF – Language Typology and Universals* 71(4): 513–538. DOI:10.1515/stuf-2018-0020
- Guirardello, Raquel (1999). A Reference Grammar of Trumai. PhD dissertation, Rice University.

- Haader, Lea (2002). Mikrodiachronie und Sprachwandel in den zusammengesetzten Sätzen. *Acta Linguistica Hungarica* 49: 51–83.
- Haiman, John (1978). Conditionals are topics. *Language* 54(3). 564–589. DOI:10.2307/412787
- Haiman, John and Sandra A. Thompson (1984). “Subordination” in Universal Grammar. *Berkeley Linguistics Society* 10: 510–523.
- Hale, Kenneth L. (1975). Gaps in grammar and culture. In M. Dale Kinkade, Kenneth L. Hale and Oswald Werner (eds.), *Linguistics and Anthropology: In Honor of C.F. Voegelin*. 295–315. Lisse: The Peter de Ridder Press.
- Hale, Kenneth L. (1983). Warlpiri and the grammar of non-configurational languages. *Natural Language & Linguistic Theory* 1: 5–74.
- Harms, Phillip Lee. (1994). *Epena Pedee Syntax*. Summer Institute of Linguistics and University of Texas at Arlington Press.
- Harris, Alice C. and Lyle Campbell (1995). *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.
- Haudry, Jean (1973). Parataxe, hypotaxe et corrélation dans la phrase latine. *Bulletin de la Société de Linguistique de Paris* 68(1): 147–186.
- Haspelmath, Martin (1989). From purposive to infinitive – a universal path of grammaticization. *Folia Linguistica Historica* 10(1-2): 287–310.
- Haspelmath, Martin. (1993). *A Grammar of Lezgian*. Berlin, New York: Mouton de Gruyter.
- Haspelmath, Martin (1995). The converb as a cross-linguistically valid category. In Martin Haspelmath and Ekkehard König (eds.), *Converbs in Cross-Linguistic Perspective*, 1–55. Berlin, New York: Mouton de Gruyter.
- Haspelmath, Martin (2013). On the cross-linguistic distribution of same-subject and different-subject ‘want’ complements: Economic vs. iconic motivation. *SKY Journal of Linguistics* 26: 41–69.
- Haspelmath, Martin (2018). Explaining grammatical coding asymmetries: Form-frequency correspondences and predictability. Ms., Leipzig University. Available online at <<https://ling.auf.net/lingbuzz/004348>> (April 2019).
- Hawkins, John A. (1994). *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
- Hawkins, John A. (2004). *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.
- Hawkins, John A. (2014). *Cross-Linguistic Variation and Efficiency*. Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780199664993.001.0001
- Heath, Jeffrey (1999). *A Grammar of Koyra Chiini*. Berlin, New York: Mouton de Gruyter.
- Heath, Jeffrey (2005). *Tamashek: The Tuareg of Mali*. Berlin, New York: Mouton de Gruyter.
- Heath, Jeffrey (2008). *A Grammar of Jamsay*. Berlin, New York: Mouton de Gruyter.
- Heine, Bernd (2009). From nominal to clausal morphosyntax: Complexity via expansion. In Talmy Givón and Masayoshi Shibatani (eds.), *Syntactic Complexity: Diachrony, Acquisition, Neuro-Cognition, Evolution*, 23–52. Amsterdam, Philadelphia: John Benjamins.
- Heine, Bernd, Ulrike Claudi and Friederike Hünemeyer (1991). *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.
- Hendery, Rachel (2012). *Relative Clauses in Time and Space*. Amsterdam, Philadelphia: John Benjamins.
- Hengeveld, Kees (1998). Adverbial clauses in the languages of Europe. In Johan van der Auwera (ed.), *Adverbial Constructions in the Languages of Europe*, 335–420. Berlin, New York: Mouton de Gruyter.
- Hetterle, Katja (2015). *Adverbial Clauses in Cross-Linguistic Perspective*. Berlin, New York: Mouton de Gruyter. DOI:10.1515/9783110409857
- Himmelman, Nikolaus P. (1997). *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Tübingen: Max Niemeyer.

- Holes, Clive (2004). *Modern Arabic: Structure, Functions, and Varieties*. Washington, D.C.: Georgetown University Press.
- Hopper, Paul J. (1991). On some principles of grammaticization. In Elizabeth C. Traugott and Bernd Heine (eds.), *Approaches to Grammaticalization. Vol. I*, 17–35. Amsterdam, Philadelphia: John Benjamins.
- Hopper, Paul J. and Elizabeth Closs Traugott (2003). *Grammaticalization*. 2<sup>nd</sup> edn. Cambridge: Cambridge University Press.
- Hualde, Jose I. and John Ortiz de Urbina (2003). *A Grammar of Basque*. Berlin, New York: Mouton de Gruyter.
- Huddleston, Rodney and Geoffrey K. Pullum (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Huttar, George L. and Mary L. Huttar (1994). *Ndyuka*. London, New York: Routledge.
- Hwang, Shin Ja J. (1997). Purpose clauses in English and Korean. *LACUS Forum* 23: 495–508.
- Ishizuka, Tomoko (2008). Restrictive and non-restrictive relative clauses in Japanese: Antisymmetric approach. Ms., University of California Los Angeles. Available online at <<https://ling.auf.net/lingbuzz/000808/current.pdf>> (July 2019).
- Jaeger, T. Florian (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology* 61(1): 23–62.
- Jaeger, T. Florian and Esteban Buz (2018). Signal reduction and linguistic encoding. In Eva M. Fernández and Helen Smith Cairns (eds.), *The Handbook of Psycholinguistics*, 38–81. Oxford: Wiley-Blackwell.
- Jaggard, Philip J. (2001). *Hausa*. Amsterdam, Philadelphia: John Benjamins.
- Kazenin, Konstantin I. (1994). Split Syntactic Ergativity: Toward an Implicational Hierarchy. *Sprachtypologie und Universalienforschung* 47: 78–98.
- Keenan, Edward L. (1985). Relative clauses. In Timothy Shopen (ed.), *Language Typology and Syntactic Description, Vol. II: Complex Constructions*, 141–170. Cambridge: Cambridge University Press.
- Keenan, Edward L. and Bernard Comrie (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8: 63–99.
- Kenesei, István, Robert M. Vago and Anna Fenyvesi (1998). *Hungarian*. London, New York: Routledge.
- Kibrik, Andrej A. (1992). Relativization in polysynthetic languages. *International Journal of American Linguistics* 58(2): 135–157.
- Kieviet, Paulus (2017). *A Grammar of Rapa Nui*. Berlin: Language Science Press.
- Kojima, Yasuhiro (2018). Georgian. In Tasaku Tsunoda (ed.), *Levels in Clause Linkage: A Cross-Linguistic Survey*, 403–450. Berlin: De Gruyter Mouton.
- Kornfilt, Jaklin (1997). *Turkish*. London, New York: Routledge.
- Kortmann, Bernd (1997). *Adverbial Subordination*. Berlin, New York: Mouton de Gruyter.
- Kuteva, Tania and Bernard Comrie (2006). The typology of relative clause formation in African languages. In F. K. Erhard Voeltz (ed.), *Studies in African Linguistic Typology*, 209–228. Amsterdam, Philadelphia: John Benjamins.
- Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog and Seongha Ree (2019). *World Lexicon of Grammaticalization*, 2<sup>nd</sup> ed. Cambridge: Cambridge University Press.
- Lakoff, George and Mark Johnson (1980). *Metaphors we Live by*. Chicago: University of Chicago Press.
- Langacker, Ronald W. (1987). *Foundations of Cognitive Grammar, Vol. I: Theoretical Prerequisites*. Stanford: Stanford University Press.
- Lefebvre, Claire and Anne-Marie Brousseau (2002). *A Grammar of Fongbe*. Berlin, New York: Mouton de Gruyter.
- Lehmann, Christian (1983). Rektion und syntaktische Relationen. *Folia Linguistica* 17: 339–378.

- Lehmann, Christian (1984). *Der Relativsatz. Typologie seiner Strukturen. Theorie seiner Funktionen. Kompendium seiner Grammatik*. Tübingen: Gunter Narr.
- Lehmann, Christian (1986). On the typology of relative clauses. *Linguistics* 24: 663–680.
- Lehmann, Christian (1988). Towards a typology of clause linkage. In John Haiman and Sandra A. Thompson (eds.), *Clause Combining in Grammar and Discourse*, 181–226. Amsterdam, Philadelphia: John Benjamins.
- Lehmann, Christian (2008). Information structure and grammaticalization. In Elena Seoane and Maria José Lopez Couso (eds.), *Theoretical and Empirical Issues in Grammaticalization*, 207–229. Amsterdam, Philadelphia: John Benjamins.
- Lehmann, Christian (2014). Review of *Relative Clauses in Languages of the Americas: A Typological Overview*, ed. by Bernard Comrie and Zarina Estrada-Fernández. *Language* 90(2): 534–536. Extended version available at <[https://www.christianlehmann.eu/publ/lehmann\\_rev\\_comrie\\_estrada\\_long\\_version.pdf](https://www.christianlehmann.eu/publ/lehmann_rev_comrie_estrada_long_version.pdf)> (May 2019).
- Li, Charles N. and Sandra A. Thompson (1981). *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.
- Lichtenberk, Frantisek (2008). *A Grammar of Toqabaqita*. Berlin, New York: Mouton de Gruyter.
- Mahootian, Sharzad (1997). *Persian*. London, New York: Routledge.
- Maienborn, Claudia and Martin Schäfer (2011). Adverbs and adverbials. In Klaus von Heusinger, Claudia Maienborn and Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning, Vol. 2*, 1390–1420. Berlin: De Gruyter Mouton.
- Malchukov, Andrej (2004). *Nominalization/Verbalization: Constraining a Typology of Transcategorical Operations*. München: LINCOM Europa.
- Martowicz, Anna (2011). Origin and Functioning of Circumstantial Clause Linkers: A Cross-Linguistic Study. PhD dissertation, University of Edinburgh.
- Maslova, Elena (2003). *A Grammar of Kolyma Yukaghir*. Berlin, New York: Mouton de Gruyter.
- Matsumoto, Yoshiko, Bernard Comrie and Peter Sells (eds.) (2017). *Noun-Modifying Clause Constructions in Languages of Eurasia: Rethinking Theoretical and Geographical Boundaries*. Amsterdam, Philadelphia: John Benjamins.
- McConvell, Patrick (2006). Grammaticalization of demonstratives as subordinate complementizers in Ngumpin-Yapa. *Australian Journal of Linguistics* 26(1): 107–137.
- McGregor, William (1990). *A Functional Grammar of Gooniyandi*. Amsterdam, Philadelphia: John Benjamins.
- Merlan, Francesca (1982). *Mangarayi*. Amsterdam: North-Holland.
- Michael, Lev (2009). The semantics of clause linking in Iquito. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *The Semantics of Clause Linkage: A Cross-linguistic Typology*, 146–166. Oxford: Oxford University Press.
- Mithun, Marianne (1984). How to avoid subordination. *Berkeley Linguistics Society* 10: 493–509.
- Mous, Maarten, and Ongaye Oda (2009). The semantics of clause linking in Konso. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *The Semantics of Clause Linking: A Cross-Linguistic Typology*, 336–355. Oxford: Oxford University Press.
- Nedjalkov, Igor' V. (1997). *Evenki*. London, New York: Routledge.
- Nedjalkov, Igor' V. (1998). Converbs in the languages of Europe. In Johan van der Auwera (ed.), *Adverbial Constructions in the Languages of Europe*, 422–455. Berlin, New York: Mouton de Gruyter.
- Nefedov, Andrey (2012). Relativization in Ket. In Volker Gast and Holger Diessel (eds.), *Clause Linkage in Cross-Linguistic Perspective*, 191–224. Amsterdam, Philadelphia: John Benjamins.
- Neukom, Lukas (2001). *Santali*. München: LINCOM Europa.
- Newman, Paul (2000). *The Hausa Language: An Encyclopedic Reference Grammar*. New Haven, London: Yale University Press.
- Noonan, Michael (1992). *A Grammar of Lango*. Berlin, New York: Mouton de Gruyter.

- Norcliffe, Elizabeth, Alice C. Harris and T. Florian Jaeger (2015). Cross-linguistic psycholinguistics and its critical role in theory development: early beginnings and recent advances. *Language, Cognition and Neuroscience* 30(9): 1009–1032. DOI:10.1080/23273798.2015.1080373
- Norcliffe, Elizabeth and T. Florian Jaeger (2016). Predicting head-marking variability in Yucatec Maya relative clause production. *Language and Cognition* 8: 167–205. DOI:10.1017/langcog.2014.39
- Nordlinger, Rachel (1998). *A Grammar of Wambaya, Northern Territory (Australia)*. Canberra: Australian National University.
- O'Dowd, Elizabeth (1992). The syntactic metaphor of subordination: A typological study. *Lingua* 86: 47–80.
- Olawsky, Knut J. (2006). *A Grammar of Urarina*. Berlin, New York: Mouton de Gruyter.
- Overall, Simon (2009). The semantics of clause linkage in Aguaruna. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *The Semantics of Clause Linkage: A Cross-Linguistic Typology*, 167–192. Oxford: Oxford University Press.
- Ramsay, Violeta (1987). The functional distribution of preposed and postposed “if” and “when” clauses in written narrative. In Russell Tomlin (ed.), *Coherence and Grounding in Discourse*, 383–408. Amsterdam, Philadelphia: John Benjamins.
- Rasekh-Mahand, Mohammad, Mojtaba Alizadeh-Sahraie and Raheleh Izadifar (2016). A corpus-based analysis of relative clause extraposition in Persian. *Ampersand* 3: 21–31.
- Reh, Mechthild (1985). *Die Krongo-Sprache (Nünò Mó-Dì)*. Berlin: Dietrich Reimer Verlag.
- Reinholtz, Charlotte (1999). On the characterization of discontinuous constituents: Evidence from Swampy Cree. *International Journal of American Linguistics* 65(2): 201–227.
- Rice, Keren (1989). *A Grammar of Slave*. Berlin, New York: Mouton de Gruyter.
- Rice, Sally and Kaori Kabata (2007). Crosslinguistic grammaticalization patterns of the allative. *Linguistic Typology* 11(3): 451–514.
- Roberts, John R. (1987). *Amele*. London: Croom Helm.
- Roberts, John R. (1988). Switch-reference in Papuan languages: A syntactic or extrasyntactic device? *Australian Journal of Linguistics* 8: 75–117.
- Romero-Figeroa, Andrés (1997). *A Reference Grammar of Warao*. München: LINCOM Europa.
- Rumsey, Alan (1982). *An Intra-Sentence Grammar of Ungarinjin, North-Western Australia*. Canberra: Australian National University.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Sakel, Jeanette (2004). *A Grammar of Mosestén*. Berlin, New York: Mouton de Gruyter.
- Saltarelli, Mario (1988). *Basque*. London, New York: Croom Helm.
- Sawka, Kenneth S. (2001). Aspects of Mayogo Grammar. M.A. thesis, The University of Texas at Arlington.
- Saxena, Anju (1995). Unidirectional grammaticalization: Diachronic and cross-linguistic evidence. *Sprachtypologie und Universalienforschung* 48(4): 350–372.
- Schackow, Diana (2015). *A Grammar of Yakkha*. Berlin: Language Science Press.
- Schmidtke-Bode, Karsten (2009). *A Typology of Purpose Clauses*. Amsterdam, Philadelphia: John Benjamins.
- Schmidtke-Bode, Karsten (2014). Complement Clauses and Complementation Systems: A Cross-Linguistic Study of Grammatical Organization. PhD dissertation, Friedrich Schiller Universität Jena.
- Schmidtke-Bode, Karsten and Holger Diessel (2017). Cross-linguistic patterns in the structure, function, and position of (object) complement clauses. *Linguistics* 55(1): 1–38. DOI:10.1515/ling-2016-0035
- Shagal, Ksenia (2019). *Participles: A Typological Study*. Berlin, Boston: De Gruyter Mouton.
- Smets, Ineke (2008). *A Grammar of Mapuche*. Berlin, New York: Mouton de Gruyter.

- Sohn, Ho-Min (1994). *Korean*. London, New York: Routledge.
- Soukka, Maria (2000). *A Descriptive Grammar of Noon*. München: LINCOM Europa.
- Strunk, Jan (2014). A statistical model of competing motivations affecting relative clause extraposition in German. In Brian MacWhinney, Andrej Malchukov and Edith A. Moravcsik, 88–106. Oxford: Oxford University Press.
- Sumbatova, Nina R. and Yury A. Lander (2014). *Darginskij Govor Selenija Tanty: Grammatičeskij Očerk, Voprosy Sintaksisa*. [The Dargwa Dialect of Tanti: A Grammatical Sketch, Syntactic Issues]. Moscow: Jazyki Slavjanskix Kul'tur.
- Suttles, Wayne (2004). *Musqueam Reference Grammar*. Vancouver, Toronto: University of British Columbia Press.
- Taylor, Charles V. (1985). *Nkore-Kiga*. London: Croom Helm.
- Terrill, Angela (2003). *A Grammar of Lavukaleve*. Berlin, New York: Mouton de Gruyter.
- Timberlake, Alan (2004). *A Reference Grammar of Russian*. Cambridge: Cambridge University Press.
- Thompson, Sandra A. (1985) Grammar and written discourse: Initial vs. final purpose clauses in English. *Text* 5(1–2): 55–84.
- Thompson, Sandra A., Robert Longacre and Shin Ja J. Hwang (2007). Adverbial clauses. In Timothy Shopen (ed.), *Language Typology and Syntactic Description, Vol. II: Complex Constructions*, 2<sup>nd</sup> ed, 237–300. Cambridge: Cambridge University Press.
- Thompson, Sandra A., Joseph S. Park & Charles N. Li (2006). *A Reference Grammar of Wappo*. Berkeley etc.: University of California Press.
- Trask, Larry L. (1997). *The History of Basque*. London, New York: Routledge.
- Tsunoda, Tasaku (ed.) (2018). *Levels in Clause Linkage: A Cross-Linguistic Survey*. Berlin: De Gruyter Mouton.
- Valentine, J. Randolph (2009). The semantics of clause linking in Ojibwe. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *The Semantics of Clause Linkage: A Cross-linguistic Typology*, 193–217. Oxford: Oxford University Press.
- van Gijn, Rik (2006). *A Grammar of Yurakaré*. PhD dissertation, Radboud Universiteit Nijmegen.
- van Lier, Eva (2009). *Parts of Speech and Dependent Clauses: A Typological Study*. Utrecht: LOT Publications.
- Van Putten, Saskia (2014). Left dislocation and subordination in Avatime (Kwa). In Rik van Gijn, Jeremy Hammond, Dejan Matić, Saskia van Putten and Ana Vilacy Galucio (eds.), *Information Structure and Reference Tracking in Complex Sentences*, 71–98. Oxford: Oxford University Press.
- Van Valin, Robert D. Jr. (1984). A typology of syntactic relations in clause linkage. *Berkeley Linguistics Society* 10: 542–558.
- Van Valin, Robert D. Jr. (2005). *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press.
- Van Valin, Robert D. Jr. and Randy LaPolla (1997). *Syntax: Structure, Meaning, Function*. Cambridge: Cambridge University Press.
- Verstraete, Jean-Christophe (2004). Initial and final position for adverbial clauses in English: The constructional basis of the discursive and syntactic differences. *Linguistics* 42(4): 819–853.
- Verstraete, Jean-Christophe (2007). *Rethinking the Coordinate-Subordinate Dichotomy: Interpersonal Grammar and the Analysis of Adverbial Clauses in English*. Berlin, New York: Mouton de Gruyter.
- Verstraete, Jean-Christophe (2008). The status of purpose, reason and intended endpoint in the typology of complex sentences. *Linguistics* 46: 757–788.
- Wakasa, Motomichi (2008). *A Descriptive Study of the Modern Wolaytta Language*. PhD dissertation, University of Tokyo.

- Wash, Suzanne (2001). Adverbial Clauses in Barbareño Chumash Narrative Discourse. PhD dissertation, University of California at Santa Barbara.
- Wasow, Thomas, T. Florian Jaeger and David M. Orr (2011). Lexical variation in relativizer frequency. In Horst J. Simon and Heike Wiese (eds.), *Expecting the Unexpected: Exceptions in Grammar*, 175–195. Berlin, New York: Mouton de Gruyter. DOI:10.1515/9783110219098.175
- Weber, David John (1983). *Relativization and Nominalized Clauses in Huallaga (Huanuco) Quechua*. Berkeley: University of California Press.
- Weber, David J. (1989). *A Grammar of Huallaga (Huánuco) Quechua*. Berkeley: University of California Press.
- Whitehead, Carl R. (2004). *A Reference Grammar of Menya*. PhD dissertation, University of Manitoba.
- Witzlack-Makarevich, Alena and Balthasar Bickel (eds.) (2019). *Argument Selectors: A New Perspective on Grammatical Relations*. Amsterdam, Philadelphia: John Benjamins.
- Wu, Tong (2011). The syntax of prenominal relative clauses: A typological study. *Linguistic Typology* 15(3): 569–623.
- Yap, Foong Ha, Karen Grunow-Hårsta and Janick Wrona (eds.) (2011). *Nominalization in Asian Languages*. Amsterdam, Philadelphia: John Benjamins.
- Ylikoski, Yussi (2003). Defining non-finites: Action nominals, converbs and infinitives. *SKY Journal of Linguistics* 16: 185–237.