The contrast between states-of-affairs and propositions in clausal complementation

Sørensen, Marie-Louise Lind

Accepted/In press: 01/01/2018

Document Version
Accepted author manuscript

The contrast between states-of-affairs and propositions in clausal complementation

Crosslinguistic perspectives

Marie-Louise Lind Sørensen

PhD thesis

Author: Marie-Louise Lind Sørensen

Supervisor: Kasper Boye

Institution: University of Copenhagen

Department: Department of Nordic Studies and Linguistics

Submitted: September 8, 2018

Summary

This thesis is a crosslinguistic study of the semantic contrast between states-of-affairs and propositions in clausal complementation. The past couple of decades have seen a rise in the interest in several aspects of complementation, but the contrast between states-of-affairs and propositions remains understudied. The purpose of the thesis is to provide a cognitive-functional analysis of grammatical contrasts between clausal complements expressing states-of-affairs and propositions. The following questions are addressed: 1) To what extent does the distinction between states-of-affairs and propositions motivate grammatical contrasts in clausal complementation? 2) What is the status of reported speech in a typology of complementation based on the contrast between states-of-affairs and propositions? and 3) How can the contrast between states-of-affairs and propositions be used as a point of departure for describing the system of complementation in a specific language? To answer these questions, the thesis contains a crosslinguistic study of a sample of 173 languages and a descriptive study of clausal complementation in the Bantu language Ruuli.

Chapter 2 reviews linguistic applications of the contrast between states-of-affairs and propositions as well as theoretical approaches to the contrast. Chapter 3 describes the semantics and morphosyntax of complement constructions and discusses the definition of complementation and the identification of complement constructions. Chapter 4 describes the methodologies used in the thesis, and Chapter 5 provides an overview of the articles and the individual contributions of the authors. Chapter 6 is an article which outlines a typology of reported speech based on layered semantic structure and the distinction between states-of-affairs and propositions. Chapter 7 is an article which presents a crosslinguistic study of utterance-predicate complementation and discusses motivations for grammatical contrasts between complements expressing states-of-affairs and propositions. Chapter 8 is an article which presents a first analysis of clausal complementation in the Bantu language Ruuli. Chapter 9 is a summary and conclusion. The thesis contributes new knowledge about clausal complementation to a growing body of literature demonstrating the significance of the contrast between states-of-affairs and propositions in crosslinguistic and language-specific descriptions.

Resumé

Denne afhandling er et tværsproligt studie af den semantiske kontrast mellem sagforhold og propositioner. I de senere årtier har der været en stigende interesse for flere forskellige aspekter af komplementering, men kontrasten mellem sagforhold og propositioner er endnu underbelyst. Formålet med afhandlingen er at fremsætte en kognitiv-funktionel analyse af grammatiske kontraster mellem komplementsætninger, der gengiver henholdsvis sagforhold og propositioner. De følgende spørgsmål søges besvaret: 1) I hvilket omfang motiverer den semantiske kontrast mellem sagforhold og propositioner grammatiske kontraster i komplementsætninger? 2) Hvilken rolle spiller gengivet tale i en komplementeringstypologi baseret på sagforhold og propositioner? og 3) Hvordan kan kontrasten mellem sagforhold og propositioner bruges i beskrivelsen af komplementering i et specifikt sprog? For at besvare disse spørgsmål anvender afhandlingen et tværsprogligt studie af et sample på 173 sprog og et deskriptivt studie af bantusproget ruuli.

Kapitel 2 gennemgår anvendelser af kontrasten mellem sagforhold og propositioner i lingvistikken samt teoretiske tilgange til kontrasten. Kapitel 3 beskriver komplementsætningskonstruktioners semantik og morfosyntaks og diskuterer definitioner af komplementering og identifikationen af komplementsætningskonstruktioner. Kapitel 4 beskriver de metoder, der er blevet anvendt, og kapitel 5 giver et overblik over artiklernes indhold og forfatternes individuelle bidrag. Kapitel 6 er en artikel som skitserer en typologi over gengivet tale baseret på lagdelt semantisk struktur og distinktionen mellem sagforhold og propositioner. Kapitel 7 er en artikel, som præsenterer et tværsproligt studie af ytringsprædikatkomplementering og diskuterer motivationer for grammatiske kontraster mellem komplementer, der gengiver henholdsvis sagforhold og propositioner. Kapitel 8 er en artikel som præsenterer en første analyse af komplementering i bantusproget Ruuli. Kapitel 9 er en opsummering og konklusion. Afhandlingen bidrager med ny viden om komplementering til en voksende mængde litteratur, som har demonstreret betydningen af kontrasten mellem sagforhold og propositioner i tværsproglig og sprogspecifik beskrivelse.

Acknowledgments

Many people played an important role for me during the past three years at the Department of Nordic Studies and Linguistics at the University of Copenhagen. I would like to thank just a few of them here.

First of all, thanks to my supervisor Kasper Boye for all of his guidance and encouragement, detailed comments on my work and for his never-ending enthusiasm about the topic of my thesis.

Many thanks also to Alena Witzlack-Makarevich for her friendship and mentorship and for welcoming me in Kiel on two occasions. Thanks are also due to Erika, Tobias, Zarina and the rest of the staff and students at Kiel University, who imade me feel at home during the three months I spent there. I would also like to thank the many wonderful and inspiring researchers, that I've had the pleasure to meet at conferences and summer schools all over Europe, in particular fellow PhD students at the time Caroline Gentens and Ivana Vrdoljak who were always keen to discuss the topic of states-of-affairs and propositions. I hope to see more of you all in the future.

Thanks also to all of my wonderful colleagues at the University of Copenhagen, who will be sorely missed. Special thanks to Byurakn, Line, Maria, Katrine, Silvia and Violaine for keeping me happy and motivated everyday – and for going out of their way to help me finish the thesis.

And last but not least thanks to my friends and family, especially Rune, my love, for their unfailing support and encouragement and for taking my mind off the thesis when necessary. Without them, this thesis would most probably not have been completed.

Contents

List of Figures

List of Tables

Abbreviations

The following abbreviations are used in the glossed examples. I have re-glossed the examples following the Leipzig Glossing Rules (Comrie et al. 2008) to improve their readability.

1	1st person
2	2nd person
3	3rd person
A.PART	active participle
ACC	accusative
ACT	actual
AG	agent marker
AOR	aorist
APPL	applicative
APPR	apprehensive
AUG	augment
AUX	auxiliary
AV	actor voice
CAUS	causative
COBL	complementizing oblique
COM	comitative
COMP	complementizer
COND	conditional
CONJ	conjunction
COP	copula
CORE	non-nominative core article
DAT	dative

DECL	declarative
DEP	dependent verb
DIR	directive auxiliary
DUR	durative
EMPH	emphatic
EPIST	epistemic
ERG	ergative
EV	evidential
FEM	feminine
FOC	focus
FUT	future
FV	final vowel
GEN	genitive
HAB	habitual
HORT	hortative
ICP	incompletive aspect
IMP	imperative
IMPERF	imperfect
INF	infinitive
INGR	ingressive
INTERJ	interjection
IRR	irrealis
J	gender agreement marker
JUS	jussive
LOC	locative
MASC	masculine
MLOC	modal locative
N	neuter
NAR	narrative
NEG	negation
NOM	nominative
NMLZ	nominalizer
OBJ	object
OBL	oblique

OBL	objective version
O.C.	objective case
PERS	personal
PFV	perfective
PFX	prefix
PL	plural
POSS	possesive
PP	pragmatic particle
PREV	preverb
PROG	progressive
PROH	prohibitive
PRS	present
PRT	particle
PS	caseless personal marker
PST	past
PRX	proximal demonstrative
QUOT	quotative
R	realis
REFL	reflexive
REL	relativizer
REM	remote
REP	repetitive
RES	resultative
SG	singular
SRP	self-reporting pronoun
SBJ	subject
SBJN	subjunctive
SUB	subordinate
TOP	topic
TR	transitive suffix
WP	witnessed past tense

Chapter 1

Introduction

The present thesis is a crosslinguistic study of the semantic contrast between *states-of-affairs* and *propositions* in clausal complementation. Propositions are traditionally understood as meaning units that have a truth value, while states-of-affairs are understood as meaning units that do not have a truth value. Although the contrast has been employed in studies of complementation (Hengeveld 1989, Svenonius 1994, Dik & Hengeveld 1991, Horie 2000, Cristofaro 2003, 2013, Noonan 2007), it does not yet have a central position in linguistics.

The past couple of decades have seen a rise in the interest in several aspects of clausal complementation such as the presence and absence of complementizers (e.g. Kaltenböck 2009, Boye et al. 2012), complement typology (e.g. Cristofaro 2003, Noonan 2007, Schmidtke-Bode 2014) and contrasts between complement clauses and independent clauses (e.g. Dixon & Aikhenvald 2006, Nordström 2010, Jensen & Christensen 2013). But few take the contrast between states-of-affairs and propositions into account. Neither do descriptive studies of complementation in specific languages systematically treat morphosyntactic contrasts correlated with the contrast between states-of-affairs and propositions (but see recent contributions to Boye & Kehavov 2016 that do, e.g. Holvoet 2016 and Serdobolskaya 2016 on complementizer semantics).

This thesis contributes new knowledge about clausal complementation to a growing body of literature that has demonstrated the significance of

the contrast between states-of-affairs and propositions in crosslinguistic and language-specific descriptions of e.g. epistemicity (Boye 2012), complementizer semantics (Boye & Kehavov 2016), factivity (Gentens 2016) and language death (Kehayov 2017).

1.1 Aims and research questions

The overarching aim of the thesis is to provide a cognitive-functional analysis of systematic grammatical contrasts in complementation in terms of the semantic contrast between states-of-affairs and propositions.

The following questions will be addressed.

1. To what extent does the semantic contrast between states-of-affairs and propositions motivate grammatical contrasts in clausal complementation?

2. What is the status of reported speech in a typology of complementation based on the contrast between states-of-affairs and propositions?

3. How can the contrast between states-of-affairs and propositions be used as a point of departure for describing the system of complementation in a specific language?

The thesis comprises three articles, each of which deals with different aspects of these questions. Articles I and II are based on crosslinguistic studies of utterance-predicate complementation. Article III is a descriptive study of clausal complementation in Ruuli (Bantu, Niger-Congo), which builds on the theoretical issues discussed in articles I and II.

The aim of the first article is to develop a functional theory of reported speech based on the contrast between states-of-affairs and propositions. The aim of the second article is to argue that a difference in conceptual complexity between states-of-affairs and propositions is an explanatory factor for grammatical contrasts in clausal complementation. The aim of the third article is to provide a first description of clausal complementation in Ruuli, while examining how the contrast between states-of-affairs and propositions is expressed.

1.2 Theoretical context

This thesis places itself within cognitive-functional linguistics and functional typology (cf. Horie & Comrie 2000, Croft 2003, Cristofaro 2003 and Fabiszak et al. 2016 for recent overviews of the field). Accordingly, it is based on a view of linguistic structure as shaped by cognitive constraints, language use and language function (in contrast to formal theories that view structure as autonomous, cf. Newmeyer 1992). In this section I will lay out the three main theoretical assumptions that underlie the thesis. They are all common assumptions in cognitive-functional linguistics and in particular of Danish Functional Linguistics (Engberg-Pedersen et al. 1996, 2005, Harder 2005).

The first assumption is that linguistic meaning amounts to *representation* rather than *denotation*. In other words, linguistic meaning is understood as conceptualization – more precisely *prompts* for conceptualization [1] (cf. Chapter 2 on denotational and cognitive approaches to the contrast between states-of-affairs and propositions).

The second assumption is that *structure* is language-specific and that languages can only be compared in terms of shared cognitive-functional *substance* (Harder 1996, cf. also Boye 2012: 7-8). The thesis treats the semantic distinction between states-of-affairs and propositions as a substance phenomenon, which entails that the expression of the distinction in complement constructions can be investigated crosslinguistically.

The third assumption is that linguistic structure is shaped by language use as well as by function. The assumption that usage events impact language structure is supported by the observation that usage frequency can influence certain structures as argued by e.g. Bybee & Hopper (2001) and Haspelmath (2008, 2017). Linguistic structure is also functionally motivated. There is a relationship between semantic content and structural expression, such that the structural expression may iconically reflect the semantic content. In some cases, the reason for this reflection is that structure is iconically motivated by semantic content (cf. e.g. Haiman 1985, Langacker 1987, Givón 2001, Croft 2003, Achard 2010).

[1] I will focus only the representational/conceptual nature of linguistic meaning, but see Harder (2007, 2009) on instructional semantics and V. Evans (2009) for a view of linguistic meaning as prompts for conceptualization.

The notion of iconicity has been central to cognitive-functional accounts of grammatical structure. According to Horie & Comrie (2000: 5), the main issues in cognitive-functional complementation research are to identify "Cognitive/Functional motivations for form-meaning (function) alignment in complementation" and "Cognitive/Functional factors governing the speaker's choice of different complement constructions".

To investigate these issues, the thesis mainly employs synchronic data on complement constructions, but diachronic developments in specific languages are no doubt important to fully understand complementation (see e.g. Cristofaro 2014a for a discussion of diachronic evidence in studies of complementation). Diachronic factors impacting the synchronic structure of complement constructions will be given only passing attention, but some hypotheses regarding the grammaticalization of complement-taking utterance predicates is presented in Article I.

1.3 Structure of the thesis

Chapters 2 and 3 lay out the theoretical background of the thesis. Chapter 2 is concerned with the contrast between states-of-affairs and propositions, while Chapter 3 reviews the literature on complementation in cognitive-functional linguistics. Chapter 4 describes the data collection procedure and discusses methodological issues in greater detail than could be done in the articles. Chapter 5 provides an overview of the articles and their common themes and describes the contribution of all authors to the articles. Chapters 6–8 consist of three articles. Article I, *Layered semantic structure in independent utterances and direct and indirect reported speech*, presents a functional theory of reported speech based on layered semantic structure (including illocutions, propositions and states-of-affairs), Article II, *Grammatical contrasts in utterance-predicate complementation: from iconicity to frequency, and back*, argues that iconicity of complexity is a motivating factor for grammatical contrasts in utterance-predicate complementation. Propositions are argued to be conceptually more complex than states-of-affairs and this complexity contrast is argued to be iconically reflected by grammatical contrasts in utterance-predicate complementation. Article III, *Clausal complementa-*

tion in Ruuli (Bantu; JE103) is a descriptive study of Ruuli (Niger-Congo) complementation. The study covers a large range of complement constructions in Ruuli and analyzes the connection between morphosyntactic contrasts and the semantic contrast between state-of-affairs and propositions. The articles are followed by a conclusion in Chapter 9.

Chapter 2

States-of-affairs and propositions

In this chapter, I first discuss some terminological issues regarding the contrast between states-of-affairs and propositions (Section 2.1). Then I review the range of grammatical phenomena to which the distinction between states-of-affairs and propositions has been applied (Section 2.2). Then I discuss how the contrast has been approached from the theoretical viewpoint of denotational semantics (Section 2.3) and cognitive linguistics (Section 2.4). In Section 2.5, I outline the cognitive-functional account of states-of-affairs and propositions, which constitutes the main theoretical foundation for the thesis (and which is extensively discussed in Article II). Finally, I discuss the contrasts between *de dicto* and *de re* and *realis* and *irrealis* and their relationship to the distinction between states-of-affairs and propositions (Section 2.6).

2.1 Terminological issues

The terms states-of-affairs and propositions will be used throughout the thesis, but it should be noted from the outset that a wide range of other terms has been used in the literature to cover the same notions.

States-of-affairs are also known as "actions" (Lees 1960) "events" (Vendler 1967, Schüle 2000, Horie 2000, Achard 1998, 2010), "second-order entities" (Lyons 1977) and "activities" (Dixon 2006).

The term "proposition" is widely used in both philosophy and linguistics, but a few other terms have been used as well, including "facts" (Lees 1960, Vendler 1967, Dixon 2006), "third-order entities" (Lyons 1977) and "propositional contents" (Dik & Hengeveld 1991, Dik 1997).

To make matters more confusing, the term "proposition" is often used in a more general sense, basically equivalent to 'sentence meaning' (cf. Saeed 2009)[2].

Additionally, it should be noted that not all authors make a distinction between only two meaning units (states-of-affairs vs. propositions). Zucchi (1993) and Polakof (2017), for example, distinguish between events, states-of-affairs and propositions and Peterson (1997) distinguishes between facts, propositions and states-of-affairs. According to Peterson (1997: 335), facts are true propositions, which entails that they are a subcategory of propositions. It may also be possible to divide states-of-affairs into subcategories (see Section 2.6). Identifying semantic subcategories of states-of-affairs and propositions, respectively, is theoretically interesting and worth pursuing in future research, but I will not attempt to identify such subcategories here.

2.2 Linguistic applications

In linguistics the distinction between states-of-affairs and propositions has been used in semantic analyses of a number of different grammatical phenomena in the world's languages. I will exemplify the most common applications here, ending with how the contrast has been applied to clausal complementation.

It has been used to describe different types of nominalizations (e.g. Lees 1960: 59-73, Vendler 1967: 122-146, Zucchi 1993 and Schüle 2000) as in (1a) and (1b). Nominalizations like the one in example (1a), *her singing of the aria* have been argued to express a state-of-affairs, while nominalizations like (1b), *her singing the area*, have been argued to express a proposition (e.g. Lees 1960: 59-73; Vendler 1967: 122-146; Schüle 2000: 49-89).

[2]According to Saeed (2009: 14) some logicians describe the common meaning of the assertion *Joan made the sorbet*, the question *Joan made the sorbet?* and the command/request *Joan make the sorbet!* as a proposition. In this thesis, only the assertion and the question are analyzed as propositional (cf. Section 2.2)

(1) a. *Her singing of the aria* (State-of-affairs nominalization)

b. *Her singing the aria* (Propositional nominalization)

States-of-affairs and propositions have also been related to NPs. In (2), for example, "claim" can be described as representing a proposition, as evidenced by the fact that it has truth value (cf. "the claim is true"). By contrast, "event" can be described as representing a state-of-affairs as it can "take place". NPs would otherwise normally be analyzed as referring to "things" in Cognitive Grammar or "first order entities" in the terms of Lyons (1977).

(2) a. *The claim that...*(Propositional NP)

b. *The event that...*(State-of-affairs NP)

Manner adverbs and non-epistemic modality have also been associated with states-of-affairs (Hengeveld 1989), while evidentiality and epistemic modality have been associated with propositions (Boye 2010, 2012), cf. also Lyons (1977: 842-843), Palmer (1979: 35) and Perkins (1983: 7-8). In (3), for example, the modal verb *may* can be interpreted in two ways. It can either be interpreted as expressing the possibility of a state-of affairs or it can be interpreted as modifying a proposition.

(3) *He may stay in that house*

a. 'It is possible for him to stay in that house.'
(State-of-affairs reading of infinitival clause)

b. 'It may be the case that he is staying in that house.'
(Propositional reading of infinitival clause)

The distinction between states-of-affairs and propositions is most often used in the analysis of sentences. It has been related to different sentence types, such as declaratives, interrogatives and imperatives (Boye 2012). According to Boye (2012: 199-206), declarative and interrogative sentences express propositions, while imperatives express state-of-affairs as illustrated in Table 2.1 (cf. also Hengeveld 1990: 7 and Boye 2012: 194-195).

Boye (2012), for example, argues that the reason why directives cannot be used pragmatically as assertions or polar questions, is that they express

Table 2.1: Sentence types and the distinction between states-of-affairs and propositions

Major illocution type	Sentence type	Meaning unit
Assertions	Declaratives	Propositions
Polar questions	Polar interrogatives	Propositions
Directives	Imperatives	States-of-affairs

states-of-affairs and states-of-affairs do not include propositions. In contrast it is possible to use declaratives and polar interrogatives pragmatically as orders/request, because propositions include states-of-affairs.

In this thesis, I am mainly concerned with the expression of the distinction between states-of-affairs and propositions in clausal complementation (however, Article I also analyzes independent main clauses). In research on clausal complementation, the semantic contrast between states-of-affairs and propositions has been used to describe morphosyntactic contrasts between complements, e.g. contrasts between the absence and presence of a complementizer or contrasts between balanced and deranked complements in the sense of Stassen (1985) and Cristofaro (2003). And the complements of different types of complement-taking predicates have been described as expressing either states-of-affairs and propositions (cf. also Chapter 3).

I will describe a few such cases here, namely complements of perception predicates, knowledge predicates, and utterance predicates. Articles I and II focus on utterance-predicate complementation, but in Article III complement constructions with perception predicates, utterance predicates and others are also taken into account.

Perception-predicate complementation has gained the most attention in studies that include the contrast between states-of-affairs and propositions. For example, Dik & Hengeveld (1991), Horie (1993), Schüle (2000) and Boye (2010) all analyze the contrast between 'direct perception' and 'indirect perception' as in (4). In (4a) *see* takes a finite complement and is used in the sense of 'indirect perception' (also called 'acquisition of knowledge'), while in example (4b), *see* takes a non-finite complement and is used in the sense of 'direct perception'. The complement in (4a) expresses a proposition, whereas the complement in (4b) expresses a state-of-affairs (Dik & Hengeveld 1991, Hengeveld & Mackenzie 2008, Boye 2010).

(4) a. *She heard that he played the piano.* (Propositional complement)

b. *She heard him play the piano.* (State-of-affairs complement)

The contrast between direct perception and indirect perception has also received quite a lot of attention in generative linguistics, cf. e.g. Kirsner & Thompson (1976) on "direct perception" vs. "indirect deduction". And reference grammars with only short descriptions of complementation tend to focus on perception-predicate complementation.

Another contrast, which has received less attention in linguistics, is the contrast between 'epistemic knowledge' and 'action knowledge' (or 'know how').[3] This contrast has also been linked to the contrast between states-of-affairs and propositions (cf. Sørensen 2013, Sørensen & Boye 2015 for a crosslinguistic study and Rentzsch & Mitkovska 2017 for a study of Turkish and Macedonian). In (5a) *know* has the sense of knowledge of information, which can be called 'epistemic knowledge', and the finite complement expresses a proposition, while in (5b), *know* has the sense of knowledge of how to do something, which can be called 'action knowledge', and occurs with a non-finite complement expressing a state-of-affairs.

(5) a. *She knows that he plays the piano.*

b. *She knows how to play the piano.*

The main focus of the thesis is on utterance-predicate complementation. Previous studies of utterance-predicate complementation have focused on contrasts between direct and indirect speech, considering morphosyntactic factors, such as tense-marking and logophoricity, or semantic factors, such as view point markers (Coulmas 1986, N. Evans 2012, Spronck 2012). Not much attention has been given to contrasts between states-of-affairs and propositions in utterance-predicate complementation until now (cf. articles I and II). Examples (6a) and (6b) illustrate a propositional complement of an utterance predicate *tell* and a state-of-affairs complement of the same verb. In (6a), *tell* introduces a finite complement and is used to report on an asser-

[3] In contrast to the literature in linguistics, this contrast is of central interest to philosophy (of mind), but a review of the philosophical literature on the topic is out of the scope of this thesis (cf. e.g. Stanley 2011 and Abbott 2013 for recent discussions).

tion expressing a proposition, whereas in (6b), *tell* reports an order/request and occurs with a non-finite complement expressing a state-of-affairs.

(6) a. *She told him that they played the piano.*

b. *She told him to play the piano.*

Articles I and II are based on a crosslinguistic study of the expression of states-of-affairs and propositions in utterance-predicate complementation. In Article III on Ruuli complementation, complement clause constructions with the most frequent complement-taking predicates are described and parallel contrasts in knowledge-predicate complementation, utterance-predicate complementation, perception-predicate complementation and complementation with propositional attitude predicates are highlighted.

2.3 The traditional account

The interest in the difference between states-of-affairs and propositions in linguistics originates in the philosophical literature (mainly overtaken from Vendler 1967 and Bolinger 1968), where the distinction is a recurrent topic in philosophical metaphysics and philosophy of mind (cf. Textor 2016 for an overview of the philosophical conception of states-of-affairs and McGrath & Devin 2018 for an overview of the philosphical conception of propositions).

Although the distinction between states-of-affairs and propositions is not widely used in linguistics, it has a place in some specific frameworks such as Functional (Discourse) Grammar (Hengeveld 1989, Dik 1997, Hengeveld & Mackenzie 2008) and Role and Reference Grammar (Foley & Van Valin 1984, Van Valin & LaPolla 1997) as well as in functional typology (Cristofaro 2003, 2013).

The traditional account of the contrast between states-of-affairs and propositions is essentially based on denotational semantics. States-of-affairs are regarded as entities without truth value, while propositions are entities with truth value. According to the philosopher Loux (1998: 132), for example, propositions are "abstract entities (...) the primary bearers of truth values" and according to Vendler (1967: 174),

> "facts (...) are not in space and time at all. They are not located, cannot move, split, or spread and they do not occur, take place, or last in any sense. Nor can they be vast or fast." (Vendler 1967: 144)

In Functional Grammar (Dik 1997, Hengeveld 1989) and in Functional Discourse Grammar (Hengeveld & Mackenzie 2008), the understanding of the contrast between states-of-affairs and propositions is based on Vendler's (1967) definition. Propositions are defined by Dik (1997) as follows.

> "Propositions are things that people can be said to believe, know or think about; they can be reason for surprise or doubt; they can be mentioned, denied, rejected, and remembered; and they can be said to be true or false." (Dik 1997: 91)

States-of-affairs, on the other hand, are not understood as bearers of truth value. Instead, Loux (1998) defines states-of-affairs as "situations that have essentially the property of obtaining or failing to obtain" and according to Vendler (1967: 144), "[e]vents and their kin are primarily temporal entities". Lyons (1977: 443) defines states–of-affairs as follows: "By second order entities we shall mean events, processes, states-of-affairs, etc., which are located in time and which, in English, are said to occur, take place, rather than to exist.". In Functional Grammar, states-of-affairs are defined as follows.

> "A SoA [State-of-affairs] is something that can be said to occur, take place, or obtain in some world; it can be located in time and space; it can be said to take a certain time (have a certain duration); and it can be seen, heard or otherwise perceived." (Dik 1997: 51)

As described in Boye (2010: 392) the distinction between states-of-affairs and propositions in these approaches is ontological in nature (cf. also Harder 1996: 236) and states-of-affairs and propositions are understood as denotable entities. According to Boye (2010: 401) a central problem with the ontological account of states-of-affairs and propositions is that if fails to account for the grammatical asymmetries found crosslinguistically between propositions and states-of-affairs.

2.4 The cognitive linguistics account

Within cognitive linguistics, states-of-affairs and propositions are understood as different construals of the same conceptual content, namely a *process*. In Cognitive Grammar a process is defined conceptually as a "a complex relationship that develops through conceived time and is scanned sequentially along this axis" (Langacker 2008: 112). Propositions are understood as grounded processes, whereas states-of-affairs are ungrounded processes (e.g. Langacker 1991: 439-440; 551; Langacker 2009: 293, cf. Boye 2010, 2012). According to Langacker (1991: 444) "Grounding locates the event with respect to the speaker's conception of reality." and it "constitutes the final step in the formation of a nominal or a finite clause." (Langacker 1991: 549). The cognitive analysis of the distinction between states-of-affairs and propositions is completely independent of the notion of extra-linguistic entities invoked by the traditional denotational accounts (cf. Boye 2012: 393, Langacker 1991: 439-440; 551; Langacker 2009: 293).

The distinction between states-of-affairs and propositions has been employed by Achard (1998) in his cognitive analysis of mood-marking in French. He argues that the subjunctive in French expresses a state-of-affairs (which he calls "event"), while the indicative expresses a proposition, and that the distinction between states-of-affairs and propositions can be understood as conceptualizations of "basic reality" and "elaborated reality", respectively (cf. also Achard 2002: 207-209).

> "the distinction between events and propositions corresponds to the distinction between basic and elaborated reality. More specifically, basic reality corresponds to events and elaborated reality corresponds to propositions." (Achard 1998: 245)

A contrast between states-of-affairs and propositions might be expressed differently in different languages, e.g. by the presence of a complementizer in a propositional complement and the absence of a complementizer in a state-of-affairs complement, or might not be expressed at all (Boye 2010: 404-407). In one language it might be possible for an emotive predicate to take a state-of-affairs complement, and in another language it might not be.

Achard (1998: 264) posits this to *construal flexibility*. Construal flexibility allows speakers of some languages to construe complements as either states-of-affairs or propositions.

2.5 A cognitive-functional account

As discussed above, states-of-affairs and propositions have been conceived of differently in denotational and cognitive semantics. In this section I outline the cognitive-functional account of states-of-affairs and propositions, which constitutes the main theoretical foundation for this thesis. I also relate states-of-affairs and propositions to so-called *layered structure* (also known as "layered clause structure" in e.g. Dik & Hengeveld 1991 and Dik 1997 or "layered semantic structure" in e.g. Harder 2010) in anticipation of Article I, in which a typology of utterance-predicate complementation based on layered structure is proposed.

Boye (2012) argues that states-of-affairs and propositions are both processes in the sense of Langacker. Both are prototypically sentence meanings and are sequentially scanned. Boye also argues that only propositions are construed as referring, while states-of-affairs are construed as non-referring. According to Boye, "A proposition is a linguistic prompt to evoke a process construed as referring." (Boye 2012: 281), while "a state-of-affairs is a linguistic prompt to evoke a process construed as non-referring" (Boye 2012: 281).[4] Figure 2.1 is a visualization of the contrast between a propositional representation and a state-of-affairs representation as conceived of in Boye (2012: 281).

Both representations are processes in the sense that they are sequentially scanned as shown by the arrow under the changing faces in the figure. The arrow pointing to the question mark in the propositional representation on the right indicates that this process refers to the world. In the state-of-affairs representation on the left, no reference is made directly to the world. In denotational conceptions of reference, it is necessary to make distinction between different types of referents, i.e. real and fictional referents. The

[4]The definitions of states-of-affairs and propositions include the definition of them as "linguistic prompts" for conceptualization, reflecting an instructional semantics (Harder 2007, 2009), cf. also footnote 1.

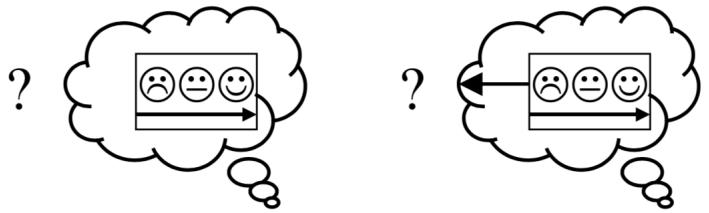

Figure 2.1: The distinction between propositional and state-of-affairs representation. States-of-affairs are processes construed as non-referring and propositions are processes construed as referring (Boye 2012: 281)

cognitive-functional analysis makes the problem of distinguishing between different types of referents obsolete. In this context, the fact that reference may be made to the world, importantly does not imply that there actually is a referent in the world as illustrated by the question mark in Figure 2.1.

> "whether or not something exists of which a given concept would be a conceptual representation – i.e. whether a not a referent of the concept exists – we may construe and entertain the concept at hand as referring (...) or we may construe and entertain it as non-referring." (Boye 2012: 280)

Having truth value in Boye's approach thus means being a referring process, which in turn reflects a "cognitive capacity for hooking concepts onto the world by referring to it" (Boye 2012: 279). In line with Cognitive Grammar and cognitive linguistics in general, the theory is based on an understanding of semantic structure as representational rather than denotational.

2.5.1 Layered structure

So far, I have discussed the distinction between states-of-affairs and propositions independently from other meaning structures. However, within some theoretical frameworks, especially functional frameworks, states-of-affairs and propositions are part of a theory of layered structure. This is most clearly the case in e.g. Danish Functional Grammar (Harder 2005) and in Functional (Discourse) Grammar (Hengeveld 1989, 1990, Dik & Hengeveld

1991, Hengeveld & Mackenzie 2008). But theories of layered structure are also found in e.g. Systemic Functional Grammar and Role and Reference Grammar, cf. Verstraete (2008: 780-784).

I will not go into details about the theoretical frameworks of Functional Grammar and Functional Discourse Grammar here (cf. Kehayov 2017: 301-307 for a concise overview), but I would like to highlight two main ideas in (Dutch) Functional Grammar that differ from the approach used in this thesis. First of all, in Functional Grammar states-of-affairs and propositions are conceived of as denotational meaning units following Vendler (1967). Second, in Functional Grammar the clause is regarded as the center point for layered structure and thus named layered *clause* structure. In contrast, Danish Functional linguistics deals with layered *semantic* structure, following Harder (1996: 228-243) who has argued that it is the semantic structure (or content structure) that is layered and not clause structure, cf. also Boye (2012: 194) and Kehayov (2017: 300).

In the conception of layered structure employed in this thesis, the layers relevant to clausal meaning are the state-of-affairs, the proposition and the layer(s) of the illocution. The illocution may be split into two: illocutionary force (or *neustic* in the terms of Hare 1970) and illocutionary value (or *tropic* in the terms of Hare 1970), as argued in Article I. Propositions always include a state-of-affairs, but states-of-affairs never include a proposition. Propositions are thus conceptually dependent on states-of-affairs (Boye 2012). In analyses of complementation that presuppose such an understanding of the relationship between states-of-affairs and propositions, it is possible to discuss the relationship between two states-of-affairs, regardless of the fact that one of them might be part of a proposition. In her typology of complementation Cristofaro (2003), for example, defines complementation in terms of the relationship between two states-of-affairs. Consequently, she makes reference to the nature of the relationship between two states-of-affairs even when one or both are actually subsumed under a proposition. For rhetorical purposes, I will use the term state-of-affairs only in cases where I do not consider the state-of-affairs to be part of a proposition.

The following table represents layered structure as it is presented in Article I and as conceived of in this thesis.

Table 2.2: Layered semantic structure

Meaning Unit	Cognitive-functional definition
Illocutionary neustic	illocutionary force
Illocutionary tropic	illocutionary value
Proposition	"a linguistic prompt to evoke a process construed as referring." (Boye 2012: 281)
State-of-affairs	"a linguistic prompt to evoke a process construed as non-referring." (Boye 2012: 281)

In summary, this thesis is based on the following key assumptions about linguistic meaning, in extension of work by e.g. Harder (1996), V. Evans (2009) and Langacker (2008).

1. Linguistic meaning is not a matter of denotation, but of representation.

2. Linguistic meaning is instructions or prompts for action, including instructions or prompts for conceptualization.

The finer details of the theoretical approach employed in the thesis is most relevant to Articles I and II, which are more theoretically oriented than Article III. Article III focuses on how the contrast between states-of-affairs and propositions is expressed in a specific language.

2.6 Related semantic notions

In this section I will discuss two notional contrast that are related to the contrast between states-of-affairs and propositions, namely the contrast between *de dicto* ("the domain of speech") and *de re* ("the domain of reality") as employed by Frajzyngier (1991) and Frajzyngier & Jasperson (1991) and the contrast between *realis* and *irrealis* (cf. e.g. Givón 1994, Mithun 1995, Chafe 1995).

2.6.1 De dicto and de re

The distinction between *de dicto* and *de re* originates in philosophy and has been extensively discussed as a metaphysical, syntactic and semantic

phenomenon (see Gallois 1998 and Mckay & Nelson 2014 for an overview), but these discussions are outside the scope of this thesis.

Within linguistics, Frajzyngier (1991) and Frajzyngier & Jasperson (1991) have described the distinction between de dicto and de re as a "fundamental distinction in semantic structure", presumably encoded in many languages (Frajzyngier & Jasperson 1991: 135-36). The term *de dicto* refers to "the domain of language" while *de re* refers to "the domain of reality".

Examples provided by Frajzyngier & Jasperson (1991), illustrating the contrast between de dicto and de re complements, are given in (7).

(7) Frajzyngier & Jasperson (1991: 139)

a. *He said that he likes apples.* (De dicto complement)

b. *He wants to eat apples.* (De re complement)

Notice here that in Section 2.2 above, parallel examples with the English utterance predicate *say* and the desiderative predicate *want* were argued to be instances of complements expressing propositions and states-of-affairs, respectively.

Frajzyngier & Jasperson (1991) use the distinction between de dicto and de re to analyze the distribution of English complements marked by the complementizer *that*, *to*-infinitives and gerunds (*ing*-form) as shown in Figure 2.2. Complements introduced by the complementizer *that* are analyzed as belonging to the de dicto domain, whereas *to*-infinitives and gerunds (*-ing* forms) are analyzed as belonging to the de re domain.

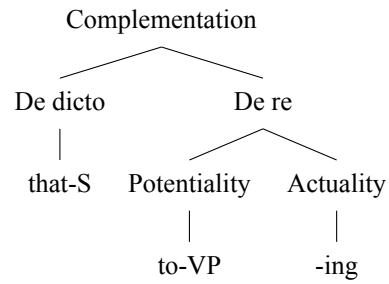

Figure 2.2: Frajzyngier and Jasperson's (1991) model of de dicto and de re in English complementation

On the basis of the examples in (7) and Frajzyngier and Jasperson's (1991) analysis of English complement types as seen in Figure 2.2, one might thus suspect that the de dicto-de re distinction is really nothing but a different terminology for the distinction between states-of-affairs and propositions, especially given the terminological diffusion in the literature (and e.g. Forker 2016: 753 appears to take this position; she lists de re and de dicto as other terms for states-of-affairs and propositions). However, Frajzyngier & Jasperson (1991) mention that they regard propositions as belonging to the domain de dicto, when they are "presented as an object of discourse", typically when they are expressed by complement clauses (in contrast to independent clauses). States-of-affairs (which they also call 'events' following Vendler 1967), on the other hand, are argued to belong to the domain de re (Frajzyngier & Jasperson 1991). Frajzyngier & Jasperson (1991) thus argue for semantic domains that encompass states-of-affairs and propositions rather than a semantic distinction that corresponds exactly to the distinction between states-of-affairs and propositions.

In later analyses of complementation, the term de dicto has been applied to complementizers that express some kind of doubt or distance towards a proposition (Frajzyngier 1995: 496, Suzuki 2000: 1613, Fagard et al. 2016: 94). Fagard et al. (2016) provide the example in (8) from Japanese as an example of a modal complementizer contrast, where the use of the complementizer *to* indicates speaker uncertainty.

(8) Japanese (Japanese; Kuno 1973: 216, quoted in Palmer 1986: 148 and Fagard et al. 2016: 93)

a. *John wa Mary ga sin-da to sinzi-na-katta*
John TOP Mary SBJ die-PST COMP believe-NEG-PST
'John did not believe that Mary was dead.'
(She might or might not have been.)

b. *John wa Mary ga sin-da koto o sinzi-na-katta*
John TOP Mary SBJ die-PST COMP OBJ believe-NEG-PST
'John did not believe that Mary was dead. (She was.)'

In Suzuki's (2000) analysis of Japanese complementizers, she argues that this shows that the Japanese complementizer *to*, as in example (8a), is a de

dicto complementizer, while *koto*, as in example (8b), which doesn't indicate any kind of doubt or psychological distance, is a de re complementizer. On the basis of the analysis of Japanese complementizers in Suzuki (2000), Achard (2010) has argued that both propositions and states-of-affairs can be described as de re. Analyses that have employed the distinction between de re and de dicto have thus not understood it as corresponding to the contrast between propositions and states-of-affairs.

The distinctions between de dicto and de re does not appear to be at odds with an analysis of complement contrasts in terms of the distinction between states-of-affairs and propositions as such, but it is unclear to what extent it would interact with it. In any case, the distinction between the domains of de dicto and de re is ultimately a denotational one and thus suffers from some of the same problems as denotational accounts of the distinction between states-of-affairs and propositions as discussed in Section 2.3.

2.6.2 Realis and Irrealis

Another semantic distinction which is related to the distinction between states-of-affairs and propositions (and the de dicto–de re distinction), is the one between *realis* and *irrealis* (cf. e.g. Givón 1994, Mithun 1995, Chafe 1995, Elliot 2000, Palmer 2001, Gijn & Gipper 2009, Haan 2012, Pietrandrea 2012), cf. also Mauri & Sansò (2012) for a recent overview of central debates regarding the realis-irrealis distinction.

According to Mithun (1995: 386), irrealis is a crosslinguistically comparable semantic category and Dixon (1995: 183) further claims that the contrast between realis and irrealis is a language universal for languages with complement clauses:

> "I suggest the following universal: for languages which have complement-clause constructions there are at least two possibilities: a 'potential (irrealis)' type, typically referring to something that has not happened but which people want or intend should happen (...); and an 'actual (realis)' type, typically referring to some existing or certain event or state." (Dixon 1995: 183)

However, exactly what the terms realis and irrealis cover is rather un-

clear (Trask 1993, Gijn & Gipper 2009) and it has been argued that it is neither a coherent nor a useful term (Palmer 1986, Bybee 1998). This is partly because of theoretical disagreements and partly because of differing descriptive traditions between researchers of different language areas. In studies of e.g. Austronesian and Amerindian languages, the terms realis and irrealis tend to be used for contrasts that in studies of European languages would be described as contrasts between indicative and subjunctive (Elliot 2000: 56, Palmer 2001: 148).

There is also no consensus regarding the relationship between the realis-irrealis distinction and the distinction between states-of-affairs and propositions. Some have related the distinction between realis and irrealis to different types of states-of-affairs (Hengeveld & Mackenzie 2008: 154, Holvoet 2016: 237,259[5], Kehayov & Boye 2016a: 813). For example Holvoet (2016), analyzes the two Latvian complementizers *lai* as in (9a) and *kaut* as in (9b) as states-of-affairs complementizers.

(9) Latvian (Indo-European; Holvoet 2016: 237)

a. *es* *gribu* *lai* *tu* *klausi*
1SG.NOM want.PRS.1SG COMP 2SG.NOM listen.PRS.2SG
man-iem *padom-iem*
my-DAT.PL.M advice-DAT.PL
'I want you to listen to my advice.'

b. *es* *vēlētos* *kaut* *tu* *klausītu* *man-iem*
1SG.NOM wish.IRR COMP 2SG.NOM listen.IRR my-DAT.PL.M
padom-iem
advice-DAT.PL
'I wish you would listen to my advice.'

Others have applied the distinction to propositions, distinguishing realis propositions from irrealis propositions (Elliot 2000, Nordström 2010, Serdobolskaya 2016). For example, Serdobolskaya (2016: 315) states that the

[5]Holvoet (2016) notes that the realis-irrealis distinction may also be found with propositional complements with reference to Russian. Whether the realis-irrealis distinction belongs to the domain of states-of-affairs or propositions is thus indirectly argued to be language-specific.

complementizer *səma* 'as if' in Ossetic "is used to encode irrealis propositions" contrasting with the complementizer *kɜj*, which is used to encode realis propositions.

(10) Ossetic (Indo-European; Serdobolskaya 2016: 315)

a. *asʲa aftɜ ɜnqɜlt:-a səma žawər ɜgaš u*
Asja so think-PST.3SG COMP Zaur alive be.PRS.3SG
'Asja thought Zaur was alive.' (But we know that he's not.)

b. *asʲa aftɜ ɜnqɜlt:-a žawər ɜgaš kɜj u*
Asja so think-PST.3SG Zaur alive COMP be.PRS.3SG
'Asja thought Zaur was alive.'
(We don't know if this is the case or not.)

As one example, Foley (1986: 158) assigns realis and irrealis to opposite ends of an epistemic scale ranging from real to possible to irreal (similar scales are proposed by Givón 1982 and Akatsuka 1985) and on the basis of a small crosslinguistic survey, Gijn & Gipper (2009) have proposed an implicational hierarchy covering the irrealis category from factual to possible to counterfactual as in Figure 2.3.

Counterfactual < Possible [- speaker commitment] < possible [+ speaker commitment] < Factual[- temp specific] < Factual [+ temp specific]

Figure 2.3: Implicational hierarchy of Irrealis (Gijn & Gipper 2009: 176)

These kinds of scales and hierarchies appear to take the realis-irrealis distinction to be an epistemic one, (and thus related to propositions, cf. Boye 2012: 33-34), but as Boye (2012) has pointed out it has nevertheless been used to describe arguably non-epistemic categories as well. In languages where there is evidence for a contrast between realis and irrealis it may be analyzed as overlapping with the contrast between states-of-affairs and propositions. It is possible that it might 1) be used for contrasts between states-of-affairs and propositions, cf. Forker (2016), 2) be used for two semantic subcategories of states-of-affairs cf. (9) or 3) be used for two semantic subcategories of propositions, cf. (10). While contrasts between realis and irrealis may find a use in language-specific descriptions, its validity as a crosslinguistic category is questionable (Bybee 1998, Exter 2012, Boye 2012).

Chapter 3

Complementation

In the previous chapter, I reviewed theoretical accounts of states-of-affairs and propositions in philosophy and linguistics as well as common linguistic applications of the contrast. I also outlined the cognitive-functional approach to complement clause semantics presupposed in this thesis and discussed the distinctions between realis and irrealis and de dicto and de re.

In this chapter, I first briefly review common definitions of clausal complementation and define central terminology that will be used in the thesis. In Section 3.2 I discuss semantic classifications of complement-taking predicates and in Section 3.3, I discuss morphosyntactic aspects of clausal complements, specifically complementizers and the distinction between balancing and deranking. Then, I discuss the problem of identifying complement constructions and their components, including complement-taking predicate clauses, complement-taking predicates and complementizers (Section 3.4). Finally, I discuss utterance-predicate complementation and motivate why utterance-predicate complementation is an especially interesting domain for studying the distinction between states-of-affairs and propositions (Section 3.5).

3.1 Defining clausal complementation

The definition of complement clauses naturally influences the range of constructions that are taking into account by different studies. However, the

definition of complement clauses is a contested issue (see e.g. Höglund et al. 2015 for a recent overview of various definitions). In typologies such as those of Givón (1980, 1990, 2001), Cristofaro (2003) and Noonan (2007), for instance, e.g. infinitive constructions are included. However, some authors like Hansen et al. (2016: 175) exclude "all complements lacking tense and mood markers like for example infinitival phrases" from their study following Dixon (2006). Some take direct speech into account (Cristofaro 2003, Noonan 2007), whereas others exclude it (Dixon 2006).

In this thesis a wide range of grammatical phenomena such as nominalizations and infinitives are all included under the heading of clausal complementation. Definitions of complementation are discussed in-depth in Article III. The aim of this chapter is merely to provide an overview of the morphosyntax and semantics of clausal complementation and discuss some issues with regards to identifying complement constructions. To do so, I use the following central terminology.

1. I reserve the term 'complementation' for all semantic complementation strategies, whether the complement can be said to be embedded, to be juxtaposed or to bear any other syntactic relation to the complement-taking predicate (clause).
2. I reserve the term 'complement construction' for a construction consisting of a complement-taking predicate (clause) and a complement.
3. I reserve the term 'complement' for clausal complements .
4. I reserve the term 'complement-taking predicate' for *verbs* that take a complement (thus excluding adjectives or NPs that may take a clausal complement from the study).
5. I reserve the term 'complement verb' for the verb in a complement.
6. I reserve the term 'complementizer' for free-standing morphemes that serve the function of marking a clause as a complement.

Sections 3.2 and 3.3 are concerned with complement-taking predicates and the morphosyntax of complements, respectively. In some sense, this is a rather artificial division as the meaning of a complement construction arguably derives from an interaction between the meanings of the complement-taking predicate, the complement clause and (optionally) complementizers.

Although useful generalizations can be made about complement-taking predicates, complement clauses and complementizers, respectively, in isolation, it will become apparent that the descriptions cannot be made completely separate.

3.2 Complement-taking predicates

Complement constructions are often analyzed in relation to complement-taking predicates. Some commonly identified complement-taking predicates are: modal predicates, e.g. 'can', phasal predicates, e.g. 'begin', manipulative predicates, e.g. 'force', knowledge predicates, e.g. 'know', perception predicates, e.g. 'see', propositional attitude predicates, e.g. 'think' and utterance predicates, e.g. 'say'. The focus of two of the articles in this thesis is utterance-predicate complementation, as certain features of utterance-predicate complementation makes it fertile ground for exploring the relevance of the distinction between states-of-affairs and propositions (cf. Section 3.5). The third article describes the entire system of complementation in a specific language and thus deals with a number of different predicate types.

Many attempts at classifying complement-taking predicates have been made. There are basically two approaches: One is to look at a specific language and describe the complement-taking predicate classes according to the distributional characteristics of complements in that language. Another is to try and classify predicates crosslinguistically. This typically results in classifications that deal with general human capabilities of interacting with the world such as perception, knowledge and thought, (see for example the semantic classifications in Givón 2001, Cristofaro 2003, Dixon 2006, 2010, Noonan 2007).

For example, Noonan (2007) defines a number of complement-taking predicates based partly on the semantics of the complement-taking predicate itself and partly on the type of complement it occurs with. His classification looks as follows:

(11) Noonan's classification of complement-taking predicates (Noonan 2007: 124-144)

- "Utterance predicates are used in sentences describing a simple transfer of information initiated by an agentive subject."
- "Propositional attitude predicates express an attitude regarding the truth of the proposition expressed as their complement."
- "Pretence predicates have as a characteristic that the world described by the proposition embodied in the complement is not the real world."
- "Commentative predicates resemble propositional attitude predicates in that, when an overt human subject appears, the subject is an experiencer since the predicate gives information about mental attitudes. They differ from propositional attitude predicates in that they provide a comment on the complement proposition."
- "Predicates of knowledge and acquisition of knowledge (...) take experiencer subjects and describe the state, or the manner of acquisition, of knowledge. Knowledge and acquisition of knowledge (KAK) predicates include *know*, *discover*, *realize*, *find out* and *forget*, as well as perception predicates such as SEE and HEAR when used in a sense other than that of immediate perception."
- "Predicates of fearing (...) are characterized semantically by having experiencer subjects and expressing an attitude of fear or concern that the complement proposition will be or has been realized"
- "Desiderative predicates, such as *want*, *wish*, *desire* and *hope* are characterized by having experiencer subjects expressing a desire that the complement proposition be realized."
- "Manipulative predicates express a relation between an agent or a situation which functions as a cause, an affectee, and a resulting situation. The affectee must be a participant in the resulting situation."
- "Modal predicates (...)[are] predicates expressing moral obligation and moral necessity, and (...) predicates of ability which resemble them closely in syntactic properties."
- "Achievement predicates can be divided into positive and negative achievement classes. Positive achievement predicates, such as *manage*, *chance*, *dare*, *remember to*, *happen to* and *get to*, refer to the

manner or realization of achievement. Negative achievement predicates, such as *try*, *forget to*, *fail* and *avoid*, refer to the manner of, or reason for, the lack of achievement in the complement predication."

– "Phasal predicates refer to the phase of an act or state: its inception, continuation, or termination"

– "Immediate perception predicates include forms such as see, hear, watch and feel where the predicate names the sensory mode by which the subject directly perceives the event coded in the complement.

– "Negative predicates (...) A CTP which takes the negated proposition as its complement."

This type of classification is used in crosslinguistic studies, e.g. also Cristofaro (1998), Givón (2001), Dixon (2006). In language-specific descriptions, such as reference grammars, classifications are based on the characteristics of the given language.

Dixon (2010) suggests considering the classes of complement-taking predicates in (12), when describing complementation in specific languages.

(12) Dixon's classification (Dixon 2010: 416-417)

– *Attention* (a) 'see', 'hear', 'notice', 'smell', 'show'; (b) 'recognize', 'discover', 'find'

– *Thinking* (a) 'think', 'consider', 'imagine', 'dream'; (b) 'assume', 'suppose'; (c) 'remember', 'forget; (d) 'know', 'understand'; (e) 'believe', 'suspect'

– *Deciding* 'decide (to)', 'resolve', 'plan', 'choose'

– *Liking* (a) 'like', 'love', 'prefer', 'regret'; (b) 'fear'; (c) 'enjoy'

– *Speaking* (if the the language has indirect speech) (a) 'say', 'inform', 'tell' (one sense); (b) 'report; (c) 'describe', 'refer to'; (d) 'promise', 'threaten'; (e) 'order', 'command', 'persuade', 'tell' (another sense)

This list is rather different than Noonan's (2007) classification and lumps together several of his categories. We also see that contrasts between states-of-affairs and propositions are implicitly highlighted for some predicate types

but not for others. For the 'speaking' category, for example, Dixon makes a distinction between 'tell' (one sense) and 'tell' (another sense), which with minimal context is a completely opaque description, but which most probably is supposed to describe that some utterance predicates like English *tell* can be used in constructions such as "*tell* (somebody) *that*..." (propositional complement) and "*tell*(somebody) *to* (do something)" (state-of-affairs complement) (cf. Chapter 2). However, there is no indication that there might be different "senses" of verbs such as 'know'. As argued in Section 2.2, complement constructions with knowledge predicates may express a contrast between states-of-affairs and propositions, such as the contrast in English between 'knowing that' (epistemic knowledge) and 'know how' (action knowledge).

Examples of such contrasts will be considered in Article III as we describe complement constructions in Ruuli. An example of the contrast between epistemic and action knowledge is given in (13).

(13) Ruuli (Niger-Congo; cf. Article III)

a. *o-maite ye-ena a-yendy-a oku-yizukiry-a*
2SG.S-know 3SG-FOC 3SG.S-need-FV INF-be.reminded-FV
'You know he also needs to be reminded.'

b. *a-maite oku-sany-a*
3SG.S-know.PFV INF-swim-FV
'He knew how to swim.'

In Article III this contrast is treated in the section on knowledge predicates since (13a) and (13b) both include the complement-taking predicate *maite* 'know'. We use the classification of predicates given in (14), which has the advantage of including the most frequent-complement-taking predicate as well as allowing an analysis of the distinction between states-of-affairs and propositions within each class in Ruuli.

(14) Classification of complement-taking predicates in Article III

– *Phasal predicates* describing the beginning, continuation or end of an event, such as 'begin', 'continue' and 'stop'

- *Modal predicates* describing possibility, necessity or ability, such as 'may' and 'can'
- *Manipulative predicates* describing the physical coercion of a participant into carrying out an action, such as 'make' and 'force'
- *Desiderative predicates* are predicates with meanings like 'want' and 'wish'. Predicates with these meanings that occur with same-subject as well as predicates that occur with different subject complements are included in this class.
- *Perception predicates* describing a physical sensation like 'see', 'hear' and 'feel'. Occur with complements signifying direct perception (also called immediate perception) as well as complements that signify information acquired via perception (Boye 2010, Dik & Hengeveld 1991).
- *Knowledge predicates* describing the state, acquisition or transfer of knowledge, such as 'know', 'learn', 'teach', 'forget', 'remember'. Predicates describing knowledge of information (epistemic knowledge) or 'know how' (action knowledge) are included in this class.
- *Propositional attitude predicates* such as 'think'
- *Utterance predicates* describing an illocutionary act and signifying information transfer, information requests or orders/directions delivered by means of speech, such as 'say','ask' and 'tell'
- Emotive predicates describing feelings, such as 'hate' and 'love'

3.3 Morphosyntactic structure of complements

There are two prominent points of interest for complement constructions, which will be the focus of this thesis: the morphological form of the complement verb and the marking of a complement clause by a complementizer. Contrasts between complementizers and contrast between complement verb forms have been related to semantic contrasts between states-of-affairs and propositions. This thesis explores these contrasts further. Section 3.3.1 discusses the semantics of complementizers and Section 3.3.2 discusses the contrast between balancing and deranking and its relation to finiteness.

3.3.1 Complementizers

Complements are often introduced by so-called *complementizers*, which are words, particles, clitics or affixes, one of whose functions is to mark clauses as complements (Noonan 2007: 55, cf. also Crystal 2008: 93). In early treatments of complementizers, complementizers were conceived of as purely syntactic phenomena without any semantic content, but later developments started appreciating that complementizers may have a number of semantic functions. According to Frajzyngier & Jasperson (1991: 134), Bolinger (1968: 122) was "one of the first to attribute semantic properties to the complementizers used". Recent studies (authors in Boye & Kehavov 2016) have also demonstrated the rich semantics of complementizers.

One of the functions complementizers may have is to distinguish propositions from states-of-affairs. For example, the propositional complement might be marked by a complementizer, whereas the state-of-affairs complement is not (complementizer absence and presence), as in (15), or both complements might be marked by a complementizer – a different one in the propositional complement and the state-of-affairs complement as in Fijian (16). In both cases the propositional complement is used to report an assertion and the state-of-affairs complement is used to report a command/request. In contrast to (15a) and (16a), (15b) and (16b) arguably do not have a truth value.

In Gulf Arabic, the propositional complement of *gaal* 'say' in (15a) is marked by a complementizer *ʔinnah*, whereas there is no complementizer in the state-of-affairs complement in (15b).

(15) Gulf Arabic (Afro-Asiatic; Holes 1990: 1;2)

a. *ir-rajjaal gaal ʔinnah Saayim*
the-man said.3SG COMP A.PART-fast
'The man said that he was fasting.'

b. *il-mudiir gaal lii ashiil il-awraag*
the-boss said.3MSG to.me 1SG.remove the-papers
'The boss told me to remove the papers.'

Example (16) from Fijian illustrates the contrast between propositions and states-of-affairs expressed as a contrast between two different complementizers *ni* and *me*.

(16) Fijian (Austronesian; Dixon 1988: 301;271)

a. *e tu'u-na vei Mere o Jone ni na la'o'*
3SG tell-TR to person art person COMP FUT go
'John told Mary that he would go/that she should go.'

b. *e tu'u-na vei Meri o Jone me na la'o'*
3SG tell-TR to person ART person COMP FUT go
'John told Mary that (she) should go.'

While complementizers can be used to mark a complement as a proposition or a state-of-affairs, they might also have other functions at the same time. As was discussed in Section 2.6, some languages have different propositional complementizers that express different degrees of certainty towards the proposition expressed by the complement. We have also seen that some states-of-affairs complements may be marked by different complementizers, possibly marking contrasts such as 'realis' and 'irrealis' (cf. discussion in Section 2.6.2).

For the sake of clarity, it should be noted that in many languages the contrast between states-of-affairs and propositions is not expressed the same way across different predicate types. In Tukang Besi for example the contrast between propositions and states of a affairs is a complementizer presence/absence contrast with perception predicates (as argued by Boye 2010, 2012) as in (17). The complementizer *kua* marks the propositional complement, while the state-of-affairs complement has no complementizer. With utterance predicates on the other hand, the contrast is expressed morphologically as a contrast between two different complementizers: *kua* versus *ka'ano* as in (18).

(17) Tukang Besi (Austronesian; Donohue 1999: 404;403)

a. *no-'ita-'e kua no-kanalako te osimpu*
3R-see-3OBJ COMP 3R-steal CORE young.coconut
'She saw that he had stolen the coconut.'

b. *no-'ita-'e* *no-kanalako* *te* *osimpu*
3R-see-3OBJ 3R-steal CORE young.coconut
'She saw him stealing the coconut.'

(18) Tukang Besi (Austronesian; Donohue 1999: 394;393)

a. *no-wuju-'e* *kua* *saba'ane* *no-lemba-'e*
3R-persuade-3OBJ COMP all 3R-carry-3OBJ
'They(i) persuaded them(j) that they(i/*j) would carry it all.'

b. *no-wuju-'e* *ka'ano* *saba'ane* *no-lemba-'e*
3R-persuade-3OBJ COMP all 3R-carry-3OBJ
'They persuaded him to carry it all.'

3.3.2 Balancing vs. deranking

The contrast between balanced and deranked complement verb forms has been proposed by Stassen (1985) and later used and developed by Cristofaro (2003) as an alternative to the contrast between finite and non-finite verb forms.

According to Stassen (1985: 76-83), languages have essentially two strategies for coding the semantic relationship between two states-of-affairs, e.g. in a complement construction. Either both the verb in the complement-taking predicate clause and the complement verb will be morphosyntactically equivalent or the verb in the complement will will differ morphosyntactically from the complement-taking predicate and for instance lack tense or person marking. The first strategy, where the verb form corresponds to that of an independent clause, is named balancing. The second strategy, where the complement verb form could not occur in an independent clause, is named deranking.

In extension of Stassen's definition of balancing and deranking, Cristofaro has argued that balanced verbs should rather be defined as verbs that are (or can be) marked by the same grammatical categories as a verb in a declarative independent clause (Cristofaro 2003: 54-55). Therefore a number of "subordinate" moods, such as subjunctives, hortatives and jussives should be regarded as deranked as well, even though for instance a subjunctive verb

form might have the same morphological structure as a verb in an independent clause. "the impossibility of a verb form occurring in independent declarative clauses taken in isolations is regarded as a sufficient condition for deranking regardless of how this is indicated" Cristofaro (2003: 57). This distinction is similar to the one between dependent and independent verb forms proposed by Haspelmath (1995).

Example (19) from Persian and (20) from Nenets illustrate complements that are considered to be deranked by Cristofaro. In Persian the complement verb is subjunctive and in Nenets the complement verb is jussive.

(19) Persian (Indo-European; Mahootian 2005: 94;32)
be to goft (ke) be-r-i
to you said (COMP) SBJN-go-2SG
'She told you to go.'

(20) Nenets (Uralic; Nikolaeva 2014: 285)
Wera-n^{o}h ma-dəm-c^{o} pida xər^{o}-ta xər^{o}-m-ta xanaə-ya
Wera-DAT say-1SG-PST he REFL-3SG knife-ACC-3SG take-JUS
'I told Wera to take his own knife.'

Whether or not moods such as subjunctives are considered deranked clearly has an impact on studies of the distribution of complementation.

A major difference between Stassen's (1985) and Cristofaro's (2003, 2013) definitions of balancing and deranking has to do with the categorization of imperative clauses. In Stassen's terms, an imperative verb form is considered to be balanced as it can occur in an independent clause (Stassen 1985: 76-83). But in Cristofaro's terms, imperative clauses must be regarded as deranked in languages where the imperative verb does not have the same morphosyntactic structure as a declarative independent clause (Cristofaro 2003: 54-55). This issue will be discussed in Article II.

3.4 Identifying complement constructions

Since defining clausal complementation is itself challenging (cf. Section 3.1), it is no surprise that the identification of language-specific instances

of clausal complementation is challenging too. In this section, I outline the main issues in identifying complement constructions in specific languages. I first discuss the issue of identifying complement-taking predicate clauses as opposed to parentheticals. Then, I turn to the distinction between complement clauses and purpose clauses, and finally, I discuss the challenges of identifying complementizers and of distinguishing complement-taking predicates from reportative evidentials.

The first challenge has to do with the relationship between a complement-taking predicate clause and a complement clause. Within usage-based linguistics, this relationship has been characterized as one of conceptual subordination, where the complement clause is an argument of a complement-taking predicate (cf. Section 3.1). However, this characterization and even the existence of complement clauses as a grammatical phenomenon has been challenged by Thompson (2002) (cf. also Thompson & Mulac 1991). Thompson argues that what is normally analyzed as a complement clause is rather an independent clause, and that what is normally considered a complement-taking predicate clause is really an epistemic or evaluative fragment. For instance, Thompson analyses complement constructions such as *wanted to say* [*I'm really happy with the stuff*] not as cases of complementation, but as combinations of an independent clause (marked in square brackets) and an evaluative fragment (*wanted to say*) (Thompson 2002: 151). However, Thompson's analysis has been criticized by e.g. Boye & Harder (2007) for conflating usage and structure.

An additional problem for Thompson's analysis which is particularly relevant in the context of this thesis, is that it only covers propositional complements. Thompson specifically discusses finite English complements, and these are arguably always propositional – they can be epistemically modified (cf. Chapter 2). Since propositions allow for epistemic modification, it is no doubt a possibility that what could be analyzed as a complement clause in English is really the main clause with an epistemic modification. However, if the aim is to describe clausal complementation as a unified phenomenon, it is simply not enough only to take propositional complements into account.

While theoretically important, the issue of distinguishing parentheticals from complement-taking predicates was not a major problem in practice for

the present study. However, one case, where we were in doubt about the analysis, was in the descriptive study of Ruuli (cf. Article III). In the Ruuli data, we almost exclusively found examples of complements preceded by a complement-taking predicate. However, we did find an example where the complement-taking predicate followed the complement clause as in (21).

(21) Ruuli (Niger-Congo; cf. Article III)
ti-bi-kya-tu-kol-a *n-dowoz-a*
NEG-8S-PERS-1PL.O-work-FV 1SG.S-think-FV
'They no longer work for us, I think.'

The non-canonical position of the complement-taking predicate clause has been argued to be an indicator of grammatical function in English and related languages (Boye & Harder 2007: 580). It may be the case, then, that example (21) is best analyzed as an independent clause followed by an epistemic fragment.

The second challenge is that complement clauses may resemble several other different types of clauses such as adverbial clauses and relative clauses (cf. e.g. Kehayov & Boye 2016b). Here I will focus on purpose clauses. Purpose clauses are generally not regarded as complement clauses, but they bear a close semantic and morphosyntactic resemblance to them. Verstraete (2008: 780), for example, has noted that the English infinitive clause to go away can function as either a purpose clause, as in (22a), or a complement clause, as in (22b). According to Verstraete the interpretation of *to go away* is entirely dependent on the semantics of the complement-taking predicate *save* vs. *tell*/*force*/*urge*.

(22) Verstraete (2008: 780)

a. *I saved money to go away.*

b. *I told/forced/urged him to go away.*

In Verstraete's conception of the difference between complement clauses and purpose clauses, there is actually not much doubt when it comes to constructions with utterance predicates, which are of main concern in this thesis.

However, in some reference grammars, potential complementizers were inconsistently glossed, and this raised doubt about the appropriate analysis. Seiler (1977), for example, glosses one potential complementizer in the Uto-Aztecan language Cahuilla as 'in.order.to' as in (23).

(23) Cahuilla (Uto-Aztecan; Seiler 1977: 347)
ʔí-vʔi *ʔékʷʔašmal* *hí-yeʔ* *pe-núʔin-qal* *p-iš*
this boy his-mother it-tell-DUR in.order.to
pe-yékaw-pi *páwiʔči-i*
it(she)-pick-for.to wild.apple-O.C.
'This boy's mother told him to pick wild apples.'

This translation could imply that *piš* is a purpose marker. However, because the construction is introduced by an utterance predicate *núʔin* 'tell' and because *piš* also appeared to be used to mark more clear cases of complements (and was accordingly sometimes glossed as a subordinate conjunction), I analyzed it as a complementizer and not a purpose marker in this case.

The problem of distinguishing between complementizers and purposive markers is more acute for bound forms that could potentially be complementizers. To avoid the problem of distinguishing between purpose markers and complementizers in these cases, I restricted the analysis to free-standing forms (cf. also Section 3.3.1 and Article II).

The third challenge has to do with the identification of complementizers. Identifying which elements can accurately be described as complementizers in different languages is not a trivial task, since complementizers may be hard to distinguish from, for example, other markers of subordinate clauses, such as relativizers (Kehayov & Boye 2016b: 83, see also the discussion in Section 8.4).

This problem is related to the question of how much a complementizer may resemble other grammatical elements in the language. Some authors have argued that the origin of a complementizer should in principle be opaque for it to be analyzed as such. Haiman (2011: 304), for example notes that the status of *tha:* as in (24a-b) as a complementizer in Khmer is controversial, since it can synchronically be understood as 'say'.

(24) Khmer (Austro-Asiatic; Haiman 2011: 227;304)

a. *kom tae lo:k prap tha: via cia pseup pul kom*
PROH only you tell say 3 be mushroom poison PROH
ej knjom mwn deung
what I not know
'If you hadn't told it was a poisonous mushroom, I wouldn't have known.'

b. *kheu:ŋj tha: rahah cia sawh sba:j me:n*
see COMP quick heal healthy well really
'See that it had really healed quickly.'

However, *tha:* is used as a complementizer with a number of complement-taking predicates, such as *kheu:ŋj* 'see' (27b), which indicates that it has taken on the general function of marking complements and not just reported speech. I therefore regard *tha:* as a complementizer (somewhat analogous to the uncontroversial analysis of *that* as a complementizer in English, even though its origin as a demonstrative pronoun is synchronically transparent).

Consider also (25a-b) from Thai. I would analyze this as an example of a contrast between utterance-predicate complements expressing a proposition (25a) and a state-of-affairs (25b). The semantic contrast is marked by a contrast between a propositional complementizer *wâa* and a state-of-affairs complementizer *hây*.

(25) Thai (Tai-Kadai; Iwasaki & Preeya 2005: 105;334)

a. *kháw bɔ̀k wâa măa nīa ná man cà mii khwam-lúsù̬k*
3 tell say/COMP dog PP PP 3 CM have PFX-feel
kɔ̀ɔn châymá
before QP
'They say that dogs sense danger first, right?'

b. *kháw bɔ̀ɔk hây thoo pay thăam*
3 tell give/CAUS phone go/DIR ask
'He told me to call and ask.'

However, in the literature on Thai, there is no consensus on this analysis. Iwasaki & Preeya (2005: 334), for example analyze *hây* as a serial verb, but they acknowledge that other studies may analyze it as a complementizer (an analysis I subscribe to).

The final challenge to be discussed here has to do with the distinction between utterance predicates and reportative evidentials. Reportative evidentials can be used express how propositional information has been acquired by the speaker. They are thus functionally equivalent to complement-taking predicate clauses, in so far as a complement-taking predicate clause such as *they say* expresses an equivalent meaning.

Consider, for instance, the expression *manu* from Ainu, which is glossed 'hearsay' by Tamura & Kamei (2000) as shown in (26). Despite the translation into an English complement construction, the Ainu example is of course not a complement construction, but rather an independent clause with a reportative evidential *manu*.

(26) Ainu (Ainu; Tamura & Kamei 2000: 121)
tan nay etokota pirika too an manu
this river upper.reaches pretty lake to.be hearsay
'They say there's a pretty lake up this river.'

In some languages reportative evidentials are derived from utterance predicates. An utterance predicate like *say* in *they say* (*that*...), which can be used in a complement construction, may grammaticalize into a reportative evidential marker, which no longer has the function of marking the clause as a complement (cf. Boye & Harder 2009). Example (27) from Paiwan is illustrative of such a case. In its first occurrence, *aya* is a complement-taking predicate 'say', while the second occurrence of *aya* is a reportative evidential ('hearsay' marker) rendered in the translation as 'it is said' (Chang 2006: 380).

(27) Paiwan (Austronesian; Chang 2006: 380)
khelem *ti* *cemedas* *tjay* *kalalu* *aya*
hit.AV NOM.PS.SG cemedas OBL.PS.SG kalalu say.AV
ti *palang* *aya*
NOM.PS.SG palang say.AV
'It is said that Palang said "Cemedas hit Kalalu".'

In practice, distinguishing between complement-taking utterance predicates and evidentials was generally a matter of paying close attention to the glossing, especially since some authors use terms such as "indirect speech" about constructions equivalent to those in (26) and (27) (cf. e.g. Derbyshire 1979 on the Hixkaryana evidential marker). In Article I, we discuss the development of evidentials from utterance predicates in more detail.

In summary, identifying complement constructions poses challenges, but in practice they are often not insurmountable. Needless to say, in cases of doubt we erred on the side of caution and generally restricted our analysis to glossed examples (cf. Article II).

3.5 Utterance-predicate complementation

As previously argued, morphosyntactic contrasts between complements expressing states-of-affairs and propositions are found within several different predicate classes, such as perception predicates and knowledge predicates. In this section, I will argue that utterance-predicate complementation, which is the main topic of Articles I and II, is of particular interest to the study of the contrast between states-of-affairs and propositions.

The first reason for studying utterance-predicate complementation is that this type of complementation provides a means for better understanding the relationship between the structural characteristics of the major sentence types: declaratives, interrogatives and imperatives, on the one hand, and those of clausal complements, on the other. As discussed in Chapter 2, declaratives and interrogatives arguably express propositions, while imperatives express states-of-affairs (Boye 2012: 199-206, cf. Table 1.1.).

In crosslinguistic typologies, declarative independent clauses have been used as the point of comparison for the verb form in complement clauses.

According to Cristofaro (2003: 54-55), deranked complements are complements that do not have the same morphosyntactic structure as declarative independent clauses (cf. Section 3.3.2). The balancing-deranking contrast is thus used to compare independent and dependent clauses, but this is especially problematic in relation to imperatives. Imperatives are typically "the simplest and the most straightforward forms in a language" and often have different TAM-marking than declarative clauses in the same language (Aikhenvald 2010: 89).

The definition of balanced complements often entails that imperatives are deranked. However, the definition focuses on subordination and not on the type of illocution reported (i.e. an assertion, a polar question or a command/request). But the question is whether deranked forms in directive utterance-predicate complements are deranked because they are subordinate or because they report imperatives. A more systematic approach to the distribution is therefore to compare the verb forms of complements reporting assertions with declaratives, complements reporting questions with interrogatives and complements reporting requests/commands with imperatives. Such an analysis is proposed in Article II.

A further reason for studying utterance-predicate complementation is that utterance-predicate complement constructions are arguably prototypical complement constructions. Most languages can reasonably be expected to have some means of reporting speech – and it is therefore likely that it will be described in reference grammars, thus providing data for the study. One crosslinguistic fact in support of the claim that utterance-predicate complement constructions are prototypical is that complementizers may develop from utterance predicates (e.g. Saxena 1988, Lord 1993, Cristofaro 2014b), typically going from marking reported speech to marking other types of complements.

In Ewe (Niger-Congo) for instance the complementizer *bé* is derived from the utterance predicate *bé* 'say'. As an utterance predicate *bé* can introduce reported speech as in (28a) and as a complementizer *bé* can mark a complement as in (28b). The same complementizer is also used for complements of other predicate types, such as the desiderative predicate *dí* want' as in (28c).

(28) Ewe (Niger-Congo; Lord 1993: 185;185;187, cf. Cristofaro 2014b: 14)

a. *me-bé me-wɔ-e*
I-say I-do-it
'I said "I did it" or 'I said that I did it.'

b. *é-gblɔ ná é bé yé dyi ye gaké ye kpe-dyí*
he-tell give him say he-EMPH bear SRP but SRP be-worthier
'He told him that he begot him but he was worthier.'

c. *me-dí bé máfle awua ɖewó*
I-want say I-SBJN-buy dress some
'I want to buy some dresses.'

Thus, utterance-predicate complementation may also provide insights into the structure of complement constructions with other types of complement-taking predicates. It should be noted that there is some dispute as to whether such a development is frequent. Noonan (2007) claims that complementizers derive more often from e.g. pronouns than from verbs, while Cristofaro (2014b) claims that it is a well-corroborated fact that this process is common.[6]

Finally, studying the domain of utterance-predicate complementation may provide new insights into reported speech. This goes for contrasts between direct and indirect reported speech such as the one illustrated in (29) in Modern Greek (Indo-European). (29a) is an example of direct speech without marking and (29b) is an example of indirect speech, which is marked by a complementizer *óti*.

(29) Modern Greek (Indo-European; Joseph & Philippaki-Warburton 1987: 3)

a. *o jánis ípe θa voiθíso*
the John.NOM said.3SG FUT help.1SG
'John said "I'll help".'

[6] I am not aware of any crosslinguistic survey of the frequency of the grammaticalization of utterance predicates. One possible explanation for the current dispute is that the frequency varies greatly within specific language families or geographic areas.

b. *o jánis ípe óti θa voiθísi/voiθúse*
the John.NOM said.3SG COMP FUT help.3SG/helped.3SG
'John said that he will help/would help.'

There is an extensive literature on reported speech (e.g. Coulmas 1986, Jäger 2007, 2010, N. Evans 2012, Spronck 2012 among many others), which I will not go into here, but it is notable that the contrast between states-of-affairs and propositions is generally not systematically investigated. This is likely due to two main factors: 1) utterance predicates are typically equated with assertive utterance predicates (that take propositional complements), as discussed in Section 3.2 (e.g. Dixon 2006, Noonan 2007, Cristofaro 2013), and 2) direct speech is sometimes defined as outside the domain of complementation - or seen as part of complementation only in some languages (e.g. Dixon 2006, Cristofaro 2013, Schmidtke-Bode 2014). Cristofaro (2013), for example, states that in her study of utterance-predicate complementation "direct speech clauses are regarded as utterance complements (…) unless the language also has some indirect speech construction". Similarly, Schmidtke-Bode (2014: 44), in his crosslinguistic study of complement constructions, includes direct speech exclusively "if the language in question lacks indirect means of reporting discourse and if the quoted clause acts as a syntactic argument of the matrix predicate". In this thesis, I take a different approach and study both direct and indirect speech as complementation, although a specific language may only have one construction. In the articles, we pay close attention to the semantic distinctions and commonalities between independent clauses and direct and indirect speech complements.

For these reasons, I focus mainly on utterance-predicate complementation, but I will continue to draw attention to other types of complementation throughout the thesis. In the following chapter, I describe the methodology of the crosslinguistic and language-specific studies presented in the following articles.

Chapter 4

Methodology

This chapter documents the data collection process and the coding procedures for the articles in Articles I-III in greater detail than was feasible in the articles themselves. Section 4.1 describes the crosslinguistic study which is used in the introductory chapters and in Articles I and II. Section 4.2 describes the corpus study of English and Danish utterance-predicate complementation, which supplements the crosslinguistic study in Article II. Section 4.3 describes the Ruuli corpus data and the coding procedures for the study in Article III.

4.1 Crosslinguistic study

The crosslinguistic study of utterance-predicate complementation is based on a sample of genetically and geographically diverse languages. The sample is a modification of the 200-language sample included in the *World Atlas of Language Structures Online* (WALS) (Dryer & Haspelmath 2013).[7] The WALS-sample has been put together by a number of language experts who have considered the variety of the sample. I removed the languages that the editors of WALS recommended removing for future typological studies to improve the genealogical balance of language families (Comrie et al. 2013).[8]

[7] For some reason the "200"-sample consists of 202 languages.

[8] The 29 removed languages that are included in the full WALS 200[202]-language sample are: Bawm, Batak (Karo), Carib, Nkore-Kiga, Koasati , Drehu, English (of a choice between English and German), Kiribati, Hebrew (Modern) (of a choice between Hebrew

The resulting sample used in the thesis includes 173 languages, which are listed in Table 4.1. The languages are sorted geographically according to which macroarea they belong to, according to the WALS. A list of the sample with genetic information (genus and family) can be found in Appendix A. The names of languages, language families and macroareas that appear in the thesis also adhere to the WALS classification.

Table 4.1: The 173-language sample sorted according to macroareas in the sample (as classified by the WALS)

Macroarea	Languages	Sum
Africa	Bagirmi, Bambara, Beja, Berber (Middle Atlas), Diola-Fogny, Ewe, Fur, Grebo, Hausa, Igbo, Iraqw, Ju|'hoan, Kanuri, Kera, Khoekhoe, Koyraboro Senni, Krongo, Kunama, Lango, Maba, Malagasy, Murle, Ngiti, Nubian (Dongolese), Oromo Boraana, Sango, Supyire, Swahili, Yoruba	29
Australia	Bunuba, Gooniyandi, Kayardild, Mangarrayi, Martuthunira, Maung, Nunggubuyu, Tiwi, Ungarinjin, Wambaya, Wardaman, Ngiyambaa	12
Eurasia	Abkhaz, Ainu, Arabic (Gulf), Armenian, Basque, Brahui, Burmese, Burushaski, Chukchi, Evenki, Finnish, French, Garo, Georgian, German, Greek (Modern), Greenlandic (West), Hindi, Hmong Njua, Hungarian, Hunzib, Ingush, Irish, Japanese, Kannada, Kayah Li (Eastern), Ket, Khalkha, Khasi, Khmer, Khmu', Korean, Ladakhi, Lak, Latvian, Lepcha, Lezgian, Mandarin, Meithei, Mundari, Muong, Nenets, Nivkh, Persian, Russian, Semelai, Thai, Turkish, Vietnamese, Yukaghir (Kolyma)	50
North America	Acoma, Bribri, Cahuilla, Chinanctec (Lealao), Comanche, Coos (Hanis), Cree (Plains), Haida, Jakaltek, Karok, Kiowa, Lakhota, Maricopa, Miwok (Southern Sierra), Mixtec (Chalcatongo), Nahuatl (Tetelcingo), Nez Perce, Oneida, Otomí (Mezquital), Pomo (Southeastern), Rama, Slave, Squamish, Tlingit, Tsimshian (Coast), Tunica, Wichita, Yaqui, Yuchi, Yurok, Zoque (Copainalá)	31
Papunesia	Alamblak, Amele, Arapesh (Mountain), Asmat, Chamorro, Daga, Dani (Lower Grand Valley), Ekari, Fijian, Imonda, Indonesian, Kewa, Lavukaleve, Marind, Maybrat, Menya, Paiwan, Sentani, Suena, Taba, Tagalog, Tukang Besi, Una, Yimas	24
South America	Abipón, Apurinã, Araona, Awa Pit, Aymara (Central), Barasano, Canela-Krahô, Cayuvava, Epena Pedee, Guaraní, Hixkaryana, Huitoto (Minica), Ika, Mapudungun, Ndyuka, Paumarí, Pirahã, Qawasqar, Quechua (Imbabura), Sanuma, Selknam, Shipibo-Konibo, Trumai, Warao, Wari', Wichí, Yagua	27

and Arabic), Koromfe, Kilivila, Kongo, Kobon, Kutenai, Luvale, Maori, Makah, Navajo, Pitjantjatjara, Paamese, Passamaquoddy-Maliseet, Rapanui, Spanish, Urubú-Kaapor, Usan Ngiyambaa, Yidiny, Yup'ik (Central), Zulu.

The WALS sample was chosen under the assumption that the languages included were well-described, thus improving the chances of finding enough examples of clausal complementation. Table 4.2 shows the main aspects of complementation surveyed using the available data sources (mainly reference grammars[9]) for each language in the sample.

Table 4.2: Variables considered for each language in the sample

Variable	Possible values
Predicate polysemy	assertive-directive/assertive-interrogative/interrogative-directive
Polysemy with other predicate type	[predicate], e.g. 'think'/NA
Propositional complement	yes/no/NA
State-of-affairs complement	yes/no/NA
Contrast between state-of-affairs and proposition	yes/no/NA
Epistemic modification possible	yes/no/NA
Direct speech	yes/no
Illocutionary value of complement	assertive/interrogative/directive
Indirect speech	yes/no
Polyfunctionality direct/indirect speech	yes/no
Complementizer presence/absence contrast	yes/no
Complementizer origin	[origin], e.g. utterance verb
Deranking/balancing of assertive complement	deranked/balanced/NA
Deranking/balancing of question	deranked/balanced/NA
Deranking/balancing of directive complement	deranked/balanced/NA

I noted whether the language had assertive, interrogative and/or directive predicates, and whether there was any polysemy between predicates (i.e. whether a language had a predicate such as English *tell*, capable of being used as an assertive or a directive predicate). It was also noted whether the language had complement-taking predicates that were polysemous with other predicate types, e.g. an utterance predicate meaning 'say', which could also mean 'think'. Next it was noted whether a propositonal complement of an utterance predicate was found and whether a state-of-affairs complement was

[9]The identification of relevant reference grammars was based Glottolog's list of complete grammars (cf. Hammarström et al. 2018).

found, and if there was a coded contrast between them. It was also noted whether epistemic modification of a complement was possible.[10] Presence of direct speech complements and presence of indirect speech complements was also noted. The last variables concern the occurrence of complementizers and balanced and deranked verb forms in complements.

I am aware that the reliance on reference grammars as the main source of information comes with a set of drawbacks. For obvious reasons, reference grammars rarely provide an exhaustive descriptions of complementation, and what is more, example sentences illustrating complement clause constructions are often spread out in the grammar. Some of the central data on Khmer, for example, was found in a section on negation (Haiman 2011: Ch. 7, Section 6). It is nevertheless often possible to find descriptions of direct or indirect speech. In some of the grammars consulted, such as the ones on Modern Greek (Joseph & Philippaki-Warburton 1987: 3) and Imbabura Quechua (Cole 1982: 13), direct speech is even the first topic in the grammar template to be covered by grammar authors in the Croom Helm Descriptive Grammars series. One of the major challenges was that none of the reference grammars consulted had analyzed complementation in terms of the contrast between states-of-affairs and propositions – articles on complementation in Japanese (Horie 2000) and Latvian (Holvoet 2016) did, however, which meant that the interpretation was generally completely left to my own and my co-authors' judgement. Of course, grammar authors in a some cases described contrasts between e.g. 'direct perception' and 'indirect perception' (like we also do in Article III). which in some cases corroborated my analysis. But in many cases it required thorough reading of entire grammars to find examples of contrasts between states-of-affairs and propositions.

4.2 Corpus study

In addition to the crosslinguistic study of utterance-predicate complementation, Article II includes a small corpus study of utterance-predicate complementation in English and Danish. Data was extracted from the British

[10] Epistemic modification of a complement, would indicate that a complement is propositional. Unfortunately, evidence of such a possibility is very rarely found in reference grammars.

National Corpus of written English and the LANCHART corpus of spoken Danish. We suspected that we would need to go through a rather extensive manual sorting process, if we did not exclude some possibilities from the outset. We therefore restricted the search to only include the past tense form of the complement-taking predicate and restricted the subject and indirect object of the sentences to the third person personal pronouns *he*, *she*, *him* and *her* for English and *han* 'he', *hun* 'she', *ham* 'he' and *hende* 'her' for Danish. The target constructions and the corresponding search strings for the corpus study of English are given in Table 4.3 and the target constructions and corresponding search strings for the corpus study of Danish are given in Table 4.4.

Table 4.3: Search strings for complement constructions with *tell* and *ask* in the British National Corpus

Target	Search String
told + finite	"he told", "she told"
told + infinitive	"he told him to", "he told her to", "she told him to", "she told her to"
asked + finite	"he asked", "she asked"
asked + infinitive	"he asked him to", "she asked him to", "he asked her to", "she asked her to"

Table 4.4: Search strings for complement constructions with *sige*, *spørge* and *bede* in the LANCHART corpus

Target	Search String
sagde + finite	"han sagde" ('he said'), "hun sagde" ('she said')
spurgte + finite	"han spurgte" ('he asked'), "hun spurgte" ('she asked')
bad + infinitive	"han bad hende" ('he asked [lit. prayed] her (to)'), "hun bad ham" ('she asked [lit. prayed] him (to)')

The best case scenario for corpus research would be to extract the example tokens one needs and nothing more than that. This is, however, only very rarely possible, and in our study too, we had to manually code the examples. Especially the hits from the search in the LANCHART corpus needed

extensive manual cleaning.

This corpus study is not meant as any substantial contribution to frequency studies of complement-taking predicates, but meant to give an impression of the frequency of deranked and balanced verb forms over different types of utterance predicates. We leave it to future research to investigate the role of frequency for complementation asymmetries in more depth. [11]

4.3 Descriptive study

Article III is a descriptive study of clausal complementation in Ruuli, a threatened Great Lakes Bantu language spoken in the Nakasongola and Kayunga districts of central Uganda by approximately 190.000 speakers.

The study was done in collaboration with Alena Witzlack-Makarevich who is a part of the collaborative research project "A comprehensive bilingual talking Luruuli/Lunyara-English dictionary with a descriptive basic grammar for language revitalisation and enhancement of mother-tongue based education" (the Luruuli/Lunyara dictionary project) funded by the Volkswagen Foundation (PI Saudha Namyalo).

The discrepancy between the name of the language in our study and in the title of the Luruuli/Lunyara dictionary project is due to the fact that there is no consensus on the name of the language and that there is a great deal of variation with regards to its pronunciation. We chose Ruuli out of a number of names in use, including Luruuli/Lunyara, Ruruuli-Runyala, Runyala and Luduuli.

The description of the morphosyntactic and semantic characteristics of clausal complementation in Ruuli is based on 1500 tokens of written and spoken language extracted from a collection of data assembled by researchers from the Luruuli/Lunyara dictionary project. As of May 2018, the corpus of spoken Ruuli contained about 150,000 words of naturalistic speech (transcribed and translated). Further 50,000 words were available from digitalized written resources produced by the speakers' community. I did not partic-

[11] Research on frequency in complementation is already underway in e.g. the project Form-frequency correspondences in grammar (FormGram) (PI Martin Haspelmath) funded by the European Research Council under the European Union's Horizon 2020 research and innovation program. cf. https://research.uni-leipzig.de/unicodas/grammatical-universals/

ipate in any data collection in Uganda myself, but I worked with the collected, transcribed and partly glossed examples contained in the corpus, which I was graciously granted access to by Alena Witzlack-Makarevich.

We aimed to cover as many aspects of complementation as possible for a first description of Ruuli complementation, while placing the results in a typological context. We also compared the results to related Bantu languages. In contrast to the crosslinguistic study presented in the thesis, the study of Ruuli complementation looked at several different semantic classes of complement-taking predicates, such as knowledge predicates and perception predicates, as described in Section 3.2.

We first extracted 1000 examples of complement constructions from the corpus and coded each example for several semantic and morphosyntactic variables as shown in Table 4.5.

Table 4.5: Variables considered for each extracted example from the Ruuli corpus

Variable	Possible values
Semantic class of CTP	[class], e.g. knowledge
Complementizer	[comp]/NA
Polarity of matrix clause	negative/positive
Illocutionary type	assertion/question/directive/NA
Direct reported speech	true/false/ambiguous
State-of-affairs/propositions	state-of-affairs/proposition
Same-subject or different subject	same-subject/different-subject
Complement verb form	indicative/subjunctive/infinitive

Most of the variables coded correspond to the variables in Table 4.2, as described above (i.e. complementizer presence/absence, illocutionary type, direct reported speech and state-of-affairs/proposition). In the Ruuli study we additionally coded for semantic class of complement-taking predicates (as we were concerned with a broad range of complement-taking predicates, not only utterance predicates), the polarity of the matrix clause and same-subject vs. different-subject. We also noted the Ruuli-specific form of the verb (in contrast to the crosslinguistic study, where we contrasted balanced and deranked verb forms).

We identified the most frequent complement-taking predicates and looked at the distribution of complement types relative to different semantic classes

of complement-taking predicates. On the basis of this analysis, we refined our initial hypotheses on the distribution and found an extra approximately 500 examples of less frequent complement-taking predicates and complementation strategies that we discovered during the first coding phase.

Chapter 5

An overview of the articles

In this chapter, I provide an overview of the three articles included in the thesis and describe the contribution of the authors. I have written two articles in collaboration with my supervisor, Kasper Boye, and one article in collaboration with Alena Witzlack-Makarevich from Kiel University, Germany.

Article I: Layered semantic structure in independent utterances and direct and indirect reported speech

The first article, *Layered structure in direct and indirect reported speech*, is co-authored with Kasper Boye. The article argues that the differences and similarities between independent utterances and direct and indirect reported speech can be accounted for in terms of layered semantic structure and the distinction between illocutions, states-of-affairs and propositions. The main aim of the article is to develop a unified cognitive-functional analysis of utterance-predicate complementation. The presented analysis is implicitly assumed in Article II, which focuses on grammatical contrasts in indirect reported speech.

Kasper Boye contributed to the conception of the article, the data and the drafting of the article. I collected the crosslinguistic data, contributed to the conception of the article, the data analysis and the drafting of the article. The article is an adapted and elaborated English language version of a previously published working paper, *Lagdelt indholdsstruktur i selvstændige ytringer*

og direkte og indirekte ytringsgengivelse [Layered content structure in independent utterances and in direct and indirect report of utterances] (Sørensen & Boye 2017).

Article II: Grammatical contrasts in utterance-predicate complementation: from iconicity to frequency, and back

The second article, *Grammatical contrasts in utterance-predicate complementation: from iconicity to frequency, and back* is co-authored with Kasper Boye. The article presents a crosslinguistic study of assertive, interrogative and directive complements of utterance-predicates based on a sample of 173 languages. It furthermore discusses the relative merits and limitations of iconicity-based and frequency-based accounts of grammatical contrasts in complementation. The main aim of the article is to argue that a difference in conceptual complexity between propositions and states-of-affairs is an explanatory factor for grammatical contrasts in clausal complementation.

Kasper Boye provided input to the design of the study and the interpretation of the data, provided theoretical input to the discussion of iconicity and frequency and to the discussion of the theories of states-of-affairs and propositions and participated in drafting the paper. I contributed to the design of the study, collected and analyzed the crosslinguistic data and carried out the corpus studies of English and Danish complementation, contributed to the discussion of iconicity and frequency and participated in drafting the paper.

Article III: Complementation in Ruuli (Bantu; JE103)

The third article, *Complementation in Ruuli (Bantu; JE103)*, is co-authored with Alena Witzlack-Makarevich. It is a study of clausal complementation in the previously undescribed language Ruuli (Great Lakes Bantu, Niger-Congo) spoken in Uganda, with a particular focus on the distinction between states-of-affairs and propositions. The main aim of the third article is to provide a first description of clausal complementation in Ruuli, while examining how the contrast between states-of-affairs and propositions is expressed.

Alena Witzlack-Makarevich collected the Ruuli data in Uganda in collaboration with colleagues from the Luruuli-Lunyara Dictionary Project (cf. Chapter 4) and provided the resulting corpus data (transcribed and translated spoken language). She contributed to the conception of the article, the extraction and coding of corpus data, glossing of examples, data analysis and drafting of the paper. I contributed to the conception of the article, the extraction and annotation of corpus data, the analysis of the data, and participated in drafting the paper. I also provided theoretical input on complementation and the contrast between states-of-affairs and propositions.

Preliminary conclusion

The three articles approach the distinction between states-of-affairs and propositions from crosslinguistic and language-specific angles. Article I focuses solely on utterance predicate complementation, Article II presents a crosslinguistic study of utterance-predicate complementation, but extends the scope of the analysis to cover complementation with other types of predicates. Article III analyzes an even wider range of predicates in an in-depth study of complementation in a single language, thus providing an example of how the distinction between state-of-affairs and propositions can be integrated in a language-specific study. Together the articles demonstrate that the contrast between states-of-affairs and propositions is pervasive and of central interest to crosslinguistic and language-specific studies of grammatical contrasts in clausal complementation.

Chapter 6

Article I: Layered semantic structure in independent utterances and direct and indirect reported speech

6.1 Introduction

Speech report presents a complex picture, and in the (vast) literature on the subject (e.g. Coulmas 1986, Jäger 2007, N. Evans 2012, Spronck 2012 to name a few), there is accordingly a tendency to focus on only part of the picture. Jäger (2007), for example, studies the crosslinguistic coding strategies of indirect report of assertions, thus leaving out of consideration direct report as well as indirect report of other speech act types. Cristofaro (2013) takes both direct and indirect report into account in her study of complement verb forms in utterance-predicate complementation, but like Jäger (2007), she is only concerned with reports of assertions. Schmidtke-Bode (2014) and Schmidtke-Bode & Diessel (2017) study complement-taking predicates used to report assertions or directive speech acts (the latter kind of predicates are named jussives); however they exclude predicates that are used to report questions.

In this paper, we aim to give a fuller picture. Based on a functional model of layered semantic structure, we outline a unified theory of:

- Semantic differences and similarities between the three sentence types declarative, interrogative and imperative, which are argued to code three illocution types: assertions, polar questions and directives (that is, the differences and similarities between (1a), (1b) and (1c) in Table 6.1 below).

- Semantic differences and similarities between main clauses and two types of complements clauses: complements used for direct report of utterances, and complements used for indirect report (that is, the differences and similarities between (1), (2) and (3) in Table 6.1 below).

For the sake of clarity, we ignore hybrids between direct and indirect reported speech and differences between the two kinds of report that have to do with tense or other types of deixis (see Togeby 2014). That is, we deal exclusively with speech report as it is reflected semantically in structurally (morphologically or syntactically) distinguishable clause types. In order to be precise about these clause types, we need to classify them along two parameters that are often conflated terminologically. One parameter has to do with the prototypical uses of the clauses. We will refer to clauses as *independent* if they are prototypically used to form utterances on their own, and as *dependent* if they are prototypically used as subordinate clauses. The other parameter has to with the actual uses of the clauses. We will refer to clauses as *main clauses* if they are actually used to form utterances on their own, and as *complement clauses* if they are actually embedded as (what can be assumed to be) complements.

This classification thus operates with four types of clauses: 1) *independent main clauses* such as (1a-c), which constitute utterances; 2) *independent complement clauses* such as (2a-c), which are used for direct utterance report; 3) *dependent complement clauses* such as (3a-c), which are used for indirect utterance report; 4) *dependent main clauses*, which correspond to one subtype of so-called insubordinate clauses (e.g. Evans & Watanabe 2016), and which are irrelevant for and will be ignored in the present paper.

Table 6.1: Independent clause structures used for independent utterances and clause structures used for reported speech

	Assertion	Polar question	Directive
Independent main clauses used for independent utterances	(1a) *She is leaving*	(1b) *Is she leaving?*	(1c) *Leave!*
Independent complement clauses used for direct utterance report	(2a) *He said:* "*She is leaving*"	(2b) *He said:* "*Is she leaving?*"	(2c) *He said:* "*Leave!*"
Dependent complement clauses used for indirect utterance report	(3a) *He told (that) she is leaving*	(3b) *He asked if she is leaving*	(3c) *He told/asked her to leave*

As for the sentence and utterance type dimension (the horizontal dimension in Table 6.1), we argue that in addition to illocutionary differences, the three types differ in terms of the distinction between proposition (truth-valued meaning unit) and state-of-affairs (non-truth-valued meaning unit). As for the report dimension (the vertical dimension in Table 6.1), we argued that independent complement clauses differ from independent main clauses in lacking what Hare (1970) called a neustic, whereas dependent complement clauses differ in additionally lacking what Hare called a tropic.

The central part of our argumentation is based on English and Danish (in contrast to English – cp. (2a) and (3a) – Danish makes a structural (syntactic) distinction between complements used for direct and indirect report of assertions). However, we intend our typology to be potentially universal and to be understood as based on what Haspelmath (2010) calls "comparative concepts". Towards the end of the paper, we therefore discuss a number of languages that might seem to present challenges to the typology.

In Section 6.2 we briefly review theories of layered semantic structure and in Section 6.3 we analyze independent main clauses in relation to layered semantic structure. In Section 6.4 we analyze the relationship between independent main clauses and independent complement clauses and in Section 6.5, we analyze the relationship between independent complement clauses and dependent complement clauses. Based on these analyses, in Section 6.6 we discuss the crosslinguistic diversity in reported speech constructions and in Section 6.7 we propose two hypotheses regarding the grammat-

icalization of utterance predicates. Section 6.8 is a brief conclusion.

6.2 Layered semantic structure

Models of layered clause structure play a central role in e.g. Functional (Discourse) Grammar (Hengeveld 1989, Dik 1997, Hengeveld & Mackenzie 2008) and in Role and Reference Grammar (Foley & Van Valin 1984, Van Valin & LaPolla 1997). They are intended to capture insights similar to those captured in formal theories by tree-structure hierarchies, but in contrast to formal approaches, especially Functional (Discourse) Grammar has emphasized that the different clause levels are associated with different meaning units. Harder (1996: 228-243) has pointed out that in fact only these semantic units are organized in hierarchical layers – the clause itself is purely linear, just take a look at one.

The present paper is based on Harder's concept and model of *layered semantic structure* (or *layered content structure*) in the form developed for clause semantics (as opposed to noun phrase semantics) in Boye (2012: Ch. 4 and 5). The relevant parts of layered structure are illustrated in the model in Figure 6.1.

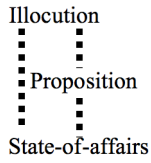

Figure 6.1: A model of layered semantic structure (based on Boye 2012: 183-315)

The model makes a distinction between three meaning units: states-of-affairs, propositions and illocutions. Illocutions are meaning units which include a specification of the type of speech act intended. Propositions and states-of-affairs are less well-known in linguistics. Roughly, states-of-affairs can be understood as meaning units that cannot be said to be true or false and cannot be epistemically evaluated (in other words, meaning units that do not have a truth value). In contrast, propositions can be understood as meaning units that can be said to be true or false and can be epistemically evaluated (in

other words, meaning units that have a truth value[12]) (cf. e.g. Vendler 1967, Ransom 1986, Hengeveld 1989, Dik & Hengeveld 1991, Horie 2000, Cristofaro 2003, Boye 2010, 2012). These three meaning units are related in the following way. Propositions always contain a state-of-affairs (something with a truth value presupposes something without), while states-of-affairs can never contain a proposition (unless of course the proposition is embedded as an argument). Illocutions either contain a proposition (including a state-of-affairs) or only a state-of-affairs, as illustrated by the dotted lines in Figure 6.1 connecting the three structural layers.

6.3 Layered semantic structure in independent main clauses

Our analysis of independent main clauses in terms of layered semantic structure is adopted from Boye (2012: 194–195, 2013). As for the illocutionary level, declaratives, polar interrogatives and imperatives code the illocutionary values of, respectively, assertion, polar question and directive (or "command"). Of course, these coded values (marked by sentence types which thus are types of what Searle (1969) called "illocutionary force indicating devices") may be overridden when a coded direct speech act is used to convey a non-coded, context-dependent indirect speech act. However, one argument in support of this "coding" analysis is that there are restrictions on which kinds of speech acts the different sentence types can be used to make. For instance, all declarative independent main clauses can be used to make assertive speech acts, but it is not possible to use a declarative clause, such as *My aunt had a dog named Benny* to make a directive speech act – at least not if only what is coded in the clause is taken into consideration (two persons could agree that this clause means 'go', but this would mean treating the clause as a holophrastic idiom rather than as a clause). This is an argument that declaratives code (i.e. are conventionally associated with) assertive illocutionary value.

As for the lower levels of layered clause structure, our analysis is this:

[12]There is reason to believe that truth value is not the optimal explanatory model, but this is not central to the present analysis (cf. Boye 2008, 2010, 2012 and Sørensen & Boye 2015 for a more thorough discussion of the distinctions)

declaratives and polar interrogatives code propositions (including states-of-affairs). In contrast, imperatives code states-of-affairs only. This analysis captures the fact that in some languages, declaratives and polar interrogatives share a feature not found in imperatives. This feature, which is also captured by Generative Grammar's notions of 'I' and 'IP' is sometimes called "indicative". According to the above analysis, indicative markers are markers of propositional status. One argument in support of this analysis has to do with the possibility of epistemic evaluation as a criterion for propositional status (for more arguments see Boye 2013): Only declaratives and polar interrogatives allow of epistemic evaluation, imperatives do not, as illustrated in (4) (note that due to the uncertainty implied by polar questions, there are restrictions on which polar interrogatives can be modified epistemically).

(4) a. *She is probably leaving*

b. *Is she possibly leaving?*

c. **Leave probably/possibly!*

The analysis of independent main clauses is as follows, then: declarative and polar interrogative clauses contain an illocutionary layer, a propositional layer and a state-of-affairs layer, while imperative clauses only contain an illocutionary layer and a state-of-affairs layer. In other words, assertions and polar questions concern propositions, while directives concern states-of-affairs (Boye 2013).

The analysis is summarized in Table 6.2, where I = illocution, P = proposition, and S = state-of-affairs.

Table 6.2: Independent main clauses and layered semantic structure (I = illocution, P = proposition, and S = state-of-affairs).

	Assertion	Polar question	Directive
Independent main clauses used for independent utterances	(1a) *She is leaving* I-P-S	(1b) *Is she leaving?* I-P-S	(1c) *Leave!* I–S

6.4 The difference between independent main clauses used for independent utterances and dependent main clauses used for direct utterance report

Since declarative, interrogative and imperative sentence types code assertive, polar interrogative and directive illocutionary value, respectively, it follows that declarative, interrogative and imperative sentences have these illocutionary values also when they are embedded in a superordinate clause and are used for direct report. In English, for instance, the sentence type is marked by means of word order, and the word order differences found when independent main clauses are used for independent utterances are maintained when utterances are reported by means of independent complement clauses (cp. (1a-c) and (2a-c)).

But if this is the case, what is the semantic difference between independent main clauses used for independent utterances and independent complement clauses used for direct utterance report? Haberland (1986) describes the central difference as follows:

> "The difference is that the illocutionary force of the model utterance is only indicated or displayed in the report, not performed or enforced" (Haberland 1986: 220)

In other words, the directly reported utterance retains the illocutionary value of the independent utterance, but lacks the potential for being performed as a speech act. This differences is captured by Hare's (1970) distinction between a "tropic" and a "neustic":

> "Now although a neustic has to be present or understood before a sentence can be used to make an assertion or perform any other speech act, it is in virtue of its tropic that it is used to make an assertion and not to perform some other speech act"
> (Hare 1970: 22)

Tropic corresponds to what was called illocutionary value above, while

neustic corresponds to the performative aspect. In other words, it is possible to split the illocution into two components, as shown in Figure 6.2.

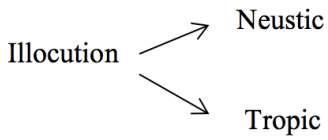

Figure 6.2: Splitting the illocution into neustic and tropic

Accordingly, corresponding independent main clauses and independent complement clauses are related as follows: they share a tropic, but only the first have a neustic. This analysis is summarized in Table 6.3, where I^n = neustic, I^t = tropic, P = proposition, and S = state-of-affairs.

Table 6.3: Independent main clause structures, used in independent utterances and clause structures used for direct utterance report (I^n = neustic, I^t = tropic, P = proposition, S = state-of-affairs).

	Assertion	Polar question	Directive
Independent main clauses used for independent utterances	(1a) *She is leaving* I^n-I^t-P-S	(1b) *Is she leaving?* I^n-I^t-P-S	(1c) *Leave!* I^n-I^t–S
Independent complement clauses used for direct utterance report	(2a) *He said:* "*She is leaving*" -I^t-P-S	(2b) *He said:* "*Is she leaving?*" -I^t-P-S	(2c) *He said:* "*Leave!*" -I^t–S

We have described our analysis as a semantic one, and by semantics we understand coded (= conventionalized) meaning. In this connection, one might ask whether a neustic can be coded in the world's languages, or in other words: do independent main clauses code a neustic, or do independent complement clauses code the absence of a neustic?

Evidently, at least the absence of a neustic can be marked. In written language, absence of a neustic can be marked by quotation marks. In sentences (2a)–(2c) in Table 6.1 and 6.3, absence of a neustic is furthermore marked by the complement-taking predicate clause he said, which shows that the complement clause reports an utterance. Disregarding the quotation marks of the written language, the absence of a neustic is not always unambiguously marked, however. In a discussion of reported speech in Danish, Hansen &

Heltoft (2011) provide the following example:

> "Pludselig var der en af tjenerne, der stak hovedet ind i køkkenet og spurgte 'Jacob, er det rigtigt, at du er naturist?' Så tak! Hvor fejt ville det ikke være at lyve? Og dumt, for hvorfor egentlig skamme sig? Så 'Jo, det er jeg da'". (Hansen & Heltoft 2011: 1655).
>
> 'Suddenly one of the waiters came into the kitchen and asked, 'Jacob, is it true that you are a nudist?'. Oh my! How cowardly wouldn't it be to lie? And how stupid, because why be ashamed? So, 'Yes, I am'"

In cases like this, the absence of a neustic is most likely not coded (disregarding the quotation marks of the written language), but a pragmatic phenomenon.

It should be noted that Hare (1970) in addition to the terms neustic and tropic operates with the term *frastic*. This term covers what we call proposition and states-of-affairs in this paper, but Hare does not distinguish between these subtypes of *frastic*.

6.5 The difference between independent complement clauses used for direct report and dependent complements used for indirect report

The next question is whether there is a semantic difference between independent complement clauses used for direct utterance report and dependent complement clauses used for indirect utterance report. In what follows, we first argue against the answers to this question provided by Hare (1970) and Harder (2016). Subsequently, we expand our own analysis, the main point of which is that dependent complement clauses lack not only a neustic, but also a tropic.

6.5.1 Hare (1970)

As mentioned in Section 6.3, Hare (1970) does not distinguish between propositions and states-of-affairs. Moreover, his focus is on English clauses with declarative word order, and these clauses may function both as independent and dependent complements and can thus be used for both direct and indirect utterance report. Accordingly, Hare does not make a distinction between these types of complement clauses or the associated types of utterance report. As a consequence of this, Hare proposes essentially the same analysis for dependent complements and indirect report as the one we have proposed for independent complements and direct report: according to Hare, dependent complement clauses have a tropic, but not a neustic – just like independent complements.

There are at least three arguments against Hare's analysis of dependent complement clauses. Firstly, it overlooks the fact that many languages use a distinct clause type for dependent complements, and that this clause type lacks the tropic marker found in independent main and complement clauses. For instance, Danish, like English, distinguishes sentence types by means of word order, as illustrated in (5).

(5) a. *Hun spiller klaver.*
she play.PRS piano
'She plays/is playing the piano.'

b. *Spiller hun klaver?*
play.PRS she piano
'Does/is she play/playing the piano?'

c. *Spil klaver!*
play.IMP piano
'Play the piano!'

This difference is maintained in independent complements used for direct utterance report.

(6) a. *Han sagde: "Hun spiller klaver".*
he say.PST she play.PRS piano
'He said: "She plays/is playing the piano".'

b. *Han sagde "Spiller hun klaver?".*
he say.PST play.IMP she piano
'He said: "Does/is she play/playing the piano?"'

c. *Han sagde: "Spil klaver!"*
he say.PST play.IMP piano
'He said: "Play the piano!"'

But the difference is not found in dependent complements used for indirect utterance report. Here, distinctions in terms of complementizers or finiteness are found instead.

(7) a. *Han fortalte (at) hun spiller klaver.*
he tell.PST COMP she play.PRS piano
'He told that she plays/is playing the piano'

b. *Han spurgte om hun spiller klaver.*
he ask.PST COMP she play.PRS piano
'He asked if she plays/is playing the piano'

c. *Han bad hende spille klaver.*
he tell/ask.PST her play.INF piano
'He told/asked her to play the piano.'

Secondly, it overlooks the fact that, at least in the languages known to us, dependent complements do not have a separate tropic marking. For example, the Danish complementizer *om* (which has a function similar to English *if)* does not mark "indirect question" (Boye 2008, Nordström & Boye 2016). If it did, one would expect that the interrogative particle *mon* (roughly, 'perhaps') could always be felicitously inserted in a complement clause marked by *om*. This is often possible, but not always. (8a), where *mon* can be read as combining harmonically with *spurgte* 'asked' is fully acceptable, but (8b) and (8c), where a harmonic reading is excluded, are only marginally ac-

ceptable, and only if *mon* is read metalinguistically as belonging to another speaker or utterance.

(8) a. *De spurgte, om der mon var mere at spise.*
they ask.PST COMP there PRT be.PST more to eat.INF
'They asked whether there was perhaps anything left to eat.'

b. * *De drak uanset om der mon var mere at spise.*
they drink.PST irrespective COMP there PRT be.PST more to eat
'They were drinking even if there was perhaps anything left to eat.'

c. * *De fortalte, om der mon var mere at spise.*
they tell.PST COMP there PRT be.PST more to eat
'They told whether there was perhaps anything left to eat.'

Thirdly, Hare's analysis is incompatible with the fact that when dependent complement clauses clearly report utterances, the illocutionary value is expressed by the complement-taking predicate.

(9) a. Report of a question
Han spurgte, om hun spiller klaver.
he ask.PST COMP she play.PRS piano
'He asked if she plays/is playing the piano.'

b. Not a report of a question
Han sagde, om hun spiller klaver.
he say.PST COMP she play.PRS piano
'He said if she plays/is playing the piano.'

In contrast, when utterances are reported directly by means of an independent complement, the complement-taking predicate does not necessarily express the illocutionary value of the original utterance. In (10a), the illocutionary value of the polar question is expressed both by the complement-taking predicate *spørge* 'ask' and by the interrogative word order in the independent complement clause. In (10b), however, the same illocutionary value is only expressed by the interrogative word order.

(10) a. Report of a polar question
Han spurgte: "Spiller hun klaver?".
he ask.PST play.PRS she piano
'He asked: "Does/is she play/playing the piano?".'

b. Report of a polar question
Han sagde: "Spiller hun klaver?".
he say.PST play.PRS she piano
'He said: "Does/is she play/playing the piano?".'

By contrast, in (9a) the tropic value of polar question is only marked by the complement-taking verb *spørge* 'ask'. Accordingly, (10b), where *spørge* is replaced with the tropically neutral *sige* 'say', can (in conservative language use, at least) only be read as a report of an assertive utterance. Based on these counterarguments, Hare's analysis must be discarded.

6.5.2 Harder (2016)

Harder (2016) follows Hare (1970) in arguing that both independent and dependent complement clauses have an illocution. But in contrast to Hare, who believes that these two types of clauses have the same type of illocution (a tropic), Harder believes that the illocution is not of the same kind for the two types of clauses. According to Harder, dependent complement clauses have a special kind of illocution, what he calls a "subillocution".

The idea of a subillocution is based on a view of subordination as a functional alternative to an illocution. This view is supported by languages like Hidatsa (Siouan), Huichol (Uto-Aztecan), and West Greenlandic (Eskimo-Aleut), where subordination markers are found in the same paradigm as illocution markers. The idea goes like this: A sentence can be launched as an independent speech act or as part of a superordinate construction. Or the other way around: Speech acts and superordinate constructions both provide a functional specification that enables us to know what to do with any given sentence and its content (this is the functional motivation behind Generative Grammar's generalization in terms of 'CP': 'C' represents subordination or illocution). Harder's idea is that subordinating conjunctions, in addition to marking subordination and possibly subordination type, may express a func-

tional specification of what hearers should make of the subordinate clause. This specification is the subillocution. For example, according to Harder, the Danish complementizer *om* and the English complementizer *whether* both mark that the proposition expressed by the complement clause is an "issue" (Harder 2016: 803) – issues are defined as "things you need to decide on in order to take action". Harder's (2016: 903) central example is the question in (11) posed by a teacher.

(11) *How does one determine whether the earth moves around the sun or not?*

According to Harder the complement clause *whether the earth moves around the sun or not* expresses an issue, as marked by *whether*.

However, there are at least three arguments against regarding subillocutions as a special phenomenon. Firstly, subillocutions are not well-defined, and therefore, they are difficult to identify. Secondly, it is unclear whether subillocutions are neustics or tropics. Thirdly, and most importantly, subillocutions can be reduced away without any theoretical or analytical consequences. At closer inspection of (11), it ought to be considered whether the meaning of "issue" is in fact due to the meaning of the complement-taking verb *determine* and the superordinate *how*-question. Consider also (12).

(12) *He asked, whether she is leaving*

Here, the *whether*-clause merely reports the content of a polar question. If the *whether*-clause expresses an issue, the question, which is reported, must also express an issue. If that is the case, however, there is nothing special about *whether*-clauses. If, on the other hand, the sentence does not express an issue, *whether* cannot be analyzed as an "issue-marker". Especially in light of the last argument, Harder's analysis and the idea of subillocutions must be discarded. Issues and subillocutions are usage-epiphenomena at best.

6.5.3 Dependent complement clauses have no illocution

As argued above, there is no reason to believe that dependent complement clauses have illocutions, neither neustic nor tropic ones. Apart from lacking

neustics and tropics, however, dependent complement clauses that are used for indirect report of assertions, polar questions or directives can straightforwardly be analyzed as having the same semantic specifications as the main clauses used in independent utterances or for direct report.

The main points of this analysis are as follows.

1. Dependent complement clauses used for indirect report designate propositions, as in (13a) and (13b), or states-of-affairs, as in (13c).

(13) a. *He told that she is leaving.*

b. *He asked if she is leaving.*

c. *He told/asked her to leave.*

2. Complementizers can, in addition to having "complementizing function", mark propositional or state-of-affairs status, or they can modify propositions or states-of-affairs (see Boye et al. 2015 and Kehayov & Boye 2016a for crosslinguistic possibilities)

3. The functional specification of a subordinate clause (what to use it for) is provided by items in the superordinate clause (e.g. complement-taking predicates) in collaboration with possible subordinators (which have neither illocutionary nor subillocutionary value). Thus, a sentence like (13b) must be analyzed as follows:

   - The complement clause expresses the propositional content of a reported utterance.
   - The complement-taking predicate *ask* expresses the tropic of the reported utterance: polar question.
   - The complementizer *if* indicates uncertainty about the proposition expressed by the complement clause (Nordström & Boye 2016: 140). Exactly this complementizer is used because its epistemic value is compatible with polar questions: also polar questions imply uncertainty about the proposition questioned (Boye 2012: 103-105).

The analysis of the difference between independent and dependent complement clauses is summarized in Table 6.4, where I^t = tropic, P = proposition and S = state-of-affairs.

Table 6.4: The differences between direct and indirect utterance report (I= Neustic illocution, I = Tropic Illocution, P = Proposition, S = State-of-affairs).

	Assertion	Polar question	Directive
Independent complement clauses used for direct utterance report	(2a) *He said:* "*She is leaving*" -I^t-P-S	(2b) *He said:* "*Is she leaving?*" -I^t-P-S	(2c) *He said:* "*Leave!*" -I^t–S
Dependent complement clauses used for indirect utterance report	(3a) *He told* *(that) she is leaving* –P-S	(3b) *He asked* *if she is leaving* –P-S	(3c) *He told/asked* *her to leave* —-S

6.6 Crosslinguistic diversity in reported speech constructions

The above analysis entails that the difference between direct and indirect utterance report is expressed through the difference between, respectively, independent and dependent complement clauses. Not all languages make such a distinction systematically. In English, the distinction is made only for the report of polar questions and directives. In the case of report of assertions, English makes no structural distinction between independent and dependent complements and thus no distinction between direct and indirect report. In Danish, on the other hand, a structural distinction is made for all types of utterances.

Something similar can be said of other, typologically distinct languages. In Epena Pedee (Choco), for example, the difference between direct and indirect report is expressed through the presence and absence, respectively, of a declarative, interrogative or imperative marker. (14a) is an example of direct report, while (14b) is an example of indirect report.

(14) Epena Pedee (Choco;Harms 1994: 173)

a. *nepiri-pa-či-dá rey táu pʰārîu bi-da a-hi-dá*
tell-HAB-PST-PL king eye dark be-DECL say-PST-PL
'They told him "The King is blind" they said.'

b. *nepiri-pa-či-dá rey táu pʰārîu bi*
tell-HAB-PST-PL king eye dark be
'They told him The King is blind they said.'

In the case of direct report (14a), there is a declarative suffix *-dá*, while there is no illocutionary marking in the case of indirect report (14b). This difference reflects the presence of a tropic in direct reported speech and the absence of a tropic in indirect reported speech.

While Danish and Epena Pedee distinguish systematically between direct and indirect report, other languages do not make a clear distinctions or do not make a distinction at all. Dunn (1999: 91), for example, notes that Chuckchi (Chukotko-Kamchatkan) "does not have any mechanism for marking indirect speech". And Derbyshire (1979: 22) notes that in Hixkaryana (Cariban) the only means of conveying indirect speech is by means of a 'hearsay' particle (and not a complement construction). In some languages, complement clauses are ambiguous and can be used for both direct and indirect report. An example is found in Pirahã.

(15) Pirahã (Mura; Everett 1986: 269)
hi gái-sai xahóápátí ti xi aagá-hóág-a
3SG say-NMLZ Xahóápátí 1SG hunger have-INGR-REM
"'Xahóápátí said: "I am hungry"'. / 'Xahóápátí said that I am hungry.'

Based on constructions like that in (15), it has been argued that direct utterance report does not involve complementation . However, Cristofaro (2003, 2013) has argued that subordination (including complementation) should be defined semantically as a situation where one state-of-affairs is connected to another state-of-affairs "such that one of them (the main one) entails that another one (the dependent one) is referred to" (Cristofaro 2003: 95). Accordingly, she maintains that direct and indirect utterance reports by means

of utterance-verb constructions are semantically distinct types of complementation. In some languages – for instance, Danish and English – this semantic distinction is coded, while in others – such as Chuckchi, Hixkaryana and Pirahã – it is not.

6.7 Grammaticalization of utterance predicates

The theory we have outlined above entails a set of hypotheses regarding the grammaticalization of utterance predicates. Basically, it entails that utterance predicates that introduce direct report and utterance predicates that introduce indirect report have distinct grammaticalization paths.

It is well known that complement-taking predicates can undergo grammaticalization. According to Boye & Harder (2007, 2009, 2012), the grammaticalization pathway of a complement-taking predicate consists in a loss, in language usage, of discursive prominence relative to the complement clause, and a subsequent conventionalization of the complement-taking predicate as discursively secondary. (16) is an often-cited example of such a grammaticalization path from Afrikaans, where a complement-taking predicate *glo* has evolved into an epistemic particle. Discursively primary elements are in bold, while discursively secondary elements are not highlighted.

(16) Afrikaans (Indo-European; Thompson & Mulac 1991: 318, Boye 2012: 211)

a. ***ek*** ***glo*** *dat* *hy* *ryk* *is*
1SG think.PRS COMP 3SG.MASC rich be.PRS
'I think he is rich.'

b. *ik* *glo* ***dat*** ***hy*** ***rik*** ***is***
1SG think.PRS COMP 3SG.MASC rich be.PRS
'I think he is rich.'

c. ***hy*** ***is,*** *glo* *ek,* ***ryk***
3SG.MASC be.PRS think.PRS 1SG rich
'He is, I think, rich.'

d. ***hy*** ***is,*** *glo'k,* ***ryk***
3SG.MASC be.PRS 1SG.think rich
'He is, I think, rich.'

e. ***hy*** ***is*** *glo* ***ryk***
3SG.MASC be.PRS EPIST rich
'He is possibly/presumably/allegedly/seemingly rich.'

Similar kinds of grammaticalization are found for utterance predicates. We will focus on two known grammaticalization paths involving utterance predicates: the grammaticalization of assertive utterance predicates like 'say' to 1) grammatical markers of reportative evidentiality, and to 2) grammatical quotative markers (Heine & Kuteva 2002: 265-268). An example of the first type of development is found in Kannada (Dravidian), where a clitic *-ante*, which is a grammaticalized form of the utterance predicate *ennu*, can function as a 'hearsay'-marker', as in (14) (Boye 2010, 2012; Sridhar 1990: 1-5; for more examples, see Boye & Harder 2009).

(17) Kannada (Dravidian; Sridhar 1990: 4)
avara *magaLu* *i:ga:gale:mayneredidda:L****ante***
their.daughter now.already reach.puberty.N.PST.2SG.FEM.hearsay
'It is said that their daughter already went into puberty.'

An example of the second type of development is found in Khmu' (Austroasiatic), where the quotative marker *làw* derives from the utterance predicate *làw* 'say'. The grammatical function of *làw* as a quotative marker is illustrated in (18). Notice that in (18) there is also a non-grammaticalized utterance-predicate *hɛ́:t* 'call'.

(18) Khmu' (Austroasiatic; Premsrirat 1987: 73)
kɔ:ninim hɛ́:t làw tá:j *tèn tak à:ŋ*
girl call say brother sit in house
'The girl called: 'Brother, come here and sit down.'

Regarding these two grammaticalization paths, the theory we have outlined above entails that assertive utterance predicates can develop into gram-

matical markers of reportative evidentiality in constructions where they introduce indirect reported speech, while they can develop into grammatical quotative markers in constructions where they introduce direct reported speech.

According to the theory outlined above, dependent complement clauses that are used for indirect report of assertions only express propositions, while independent complement clauses that are used for direct report of assertions also express a tropic. Boye (2010, 2012: 35) has has argued that evidential markers have propositional scope, while quotation markers have illocutionary scope – or more precisely tropic scope. Based on this, the hypotheses we wish to put forward are simply that utterance predicates with propositional scope can grammaticalize into elements that also have propositional scope, while utterance predicate with tropic scope can grammaticalize into elements that also have tropic scope. This entails that the grammaticalization of utterance predicates retains part of the meaning of the clauses they co-occur with.

6.8 Conclusion

Based on a functional model of layered semantic structure, we have outlined a unified theory of clause types, the illocutions coded by different clause types and the utterance-predicate complement constructions that can be used to report such illocutions. We argued that the crucial difference between them, with regard to layering, is that independent main clauses, that are used to make independent utterances, have both a neustic and a tropic, while independent complement clauses, which are used for direct utterance report lack a neustic, and dependent complement clauses, which are used for indirect utterance report, lack both a neustic and a tropic (the tropic is expressed by the complement-taking predicate). What is common to all sentences that code, or report assertive and interrogative utterances, is that they contain both a proposition and a state-of-affairs. In contrast, clauses that code or are used to relay directive utterances only contain a state-of-affairs. Table 6.5 summarizes the complete analysis and can be read as a potentially universal semantic typology.

Table 6.5: The differences and similarities between independent utterances, direct report of utterances and indirect report of utterances (I^n = Neustic, I^t = Tropic, P = Proposition, S = State-of-affairs).

	Assertion	Polar question	Directive
Independent main clauses used for independent utterances	(1a) *She is leaving* I^n-I^t-P-S	(1b) *Is she leaving?* I^n-I^t-P-S	(1c) *Leave!* I^n-I^t–S
Independent complement clauses used for direct utterance report	(2a) *He said:* "*She is leaving*" -I^t-P-S	(2b) *He said:* "*Is she leaving?*" -I^t-P-S	(2c) *He said:* "*Leave!*" -I^t–S
Dependent complement clauses used for indirect utterance report	(3a) *He told* *(that) she is leaving* –P-S	(3b) *He asked* *if she is leaving* –P-S	(3c) *He told/asked* *her to leave* —-S

Chapter 7

Article II: Grammatical contrasts in utterance-predicate complementation: from iconicity to frequency, and back

7.1 Introduction

In cognitive and functional linguistics the canonical way of accounting for grammatical contrasts is in terms of iconicity (e.g. Haiman 1985, Langacker 1987, Croft 2003, Achard 2010. In recent years, however, frequency of use has been shown to be a major factor motivating such contrasts (Zipf 1935, Haiman 1985, Bybee & Hopper 2001), and accounts in terms of frequency have been gaining ground to the extent that iconicity accounts have been discarded (Haspelmath 2008).

This paper compares existing accounts of grammatical contrasts in terms of iconicity of cohesion and frequency of use. The basis for the comparison

is a study of grammatical contrasts in utterance-predicate complements. Existing studies of utterance-predicate complementation tend to focus on the distinction between so-called *direct speech* and *indirect speech* (e.g. Coulmas 1986, Jäger 2007, N. Evans 2012, Spronck 2012). In this paper we shift focus to contrasts between three kinds of utterance-predicate complements: assertive, interrogative and directive complements.

Assertive complements are complements that can be used to report assertions, as in (1) from Swahili.

(1) Swahili (Niger-Congo; Massamba 1986: 100)
maganga a-li-sema kwamba baba yake a-li-kuwa monjwa
Maganga he-PST-say COMP father his he-PST-be sick
'Mananga said that his father was sick.'

Interrogative complements are complements that can be used to report questions. In this paper, we deal exclusively with complements that can be used to report polar questions, as in (2) from Ndyuka.

(2) Ndyuka (Creole; Huttar & Huttar 1994: 81)
a dataa akisi mi ofu a gi mi kolusu
the.SG doctor/nurse ask 1SG COMP 3SG give 1SG fever
'The doctor asked me whether it was making me feverish.'

Directive complements are complements that can be used to report commands/requests, as in (3) from Barasano.

(3) Barasano (Tucanoan; Jones & Jones 1991: 28)
so-re budi roti-bĩ
3SG.FEM-OBJ exit order-3SG.MASC
'He told her to leave.'

The contrasts under scrutiny are of two kinds. The first kind has to do with the distinction between deranked and balanced complements (in the sense of Cristofaro 2003; see Section 7.2.2 below). In some languages, there is an asymmetry between different types of complements such that some

types are deranked whereas others are balanced. For instance, Garo has balanced assertive complements and deranked directive complements, as illustrated in (4a) and (4b) respectively

(4) Garo (Sino-Tibetan; Burling 2003: 330; 323)

a. *ang-a mi cha-jok ine na-a babulchi-na agan-bo*
I-NOM rice eat-PFV COMP you-NOM cook-DAT tell-IMP
'Tell the cook that I have eaten.'

b. *magipa bisako mi cha-kana (ine) agan-a*
mother child-ACC rice eat-SUB COMP tell-PRS
'The mother tells the child to eat rice.'

The second kind of contrasts concerns the distinction between presence and absence of complementizers. In some languages, there is an asymmetry between complement types such that some types have an obligatory or optional complementizer whereas others have, respectively, an optional complementizer or no complementizer at all. As an example, Finnish has assertive complements with complementizer *että* and directive complements without complementizers, as illustrated in (5a) and (5b) respectively.

(5) Finnish (Uralic; Sulkala & Karjalainen 1992: 30; 34)

a. *sairas valitti että kurkku on kipeä*
patient complain.3SG.IMPERF COMP throat be.3SG sore
'A patient complained that his throat was sore.'

b. *käskin hänen mennä*
command.1SG.IMPERF s/he.GEN go.INF
'I commanded him/her to go.'

We first present a crosslinguistic study which examines crosslinguistic tendencies pertaining to contrasts between these complement types. Subsequently, we examine to which extent these tendencies can be accounted for in terms of respectively iconicity of cohesion (e.g. Schüle 2000, Givón 2001, Cristofaro 2003, Noonan 2007) and frequency of use (Haspelmath 2008).

We argue – partly based on frequency data from supplementary corpus studies of Danish and English complement constructions – that a frequency account is inadequate, and that the account in terms of iconicity of cohesion is superior to it. However, we also point out limitations to the latter account. As an alternative to these two kinds of accounts we propose a new account based on the semantic distinction between propositions (truth-valued meaning units) and states-of-affairs (non-truth valued meaning units) (Vendler 1967, Ransom 1986, Hengeveld 1989, Dik & Hengeveld 1991, Horie 2000, Cristofaro 2003, Boye 2010, 2012). We argue that this account captures all crosslinguistic tendencies documented and that it enables a generalization, missed by the other accounts, over utterance-predicate complements and their corresponding main clauses.

The remainder of the paper is structured as follows: Section 7.2 presents the crosslinguistic study and the general results of it: 8 crosslinguistic tendencies pertaining to the grammatical marking of utterance-predicate complements. Sections 7.3 and 7.4 review the accounts of grammatical asymmetries in terms of respectively iconicity of cohesion and frequency, and critically examine these accounts in light of the crosslinguistic study and the supplementary corpus studies. Section 7.5 outlines the alternative account in terms of iconicity of complexity, and Section 6 compares the three types of accounts. Section 7.6 is a brief conclusion.

7.2 Grammatical contrasts in utterance-predicate complementation

In the following sections, we present a crosslinguistic study of utterance-predicate complementation, focused on two types of grammatical contrasts, one having to do with the distinction between deranking and balancing, the other one pertaining to the distribution of complementizers. The study covers all types of predicates that describe a speech act – that is, predicates with meanings such as ‘say’, ‘tell’, ‘ask’, ‘order’ and ‘write’ – and it covers complements that can be used to report assertions, questions and commands.

The study differs from earlier crosslinguistic studies of utterance complementation in two respects. Firstly, earlier studies have not investigated the

distribution of complementizers, at least not systematically. Secondly, the distribution of deranking and balancing has been studied only in assertive complements (see, for instance, Cristofaro 2003); interrogative and directive complements have been ignored. The emphasis in earlier studies on assertive complements is partly a consequence of the fact that prevailing functional typologies (Givón 2001, Cristofaro 2003, Noonan 2007, Cristofaro 2013) classify directive complements (like those in examples (1), (2) and (3) above) not as utterance-predicate complements, but as complements of "manipulative" predicates".

7.2.1 Crosslinguistic data

The crosslinguistic study presented in this paper is based on data from 84 genetically diverse languages. Our language sample initially consisted of 173 language constructed by adopting the 200-language sample in the World Atlas of Languages Structures Online (WALS) (Dryer & Haspelmath 2013) and removing the recommended 29 languages[13] to create a more genealogically balanced sample. Due to lack of reliable information, we additionally removed 89 languages from the sample and arrived at a final number of 84 languages. The high number of languages with insufficient information reflects the fact that many grammars do not treat complementation exhaustively, and the fact that in order to be as accurate as possible we wanted to study only glossed examples. The 84 languages from which we obtained information are listed in the appendix [Appendix B]. The data were primarily collected from reference grammars, but in a few cases example sentences were provided by language experts.

7.2.2 Balancing vs. deranking

We will first discuss the contrast that has to do with the distinction between deranking and balancing. As originally defined by Stassen (1985: 76-83), balanced complements have verbs that are marked for the same TAM-distinctions as verbs in independent clauses, whereas deranked complements have verbs that are marked for less TAM-categories than verbs in independent clauses.

[13] The so-called 200-language sample consists of 202 languages at present.

Cristofaro (2003: 54-55) later expanded the notion of deranking to include complements with subjunctives and other "dependent moods", an expansion which we will adopt in the present study (This distinction is similar to the one between dependent and independent verb forms proposed by Haspelmath 1995). Example (5) above illustrated the contrast between balanced and deranked complements (in addition to a complementizer contrast) in the context of utterance-predicate complementation. The complement verb in (5a) can be described as balanced, since it has the TAM-marking of a declarative main clause, whereas the complement verb in (5b) can be described as deranked, since the verb is an infinitive without the TAM-marking of a declarative main clause verb.

According to the definition of balancing and deranking proposed by Cristofaro (2003: 54-55) deranked verbs may also be hortatives or any other type of dependent mood, as in Kayardild, which contrasts indicative (balanced) complements and hortative (deranked) complements as in (6a) and (6b), respectively.

(6) Kayardild (Tangkic; N. Evans 1995: 516; 507)

a. *ngada* *kamburi-ja* *niwan-ji* *walbu-ntha* *dathin-inja*
1SG.NOM say-ACT 3SG-MLOC raft-COBL that-COBL
barji-nyarra-nth
capsize-APPR-COBL
'I told him the raft would capsize.'

b. *kamburi-ja* *dathin-a* *dangka-a* *warra-nanginj*
speak-IMP that-NOM man-NOM go-NEG.HORT
'Tell that man not to go.'

Cristofaro (2013) surveys the distribution of deranking and balancing for complements of utterance predicates. According to her survey, there are far more languages that exclusively have balanced utterance-predicate complements than there are languages that exclusively have deranked utterance-predicate complements. However, the survey is based on a definition of complement-taking utterance predicates as predicates that take propositional complements (see Section 7.5 on the contrast between propositions and states-

of-affairs), and consequently only assertive complements are taken into account, as mentioned earlier. In this section we will look at which patterns exist for all major complement types of utterance predicates: assertive complements, interrogative complements and directive complements.

7.2.2.1 Unrestricted tendencies

We first look for possible unrestricted tendencies, that is, tendencies pertaining to each of the complement types in isolation. In the case of assertive complements, our findings are in line with the results of the study in Cristofaro (2013). As shown in Table 7.1, 62 languages in our sample have only a balanced assertive complement, 2 languages have both a balanced and a deranked assertive complement, and 9 languages have only a deranked complement.

Table 7.1: Deranking-balancing contrasts – assertive complements

	Number of languages
Assertive balanced	62
Assertive balanced/deranked	2
Assertive deranked	9
Total	73

As for interrogative utterance complements, a similar pattern emerges (despite a lower number of data points), as shown in Table 7.2. Most languages have balanced interrogative complements, while only 1 language has both balanced and deranked constructions, and 3 have only a deranked construction.

Table 7.2: Deranking-balancing contrasts – interrogative complements

Construction	Number of languages
Interrogative	35
Interrogative balanced/deranked	1
Interrogative deranked	3
Total	39

In contrast, directive complements tend to be deranked. As shown in

Table 7.3, 14 languages in the sample have a balanced directive complement while 5 have a balanced as well as a deranked complement, and 39 have a deranked construction only.

Table 7.3: Deranking-balancing contrasts – directive complements

Construction	Number of languages
Directive balanced	14
Directive balanced/deranked	5
Directive deranked	39
Total	58

At this point it should be noted that while the balancing-deranking contrast is intended as a means for comparing main clauses with subordinate ones, in the case of directive complements (and to a lesser extent interrogative complements), it cannot unproblematically be used this way. The problem is that deranking is defined relative to declarative clauses, but that directive (and interrogative) complements do not only differ from declarative clauses in terms of subordination: declarative clauses arguably code assertions, whereas directive complements report commands/requests (and interrogative complements polar questions). Thus, the tendency for directive complements to be deranked may be simply an artifact of the definition of deranking offered by Cristofaro (2003: 57): it may reflect the fact that directive complements report another kind of illocution than that coded by declarative clauses. To avoid this problem, we can instead compare complements reporting assertions, questions and commands to their corresponding main clauses. For example, if a directive complement retains the TAM-marking of an imperative main clause, we will call it balanced; if not, we will call it deranked.

Some languages have the same imperative verb form in both independent imperatives and in directive complements. In these cases, the complements would be deranked on Cristofaro's definition, but on our revised definition it is balanced. Consider for instance the following example from Ingush.

(7) Ingush (Nakh-Daghestanian; Nichols 2011: 542)
Waishietaz Muusaaiga pwieghazh jyla ealar
Aisha.ERG Musha.all dish.PL J.wash.IMP say.WP
'Aisha told Musha to wash the dishes.'

The complement verb in (7) is an imperative form which is naturally not found in declaratives. Yet, the exact same verb form is used in main clause imperatives, as in (8).

(8) Ingush (Nakh-Daghestanian; Nichols 2011: 107)
juxa aala
back say.IMP
'Say it again.'

The crosslinguistic frequencies with which utterance-predicate complements are deranked on our revised definition, differ from (and are smaller than) the frequencies with which they are deranked in the sense of Cristofaro (2003). In the sample there are 7 languages which appear to use an imperative verb form in a directive complement (Amele, Hungarian, Ingush, Lezgian, Menya, Paiwan and Ungarinjin). This increases the ratio of balanced to deranked directive complements relative to the ratio of balanced to deranked complements (cf. Table 7.3 above and Table 7.4 below). However, also based on our revised definition, there is a tendency for directive complements to be deranked.

Table 7.4: Deranking-balancing contrasts – directive complements, when imperative verb form is considered to be balanced

Construction	Number of languages
Directive balanced	21
Directive balanced/deranked	5
Directive deranked	32
Total	58

To sum up our results for complement types considered in isolation, the following unrestricted tendencies can be formulated:

(9) Unrestricted tendencies

a. There is a strong tendency for assertive complements to be balanced.

b. There is a strong tendency for interrogative complements to be balanced.

c. There is a weak tendency for directive complements to be deranked

7.2.2.2 Restricted tendencies

We now take a closer look at languages for which we have data on more than one type of complement. Doing so allows us to get a precise idea of the extent to which contrasts between different types of complements are expressed in terms of balancing vs. deranking. For briefness and clarity, and because we have relatively little data on contrasts involving interrogative complements, we will only consider contrasts between assertive and directive complements (this applies to the study of complementizer contrasts in Section 7.2.3 as well). For both these types of complements, we distinguished three subtypes: i) only balanced, ii) either balanced or deranked; iii) only deranked. This means that there are 9 different possible relations between assertive and directive complements. These 9 possibilities are depicted in Table 7.5, and for each possible relation, it is marked in how many languages it was attested:

Table 7.5: Comparison of deranking and balancing in assertive and directive complements

	Directive B	Directive B/D	Directive D
Assertive B	12	1	31
Assertive B/D	0	1	0
Assertive D	0	0	4

The most frequent situation by far is that the assertive complement is balanced and the directive complement is deranked (see e.g. (6) above for an example). This pattern is attested in 31 languages out of 49 languages. More importantly, there are no languages in the sample in which assertive

complements are deranked while the directive complements are balanced. In other words, our data supports the following restricted, unidirectional implicational tendency:

(10) In languages with a contrast between deranked and balanced complement verbs, there is a strong tendency that if the directive complement is balanced, the contrasting assertive complement is balanced as well.

7.2.3 Complementizer contrasts

We will now turn to the grammatical contrasts that have to do with the distribution of complementizers. Since the presence of a complementizer adds additional morphological material to a complement, complements introduced by a complementizer are morphologically more complex than complements not introduced by a complementizer. This type of contrasts has not to our knowledge been studied in relation to utterance-predicate complementation, and in general it has received less attention than the deranking-balancing contrasts described in Section 7.2.2.[14].

In parallel to Section 7.2.2 on balancing and deranking, we will first look at tendencies pertaining to complementizer marking for assertive, interrogative and directive complements individually, before proceeding to complementizer contrasts between assertive and directive complements.

When identifying complementizers in our materials, we basically adhered to Noonan's (2007: 55) definition according to which a complementizer is "a word, particle, clitic or affix, one of whose functions it is to identify the entity as a complement". As discussed in Boye et al. (2015: 3) and Kehayov & Boye (2016b: 9-10), this definition does not entail that it is a straightforward task to identify complementizers. One among several problems is that it may be hard to distinguish the function of identifying a complement from the related functions of nominalization and adjectivization, and views may differ as to what is the most accurate description of a given function. For instance, Modern Greek has *na* as in (11a), which is considered a

[14] Givón (2001) has a brief description of subordinating morphemes in relation to semantic integration. Cf. also Boye (2010) on complementizer contrasts in perception-predicate complementation and Boye & Kehavov (2016) for recent crosslinguistic and language-specific studies of complementizers in European Languages

complementizer by Dixon (2010: 391), but a subjunctive particle by Joseph & Philippaki-Warburton (1987: 23) – as opposed to a complementizer, such as *óti* in (11b).

(11) Modern Greek (Indo-European; Joseph & Philippaki-Warburton 1987: 22; 23)

a. *tu ípa na érθi amésos*
him.GEN told.1SG PRT come.3SG immediately
'I told him to come at once.'

b. *o jánis ípe óti θa voiθísi*
the John.NOM said.3SG COMP FUT help.3SG
'John said that he will help.'

Another problem is that the definition of complementizers depends on the definition of complements. A contested issue is whether purpose clauses are complements or adverbial clauses and thus whether purposive markers are complementizers or adverbial subordinators. To minimize these problems, we decided to include only free-standing items forms (as in the case of bound forms, we were sometimes in doubt of whether they actually had a complementizing function), and to exclude items that are clearly purposive markers. For instance, we excluded the Martuthunira (Australian) "purposive marker" *–waa*, which attaches to the complement verb in constructions such as the one in (12). Though described as a purposive marker the use of the item in (12) could very well be complementizing. However, we excluded it because it is not a free-standing form.

(12) Martuthunira (Australian; Dench 1995: 224)
ngayu wanka-lha pawulu-u manku-waa nganaju-u
1SG.NOM tell-PST child-ACC get-PURP=O 1SG.GEN-ACC
ngamari-i
tobacco-ACC
'I told the child to get my tobacco.'

7.2.3.1 Unrestricted tendencies

As shown in Table 7.6, there is a strong tendency for assertive complements to be marked by a complementizer – this is the case for 40 languages out of 47.

Table 7.6: Complementizer distribution – assertive complements

Construction	Number of languages
Assertive COMP	40
Assertive COMP/no COMP	0
Assertive no COMP	7
Total	47

There is less information on interrogative complements in our data, but the tendency for them to be marked by a complementizer is nevertheless clear. As shown in Table 7.7, in 15 out of 20 languages the interrogative complement is introduced by a complementizer.

Table 7.7: Complementizer distribution – interrogative complements

Construction	Number of languages
Interrogative COMP	15
Interrogative COMP/no COMP	1
Interrogative no COMP	4
Total	20

Like assertive and interrogative complements, directive complements are also frequently marked by a complementizer, but the tendency is not as strong as for assertive and interrogative complements. As seen in Table 7.8, 16 languages out of 35 have directive complements introduced by a complementizer and an additional 7 languages have directive complements optionally introduced by a complementizer.

To sum up our results for complement types considered in isolation, the following unrestricted tendencies can be formulated based on our results:

(13) a. There is a strong tendency for assertive complements to be marked by complementizers.

Table 7.8: Complementizer distribution – directive complements

Construction	Number of languages
Directive COMP	16
Directive COMP/no COMP	7
Directive no COMP	12
Total	35

b. There is a strong tendency for interrogative complements to be marked by complementizers.

c. There is a tendency – though only remote – for directive complements to be marked by complementizers.

7.2.3.2 Restricted tendencies

We now take a closer look at languages for which we have data on complementizer marking for more than one type of complement. Such languages allow us to get a precise idea of the extent to which contrasts between complement types are marked in terms of complementizer distribution. For briefness and clarity, and because we have relatively little data on contrasts involving interrogative complement, we will only consider contrasts between assertive and directive complements. There are in principle 9 possible contrast patterns. For both these types of complements, we distinguished three subtypes: i) obligatory complementizer ("COMP"), ii) optional complementizer ("COMP/no COMP"); iii) no complementizer ("no COMP"). These 9 possibilities are depicted in Table 7.9, and for each possible relation, it is indicated in how many languages it is attested:

Table 7.9: Comparison of complementizer distribution in assertive and directive complements

	Directive COMP	Directive COMP/ no COMP	Directive no COMP
Assertive COMP	15	7	8
Assertive COMP/no COMP	0	0	0
Assertive no COMP	1	0	4

As can be seen in Table 7.9, only five of the possible patterns were attested. Three patterns account for the majority of contrasts attested. As the most frequent pattern, in 15 out of 35 languages the assertive and the directive complement are both marked by a complementizer.

An example of this pattern is the contrast in (14) from Khmer. Assertive complements in Khmer are marked by the complementizer *tha:* as in (14a) and directive complements are marked by *aoj* as in (14b).

(14) Khmer (Austro-Asiatic; Haiman 2011: 227; 289)

a. *kom tae lo:k prap tha: via cia pseup pul kom*
PROH only you tell COMP 3 be mushroom poison PROH
ej knjom mwn deung
what I not know
'If you hadn't told it was a poisonous mushroom, I wouldn't have known.'

b. *knjom trey ba:n kee: prap aoj te sala:rian*
I PASS PST 3 tell COMP go school
'I was told to go to school'

As the second most frequent pattern, in 8 languages out of 35 the assertive complement is marked by a complementizer while the directive complement is not, and as the third most frequent pattern, in 7 languages the assertive complement is marked by a complementizer while directive complements are only optionally marked by a complementizer. Examples of these patterns are found in Hindi (15) and Georgian (16), respectively. In Hindi, the assertive complement is marked by the complementizer *ki* (15a), whereas the directive complement is not marked by a complementizer (15b).

(15) Hindi (Indo-European; Kachru 2006: 135)

a. *rətən ne kəha ki sureʃ səb se naraz*
Ratan AG say.PFV.MASC.SG COMP Suresh all.OBL with angry
hɛ
be.PRS.SG
'Ratan said that Suresh is angry with everyone.'

b. *usne* *bʰaĩse* *dehradũn* *me* *məkan* *bənvane*
s(he).AG brother.with Dehradoon.OBL in house make.CAUS.INF.OBL
kĩ *bat* *kə* *rəkʰĩ* *hɛ*
of.FEM matter.FEM.SG say keep.PFV.FEM.SG PRES.SG
'S(he) has told her brother to construct a house in Dehradoon (for herself).'

In Georgian, assertive and directive complements can be marked by the same complementizer *rom*, but whereas assertive complements as in (16a) must be introduced by *rom*, in directive complements as in (16b) it is optional.

(16) Georgian (Kartvelian; Hewitt 1995: 614; Hewitt & Crisp 1986: 124)

a. *iza-m* *(ø-)gv-i-txr-a* *rom* *meore* *dġe-s*
iza-ERG (it)us-OV-tell-she(AOR) COMP second day-DAT
mo-vid-od-a, *da* *xom* *mo-vid-a*
PREV-come-IMPERF-she(COND) and surely PREV-come-she(AOR)
'Iza told us she would come the next day, and she came, didn't she?'

b. *brʒana,* *(rom)* *kučebši* *mxiaruloba* *šec'q'des*
he.ordered.it (COMP) in.the.streets merry.making it.should.stop
šemc'q'dariq'o
it.had.stopped
'He ordered the merry-making in the streets to stop.'

The only language that does not follow the tendency is Taba, which according to Bowden (1997: 454-455) uses the resultative marker *de* 'so that/in order that' as a complementizer in directive utterance-predicate complements, as in (17b). Assertive complements on the other hand are not marked by a complementizers, as shown in (17a).

(17) Taba (Austronesian; Bowden 1997: 455; 442)

a. *n=sul-ak wang=si de l=mul ak-le*
3SG=order-APPL child=PL RES 3PL=return all-land
'He told the children to go home.'

b. *n=ha-lusa n=han*
3SG=CAUS-say 1SG=go
'He said that he was going.'

To sum up our findings for complement contrasts, they point to the restricted, unidirectional implicational tendency in (18):

(18) In languages with a complementizer contrast, there is a strong tendency that if the directive complement is marked by a complementizer, the contrasting assertive complement is marked by a complementizer as well.

7.2.4 Comparison

To some extent, our findings for the distribution of balanced and deranked complements and our findings for complementizer distribution are parallel. The restricted tendency for directive complements to be balanced only if contrasting assertive complements are also balanced (10) is parallel to the likewise restricted tendency for directive complements to be marked by a complementizer only if contrasting assertive complements are also marked by a complementizer (18). Similarly, the unrestricted tendencies for assertive and interrogative complements to be balanced (9a, 9b) are parallel to the unrestricted tendencies for these types of complements to be marked by a complementizer (13a, 13b).[15]

The parallelism is not complete, however. Directive complements tend to be deranked, but contrary to what might be expected based on the general picture, they also tend to be marked by complementizers.

[15]This raises the question to which extent balancing-deranking contrasts combine with complementizer contrasts in language-specific constructions. This question, while interesting, is beyond the scope of this paper.

7.3 Accounts in terms of iconicity of cohesion

In cognitive-functional linguistics, the standard account of the grammatical asymmetries discussed in Section 7.2 is in terms of iconicity of cohesion, also known as, for instance, "binding", "semantic integration", and "event integration" (cf. Haspelmath 2008). Iconicity of cohesion covers the idea that a 'strong syntactic integration" of two clauses into one complex clause reflects and is motivated by a strong "semantic bond" between events (Givón 2001). This idea has been used primarily to account for the distribution of deranked and balanced complements: deranked complements are more syntactically integrated in the complement-taking clause than balanced complements, in that the former lack properties of independent clauses; this is seen as reflecting and as being motivated by a higher degree of semantic integration of deranked complements (Givón 1980, Verspoor 2000, Cristofaro 2003, Noonan 2007). To a lesser extent the idea has been used to account for the distribution of complementizers: complements without complementizers can be seen as more syntactically integrated in the complement-taking clause than complements with complements, in that complementizers mark a clause boundary; as in the case of balancing and deranking, this is seen as reflecting and as being motivated by a higher degree of semantic integration of deranked complements (Givón 2001).

In the case of utterance-predicate complementation, Givón (2001) suggests that assertive complements are semantically rather disintegrated, while directive ones are relatively integrated (see also Cristofaro 2003, 2013, Noonan 2007). Dealing with directive complements as a subtype of manipulative complements, he locates them at the top end of a scale of semantic integration, and assertive complements at the bottom end (he does not take interrogative complements into consideration). This raises the question of how more precisely directive complements are semantically more integrated than assertive and interrogative ones. Givón does not address this question specifically, but proposes several dimensions of integration that may be seen as candidate answers (Givón 2001: 44ff). Some of these dimensions are not relevant for distinguishing utterance-predicate complements. For instance, the dimension of "successful vs. intended manipulation" and the dimensions

of "degree of success", "agentive control", "manipulee resistance", and "directness of the manipulation" apply only to directive and (other kinds of) manipulative complements (Givón 2001: 44–45; 47–49). Other dimensions are in principle relevant, but the different types of utterance-predicate complements do not differ in terms of them. This goes, for instance, for "degree of intentionality": directives, assertions and polar questions are all intentional speech acts, and complement-taking predicates such as 'say', 'ask' and 'tell to' thus do not differ in terms of degree of intentionality.

Cristofaro (2013) suggests an answer to the question raised above related to what Givón (2001: 44) calls temporal integration. According to Cristofaro (2013), assertive utterance predicates (simply "utterance predicates" in her terms) "do not involve predetermination of the time reference or of the participants of the dependent event". In contrast, directive utterance predicates (grouped under manipulative predicates) have been described as having determined time reference by e.g. Noonan (2007: 103).

Whereas assertions and polar questions can concern events that take place later than, simultaneous with, or prior to, the act of asserting or questioning, directive speech acts necessarily concern events that take place later than the act of posing them. This would then motivate a tendency for directive complements to be deranked and at least lack Tense specification, and a tendency for assertive (and interrogative) complements to be balanced; the time location of the event described in a directive complement follows from the time specification of the complement-taking clause reporting the directive speech act. In contrast, the time location of the event described in an assertive or interrogative complement is not determined by the time specification of the complement-taking clause.

Such an analysis presents two limitations, however. Firstly, the account in terms of predetermination of time is restricted to languages that have Tense-marking, and there are languages without Tense-marking in which the contrast between assertive and directive complements is still marked in a contrast between balanced and deranked verbs. Secondly, tendencies pertaining to complementizers cannot straightforwardly be accounted for in terms of determined time-reference, as complementizers are – unlike balancing vs. deranking – not straightforwardly related to Tense considerations. In our

sample, there are languages without Tense-marking in which the contrast between assertive and directive complements is still marked in a contrast between presence and the (optional) absence of a complementizer. Persian is a case in point. According to Mahootian (2005), both the assertive complement in (19a) and the directive complement in (19b) can be marked by the complementizer *ke*, but the complementizer is only optional in the directive complement (as indicated by the parentheses in example (19b).

(19) Persian (Indo-European; Mahootian 2005: 94; 32)

a. *mænsur goft ke xod-es m-i-re*
Mansur said COMP self-3SG.PC DUR-go-3SG
'Mansur said that he will go himself.'

b. *be to goft (ke) be-r-i*
to you said (COMP) SBJN-go-2SG
'She told you to go.'

Another candidate for an understanding of how exactly directive complements are more semantically integrated than assertive ones is what Givón (2001) refers to as "referential integration": "The more two events share their referents, the more likely they are to be construed as a single event" (Givón 2001: 50). Assertive complement events clearly do not necessarily share referents with the utterance events described by their main clauses. It is easy to conceive of a report of an assertion made by somebody about an event involving somebody else. Accordingly, Givón considers (assertive) utterance complements the least semantically integrated complement type. In contrast, directive complement events standardly share at least one referent with the utterance events described by their main clauses. With the exception of by proxy directives, directives entail that the addressee of the utterance event is involved in the event to be carried out.

We assume with Givón that assertive and interrogative complements are referentially rather disintegrated, whereas directive complements are referentially relatively integrated. Based on this assumption, we discuss below to which extent the findings presented in Section 7.2 can be accounted for in terms of iconicity of cohesion.

We first deal with the distribution of balancing and deranking in Section 7.2.2 and then with complementizer distribution in Section 7.2.3.

7.3.1 Iconicity of cohesion and balancing vs. deranking

As discussed in Section 7.2, the crosslinguistic tendencies pertaining to the distinction between balancing and deranking point in the same direction: assertive and interrogative complements tend to be balanced, while directive ones tend to be deranked. Since balancing reflects lack of syntactic integration, while deranking reflects integration, this amounts to saying that assertive and interrogative complements tend to be less syntactically integrated than directive ones.

This can be straightforwardly accounted for in terms of iconicity of cohesion, in so far as assertive and interrogative complements are also semantically (specifically "referentially") less integrated than directive ones.

7.3.2 Iconicity of cohesion and complementizer marking

Some of the tendencies pertaining to complementizer distribution can be accounted for in terms of iconicity of cohesion. If the presence of complementizers reflects syntactic disintegration, and the absence of complementizers reflects syntactic integration, we would expect the semantically disintegrated assertive and interrogative complements to have complementizers and the semantically integrated directive ones to lack complementizers.

This is to some extent what we find: assertive and interrogative complements tend to have complementizers (cf. tendencies 13a and 13b), and directive complements tend to be marked by a complementizer only if contrasting assertive complements are also marked by a complementizer (cf. tendency 18).

However, the unrestricted tendency for directive complements to have complementizers runs counter to expectations. If iconicity of cohesion were at work here, there ought to be a tendency for these semantically integrated complements not to be syntactically disintegrated by means of a complementizer. The tendency in (13c) thus presents a serious problem for iconicity of cohesion as a general account of grammatical asymmetries in utterance-

predicate complements. It is therefore natural to search for alternative accounts.

7.4 An account in terms of frequency

One candidate for an alternative to iconicity of cohesion is frequency of use. Frequency of use has been shown to be strong factor motivating grammatical asymmetries (Zipf 1935, Bybee & Hopper 2001, Haspelmath 2008), and the notion plays an important role in most cognitive and functional approaches to language. An often-cited example of asymmetries that are shaped by frequency differences is the contrast between singular and plural nouns. There is a crosslinguistic tendency for plurals to be expressed by more morphosyntactic material than singulars. Many scholars, including Jakobson (1965 [1971]), have analyzed this tendency as being motivated in terms of iconicity of quantity: "Greater quantities in meaning are expressed by greater quantities of form" (Haspelmath 2008: 2). But as argued by e.g. Haspelmath (2008: 5), this analysis is wrong; the tendency follows straightforwardly from the fact that singulars tend to be more frequent than plurals. Haspelmath (2008) suggests a similar story for some cases of complement asymmetries: they were originally understood in terms of iconicity of cohesion, but this understanding can be straightforwardly replaced, he argues, by an explanation in terms of frequency differences. The part of his argumentation which is relevant here is centered on the following examples, adapted from Givón's (2001) discussion of "event integration and clause union".

(20) Givón (2001: 48); Haspelmath (2008: 18)

a. *She told him to leave.*

b. *She insisted that he must leave.*

Both *tell* and *insist* are utterance predicates on our definition, and both the complement in (20a) and that in (20b) are thus utterance-predicate complements. The contrast between them clearly displays a balanced-deranked contrast of the sort discussed in Section 7.2.2: the complement of *insist* has a balanced, tensed verb, whereas the complement of *tell* has only a root or infinitival verb (as is well-known, English does not distinguish infinitives

from verbal roots). On the analysis of *to* as an infinitival marker rather than a complementizer, (20) also displays a complementizer asymmetry of the sort discussed in Section 7.2.3: the complement of *insist* has a complementizer, whereas the complement of *tell* does not.

On the face of it, (20) might be seen as a contrast between two directive utterance predicates and associated directive complements: in one reading of (20b), it reports a directive speech act (she to him: leave!), just like (20a). However, while the construction in (20a) is exclusively used to report directive speech acts, the construction in (20b) is standardly used for reporting (emphatic) assertions, as illustrated in (21)–(23).

(21) a. *Two plus two equals four.* (Assertion)

b. *She insisted (that) two plus two equals four.* (Report of assertion)

(22) a. *The butler did it.* (Assertion)

b. *She insisted (that) the butler did it.* (Report of assertion)

(23) a. *He must leave.* (Assertion)

b. *She insisted (that) he must leave.* (Report of assertion)

What makes the complements in (20b) and (23b) special is that they include a deontic modal *must*. Because a non-epistemic modal designates social force dynamics, it can easily be used to report directives in addition to assertions. This means that the directive potential resides in the modal, not in the complement as such. (Note that *insist* plus *that*-clause can also be used to report directive speech acts if it involves a mandative subjunctive, as in *she insisted that he be here by ten*, but also that mandative subjunctive is clearly a deranked verb form). Thus, as far as complement types are concerned, Haspelmath's example involves a contrast between directive and assertive complements similar to the ones studied in this paper.

The centerpiece of Haspelmath's account of asymmetries like that in (20) (and a couple of other non-utterance complement pairs that he discusses) is that,

> "[t]here are also obvious frequency asymmetries between the

> pairs (…) which suffice to explain the shorter coding of the first member of each pair" (Haspelmath 2008: 24-25).

Before this claim can be evaluated properly, it must be noted that by "shorter coding" Haspelmath does not refer to the fact that the complement in (20a) lacks a modal and a subject, but to its deranked status and the fact that it lacks a complementizer (on the analysis that *to* is not a complementizer) (Haspelmath 2008: 18). Thus, Haspelmath's claim is that exactly the two kinds of grammatical contrasts under scrutiny here can be accounted for in terms of frequency of use.

Haspelmath (2008) recognizes that there is a semantic difference between (20a) and (20b). In fact, he argues that the morphosyntactic differences between perception complements like (24a) and (24b) cannot be explained in terms of frequency, but that an account in terms of semantics and iconicity is required.

(24) a. *She saw him coming out of the theater.*

b. *She saw that he had left two hours earlier.*

But although the contrast in (24) is morphosyntactically and arguably also semantically similar to the utterance complement contrasts discussed above (cf. Boye 2012: 188-191), he deems the semantic difference irrelevant for utterance complement asymmetries, and places the burden of proof on those who wish to argue that this difference is what shapes the morphosyntactic asymmetry rather than on those who wish to argue for a frequency account:

> "Givón is right that in each case there is also a semantic contrast, but in order to show that the semantic contrast is indeed responsible for the formal contrast, he should provide contrasting examples of constructions with roughly equal frequency" (Haspelmath 2008: 25)

In what follows we attempt to lift this burden of proof. We do not dispute that the construction in (20a) is more frequent than that in (20b). Tell is semantically more general than insist, and therefore convenient in more

contexts. In order to get an idea of the role of frequency, however, general or abstract utterance predicates like *tell* should not be contrasted with semantically narrow predicates like *insist*, but with other abstract predicates. In what follows we therefore shift attention to the contrasts in (25)–(27) from English and Danish.

(25) English

a. *She told him (that) she had left.* (Assertive complement)

b. *She told him to leave.* (Directive complement)

(26) a. *She asked him if he had left.* (Assertive complement)

b. *She asked him to leave.* (Directive complement)

(27) Danish

a. Assertive complement

Hun	*sagde*	*(at)*	*han*	*gik*
3SG.FEM.NOM	say.PST	COMP	3SG.MASC.NOM	leave.PST

'She said that he left.'

b. Interrogative complement

Hun	*spurgte*	*om*	*han*	*gik*
3SG.FEM.NOM	ask.PST	COMP	3SG.MASC.NOM	leave.pst

'She asked whether he left.'

c. Directive complement

Hun	*bad*	*ham*	*gå.*
3SG.FEM.NOM	request.PST	3SG.MASC.OBL	leave.INF

'She told him to leave.'

Examples (25a) and (27a) are balanced, have an optional complementizer and report assertive speech acts. Examples (26a) and (27b) are likewise balanced, but have an obligatory complementizer and report polar questions. Examples (25b), (26b) and (27c) are deranked, lack a complementizer and report directive speech acts. In the English examples (25) and (26), complements with the same abstract complement-taking predicates, *tell* and *ask*, respectively, are contrasted. These contrasts thus differ from the contrast

in (20), given by Haspelmath (2008) in being more minimal. In the Danish examples (27), complement contrasts are associated with contrasts between different abstract predicates *sige* 'say', *spørge* 'ask' and *bede* 'request'.

Haspelmath's proposal entails that deranked complements without complementizers are more frequent than balanced complements with (optional) complementizers. In order to test whether this is the case, we conducted two corpus studies of utterance-predicate complement constructions in English and Danish.

For English, we extracted sentences from the British National Corpus British National Corpus[16] and for Danish we extracted sentences from the LANCHART corpus. In order to get a reasonable amount of data, we restricted the search to only include the past tense form of the complement-taking predicate and to include only subjects and indirect objects that are 3rd person singular personal pronouns (*he*, *him* and *she*, *her* for English; *han* 'he', *ham* 'him', *hun* 'she', *hende* 'her' for Danish).

The results of the survey are given in tables 7.10 (English) and 7.11 (Danish).

Table 7.10: Distribution of balanced and deranked complements in English

Construction	COMP	no COMP	Total
tell + balanced	59	52	111
tell + deranked	0	31	31
ask + balanced	26	0	26
ask + deranked	0	43	43

Table 7.11: Distribution of balanced and deranked complements in Danish

Construction	COMP	no COMP	Total
sige + balanced	94	123	217
spørge + balanced	17	0	17
bede + deranked	0	0	0

For the *ask*-contrast, the figures point in Haspelmath's direction, but for the other contrasts, they point in the opposite direction. This means that

[16]The British National Corpus, version3(BNCXMLEdition). 2007. Distributed by Bodleian Libraries, University of Oxford, on behalf of the BNC Consortium.
URL: http://www.natcorp.ox.ac.uk/

not only is Haspelmath's specific proposal falsified by our findings, there is in fact no systematic relationship between frequency on the one hand and balancing-deranking contrasts complementizer contrasts on the other. These findings suggest that grammatical contrasts in utterance complements cannot be explained in terms of frequency. While studies of frequency patterns provide plausible explanations of a number of morphological contrasts that were previously thought to be motivated by iconicity (or simply described as 'markedness' contrasts), frequency of use does not seem to capture the crosslinguistic variation in utterance-predicate complementation. Our findings thus support Haiman (2008), who in his response to Haspelmath (2008), argues that frequency is not enough to explain all asymmetrical grammatical phenomena and that in certain contexts a longer form might be used, even though it is the most frequent.

7.5 An account in terms of conceptual semantics and iconicity of complexity

Below, we present an alternative to the accounts in terms of frequency and iconicity of cohesion discussed in sections 7.3 and 7.4 above. We argue that the crosslinguistic tendencies documented in Section 7.2 can be accounted for in terms of the conceptual semantics of the different complement types and in terms of iconicity of complexity.

Central to our proposal is the distinction between propositions – also sometimes referred to as "third-order entities" (Lyons 1977) or "facts" (Lees 1960, Vendler 1967, Dixon 2006) – and states-of-affairs – also known as "second-order entities" (Lyons 1977) or "events" (Vendler 1967); see Boye 2010 for an overview. This distinction can be thought of as pertaining to meanings prototypically expressed by clauses. According to standard conceptions of the distinction, propositions are truth-valued; they can thus be said to be true or false, and they can be evaluated epistemically. In contrast, states-of-affairs are not truth-valued; they cannot be said to be true or false, and they cannot be epistemically modified, but they can be said 'to occur' or 'take place' (e.g. Vendler 1967: 174, Lyons 1977: 443, Dik 1997: 51).

It is well-established that complement contrasts can be analyzed seman-

tically in terms of this distinction (e.g. Vendler 1967, Hengeveld 1989, Dik & Hengeveld 1991, Horie 2000, Cristofaro 2003, Boye 2010, 2012). For instance, knowledge verbs that express epistemic (or "declarative") knowledge take propositional complements, cf. example (28a), whereas knowledge verbs that express "knowing how" (or "procedural" knowledge") take state-of-affairs complements, cf. example (28b) (e.g. Sørensen & Boye 2015).

(28) Latvian (Indo-European; Kehayov & Boye 2016a: 814)

a. *viņš zināja, ka viņa spēlē klavieres*
he know.PST.3 COMP she play.PRS.3 piano.ACC[PL]
'He knew that she played the piano.'

b. *viņš zināja, kā spēlēt klavieres*
he know.PST.3 COMP play.INF piano.ACC[PL]
'He knew how to play the piano.'

Similarly, perception verbs that express "indirect perception" (or "acquisition of knowledge") take propositional complements (29a), while perception verbs that express "direct" (or "immediate") perception take state-of-affairs complements (29b) (e.g. Dik & Hengeveld 1991, Boye 2010).

(29) Latvian (Indo-European; Kehayov & Boye 2016a: 815)

a. *viņš redzēja, ka viņa spēlē klavieres*
he see.PST.3 COMP she play.PRS.3 piano.ACC[PL]
'He saw that she played the piano.'

b. *viņš redzēja, kā viņa spēlē klavieres*
he see.PST.3 COMP she play.PRS.3 piano.ACC[PL]
'He saw how she played the piano.'

As illustrated for knowledge complements in (30), only the propositional complements can be epistemically evaluated. (30b) is unacceptable if the epistemic adverb *visticamāk* 'probably' is read as having scope over the whole state-of-affairs '(he) play the piano'.

(30) Latvian (Indo-European; Axel Holvoet p.c.)

a. *viņš zināja ka visticamāk, viņa spēlē klavieres*
he know.PST.3 COMP probably she play.PRS.3 piano.ACC[PL]
'He knew that she probably played the piano.'

b. **viņš zināja kā, visticamāk, spēlēt klavieres*
he know.PST.3 COMP probably play.INF piano.ACC[PL]
Intended reading 'He knew how to probably play the piano.'

In the case of utterance complements, assertive complements are standardly analyzed as propositional, and directive complements (often treated as complements of manipulative predicates) as state-of-affairs designating (Noonan 2007, Cristofaro 2003, 2013). Accordingly, the former readily allow of epistemic modification while the latter do not: (31b) is unacceptable, at least if *probably* is read as taking the whole state-of-affairs '(her) to drink' in its scope.

(31) a. *He told her that she had probably been drinking.*

b. *He told her to probably drink.*

To our knowledge, interrogative complements have have not been discussed in relation to the proposition vs. state-of-affairs distinction, but they are clearly propositional (cf. Boye 2012: 194-195, 200-201). Firstly, they report polar questions, and polar questions are clearly not about actions, but about propositions: a polar question can be seen as a presentation of a proposition plus a request that the truth value of it is confirmed or disconfirmed. Secondly, they readily allow of epistemic modification – though, because of the uncertainty implied by the interrogative predicate, they occur most naturally with expressions of uncertainty, possibility or other values at the low end of a scale of degree of certainty (see Boye 2012: 312-313 for detailed discussion).

(32) **He asked her if she had maybe/perhaps/possibly been drinking.*

Different approaches to propositions and states-of-affairs converge on analyzing the two types of meaning as differing in terms of complexity such that there is a privative opposition between them: propositions are states-of-affairs plus something extra. This analysis is evident in frame-work neutral characterizations of propositions as "truth-valued" and states-of-affairs as "non-truth-valued" (e.g. Holvoet 2016: 228), but also in more elaborate theories of the opposition.

Functional Grammar (FG) and Functional Discourse (FDG) grammar adopt a denotational approach to the opposition and defines propositions and states-of-affairs as distinct kinds of denotable entities. Propositional entities (what FG and FDG refer to as "propositional contents") are defined in line with Lyons' (1977) "third-order entities" as "things that people can be said to believe, know or think about" (Dik 1997: 91), and state-of-affairs entities in line with his "second-order entities" as "something that can be said to occur, take place, or obtain in some world" (Dik 1997: 51), and it is of course not clear from such definitions that propositions would be more complex than states-of-affairs. However, the privative aspect is clear from the conception of the grammatical units that express proposition(al content)s and states-of-affairs: in FG's and FDG's layered clause structure propositional units are allocated to a level higher and thus more complex than state-of-affairs designating units (Boye 2012: Ch. 4, Kehayov 2017: 302).

Within Cognitive Grammar (CG), the privative nature of the distinction is evident in Langacker's (1991, 2009) and Achard's (2002) accounts. Langacker deals with the distinction in terms of the wider phenomenon of grounding: propositions are grounded processes, whereas states-of-affairs are covered by his ungrounded processes (e.g. Langacker 1991: 439-440; 551; Langacker 2009: 293). Thus, propositions are states-of-affairs plus grounding. Elaborating on this, Achard analyzes propositions as located in respectively elaborated and basic reality. This analysis is privative in so far as elaborated reality presupposes basic reality and adds something extra (Achard 2002: 207-209).

Finally, the privative nature of the distinction is evident in a recent cognitive-functional linguistic theory of the distinction which also builds on CG (Boye 2010, 2012). According to this theory, propositions and state-of-affairs both

evoke processes in the sense of Langacker (1987, 1991), that is, relational concepts with a temporal profile prototypically evoked by clauses. However, they differ in terms of reference (in the sense of Lyons 1977: 177–197), which may be seen as a specific type of grounding: propositions evoke processes construed as referring to states or events in the (real or imagined) world, while states-of-affairs evoke processes not construed as referring. Thus, propositions are states-of-affairs plus reference.

7.5.1 Conceptual semantics and balancing vs. deranking

Recall that all the crosslinguistic tendencies pertaining to the distinction between balancing and deranking (tendencies 9a, 9b and 9c) point in the same direction: assertive and interrogative complements tend to be balanced, while directive ones tend to be deranked. These tendencies can be accounted for at least in terms of Langacker's and Boye's conception of the proposition vs. state-of-affairs distinction.

A Langackerian account would be that the linguistic elements that distinguish balanced from deranked complements are grounding predications. As mentioned, Langacker deals with propositions as grounded processes, whereas states-of-affairs are covered by his ungrounded processes. Many of the linguistic elements that are unique to balanced clauses are TAM-markers and can thus straightforwardly be classified as "grounding predications", i.e. elements such as tense and mood markers that establish a "grounding relationship" between a process and the speaker's conception of the ground (e.g. Langacker 1987: 126-127, Langacker 1991: 439-440, Langacker 2009: 293; see also Achard 1998: 222-226).

Based on Boye (2010, 2012), the elements that distinguish balanced from deranked complements can be understood more precisely as markers of referential status. As argued by Boye (2010: 421-423), there is a close relationship between – at least temporal – grounding and reference in that temporal grounding provides a specification of the deictic coordinates for reference. Thus, markers of deictic tense can be seen as elements that both assign reference and provide the temporal coordinates for the reference. However, an analysis in terms of referential status would account also for balanced clause marking elements that are not TAM-markers and cannot immediately

be seen as grounding predications. An example is the Kayardild indicative, illustrated in (6a) above. This indicative can be understood as a marker that assigns reference to the process designated by the complement verb, but it clearly does not provide temporal coordinates for the reference; there is no tense marking. The contrasting hortative in (6b) must be analyzed as blocking reference.

7.5.2 Conceptual semantics and complementizer marking

Turning now to complementizer marking, we found a strong restricted tendency (the tendency in 18) that if the directive complement is marked by a complementizer, the contrasting assertive complement is marked by a complementizer as well. Langacker's (1987, 1991, 2009), Achard's (1998, 2002) and Boye's (2010, 2012) theories all make possible an account of this tendency in terms of iconicity of complexity. All theories entail that the contrast between proposition and state-of-affairs is privative, such that propositions are conceptually more complex than states-of-affairs. Thus, the restricted tendency for propositions to be marked by at least as much complementizer material as states-of-affairs can be seen as iconically reflecting (and possibly being motivated by) this privative relationship.

In contrast, the unrestricted tendencies in (13a-c) cannot be accounted for in terms of iconicity of complexity. This is because 1) unrestricted tendencies (unlike restricted ones) do not pertain to contrasts, and 2) accounts in terms of iconicity of complexity – just like accounts in terms of frequency, but unlike accounts in terms of iconicity of cohesion – presuppose a contrast: they relate morphosyntactic contrasts to conceptual contrasts (just as frequency accounts relate morphosyntactic contrasts to frequency contrasts).

This does not mean, however, that the unrestricted tendencies in (13) cannot be accounted for in terms of the contrast between propositions and states-of-affairs. The tendencies are as follows: assertive, interrogative and directive complements all tend to be marked by complementizers (though for directive complements the tendency is only weak). These tendencies can be accounted for as reflecting a need to indicate whether a complement designates a proposition or a state-of-affairs. Since Bolinger (1968: 122), there has been an increasing awareness that complementizers are not purely

formal elements devoid of function. Frajzyngier in particular has argued that they contribute semantically to the complement clause (e.g. Frajzyngier 1991, 1995, Frajzyngier & Jasperson 1991; see also e.g. Boye et al. 2015; Boye & Kehavov 2016), and in a survey of complementizer semantics in European languages, Kehayov & Boye (2016a) found that a frequent function of complementizers is to mark the contrast between propositions and states-of-affairs.

7.6 Conclusion

In a survey of the three major types of utterance-predicate complements, based on data from 84 genetically diverse languages, we documented two sets of crosslinguistic tendencies. One set has to do with the distribution of balanced vs. deranked complements across the three complement types:

- Tendency (9a): There is a strong tendency for assertive complements to be balanced.
- Tendency (9b): There is a strong tendency for interrogative complements to be balanced.
- Tendency (9c): There is a weak tendency for directive complements to be deranked.
- Tendency (10): In languages with a contrast between deranked and balanced complements , there is a strong tendency that if the directive complement is balanced, the contrasting assertive complement is balanced as well.

The other set concerns the distribution of complementizers:

- Tendency (13a): There is a strong tendency for assertive complements to be marked by.
- Tendency (13b): There is a strong tendency for interrogative complements to be marked by complementizers.
- Tendency (13c): There is a tendency – though only weak – for directive complements to be marked by complementizers.

- Tendency (18): In languages with a complementizer contrast, there is a strong tendency that if the directive complement is marked by a complementizer, the contrasting assertive complement is marked by a complementizer as well.

We contrasted three types of accounts of these tendencies. Based on corpus studies of English (British National Corpus; predominantly written) and Danish text collections (the LANCHART corpus; exclusively spoken), we can dismiss an account of any of the tendencies in terms of frequency of use. While frequency of use is undoubtedly a strong factor motivating grammatical asymmetries, it does not seem to play any role in the asymmetries under scrutiny here.

In contrast, the tendencies in (9), (10), (13) and (18) can be straightforwardly accounted for in terms of iconicity of cohesion, on the analysis that assertive and interrogative complements are semantically highly disintegrated, whereas directive complements are highly integrated with the complement-taking clause. However, the tendency in (13c) runs counter to what one would expect based on this analysis.

All tendencies can be accounted for simultaneously based on a cognitive analysis of the semantics of the three complement types. Assertive and interrogative complements designate propositions, while directive complements designate states-of-affairs. Cognitive linguistic theories converge on seeing the contrast between proposition and state-of-affairs as a privative one such that both propositions and states-of-affairs evoke a Langackerian process, but only propositions evoke something extra (grounding or reference). The privative nature of this contrast allows us to account for the tendency in (18) in terms of iconicity of complexity, and for the tendencies in (9a–c) and (10) in terms of a link between especially TAM-markers and grounding or reference. The tendencies in (13a–c) can be explained as reflecting the functional importance, hence frequent coding, of the proposition vs. state-of-affairs contrast.

Chapter 8

Article III: Clausal Complementation in Ruuli (Bantu; JE103)

8.1 Introduction

This paper describes the morphosyntactic and semantic characteristics of clausal object complementation in the Great Lakes Bantu language Ruuli (also known as Luruuli/Lunyara, ISO 639-3; ruc; JE103). We will analyze the features of Ruuli complementation and compare them to those of other Bantu languages as well as to common crosslinguistic patterns in clausal complementation.

Ruuli employs several different complementation strategies, including indicative, subjunctive and infinitive constructions. Complement clauses can be either unmarked or marked with a complementizer, the most common of which is *nti*. These two options are also available for direct speech. Other less common complementizers are *oba*, *nga* and *ni*, which cannot be used to introduce direct speech. As individual complement-taking predicates do not allow for every complementation strategy and have preferences for specific complementation strategies, we will explore the morphosyntactic and semantic conditions which predict the choice of complement. We investigate the restrictions imposed by various complement-taking predicate

types, e.g. knowledge predicates, phasal predicates and utterance predicates. Then we consider whether the complement refers to a proposition (a truth-valued meaning unit) or a state-of-affairs (a non-truth valued meaning unit) and whether the subject arguments in the two clauses are identical. We also consider the polarity of the two clauses

The paper is structured as follows. Section 8.2 provides at typological overview of clausal complementation. Section 8.3 introduces the Ruuli language and its speakers, as well as outlines the sample and methods used in the present study. Section 8.4 describes the main complement types in Ruuli and Section 8.5 analyzes the distribution of different complement types relative to different semantic classes of complement-taking predicates. Section 8.6 discusses the main predictors of complement choice. A conclusion and outlook follows in Section 8.7.

8.2 A typological overview of clausal complementation

This section provides a brief overview of central morphosyntactic and semantic features of clausal complementation from a crosslinguistic perspective and serves as a preliminary to placing the complementation strategies found in Ruuli in a broader typological context. We review definitions of complement clauses and complementation strategies in Section 8.2.1, complement-taking predicates in Section 8.2.2, aspects of complement verb forms (e.g. TAM-marking) in Section 8.2.3, complementizers in Section 8.2.4, as well as the semantic distinction between states-of-affairs and propositions in Section 8.2.5.

8.2.1 Complement clauses and complementation strategies

Complement constructions have been defined in both syntactic and semantic terms. In studies based on semantic definitions, terms such as *complement relations* and *complementation strategies* tend to be preferred over *complement clauses*.

Syntactic definitions of complement constructions define complement clauses as core syntactic arguments (i.e. subject or object) of complement-

taking predicates in clausal form (as opposed to NP-complements) (Dixon 2006: 15). On this account, clauses such as the one in brackets in (1) can be described as a complement clause, because it functions as the object of the matrix clause *She thinks*.

(1) *She thinks [they are gone]*.

Semantically oriented analyses of complement constructions, on the other hand, define complements as semantic (rather than syntactic) arguments of complement-taking predicates (Noonan 2007) and in terms of the relationship between two states-of-affairs Cristofaro (2003: 95). The motivation for the latter definition, in particular, is that definitions of complement clauses as syntactic arguments is too narrow for typological purposes and implies that complements are embedded. English complement clauses like the one in (1) can be analyzed as embedded syntactic arguments of a complement-taking predicate, and similar constructions exist in other languages. However, complements in certain other languages cannot be described as embedded, even though they express similar semantic relations.

A case in point are languages in which reported speech constructions are ambiguous between direct reported speech and indirect reported speech. Although direct reported speech is not a syntactic argument of a complement-taking predicate and therefore not what is traditionally meant by a complement clause, functionally it constitutes a *complement relation* (Cristofaro 2003, 2013) or in the terms of Dixon (2006) a *complement strategy*. Because of the existence of constructions which are arguably instances of complementation, despite not being syntactic arguments, Cristofaro (2003: 95) prefers the term *complement relations* which she defines as a situation where two states-of-affairs are linked "such that one of them (the main one) entails that another one (the dependent one) is referred to." (cf. Section 8.2.5 below on states-of-affairs vs. propositions). The advantage of this definition is that it functionally covers complement relations in all languages, while remaining compatible with the definition in terms of argument status.

In this article we adopt this broader conception of complementation in order to achieve the most comprehensive description of the structure and semantics of complementation in Ruuli. The analysis will thus cover more

traditional examples of complement constructions as well as reported speech, which would be excluded under a narrow syntactic definition of complement clauses. We use the term *complement* to refer to clausal semantic arguments of complement-taking predicates, including infinitives as well as reported speech.

8.2.2 Complement-taking predicates

In crosslinguistic studies of complementation, a number of complement-taking predicate classes have been identified. Commonly identified classes include modal predicates (e.g. *can*, *may*), phasal predicates (e.g. *begin*, *continue*), manipulative predicates (e.g. *force*, *make*), desiderative predicates (e.g. *want*, *wish*), perception predicates (e.g. *see*, *hear*), knowledge predicates (e.g. *know, forget*), propositional attitude predicates (e.g. *think*, *believe*) and utterance predicates (e.g. *say*, *ask*) (Ransom 1986, Givón 2001, Cristofaro 2003, Dixon 2006, Noonan 2007).

Existing crosslinguistic classifications of complement-taking predicates are typically based on a combination of predicate semantics and the type of complement that occurs with a given complement-taking predicate. A downside of such classifications is that they downplay the fact that certain predicates — such as *see*, *know* and *tell* in English – are able to take more than one complement, as pointed out by e.g. Ransom (1986), Verspoor (2000), Boye (2012) and Serdobolskaya (2016). Consider the contrast between the gerund and the indicative complement of *see* in example (2).

(2) a. *She saw him playing the piano.*

b. *She saw that he played the piano.*

The term *perception predicate* would be used to describe complement-taking predicates, such as *see*, only in cases where they occur with a complement describing direct perception as in (2a). In contrast, *see* in example (2b) would be described as a knowledge predicate. This classification might have the unwanted side-effect of language-specific descriptive studies being led to focus only on direct-perception constructions with complement-taking predicates meaning ‘see’ and overlook other complement types used with the

same predicate. In the analysis of Ruuli, we make an attempt at highlighting contrasts such as the one in example (2), as we describe the distribution of complement types over different predicate classes.

8.2.3 Complement verb forms

Typical contrasts between complement verb forms include contrasts between those forms that can be described as finite vs. non-finite ones or between balanced vs. deranked ones (cf. Stassen 1985, Cristofaro 2003). Balanced verbs are verbs that correspond to verbs in independent declarative clauses with regards to TAM-marking and/or agreement, while deranked verbs are verbs that are different from verbs in an independent declarative clause in some way in terms of tense, aspect, mood or person agreement or by having markers that are not found in independent declarative clauses (Cristofaro 2013). In contrast to Stassen (1985), Cristofaro (2003: 57) furthermore considers a number of moods such as subjunctives and hortatives to be deranked by default.

Languages vary considerably as to whether — and to what extent — TAM-marking is obligatory on complement verbs. On one end of the spectrum, there are verbs that need to be marked for the exact same TAM-categories as verbs in independent clauses. On the other end of the spectrum, there are verbs that do not allow any of the TAM-marking available to an independent clause verb. In intermediate cases, verbs may display a large number of TAM-markers, even when the complement does not allow all of the TAM-markers possible for an independent clause verb. In the latter case, the complement would still be considered deranked in the sense of Cristofaro (2003), but finite in traditional terms.

8.2.4 Complementizers and quotatives

Another important point of variation in complementation is the distribution and function of *complementizers*. According to Noonan (2007: 55), a complementizer is "a word, particle, clitic or affix, one of whose functions it is to identify [a clause] as a complement", such as English *that* and *whether*. Diachronically, complementizers are often derived from elements such as demonstratives or case markers and more rarely from verbs (Heine & Kuteva

2002, Dixon 2006, Noonan 2007). In deviation from the crosslinguistic tendency, complementizers in Bantu languages are rather frequently derived from speech verbs, but can also have other sources such as personal pronouns (Kawasha 2007: 181).

While some languages have rich systems of complementizers expressing many different morphosyntactic and semantic features, other languages have few complementizers or lack them altogether. In addition to identifying a clause as a complement, complementizers may have additional semantic functions, for instance modal functions, indicating epistemic certainty or uncertainty (Frajzyngier 1995, Frajzyngier & Jasperson 1991), or the function of distinguishing between propositions (truth-valued meaning units) and states-of-affairs (non-truth valued meaning units) (Kehayov & Boye 2016a: 812–818, cf. Section 8.2.5).

The identification of complementizers in any individual language is complicated by the fact that a candidate for complementizer status might also function as e.g. a relativizer or adverbializer or it might be synchronically identifiable as an adverb, verb or noun (Kehayov & Boye 2016b: 83).

In analyses of complementation, the term *quotative* rather than complementizer is sometimes encountered (note that the same term has also been used to describe items that are not complementizers, but rather reportative evidentials (also known as"hearsay-markers"), cf. Boye 2012: 20). This term is often used for complementizers that introduce direct reported speech – probably motivated by the fact that within some analytical approaches to complementation, direct reported speech is not considered a type of complement. In languages with a complementizer that only occurs with direct reported speech introduced by an utterance predicate, this can make sense. However, if the so-called quotative is also found with other complement-taking predicate types and/or with indirect reported speech (or complements that do not report speech at all for that matter), the line between quotatives and complementizers begins to blur. As Güldemann (2008: 456) concludes, there is generally "no principled distinction between a quotative and a complementizer".

In the remaining part of the paper, we will use the term *complementizer* to cover a morpheme which has among its functions to introduce the fol-

lowing clause as a complement no matter whether the complement can be characterized as reported speech or not.

8.2.5 Complement semantics

A central motivation for complement contrasts, such as balancing-deranking contrasts and complementizer contrasts, is the semantic contrast between states-of-affairs and propositions. This contrast has traditionally been understood as a contrast between non-truth valued and truth-valued meaning units (but see Boye 2012 and the references therein for alternative cognitive-functional analyses). The terms states-of-affairs and propositions are also known as *actions* and *facts* (Lees 1960), *second-order entities* and *third-order entities* (Lyons 1977), and *events* and *propositions* (e.g. Palmer 1979, Perkins 1983. This distinction plays a central role in functional frameworks such as Functional Grammar (Hengeveld 1989) and have been shown to be analytically applicable to a number of genealogically unrelated languages across the world. Language-specific studies include analyses of clausal complementation in Mayan languages (e.g. Schüle 2000 on Akateko), Koreanic and Japonic languages (e.g. Horie 2000 on Japanese and Korean) and Indo-European languages (e.g. Holvoet 2016 on Latvian) and Turkic (e.g. Rentzsch & Mitkovska 2017 on Balkan and Standard Turkish).

Examples (3)–(5) illustrate some morphosyntactic contrasts from English that have been linked to the semantic contrast between states-of-affairs and propositions. In (3a) *see* takes a finite complement and is used in the sense of acquisition of knowledge (also called 'indirect perception'), while in example (3b), *see* takes a non-finite complement and is used in the sense of direct/immediate perception. The complement in (3a) expresses a proposition, whereas the complement in (3b) expresses a state-of-affairs (Dik & Hengeveld 1991, Boye 2010).

(3) a. *She saw that he played the piano.*

b. *She saw him playing the piano.*

In (4a) *know* has the sense of knowledge of information, which can be called *epistemic knowledge*, and the finite complement expresses a proposition, while in (4b), *know* has the sense of knowledge of how to do something,

which can be called *action knowledge*, and occurs with a non-finite complement expressing a state-of-affairs (cf. Sørensen & Boye 2015).

(4) a. *She knows that he plays the piano.*

b. *She knows how to play the piano.*

In (5a), *tell* introduces a finite complement and is used to report an assertion expressing a proposition, whereas in (5b), *tell* describes an order/request and occurs with a non-finite complement expressing a state-of-affairs.

(5) a. *She told him that they played the piano.*

b. *She told him to play the piano.*

The English predicates *see*, *know* and *tell* all occur with both state-of-affairs and propositional complements, as exemplified in (3)–(5). However, in other languages that distinguish between states-of-affairs and propositions, complement-taking predicates need not be polyfunctional in this respect. Instead some languages employ different complement-taking predicates to introduce propositional and states-of-affairs complements. We will make reference to this distinction throughout the paper and in Section 8.5, we exemplify contrasts between states-of-affairs and propositions within different semantic classes of predicates.

Crosslinguistically there is a great deal of variation in the morphosyntactic features of complement constructions and a given languages may employ several complementation strategies, each with its own semantic and distributional properties (Noonan 2007). But there are also recurrent semantic patterns, such the contrast between states-of-affairs and the semantic classification of predicates, that have proven useful for the analysis of most languages.

8.3 Ruuli

Ruuli (or Luruuli/Lunyara or Ruruuli-Runyala, E.103, ISO 639-3: ruc) is a Great Lakes Bantu language mainly spoken in the Nakasongola and Kayunga districts of central Uganda. Following Schoenbrun (1994), Hammarström

et al. (2018) classify Ruuli as a West Nyanza/Rutara language of the Great Lakes Bantu group of languages. However, it should be noted that Schoenbrun's genealogical study did not include any data from Ruuli as there were none available at the time (Schoenbrun 1994: 118–119).

The number of ethnic Baruuli/Banyara is 190,000 according to the 2014 census, but the actual number of Ruuli speakers is difficult to estimate at the moment. Most speakers of Ruuli are multilingual. In addition to often being fluent in Ganda (West Nyanza/ North Nyanza, JE.15, the dominant language of the area) and English (the official language of instruction in Uganda), many speakers interviewed for the corpus used in the present study indicated that they also speak other, mostly Bantu languages of Uganda.

Ruuli is a typical Bantu language: The dominant constituent order is SVO. Nominal and verbal inflectional morphology is primarily prefixing. Nominal morphology is characterized by a system of noun class prefixes (cf. Katamba 2003). Every noun in singular and plural is assigned to one of the twenty noun classes numbered from 1 to 23. The class numbers correspond to the reconstructed Proto-Bantu classes with their respective noun prefixes and are used to label noun classes in all Bantu languages (cf. Katamba 2003: 104–105 and the references therein). Modern Bantu languages do not have all of the Proto-Bantu noun classes, and this explains the gaps in the numbering. Ruuli lacks classes 13, 19, and 21. Most nouns have both singular and plural forms and thus belong to two noun classes. As in many other Bantu languages, the odd-numbered classes 1, 3, 5, 7, and 9 contain singular nouns, their corresponding plural forms usually belong to the even-numbered classes 2, 4, 6, 8, and 10, respectively (cf. Katamba 2003: 109–110). To enhance readability we do not segment nominal prefixes on nouns in the examples and indicate noun classes in brackets, as e.g. on *nkodole* 'francolin(9)' in example (6) which has a homorganic nasal prefix of class 9 realized as *n-*.

The nominal class determines the shape of the agreement prefixes on dependents in a noun phrase, on the verb, as well as on a number of other constituents. In addition, in many cases the shape of the nominal prefix on nouns is identical to the shape of the agreement prefix on various dependents. We indicate the class agreement prefixes on dependents by segmenting them

and providing the respective class number in Arabic numerals, as in the case of the object index *li-* '5O' on the verb in (7). We also use Arabic numerals to indicate person indexing on the verb and person information on pronouns. Notice that in these cases the Arabic numerals are always followed by the indication of number (SG or PL), for instance, both on the verb and the pronoun in (6).

As in some Bantu languages, the noun-class prefixes in Ruuli are often preceded by another prefix – the so-called augment, also referred to as the pre-prefix or the initial vowel (cf. Katamba 2003: 107–108). For instance, the noun *i-sumu* '(5)-spear' is used with the augment *e-* in (7). The augment is not segmented and not glossed in the examples in order to enhance readability.

The verb in Ruuli has about nine prefix slots and five suffix slots (the final analysis of the verb morphology is still pending). The verb of an independent clause obligatorily indexes its subject, as in (6), as well as optionally its object, as in (7).

(6) *nje n-lia nkodole*
1SG 1SG.S-eat francolin(9)
'I eat a francolin.'

(7) *naye nje eisumu n-a-li-zw-ire=ku*
but 1SG spear(5) 1SG.S-PST-5O-abandon-PFV=17LOC
'But I abandoned the spear.'

The data used in the present study come from a corpus collected within the project *A comprehensive bilingual talking Luruuli/Lunyara-English dictionary with a descriptive basic grammar for language revitalisation and enhancement of mother-tongue based education* (PI Saudha Namyalo, 2017–2020, funded by the Volkswagen Foundation). Before the beginning of this project no description of Ruuli existed. As of May 2018 the corpus contained about 150,000 words of naturalistic speech (transcribed and translated). Further 50,000 were available from digitalized written resources produced by the speakers' community. We first sampled about 1,000 tokens of complementation by exhaustively identifying all potential complementation strategies and complement taking verbs in a number of texts. These 1,000 tokens

were annotated for such variables as the presence and form of complementizers, polarity of the two clauses, the identity of subjects in the two clauses, the semantic class of the complement-taking predicates, etc. Once the major complement-taking predicates and complementation strategies were identified and initial hypothesis about their distribution were made, we focused on obtaining further tokens (about 500) of less common complement-taking predicates and complementation strategies.

8.4 Complementation in Ruuli

In this section we will describe and exemplify the complement types found in Ruuli. There are three main complement types, given in (8).

(8) Complement types

a. indicative complements (suffix *-a*)

b. subjunctive complements (suffix *-e*)

c. infinitive complements (class marker *(o)ku-*)

This distinction between indicative, subjunctive and infinitive complements is typical for Bantu languages and there is furthermore a strong tendency for it to be expressed in morphologically similar ways across individual languages (Myers 1975: 185), i.e. as affixes that are identical or near-identical to *-a*, *-e* and *(o)ku* in Ruuli. In Ruuli the most frequent and versatile complement type is the infinitive, which occurs with all classes of complement-taking predicates except for perception predicates (see Section 8.5).

More diversity is found between individual Bantu languages when it comes to the form, distribution and function of complementizers. In Ruuli there are at least three complementizers, given in (9). Complementizers mark indicative complements only.

(9) Complementizers

a. *nti*

b. *oba*

c. *nga*

In sections 8.4.1-8.6.1 we consider each individual complement type and complementizer in turn. In Section 8.4.7 we furthermore consider an additional potential candidate for complementizer status, *ni*. Reported speech will be considered in Section 8.5.9.

Complements generally follow the complement-taking predicate. However, on rare occasions, complements precede the complement-taking predicate, as in (10).

(10) *ti-bi-kya-tu-kol-a* *n-dowoz-a.*
NEG-8S-PERS-1PL.O-work-FV 1SG.S-think-FV
'They no longer work for us, I think.'

In this case, however, *lowooz* 'think' is arguably a parenthetical verb rather than a complement-taking predicate, cf. discussions of parentheticals versus complement-taking predicates in Thompson (2002), Boye & Harder (2007) and Newmeyer (2015). Such constructions will therefore be exempt from our analysis of Ruuli complementation.

8.4.1 Indicative complements

The verb form in indicative complements corresponds to that of the verb in an independent declarative clause as described in Section 8.3. The verb obligatorily indexes the subject and optionally the object and can occur with any TAM-categories the verb in an independent clause can occur with. It can be negated in the same way verbs in independent clauses are negated. Indicative complements follow perception predicates, knowledge predicates, propositional attitude predicates and utterance predicates and can be marked by a complementizer (*nti*, *oba* or *nga*, cf. Sections 8.4.4, 8.4.5 and 8.4.6, respectively). Examples of indicative complements can be found throughout the article, e.g. (37b), (38), and (46) and (61a).

8.4.2 Subjunctive complements

The subjunctive in Ruuli is formed by adding the suffix *-e* to a verb stem replacing the neutral final vowel *-a*. Both the form and the function of this

suffix are similar to the cognate ones in closely related Great Lakes Bantu languages (Nurse & Muzale 1999), as well as in many other Bantu languages (Nurse 2008: 44, 192). Final **-e* has also been reconstructed for Proto-Bantu (Meussen 1967).

In Ruuli independent clauses the subjunctive is used to express hortative, optative and modal meanings, as in (11) and (55) below.

(11) a. *tu-somesy-e baana.*
1PL.S-teach-SBJN child(2)
'Let us educate children.'

b. *okatonda a-tu-jun-e.*
god(3) 3SG.S-1PL.O-help-SBJN
'May god help us.'

Similarly to the finite indicative forms (Section 8.4.1) and in contrast to infinitives (Section 8.4.3), the subjunctive shows obligatory subject agreement, as in (12). In contrast to the finite forms, the subjunctive is not marked for either tense or aspect.

In complement constructions, the subjunctive is found with modals, desideratives and utterance predicates. Complements with subjunctive verbs are never marked by complementizers. Subjunctives occur both in same-subject (12a) and different-subject (12b) constructions.

(12) a. *o-tak-a ate [o-ta-e=wo olukonko].*
2SG.S-want-FV FOC 2SGS-put-SUBJ=LOC rift(11)
'You want to cause a rift.'

b. *n-ku-tak-a [ansemu ya Bunyala*
1SG.S-PROG-want-FV anthem(9) 9.GEN Bunyala(14)
e-bba-e omu isomero lya-amu].
9S-be-SBJN 18LOC school(5) 5-2pl.POSS
'I want the Bunyala anthem to be in your school.'

8.4.3 Infinitive complements

The infinitive is formed by adding the class 15 prefix *ku-* to the verb stem (followed by a final vowel). The class 15 infinitive morpheme **ku* is reconstructed for Proto-Bantu (see Nurse 2008: 141) and is found in many other Bantu languages, e.g. in the closely related Soga (Nabirye 2016: 309) and Nkore-Kiga (Taylor 1985: 20). The prefix *ku-* is realized as *kw-* when the following verb stem begins with a vowel.

In most instances, the respective augment prefix *o-* precedes the class 15 prefix *ku-*, as with *oku-emb-a* 'to sing' in (13). The distribution of *ku-* vs. *oku-* cannot be described definitively as of now, but one generalization is worth noticing: With negated complement-taking predicates only *ku-* occurs. However, when the complement-taking predicate is not negated, both *oku-* and *ku-* are possible.[17]

(13) *n-yend-a [muno oku-emb-a].*
1SG.S-like-FV much INF-sing-FV
'I liked to sing very much.'

(14) *nbantu ti-ba-kya-yendy-a [ku-kol-a].*
person(2) NEG-3PL.S-PERS-want-FV INF-work-FV
'People no longer want to work.'

The infinitive does not show subject indexing and does not take any TAM-marking. The infinitive can index objects, as well as take valency-changing affixes, as in (15). The infinitive can be negated with the prefix *ta-*. Instead of the regular class 15 prefix *ku-*, negative infinitives take class 14 prefix *bu-* frequently accompanied by the respective augment *o-*, as in (16).

(15) *ni-ba-tak-a [oku-tu-band-isy-a o-Kawumpuli].*
NAR-3PL.S-want-FV INF-1PL.O-worship-CAUS-FV AUG-Kawumpuli
'They wanted to make us worship Kawumpuli.'

[17] This is different from some other Bantu languages, which reportedly show free variation between *oku* and *ku* infinitives. Describing Nkore-Kiga, for example, Taylor (1985: 28) notes that "[t]he form *ku* is normal, and *oku* gives a more general force".

(16) *omusaiza a-yinz-a [obu-ta-leet-a=wo mukali].*
man(1) 3SG.S-be.able-FV INF-NEG-bring-FV=LOC wife(1)
'The man may not bring there a wife.'

The infinitive can follow all predicate types described in Section 8.5 below except for perception predicates, that is modals, phasals, desideratives, knowledge predicates, propositional attitude predicates, emotive predicates and utterance predicates. Infinitives are used exclusively to express states-of-affairs and occur both in same-subject and different-subject constructions, cf. sections 8.5 and 8.6.

8.4.4 Complementizer *nti*

The most frequent complementizer in Ruuli is *nti*, which optionally marks indicative complements introduced by perception predicates, such as 'see' as in (17a), knowledge predicates, such as 'know' as in (18a), propositional attitude predicates, such as 'think', and utterance predicates, such as 'say'. The cognate form is used in similar contexts in the closely related Ganda (Ashton et al. 1954: 502) and Soga (Nabirye 2016: 390).

(17) Perception

a. *o-ku-bon-a [nti te-tu-ku-sigal-a mabega].*
2SG.S-PROG-see-FV COMP NEG-1PL.S-PROG-stay-FV back(6)
'You see that we are not staying behind.'

b. *nje n-ku-bon-a [buli kimwei o-mwana*
1SG 1SG.S-PROG-see-FV every thing(7) child(1)
a-ku-sobol-a oku-ki-tambuly-a].
3SG.S-PROG-can-FV INF-7O-perform-FV.
'I see (that) a child can perform everything.'

(18) Knowledge

a. *o-maite [nti e Ibbaale tu-tandik-ire*
2SG.S-know.PFV COMP 23LOC Ibbaale 1PL.S-start-PFV
oku-somesy-a abaana].
INF-teach-FV child(2)
'Do you know that we started educating children at Ibbaale?'

b. *o-maite [ye-ena a-yendy-a oku-yizukiry-a].*
2SG.S-know.PFV 3sg-FOC 3SG.S-need-FV INF-be.reminded-FV
'You know (that) he also needs to be reminded.'

The absence or presence of *nti* does not appear to be correlated with any semantic contrast, nor does *nti* disambiguate direct reported speech from indirect reported speech (cf. Section 8.5.9). *Nti* is equally optional with indirect reported speech, as shown in (19a), and with direct reported speech, as in (19b).

(19) a. *ni-a-kob-a [ba-a-bi-gul-ire mpani].*
NAR-3SG.S-say-FV 3PL.S-PST-8O-buy-PFV here
'He said (that) they bought them here.

b. *ni-bi-simool-a [aah iswe ba-tu-leet-ire*
NAR-8S-say-FV INTERJ 1PL 3PL.S-1PL.O-bring-PFV
ku-kol-a sente].
INF-make-FV money(10)
'They said, "Aah, we were brought to make money.'

In addition to its use as a complementizer, *nti* can be used as a quotative particle independent of the presence of a matrix predicate, which serves to present a word as a direct quote as in (20).

(20) *nti bbe.*
QUOT no
'she answered) "No."'

8.4.5 Complementizer *oba*

A less frequent complementizer is *oba* which marks indicative complements and expresses doubt towards the proposition, as in (21). This morpheme is otherwise used with the meaning 'or' to coordinate two noun phrases, verbs, and other units of the same type.

(21) *ti-maite [oba ki-kol-a].*
NEG.1SG.S-know COMP 7S-work-FV
'I don't know whether it works.'

Oba is also used for indirect reports of polar questions, as illustrated in (55) in Section 8.5.9.

8.4.6 Complementizer *nga*

Another less frequent complementizer is *nga* (sometimes reduced to *ng'* when preceding a vowel). The morpheme *nga* is found in a large number of Bantu languages, but with language-specific functions (Kimenyi 2018).

(22) *abaana o-yiz-a ku-bon-a [nga ba-ku-yikiriz-a].*
child(2) 2SG.S-AUX-FV INF-see-FV CONJ 3PL.S-2SG.O-believe-FV
'You will see the children believe you.'

As a complementizer its use is restricted to perception predicates *wuur* 'hear' and *bon* 'see', as well as a single phasal predicate *sigal* 'continue, remain'. The semantics of *nga* is not yet clear to us, but possible semantic motivations for the distribution will be discussed in sections 8.5.2 on phasal predicates and 8.5.5 on perception predicates.

In addition to its use as a complementizer, as in (22), *nga* can also be used as a temporal conjunction 'while, when', as in (23) (see also Nabirye 2016: 390–391 on the use of the cognate form in the closely related Soga with a similar range of functions).

(23) *n-a-som-ere nga n-kya-li mu-to.*
1SG.S-PST-study-PFV CONJ 1SG.S-PERS-be 1-young
'I studied, when I was still young.'

8.4.7 Other complementizers

In addition to *nti*, *oba* and *nga*, there are a few other less frequent complementizers, which need further study. One of them is *ni*, which is primarily used as the temporal and conditional conjunction. We have found examples of *ni* occurring with the verbs *bon* 'see' and *izukir* 'remember', as in (24) and (25).

(24) *o-a-bwon-ire* *[ni* *bu-ku-emba* *olulimi*
2SG.S-PST-watch-FV CONJ 14S-PROG-sing-FV language(11)
lwa-abwe *o-lwa* *Bunyala]?*
14-3PL.POSS AUG-11.GEN Bunyala
'Did you watch how [=when] they sang their song about Bunyala?'

(25) *izukir-a* *[ni* *tu-a-iruk-ir-ire* *oku* *Lango*
remember-FV CONJ 1PL.S-PST-run-APPL-PFV 17LOC Lango
eyo].
therefore
'Remember when we ran to Lango.'

8.5 Complement-taking predicates and complement semantics

Below, we will define a number of complement-taking predicate classes and describe the distributional variation of complements within each class. The point of departure for the analysis of Ruuli has been the classes of complement-taking predicates defined in Cristofaro (2003). However, it should be noted that we work with slightly different definitions of specifically modal, manipulative, perception, knowledge and utterance predicates. We choose not to place complement-taking predicates occurring with several complement types in more than one category. Thus, Section 8.5.5 labeled *perception predicates*, for example, covers all constructions with complement-taking predicates that describe perception such as *bon* 'see' in constructions that describe 'direct perception' (*I saw him leave*) as well constructions that

describe 'indirect perception' (*I saw that he left*) and not only the former. In this way, we wish to display the versatility of individual complement-taking predicates, as well as systematically discuss the prevalence of contrasts between states-of-affairs and propositions in Ruuli complementation.

8.5.1 Modals

Modal predicates have meanings such as 'may' and 'can' and describe likelihood, possibility, ability, permission and obligation (Palmer 2001: 33). We follow Van der Auwera & Plungian (1998) and distinguish two major types of modal expression, viz. possibility and necessity, as well as four modality domains viz. participant-internal modality, participant-external modality with a subtype of deontic modality, as well as epistemic modality.

In Ruuli at least six different verbs are regularly used as modal predicates. Some of these verbs are attested in other West Nyanza languages with similar meanings (see e.g. Nabirye 2016: 313, Kawalya et al. 2018).

Possibility is expressed in Ruuli by the modal verbs *sobol* and *yinz*[18] both translated as 'can, may, be able'. These two verbs can express all possibility domains: participant-internal possibility, as in (26a) and (27a), non-deontic participant-external possibility, deontic participant-external possibility, as in (26b), as well as epistemic possibility, as in (27b).

(26) a. *n-sobola [oku-sosoitoor-a omuntu ekiibulo].*
1SG.S-can-FV INF-serve-FV person(1) meal(7)
'I can serve a person a meal'

b. *o-sobol-a [oku-yab-a omu kanisa].*
2SG.S-can-FV INF-go-FV 18.LOC church(9)
'You can go to church.'

[18] Kawalya et al. (2018) claim that the verb *yinz* is only attested in North Nyanza languages and is not found in Rutara, the branch of West Nyanza to which Ruuli belongs. However, Ruuli data, which were not available to the authors of that paper, contradict this claim.Though the possibility of a borrowing from the dominant Ganda cannot be excluded, the verb *yinz* is frequently used in the corpus and occurs in a wide array of contexts building an integral part of Ruuli grammar.

(27) a. *o-yinz-a [ku-n-fun-ir-a=yo o-nkowu*
2SG.S-can-FV INF-1SG.O-get-APPL-FV=23LOC AUG-guinea.fowl(1)
o-mwei]?
1-one
'Can you get me one guinea fowl?'

b. *oba e Kidera eyo omu katale ga-yinz-a*
perhaps 23LOC Kidera there 18LOC market(12) 6S-can-FV
[oku-bbaa=yo=ku].
INF-be-FV=23LOC=17LOC
'There may be some (spears to buy) at Kidera in the market.'

Necessity is expressed by the modal verbs *lin* and *teek(w)* both translated as 'must, have to'. Both these verbs are used to express deontic (28) and non-deontic participant-external necessity. Only *teek(w)* is found in our corpus to express epistemic necessity, as in (29).

(28) *o-lin-a [ku-sal-a musaayi].*
2SGS-have.to-FV INF-sacrifice-FV blood(3)
'You have to sacrifice blood.'

(29) *ka-teek-a [oku-bba=mu omu kidoodolo].*
12S-must-FV INF-be=18LOC 18LOC granary(7)
'It must be there in the granary.'

All the modal verbs discussed above occur exclusively with infinitive complements (introduced in Section 8.4.3). To express participant-internal necessity the verb *yendy* and occasionally *tak* are used. In most cases they are used with the meaning 'want' and in this meaning they take both infinitive and subjunctive complements primarily conditioned by whether the subject of the matrix clause and the complement verb are identical (see Section 8.5.4). In their modal meaning, they also allow for these two complementation strategies under the same conditions. With the same subject in the two clauses we find an infinitive complement, as in (30).

In addition to these two verbs, the corpus contains a few tokens of the verb *etaag* 'need', which is also used to express participant-internal necessity.

Cognates of this verb are found in a number of closely related languages (see e.g. Nabirye 2016: 313, Kawalya et al. 2018).

(30) *n-kw-endya [ku-tanaka].*
1SG.S-PROG-need INF-vomit
'I need to vomit.'

Of the verbs discussed above, only some are also found as lexical verbs taking nominal objects. *sobol* has the meaning 'manage smth./smb.' when used as the main verb, *eetag* is used as 'need smth.', *lin* is common as 'have smth./smb.', *tak* and *yendy* both meaning 'want smth./smb.',

8.5.2 Phasals

Phasal predicates are predicates that describe the beginning, continuation or end of an event. There are at least five phasal predicates in Ruuli: *tandik* 'start/begin', *mal* 'finish', *lek* 'stop/leave', *onger* 'continue' and *sigal* 'remain, continue'.

Phasal predicates are primarily used with infinitives, as in (31).

(31) a. *ba-tandik-ire [oku-tu-bulyabuly-a].*
3PL.S-start-PFV INF-1PL.O-confuse-FV
'They started confusing us.'

b. *Tu-lek-e [ku-yendek-a].*
1PL.S-stop-SBJN INF-burden.oneself-FV
'We stop burdening ourselves.'

In contrast to other phasal predicates *sigal* 'remain, continue' prefers a different construction, namely a complement marked by the complementizer *nga* presented in Section 8.4.6, as in (32a) and (32b). The complementizer *nga* is otherwise only found with perception predicates (cf. Section 8.5.5).

(32) a. *ni-n-sigal-a [nga n-e-gomb-a].*
NAR-1SG.S-remain-FV COMP 1SG.S-admire-FV
'I was left admiring.'

b. *ba-sigal-ire* *[nga* *ba-iz-a* *oku-n-syom-a]*.
3PL.S-remain-PFV COMP 3PL.S-AUX-FV INF-1SG.O-hire-FV
'They could still hire me.'

8.5.3 Manipulatives

Manipulative predicates are predicates with meanings such as 'make' and 'force'. They describe the coercion of a participant into carrying out an action. Ruuli appears to lack specialized manipulative predicates like English *force* or *prevent*. Functions of constructions with these kinds of complement-taking predicates are primarily carried out by the morphological causative constructions (using verbs with the causative suffix). The best candidate for a manipulative predicate in Ruuli is *yamb* 'help', as in (33), which occurs with infinitive complements.

(33) *tu-ba-yamb-a* *[oku-lim-a* *emwani]*.
1PL.S-3PL.O-help-FV INF-cultivate-FV coffee(10)
'We help them to cultivate coffee plantations.'

Note also that it is possible to use utterance predicates in a manipulative sense as described in Section 8.5.9.

8.5.4 Desideratives

Desiderative predicates are predicates with meanings like 'want' and 'wish'. Ruuli frequently employs two desiderative predicates, viz. *tak* 'want, desire' and *yendy* 'want, like'. These predicates *tak* 'want, desire' and *yendy* 'want, like' most often introduce infinitive complements, as in (34a), but subjunctive complements, as in (34b), are also frequent. The two predicates are also used to express participant-internal necessity and in these function they were discussed and exemplified in Section 8.5.1.

(34) a. *tu-ku-taka* *[ku-ki-yindula]*.
1PL.S-PROG-want INF-7O-change
'We want to change it.'

b. *tu-ku-taka* *[mu-ta-e=wo* *elesoni]*.
1PL.S-PROG-want 2PL.S-introduce-SBJN=LOC lesson(9)
'We want you to introduce a lesson (of teaching our children).'

According to Haspelmath (1999: 41–42) there is a crosslinguistic tendency for same-subject and different-subject constructions with so-called 'want' complements to take morphosyntactically asymmetric complements. Among others, Swahili (Bantu) is given as an example of a language, where infinitives are used in same-subject constructions, while subjunctives are used in different-subject constructions. Also in Ruuli we find a correlation between same vs. different subject and the type of complement: Most of the different-subject constructions found with *tak* 'want, desire' and *yendy* 'want, like' are subjunctive, whereas same-subject constructions tend to have infinitive complements. However, there is still a non-negligible amount of examples of different-subject constructions involving infinitives and same-subject constructions involving subjunctives. Possible motivating factors for the distribution of infinitives and subjunctives are discussed further in Section 8.6.1 below.

8.5.5 Perception predicates

Perception predicates are predicates describing a physical sensation like 'see', 'hear' and 'feel'. Ruuli has at least two perception predicates, viz. *bon* 'see' and *wuur* 'hear'. They can occur with complements that express states-of-affairs and signify 'direct perception' (also called 'immediate perception'), as well as complements expressing propositions that signify information acquired via perception (known as 'indirect perception' or 'knowledge acquired') (Dik & Hengeveld 1991, Boye 2010). Indicative complements as in (35) – with or without complementizer *nti* – are most common. The complementizer *nti* can only be used to indicate an 'indirect perception'-reading.

(35) *n-wulir-a [nti ndowo mayembe a-ga-simool-a*
1SG.S-hear-FV COMP there.aren't 1.spirit REL-6S-speak-FV
o-lulimi lundi okuiyaku oluganda lwonkai].
AUG-11.language other except Ganda(11) 11.FOC
'I hear that there are no spirits that speak any other language except Ganda.'

In contrast to other predicate types, perception predicates often take complements with *nga*, but it is not yet clear whether there is a morphosyntactic or semantic explanation for the distribution of *nga*. With *wuur* 'hear' it seems that the presence and absence of *nga* contrasts direct and indirect perception as in (36a) and (36b), respectively. Example (36a) describes the perception of the sound of the wind, whereas example (36b) describes knowledge acquired through hearsay rather than the perception of sound.

(36) a. *m-puur-a [empewo nga e-ku-n-yakal-a=ku].*
1SG.S-hear-FV wind(9) COMP 9S-PROG-1SG.O-pass–FV=17LOC
'I heard/felt wind passing over me.'

b. *m-puura [a-zwamu alubaawo].*
1SG.S-hear 3SG.S-produce timber(11)
'I hear it produces timber.'

However, with *bon* 'see' we find complements that express indirect perception marked by *nga*, as in (37a), parallel to examples without a complementizer, as in (37b).

(37) a. *n-ku-ki-lool-era [nga ki-yinz-a [oku-bba eki-zibu].*
1SG.S-PROG-7O-see-APPL COMP 7S-may INF-be 7-difficult
'I am seeing (observing) that it may be difficult.'

b. *ba-bon-ire [ti-e-kya-li ya mugaso].*
3PL.S-see-PFV NEG-9S-PERS-be 9.GEN importance(3)
'They have seen they are no longer of importance.'

It is clear that perception predicates differ from other predicate classes due to their relatively frequent co-occurrence with the complementizer *nga*,

which is otherwise only found with the phasal predicate *sigal* 'remain/continue'. The contrast between direct and indirect perception seems to play a role in constructions with *wuur* 'hear', but not as much with *bon* 'see'. At present, it is not yet clear whether a comprehensive semantic analysis of *nga* is really tenable or whether the synchronic distribution of *nga* is simply what is left of a previously more widespread use.

8.5.6 Knowledge predicates

Knowledge predicates are predicates with meanings such as 'know', 'learn' and 'teach' that signify the state, acquisition or transfer of knowledge. Predicates describing knowledge of information (epistemic knowledge) or 'know how' (action knowledge) are included in this class. There are at least seven knowledge predicates in Ruuli: *many* 'know', *izukir* 'remember', *rabir* 'forget', *etejery* 'realize', *egesy* 'teach', *somesy* 'teach' and *lang* 'show'.

As a group, knowledge predicates most often take indicative complements with or without *nti* as in (38) and (39), in which case they describe epistemic knowledge.

(38) *o-maite [nti abantu ba-kom-ire muni abantu*
2SG.S-know COMP person(2) 3PL.S-select-PFV here person(2)
abakulu, ba-li Kayunga mu Katikoomu]?
elder(2) 3PL.S-be Kayunga 18LOC Katikoomu
'Do you know that people have selected elders and they are in Kayunga at Katikoomu?'

(39) *naye izukir-a [nti omwana a-a-li mwojo].*
but remember-FV COMP child(1) 3SG.S-PST-be boy(1)
'But remember that the child was a boy.'

As for the complementizer *nti*, we find an equal number of complements with and without *nti* with complements of *many* 'know' and it thus appears that it is completely optional. On the other hand, *izukir* 'remember' is more frequently used without the complementizer *nti*. Finally, complements of *lang* 'show' are always marked by *nti*, as in (40).

(40) *nga ba-tu-lang-a [nti Kyamuganwa*
CONJ 3.PL.S-1.PL.O-show-FV COMP Kyamuganwa
wa-ba-nga=wo enjazi]
16SG.S-be-HAB=LOC well(?)
'And they showed to us that at Kyamuganwa there were wells.'

Infinitives, though infrequent with knowledge predicates, are used when the complement-taking predicate means 'know how to', as in (43), 'teach how to', as in (41) and (42), and 'forget to', as in (44). Note that the same complement-taking predicate *many* is used in (38) and (43). It can be seen how the choice of complement type (indicative vs. infinitive) makes a difference to the meaning of the complement construction (epistemic knowledge vs. know how).

(41) *omuwala tu-mu-egesy-a [oku-lamuc-a].*
girl(1) 1PL.S-SG.O-teach-FV INF-greet-FV
'We teach the girl to greet.'

(42) *ba-ku-somesy-e abantu [oku-yindul-a edikisonare*
3PL.S-PROG-teach-SBJN person(2) INF-translate-FV dictionary(9)
ba-gi-ta-e omu lulimi lwa-aiswe oLunyala].
3PL.S-9O-put-SBJN 18.LOC language(11) 11-1PL.POSS Lunyala(11)
'They may teach people how to translate a dictionary into Lunyala, our language.'

(43) *omunyala y-a-maite [oku-ly-a, oku-lisy-a amaka].*
Munyala(1) 3SG.S-PST-know.PFV INF-eat-FV INF-feed-FV home(6)
'A Munyala knew how to eat, how to feed the home.'

(44) *ni tw-erabir-a [oku-sumb-isy-a].*
CONJ 1PL.S-forget-FV INF-cook-APPL-FV
'And we forgot how to cook food (with firewood).'

8.5.7 Propositional attitude predicates

Propositional attitude-predicates are predicates with meanings like ‘think’ and ‘believe’. There are at least two propositional attitude predicates in Ruuli: *lowooz* ‘think’ and the polysemous *ikiriz*, which means ‘accept, agree to, allow, believe, permit’. Propositional attitude predicates most often occur with indicative complements with or without *nti*, as in (45b), and (45a).

(45) a. *n-ku-lowooz-a [nti tu-ku-tandik-a=wo*
1SG.S-PROG-think-FV COMPL 1PL.S-PROG-start-FV=LOC
eprojects].
project(?)
‘I was thinking that we start up projects.’

b. *n-lowooz-a [ba-ku-fun-a=mu kidooli]*.
1SG.S-think-FV 3PL.S-PROG-get-FV=18LOC little
‘I think they benefit little.’

(46) *n-kw-ikiriz-a [nti wa-bba-a=wo=ku eibbaale]*.
1SG.S-PROG-believe-FV COMPL 16SG.S-be-FV stone(5/6)
‘I believe that there will be a stone’

Both *lowooz* and *ikiriz* also occur with infinitive complements.

The infinitive is possible with *lowooz* ‘think’ when the complement expresses something that is planned to happen as in (47).

(47) *a-lowooz-a [oku-aba n’-omwibi]*.
3SG.S-think-FV INF-go COM-thief(1)
‘He thinks (i.e. intends) to go with the thief.’

Thus, the complement-taking predicate in (47) does not actually express an ‘attitude’ towards a proposition as the name of the class would suggest. It means ‘intend’ rather than ‘think’. However, we chose to keep the class name to mirror the ones used in Cristofaro (2003).

The infinitive is very common with *ikiriz* and then *ikiriz* has the meaning of ‘allow, permit’ as in (48).

(48) *ekuruhani e-kiriz-a [oku-gul-a omukazi].*
Koran(9/10) 9S-accept-FV INF-buy-FV woman(1)
'The Koran accepts (i.e. permits) buying a woman.'

8.5.8 Emotive predicates

Emotive predicates are predicates describing an emotion with meanings such as 'hate', 'love' etc. There are at least four emotive predicates in Ruuli: *tiin* 'fear, be afraid' (49), *sanyuk* 'be happy, be glad' (50), *semeer* 'be happy, be glad' (51), and *eyanz* 'appreciate, thank, be grateful' (52). Emotive predicates always occur with infinitives.

(49) *nje n-tiin-a [oku-nyakalya omu kanwa*
1SG 1SG.S-be.afraid-FV INF-pass-FV 18LOC mouth(12)
kange].
12-1SG.POSS
'As for me, I fear to pass it through my mouth.'

(50) *okyooto ni-a-sanyuk-a [oku-fun-a abaana].*
Kyooto(1) NAR-3SG.S-be.happy-FV INFget-FV child(2)
'And Kyooto was happy to get children.'

(51) *njeena n-semereirwe [oku-bb-a ani].*
1SG.S 1SG.S-be.happy.PFV INF-be-FV here
'I am also glad to be here.'

(52) *n-eyanz-ire muno [oku-ku-sang-a].*
1SG.S-appreciate-PFV much INF-2SG.O-meet-FV
'I am pleased to meet you.'

8.5.9 Utterance predicates

Utterance predicates have meanings such as 'say', 'ask' and 'tell' and signify information transfer, information requests or orders and directions delivered by means of speech. There is more than a dozen utterance predicates in Ruuli,

the most common are *kob* ‘say, tell’, *buuly* ‘ask (a question)’, *sab* ‘ask (for), request, pray (for)’ and *gaan* ‘refuse, reject’.

In principle, it is possible to distinguish between direct and indirect reported speech in Ruuli, although many tokens in the corpus are ambiguous due to ambiguous reference of deictic elements. The most reliable criteria are shifts in pronominal reference, as in (53a) vs. (53b), as well as the use of imperatives. Temporal deixis is less reliable. Further investigations are needed to assess the reliability of spatial deixis and of other discourse features in distinguishing between the two types of reported discourse. (For the discussion of criteria for the distinction between direct and indirect (see e.g. Güldemann 2008: 27–28.) Direct speech is especially frequent with *kob* ‘say, tell’ and *buuly* ‘ask (a question)’.

Utterance predicates are the most versatile of all complement-taking predicates in Ruuli. For instance, the most common utterance predicate by far *kob* ‘say, tell’ occurs with direct and indirect reports of assertions, questions and commands. Structurally, it can occur with indicatives, subjunctives and infinitives. In what follows we provide examples of direct and indirect reports of assertions, direct and indirect reports of questions, as well as of direct and indirect reports of commands introduced by utterance predicates.

Reported assertions are often introduced by *nti*, but the presence or absence of *nti* is not indicative of the complement being a direct or indirect report, cf. (53a) and (53b). Direct speech complements are marked by *nti* slightly more often than indirect speech complements.

(53) a. Direct report of assertion
a-ku-kob-a [nje n-li musajja wa kabaka].
3SG.S-2SG.O-tell-FV 1sg 1SG.S-be man(1) 1.GEN king(1)
‘He tells you, “I am Kabaka’s man”.’

b. Indirect report of assertion
a-a-salewo ku-kob-a [nti wakiri a-zw-e=wo
3SG.S-PST-decide INF-say-FV COMP at.least 3SG.S-leave=16LOC
omu maka a-yab-e].
18.LOC home(6) 3SG.S-go-SBJN
‘He decided to say that he’d rather leave home and go away.’

Indicative complements of utterance predicates marked by *nti* do not have any special characteristics in comparison to indicative complements of other complement-taking predicates that co-occur with *nti* (perception, knowledge and propositional attitude predicates).

Direct reports of polar questions (yes/no-questions) do not have any special characteristics distinguishing them from reported assertions. This follows naturally from the fact that independent clause polar questions are morphosyntactically identical to independent clause assertions in Ruuli. Direct reports of questions can be marked by *nti* as in (54).

(54) Direct report of question
ni-a-n-buuly-a *[nti* *mwana* *wa-ange*
NAR-3SG.S-1SG.O-ask-FV COMP child(1) 1-1SG.POSS
o-yab-ire=ku *omu* *nkwi].*
2SG.S-go-PFV=17LOC 18LOC firewood(10)
'She asked me, "My child have you ever gone to collect firewood?"'

Indirect reports of questions, as in (55) are introduced by *oba*. In such cases *oba*, which is otherwise used as a conjunction 'or' in Ruuli, functions as a complementizer somewhat similar in meaning to English *whether* (cf. Section 8.4.5). The occurrence of a morpheme functioning both as a conjunction 'or' and as a complementizer used to report polar questions is not surprising: Meanings such as 'or' are related to uncertainty and uncertainty is related to polar questions (cf. Boye 2012).

(55) Indirect report of question
tu-kol-e *ekintu* *nga* *na-opurezidenti* *tu-ku-mu-sab-a*
1PL.S-do-SBJN thing(7) CONJ ?-president(1) 1PL.S-PROG-3SG.O-ask-FV
[oba *a-ku-tu-weery-a=yo* *akaseera* *ka-dyoli].*
whether 3SG.S-PROG-1PL.O-give-FV=23LOC time(12) 12-little
'Let us do something as we ask the president whether he gives us a moment.'

In direct reports of commands and requests the imperative form of the verb (stem and final vowel, no subject indexing) is used, exactly as in the imperative main clause.

(56) Direct report of command/request
ni-a-kob-a *[seny-a* *enkwi]*.
NAR-3SG.S-say-FV collect-FV firewood(10)
'She said, "Collect firewood!"'

Indirect reports of commands are achieved by employing a subjunctive verb form as in (57) or – less frequently – by an infinitive.

(57) Indirect report of command/request
naye *tu-a-mu-kob-ire* *[a-tu-weery-e=yo*
but 1PL.S-PST-3SG.O-say-PFV 3SG.S-1PL.O-give-SBJN=23LOC
omusaayi].
blood(3)
'But we told him to give us some blood.'

8.5.10 Other complement-taking predicates

In this final section we provide some examples of less frequent complement-taking predicates *yombok* 'struggle' and *byal* 'thank'. These predicates occur with infinitive complements.

(58) *tu-ku-gad-a* *[oku-yombok-a* *olubiri]*.
1PL.S-PROG-struggle-FV INF-build-FV palace(11)
'We are struggling to build a palace.'

(59) *webale* *[ku-byal-a]*.
thank.you INF-give.birth-FV
'Thank you for giving birth (to children).'

8.6 Discussion

In the previous sections we have pointed out two main predictors of complement choice with different predicate classes: 1) different-subject and same-subject constructions and 2) the semantic distinction between states-of-affairs and propositions. This section summarizes the most important analytic points and discusses the predictions further.

8.6.1 Comparison of the subjunctive and the infinitive: different-subject vs. same-subject constructions

As described in Section 8.4, subjunctive complements follow a limited number of predicates, namely modals, phasals, desideratives, and utterance predicates, but infinitives combine with all predicate classes except for perception predicates. Some complement-taking predicates in Ruuli thus allow both infinitive and subjunctive complements.

Several other Bantu languages have complement-taking predicates that can introduce both subjunctive and infinitive complements, but the motivation for choosing one over the other is likely language-specific. For Bemba (East Bantu, M.42; Zambia), for example, Givón (1969: 224) has noted that the meaning difference between subjunctive and infinitive complements of manipulative predicates is that with subjunctives the event in the complement may or may not have happened, whereas with infinitives it has happened (at some point). In Nzadi (Narrow Bantu, B.30; DR Congo) infinitives are generally used in same-subject constructions, while subjunctives are used in different-subject constructions (Crane et al. 2011: 180-182). Such an analysis does not comprehensively account for the distribution of infinitives and subjunctives in Ruuli. Realis status does not seem to be a motivating factor at all, and while the account in terms of the distinction between different-subject and same-subject constructions does indeed motivate the distribution to some degree, it cannot stand on its own. Complements in different-subject constructions with desiderative predicates, for example, are almost always subjunctive, but same-subject complements are not always infinitives as shown in (60a).

(60) a. *a-ku-yendy-a [a-yab-e oku university].*
3SG.S-PROG-want-FV 3SG.S-go-SBJN 17.LOC university
'She wants to go to university.'

b. *nje n-ku-yendy-a [ku-som-a catering].*
1SG 1SG.S-PROG-want-FV INF-study-FV catering
'For me, I want to study catering'

It is not yet entirely clear what motivates the use of a subjunctive instead

of an infinitive in same-subject constructions, but one possibility is that the choice is motivated by a rather subtle semantic difference: when *yendy* introduces an infinitive complement as in (60b) it resembles English 'want', whereas when *yendy* introduces subjunctive complements the meaning of *yendy* is more akin to 'would like to', as in (60a).

8.6.2 States-of-affairs and propositions

In Section 8.2 above we discussed the connection between semantic contrasts and morphosyntactic contrasts in complementation, in particular the contrasts between states-of-affairs and propositions as found in complement constructions with perception predicates (direct perception vs. indirect perception), knowledge predicates (epistemic knowledge vs. action knowledge/know how) and utterance predicates (reports of assertions or questions vs. reports of commands/requests). In this section we relate the contrast between state-of-affairs and propositions to Ruuli complementation.

Several tests for identifying whether a complement clause expresses a state-of-affairs or a proposition have been proposed in the literature. Among them are the distribution of complements on specific predicate types and the acceptability of epistemic modification of the complement (e.g. insertion of *maybe* in English) (for discussions of appropriate tests cf. Boye 2012, Serdobolskaya 2016). We will focus on the distribution of complement types and complementizers over predicate classes.

In Ruuli the contrast between propositions and states-of-affairs is found within different complement-taking predicate classes, namely knowledge predicates, propositional attitude predicates, utterance predicates and perception predicates.

For example, (61a) and (61b) illustrate the contrast between propositional and state-of-affairs complements in Ruuli knowledge-predicate complementation. The indicative complement of *mait* 'know' in (61a) represents epistemic knowledge, whereas the infinitive complement of *mait* 'know' in (61b) designates 'know how' (note that the infinitive construction does not have any element corresponding to English *how*).

(61) a. Proposition (epistemic knowledge)
o-maite [ye-ena a-yendy-a oku-yizukiry-a].
2SG.S-know 3SG-FOC 3SG.S-need-FV INF-be.reminded-FV
'You know he also needs to be reminded.'

b. State-of-affairs (action knowledge/know how)
a-maite [oku-sany-a].
3SG.S-know.PFV INF-swim-FV
'He knew how to swim.'

A parallel contrast from utterance-predicate complementation is given in example (62). In this case too, a morphosyntactic contrast is accompanied by a semantic contrast. Example (62a) illustrates an indicative complement (marked by the complementizer *nti*) contrasting with a state-of-affairs complements (a subjunctive) in (62b)

(62) a. Proposition (report of assertion)
abaganda ni-ba-kob-a [nti bamafumbe
Baganda(2) NAR-3PL.S-say-FV COMP African.civet.clan(2)
ni-bo ba-byal-a oKawumpuli]
COP-3PL 3PL.S-produce-FV Kawumpuli
'And the Baganda said that the African Civet clan are the mothers of Kawumpuli.'

b. State-of-affairs (report of command/request)
ni-ba-kob-a [akiri tu-byal-e].
NAR-3PL.S-say-FV at.least 1PL.S-give.birth-SBJN
'They said, (that) we should at least produce.'

The morphosyntactic and semantic contrasts in examples (61), (62) as well as those found with propositional attitude predicates (propositional attitude vs. intention) and perception predicates (direct vs. indirect perception) (see sections 8.5.7 and 8.5.5) are comparable to e.g. English contrasts such as *She knows that he plays the piano* vs. *She knows how to play the piano* (see Section 8.2.5). As discussed in Section 8.5.5, the picture is somewhat less clear for perception-predicate complementation, a predicate class that has otherwise received much attention in studies of contrasts between

states-of-affairs and propositions (specifically direct and indirect perception) in relation to other languages (Dik & Hengeveld 1991, Schüle 2000, Boye 2010).

Indicative complements and the complementizers *nti* and *oba* occur in constructions where the complement is arguably propositional (for an example with *oba* see Section 8.5.9). When the complement is indicative the complement-taking predicate functions as the kind of predicates that have been related to propositions, e.g. epistemic knowledge predicates, propositional attitude predicates and assertive utterance predicates (Cristofaro 2003, 2013, Noonan 2007, Boye 2012, Sørensen 2013, Sørensen & Boye 2015). Infinitives and subjunctives on the other hand generally express states-of-affairs. Non-epistemic modals as well as phasals, desideratives and directive utterance predicates, which have been related to state-of-affairs (Cristofaro 2003, Noonan 2007, Boye 2012), occur with infinitives and/or subjunctives and do not occur with complementizers.

On the basis of the available data, the contrast between states-of-affairs and propositions can be said to be expressed quite systematically in Ruuli. Infinitive and subjunctive complements express states-of-affairs in contrast to indicative complements which express propositions. Thus, no one complement type appears to be completely polyfunctional between the two readings as is the case in some languages (Boye 2010: 407) – with the exception that epistemic modals can take infinitives expressing a proposition (but note that epistemic modals are not included in the class of modals in Cristofaro 2003, Noonan 2007). But there is tendency for complement-taking predicates to be polyfunctional and take more than one type of complement.

8.7 Conclusion and outlook

We have made a first attempt at a description of the morphosyntax and semantics of clausal complementation in Ruuli. We have identified and characterized the main complement-taking predicates and complement types and discussed the distribution of complements and complementizers. We have also discussed semantic features motivating complement structures, in particular the contrast between states-of-affairs and propositions.

Ruuli has a wide range of complement-taking predicates spanning most of the semantic classes identified crosslinguistically (Givón 2001, Cristofaro 2003, Dixon 2006, Noonan 2007). As in many other Bantu languages the main complement types are indicatives, subjunctives and infinitives, marked by final vowel *-a*, final vowel *-e* and class marker *(o)ku-*, respectively. We found that the infinitive affix has the form *ku* when the complement-taking predicate is negated and that there is free variation between *oku* and *ku* when the complement-taking predicate is not negated.

The most frequent complementizer is *nti*, which optionally introduces indicative complements, followed by *nga*, which typically follows perception predicates, and *oba* which indicates doubt towards the proposition in the complement.

We have shown that many types of complement-taking predicates systematically take morphosyntactically different complement types according to whether the complement expresses a proposition or a state-of-affairs. The classes of complement-taking predicates which have been associated with state-of-affairs complements crosslinguistically, like modals and phasals, take infinitive complements in Ruuli, while other predicate classes like knowledge predicates and utterance predicates take both indicative and infinitive complements with an associated semantic contrast between propositions and state-of-affairs. Indicative complements are generally used to express propositions while subjunctive and infinitive complements express state-of-affairs (except perhaps for perception-predicate complementation). The distribution of complement types relative to predicate types thus appears to be semantically motivated by the contrast between states-of-affairs an propositions and thus adds evidence to already observed crosslinguistic trends.

Similar complement contrasts, such as contrasts between indicatives and infinitives used to describe epistemic knowledge and ‘know how’, respectively, are found in existing descriptions of Bantu languages. However, future studies might benefit from systematically looking for possible contrasts with all predicate classes although they might be rare.

Chapter 9

Conclusion

In this thesis, I have studied the crosslinguistic expression of the semantic contrast between states-of-affairs and propositions in clausal complementation. In this chapter, I will briefly summarize and discuss the main results of the thesis in relation to the research questions presented in the introduction: 1) To what extent does the semantic contrast between states-of-affairs and propositions motivate grammatical contrasts in clausal complementation? 2) What is the status of reported speech in a typology of complementation based on the contrast between states-of-affairs and propositions? and 3) How can the contrast between states-of-affairs and propositions be used as a point of departure for describing the system of complementation in a specific language?

In Section 9.1 I discuss the role of states-of-affairs and propositions in the analysis of reported speech, in Section 9.2 I discuss states-of-affairs and propositions as motivations for grammatical contrasts and in Section 9.3 I discuss how the distinction may be used in language-specific descriptive work.

9.1 States-of-affairs and propositions and reported speech

Article I focused on utterance-predicate complementation. We proposed an analysis of the differences between independent main clauses, independent complement clauses used for direct utterance report, and dependent comple-

ment clauses used for indirect utterance report, based on layered semantic structure.

We first argued that the layer of illocution can be divided into a neustic and a tropic based on a modified version of Hare's (1970) understanding of these notions. We then argued that the main difference between independent main clauses and independent complement clauses is the lack of a neustic in the latter. We also argued that the main difference between independent complement clauses and dependent complement clauses is that the latter lacks a tropic. It follows from the analysis that it is possible to study the expression of states-of-affairs and propositions in independent main clauses as well as in complements used for direct and indirect reported speech.

In Article II we focused on complements of utterance predicates used for indirect utterance report. Based on the analysis in Article I, we assumed that the extent to which states-of-affairs and propositions motivate grammatical contrasts could best be investigated in dependent complements, because it would be easier to ascertain the coding of states-of-affairs and propositions in the absence of a neustic and a tropic.

In Article III we studied both independent and dependent complements to achieve a more comprehensive analysis of the entire complementation system in Ruuli.

9.2 States-of-affairs and propositions and grammatical structure

On the basis of the analysis of reported speech, we also proposed a hypothesis in Article I regarding the grammaticalization of utterance predicates into reportative evidentials. We proposed that utterance verbs that function as complement-taking predicates of propositional complements can grammaticalize into reportative evidentials. However, the collection of extensive crosslinguistic data supporting the hypothesis was out of the scope of the thesis, since it would have required a different set of data. The reference grammars used for the data collection did not always include information of the grammaticalization of utterance predicates (and as is well-known, this kind of information is not even available for all languages). Testing the hy-

pothesis would require gathering evidence directly from language experts, e.g. through questionnaires. It is nevertheless clearly the case that the contrast between states-of-affairs and propositions should be considered in diachronic studies (and as previously mentioned, this has been done in recent work such as Kehayov 2017).

In Article II, we presented a crosslinguistic study of utterance-predicate complementation, comparing the morphosyntax of assertive, interrogative and directive complements as well as grammatical contrasts between them. We looked at two main variables: balanced and deranked verb forms and complementizers. Then we discussed the merits and limitations of previous accounts of such grammatical contrasts: iconicity of cohesion (e.g. Givón 2001, Cristofaro 2003) and frequency (e.g. Haspelmath 2008). We reached the conclusion that frequency of use seems to be of limited value, but that iconicity of cohesion makes some correct predictions. Finally, we argued that the hypothesis that propositions are conceptually more complex than states-of-affairs (as proposed by Boye 2010, 2012) to some extent has predictive value regarding the distribution of balanced and deranked verbs forms and complementizers and that this can be described as a case of iconicity of complexity.

The analyses presented in Articles I and II thus suggest that the conceptual nature of states-of-affairs and propositions may influence grammatical structure. Article I suggested that the cooccurence of propositional complements with certain complement-taking predicates may influence their potential grammaticalization path and Article II suggested that the distinction between states-of-affairs and propositions may motivate the grammatical structure of utterance-predicate complements with regards to the distribution of deranked and balanced complements and complementizers.

9.3 States-of-affairs and propositions and language description

In Article III, we presented af first analysis of the morphosyntactic and semantic features of clausal complementation in the Bantu language Ruuli. One of the motivations behind the article was to see whether the contrast

between states-of-affairs and propositions might play a role in Ruuli complementation.

It turned out that there were indeed systematic grammatical contrasts between complements expressing states-of-affairs and propositions within different predicate classes. Interestingly, state-of-affairs complements were quite infrequent with some types of predicates, such as knowledge predicates, and complements expressing states-of-affairs might thus have been overlooked, were it not for the attempt to identify contrasts between states-of-affairs and propositions. This suggests the need for more language-specific descriptions of the expression of states-of-affairs and propositions in complementation. Crosslinguistic studies, such as the one in this thesis, can serve to move the attention of researchers to the contrast between states-of-affairs and propositions and the variety of ways in which this contrast can be expressed crosslinguistically.

9.4 Concluding remarks

The three main aims of the thesis have been reached. We have developed a functional theory of the semantics of reported speech based on the contrast between states-of-affairs and propositions, we have substantiated the claim that a difference in conceptual complexity between propositions and states-of-affairs motivates grammatical contrasts in clausal complementation and we have provided a first description of clausal complementation in Ruuli. The collective results of the thesis suggest that the distinction between states-of-affairs and propositions needs more attention in linguistics and that it is especially relevant to crosslinguistic and language-specific studies of clausal complementation. The thesis has thus contributed new knowledge to the literature on the interplay between the morphosyntax and semantics of complementation in general and to the study of the role of states-of-affairs and propositions in complementation in particular.

Bibliography

Abbott, Barbara. 2013. Linguistic solutions to philosophical problems. *Philosophical Perspectives* 27(1). 1–21.

Achard, Michel. 1998. *Representation of cognitive structures*. Berlin/New York: Mouton de Gruyter.

Achard, Michel. 2002. The meaning and distribution of French mood inflections. In Frank Brisard (ed.), *Grounding: The epistemic footing of deixis and reference*, 197–249. Berlin: Mouton de Gruyter.

Achard, Michel. 2010. Complementation. In Dirk Geeraerts & Hubert Cuyckens (eds.), *The Oxford handbook of cognitive linguistics*. Oxford: Oxford University Press.

Aikhenvald, Alexandra. 2010. *Imperatives and commands* (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press.

Akatsuka, Noriko. 1985. Conditionals and the epistemic scale. *Language* 61(3). 625–639.

Ashton, Ethel O. et al. 1954. *A Luganda grammar*. London: Longmans, Green & Company.

Bolinger, Dwight. 1968. Entailment and the meaning of structures. *Glossa* 2(2). 119–127.

Bowden, John. 1997. *Taba (Makian Dalam): Description of an Austronesian language from Eastern Indonesia*. Melbourne: The University of Melbourne dissertation.

Boye, Kasper. 2008. Om *om*. *Ny Forskning i Grammatik* 15. 7–29.

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

Boye, Kasper. 2012. *Epistemic meaning: A crosslinguistic and functional-cognitive study*. Berlin: De Gruyter Mouton.

Boye, Kasper. 2013. Sagforhold, propositioner og den danske indikativ. *Ny Forskning i Grammatik* 20. 31–52.

Boye, Kasper & Peter Harder. 2007. Complement-taking predicates: Usage and linguistic structure. *Studies in Language* 31(3). 569–606.

Boye, Kasper & Peter Harder. 2009. Linguistic categories and grammaticalization. *Functions of Language* 16(1). 9–34.

Boye, Kasper & Peter Harder. 2012. A usage-based theory of grammatical status and grammaticalization. *Language* 88(1). 1–44.

Boye, Kasper & Petar Kehavov. 2016. *Complementizer semantics in European languages* (Empirical approaches to language typology 57). Berlin: De Gruyter Mouton.

Boye, Kasper, Eva van Lier & Eva Theilgaard Brink. 2015. Epistemic complementizers: a crosslinguistic survey. *Language Sciences* 51(5). 1–17.

Boye, Kasper et al. 2012. At eller ikke at i tale og skrift. *NyS* 42. 41–61.

Burling, Robbins. 2003. *Language of the Modhupur Mandi (Garo)*. Michigan: Michigan Publishing, University of Michigan Library.

Bybee, Joan. 1998. 'Irrealis' as a grammatical category. *Anthropological Linguistics* 40(2). 257–271.

Bybee, Joan & Paul Hopper. 2001. Introduction to frequency and the emergence of linguistic structure. In Joan Bybee & Paul Hopper (eds.), *Frequency and the emergence of linguistic structure*, 1–24. Amsterdam: John Benjamins.

Chafe, Wallace. 1995. The realis-irrealis distinction in Caddo, the Northern Iroquoian languages and English. In Joan Bybee & Suzanne Fleischmann (eds.), *Modality in grammar and discourse*, 349–365. Amsterdam: John Benjamins.

Chang, Anna Hsiou-chuan. 2006. *A reference grammar of Paiwan*. Canberra: Australian National University dissertation.

Cole, Peter. 1982. *Imbabura Quechua*. Amsterdam: North-Holland.

Comrie, Bernard, Martin Haspelmath & Balthasar Bickel. 2008. *The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme*

glosses. http://www.eva.mpg.de/lingua/resources/glossing-rules.php (22 August, 2018).

Comrie, Bernard et al. 2013. Introduction. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/s1 (8 September, 2018).

Coulmas, Florian. 1986. *Direct and indirect speech*. Berlin; New York: De Gruyter Mouton.

Crane, Thera Marie, Larry M. Hyman & Simon Nsielanga Tukumu. 2011. *A grammar of Nzadi* (University of California Publications in Linguistics 147). Berkely and Los Angeles, CA: University of California Press.

Cristofaro, Sonia. 1998. *Subordination strategies: A typological study*. Pavia: University of Pavia dissertation.

Cristofaro, Sonia. 2003. *Subordination* (Oxford studies in typology and linguistic theory). Oxford: Oxford University Press.

Cristofaro, Sonia. 2013. Utterance complement clauses. In *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://wals.info/chapter/128 (22 August, 2018).

Cristofaro, Sonia. 2014a. Cognitive explanations, distributional evidence and diachrony. In Nikolas Gisborne & Willem B. Hollmann (eds.), *Theory and data in cognitive linguistics* (Benjamins Current Topics 64), 185–210. Amsterdam: John Benjamins.

Cristofaro, Sonia. 2014b. Is there really syntactic subordination? In Laura Visapää, Kallioskoski Jyrki & Sorva Helena (eds.), *Contexts of subordination*, 73–91. Amsterdam: John Benjamins.

Croft, William. 2003. *Typology and universals*. Cambridge: Cambridge University Press.

Crystal, David. 2008. *A dictionary of linguistics and phonetics*. Oxford: Blackwell.

Dench, Alan Charles. 1995. *Martuthunira: a language of the Pilbara region of Western Australia* (Pacific Linguistics, Series C, 125). Canberra: Australian National University.

Derbyshire, Desmond C. 1979. *Hixkaryana* (Lingua Descriptive Studies 1). Amsterdam: North-Holland.

Dik, Simon C. 1997. *The theory of functional grammar* (Functional Grammar Series 20). Berlin: Mouton de Gruyter.

Dik, Simon C. & Kees Hengeveld. 1991. The hierarchichal structure of the clause and the typology of perception-verb complements. *Linguistics* 21. 231–259.

Dixon, R. M. W. 1988. *A grammar of Boumaa Fijian*. Chicago: University of Chicago Press.

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

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

Dixon, R. M. W. 2010. *Basic linguistic theory*. Vol. 2. Oxford: Oxford University Press.

Dixon, R. M. W. & Alexandra Y. Aikhenvald. 2006. *Complementation: A cross-linguistic typology*. Oxford: Oxford University Press.

Donohue, Mark. 1999. *A grammar of Tukang Besi*. Berlin: Mouton de Gruyter.

Dryer, Matthew S. & Martin Haspelmath (eds.). 2013. *Wals online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. `http://wals.info/` (22 August, 2018).

Dunn, Michael John. 1999. *A grammar of Chukchi*. Canberra: Australian National University dissertation.

Elliot, Jennifer R. 2000. Realis and irrealis: forms and concepts of the grammaticalization of reality. *Linguistic Typology* 4(1). 55–90.

Engberg-Pedersen, Elisabeth et al. (eds.). 1996. *Content expression and structure: Studies in Danish Functional Grammar*. Amsterdam: John Benjamins.

Engberg-Pedersen, Elisabeth et al. (eds.). 2005. *Dansk Funktionel Lingvistik: en helhedsforståelse af forholdet mellem sprogstruktur, sprogbrug og kognition*. Copenhagen: Copenhagen University.

Evans, Nicholas. 1995. *A grammar of Kayardild, with historical-comparative notes on Tangkic* (Mouton Grammar Library 15). Berlin; New York: Mouton de Gruyter.

Evans, Nicholas. 2012. Some problems in the typology of quotation: A canonical approach. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*. Oxford: Oxford University Press.

Evans, Nicholas & Honoré Watanabe (eds.). 2016. *Insubordination* (Typological Studies in Language 115). Amsterdam: John Benjamins.

Evans, Vyvyan. 2009. Semantic representation in LCCM theory. In Vyvyan Evans & Stephanie Pourcel (eds.), *New directions in cognitive linguistics*, 27–55. Amsterdam: John Benjamins.

Everett, Daniel L. 1986. Pirahã. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, vol. 1, 200–325. Berlin: Mouton de Gruyter.

Exter, Mats. 2012. 'realis' and 'irrealis' in Wogeo: a valid category?: Challenges for the 21st century. In Nicolas Evans & Marian Klamer (eds.), *Melanesian languages on the edge of Asia* (Language Documentation & Conservation Special Publication No. 5), 174–190.

Fabiszak, Małgorzata, Martin Hilpert & Karolina Krawczak. 2016. Usage-based cognitive-functional linguistics: From theory to method and back again. *Folia Linguistica* 50(2). 345–353.

Fagard, Benjamin, Paola Pietrandrea & Julia Glikman. 2016. Syntactic and semantic aspects of Romance complementizers. In Kasper Boye & Petar Kehayov (eds.), *Complementizer semantics in European languages*, vol. 57 (Empirical approaches to language typology), 75–130. Berlin: De Gruyter Mouton.

Foley, William A. 1986. *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.

Foley, William A. & Robert D. Van Valin. 1984. *Functional syntax and universal grammar* (Cambridge studies in linguistics 38). Cambridge: Cambridge University Press.

Forker, Diana. 2016. Complementizers in Hinuq. In Kasper Boye & Petar Kehayov (eds.), *Complementizer semantics in European languages* (Empir-

ical approaches to language typology 57), 745–792. Berlin: De Gruyter Mouton.

Frajzyngier, Zygmunt. 1991. The de dicto domain in language. In Elisabeth Closs Traugoot & Bernd Heine (eds.), *Approaches to grammaticalization*, vol. I: Focus on theoretical and methodological issues, 219–251. Amsterdam: John Benjamins.

Frajzyngier, Zygmunt. 1995. A functional theory of complementizers. In Joan Bybee & Suzanne Fleischman (eds.), *Modality in grammar and discourse*, 473–502. Amsterdam: John Benjamins.

Frajzyngier, Zygmunt & Robert Jasperson. 1991. *That*-clauses and other complements. *Lingua* 83. 133–153.

Gallois, Andre. 1998. De re/de de dicto. In *Routledge encyclopedia of philosophy*. Taylor & Francis. `https://www.rep.routledge.com/articles/thematic/de-re-de-dicto/v-1` (22 August, 2018).

Gentens, Caroline. 2016. *The factive-reported distinction in English. Representational and interpersonal semantics*. Leuven: KU Leuven dissertation.

Gijn, Rik van & Sonia Gipper. 2009. Irrealis in Yurakare: On the crosslinguistic consistency of an elusive category. In *Cross-linguistic semantics of tense, aspect, and modality* (Linguistics Today), 155–178. Amsterdam: John Benjamins.

Givón, Talmy. 1969. *Studies in Chibemba and Bantu grammar*. Los Angeles: University of California, Los Angeles dissertation.

Givón, Talmy. 1980. The binding hierarchy and the typology of complements. *Studies in Language* 4. 333–77.

Givón, Talmy. 1982. Evidentiality and epistemic space. *Studies in Language* 6(1). 23–49.

Givón, Talmy. 1990. *Syntax: A functional-typological introduction*. Vol. 2. Amsterdam: John Benjamins.

Givón, Talmy. 1994. Irrealis and the subjunctive. *Studies in language* 18(2). 265–337.

Givón, Talmy. 2001. *Syntax: Vol I: a functional-typological introduction*. revised ed. Amsterdam: John Benjamins.

Güldemann, Tom. 2008. *Quotative indexes in African languages: A synchronic and diachronic survey* (34). Mouton de Gruyter.

Haan, Ferdinand de. 2012. Irrealis: Fact or fiction? *Language Sciences* 34. 107–130.

Haberland, Hartmut. 1986. Reported speech in Danish. In Florian Coulmas (ed.), *Direct and indirect speech*, 219–253. Berlin; New York: De Gruyter Mouton.

Haiman, John. 1985. *Natural syntax: Iconicity and erosion*. Cambridge: Cambridge University Press.

Haiman, John. 2008. In defence of iconicity. *Cognitive Linguistics* 19(1). 35–48.

Haiman, John. 2011. *Cambodian Khmer* (London Oriental and African language library 16). Amsterdam: John Benjamins.

Hammarström, Harald, Robert Forkel & Martin Haspelmath. 2018. *Glottolog 3.3*. Max Planck Institute for the Science of Human History. Jena. `http://glottolog.org/` (22 August, 2018).

Hansen, Björn, Alexander Letuchiy & Izabela Błaszczyk. 2016. Complementizers in Slavonic (Russian, Polish, and Bulgarian). In Kasper Boye & Petar Kehayov (eds.), *Complementizer semantics in European languages* (Empirical approaches to language typology 57). Berlin: De Gruyter Mouton.

Hansen, Erik & Lars Heltoft. 2011. *Grammatik over det danske sprog*. Copenhagen: Det danske sprog- og litteraturselskab.

Harder, Peter. 1996. *Functional semantics: A theory of meaning, structure and tense in English*. Berlin: Mouton de Gruyter.

Harder, Peter. 2005. Lagdelt sætningsstruktur. In Elisabeth Engberg-Pedersen et al. (eds.), *Dansk funktionel lingvistik: En helhedsforståelse af forholdet mellem sprogstruktur, sprogbrug og kognition*, 101–111. Copenhagen: University of Copenhagen.

Harder, Peter. 2007. Shaping the interactive flow: Language as input, process and product. *Acta Linguistica Hafniensia* (39). 7–36.

Harder, Peter. 2009. Meaning as input: The instructional perspective. In Vyvyan Evans & Stéphanie Pourcel (eds.), *New directions in cognitive linguistics*, 15–26. Amsterdam: John Benjamins.

Harder, Peter. 2010. *Meaning in mind and society: A functional contribution to the social turn in cognitive linguistics* (Cognitive linguistics research 41). De Gruyter Mouton.

Harder, Peter. 2016. Complementizers and the complementizing function: Structure and pragmatics in panchronic perspective. In Kasper Boye & Kehayov Petar (eds.), *Complementizer semantics in European languages* (Empirical approaches to language typology 57), 793–808. Berlin: De Gruyter Mouton.

Hare, R. M. 1970. Meaning and speech acts. *The Philosophical Review* 79(1).

Harms, Phillip Lee. 1994. *Epena Pedee syntax* (Studies in the languages of Colombia 4). Summer Institute of Linguistics: University of Texas at Arlington.

Haspelmath, Martin. 1995. The converb as a cross-linguistic valid category. In *Converbs in cross-linguistic perspective*, 1–56. Berlin/New York: Mouton de Gruyter.

Haspelmath, Martin. 1999. On the cross-linguistic distribution of same-subject and different-subject complement clauses: Economic vs. iconic motivation. *ICLC, Stockholm*. (18 August, 2015).

Haspelmath, Martin. 2008. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19(1). 1–33.

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

Haspelmath, Martin. 2017. Explaining alienability contrasts in adpossessive constructions: Predictability vs. iconicity. *Zeitschrift für Sprachwissenschaft* (36). 193–231.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Hengeveld, Kees. 1989. Layers and operators in Functional Grammar. *Journal of Linguistics* 25(1). 127–157.

Hengeveld, Kees. 1990. The hierarchichal structure of utterances. In Jan Nuyts, A. Machtelt Bolkestein & Vet Co (eds.), *Layers and levels of representation in language theory: A functional view*. Amsterdam: John Benjamins.

Hengeveld, Kees & J. Lachlan Mackenzie. 2008. *Functional Discourse Grammar: A typologically-based theory of language structure* (Oxford linguistics). Oxford University Press.

Hewitt, B. G. & S. R. Crisp. 1986. Speech reporting in the Caucasus. In Florian Coulmas (ed.), *Direct and indirect speech*. Berlin, New York: De Gruyter Mouton.

Hewitt, Brian G. 1995. *Georgian, a structural reference grammar*. .2 (London Oriental and African language library). Amsterdam: John Benjamins.

Höglund, Mikko et al. (eds.). 2015. *Perspectives on complementation: Structure, variation and boundaries*. Basingstoke: Palgrave Macmillan.

Holes, Clive. 1990. *Gulf Arabic* (Croom Helm Descriptive Grammars). London/New York: Routledge.

Holvoet, Axel. 2016. Semantic functions of complementizers in Latvian. In Kasper Boye & Petar Kehayov (eds.), *Complementizer semantics in European languages* (Empirical approaches to language typology 57). Berlin: De Gruyter Mouton.

Horie, Kaoru. 1993. *A cross-linguistic study of perception and cognition verb complements: A cognitive perspective*. Los Angeles: University of Southern California dissertation.

Horie, Kaoru. 2000. Complementation in Japanese and Korean. In Kaoru Horie (ed.), *Complementation: Cognitive and functional perspectives* (Converging evidence in language and communication research 1), 11–31. Amsterdam: John Benjamins.

Horie, Kaoru & Bernard Comrie. 2000. Introduction. In *Complementation: Cognitive and functional perspectives* (Converging evidence in language and communication research 1), 1–31. Amsterdam: John Benjamins.

Huttar, George L. & Mary L. Huttar. 1994. *Ndyuka* (Descriptive Grammars). London; New York: Routledge.

Iwasaki, Shoichi & Ingkaphirom Preeya. 2005. *A reference grammar of Thai*. Cambridge: Cambridge University Press.

Jäger, Andreas. 2007. Coding strategies of indirect reported speech in typological perspective. *The University of Queensland Working Papers in Linguistics* 1.

Jäger, Andreas. 2010. Reported speech constructions and the grammaticalization of hearsay evidentiality: a cross-linguistic survey. *STUF-Language Typology and Universals Sprachtypologie und Universalienforschung* 63(3). 177–195.

Jakobson, Roman. 1965 [1971]. Quest for the essence of language. In *Selected writings*, vol. II, 345–359. The Hague: De Gruyter Mouton.

Jensen, Torben Juel & Tanya Karoli Christensen. 2013. Promoting the demoted: The distribution and semantics of "main clause word order" in spoken Danish complement clauses. *Lingua* 137. 38–58.

Jones, Wendell & Paula Jones. 1991. *Barasano syntax: Studies in the languages of Colombia 2* (Summer Institute of Linguistics and the University of Texas at Arlington Publications in Linguistics (Book101). Summer Institute of Linguistics.

Joseph, Brian D. & Irene Philippaki-Warburton. 1987. *Modern Greek* (Croom Helm Descriptive Grammars). London/New York: Croom Helm.

Kachru, Yamuna. 2006. *Hindi* (London Oriental and African language library 12). Amsterdam: John Benjamins.

Kaltenböck, Gunther. 2009. Is *that* a filler? on complement clauses in spoken object clauses. *Vienna English Working Papers* 18(1). 28–63.

Katamba, Francis. 2003. The Bantu languages. In Derek Nurse & Gérard Philippson (eds.), *Bantu nominal morphology*, 103–120. London: Routledge.

Kawalya, Deo, Gilles-Maurice de Schryver & Koen Bostoen. 2018. Reconstructing the origins of the Luganda (JE15) modal auxiliaries -sóból- and -yînz-: a historical-comparative study across the West Nyanza Bantu cluster. *South African Journal of African Languages* 38(1). 13–25.

Kawasha, Boniface. 2007. Subject-agreeing complementizers and their functions in Chokwe, Luchazi, Lunda and Luvale. *Selected Proceedings of the 37th Annual Conference on African Linguistics*. 180–190.

Kehayov, Petar. 2017. *The fate of mood and modality in language death: Evidence from Minor Finnic* (Trends in Linguistics. Studies and Monographs 307). Berlin/Boston: De Gruyter Mouton.

Kehayov, Petar & Kasper Boye. 2016a. Complementizers in European languages: Overview and generalizations. In Kasper Boye & Petar Kehayov

(eds.), *Complementizer semantics in European languages* (Empirical approaches to language typology 57), 809–878. Berlin: De Gruyter Mouton.

Kehayov, Petar & Kasper Boye. 2016b. Complementizers in European languages: An introduction. In Kasper Boye & Petar Kehayov (eds.), *Complementizer semantics in European languages* (Empirical approaches to language typology 57), 1–11. Berlin: De Gruyter Mouton.

Kimenyi, Alexandre. 2018. *The Multiple Syntagmatic Positions and Various Meanings and Functions of the Morpheme Nga in Bantu*. `http://www.kimenyi.com/the-nga-morpheme-in-bantu.php` (26 August, 2018).

Kirsner, Robert & Sandra A. Thompson. 1976. The role of pragmatic inference in semantics: A study of sensory verb complements in English. *Glossa*. 200–240.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA & London: MIT Press.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*. Vol. I: Theoretical prerequisites. Stanford, CA: Stanford University Press.

Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar*. Vol. II: Descriptive application. Stanford, CA: Stanford University Press.

Langacker, Ronald W. 2008. *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.

Langacker, Ronald W. 2009. *Investigations in Cognitive Grammar*. Berlin; New York: Mouton de Gruyter.

Lees, Robert B. 1960. *The grammar of English nominalization*. Bloomington, Indiana: University Research Center in Anthropology, Folklore, and Linguistics.

Lord, Carol. 1993. *Historical change in serial verb constructions*. Amsterdam: John Benjamins.

Loux, Michael J. 1998. *Metaphysics, a contemporary introduction* (Routledge contemporary introductions to philosophy 1). London: Routledge.

Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.

Mahootian, Shahrazad. 2005. *Persian*. In collab. with Lewis Gebhardt. Digital print. (Descriptive grammars). Routledge.

Massamba, David P. B. 1986. Reported speech in Swahili. In *Direct and indirect speech* (Trends in Linguistics. Studies and Monographs 31). Berlin; New York; Amsterdam: Mouton de Gruyter.

Mauri, Caterina & Andrea Sansò. 2012. What do languages encode when they encode reality status? *Language sciences* 34. 99–106.

McGrath, Matthew & Frank Devin. 2018. Propositions. In Edward N. Zalta (ed.), *The Stanford encyclopedia of philosphy*, Spring 2018. `https://plato.stanford.edu/archives/spr2018/entries/propositions` (22 August, 2018).

Mckay, Thomas & Michael Nelson. 2014. The de re/de dicto distinction. In *The Stanford encyclopedia of philosphy*, Spring 2014. `https://plato.stanford.edu/entries/prop-attitude-reports/dere.html` (22 August, 2018).

Meussen, Achille E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 79–121.

Mithun, Marianne. 1995. On the relativity of irreality. In Joan Bybee & Suzanne Fleischmann (eds.), *Modality in grammar and discourse*, 367–388. Amsterdam: John Benjamins.

Myers, Amy. 1975. Complementizer choice in selected Eastern Bantu languages. *Proceedings of the Sixth Conference on African Linguistics. DSU WPL 20*. 184–193.

Nabirye, Minah. 2016. *A corpus-based grammar of Lusoga*. Gent: Universiteit Gent dissertation.

Newmeyer, Frederick. 1992. Iconicity and generative grammar. *Language* 68. 756–796.

Newmeyer, Frederick. 2015. Parentheticals and the grammar of complementation. In Stefan Schneider, Julie Glikman & Mathieu Avanzi (eds.), *Parenthetical verbs*. Berlin: De Gruyter.

Nichols, Johanna. 2011. *Ingush grammar* (University of Califormnia publications in linguistics). Berkely, CA: University of California Press.

Nikolaeva, Irina. 2014. *A grammar of Tundra Nenets* (Mouton Grammar Library 65). Berlin: De Gruyter.

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

Nordström, Jackie. 2010. *Modality and subordinators*. Amsterdam: John Benjamins.

Nordström, Jackie & Kasper Boye. 2016. Complementizer semantics in Germanic languages. In *Complementizer semantics in European languages* (Empirical approaches to language typology 57), 131–174. Berlin: De Gruyter Mouton.

Nurse, Derek. 2008. *Tense and aspect in Bantu*. Oxford: Oxford University Press.

Nurse, Derek & Henry R. T. Muzale. 1999. Tense and aspect in Great Lakes Bantu languages. In Jean-Marie Hombert & Larry Michael Hyman (eds.), *Bantu historical linguistics: Theoretical and empirical perspectives*, vol. 99 (Lecture notes), 517–544. Stanford: CSLI (Center for the Study of Language & Information) Publications.

Palmer, Frank Robert. 1979. *Modality and the English modals*. London: Longman.

Palmer, Frank Robert. 1986. *Mood and modality*. 1st edn. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Palmer, Frank Robert. 2001. *Mood and modality*. 2nd edn. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Perkins, Mick R. 1983. *Modal expressions in English*. Frankfurt am Main: Peter Lang.

Peterson, Philip L. 1997. *Fact proposition event* (Studies in linguistics and philosophy 66). Springer Netherlands.

Pietrandrea, Paula. 2012. The conceptual structure of irreality: A focus on non-exclusion of-factuality as a conceptual and a linguistic category. *Language Sciences* 34. 184–199.

Polakof, Ana Clara. 2017. Why are events, facts and states-of-affairs different? *Disputatio* IX(44). 99–122.

Premsrirat, Suwilai. 1987. A Khmu grammar. *Papers in South-east Asian linguistics* 10. 1–143.

Ransom, Evelyn N. 1986. *Complementation, its meanings and forms* (Typological studies in language 10). Amsterdam: John Benjamins.

Rentzsch, Julian & Liljana Mitkovska. 2017. Complementation strategies with the verb 'know' in Balkan Turkish, Standard Turkish and Macedonian. In. Saarland University, Saarbrücken: 39th Annual Conference of the German Linguistic Society, March 8th-10th.

Saeed, John. 2009. *Semantics*. 3rd edn. Chichester: Wiley-Blackwell.

Saxena, Anju. 1988. Historical change in serial verb constructions. In *Proceedings of the fourteenth annual meeting of the Berkeley linguistic society*, 375–388.

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

Schmidtke-Bode, Karsten & Holger Diessel. 2017. Cross-linguistic patterns in the structure, function, and position of (object) complement clauses. *Linguistics* 1(55). 233–269.

Schoenbrun, David L. 1994. Great Lakes Bantu: Classification and settlement chronology. *Sprache und Geschichte in Afrika* (15). 91–152.

Schüle, Susanne. 2000. *Perception verb complements in Akatek, a Mayan language*. Tübingen: University of Tübingen dissertation.

Searle, John R. 1969. *Speech acts: An essay in philosophy of language*. Cambridge: Cambridge University Press.

Seiler, Hans Jakob. 1977. *Cahuilla grammar*. Banning, CA: Malki Museum Press.

Serdobolskaya, Natalia. 2016. Semantics of complementation in Ossetic. In *Complementizer semantics in European languages* (Empirical approaches to language typology 57). Berlin: De Gruyter Mouton.

Sørensen, Marie-Louise Lind. 2013. *A cross-linguistic study of knowledge-predicate complementation*. Copenhagen: University of Copenhagen MA thesis.

Sørensen, Marie-Louise Lind & Kasper Boye. 2015. Vidensprædikatskomplementering. *Ny Forskning i Grammatik* 22. 233–253.

Sørensen, Marie-Louise Lind & Kasper Boye. 2017. Lagdelt indholdsstruktur i selvstændige ytringer og direkte og indirekte ytringsgengivelse. *Ny Forskning i Grammatik* 24. 121–140.

Spronck, Stef. 2012. Minds divided: Speaker attitudes in quotatives. In Isabelle Buchstaller & Ingrid van Alphen (eds.), *Quotatives, cross-linguistic and cross-disciplinary perspectives* (Converging evidence in language and communication research 15), 71–113. Amsterdam: John Benjamins.

Sridhar, S. N. 1990. *Kannada: Descriptive grammar* (Croom Helm Descriptive Grammars). London: Routledge.

Stanley, Jason. 2011. *Know how*. Oxford: Oxford University Press.

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

Sulkala, Helena & Merja Karjalainen. 1992. *Finnish* (Descriptive Grammar Series). London: Routledge.

Suzuki, Satoko. 2000. De dicto complementation in japanese. In Kaoru Horie (ed.), *Complementation: Cognitive and functional perspectives* (Converging Evidence in Language and Communication Research 1), 33–57. Amsterdam: John Benjamins.

Svenonius, Peter. 1994. *Dependent nexus: subordinate predication structures in English and the Scandinavian languages*. Santa Cruz: University of California, Santa Cruz dissertation.

Tamura, Suzuko & Takashi Kamei. 2000. *The Ainu language*. 1st edn. (ICHEL Linguistic Studies vol. 2). Tokyo: Sanseido.

Taylor, Charles V. 1985. *Nkore-Kiga* (Croom Helm Descriptive Grammars). London: Croom Helm.

Textor, Mark. 2016. States of affairs. In Edward N. Zalta (ed.), *The Stanford encyclopedia of philosphy*, Winter 2016. `https://plato.stanford.edu/archives/win2016/entries/states-of-affairs` (22 August, 2018).

Thompson, Sandra A. 2002. "Object complements" and conversation: Towards a realistic account. *Studies in Language* 26(1). 125–163.

Thompson, Sandra & Anthony Mulac. 1991. A quantitative perspective on the grammaticization of epistemic parentheticals in English. In Elisabeth

Closs Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*, vol. 2, 313–329. Amsterdam: John Benjamins.

Togeby, Ole. 2014. Tidernes følge – hos Mikkelsen. In Ole Togeby, Sten Vikner & Henrik Jørgensen (eds.), *Problemer og perspektiver i dansk syntaks*, 137–170. Odense: University Press of Southern Denmark.

Trask, Robert L. 1993. *A dictionary of grammatical terms in linguistics*. London: Routledge.

Van der Auwera, Johan & Vladimir A. Plungian. 1998. Modality's semantic map. *Linguistic Typology* 2. 79–144.

Van Valin, Robert & Randy LaPolla. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.

Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca, New York: Cornell University Press.

Verspoor, Marjolijn. 2000. Iconicity in English complement constructions. In Kaoru Horie (ed.), *Complementation: Cognitive and functional perspectives* (Converging Evidence in Language and Communication Research 1). Amsterdam: John Benjamins.

Verstraete, Jean-Cristophe. 2008. The status of purpose, reason, and intended endpoint in the typology of complex sentences; implications for layered models of clause structure. *Linguistics* 46(4). 757–788.

Zipf, George Kingsley. 1935. *The psychobiology of language*. Boston: Houghton-Mifflin.

Zucchi, Alessandro. 1993. *The language of propositions and events* (Studies in linguistics and philosophy 51). Dordrecht: Kluwer Academic Publishers.

Appendix A

173-Language sample

173-language sample

Language	Genus	Language Family
Abipón	South Guaicuruan	Guaicuruan
Abkhaz	Northwest Caucasian	Northwest Caucasian
Acoma	Keresan	Keresan
Ainu	Ainu	Ainu
Alamblak	Sepik Hill	Sepik
Amele	Madang	Trans-New Guinea
Apurinã	Purus	Arawakan
Arabic (Gulf)	Semitic	Afro-Asiatic
Araona	Tacanan	Tacanan
Arapesh (Mountain)	Torricelli	Torricelli
Armenian (Eastern)	Armenian	Indo-European
Asmat	Asmat-Kamoro	Trans-New Guinea
Awa Pit	Barbacoan	Barbacoan
Aymara (Central)	Aymaran	Aymaran
Bagirmi	Bongo-Bagirmi	Central Sudanic
Bambara	Western Mande	Mande
Barasano	Tucanoan	Tucanoan
Basque	Basque	Basque
Beja	Beja	Afro-Asiatic
Berber (Middle Atlas)	Berber	Afro-Asiatic
Brahui	Northern Dravidian	Dravidian
Bribri	Talamanca	Chibchan
Bunuba	Bunuban	Australian
Burmese	Burmese-Lolo	Sino-Tibetan
Burushaski	Burushaski	Burushaski
Cahuilla	Takic	Uto-Aztecan
Canela-Krahô	Ge-Kaingang	Macro-Ge

173-language sample (continued)

Language	Genus	Language Family
Cayuvava	Cayuvava	Cayuvava
Chamorro	Chamorro	Austronesian
Chinantec (Lealao)	Chinantecan	Oto-Manguean
Chukchi	Northern Chukotko-Kamchatkan	Chukotko-Kamchatkan
Comanche	Numic	Uto-Aztecan
Coos (Hanis)	Coosan	Oregon Coast
Cree (Plains)	Algonquian	Algic
Daga	Dagan	Dagan
Dani (Lower Grand Valley)	Dani	Trans-New Guinea
Diola-Fogny	Northern Atlantic	Niger-Congo
Ekari	Wissel Lakes-Kemanche	Trans-New Guinea
Epena Pedee	Choco	Choco
Evenki	Tungusic	Altaic
Ewe	Kwa	Niger-Congo
Fijian	Oceanic	Austronesian
Finnish	Finnic	Uralic
French	Romance	Indo-European
Fur	Fur	Fur
Garo	Bodo-Garo	Sino-Tibetan
Georgian	Kartvelian	Kartvelian
German	Germanic	Indo-European
Gooniyandi	Bunuban	Australian
Grebo	Kru	Niger-Congo
Greek (Modern)	Greek	Indo-European
Greenlandic (West)	Eskimo	Eskimo-Aleut
Guaraní	Tupi-Guaraní	Tupian
Haida	Haida	Haida
Hausa	West Chadic	Afro-Asiatic
Hindi	Indic	Indo-European
Hixkaryana	Cariban	Cariban
Hmong Njua	Hmong-Mien	Hmong-Mien
Huitoto (Minica)	Huitoto	Huitotoan
Hungarian	Ugric	Uralic
Hunzib	Avar-Andic-Tsezic	Nakh-Daghestanian
Igbo	Igboid	Niger-Congo
Ika	Arhuacic	Chibchan
Imonda	Border	Border
Indonesian	Malayo-Sumbawan	Austronesian
Ingush	Nakh	Nakh-Daghestanian
Iraqw	Southern Cushitic	Afro-Asiatic
Irish	Celtic	Indo-European
Jakaltek	Mayan	Mayan
Japanese	Japanese	Japanese
Ju|'hoan	Ju-Kung	Kxa
Kannada	Southern Dravidian	Dravidian
Kanuri	Western Saharan	Saharan
Karok	Karok	Karok
Kayah Li (Eastern)	Karen	Sino-TIbetan
Kayardild	Tangkic	Tangkic

173-language sample (continued)

Language	Genus	Language Family
Kera	East Chadic	Afro-Asiatic
Ket	Yeniseian	Yeniseian
Kewa	Engan	Trans-New Guinea
Khalkha	Mongolic	Altaic
Khasi	Khasian	Austro-Asiatic
Khmer	Khmer	Austro-Asiatic
Khmu'	Palaung-Khmuic	Austro-Asiatic
Khoekhoe	Khoe-Kwadi	Khoe-Kwadi
Kiowa	Kiowa-Tanoan	Kiowa-Tanoan
Korean	Korean	Korean
Koyraboro Senni	Songhay	Songhay
Krongo	Kadugli	Kadu
Kunama	Kunama	Kunama
Ladakhi	Bodic	Sino-Tibetan
Lak	Lak-Dargwa	Nakh-Daghestanian
Lakhota	Siouan	Siouan
Lango	Nilotic	Eastern Sudanic
Latvian	Baltic	Indo-European
Lavukaleve	Lavukaleve	Solomons East Papuan
Lepcha	Lepcha	Sino-Tibetan
Lezgian	Lezgic	Nakh-Daghestanian
Maba	Maban	Maban
Malagasy	Barito	Austronesian
Mandarin	Chinese	Sino-Tibetan
Mangarrayi	Mangarrayi	Australian
Mapudungun	Araucanian	Araucanian
Maranungku	Western Daly	Australian
Maricopa	Yuman	Hokan
Marind	Marind Proper	Marind
Martuthunira	Pama-Nyungan	Australian
Maung	Iwaidjan	Australian
Maybrat	North-Central Bird's Head	West Papuan
Meithei	Kuki-Chin	Sino-Tibetan
Menya	Angan	Trans-New Guinea
Miwok (Southern Sierra)	Miwok	Penutian
Mixtec (Chalcatongo)	Mixtecan	Oto-Manguean
Mundari	Munda	Austro-Asiatic
Muong	Viet-Muong	Austro-Asiatic
Murle	Surmic	Eastern Sudanic
Nahuatl (Tetelcingo)	Aztecan	Uto-Aztecan
Ndyuka	Creoles and Pidgins	other
Nenets	Samoyedic	Uralic
Nez Perce	Sahaptian	Penutian
Ngiti	Lendu	Central Sudanic
Nivkh	Nivkh	Nivkh
Nubian (Dongolese)	Nubian	Eastern Sudanic
Nunggubuyu	Nunggubuyu	Australian
Oneida	Northern Iroquoian	Iroquoian
Oromo Boraana	Lowland East Cushitic	Afro-Asiatic

173-language sample (continued)

Language	Genus	Language Family
Otomí (Mezquital)	Otomian	Oto-Manguean
Paiwan	Paiwan	Austronesian
Paumarí	Arauan	Arauan
Persian	Iranian	Indo-European
Pirahã	Mura	Mura
Pomo (Southeastern)	Pomoan	Hokan
Qawasqar	Alacalufan	Alacalufan
Quechua (Imbabura)	Quechuan	Quechuan
Rama	Rama	Chibchan
Russian	Slavic	Indo-European
Sango	Ubangi	Niger-Congo
Sanuma	Yanomam	Yanomam
Selknam	Chon Proper	Chon
Semelai	Aslian	Austro-Asiatic
Sentani	Sentani	Sentani
Shipibo-Konibo	Panoan	Panoan
Slave	Athapaskan	Na-Dene
Squamish	Central Salish	Salishan
Suena	Binanderean	Trans-New Guinea
Supyire	Gur	Niger-Congo
Swahili	Bantoid	Niger-Congo
Taba	South Halmahera - West New Guinea	Austronesian
Tagalog	Greater Central Philippine	Austronesian
Thai	Kam-Tai	Tai-Kadai
Tiwi	Tiwian	Australian
Tlingit	Tlingit	Na-Dene
Trumai	Trumai	Trumai
Tsimshian (Coast)	Tsimshianic	Penutian
Tukang Besi	Celebic	Austronesian
Tunica	Tunica	Tunica
Turkish	Turkic	Altaic
Una	Mek	Trans-New
Ungarinjin	Wororan	Australian
Vietnamese	Viet-Muong	Austro-Asiatic
Wambaya	West Barkly	Australian
Warao	Warao	Warao
Wardaman	Yangmanic	Australian
Wari’	Chapacura-Wanham	Chapacura-Wanham
Wichí	Matacoan	Matacoan
Wichita	Caddoan	Caddoan
Yagua	Peba-Yaguan	Peba-Yaguan
Yaqui	Cahita	Uto-Aztecan
Yimas	Lower Sepik	Lower Sepik-Ramu
Yoruba	Defoid	Niger-Congo
Yuchi	Yuchi	Yuchi
Yukaghir (Kolyma)	Yukaghir	Yukaghir
Yurok	Yurok	Algic
Zoque (Copainalá)	Mixe-Zoque	Mixe-Zoque

Appendix B

84-language sample (Article II)

84-language sample used in Article II

Language	Genus	Language Family
Abkhaz	Northwest Caucasian	Northwest Caucasian
Ainu	Ainu	Ainu
Amele	Madang	Trans-New Guinea
Arabic	Semitic	Afro-Asiatic
Armenian (Eastern)	Armenian	Indo-European
Awa Pit	Barbacoan	Barbacoan
Barasano	Tucanoan	Tucanoan
Basque	Basque	Basque
Bunuba	Bunuban	Australian
Burmese	Burmese-Lolo	Sino-Tibetan
Burushaski	Burushaski	Burushaski
Cahuilla	Takic	Uto-Aztecan
Chukchi	Northern Chukotko-Kamchatkan	Chukotko-Kamchatkan
Comanche	Numic	Uto-Aztecan
Cree (Plains)	Algonquian	Algic
Daga	Dagan	Dagan
Epena Pedee	Choco	Choco
Evenki	Tungusic	Altaic
Ewe	Kwa	Niger-Congo
Fijian	Oceanic	Austronesian
Finnish	Finnic	Uralic
Garo	Bodo-Garo	Sino-Tibetan
Georgian	Kartvelian	Kartvelian
Gooniyandi	Bunuban	Australian
Greek (Modern)	Greek	Indo-European

84-language sample used in Article II (continued)

Language	Genus	Language Family
Greenlandic (West)	Eskimo	Eskimo-Aleut
Guaraní	Tupi-Guaraní	Tupian
Hindi	Indic	Indo-European
Hmong Njua	Hmong-Mien	Hmong-Mien
Hungarian	Ugric	Uralic
Ingush	Nakh	Nakh-Daghestanian
Iraqw	Southern Cushitic	Afro-Asiatic
Jakaltek	Mayan	Mayan
Japanese	Japanese	Japanese
Kannada	Southern Dravidian	Dravidian
Karok	Karok	Karok
Kayah Li (Eastern)	Karen	Sino-TIbetan
Kayardild	Tangkic	Tangkic
Ket	Yeniseian	Yeniseian
Kewa	Engan	Trans-New Guinea
Khmer	Khmer	Austro-Asiatic
Khmu’	Palaung-Khmuic	Austro-Asiatic
Kiowa	Kiowa-Tanoan	Kiowa-Tanoan
Korean	Korean	Korean
Krongo	Kadugli	Kadu
Lakhota	Siouan	Siouan
Lango	Nilotic	Eastern Sudanic
Latvian	Baltic	Indo-European
Lepcha	Lepcha	Sino-Tibetan
Lezgian	Lezgic	Nakh-Daghestanian
Maba	Maban	Maban
Mandarin	Chinese	Sino-Tibetan
Maricopa	Yuman	Hokan
Maybrat	North-Central Bird’s Head	West Papuan
Meithei	Kuki-Chin	Sino-Tibetan
Menya	Angan	Trans-New Guinea
Mixtec (Chalcatongo)	Mixtecan	Oto-Manguean
Ndyuka	Creoles and Pidgins	other
Nenets	Samoyedic	Uralic
Nivkh	Nivkh	Nivkh
Paiwan	Paiwan	Austronesian
Paumarí	Arauan	Arauan
Persian	Iranian	Indo-European
Pirahã	Mura	Mura
Quechua (Imbabura)	Quechuan	Quechuan
Rama	Rama	Chibchan
Russian	Slavic	Indo-European
Semelai	Aslian	Austro-Asiatic
Slave	Athapaskan	Na-Dene
Suena	Binanderean	Trans-New Guinea

84-language sample used in Article II (continued)

Language	Genus	Language Family
Supyire	Gur	Niger-Congo
Swahili	Bantoid	Niger-Congo
Taba	South Halmahera - West New Guinea	Austronesian
Thai	Kam-Tai	Tai-Kadai
Trumai	Trumai	Trumai
Tukang Besi	Celebic	Austronesian
Turkish	Turkic	Altaic
Una	Mek	Trans-New
Ungarinjin	Wororan	Australian
Wambaya	West Barkly	Australian
Yagua	Peba-Yaguan	Peba-Yaguan
Yuchi	Yuchi	Yuchi
Zoque (Copainalá)	Mixe-Zoque	Mixe-Zoque