

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Epistemic modality across syntactic categories in Kipsigis

Permalink

<https://escholarship.org/uc/item/7d70r63x>

Author

Bossi, Madeline

Publication Date

2023

Peer reviewed|Thesis/dissertation

Epistemic modality across syntactic categories in Kipsigis

by

Madeline Claire Bossi

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Linguistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Amy Rose Deal, Chair

Professor Maria Kouneli

Professor Line Mikkelsen

Professor Peter Jenks

Summer 2023

Epistemic modality across syntactic categories in Kipsigis

Copyright 2023
by
Madeline Claire Bossi

Abstract

Epistemic modality across syntactic categories in Kipsigis

by

Madeline Claire Bossi

Doctor of Philosophy in Linguistics

University of California, Berkeley

Professor Amy Rose Deal, Chair

This dissertation concerns the expression of epistemic modality across syntactic categories in Kipsigis, an understudied Kalenjin language of Western Kenya. In addition to describing a range of morphosyntactic and semantic properties of Kipsigis, this dissertation takes as its empirical focus two case studies of epistemic modality in the nominal and verbal domains. Through these case studies, this dissertation contributes to the integration of data from understudied languages into semantic theory and suggests that the derivation of epistemic modal content—both across languages and across syntactic domains—must be achieved in a range of different ways.

In the nominal domain, I describe and analyze Kipsigis epistemic indefinites, highlighting the ways in which these forms pose challenges for existing analyses of epistemic indefinites across languages. In particular, I show that Kipsigis epistemic indefinites can signal both first order and higher order ignorance but do not take obligatory wide scope, which challenges the correlation between ignorance type and scope in Dawson (2020). To capture the Kipsigis pattern, I propose that the ignorance effects triggered by use of an epistemic indefinite in Kipsigis arise pragmatically due to competition with another type of Kipsigis nominal, which signals speaker knowledge about the noun's referent (following work on Russian *to* and *koe* indefinites by Geist & Onea 2007).

In the verbal domain, I consider biased belief reports with the Kipsigis verb *par* 'think', which can be used in two very different contexts: 1) to suggest that the reported belief is false or unlikely, and 2) to remind the addressee that the reported belief is true. While these negative bias and reminding functions are independently attested in other languages, the Kipsigis pattern is unique in that it combines these two functions in a single lexical item. To account for this behavior, I propose that, in addition to its basic belief semantics, *par* contributes an instruction for Common Ground management (Krifka 2008): the reported belief must not be *added* to the Common Ground.

To Linus, Kip, and all the other Kipsigis who made this work possible

Contents

Contents	ii
List of Figures	v
List of Tables	vi
List of Abbreviations	vii
1 Introduction	1
1.1 Dissertation structure	4
1.2 Kipsigis background and data collection	5
2 Kipsigis grammar sketch	10
2.1 Introduction	10
2.2 Phonology	10
2.2.1 Consonants	11
2.2.2 Vowels	11
2.2.3 Tone	12
2.3 Nominal structure and interpretation	13
2.3.1 Basic noun structure	13
2.3.2 Interpretation of nouns with a secondary suffix	15
2.3.2.1 Interpretation of mass and plural nouns	16
2.3.2.2 Interpretation of singular nouns	19
2.3.3 Dedicated indefinites	23
2.3.4 Possession	26
2.3.5 Relative clauses and adjectives	27
2.3.6 Demonstratives	29
2.3.7 Pronouns	31
2.4 Verbal structure	33
2.4.1 Overview of Kipsigis verbs	33
2.4.2 Tense	33
2.4.3 Aspect	34

2.4.4	Negation	37
2.4.5	Subject marking	37
2.4.6	Object marking	40
2.4.7	Valence-altering morphology	43
2.4.8	Object marking in ditransitives	46
2.5	Clausal structure	50
2.5.1	Word order	50
2.5.2	Case marking	54
2.5.3	Coordination	56
2.5.3.1	Conjunction	56
2.5.3.2	Disjunction	57
2.5.4	Questions	57
2.5.4.1	Polar questions	57
2.5.4.2	Constituent questions	59
2.5.5	Preverbal constituents	61
2.5.5.1	Pseudoclefts with the relativizers <i>ne</i> , <i>che</i> , or <i>ele</i>	62
2.5.5.2	Topic fronting with <i>ko</i>	69
2.5.6	Clausal subordination	73
2.5.6.1	Subjunctive embedded clauses	74
2.5.6.2	Indicative embedded clauses with no overt complementizer	75
2.5.6.3	Indicative embedded clauses with the element <i>kole</i>	75
2.5.6.4	Embedded questions	78
2.5.7	Modal verbs	79
2.6	Summary	84
3	Higher order ignorance in Kipsigis epistemic indefinites	85
3.1	Introduction	85
3.2	The Kipsigis nominal inventory	87
3.3	Characterizing the ignorance effects triggered by <i>-yan</i>	89
3.4	Ignorance effects are pragmatically derived	92
3.5	Semantic behavior of <i>-yan</i> forms	95
3.5.1	Domain widening analyses	95
3.5.2	Choice functional analyses	98
3.6	Nouns with a secondary suffix in Kipsigis	102
3.6.1	Distribution of nouns with a secondary suffix	102
3.6.2	Semantics of nouns with a secondary suffix	105
3.7	An analysis of <i>-yan</i>	109
3.7.1	Semantics of <i>-yan</i> forms	109
3.7.2	Pragmatic competition	111
3.8	Existential quantification, choice functions, and cross-linguistic variation	113
3.9	Conclusion	116

4	Negative bias, reminding, and pragmatic reasoning in Kipsigis belief reports	117
4.1	Introduction	117
4.2	The Kipsigis pattern	119
4.2.1	Interpretation with non-first person belief holders	119
4.2.2	Interpretation with first person belief holders	123
4.3	Negatively biased belief and reminding across languages	127
4.4	Analyzing Kipsigis <i>par</i>	130
4.4.1	Semantics of <i>par</i>	130
4.4.1.1	Evidence for not-at-issue content	132
4.4.1.2	The distribution of <i>par</i> and ruling out existing analyses	133
4.4.2	Pragmatics of <i>par</i>	135
4.4.2.1	<i>p</i> is false	136
4.4.2.2	Speaker is biased against <i>p</i>	136
4.4.2.3	<i>x</i> is unreliable	137
4.4.2.4	Addressee should know <i>p</i>	138
4.4.2.5	Restricting <i>par</i> 's reminding function	139
4.4.2.6	Evidence for pragmatic inferences	141
4.4.3	The grammatical status of first person present <i>par</i>	142
4.4.4	Against a syntactic alternative	143
4.5	Conclusion	146
5	Conclusion	148
5.1	Empirical and theoretical contributions	148
5.1.1	Epistemic indefinites	148
5.1.2	Biased belief reports	149
5.2	Future directions	151
	Bibliography	153

List of Figures

1.1	Kipsigis-speaking area in Western Kenya (Eberhard et al. 2021)	5
1.2	Cattle on a Kipsigis farm in Choronok Village, Bomet County	6
1.3	Tea fields in Choronok Village, Bomet County	7

List of Tables

2.1	Phonemic consonant inventory (from Kouneli 2019:20)	11
2.2	Phonemic vowel inventory (from Kouneli 2019:22)	12
2.3	Interpretation of mass and plural nouns with a secondary suffix	19
2.4	Interpretation of nouns with a secondary suffix	23
2.5	Kipsigis possessive morphemes (from Kouneli & Nie 2021:ex. 32)	26
2.6	Kipsigis demonstrative suffixes	30
2.7	Kipsigis temporal demonstrative suffixes (from Kouneli 2019:exs. 15-17)	31
2.8	Short and long pronouns in Kipsigis	32
2.9	Perfect subject markers in Kipsigis	36
2.10	Class I indicative subject markers	37
2.11	Class II indicative subject markers	38
2.12	Class I subjunctive subject markers	39
2.13	Class II subjunctive subject markers	39
2.14	Object markers	40
2.15	Nominative L-H*-L tonal melody on nouns	55
2.16	Modal verbs in Kipsigis	79
3.1	Kipsigis <i>-yan</i> vs. predictions of existing analyses and documented patterns	102
3.2	Distribution of noun types and ignorance effects in Kipsigis	105
3.3	Typology of epistemic indefinites and their properties	114
3.4	Comparison of the Tiwa and Kipsigis systems	115
4.1	Input and output CGs compatible with <i>par</i> statements (with respect to <i>p</i>)	134

List of Abbreviations

1	first person	NEUT	neutral aspect
2	second person	NMLZ	nominalizer
3	third person	NOM	nominative
ABS	absolutive	O	object
ACC	accusative	OBL	oblique
AND	andative	PERF	perfect aspect
AP	antipassive	PL	plural
APPL	applicative	POSS	possessive pronoun
AUX	auxiliary verb	PROX	proximal demonstrative
BIAS	bias particle	PST	past tense
C	complementizer	PST1	recent past tense
COND	conditional	PST2	past tense
COP	copula	PST3	distant past tense
DAT	dative	PURP	purpose clause
DO	direct object	REFL	reflexive
EP	epenthetic vowel	SBJV	subjunctive mood
GEN	genitive	SEC	secondary suffix
INF	infinitive	SG	singular
INST	instrumental	TH	theme vowel
IO	indirect object	TOP	topic
IPFV	imperfective aspect	Q	polar question particle
LP	local person subject	VBLZ	verbalizer
MOD	modal verb	VEN	venitive
NEG	negation		

Acknowledgments

This dissertation is the result of conversations with more people than I could possibly list here. Thank you to every single person who has supported me in this process, especially those who patiently endured my rambling about all things Kipsigis over the past five years. Whether or not your name appears below, know that I'm incredibly grateful for your help and wouldn't be here without you.

When I first started learning about Kipsigis in 2015, I never would've guessed that I'd be writing my dissertation about it eight years later. My continued interest in the language is due to the wonderful Kipsigis people who I've gotten to meet and collaborate with while studying the language. This dissertation wouldn't exist without the kindness and patience of Linus Kipkoech. From meeting with me on Zoom every Saturday morning during lockdown to introducing me to his entire—incredibly large—family back in Kenya, this dissertation only came about thanks to his dedication and generosity. Kiplangat Yegon also deserves special thanks for sharing his passion about Kipsigis language and culture with me, for working with me remotely even though it meant spending more time on the computer, and for cooking some of the best Kenyan food I've ever tasted. Linus and Kip, you're the reason this dissertation exists, and I'm so grateful for your support and friendship! And finally, to the rest of the Kipsigis community in Kenya who welcomed me with open arms, a huge *koongoi* to you all. Lydia, Beatrice, John, Memo, Ezra, the entire Ghovel family, thank you, and I'll see my Kenyan family again soon!

I'm also incredibly grateful to my dissertation committee. Not everyone is lucky enough to have four committee members who each bring such valuable and complementary perspectives to a project, but somehow I'm one of the lucky ones. I thank my advisor Amy Rose for all of her support over the last five years. Who knew that this syntactician would also become a semanticist, but this generally semantic dissertation would absolutely not exist without your mentorship. Your ability to recognize and articulate what's really at stake given an empirical puzzle is inspiring and made all of my work so much stronger. Thank you also to Line Mikkelsen. Conversations with you are some of the first experiences that made me feel like a real, grown-up linguist with interesting ideas of my own, and I'm so grateful to have been able to work with you. I still remember meeting you at BLS back in 2017, and you're a significant part of why I ended up here at Berkeley. I couldn't be happier with my decision! I'm also grateful for Peter Jenks, whose enthusiasm about all things linguistic always reminded me that there were so many more puzzles to explore and kept me excited, even when dissertation writing got exhausting. I always knew that, even if I didn't know what to work on next, you would have an idea, and I'm so grateful for that. Last but certainly not least, I thank Maria Kouneli for being such a kind and generous support on all things Kipsigis. Your deep knowledge of so many different aspects of the language is inspiring to me, and I'm so grateful for the Zoom brainstorming sessions we've been able to have over the past few years. Thank you all so much!

I also owe so much to the other linguists who have helped me develop as a linguist, researcher, and person. Mike Diercks is the reason that I became a linguist in the first place, and he continues to be a wonderful support to me. Apparently I'm not an easy advisee to get rid of, since I'm still texting you at random intervals six years later, but I'm so happy that you still respond! The

UC Berkeley linguistics department has been an incredible home to me over the last five years, and I'm grateful for the inspiration and support that I've received from so many department members. To Wesley and Emily Grabowski, you're the best cohortmates that I could've asked for. From struggling through problem sets together in first year to wrapping up our dissertations together, you've been a great source of support for me, and I'm so excited to see where we all land next. For making my time in grad school so much more enjoyable—despite the world falling apart right in the middle of it—I thank my linguist friends at Berkeley and beyond: Zach, Tessa, Myriam, Ginny, Emily Clem, Eric, Justin, Richard, Dakota, Allegra, Becky, Katie, Shweta, Franco, Amber, Julianne, Irene, and so many others. You all are simply the best, and I'm lucky to be your friend!

This dissertation also wouldn't have come about without the support of the many non-linguists in my life. You all keep me sane, and I'd be a much more annoying person without you. Madison, Danny, Sabrina, Michael, Ben, Maddie, Alex, Callie, Emma, Jackie, Angelica, and the rest of the Berkeley crew, I'll miss our impromptu game nights and Sunday dinners, but I can't wait to make new traditions with you all in this next chapter of my life. To Emma and Anne-Marie, I can't believe we're all about to live in the same city, but I couldn't be more excited. You're some of the kindest, funniest, and generally best humans I know, and I'm so excited to relive our Paris adventures together in a new place.

I'm also thankful for my family who support me unconditionally through all of my adventures. Mom and Dad, you didn't bat an eye when I decided to become a linguist—despite not knowing exactly what linguists do—and I really appreciate all the love and laughs you provided when things felt overwhelming. I'm sorry to leave California and take away Dad's excuse to come visit whenever he got bored, but I'm so happy to be living closer to you now, since it means that I'll get to see you and Heather more often!

Finally, I owe so much to Emily Drummond whose intelligence, passion, and kindness helped make this dissertation what it is and helped me become the person who I am today. I didn't quite know where to stick you in these acknowledgments, since you're my friend and cohortmate, but most importantly my partner. You support me in every possible way—as a linguist, a person, a cat mom, and so much more. Thank you for hugging me when hitting milestones didn't feel possible, for taking snail shots with me when they were done, and for being by my side for the next chapter of our lives and every other one down the line.

Chapter 1

Introduction

In a conversation, discourse participants often have many goals, though a primary one is to share information with each other to mutually construct an understanding about how the world is. An important part of developing this shared understanding involves communicating what one does and does not know or believe—a process that relates directly to the concept of epistemic modality. In a broad sense, epistemic modality concerns how individuals talk about what they know or believe. All languages have tools for expressing epistemic modal claims, ranging from modal verbs like English *might* and *must* to evidential markers like St’át’imcets *k’a*, *-an’*, and *ku7* (Matthewson et al. 2007) to epistemic indefinites like Spanish *algún* (Alonso-Ovalle & Menéndez-Benito 2010, 2013, 2017). As a result, the realm of epistemic modality offers a rich empirical landscape for cross-linguistic investigation and the integration of data from understudied languages into semantic theory.

Furthermore, as noted above, epistemic modal content can be communicated across a range of syntactic domains in these different languages. While modal verbs across languages have received a good deal of attention in the literature (for a useful overview, see Matthewson 2016 and references therein), there are many other strategies for expressing epistemic modal content beyond the verbal domain, as highlighted in Arregui et al. (2017). For instance, epistemic indefinites—which are indefinite pronouns or determiners that convey speaker ignorance with respect to the witness to the indefinite—are found in a wide range of languages, including English (Farkas 2002), German (Kratzer & Shimoyama 2002), Italian (Aloni & van Rooij 2004; Chierchia 2006; Zamparelli 2007), French (Jayez & Tovenca 2006, 2007), Russian (Geist & Onea 2007; Kagan 2011), Romanian (Farkas 2002; Fălăuş 2014), Japanese (Alonso-Ovalle & Shimoyama 2014), Czech (Šimík 2015), Tiwa (Dawson 2018, 2020), and Akan (Owusu 2019). In this way, it is possible to identify both cross-linguistic variation within a particular phenomenon related to epistemic modality, as well as within-language variation in the expression of epistemic modal content across syntactic categories.

A natural question that arises in light of this varied empirical landscape concerns how epistemic modality should be modeled across different languages and domains. In particular, do all expressions that convey epistemic modal content warrant a similar type of analysis, regardless of their syntactic category? To get at this question, this dissertation explores two case studies related

to epistemic modality in the nominal and verbal domains in Kipsigis—a Nilo-Saharan language of the Kalenjin subgroup spoken in Western Kenya. I offer an in-depth description of Kipsigis epistemic indefinites and biased belief reports. In light of these empirical patterns, I draw on existing work from other languages to provide analyses of these phenomena in Kipsigis. Perhaps unsurprisingly, these analyses differ quite significantly from each other; that is, the derivation of speaker ignorance with epistemic indefinites is quite different from the derivation of speaker bias with biased belief reports. Yet even so, neither of these accounts relies on the standard sort of machinery assumed for epistemic modal verbs like English *might* and *must*. In this way, this dissertation suggests that the derivation of epistemic modal content—both across languages and across syntactic domains—must be achieved in a range of ways.

As a starting point, consider the English epistemic modal verbs *might* and *must* in (1) - (2). Use of *might* in (1a) conveys that Emily's being at home is a possibility given the information that is currently available to the speaker. This type of epistemic possibility claim can be, as a starting place, modeled with the truth conditions in (1b), which express that there exists some possible world w' that is compatible with the speaker's epistemic state in w and that Emily is at home in w' . By contrast, use of *must* in (2a) conveys that Emily's being at home is a necessity given the information that is currently available to the speaker. This type of epistemic necessity claim can be modeled with the truth conditions in (2b), which express that in all worlds w' that are compatible with the speaker's epistemic state in w , Emily is at home in w' . In this way, one standard view of verbal epistemic modality is that it involves quantification—either existential or universal depending on the force of the modal—over possible worlds that are compatible with what the speaker knows or believes.¹

- (1) a. Emily **might** be at home.
 b. $\exists w'[w' \in Epi(sp, w) \ \& \ \text{Emily is at home in } w']$
- (2) a. Emily **must** be at home.
 b. $\forall w'[w' \in Epi(sp, w) \ \rightarrow \ \text{Emily is at home in } w']$

In this dissertation, I explore to what extent this type of analysis involving quantification over possible worlds applies to other kinds of epistemic modality—in particular, epistemic indefinites and biased belief reports in Kipsigis. In doing so, I show that different sorts of epistemic effects like speaker ignorance and speaker bias can be derived without recourse to quantification over possible worlds, using other mechanisms like pragmatic competition and instructions for Common Ground management (Krifka 2008).

Specifically, in Chapter 3, I consider epistemic modality in the nominal domain, describing and analyzing Kipsigis epistemic indefinites marked with the suffix *-yan*. I show that these *-yan* forms can signal first order ignorance about the individual witness to the indefinite (3a), as well as higher order ignorance about some contextually relevant property of the individual witness

¹The picture painted in (1) - (2) is a simplification; there is significant variation in the literature about how epistemic modality should be described and analyzed. As von Stechow & Gillies (2008) point out, there are different views about whose information is relevant when making an epistemic modal claim and about what exact kind of information this is, among other points of variation.

to the indefinite (3b). These forms are only ruled out when the speaker is not ignorant about the individual witness to the indefinite or any of their contextually relevant properties (3c).

- (3) Kibet and Chepkoech are playing hide-and-seek—Kibet is the seeker and Chepkoech is hiding. Kibet says to his babysitter:

∅-Uny-e-kεε Chεεpkεεch εεn rɔɔm-i-yan.

3-hide-IPFV-REFL C. in room-TH-YAN

‘Chepkoech is hiding in some room.’

(context adapted from Alonso-Ovalle & Menéndez-Benito 2010)

- a. ✓ Kibet knows that Chepkoech is in the house, but he doesn’t know which room she’s in, so he can’t find her.
- b. ✓ Kibet knows that Chepkoech is in the living room, but he doesn’t know where in the house the living room is, so he can’t find her.
- c. # Kibet cheated, so he knows that Chepkoech is in the living room. He also knows where in the house the living room is, so he can find her.

However, unlike other epistemic indefinites discussed in the literature that are compatible with higher order ignorance (e.g. Tiwa *khi* indefinites; Dawson 2018, 2020), Kipsigis *-yan* forms do not take obligatory wide scope (4), which challenges the correlation between ignorance type and scope proposed in Dawson (2020).

- (4) You overhear your mom complaining and ask your sibling why she’s so upset. Your sibling replies:

Ma-i-buch Kibeet rɔɔm-i-yan.

NEG-3-sweep K. room-TH-YAN

‘Kibet didn’t sweep any/some room’

✓ Kibet didn’t sweep any room.

¬ > ∃

✓ Kibet forgot to sweep a particular, unknown room.

∃ > ¬

To capture the Kipsigis pattern, I propose that the ignorance effects triggered by use of a *-yan* form arise as a conversational implicature due to competition with other Kipsigis nominals, which signal speaker knowledge about the noun’s referent (in line with work on Russian *to* and *koe* indefinites by Geist & Onea 2007). I offer a relatively simple semantics for *-yan*, according to which it contributes ordinary existential quantification over individuals. However, I suggest that other Kipsigis nouns with a secondary suffix (Kouneli 2019, 2021) introduce free choice function variables that are resolved via a contextually supplied assignment function and, consequently, signal speaker knowledge about how this variable is to be resolved. A speaker’s choice to existentially quantify with a *-yan* form—rather than refer using a noun with a secondary suffix—therefore gives rise to first order and higher order ignorance with *-yan* forms. In this way, the epistemic effects associated with Kipsigis nominals do not necessitate any kind of quantification over possible worlds; knowledge effects with nouns with a secondary suffix arise because of how free choice function variables are resolved, and ignorance effects arise via pragmatic reasoning about a speaker’s choice to existentially quantify rather than refer.

In Chapter 4, I explore verbal epistemic modality through the lens of biased belief reports in Kipsigis. The belief verb *par* in Kipsigis generally translates to ‘think’ but triggers additional inferences about the reported belief’s status with respect to the Common Ground (CG). In particular, with non-first person belief holders, use of *par* suggests that the speaker views the reported belief as false or unlikely (5).

- (5) I’m healthy, but my mother thinks that I’m sick because I fooled her to skip school. I say:
 Ø-**Par**-e kaamεε-nyʊʊn aa-mnyon-i.
 3-think-IPFV mother-1SG.POSS 1SG-be.sick-IPFV
 ‘My mother is **under the impression** that I’m sick.’

However, in the present tense with a first person belief holder, use of *par* triggers a very different inference; here, it suggests that the addressee should already know the reported belief to be true (6). In other words, *par* serves a reminding function here.

- (6) I arrive home and see a guest. I don’t know who the guest is, so I ask my mother who they are. She replies:
 A-**par**-e abuleyaanit.
 1SG-think-IPFV uncle
 ‘**Remember**, it’s your uncle.’ (Lit: ‘I think that it’s your uncle.’)

The negative bias and reminding functions seen in (5) - (6) are independently attested; a number of verbs accomplish only negative bias (e.g. Spanish *creerse*, Anvari et al. 2019; Mandarin *yǐwéi*, Glass 2022), while the German discourse particles *ja* and *doch* serve only a reminding function (e.g. Grosz 2016; Döring 2016; among many others). The Kipsigis pattern is novel in that it combines these two, seemingly contradictory phenomena in a single lexical item. In order to capture the full range of *par*’s interpretive effects, I propose that, in addition to its basic belief semantics, *par* contributes an instruction for CG management (Krifka 2008): the reported belief must not be *added* to the CG. Unlike existing analyses of just negative bias or reminding, this instruction is compatible with the reported belief not being in the CG prior to the belief report (5) or with it being in the CG prior to the report (6). Together with context-sensitive pragmatic reasoning about why a speaker would block addition of the reported belief to the CG, this instruction explains the case of a verb that can be used both to suggest that the reported belief is false and to remind the addressee that it is true. Notably, while the basic belief semantics associated with *par* involve quantification over possible worlds, the negative bias and reminding inferences do not require direct reference to the speaker’s epistemic state in any way; instead, they arise as a result of how discourse participants reason about why a speaker might not want a particular proposition to be added to the CG.

1.1 Dissertation structure

The dissertation is structured as follows. The remainder of this chapter offers background information about Kipsigis and the data that are found in the dissertation. Chapter 2 is a grammar

sketch of Kipsigis, with a particular focus on the morphosyntactic and semantic properties of the language. This chapter documents a range of grammatical properties of Kipsigis that have not received sufficient attention elsewhere in the literature and provides important context for understanding the data examples found in the rest of the dissertation. Then, the dissertation turns to two case studies related to epistemic modality in the nominal and verbal domains in Kipsigis. As noted above, Chapter 3 describes and analyzes epistemic indefinites in Kipsigis, highlighting the ways in which these forms pose challenges for existing analyses of epistemic indefinites across languages. Then, in Chapter 4, I consider biased belief reports with the belief verb *par* in Kipsigis, suggesting that the epistemic effects seen with this verb arise as a result of a Common Ground management instruction that it provides—namely, that the reported belief must not be *added* to the Common Ground. Chapter 5 concludes by summarizing the findings of the dissertation and highlighting directions for future research.

1.2 Kipsigis background and data collection

Kipsigis (ISO 639: *sgc*) is a Nilo-Saharan language of the Kalenjin subgroup spoken by a reported 1.9 million people (Eberhard et al. 2021). Kipsigis is traditionally spoken in the Rift Valley of Western Kenya, in the area shaded in orange in Figure 1.1. However, there is a growing diaspora population of Kipsigis people in urban centers like Nairobi, Kenya, as well as in other parts of the world including the United States.



Figure 1.1: Kipsigis-speaking area in Western Kenya (Eberhard et al. 2021)

Kipsigis people are traditionally pastoralists, and cattle play an important role in Kipsigis culture and economy. Figure 1.2 shows cattle on a Kipsigis farm in Choronok Village, Menet Sublocation, Kiromwok Location, Bomet County.



Figure 1.2: Cattle on a Kipsigis farm in Choronok Village, Bomet County

Agriculturally, the Rift Valley grows a significant amount of tea, which is a primary cash crop in the region. Tea fields in Choronok Village are shown in Figure 1.3.



Figure 1.3: Tea fields in Choronok Village, Bomet County

Finally, Kalenjin people have gained international fame due to their impressive running abilities. The vast majority of long-distance runners who have earned Kenya a reputation in world athletics are Kalenjin, including Eliud Kipchoge and Brigid Kosgei who currently hold world records in marathon running. As a result, there is some research that seeks to understand the source of these impressive athletic abilities (e.g. Saltin et al. 1995).

As mentioned above, Ethnologue reports that there are about 1.9 million Kipsigis speakers, citing 2009 census data (Eberhard et al. 2021). Kalenjin people constitute a significant ethnic group in Kenya, making up about 14% of the Kenyan population (Kouneli 2019) and holding a large degree of political power; for instance, Daniel arap Moi—the longest-serving Kenyan president since independence—is Tugen, and the current president William Ruto is Kalenjin. However, the source of the speaker count in Eberhard et al. (2021) is uncertain, as neither the 2009 nor the 2019 census contained a language question, and data gathered about respondents' ethnicity was not made public due to the sensitive nature of this information (KNBS 2009, 2019).² At best, then, the number reported in Eberhard et al. (2021) represents ethnic Kipsigis rather than speakers, who make up a large part but not the entirety of this ethnic population.

Regardless, speaker numbers are an unreliable metric in understanding language vitality in Africa (Dimmendaal & Voeltz 2008:595). The population of Western Kenya has doubled within the past 30 years, which means that there are likely far more Kipsigis people today than in the past. Despite this population growth, it is not necessarily the case that the number of Kipsigis speakers has doubled in this time period. Instead, Dimmendaal & Voeltz (2008:598) suggest that

²There is a history of tension between different ethnic groups in Kenya, which makes census data about ethnicity sensitive. The largest and most recent of these clashes occurred in 2007-2008—right before the launch of the 2009 census—as a result of the presidential election.

various sociolinguistic factors are relevant in determining an African language's endangerment status, including: patterns of multilingualism, the colonial legacy, the educational policy, and urbanization. Here, I discuss each of these factors in turn and show that Kipsigis is in a challenging position with respect to each of them. As is the case throughout Africa, there are extremely few sociolinguistic studies of Kipsigis language attitudes and use (see Muthwii 2004 on Kalenjin languages and Fink 2005 on Swahili and English in Kenya). The observations reported here come primarily from discussion with Kipsigis speakers and other researchers who work on Kipsigis and/or in Western Kenya.

First, the vast majority of Kipsigis speakers are multilingual in at least Kipsigis and Swahili. While many children—especially those living in more rural areas—still learn Kipsigis in the home, most Kipsigis under the age of forty also speak English. In fact, English and Swahili are the only languages given official recognition in the Kenyan constitution, which technically provides protections for indigenous languages (Republic of Kenya 2010), though in practice, there are very few actionable policies on this front (Kouneli 2019). Second, British colonial rule only ended in Kenya in 1963. English remains prestigious throughout the country, and many parents prioritize teaching their children English over Kipsigis (Muthwii 2004), as it offers more international mobility. Recently, attitudes toward English have soured slightly (Fink 2005), but—even so—Kipsigis has not noticeably increased in prestige. Instead, it is being replaced by Swahili, which is often seen as an African alternative over English. Third, all schooling beginning in Grade 1 is in English except for Swahili class, which is in Swahili. There is no established Kipsigis or Kalenjin orthography (see Jerono et al. 2012 for a recent attempt with no official recognition). Consequently, there are no pedagogical materials in Kipsigis, and even highly educated speakers do not read and write in Kalenjin (Kouneli 2019). In fact, children are often punished for speaking indigenous languages at school (Bunyi 1999), and because students often come from different ethnic backgrounds, they typically communicate with each other in Swahili rather than their native tongue. Fourth, more and more Kipsigis have moved from villages to cities within the past several decades, since there is more opportunity for economic advancement in urban areas like Nairobi, Kenya. Consequently, use of languages like Kipsigis has declined.

Within the Nilo-Saharan language family, Kipsigis is part of the Nilotic branch (Dimmendaal 2000, 2017), which includes languages like Turkana, Maasai, Dinka, and Luo (among others). Within Nilotic, Kipsigis is classified as a Southern Nilotic language, which is a branch that only contains Kalenjin languages and Datooga (see Ehret 1971; Rottland 1982 for the only comparative, historical work on Southern Nilotic). Within the Kalenjin subgroup specifically, Kipsigis is the most widely spoken language. The other Kalenjin languages include Nandi, Keiyo, Tugen, Terik, Endo-Marakwet, Pökoot, Sebei, and Okiek—listed roughly in order of highest to lowest mutual intelligibility with Kipsigis (Franciscar & Phylis 2012).

Kipsigis and all other Kalenjin varieties are highly understudied with particularly little work on their syntactic and semantic structure. Existing materials on Kipsigis include: Toweett's (1979) morphophonological sketch, descriptive work in German by Rottland (1982), and research on Kipsigis morphosyntax by Maria Kouneli and collaborators (e.g. Kouneli 2019, 2021, 2022; Kouneli & Nie 2021; Driemel & Kouneli 2022a). Considering Kalenjin more broadly, Creider & Creider (1989) offer a description of Nandi, Baroja et al. (1989) discuss Pökoot, and Jerono (2012) provides

a description of Tugen. In this way, a central goal of this dissertation is to help fill in this empirical gap by providing documentation and description of a wide range of syntactic, semantic, and pragmatic properties of Kipsigis.

Unless otherwise noted, the Kipsigis data in this dissertation come from my research on the language, which began in 2016 and continues through the present. Most of the data were collected through elicitation sessions conducted in-person or remotely over Zoom with Kipsigis speakers based in the United States—in particular, Linus Kipkoech and Kiplangat Yegon. Both Linus and Kiplangat were born and raised in Bomet County, Kenya and moved to the US to attend university: Linus at the University of California, Berkeley and Kiplangat at Lehigh University in Pennsylvania. They are both native speakers of Kipsigis and are also fluent in English and Swahili. In addition to this US-based work, I spent a month in Nairobi and Choronok Village, Kenya during Summer 2022, during which I conducted additional elicitation sessions, collected naturalistic texts, and learned more about Kipsigis culture. All original data in this dissertation is archived with the California Language Archive and is publicly available online: <http://dx.doi.org/doi:10.7297/X2D79918>.

The bulk of this dissertation addresses questions of semantic and pragmatic meaning. To get at these sorts of questions, I adopted the methodology for semantic fieldwork developed in Matthewson (2004) and Bochnak & Matthewson (2015). In particular, I presented speakers with sentences that I knew to be grammatical in elaborated contexts. These contexts were typically offered in English, but also included more extended Kipsigis discourses. Speakers were then asked to provide felicity judgements regarding these sentences in their respective contexts and/or felicity judgements of follow-up statements in Kipsigis. As is standard practice in semantics, infelicitous sentences are marked with #, while felicitous sentences are unmarked or marked with ✓. In addition to these contexts, speaker translations and comments were noted and taken to reveal information about the meaning of particular utterances.³ I am extremely lucky to work with individuals with a remarkable degree of metalinguistic awareness, and many of the ideas presented in this dissertation stem from seemingly small remarks that speakers have made over the years.

³At any point in the dissertation where this basic methodology was modified, I include explicit discussion of these modifications.

Chapter 2

Kipsigis grammar sketch

2.1 Introduction

This chapter provides an overview of the grammar of Kipsigis, with a focus on the morphosyntactic and semantic properties of the language. The goals of this chapter are two-fold. First, it provides background on my assumptions about Kipsigis grammar and helps contextualize the rest of the dissertation, which focuses on epistemic modality in two different syntactic domains in Kipsigis. Second, it describes many aspects of Kipsigis grammar that have not been adequately examined in the existing documentation of the language. The description and analysis in this chapter draws primarily on my own work, as well as that of Maria Kouneli and colleagues. Although this grammar sketch sets the stage for the rest of the dissertation, the following chapters are designed to stand on their own. The most relevant background information is repeated in those chapters, with references to this overview when applicable.

This chapter is structured as follows. §2.2 summarizes key phonological properties of Kipsigis, drawing on Kouneli (2019) and Kouneli & Nie (2021). In §2.3, I describe the structure and interpretation of Kipsigis nouns, including so-called “bare” nouns (§2.3.1-2.3.2), indefinites (§2.3.3), possessives (§2.3.4), demonstratives (§2.3.6), adjectives and relative clauses (§2.3.5), and pronouns (§2.3.7). Then, §2.4 outlines the structure of Kipsigis verbs, describing the system of tense (§2.4.2), aspect (§2.4.3), polarity (§2.4.4), subject marking (§2.4.5), transitive object marking (§2.4.6), valence-altering morphology (§2.4.7), and ditransitive object marking (§2.4.8). In §2.5, I turn to Kipsigis clause structure and describe word order variation (§2.5.1), case marking (§2.5.2), coordination (§2.5.3), questions (§2.5.4), constructions with preverbal constituents (§2.5.5), clausal embedding (§2.5.6), and verbal modality (§2.5.7).

2.2 Phonology

This section summarizes key aspects of Kipsigis phonology based primarily on Kouneli (2019) and Kouneli & Nie (2021). I refer readers directly to their work for data examples and for significantly more discussion of Kipsigis phonology.

2.2.1 Consonants

Kipsigis has thirteen phonemic consonants, which are listed in Table 2.1. While I generally transcribe data examples in this dissertation using IPA characters, there are a few consonants for which my spelling convention differs from the IPA symbol. For these sounds, my spelling convention is provided in [square brackets] in Table 2.1 (following Kouneli 2019). The deviation from standard IPA transcription in these cases is motivated by my desire to make the data examples in this dissertation accessible to as broad an audience as possible, including those who do not have extensive experience with the IPA.

	bilabial	alveolar	palatal	velar
stops	p	t	c [ch]	k
nasals	m	n	ɲ [ny]	ŋ [ngʻ]
trills		r		
fricatives		s		
glides			j [y]	w (labio-velar)
lateral approximants		l		

Table 2.1: Phonemic consonant inventory (from Kouneli 2019:20)

There are a number of phonological processes that target consonants in Kipsigis. First, the palatal stop /c/ is almost always realized as the post-alveolar affricate [tʃ] (Kouneli 2019:20). Second, the alveolar trill /r/ is variably realized as the trill [r] or the tap [ɾ] (Kouneli 2019:20). Third, all voiceless stops have voiced allophones that appear in specific environments, as summarized in (7) from Kouneli (2019).

- (7) Consonant phonological processes in Kipsigis (Kouneli 2019:21)
- a. voiceless stops → voiced / [l], [+nasal] ____
 - b. /p/, /k/ → voiced / V ____ V
 - c. /p/, /k/ → voiced / [r] ____

The voiced allophone [g] of /k/ often becomes the voiced fricative [ɣ] between vowels.

2.2.2 Vowels

Kipsigis has ten phonemic vowels, which are listed in Table 2.2.

	front	central	back
high	i ɪ		ɯ u
mid	e ɛ		ɔ o
low		a	ɑ

Table 2.2: Phonemic vowel inventory (from Kouneli 2019:22)

Advanced Tongue Root (ATR) and vowel length are distinctive features in the Kipsigis vowel system (Kouneli 2019:21). The [+ATR] vowels in Kipsigis are: i, u, e, o, a. The [-ATR] vowels are: ɪ, ʊ, ɛ, ɔ, a. In addition to [ATR], vowels are also contrastive in terms of length (Kouneli 2019:27). In this dissertation, I notate long vowels with a double vowel (e.g. [ɑ:] is represented [ɑɑ]).

Like many other African languages, Kipsigis displays ATR harmony—specifically, a dominant-recessive ATR harmony system (Kouneli 2019:23). In fact, although Kalenjin languages are extremely understudied, their system of ATR harmony has featured in much work on theoretical phonology (Hall et al. 1974; Halle & Vergnaud 1981; Lodge 1995; Bakovič 2000; Local & Lodge 2004; Nevins 2010). In the Kipsigis system, a morpheme specified as [+ATR] within the phonological word will cause all other vowels to become [+ATR]. As a result, in most data examples in this dissertation, all vowels within a word match in their [ATR] specification. There are, however, a few morphemes—like pronominal possessive suffixes and *-keɛ* ‘REFL’—that do not participate in ATR harmony; Kouneli (2019:24) attributes this behavior to their status as phonological clitics rather than true affixes.

When two identical short vowels occur side-by-side, they form a long vowel (Kouneli 2019:27). When two non-identical short vowels occur side-by-side, no change occurs unless they appear in one of the combinations in (8)—in which case, the change depends on the exact vowel combination (Kouneli 2019:27-28). [ATR] value is not relevant to the coalescence patterns in (8).

- (8) Vowel coalescence in Kipsigis (Kouneli 2019:ex. 10)
- a. /ɑ/ + /i/ → [ee]
 - b. /e/ + /i/ → [ee]
 - c. /u/ + /i/ → [uu]

Kipsigis also displays a process of vowel length dissimilation, in which some underlying long vowels are realized as short vowels. Because these patterns are quite complicated, are only attested in the nominal domain so far, and do not directly relate to the topics discussed in this dissertation, I do not summarize them here, but see Kouneli (2019:28-30) for discussion.

2.2.3 Tone

Kipsigis is a tonal language. There are three surface tones in Kipsigis: high (H), low (L), and a contour high-falling tone (HL) (Kouneli & Nie 2021:117). H and L can appear on syllables with long or short vowels, while HL can only surface on long vowels and in some syllables with a short

vowel and a sonorant coda (Kouneli & Nie 2021:117). All sequences of tones across two syllables are grammatical in Kipsigis, except sequences of contour tones (Kouneli 2019:33).

Tonal phonological processes in Kipsigis include rising tone simplification (9a) and high tone lowering (9b), as described in Kouneli & Nie (2021).

- (9) Tonal phonological processes in Kipsigis (Kouneli & Nie 2021:119-120)
- a. Rising tone simplification: LH contour within a syllable due to morphological concatenation surfaces as H.
 - b. High tone lowering: HH sequence within a syllable due to morphological concatenation surfaces as HL contour.

Tone serves a range of grammatical purposes in Kipsigis, including (but likely not limited to) verbal inflection (§2.4) and case marking (§2.5.2). At present, the grammatical functions of tone, as well as the exact tonal reflexes associated with these different functions, are not particularly well-understood. In this dissertation, I transcribe tone when it seems grammatically relevant, but I omit it elsewhere to avoid any potential misrepresentation of the data. Existing work that considers the tone systems of Kalenjin languages includes Creider (1982); Zwarts (2004); Dimmendaal (2012); Rolle (2018); Kouneli & Nie (2021); Sande (To Appear).

2.3 Nominal structure and interpretation

In this section, I provide a detailed overview of the structure and interpretation of Kipsigis nouns. The reason for this level of detail is two-fold. First, while Kipsigis is generally quite understudied, the structure of Kipsigis nouns has received a good amount of attention in the theoretical literature (Kouneli 2019, 2021). Second, Chapter 3 of this dissertation offers an even deeper dive into the interpretation of two types of Kipsigis nominals, rendering the background information presented here particularly important for the rest of the dissertation.

2.3.1 Basic noun structure

Kipsigis has been classified as a bare noun language (Kouneli 2019), since nouns in the language can appear in both indefinite (10) and definite contexts (11) without the need for articles (see e.g. Dayal & Sağ 2020 for an overview on bare nouns). In the following mini-discourse, *kaaneetiindet* ‘teacher’, *laakwæt* ‘child’, and *ng’ookta* ‘dog’ introduce new discourse referents in (10). Then, in (11), these same nouns act as anaphoric definites, referring back to the individuals introduced previously.

- (10) Ø-Sooman-chiin **kaaneetiindet laakwæt** kɪtabʊʊt agɔbɔ **ng’ookta**.
 3-read-APPL.IPFV teacher child book about dog
 ‘A teacher is reading a book about a dog to a child.’

- (11) Ø-Tep-**een laakwæt kaaneetiindet** kole tyan oo ng'ookta.
 3-ask-INST child teacher C how.much big dog
 'The child asks the teacher how big the dog is.'

Following Kouneli (2019, 2021), I assume that these Kipsigis nouns are morphologically complex and fall into three classes based on number: 1) inherently singular nouns, 2) inherently plural nouns, and 3) numberless nouns. Inherently singular nouns only contain a number morpheme in the plural (12), whereas inherently plural nouns only contain a number morpheme in the singular (13). Numberless nouns contain a number morpheme in the singular and plural (14).

- (12) Inherently singular noun (Kouneli 2021:4)

- | | |
|--|---|
| a. peet-u-it → pêtúut
day-TH-SEC
'day' | b. peet-uus-ya-ik → pêtúusyéék
day-PL-TH-SEC
'days' |
|--|---|

- (13) Inherently plural noun (Kouneli 2021:4)

- | | |
|---|--|
| a. ngeend-yaan-ta-it → ngéendyáat
bean-SG-TH-SEC
'bean' | b. ngeend-a-ik → ngéendéék
bean-TH-SEC
'beans' |
|---|--|

- (14) Numberless noun (Kouneli 2021:4)

- | | |
|---|--|
| a. sigis-yaan-ta-it → sigìsyáat
sock-SG-TH-SEC
'sock' | b. sigis-iin-ik → sigìsìiník
sock-PL-SEC
'socks' |
|---|--|

As seen in (12) - (14), Kipsigis nouns contain a nominal root followed by a number suffix, then a thematic vowel, then a secondary suffix. Nouns do not necessarily contain all of these morphemes; while all nouns contain a secondary suffix (or an equivalent; §2.3.3), not every nominal requires a number morpheme (12a) or a thematic vowel (14b).¹ However, when all of these morphemes co-occur, they surface in the order in (15).

- (15) Root - Number suffix - Theme vowel - Secondary suffix

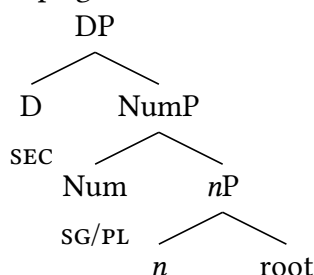
Based on this system of nominal morphology, Kouneli (2019, 2021) proposes that Kipsigis nouns have the structure in (16). A noun's inherent number feature is realized in *n*, which determines the noun's class (i.e. whether it is inherently singular, plural, or numberless). Num is responsible for the interpretable number feature associated with a noun, which dictates whether the noun is interpreted as singular or plural. Finally, the secondary suffix is realized in *D*, which

¹See Kouneli (2021:1201) for more detailed discussion of the distribution of thematic vowels. Following Kouneli (2021), I assume that thematic suffixes are inserted post-syntactically, meaning that their presence or absence does not have structural ramifications.

agrees with Num in number features, since secondary suffixes have different singular vs. plural realizations, as seen with *-it* vs. *-ik* in (12) - (14). Evidence that the secondary suffix is in D comes from the fact that it alternates with another determiner-like element—namely, the indefinite marker *-yan* discussed in §2.3.3 and analyzed in depth in Chapter 3. The appropriate morpheme order is generated via head movement of the nominal root to D through the intervening heads. In this way, the nouns that have been described as “bare” in Kipsigis are notably different from true bare nouns in other languages, since they necessarily contain DP structure. For this reason, in the rest of the dissertation, I call these forms “nouns with a secondary suffix” rather than bare nouns.

(16) Kipsigis basic noun structure

(Kouneli 2021:24)



2.3.2 Interpretation of nouns with a secondary suffix

This section describes the range of interpretations available for nouns with a secondary suffix in Kipsigis. In order to describe these interpretations, I divide Kipsigis nouns into three categories: 1) mass nouns, 2) plural nouns, and 3) singular nouns. Mass nouns cannot be counted in their standard form and typically denote things without obvious sub-parts like water, mud, sand, and blood. Plural and singular nouns, on the other hand, are countable and denote things that can be individuated like tables, cookies, elephants, and flowers.

In terms of possible interpretations, I focus on kind-level, definite, and existential readings of these noun types. Kind-level readings make a statement about the kind of thing denoted by the noun. For instance, the English predicate *extinct* is a kind-level predicate because it is only possible for a kind to be extinct; when someone says *Dinosaurs are extinct*, they are making a claim about the natural kind—they are not talking about a particular dinosaur or even a particular group of dinosaurs. A definite interpretation of a noun arises when it is used in a context where its intended reference is recoverable, as seen previously in (11) repeated here as (18). By contrast, existential readings arise when the noun is used to introduce a new discourse referent, as seen above in (10) repeated here as (17).

- (17) Ø-Sooman-chiin **kaaneetiindet laakwæet** kɪtabʊʊt agɔbɔ **ng'ookta**.
 3-read-APPL.IPFV teacher child book about dog
 ‘A teacher is reading a book about a dog to a child.’

- (18) Ø-Tep-**een laakwæt kaaneetiindet** kole tyan oo **ng'ookta**.
 3-ask-INST child teacher C how.much big dog
 'The child asks the teacher how big the dog is.'

The discussion in this section draws on work on the interpretation of bare nouns across languages (Chierchia 1998; Dayal 2004; Deal & Nee 2018; Collins 2019; Little 2020; Moroney 2021), given that Kipsigis nouns with a secondary suffix are like bare nouns in some ways. Cross-linguistically, bare mass and plural nouns typically pattern together in terms of their possible interpretations, while bare singular nouns show a more restricted range of interpretations. This same patterning is found in Kipsigis, and so I describe the interpretation of mass and plural nouns together in §2.3.2.1, then consider the interpretation of singular nouns in §2.3.2.2.

2.3.2.1 Interpretation of mass and plural nouns

Mass (19) and plural (20) nouns have kind-level readings in Kipsigis. Testing kind-level readings is difficult in the language because many kind-level predicates like 'common', 'widespread', 'rare', 'extinct', etc. do not translate directly into Kipsigis. For this reason, my diagnostics for kind-level readings rely on a small handful of predicates including *tau* 'start, invent' and *it* 'arrive'.

- (19) Kii-Ø-it **sugarwək** Amerika keny.
 PST3-3-arrive sugar America long.ago
 'Sugar arrived in America long ago.'
- (20) Ng'oo ne kii-Ø-tau **mɔnaɔk?**
 who REL.SG PST3-3-start glasses
 'Who invented glasses?' (Lit: 'Who started glasses?')

Mass and plural nouns also have definite readings—including both anaphoric and maximal definites.² The sentences in (21) - (22) show anaphoric definite uses of mass and plural nouns. The first sentence introduces the referent, and the second sentence shows that the same noun can be used to refer back to this referent.

- (21) a. Ø-Tiny-e Kiproono **cheega** ak **peek**.
 3-have-IPFV K. milk and water
 'Kiproono has milk and water.'
- b. I-yaam-toos **peek** ak **cheega** Kiproono.
 3-mix-IPFV.PL water and milk K.
 'Kiproono is mixing the water with the milk.'
- (22) a. Ø-Cheesan-toos **ng'eetiik** ak **tiibiik** een pii.
 3-play-IPFV.PL boys and girls in outside
 'Boys and girls are playing outside.'

²Following Chierchia (1998); Dayal (2004); Geist (2010); Schwarz (2013); and others, I treat maximal definites as the plural and mass counterpart to uniqueness definites in the singular. Just as singular uniqueness definites pick out the unique noun, maximal definites pick out the unique maximality of the relevant mass or plural noun.

- b. Ø-Pir-e mbireet **tiibiik** ko-rwaat-e **ng'eetiik**.
 3-kick-IPFV ball girls 3.SBJV-run.around-IPFV boys
 'The girls are kicking a ball and the boys are running around.'

The sentences in (23) - (24) show maximal definite uses of mass and plural nouns.

- (23) When Cherono went to the kitchen, she noticed sugar on the table. Her son Kibet loves sugar, and Cherono knows that he often sneaks into the kitchen to eat sugar when she's not paying attention. Cherono went to go look for Kibet.
 Ka-Ø-manda Cheeroono εεn chigeet ko-pwaat-e kole ka-Ø-am Kibeet **sugarʊʊk**.
 PST1-3-leave C. in kitchen 3.SBJV-think-IPFV C PST-3-eat K. sugar
 'Cherono left the kitchen, thinking that Kibet ate the sugar.'
- (24) We have 4 dogs, and they're all outside barking right now.
 Ø-Po-toos **ng'oogik** εεn pii.
 3-bark-IPFV.PL dogs in outside
 'The dogs are barking outside.'

Finally, mass (25) and plural (26) nouns can have existential readings. Note that the relevant nouns in (25) - (26) are external arguments, which limits the possibility that these existential interpretations are due to semantic incorporation, since semantic incorporation can generally only apply to internal arguments like objects.

- (25) a. Koo-Ø-tum-chi-kεε **cheega** shiati-ng'ʊʊng'-i?
 PST2-3-spill-APPL-REFL milk shirt-2SG.POSS.SG-Q
 'Did milk spill on your (sg) shirt?'
 (context adapted from Deal & Nee 2018:ex. 23)
- b. You notice a water leak.
 Ø-Peend-i **peek**.
 3-go.PL-IPFV water
 'Water is spilling.'
 (context adapted from Deal & Nee 2018:ex. 22)
- (26) a. You hear some dogs barking.
 Ø-Po-toos **ng'oogiik** εεn pii.
 3-bark-IPFV.PL dogs in outside
 'Dogs are barking outside.'
- b. You notice a pot boiling on the stove.
 Ø-Rur-toos **piaasiniik**.
 3-boil-IPFV.PL potatoes
 'Potatoes are boiling.'

Crucially, though, the existential quantification seen in (25) - (26) must take narrow scope with respect to other operators in the sentence. This pattern is illustrated here with negation. Only the narrow scope $\neg > \exists$ reading is appropriate in response to the question in (25a); such a

response would indicate that no milk spilled on the speaker's shirt—not that some milk did not spill on their shirt, even if some other milk did (i.e. the wide scope $\exists > \neg$ reading). As seen in (27), the mass noun subject is felicitous in response to the question in (25a).

(27) In response to (25a):

Aacha, ma- \emptyset -tum-chi-kεε **cheega** shratr-nyʊʊn.

no NEG-3-spill-APPL-REFL milk shirt-1SG.POSS.SG

'No, milk didn't spill on my shirt.'

$\neg > \exists$

(context adapted from Deal & Nee 2018:ex. 24)

By contrast, only the wide scope $\exists > \neg$ reading is appropriate given the context in (28). The sentence in (28) states that water poured out of the truck, and the continuation in (29) is intended to explain that not *all* of the water spilled—only *some* of it did. However, the mass noun subject is not felicitous here (29).

(28) Koo- \emptyset -lul loorit ago tumnda-kεε **peek**.

PST2-3-fall truck then pour-REFL water

'A truck fell over and water poured out.'

(context adapted from Deal & Nee 2018:ex. 25)

(29) #Lakini ma- \emptyset -tumnda-kεε **peek**.

but NEG-3-pour-REFL water

Intended: 'But some water didn't pour out.'

* $\exists > \neg$

Speakers report that (29) sounds like a contradiction in light of (28). To express the wide scope existential reading, speakers quantify the mass noun (30) or restructure the sentence altogether (31).

(30) Lakini ma- \emptyset -tumnda-kεε **peek** tʊgʊl.

but NEG-3-pour-REFL water all

'But not all the water poured out.'

(31) Lakini mii **peek** che ma- \emptyset -tumnda-kεε.

but COP water REL.PL NEG-3-pour-REFL

'But there's some water that didn't pour out.'

An identical pattern is seen with plural nouns, which are likewise restricted to narrow scope existential readings. In contexts like (32a), which support a narrow scope $\neg > \exists$ interpretation, plural nouns are felicitous. However, in contexts like (32b), which require a wide scope $\exists > \neg$ interpretation, these same nouns are no longer appropriate.

(32) Ma- \emptyset -po-toos **ng'oogiik** εen pii.

NEG-3-bark-IPFV.PL dogs in outside

'Dogs aren't barking outside.'

a. ✓ There are no dogs barking outside.

$\neg > \exists$

- b. # There are six dogs outside; four are barking, but two are not. *∃ > ¬
 (context adapted from Deal & Nee 2018:ex. 27)

As with mass nouns, speakers paraphrase to get a wide scope existential reading, which involves syntactic restructuring (33) or the addition of numerals (34).

- (33) Mii ng'oogiik che ma-∅-po-toos.
 COP dogs REL.PL NEG-3-bark-IPFV.PL
 'There are dogs that aren't barking.'
- (34) ∅-Po-toos ng'oogiik ang'wan-u lakini ma-po-toos oeeng'-u.
 3-bark-IPFV.PL dogs four-NOM but NEG-bark-IPFV.PL two-NOM
 'Four dogs are barking but two aren't.'

In this way, mass and plural nouns in Kipsigis have the range of possible interpretations summarized in Table 2.3. Interestingly, this is the same pattern of interpretations found in true bare noun languages like Hindi, Russian, Mandarin, and many others (see e.g. Chierchia 1998; Dayal 2004; Deal & Nee 2018; Collins 2019; Little 2020; Moroney 2021). Despite the syntactic complexity of mass and plural nouns in Kipsigis (§2.3.1), they seem to behave semantically like bare nouns in other languages. This fact differentiates them from singular nouns in Kipsigis, which have a unique range of interpretations that is the focus of the following section.

Reading	Mass noun	Plural noun
a) Kind-level	✓	✓
b) Definite	✓	✓
c) Narrow scope existential	✓	✓
d) Wide scope existential	✗	✗

Table 2.3: Interpretation of mass and plural nouns with a secondary suffix

2.3.2.2 Interpretation of singular nouns

Turning to singular nouns with a secondary suffix, in this section, I show that they have a range of interpretations that is restricted differently than with mass and plural nouns. While kind-level and definite interpretations are available for Kipsigis singular nouns, existential readings are possible but necessarily take wide scope. This pattern is the opposite of the one seen with existential readings of mass and plural nouns, where only narrow scope interpretations are possible.

First, however, I consider kind-level and definite readings of these noun forms. Like mass and plural nouns, Kipsigis singular nouns allow kind-level readings (35).

- (35) Kii-∅-tau ng'oo karɪt?
 PST3-3-start who car
 'Who invented the car?' (Lit: 'Who started the car?')

Likewise, singular nouns have anaphoric and unique definite readings. As discussed at the beginning of this section, in the mini-discourse in (36), the sentence in (36a) introduces a number of new discourse referents, and the following sentence in (36b) uses the same nouns to refer anaphorically back to these individuals.

- (36) a. \emptyset -Sooman-chiin **kaaneetiindet laakwæet** kɪtabʊʊt agɔbɔ **ng'ookta**.
 3-read-APPL.IPFV teacher child book about dog
 'A teacher is reading a book about a dog to a child.'
- b. \emptyset -Tep-**een laakwæet kaaneetiindet** kole tyan oo **ng'ookta**.
 3-ask-INST child teacher C how.much big dog
 'The child asks the teacher how big the dog is.'

Similarly, the singular noun in (37) has a unique definite reading, as it refers to the sun, which is necessarily unique.

- (37) \emptyset -Lool-e **asiista-i?**
 3-shine-IPFV sun-Q
 'Is the sun shining?'

Finally, singular nouns can have existential readings, as seen above in (36a) and again in (38). These examples illustrate that Kipsigis singular nouns can introduce new discourse referents without the numeral *agɛɛngɛ* 'one.' This Kipsigis pattern differs from the general cross-linguistic picture, in which bare singular nouns in many bare noun languages require the numeral 'one' in order for existential readings of these nouns to be possible (see e.g. Dayal 2004; Deal & Nee 2018; Collins 2019; Moroney 2021).

- (38) There are six dogs outside. I hear lots of barking and see you come inside looking disheveled. I ask what happened. You say:
 Ka- \emptyset -sʊs-an **ng'ookta**.
 PST1-3-bite-1SG.O dog
 'A dog bit me.'

The singular noun in (38) is an external argument, meaning that its existential reading is unlikely to be due to semantic incorporation. Furthermore, it cannot be any standard kind of definite; there is no previous mention of dogs in the discourse to license an anaphoric definite, and there is no unique dog in the context to license a unique definite. Instead, it seems as though the singular noun can truly introduce a new discourse referent—contrary to the cross-linguistic pattern reported for bare noun languages in Dayal (2004). In fact, this use of singular nouns is common in Kipsigis; stories often start with singular nouns, as in (39) - (40) taken from the beginning of different Kipsigis texts.³

³These texts were written by Kipsigis speakers, then reviewed by me for morpheme identification and minor phonetic editing based on the word list in Toweett (1975). As a result, it is likely that some of the phonological properties represented here (e.g. the [ATR] value and length of vowels) are not completely accurate.

- (39) Kii-mii **mooseet** ne kii-∅-am-e logoeek εen **keetit** parak ne kii-mii
 PST3-COP monkey REL.SG PST3-3-eat-IPFV fruits in tree top REL.SG PST3-COP
 taband-aap **aineet**.
 bank-of river
 ‘There once was a monkey that ate fruits on top of a tree that was on a river bank.’
- (40) Kii-∅-leen-j-eech abule kii-mii **cheepkelelyaat** ne kii-ki-kuren
 PST3-3-say.IPFV-APPL-1PL.IO uncle PST3-COP young.lady REL.SG PST3-1PL-call
 Tabutany ne kii-∅-meny-e **kookweet** ne kii-negit **cheeplangeet**.
 T. REL.SG PST3-3-live-IPFV village REL.SG PST3-near leopard
 ‘The uncle told us that there was a young lady called Tabutany who lived in a village that was close to a leopard.’

However, when the sentence contains other scope-taking elements, the existential quantification seen above must take wide scope with respect to these other operators. This pattern is illustrated here with the attitude verb *mach* ‘want’ (41), the necessity modal *nyaalu* (42), and negation (43). Note that the singular nouns in (42) - (43) are the object of the verb. The unavailability of narrow scope existential readings even with singular nouns in object position suggests that semantic incorporation is unavailable in Kipsigis; if it were available, narrow scope existential readings should be possible in (42) - (43).⁴

- (41) ∅-Mach-e Kibeet ko-nyoo **cheepta** paarti.
 3-want-IPFV K. 3.SBJV-come girl party
 ‘Kibet wants a certain girl to come to the party.’
- a. # There are many girls in Kibet’s class at school. He’d be happy if any of them come to his party. *want > ∃
- b. ✓ There are many girls in Kibet’s class at school. He dislikes most of these girls and doesn’t want them to come to his party. There’s just one particular girl who he wants to come: he wants Sharon to come because he has a crush on her. ∃ > want
- (42) Nyaal-u ko-buch Kibeet **rɔvmit**.
 MOD-IPFV 3.SBJV-sweep K. room
 ‘Kibet must sweep a certain room.’
- a. # There are 5 rooms in the house. As part of his chores, Kibet has to sweep any one of those rooms. It could be the kitchen or the bathroom or the living room. *□ > ∃
- b. ✓ There are 5 rooms in the house. As part of his chores, Kibet has to sweep the kitchen—the messiest room in the house. ∃ > □

⁴The unavailability of semantic incorporation of nouns in object position is typologically unusual in the space of bare noun languages. However, Kipsigis nouns are more morphologically complex than those in most bare noun languages (§2.3.1), which might offer an explanation for why they cannot undergo semantic incorporation. Furthermore, the general unavailability of semantic incorporation has no consequences for the discussion prior to this point, since all data examples presented previously include nouns in subject position, which generally cannot semantically incorporate even in languages where this is possible.

(43) Ma-i-buch Kibeet rɔʊmɪt.

NEG-3-sweep K. room

‘Kibet didn’t sweep a certain room.’

a. # Kibet didn’t do any of his chores—he didn’t sweep a single room! $*\neg > \exists$

b. ✓ Kibet did some of his chores but not all of them. There are 5 rooms that he needs to sweep. Today, he swept 4 of them, but he didn’t sweep 1: he skipped the kitchen because it’s the messiest room in the house and too much work to sweep. $\exists > \neg$

Speaker comments in response to these kinds of sentences suggest that these singular nouns are most natural when the speaker has a particular instance of the noun in mind (i.e. a particular girl or room in the examples above). In this way, Kipsigis singular nouns appear to have specific indefinite interpretations. For instance, in response to (42), speakers suggest that “you’re singling out a room” and that an interlocutor might ask the speaker which particular room Kibet has to sweep. In fact, a natural follow-up to the statement in (42) is shown in (44), where the speaker provides more information about which room Kibet has to sweep.

(44) Following (42):

Nyaal-u ko-buch chiigeet amɔn mur nra.

MOD-IPFV 3.SBJV-sweep kitchen because dirty very

‘He has to sweep the kitchen because it’s very dirty.’

To express narrow scope existential readings, speakers typically use a form of quantification that translates to ‘any N at all’ (45).

(45) In context (41a):

∅-Mach-e Kibeet ko-nyoo chepta age tɔgɔl paartɪ.

3-want-IPFV K. 3.SBJV-come girl any all party

‘Kibet wants any girl to come to the party.’

In their indefinite use, then, Kipsigis singular nouns necessarily take wide scope and are most natural when the speaker has a particular referent in mind. While I do not develop a full account of these forms here, see Chapter 3 for more empirical discussion and an analysis.

To conclude this section, Table 2.4 builds upon Table 2.3 to summarize the range of interpretations available for all nouns with a secondary suffix in Kipsigis. This distribution parallels the cross-linguistic picture with true bare noun languages in many ways (see e.g. Dayal 2004). However, there are two key semantic properties that differentiate Kipsigis from bare noun languages—both of which relate to the interpretation of singular nouns. First, Kipsigis speakers can use singular nouns to introduce new discourse referents without a numeral and second, singular nouns must take wide scope on their indefinite readings.

Reading	Mass noun	Plural noun	Singular noun
a) Kind-level	✓	✓	✓
b) Definite	✓	✓	✓
c) Narrow scope existential	✓	✓	✗
d) Wide scope existential	✗	✗	✓

Table 2.4: Interpretation of nouns with a secondary suffix

2.3.3 Dedicated indefinites

In addition to the noun forms discussed in §2.3.1 - 2.3.2, the suffix *-yan* can replace a noun's secondary suffix to form a dedicated indefinite, as in the b. examples in (46) - (47), which are both built upon inherently singular nouns. While speakers translate nouns with a secondary suffix as 'the N' or 'a N' depending on the context, forms with *-yan* are consistently translated as 'some N'. In this section, I focus on the structure of these *-yan* forms and refer readers to Chapter 3 for a full description of their interpretation and a semantic analysis.

- | | |
|---|---|
| <p>(46) a. kar-I-IT → kàríit
 car-TH-SEC
 'the/a car' (Kouneli 2019:ex. 5c)</p> <p>b. kar-I-YAN
 car-TH-YAN
 'some car'</p> | <p>(47) a. laak-wa-IT → làakwéet
 child-TH-SEC
 'the/a child' (Kouneli 2019:ex. 5b)</p> <p>b. laak-wa-YAN
 child-TH-YAN
 'some child'</p> |
|---|---|

Some evidence that these *-yan* forms are dedicated indefinites comes from the fact that they cannot co-occur with expressions that often encode definiteness across languages, including demonstrative suffixes (48) and possessive pronouns (49; §2.3.4).

- (48) a. *kar-I-YAN-NI → *kariyaní
car-TH-YAN-PROX
Intended: 'this some car'
- b. *kar-I-NI-YAN → *karníyan
car-TH-PROX-YAN
Intended: 'this some car'
- (49) *kar-I-YAN-NG'ʊŋg' → *kariyang'ʊŋg'
car-TH-YAN-2SG.POSS.SG
Intended: 'some car of yours'

To express possession with *-yan* forms, speakers use an alternative possession structure that includes a relative clause (for more discussion, see §2.3.4).

- (50) choor-wa-**yan** ne po Kibeet
 friend-TH-YAN REL.SG belong.to K.
 ‘some friend of Kibet’s’

Turning to the structure of *-yan* forms, the examples in (46) - (47) show that *-yan* replaces a noun’s secondary suffix to create a dedicated indefinite. In a similar vein, the ungrammatical examples in (51) show that it is impossible for *-yan* to surface alongside a noun’s typical secondary suffix—be it before (51a) or after (51b) the secondary suffix.

- (51) a. *laak-wa-**yan-it** → *laakwayanɪt
 room-TH-YAN-SEC
 Intended: ‘some child’
 b. *laak-wa-**it-yan** → *laakwεε(t)yan
 room-TH-SEC-YAN
 Intended: ‘some child’

Not all nouns can take the *-yan* suffix. First, only singular nouns have *-yan* forms. All nouns with *-yan* are interpreted as singular, and it is impossible for *-yan* to replace the secondary suffix on a plural noun. This restriction is particularly clear in (52), where *-yan* cannot surface alongside the plural morpheme *-oy*; instead, it must surface after the thematic vowel in the singular form, as seen in (47b).

- (52) a. laak-oy-ɪk → là̀góók
 child-PL-SEC
 ‘children’ (Kouneli 2019:49)
 b. *laak-oy-**yan**
 child-PL-YAN
 Intended: ‘some children’

Second, the set of nouns that *-yan* attaches to is lexically restricted. So far, I have found that the forms in (53) can contain *-yan*. While this list is not exhaustive, it highlights the lexically restricted nature of the *-yan* suffix. However, note that the addition of *-yan* applies to native Kipsigis words but is also common with borrowed words, which speaks to its synchronic productivity. For instance, the Kipsigis noun *tagitarɪt* ‘doctor’ is borrowed from English *doctor* and has the *-yan* form *tagitarɪyan*. Similarly, the Kipsigis noun *ndisiyaat* ‘banana’ is borrowed from Swahili *ndizi* ‘banana’ and has the *-yan* form *ndisiyan*.

(53) Non-exhaustive list of nouns with *-yan* forms⁵

tagitariyan*	'doctor'	araawayan	'month'	tɔrayan	'boar'
kaaneetiindayan	'teacher'	sɛgɛnɛniyan	'stream'	ngɔkyayan	'chicken'
poiyan	'man'	pɔrɔstayan	'forest'	ngurwayan	'pig'
murenyan	'man'	keetiyān	'branch'	kariyan*	'car'
choorwayan	'friend'	uyan	'somewhere'	ndisiyan**	'banana'
cheepkerichan	'healer'	eetiyān	'bridge'	kitunguyan	'onion'
kirwaagiindayan	'chief'	marɪndayan	'dress'	sɔgɔliyan*	'school'
laakwayan	'child'	artayan	'goat/sheep'	rɔɔmiyan*	'room'
kaandoiindayan	'leader'	mɔriyan	'calf'		
kɪplagɔriyan	'hunter'	nyuumbuyan	'mule'		

In addition, *-yan* forms crosscut Kipsigis noun classes, applying to inherently singular nouns and inherently plural nouns alike. Inherently singular nouns like *laakweet* 'child' and *arteet* 'goat/sheep' have *-yan* forms, specifically *laakwayan* 'some child' and *artayan* 'some goat/sheep'. In these cases, *-yan* does not co-occur alongside number morphology, since inherently singular nouns are not marked for number in the singular. The inherently plural nouns *kaaneetiindeet* 'teacher' and *kirwaagiindeet* 'chief' also have *-yan* forms, where *-yan* co-occurs alongside the singulative suffix *-iin* (54). In this way, *-yan*'s distribution suggests that it acts independently of the Kipsigis noun class system.

- (54) a. *kaaneet-iin-da-yan* → *kaaneetiindayan*
 teach-SG-TH-YAN
 'some teacher'
- b. *kirwaak-iin-da-yan* → *kirwaagiindayan*
 chief-SG-TH-YAN
 'some chief'

An apparent exception to this generalization is that *-yan* cannot co-occur with singulative *-yaan*; it can only co-occur with singulative *-iin*, which is the only other singulative morpheme in the language. Compare for instance, the ungrammaticality of (55) to the grammaticality of the examples in (54).

- (55) * *peel-yaan-yan* → **peelyaanyan*
 elephant-SG-YAN
 Intended: 'some elephant'

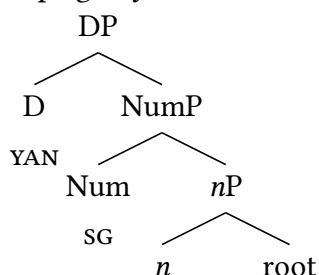
Because indefinite *-yan* can surface alongside singulative *-iin*, I do not take the ungrammaticality of (55) as evidence that *-yan* is a realization of Num. Instead, I assume that there is a dispreference in Kipsigis for adjacent, phonologically similar morphemes (i.e. haplology). In fact, speakers comment that *peelyan*—the form in (55) with *-yaan* 'SG' deleted—can mean 'some elephant' as predicted for *-yan* forms.⁶

⁵Forms marked with * and ** are English and Swahili borrowings, respectively.

⁶See Chapter 3 for comments on the historical development of indefinite *-yan* from singulative *-yaan*.

Given these empirical facts and building on the structure of Kipsigis nouns in (16), I suggest that *-yan* is an alternative realization of D, such that *-yan* forms have the structure in (56). This structural analysis captures the fact that *-yan* surfaces after number morphology (i.e. singulative *-iin*) and is in complementary distribution with the secondary suffix, which is also realized in D. As with nouns with a secondary suffix, I assume that the appropriate morpheme order is generated via head movement of the nominal root to D through the intervening heads.

(56) Kipsigis *-yan* form structure



2.3.4 Possession

Possessive pronouns in Kipsigis are suffixes that attach to the nominal that they modify. These pronouns are sensitive to the person and number of the possessor, as well as the number of the possessum. Table 2.5 lists the possessive pronouns found in Kipsigis. As mentioned in §2.2.2, these morphemes are not within the ATR harmony domain of the noun that they attach to, which might suggest that they are phonological clitics rather than true affixes.

	SG _{possessum}	PL _{possessum}
1SG	-nyʊʊn	-chʊʊk
2SG	-ng'uung'	-kuuk
3SG	-nyim	-chɪk
1PL	-naan	-chaak
2PL	-ng'waang'	-kwaak
3PL	-nywaan	-chwaak

Table 2.5: Kipsigis possessive morphemes (from Kouneli & Nie 2021:ex. 32)

To express possession by a non-pronominal possessor, two different structures can be used. One construction involves attaching the morpheme *-aap* to the possessed nominal, followed by the possessor (57). As with possessive pronouns, *-aap* is outside the ATR harmony domain of the noun and remains [-ATR] even when attached to a [+ATR] noun, as seen in (57) with the [+ATR] noun *chuumbiik* 'salt'. An alternative construction uses the relative clause *ne pa* 'that belongs to' followed by the possessor (58). For more on the structure of relative clauses see §2.3.5.

- (57) \emptyset -Al-e Taputany chuumbiik-aap tuuga.
 3-buy-IPFV T. salt-of cows
 ‘Taputany is buying cows’ salt.’
- (58) choorwayan **ne** \emptyset -pa Kibeet
 friend REL.SG 3-belong.to K.
 ‘a friend of Kibet’s’

2.3.5 Relative clauses and adjectives

Kipsigis relative clauses are head-external, meaning that the noun that the relative clause modifies does not appear in its typical position within the clause, but rather appears to the left of the clause. These relative clauses are introduced by the relativizers *ne* (59), *che* (60), and *ele* (61). *Ne* and *che* are used when the head of the relative clause is a core argument of the verb (e.g. subject, direct object, indirect object), whereas *ele* is used when the head of the relative clause is an oblique. *Ne* is used when the relative clause head is singular, while *che* is used when it is plural.

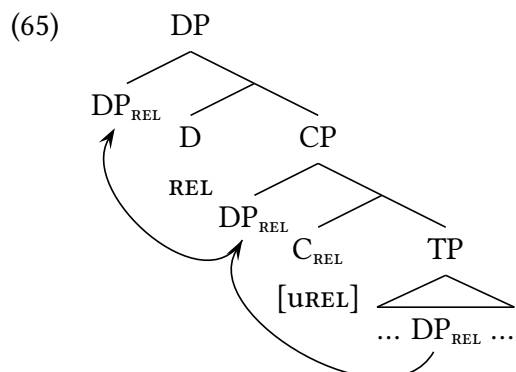
- (59) kwaanda **ne** koo- \emptyset -keer ng’ookta
 woman REL.SG PST2-3-see dog
 ‘the woman who saw a dog’
- (60) ng’oogiik tugul **che** mii pii
 dogs all REL.PL COP outside
 ‘all the dogs that are outside’
- (61) kiwaanyjeet **ele** koo- \emptyset -keer-een kwaanda ng’ookta
 field REL.OBL PST2-3-see-in woman dog
 ‘the field where the woman saw the dog’

In addition to the relative clauses in (59) - (61), where the relative head is overt, Kipsigis also has headless relative clauses, where the head is null. Headless relative clauses use the same relativizers *ne* (62), *che* (63), and *ele* (64), but no overt material precedes the relativizer.

- (62) **ne** \emptyset -cham-e Kibeet
 REL.SG 3-like-IPFV K.
 ‘the one that Kibet likes’
- (63) **che** ka- \emptyset -chap Kibeet
 REL.PL PST1-3-make K.
 ‘the ones that Kibet made (recently)’
- (64) **ele** \emptyset -weend-i Kipchirchir
 REL.OBL 3-go.SG-IPFV K.
 ‘where Kipchirchir is walking’

A possible structure for Kipsigis relative clauses is shown in (65), following Kouneli (2019). The relative head begins the derivation within the relative clause and moves to its surface position.

Again following Kouneli (2019), I assume that the relativizer is a realization of D; for instance, it agrees in number with the noun that it modifies, which is characteristic of D in Kipsigis (§2.3.1), and it bears tonal case morphology (§2.5.2). In order to derive the correct word order in Kipsigis relative clauses—in which the relative head is to the left of the relativizer—I assume that the relative head moves to Spec,CP, then undergoes an additional movement step to Spec,DP above the relativizer that is realized in D, as suggested in Kouneli (2019).



Evidence in support of this style of raising analysis for Kipsigis relative clauses comes from island sensitivity and reconstruction effects and is presented in §2.5.5 for headless relative clauses in a particular type of focus construction in Kipsigis. However, I do not assume that this relative clause structure is necessarily the only possible one found in Kipsigis; it remains an open question whether Kipsigis also has matching relative clauses (Chomsky 1965; Sauerland 1998, 2000, 2003; a.o.), in which the relative head is base-generated at the edge of the relative clause.

Relative clauses are extremely frequent in Kipsigis because all adjectival modification uses a relative clause structure. Examples of adjectival modification via relative clauses are shown in (66). It is impossible to modify nouns directly with adjectives without this sort of relative clause structure.

- (66) a. ng'ookta *(ne) tʊʊy
 dog REL.SG black
 'a black dog' (Kouneli 2019:127)
- b. peeleeek somok *(che) eecheen
 elephants three REL.PL big.PL
 'the three big elephants' (Kouneli 2019:146)

Adjectives strongly resemble verbs in Kipsigis. For instance, tense morphology (§2.4.2) and subject marking (§2.4.5) can directly attach to adjectives, which can serve as predicates without an overt copula (67).

- (67) **Kii-a**paypay.
 PST3-1SG-happy
 'I was happy (long ago).' (Kouneli 2019:107)

Despite this similarity, Kouneli (2019) argues that adjectives represent a distinct lexical category in Kipsigis because there are several properties that differentiate adjectives from verbs in Kipsigis. I summarize three of her arguments here, but see Kouneli (2019: Chapter 4) for discussion of additional empirical differences between adjectives and verbs. First, adjectives in Kipsigis relative clauses must agree with the head noun in case, whereas verbs in relative clauses do not bear any kind of case morphology (see §2.5.2 for more on the case system of Kipsigis). Second, Kipsigis verbs display a productive system of reduplication, which indicates repeated action; however, adjectives are unable to undergo this same type of reduplication. Third, Kipsigis adjectives are compatible with the intensifier *kot* ‘very’ (68a), while verbs are not (68b).

- (68) a. Chepta ko karaaran **kot**.
 girl TOP beautiful very
 ‘The girl is very beautiful.’ (Kouneli 2019:119)
- b. *Ru-e **kot**.
 sleep-IPFV very
 Intended: ‘He/she sleeps a lot.’

While this is only a sample of the empirical patterns discussed in Kouneli (2019), it suggests that adjectives are morphosyntactically distinct from verbs in Kipsigis.

Kouneli (2019) also notes that all adjectives have different singular vs. plural realizations. The suffix *-een* often marks plural adjectives, as seen in (69).

- (69) (69) Piriir ngecherεet.
 red chair
 ‘The chair is red.’ (Kouneli 2019:114)
- (70) Piriir-**een** ngecherook.
 red-PL chairs
 ‘The chairs are red.’ (Kouneli 2019:118)

However, it is also possible for the number distinction on Kipsigis adjectives to be marked purely in terms of the [ATR] specification of the vowels in the adjective. In (71a) for instance, the singular form of the adjective is [-ATR], while the plural form of the adjective is [+ATR], as in (71b).

- (71) a. laakwεet ne **karaaran**
 girl REL.SG beautiful.SG
 ‘a beautiful girl’ (Kouneli 2019:108)
- b. laag-oo-chu somok chu **karaaran**
 girl-PL-PROX three PROX beautiful.PL
 ‘these three beautiful girls’ (Kouneli 2019:136)

2.3.6 Demonstratives

Kipsigis makes a three-way distinction in the demonstrative domain between proximal, medial, and distal demonstratives. These forms—listed in Table 2.6—are suffixes that attach to nouns.

The proximal demonstratives *-ni* and *-chʊ* are in the ATR harmony domain of the noun, while the medial and distal demonstratives do not participate in ATR harmony (Kouneli 2019). In this section, I summarize some key properties of the Kipsigis demonstrative system, but I refer readers to Kouneli (2019) for more discussion.

	SG	PL
PROX	-ni	-chʊ
MED	-naan	-chaan
DIST	-nɪn	-chʊʊn

Table 2.6: Kipsigis demonstrative suffixes

These demonstrative suffixes attach to different amounts of nominal structure. The singular demonstrative suffixes replace the noun’s secondary suffix (§2.3.1), while the plural demonstrative suffixes generally co-occur with it.⁷ This pattern can be seen in (72) for singular nouns and in (73) for plural nouns.

- (72) laak-waa-**ni**/**-naan**/**-nɪn** → laakwaani/laakwaanaan/laakwaanɪn
 child-TH-PROX.SG/MED.SG/DIST.SG
 ‘this/that child’ (Kouneli 2019:ex. 4b)
- (73) peel-a-**ik-chʊ**/**-chaan**/**-chʊʊn** → peeleechu/peeleechaan/peeleechʊʊn
 elephant-TH-SEC-PROX.PL/MED.PL/DIST.PL
 ‘these/those elephants’

The [k] that is typically found in the plural secondary suffix deletes in the presence of a demonstrative suffix. However, it is possible to diagnose the presence of the secondary suffix in (73) because of the vowel quality and length in the noun’s second syllable; the second [ee] in forms like *peeleechu* ‘these elephants’ arises due to regular vowel coalescence between [a] from the thematic suffix and [i] from the secondary suffix (§2.2.2; see also Kouneli 2019:fn. 98).

Kipsigis also displays a pattern of determiner spreading, which applies to the demonstrative suffixes in Table 2.6. While the full empirical picture is complicated and thoroughly described in Kouneli (2019:Chap. 5), I outline the basic facts here. As discussed in §2.3.5, nouns can be modified by a relative clause introduced by the relativizers *ne*, *che*, or *ele*. However, a demonstrative suffix can replace the relativizer, as seen in (74) with adjectival modification.

- (74) pagaa-ni (*ne) leel
 cat-PROX.SG REL.SG white.SG
 ‘this white cat’ (Kouneli 2019:ex. 28b)

⁷This is a slight simplification of the pattern. While most nouns display the type of co-occurrence in (73), the plural demonstrative suffixes replace the secondary suffix in nouns that take the *-ka* allomorph of the secondary suffix. See Kouneli (2019:fn. 98) for more information.

When a demonstrative replaces the relativizer and there are multiple adjectives modifying a single noun, the demonstrative must be repeated before every adjective (75); this is the pattern known as determiner spreading, since the demonstrative appears multiple times but only seems to be interpreted once.

- (75) pagaa-nɪ tuuy *(nɪ) oo
 cat-PROX.SG black.SG PROX.SG big.SG
 ‘this big black cat’ (Kouneli 2019:ex. 35a)

Finally, in addition to the demonstratives in Table 2.6, Kipsigis also has three temporal demonstratives, listed in Table 2.7. These suffixes can only attach to a noun that already bears a proximal demonstrative suffix, in which case the proximal demonstrative serves as a definiteness marker rather than a true demonstrative (Kouneli 2019:143).

	SG	PL
PST1	-kaan	-kaan
PST2	-kɔɔnyɛ	-kɔɔchɛ
PST3	-kɪnyɛ	-kɪchɛ

Table 2.7: Kipsigis temporal demonstrative suffixes (from Kouneli 2019:exs. 15-17)

The temporal demonstratives in Table 2.7 are the nominal counterparts of the verbal graded tense system found in Kipsigis (§2.4.2). In particular, they locate an individual in discourse time; for instance, use of the suffix *-kɪnyɛ* in (76) indicates that the speaker is talking about the girl from a long time ago, rather than yesterday or earlier today.

- (76) laak-waa-nɪ-kɪnyɛ
 girl-TH-PROX.SG-PST3.SG
 ‘this girl from long ago’ (Kouneli 2019:ex. 17a)

2.3.7 Pronouns

Kipsigis has two pronoun series for all person and number combinations, which are summarized in Table 2.8. I refer to these pronoun types as short vs. long forms, since the long forms are morphologically built upon the short forms. In general, speakers prefer to use the short forms with first and second person referents and the long forms with third person referents.

	SG	PL
1	anɛɛ(ndɛt)	ɛɛchɛɛk(ɛt)
2	inyee(ndɛt)	ookweek(ɛt)
3	mɛɛ(ndɛt)	ɪchɛɛk(ɛt)

Table 2.8: Short and long pronouns in Kipsigis

The pronouns in Table 2.8 can only refer to animate referents. This restriction is difficult to see with first and second person pronouns because these pronouns necessarily refer to discourse participants, who are generally animate. However, this restriction becomes apparent with third person pronouns, which cannot be anteceded by inanimate objects. This pattern is shown in (77) where the pronoun *mɛɛndɛt* ‘3SG’ is intended to refer to the speaker’s new jacket. Speakers deem this sentence unnatural and suggest that it would be appropriate if *mɛɛndɛt* referred to a human, but is not felicitous as-is in (77).

- (77) In response to a complement about my new jacket:
 Koongoi! Koo-a-al-e (#mɛɛndɛt) ɛɛn tʊʊgɛt.
 thanks PST2-1SG-buy-IPFV 3SG in store
 ‘Thanks! I bought it at the store.’

Full pronouns in Kipsigis have a relatively limited distribution. Full pronouns do not typically appear in subject and object position; instead, they usually undergo *pro*-drop, and the subject and object features are indexed on the verb (see §2.4 for more discussion). However, the pronouns in Table 2.8 can appear in three main sorts of constructions. First, they can be used as stand-alone responses to *wh*-questions, as seen in (78).

- (78) a. Koo-∅-keer ng’oo Kiproono?
 PST2-3-see who K.
 ‘Who did Kiproono see?’
 b. Anɛɛ.
 1SG
 ‘Me.’

They can also surface in cases where the standard subject or object marker on the verb is doubled, as in (79), which shows object doubling. Kipsigis subject marking is discussed in §2.4.5, while object marking and the contexts that allow the type of doubling seen in (79) are discussed in §2.4.8.

- (79) In response to *Kootaraach ng’oo Lidya?* ‘Who did Lydia hug?’:
 Koo-∅-taraach-an anɛɛ(ndɛt) Lidya amut.
 PST2-3-hug-1SG.O 1SG L. yesterday
 ‘Lydia hugged me yesterday.’

Finally, full pronouns can also appear in clause-initial position in constructions like the one in (80), where non-verb-initial word order is possible (see §2.5.5 for more discussion).

- (80) a. **Eεchεεk** che ka-∅-ba sɔgɔl.
 1PL REL.PL PST1-3-go.PL school
 ‘It’s us who who went to school (recently).’
 b. **Anεε** ko koo-∅-taraach-an Chepta.
 1SG TOP PST2-3-hug-1SG.O C.
 ‘Cheptoo hugged me.’

2.4 Verbal structure

2.4.1 Overview of Kipsigis verbs

Kipsigis verbs are inflected with information about the tense, aspect, polarity, mood, subject, object, and valence of the verb. In this section, I outline key properties of Kipsigis verbs, though more work remains to develop a complete understanding of verbal inflection in Kipsigis.

Kipsigis and Kalenjin languages more broadly have two verbal inflection classes, which have different morphological profiles in terms of aspect, subject marking, and tonal melody. Many verbs alternate between the two classes with a semantic effect—specifically, anticausative vs. causative interpretation—while many others appear exclusively in one class with no semantic generalization predicting which class a verb is in. Throughout this section, I point out the empirical differences that hold across inflection classes as they become relevant. For more discussion of Kalenjin verb classes, see Toweett (1979); Creider & Creider (1989); Kouneli (2022). In fact, Kouneli (2022) offers a unified analysis of Kipsigis verbal inflection, arguing that the appearance of inflection classes is an epiphenomenon arising from the presence of additional functional structure with one class of verbs. Here, I describe the empirical patterns seen with verbal inflection in terms of inflection class, while acknowledging that these differences do not necessarily reflect underlying inflection classes as morphological primitives.

2.4.2 Tense

Kipsigis distinguishes between non-past (81) and past tense (82). The unmarked form in (81) is appropriate in present and future contexts, while the tense morphemes in (82) are appropriate in different past contexts. Specifically, *ka-* (82a) indicates recent past tense (i.e. events that happened within the past few hours), *kɔɔ-* indicates intermediate past tense (i.e. events that happened yesterday), and *kII-* (82c) indicates distant past tense (i.e. events that happened many years ago).

- (81) ∅-Tyén-ì.
 3-sing/dance-IPFV
 ‘He/she is singing/dancing.’
 ‘He/she will sing/dance.’

- (82) a. **Kà-∅-tyén-ì.**
 PST1-3-sing/dance-IPFV
 ‘He/she was singing/dancing (recently).’
- b. **Kòo-∅-tyén-ì.**
 PST2-3-sing/dance-IPFV
 ‘He/she was singing/dancing.’
- c. **Kii-∅-tyén-ì.**
 PST3-3-sing/dance-IPFV
 ‘He/she was singing/dancing (long ago).’

While *kòo-* is most often used to describe events that happened during the previous day (83a), it is also possible for verbs with *kòo-* to describe events that happened more distantly in the past (83b). For this reason, I consider *kòo-* to be a default past tense within the Kipsigis graded tense system and, as a result, I do not include a parenthetical adverbial indicating when in the past the event occurred in translations of sentences with *kòo-*.

- (83) a. **Kòo-mà-∅-márár Kibéèt ámùt.**
 PST2-NEG-3-dance K. yesterday
 ‘Kibet didn’t dance yesterday.’
- b. **Kòo-á-márár-ì áráwà nè kòo-∅-nyé.**
 PST2-1SG-dance-IPFV month REL.SG PST2-3-pass
 ‘I was dancing last month.’

2.4.3 Aspect

Turning to aspect, Kipsigis distinguishes between perfective, imperfective, and perfect aspect. Perfective aspect is unmarked, as seen in (84).

- (84) a. **Kà-∅-tyén.**
 PST1-3-sing/dance
 ‘He/she sang/danced (recently).’
- b. **Kòo-∅-tyén.**
 PST2-3-sing/dance
 ‘He/she sang/danced.’
- c. **Kii-∅-tyén.**
 PST3-3-sing/dance
 ‘He/she sang/danced (long ago).’

Imperfective aspect is typically realized as the suffix *-e* or *-i* depending on the number of syllables in the verb root and the length of the last vowel in the verb root. Following Kouneli (2022), I assume that the default realization of imperfective aspect is the suffix *-e*; however, in verbs with two or more syllables and with a short vowel in the final syllable, the *-i* allomorph of imperfective aspect is used. For instance, the monosyllabic verb *cham* ‘like’ uses the default *-e*

allomorph of imperfective aspect (85). By contrast, the polysyllabic verb *kanap* ‘lift’ with a short vowel in its final syllable requires the *-i* allomorph of imperfective aspect (86). Imperfective aspect can be seen alongside all Kipsigis tenses in (81) - (82).⁸ Non-past tense requires imperfective aspect, while past tenses are compatible with all aspectual distinctions.

- | | |
|---|--|
| (85) Ø-Chám-è.
3-like-IPFV
‘He/she likes.’
(Kouneli 2022:13) | (86) Ø-Kánàp-ì.
3-lift-IPFV
‘He/she lifts.’
(Kouneli 2022:13) |
|---|--|

There are also several other ways to realize imperfective aspect, which appear in particular contexts and replace the standard *-e* and *-i* imperfective suffixes. For instance, imperfective aspect can also be realized as the suffix *-u* (87). In my data, the *-u* allomorph of imperfective aspect only appears on a subset of verbalized adjectives with the suffix *-iit*, as in (87), or in the modal verb *nyaalu*, which is discussed in §2.5.7.

- (87) Pàypày-íit-ù Nàansí kòt kó-nyór kàanèetìindáyàn.
happy-VBLZ-IPFV N. if 3.SBJV-meet teacher
‘Nancy will be happy if she meets a teacher.’

In a similar vein, the morpheme *-toos* realizes imperfective aspect alongside plural number of the subject (88), while *-chiin* generally realizes the applicative plus imperfective aspect (89).

- (88) Ø-Tyén-tóos làagóok.
3-sing/dance-PL.IPFV children
‘Children are singing/dancing.’
- (89) Kòo-í-kòo-chiin Kibéet Chèepkòèch kitàbóot.
PST2-3-give-APPL.IPFV K. C. book
‘Kibet was giving Chepkoech a book.’

Finally, imperfective aspect is suprasegmentally marked on object markers; while object markers typically bear low tone and assimilate to the verb that they attach to in terms of [ATR] specification (90a), they are marked with high tone and necessarily make the verb [+ATR] in the imperfective aspect (90b). For more on the various allomorphs of imperfective aspect found in Kipsigis, see Toweett (1979:185-194).

- (90) a. Kà-Ø-sús-àn ng’òoktà.
PST1-3-bite-1SG.O dog
‘The dog bit me (recently).’

⁸Based on the discussion here, it is surprising that *tyen* ‘sing/dance’ in (81) - (82) takes the *-i* allomorph of imperfective aspect; because it is monosyllabic, one would expect it to take *-e*. As noted in Toweett (1979:192), *tyen* is simply an exception to the more general pattern, likely because the root is underlyingly polysyllabic *tien* with a short vowel in the final syllable, which then undergoes gliding to *tyen*.

- b. Ø-Sús-án ng'òoktà.
 3-bite-1SG.O.IPFV dog
 'The dog is biting me.'

Unlike imperfective aspect, which is realized via a range of verbal suffixes, perfect aspect is realized using a distinct set of subject markers (§2.4.5), which are summarized in Table 2.9. These subject markers are identical to the subjunctive subject markers (§2.4.5), with an added *r* at the beginning of those that would otherwise be verb-initial.⁹ Examples illustrating perfect aspect are provided in (91). Note that—like all other subject markers—the perfect subject markers appear after tense but before the verb root.

	SG	PL
1	raa-	kεε-
2	rii-	rɔɔ-
3	kɔ-	

Table 2.9: Perfect subject markers in Kipsigis

- (91) a. Kòo-**rúú**-kéer sùgùl.
 PST2-1SG.PERF-see school
 'I had seen the school.'
- b. Kòo-**rú**-kéer sùgùl-í?
 PST2-2SG.PERF-see school-Q
 'Had you (sg) seen the school?'
- c. Kòo-**kó**-kéèr Kibéèt sùgùl.
 PST2-3.PERF-see K. school
 'Kibet had seen the school.'
- d. Kòo-**kéè**-kéer sùgùl.
 PST2-1PL.PERF-see school
 'We had seen the school.'
- e. Kòo-**ròò**-kéer sùgùl-í?
 PST2-2PL.PERF-see school-Q
 'Had you (pl) seen the school?'
- f. Kòo-**kó**-kéèr Kibéèt àk Kíròonò sùgùl-í?
 PST2-3.PERF-see K. and K. school-Q
 'Had Kibet and Kiprono seen the school?'

⁹Kouneli (p.c.) notes that this *r* is optional for some of the speakers who she works with, though I have not explored this optionality in my own work.

2.4.4 Negation

Negation in Kipsigis is marked on the verb with the prefix *ma-* (92).

- (92) a. Kà-**mà**-∅-tyén-ì Làmás.
 PST1-NEG-3-sing/dance-IPFV L.
 ‘Linus wasn’t singing/dancing (recently).’
- b. Kòo-**mà**-∅-tyén-ì Làmás.
 PST2-NEG-3-sing/dance-IPFV L.
 ‘Linus wasn’t singing/dancing.’
- c. Kii-**mà**-∅-tyén-ì Làmás.
 PST3-NEG-3-sing/dance-IPFV L.
 ‘Linus wasn’t singing/dancing (long ago).’

Negation surfaces after tense but before subject marking on the verb (93).¹⁰

- (93) Kìimétyèní.
 Kii-**mà**-í-tyèn-í.
 PST3-NEG-2SG-sing/dance-IPFV
 ‘You (sg) weren’t singing/dancing (long ago).’

2.4.5 Subject marking

Subject markers in Kipsigis are obligatory and sensitive to the inflection class, aspect, and mood of the verb. In the indicative mood, Class I subject markers contain short vowels for first and second person subjects (i.e. local person subjects), and the third person subject is not indexed on the verb. These markers are summarized in Table 2.10. Class II subject markers, on the other hand, contain long vowels for local person subjects, and the third person subject marker is *ɪ-*.¹¹ These markers are summarized in Table 2.11. Third person subject marking is not sensitive to number in Class I or Class II.

	SG	PL
1	a-	kɪ-
2	ɪ-	ɔ-
3		∅

Table 2.10: Class I indicative subject markers

¹⁰While I generally provide three-line glosses for data examples, I use four-line glosses when phonological processes obscure the presence of all of the underlying morphemes. For instance, I use a four-line gloss in (93) because coalescence of /a/ and /i/ makes it impossible to see both the negation morpheme *ma-* and the subject marker *ɪ-*.

¹¹See Kouneli (2022) for arguments that *ɪ-* is not a subject marker, but rather an epenthetic vowel that is inserted to realize additional functional structure that is present in Class II—but not Class I—verbs. Here, I call *ɪ-* a third person subject marker in a purely descriptive sense, since Class II verbs with third person subjects bear the *ɪ-* prefix.

	SG	PL
1	aa-	kɪ-
2	ɪ-	ɔɔ-
3		ɪ-

Table 2.11: Class II indicative subject markers

A sample indicative verb paradigm is provided in (94) for a Class I verb and in (95) for a Class II verb. In (95), the verb *put* ‘fall’ bears an additional suffix *-ye* in the plural, which is a dedicated plural suffix for intransitive verbs. As mentioned in §2.3.7 and seen in (94) - (96), Kipsigis is a *pro*-drop language, meaning that subject pronouns do not generally co-occur alongside subject markers.

- | | |
|--|--|
| <p>(94) a. Kòo-á-tàràach Lídyà ámùt.
PST2-1SG-hug L. yesterday
‘I hugged Lydia yesterday.’</p> <p>b. Kòo-í-tàràach Lídyà ámùt.
PST2-2SG-hug L. yesterday
‘You (sg) hugged Lydia yesterday.’</p> <p>c. Kòo-∅-tàràach Lídyà ámùt.
PST2-3-hug L. yesterday
‘He/she hugged Lydia yesterday.’</p> <p>d. Kòo-kí-tàràach Lídyà ámùt.
PST2-1PL-hug L. yesterday
‘We hugged Lydia yesterday.’</p> <p>e. Kòo-ó-tàràach Lídyà ámùt.
PST2-2PL-hug L. yesterday
‘You (pl) hugged Lydia yesterday.’</p> <p>f. Kòo-∅-tàràach Lídyà ámùt.
PST2-3-hug L. yesterday
‘They hugged Lydia yesterday.’</p> | <p>(95) a. Kòo-áà-pùt ámùt.
PST2-1SG-fall yesterday
‘I fell yesterday.’</p> <p>b. Kòo-íi-pùt ámùt.
PST2-2SG-fall yesterday
‘You (sg) fell yesterday.’</p> <p>c. Kòo-í-pùt ámùt.
PST2-3-fall yesterday
‘He/she fell yesterday.’</p> <p>d. Kòo-kíi-pùt-yè ámùt.
PST2-1PL-fall-PL yesterday
‘We fell yesterday.’</p> <p>e. Kòo-óò-pùt-yè ámùt.
PST2-2PL-fall-PL yesterday
‘You (pl) fell yesterday.’</p> <p>f. Kòo-í-pùt-yé ámùt.
PST2-3-fall-PL yesterday
‘They fell yesterday.’</p> |
|--|--|

As seen in (94) - (95), verbs display a different tonal melody depending on the person value of the subject; for both Class I and Class II verbs, local person subjects co-occur with one tonal melody, while third person subjects co-occur with a different tonal melody. It is not entirely clear how to characterize these two different tone patterns, and it is possible that different tonal melodies appear across different tense and aspect combinations. In (94) - (95), for instance, all verbs are in the default past tense with perfective aspect. In (94), the Class I verb *taraach* ‘hug’ bears a L.H melody with local person subjects and a L.L melody with third person subjects. In (95), the Class II verb *put* ‘fall’ is L with local person subjects and H with third person subjects.

On the other hand, according to Kouneli (2022), monosyllabic Class I verbs are uniformly H and monosyllabic Class II verbs uniformly L in the present tense with imperfective aspect. Fleshing out the tonal melodies that occur with different types of subjects and verb roots across tenses and aspects remains an important area for future research.

In the subjunctive mood, whose distribution is discussed in §2.5.6, subject markers are largely the same across verb classes. Subjunctive Class I and Class II subject markers only differ in the length of the vowel of the third person subject marker: *kɔ-* for Class I verbs and *kɔɔ-* for Class II verbs. In the remaining cells of the paradigm, the subjunctive subject markers are identical to the Class II indicative subject markers. These markers are summarized in Table 2.12 for Class I and in Table 2.13 for Class II.

	SG	PL
1	a(a)-	kεε-
2	ɪɪ-	ɔɔ-
3		kɔ-

Table 2.12: Class I subjunctive subject markers

	SG	PL
1	a(a)-	kεε-
2	ɪɪ-	ɔɔ-
3		kɔɔ-

Table 2.13: Class II subjunctive subject markers

Additionally, according to Driemel & Kouneli (2022a), there are two types of subjunctive mood in the language, which differ only in the length of the first person singular subject marker. In the most common type of subjunctive (termed “Type II” by Driemel & Kouneli 2022a), this morpheme is *a-*, whereas in the other type of subjunctive (termed “Type I”), this morpheme is *aa-*. The distribution of these two different types of subjunctive mood is discussed in §2.5.6.

A sample subjunctive verb paradigm is provided in (96), where the relevant subjunctive verb is embedded under the matrix verb *mach* ‘want’. This construction uses the Type II subjunctive and, as a result, the subjunctive subject marker in (96a) is *a-* with a short vowel. Note also the presence of the suffix *-i* in several sentences in (96); this suffix can attach to verbs with local person subjects, hence its glossing as ‘L(OCAL) P(ERSON)’.¹²

¹²What conditions the distribution of *-i* ‘LP’ is an open question. In some cases (e.g. with certain verbs and aspects), this morpheme seems truly optional. Yet in other cases (e.g. after morphemes like *iisye* ‘AP.IPFV’), *-i* is preferred if not obligatory. Likewise, after other suffixal verbal morphemes, it seems as though *-i* is dispreferred if not impossible.

- (96) a. Máchè Kíbèèt áwèè s̀pítáàlì.
 Ø-Mách-è Kíbèèt á-wà-ì s̀pítáàlì.
 3-want-IPFV K. 1SG.SBJV-go.SG-LP hospital
 ‘Kibet wants me to go to the hospital.’
- b. Máchè Kíbèèt ìiwèè s̀pítáàlì.
 Ø-Mách-è Kíbèèt ìi-wà-ì s̀pítáàlì.
 3-want-IPFV K. 2SG.SBJV-go.SG-LP hospital
 ‘Kibet wants you (sg) to go to the hospital.’
- c. Ø-Mách-è Kíbèèt kó-pèl-íis chòorwé-nyóʊn èen làbàtéet.
 3-want-IPFV K. 3.SBJV-win-AP friend-1SG.POSS.SG in race
 ‘Kibet wants my friend to win the race.’
- d. Máchè Kíbèèt kéèbèè s̀pítáàlì.
 Ø-Mách-è Kíbèèt kéè-bà-ì s̀pítáàlì.
 3-want-IPFV K. 1PL.SBJV-go.PL-LP hospital
 ‘Kibet wants us to go to the hospital.’
- e. Máchè Kíbèèt óòbèè s̀pítáàlì.
 Ø-Mách-è Kíbèèt óò-bà-ì s̀pítáàlì.
 3-want-IPFV K. 2PL.SBJV-go.PL-LP hospital
 ‘Kibet wants you (pl) to go to the hospital.’
- f. Ø-Mách-è Kíbèèt kó-pèl-íis chòoròndò-kyóʊk èen làbàtéet.
 3-want-IPFV K. 3.SBJV-win-AP friends-1SG.POSS.PL in race
 ‘Kibet wants my friends to win the race.’

2.4.6 Object marking

Like subject markers, object markers are also obligatory in Kipsigis. Here I outline the object marking pattern seen in basic transitive sentences, but see §2.4.8 for discussion of object marking in ditransitives. Kipsigis object markers are not sensitive to the inflection class of the verb. Table 2.14 outlines the object markers found in Kipsigis. As with Class I subject marking, third person objects are not overtly indexed on the verb.

	SG	PL
1	-an	-εεch
2	-m	-aak
3		∅

Table 2.14: Object markers

A sample verb paradigm showing object marking is provided in (97). Note that sentences with third person objects can have an independent object pronoun, though it is typically *pro*-dropped.

- (97) a. Kòo-∅-tàràach-**àn** Lìdyá ámùt.
PST2-3-hug-1SG.O L. yesterday
'Lydia hugged me yesterday.'
- b. Kòo-∅-tàràach-**in** Lìdyá ámùt.
PST2-3-hug-2SG.O L. yesterday
'Lydia hugged you (sg) yesterday.'
- c. Kòo-∅-tàràach Lìdyá ({ínéè / ínèèndèt}) ámùt.
PST2-3-hug L. 3SG 3SG yesterday
'Lydia hugged him/her yesterday.'
- d. Kòo-∅-tàràach-**èech** Lìdyá ámùt.
PST2-3-hug-1PL.O L. yesterday
'Lydia hugged us yesterday.'
- e. Kòo-∅-tàràach-**àak** Lìdyá ámùt.
PST2-3-hug-2PL.O L. yesterday
'Lydia hugged you (pl) yesterday.'
- f. Kòo-∅-tàràach Lìdyá ({íchéèk / íchéègèt}) ámùt.
PST2-3-hug L. 3PL 3PL yesterday
'Lydia hugged them yesterday.'

Local person (i.e. first and second person) object markers are obligatory; local person pronouns cannot be used as postverbal objects in lieu of object markers, even when focused. Examples (98) - (99) show this generalization with new information and corrective focus, respectively. Here, it is impossible to use full local person pronouns instead of object markers, as shown in the b. examples. Instead, in these contexts, speakers use the standard object marker (the a. examples) or invoke an alternative focus construction (the c. examples; see §2.5.5 for more discussion).

- (98) In response to *Kataraach ng'oo Kibeet?* 'Who did Kibet hug?':
- a. Kà-∅-tàràach-**àn**.
PST3-3-hug-1SG.O
'He hugged ME.'
- b. *Kà-∅-tàràach {**ánéè / ánèèndèt**}.
PST3-3-hug 1SG 1SG
Intended: 'He hugged ME.'
- c. {**Ánéè / ánèèndèt**} né kà-∅-tàràach.
1SG 1SG REL.SG PST3-3-hug
'It's ME who he hugged.'
- (99) I want to correct Linus, who mistakenly thinks that Kibet hugged Nancy. I say *Aacha, mataraach Kibeet Naansi...* 'No, Kibet didn't hug Nancy...' and continue:
- a. Kà-∅-tàràach-**àn**.
PST3-3-hug-1SG.O
'He hugged ME.'

- b. *Kà-∅-tàràach {ánéè / ánéèndèt}.
 PST3-3-hug 1SG 1SG
 Intended: ‘He hugged ME.’
- c. {Ánéè / ánéèndèt} né kà-∅-tàràach Kibéèt.
 1SG 1SG REL.SG PST3-3-hug Kibet
 ‘It’s ME who Kibet hugged.’

By contrast, as seen in (97c) and (97f), third person pronouns can serve as postverbal objects, even in the absence of focus or any other clear discourse considerations.

In addition, local person object markers cannot generally co-occur alongside full pronouns. Doubling of this sort is disallowed in out-of-the-blue contexts and those in which the object is not focused. For instance, the context in (100) establishes the first person object as an aboutness topic (Frey 2004), which is incompatible with focus. In such a context, doubling of the first person object marker with a full pronoun is infelicitous (100b). Speakers report that doubling the object marker with the long form of the pronoun is highly unnatural, while the short form of the pronoun is slightly more natural but still dispreferred to (100a).

(100) Kipkoech says to me *Mwawan agɔba inyee!* ‘Tell me about yourself!’ so I reply:

- a. Kòo-∅-tàràach-àn Lidyá ámùt.
 PST2-3-hug-1SG.O L. yesterday
 ‘Lydia hugged me yesterday.’
- b. #Kòo-∅-tàràach-àn {ánéè / ánéèndèt} Lidyá ámùt.
 PST2-3-hug-1SG.O 1SG 1SG Lydia yesterday
 ‘Lydia hugged me yesterday.’

On the other hand, when the object is focused, doubling of the local person object marker with a full pronoun is possible, though it is still dispreferred to the version of the sentence with just the object marker. Speakers accept sentences like (101b) when prompted, but consistently offer sentences like (101a) first and express a preference for this type of construction over the one with doubling.

(101) In response to *Kootaraach ng’oo Lidya?* ‘Who did Lydia hug?’:

- a. Kòo-∅-tàràach-àn Lidyá ámùt.
 PST2-3-hug-1SG.O L. yesterday
 ‘Lydia hugged me yesterday.’
- b. Kòo-∅-tàràach-àn {ánéè / ánéèndèt} Lidyá ámùt.
 PST2-3-hug-1SG.O 1SG 1SG Lydia yesterday
 ‘Lydia hugged ME yesterday.’

When a third person object is focused, there is no possible doubling, given that third person objects are never marked on the verb. Instead, speakers use full third person pronouns in these contexts (102), and *pro*-drop is dispreferred. Both types of third person pronouns are possible in these focus contexts.

- (102) In response to *Kootaraach ng'oo Lidya?* 'Who did Lydia hug?':
 Kòo-Ø-tàràach {ínéè / ínéèndèt} Lìdyá àmùt.
 PST2-3-hug 3SG 3SG L. yesterday
 'Lydia hugged HIM/HER yesterday.'

Taken together, the data in (97) - (102) show that local person objects are necessarily marked on the verb (optionally alongside full pronouns in some contexts), while third person objects are never marked on the verb and can freely be realized as full pronouns.

The facts presented so far in this section suggest the general verbal template in (103). Tense, negation, and subject marking precede the verb root, while imperfective aspect and object marking follow the verb root. Imperfective aspect and object marking, however, are never both realized segmentally after the verb root. When imperfective aspect and object marking co-occur, imperfective aspect is realized tonally, as seen in (90) in §2.4.3.

- (103) Tense - Negation - Subject marker - Root - {Imperfective, Object marker}

2.4.7 Valence-altering morphology

Kipsigis has a number of processes that alter the valence of the verb. Impersonal constructions remove the subject of a transitive verb (104). The morphological realization of the impersonal involves the addition of first person plural subject marking and the third person tonal melody to the verb. A baseline active example with a first person plural subject and local person tone on the verb is shown in (104a), while the corresponding impersonal is shown in (104b). The main difference between these two examples is the change in the verb's tonal melody, as well as the presence of the local person suffix in (104a).

- (104) a. Kòokyáldèè àrtéet.
 Kòo-kí-ál-dà-ì àrtèet.
 PST2-1PL-buy-IT-LP goat/sheep
 'We sold a goat/sheep.'
- b. Kòokyàldà àrtéet.
 Kòo-kì-àl-dà àrtéet.
 PST2-1PL-buy-IT cow
 'The sheep/goat was sold.'

I classify this construction as an impersonal—rather than a passive—because the notional object remains an object syntactically (105), and it is impossible to represent the notional agent within any kind of prepositional phrase (106).

- (105) Kí-chám-án.
 1PL-love-1SG.O
 'I'm loved.'

- (106) *Kí-chám-án {àk / èen} Chéeróonò.
 1PL-love-1SG.O with in C.
 Intended: ‘I’m loved by Cheronono.’

In a similar vein, the antipassive *-iis* and its allomorphs remove the object of a transitive verb (107). While the verb *am* ‘eat’ can generally take an object (108), the addition of the imperfective allomorph of *-iis* in (109) rules out the presence of an overt object.

- (107) Ø-Ám-iisyé Kìpkòèch.
 3-eat-AP.IPFV K.
 ‘Kipkoech is eating.’
- (108) Ø-Ám-è Kìpkòèch bàandéek.
 3-eat-IPFV K. maize
 ‘Kipkoech is eating maize.’
- (109) *Ø-Ám-iisyé Kìpkòèch bàandéek.
 3-eat-AP.IPFV K. maize
 Intended: ‘Kipkoech is eating maize.’

Reflexive constructions as well as reciprocals can be built using the suffix *-kèè*. Objects of basic transitive verbs can be reflexivized (110), as can IOs (111a) and DOs (111b) in ditransitives. To my knowledge, Kipsigis reflexives with *-kèè* are subject-oriented (i.e. the subject of the sentence is the antecedent to the reflexive argument).

- (110) Kòo-Ø-máas-kèè Chéebéèt.
 PST2-3-hit-REFL C.
 ‘Chebet hit herself.’
- (111) a. Kòo-áà-kóo-chì-kèè kitàbóót.
 PST2-1SG-give-APPL-REFL book
 ‘I gave myself a book.’
- b. Kòo-áà-mùt-í-kèè Kibéèt.
 PST2-1SG-bring-APPL-REFL K.
 ‘I brought myself to Kibet.’

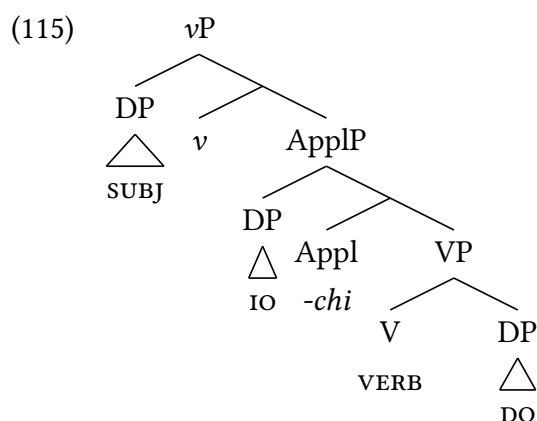
Now, I turn to a process that increases the verb’s number of arguments. The applicative morpheme *-chi* introduces a second object to create a ditransitive structure. (112) illustrates an applied recipient, while (113) shows an applied beneficiary. Both objects in (112) - (113) receive the same unmarked case, which differentiates them from the tonally-marked subject (§2.5.2).

- (112) Kòo-í-kòo-chì Kibéèt Chéeróonò kitàbóót.
 PST2-3-give-APPL K. C. book
 ‘Kibet gave Cheronono a book.’
- (113) Tóos ù-chèeng’-chì Chèeptá màrîndáyán-í?
 MOD 2SG.SBJV-look.for-APPL C. dress-Q
 ‘Could you look for a dress for Cheptoo?’

The suffix *-chi* disappears when there is no applied argument (114).

- (114) Á-chéeng’-è chóorwáyán.
 1SG-look.for-IPFV friend
 ‘I’m looking for a friend.’

Alternations like the one in (113) - (114) suggest the existence of additional functional structure in ditransitives that houses *-chi* and is absent in basic transitives. I identify this structure as ApplP following Pykkänen (2008) and suggest that Kipsigis ditransitives have the basic structure in (115). The DO is base-generated as the complement of the verb, while the IO is base-generated in Spec,ApplP.



Evidence in support of the Kipsigis ditransitive structure in (115) comes from quantifier-variable binding. In applicative sentences with IO-DO word order, IO quantifiers can bind DO variables (116a), but DO quantifiers cannot bind IO variables (116b). In (116a), the quantified IO *laakweet age tógól* ‘every child’ can bind the pronoun within the DO *kitabónyım* ‘their book’. However, in (116b), the quantified DO *kitabóvst age tógól* ‘every book’ cannot bind the pronoun within the IO *siiriındenyım* ‘its author’.

- (116) a. Kòo-áà-kóo-chì làakwéet àgé tógól kítábó-nyım.
 PST2-1SG-give-APPL child every book-3SG.POSS.SG
 ‘I gave [every child]_i their_i book.’
- b. *Kòo-áà-kóo-chì siiriındé-nyım kítábóvst àgé tógól.
 PST2-1SG-give-APPL author-3SG.POSS.SG book every
 Intended: ‘I gave [every book]_i to its_i author.’

The pattern in (116) follows if the IO asymmetrically c-commands the DO as in (115).

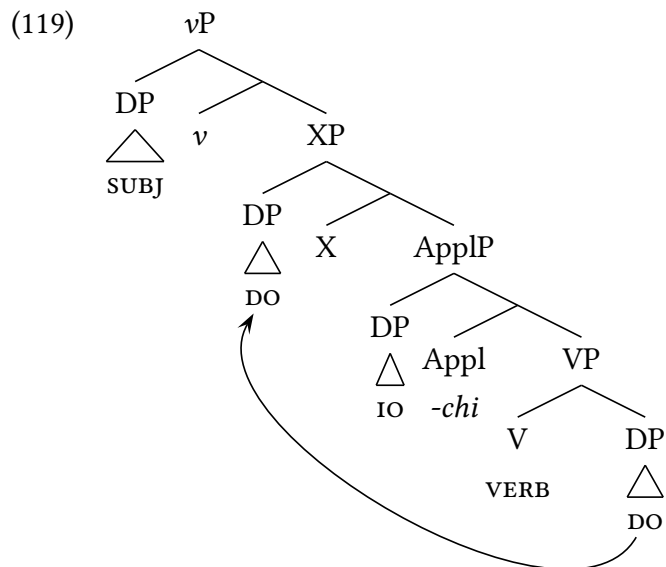
However, the binding pattern reverses in applicative sentences with DO-IO word order. With this object order, IO quantifiers cannot bind DO variables (117a), but DO quantifiers can bind IO variables (117b). In (117a), the quantified IO *laakweet age tógól* ‘every child’ cannot bind the pronoun within the DO *kitabónyım* ‘his book’. Likewise, in (117b), the quantified DO *kitabóvst age tógól* ‘every book’ can bind the pronoun within the IO *siiriındenyım* ‘its author’.

- (117) a. *Kòo-áà-kóo-**chì** kítàbó-nyìm làakwéet àgé tógòl.
 PST2-1SG-give-APPL book-3SG.POSS.SG child every
 Intended: ‘I gave [every child]_i their_i book.’
- b. Kòo-áà-kóo-**chì** kítàbówt àgé tógòl sirîndé-nyìm.
 PST2-1SG-give-APPL book every author-3SG.POSS.SG
 ‘I gave [every book]_i to its_i author.’

In this way, the binding pattern in (117) is the reverse of that in (116). Importantly, this binding pattern does not arise simply to avoid cataphora because there is no general ban on cataphora in Kipsigis; pronouns can precede R-expressions that antecede them, as seen in (118).

- (118) Kán kòo-∅_i-kèer káamée-nyìn_i, kòo-∅-wà ndòonyó Kípròndò_i.
 when PST2-3-see mother-3SG.POSS.SG PST2-3-go.SG market K.
 ‘When he_i saw his_i mother, Kiprono_i went to the market.’

In light of these facts, to derive DO-IO order, I assume short scrambling of the DO to a position above the IO, as illustrated in (119).



In (117b), the DO is interpreted in its derived position where it c-commands the IO. A DO quantifier is, then, able to bind an IO variable in this—and only this—configuration. For more on Kipsigis word order flexibility and scrambling, see §2.5.1.

2.4.8 Object marking in ditransitives

In ditransitives, when both objects are third person, neither object is indexed on the verb (120). This pattern parallels the one seen in §2.4.6 for monotransitives. In most examples in this section, the applicative *-chi* is realized as *-i*, which is an instance of phonologically-conditioned allomorphy; *-chi* surfaces as *-i* after alveolar obstruents and triggers palatalization of the preceding consonant, if it is a stop.

- (120) a. Kòo-∅-máas-í Nàansí (ínéèndèt) Kibéèt.
 PST2-3-hit-APPL N. 3SG K.
 ‘Nancy hit Kibet for him/her.’
- b. Kòo-í-mút¹-í Nàansí Kibéèt (ínéèndèt).
 PST2-3-bring-APPL N. K. 3SG
 ‘Nancy brought him/her to Kibet.’

By contrast, when one of either the IO or the DO is a local person (i.e. first or second person), it surfaces as an object marker on the verb alongside the applicative suffix, regardless of its grammatical function. In these cases, the applicative surfaces as one of two allomorphs: *-i* or *-u*. In configurations with a third person IO and a local person DO—notated 3 > local—the object marker tracks the DO and the applicative is realized as *-i* (121).

- (121) 3 > local → DO marked on verb
- a. Kòoímútyàn Nàansí ínéèndèt.
 Kòo-í-mút-í-**àn** Nàansí ínéèndèt.
 PST2-3-bring-APPL-1SG.DO N. 3SG
 ‘Nancy brought me to him/her.’
- b. Kòoímútyìn Nàansí ínéèndèt.
 Kòo-í-mút-í-**ìn** Nàansí ínéèndèt.
 PST2-3-bring-APPL-2SG.DO N. 3SG
 ‘Nancy brought you (sg) to him/her.’
- c. Kòoímútyèech Nàansí ínéèndèt.
 Kòo-í-mút-í-**èech** Nàansí ínéèndèt.
 PST2-3-bring-APPL-1PL.DO N. 3SG
 ‘Nancy brought us to him/her.’
- d. Kòoímútyàak Nàansí ínéèndèt.
 Kòo-í-mút-í-**àak** Nàansí ínéèndèt.
 PST2-3-bring-APPL-2PL.DO N. 3SG
 ‘Nancy brought you (pl) to him/her.’

On the other hand, in configurations with a local person IO and a third person DO—notated local > 3—the object marker tracks the IO and the applicative is realized as *-u* (122), which is the venitive morpheme in Kipsigis.¹³

- (122) Local > 3 → IO marked on verb
- a. Kòoímútwàn Nàansí ínéèndèt.
 Kòo-í-mút-ú-**àn** Nàansí ínéèndèt.
 PST2-3-bring-VEN-1SG.IO N. 3SG
 ‘Nancy brought him/her to me.’

¹³Directional morphology like the venitive is used in ditransitives in a wide range of languages; in fact, Driemel et al. (2020) incorporate these morphemes into an analysis of ditransitive person restrictions across languages.

- b. Kòoímùtùun Nàansí ínéèndèt.
 Kòo-í-mút-ú-ìn Nàansí ínéèndèt.
 PST2-3-bring-VEN-2SG.IO N. 3SG
 ‘Nancy brought him/her to you (sg).’
- c. Kòoímútùwèech Nàansí ínéèndèt.
 Kòo-í-mút-ú-èech Nàansí ínéèndèt.
 PST2-3-bring-VEN-1PL.IO N. 3SG
 ‘Nancy brought him/her to us.’
- d. Kòoímútùwàak Nàansí ínéèndèt.
 Kòo-í-mút-ú-àak Nàansí ínéèndèt.
 PST2-3-bring-VEN-2PL.IO N. 3SG
 ‘Nancy brought him/her to you (pl).’

Just as with basic transitives, these local person object markers are obligatory in ditransitives; full local person pronouns cannot be postverbal IOs or DOs in lieu of object markers (123).

- (123) a. *Kòo-í-mút-í-í Chéebéèt {ánéè / ánéèndèt} ínéèndèt.
 PST2-3-bring-APPL C. 1SG 1SG 3SG
 Intended: ‘Chebet brought me to him / him to me.’
- b. *Kòo-í-mút-í-í Chéebéèt {ínyèè / ínyèèndèt} ínéèndèt.
 PST2-3-bring-APPL C. 2SG 2SG 3SG
 Intended: ‘Chebet brought you (sg) to him / him to you (sg).’
- c. *Kòo-í-mút-í-í Chéebéèt {éechéèk / éechéègèt} ínéèndèt.
 PST2-3-bring-APPL C. 1PL 1PL 3SG
 Intended: ‘Chebet brought us to him / him to us.’
- d. *Kòo-í-mút-í-í Chéebéèt {óokwèèk / óokwèègèt} ínéèndèt.
 PST2-3-bring-APPL C. 2PL 2PL 3SG
 Intended: ‘Chebet brought you (pl) to him / him to you (pl).’

Finally, local person object markers in ditransitives cannot be doubled by full pronouns. While speakers simply disprefer this type of doubling in basic transitives (100) - (101), they consistently deem sentences like those in (124) ungrammatical, even in focus contexts.

- (124) a. In response to *Káimúti ng’òò Làmás?* ‘Who did Linus bring to him/her?’:
 *Kà-í-mút-í-àn Làmás ínéèndèt {ánéè / ánéèndèt}.
 PST3-3-bring-APPL-1SG.DO L. 3SG 1SG 1SG
 Intended: ‘Linus brought me to him/her.’
- b. In response to *Káimúti ng’òò Làmás?* ‘Who did Linus bring him/her to?’:
 *Kà-í-mút-ú-àn Làmás {ánéè / ánéèndèt} ínéèndèt.
 PST3-3-bring-APPL-1SG.IO L. 1SG 1SG 3SG
 Intended: ‘Linus brought him/her to me.’

An additional empirical complication arises when there are two local person objects in a ditransitive; local > local object configurations are ungrammatical in Kipsigis (Bossi 2021). (125) - (126) illustrate this pattern for singular local > local contexts, though the same is true for plural ones and ones with a combination of singular and plural objects. Regardless of which object is realized as an object marker—be it the IO (a) or the DO (b)—local > local structures are ungrammatical. As before, I use IO > DO notation to label the data examples.

(125) *1 > 2

- a. *Kòo-í-mút-ú-**àn** Kíbèèt {ínyèè / ingyèèndèt}.
 PST2-3-bring-*VEN*-1SG.IO K. 2SG 2SG
 Intended: ‘Kibet brought you (sg) to me.’
- b. *Kòo-í-mút-í-**ìn** Kíbèèt {ánéè / ánéèndèt}.
 PST2-3-bring-*APPL*-2SG.DO K. 1SG 1SG
 Intended: ‘Kibet brought you (sg) to me.’

(126) *2 > 1

- a. *Kòo-í-mút-ú-**ìn** Kíbèèt {ánéè / ánéèndèt}.
 PST2-3-bring-*VEN*-2SG.IO K. 1SG 1SG
 Intended: ‘Kibet brought me to you (sg).’
- b. *Kòo-í-mút-í-**àn** Kíbèèt {ínyèè / ingyèèndèt}.
 PST2-3-bring-*APPL*-1SG.DO K. 2SG 2SG
 Intended: ‘Kibet brought me to you (sg).’

Stacking of local person object markers is likewise impossible, regardless of their order (127).

- (127) a. *Kòo-í-mút-ú-**ìn-àn** Kíbèèt.
 PST2-3-bring-*VEN*-2SG.IO-1SG.DO K.
 Intended: ‘Kibet brought me to you (sg).’
- b. *Kòo-í-mút-í-**àn-ìn** Kíbèèt.
 PST2-3-bring-*APPL*-1SG.DO-2SG.IO K.
 Intended: ‘Kibet brought me to you (sg).’

As seen previously, there are no such restrictions on 3 > local or local > 3 configurations; these object combinations are entirely grammatical. In this way, the restriction only arises in local > local contexts.

To express these local > local object combinations—as in ‘Nancy called me for you (sg)’—speakers offer basic transitive sentences with contextually recoverable beneficiaries (128) or paraphrase to avoid the ditransitive construction altogether (129).

- (128) Kòo-∅-kúr-**àn** Nàansí.
 PST2-3-call-1SG.O N.
 ‘Nancy called me (for you (sg)).’

- (129) Kòo-∅-kúr-ǎn Nàansí áà-nyòon áà-tàarèt-ìn.
 PST2-3-call-1SG.O N. 1SG.SBJV-come 1SG.SBJV-help-2SG.O
 ‘Nancy called me to come help you (sg).’

The restriction on the co-occurrence of local person objects is also obviated when either object surfaces clause-initially for focus, in which case the object that is not clause-initial is realized as an object marker on the verb. This empirical pattern is discussed in more detail in §2.5.5.1 and in Bossi (2021).

2.5 Clausal structure

2.5.1 Word order

This section summarizes key properties of Kipsigis word order, drawing heavily on Bossi & Diercks (2019). While I discuss the main empirical patterns and provide the relevant data examples here, I refer readers to Bossi & Diercks (2019) for more detailed discussion and an analysis of Kipsigis clause structure. As seen throughout this grammar sketch, Kipsigis is a verb-initial language that allows significant postverbal word order flexibility; VSO and VOS word orders are both possible in the language, and speakers translate these sentences in the same way (130).¹⁴

- (130) a. **Kii-∅-keer** chiita tɛɛta. VSO
 PST3-3-see person cow
 ‘A person saw a cow (long ago).’
 (Bossi & Diercks 2019:ex. 2)
- b. **Kii-∅-keer** tɛɛta chiita. VOS
 PST3-3-see cow person
 ‘A person saw a cow (long ago).’

Non-verb-initial word orders, on the other hand, are ungrammatical (131), though see §2.5.5 for discussion of two constructions that allow a constituent to surface pre-verbally. The sentences in (131) have the same intended meaning as those in (130).

- (131) a. *Chiita **kii-∅-keer** tɛɛta. SVO
 (Bossi & Diercks 2019:ex. 3)
- b. *Tɛɛta **kii-∅-keer** chiita. OVS
- c. *Chiita tɛɛta **kii-∅-keer**. SOV
- d. *Tɛɛta chiita **kii-∅-keer**. OSV

Similarly, non-verbal predicates like predicate nominals (132) and adjectives (133) also appear in clause-initial position. In this way, Kipsigis is more accurately classified as a predicate-initial language—rather than just a verb-initial one.

¹⁴Transcription of the data examples from Bossi & Diercks (2019) has been updated here to reflect the phonetic properties of Kipsigis more accurately, given my deepened understanding of the language since 2019.

- (132) **Kaaneetiindet** ne myε Kiproono.
 teacher REL.SG good K.
 ‘Kiproono is a good teacher.’
 (Bossi & Diercks 2019:ex. 8)
- (133) **Eecheen** arto-chu.
 big.PL goats-those
 ‘Those goats are big.’

The word order flexibility seen in (130) is widespread in the language. In ditransitive sentences with three arguments, all six possible verb-initial word orders are permitted (134).

- (134) a. **Koo-∅-koo-chi** laakwεet tεeta baandεek. V-S-IO-DO
 PST2-3-give-APPL child cow maize
 ‘The child gave the cow some maize.’
 (Bossi & Diercks 2019:ex. 5)
- b. **Koo-∅-koo-chi** laakwεet baandεek tεeta. V-S-DO-IO
 PST2-3-give-APPL child maize cow
 ‘The child gave the cow some maize.’
- c. **Koo-∅-koo-chi** baandεek laakwεet tεeta. V-DO-S-IO
 PST2-3-give-APPL maize child cow
 ‘The child gave the cow some maize.’
- d. **Koo-∅-koo-chi** baandεek tεeta laakwεet. V-DO-IO-S
 PST2-3-give-APPL maize cow child
 ‘The child gave the cow some maize.’
- e. **Koo-∅-koo-chi** tεeta laakwεet baandεek. V-IO-S-DO
 PST2-3-give-APPL cow child maize
 ‘The child gave the cow some maize.’
- f. **Koo-∅-koo-chi** tεeta baandεek laakwεet. V-IO-DO-S
 PST2-3-give-APPL cow maize child
 ‘The child gave the cow some maize.’

The same pattern is seen with transitive sentences that contain an adverb, where the six verb-initial word orders are also all possible (Bossi & Diercks 2019:ex. 6).

Despite this word order flexibility, VSO is the pragmatically neutral word order in the language, as evidenced by “out-of-the-blue” diagnostics. In contexts like the one in (135), the assumption is that the newscaster does not know the audience’s familiarity with the current topic, so they cannot assume any shared knowledge. In these situations, VSO word order is preferred, and all other word order permutations are deemed less appropriate.

- (135) A newscaster for a Kipsigis broadcast gets on the air and wants to announce new information to the audience.

- a. Koo-∅-min laagook baandɛɛk komyɛ. VSO-Adv
 PST2-3-plant child maize well
 ‘The children planted the maize well.’
 (Bossi & Diercks 2019:ex. 7)
- b. #Koo-∅-min laagook komyɛ baandɛɛk. #VS-Adv-O
- c. #?Koo-∅-min baandɛɛk laagook komyɛ. #?VOS-Adv
- d. #?Koo-∅-min baandɛɛk komyɛ laagook. #?VO-Adv-S
- e. #Koo-∅-min komyɛ laagook baandɛɛk. #V-Adv-SO
- f. #Koo-∅-min komyɛ baandɛɛk laagook. #V-Adv-OS

This word order flexibility does not appear constrained by the phonological weight of arguments or their other morphological features like definiteness or animacy (Bossi & Diercks 2019:4-5). Instead, the VSO/VOS alternation is sensitive to information structure; constituents bearing any kind of information structural designation—including focus, contrast, and topic—surface most naturally in the position immediately after the verb. For instance, *wh*-words, which bear inherent focus, appear most naturally directly after the verb (136). For more on constituent questions like those in (136), see §2.5.4.2.

- (136) a. Kɔɔ-∅-al ng’oo artɛɛt amut?
 PST2-3-buy who goat yesterday
 ‘Who bought a goat yesterday?’
 (Bossi & Diercks 2019:ex. 21)
- b. Kɔɔ-∅-al nee kiirwaagiindet amut?
 PST2-3-buy what chief yesterday
 ‘What did the chief buy yesterday?’
- c. Kɔɔ-∅-al au kiirwaagiindet artɛɛt?
 PST2-3-buy when chief goat
 ‘When did the chief buy a goat?’

Likewise, the constituent targeted by a *wh*-question bears new information focus and surfaces most naturally immediately after the verb. The question in (137) targets the object of the verb; as a result, VOS word orders are deemed natural in response to the *wh*-question (137b) - (137c). VSO word order is also possible here because it is pragmatically neutral (137a). All other word orders are considered less natural in response to the *wh*-question.

- (137) In response to *Koomin nee laagook komyɛ?* ‘What did the children plant well?’:
- a. Koo-∅-min laagook **baandɛɛk** komyɛ. VSO-Adv
 PST2-3-plant children maize well
 ‘The children planted the maize well.’
 (Bossi & Diercks 2019:ex. 23)
- b. Koo-∅-min **baandɛɛk** laagook komyɛ. VOS-Adv
 PST2-3-plant maize children well
 ‘The children planted the maize well.’

- c. Koo-∅-min **baandεek** komyε laagook. VO-Adv-S
 PST2-3-plant maize well children
 ‘The children planted the maize well.’

Beyond *wh*-words and target constituents of *wh*-questions, which are both focused, constituents with other information structural designations are also preferred in the immediately postverbal position. These include contrastive constituents (138) as well as aboutness topics (139). As previously, word orders in which the contrastive or topical constituent is not immediately postverbal are dispreferred (see Bossi & Diercks 2019:exs. 57, 40 for the other word order judgements).

- (138) In response to *Kɔɔam mcheeleek Kiproono?* ‘Did Kiproono eat the rice?’:
 Aachicha, kɔɔ-∅-am **Chεεpkɔεch** mcheeleek.
 no PST2-3-eat C. rice
 ‘No, Chepkoech ate the rice.’
 (Bossi & Diercks 2019:ex. 57)
- (139) In response to *Mwawan kiit agɔbɔ baandεek*. ‘Tell me something about the maize.’:
 Kii-∅-min **baandεek** Kiproono.
 PST3-3-plant maize K.
 ‘Kiproono planted the maize.’
 (Bossi & Diercks 2019:ex. 40)

Given the wide range of constituents that can appear in the immediately postverbal position, Bossi & Diercks (2019) conclude that this position is reserved for “discourse differentiated” constituents, which is a categorization that crosscuts many different information structural notions.

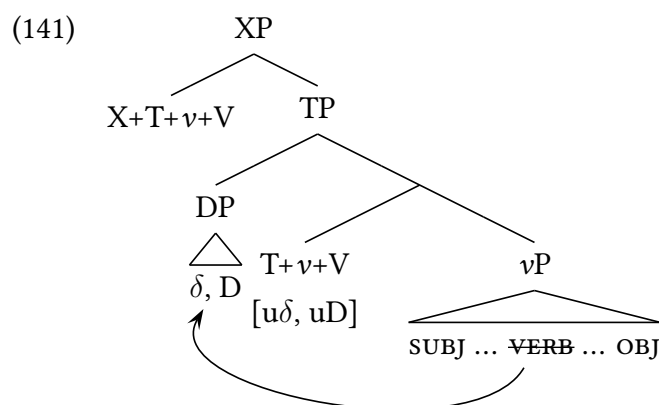
However, an additional empirical complication is the fact that constituents that are not nominal are dispreferred in the immediately postverbal position, even when they are discourse differentiated. For instance, in response to the manner *wh*-question in (140), VS-Adv-O word order is preferred (140a) over word orders in which the adverb is immediately postverbal (140b) - (140c). In this context, then, the constituent with new information focus is in the second position after the verb—rather than immediately postverbal.

- (140) In response to *Koominda ano laagook baandεek?* ‘How did the children plant the maize?’:
 a. Koo-∅-min laagook **komyε** baandεek. VS-Adv-O
 PST2-3-plant children well maize
 ‘The children planted the maize well.’
 (Bossi & Diercks 2019:ex. 54a)
- b. #Koo-∅-min **komyε** laagook baandεek. V-Adv-SO
 PST2-3-plant well children maize
 ‘The children planted the maize well.’
 (Bossi & Diercks 2019:ex. 33)

- c. #Koo-∅-min komyɛ baandɛɛk laagook.
 PST2-3-plant well maize children
 ‘The children planted the maize well.’

V-Adv-OS

Based on these observations, Bossi & Diercks (2019) conclude that the immediately postverbal position houses constituents that are discourse differentiated and that have nominal properties. When the discourse differentiated constituent is not nominal, it surfaces in the second position after the verb and the subject appears immediately postverbally. In order to account for these word order patterns, Bossi & Diercks (2019) propose the structure in (141).



The verb undergoes head movement to a functional projection above TP but below CP—labeled XP in (141). Constituent movement to the immediately postverbal position is driven by a discourse feature $[\delta]$, as well as a nominal feature $[D]$. When the discourse differentiated constituent has a $[\delta]$ feature and a $[D]$ feature, it alone moves to the immediately postverbal position. However, when the discourse differentiated constituent only has a $[\delta]$ feature, it moves to Spec,TP, and the closest nominal with a $[D]$ feature (i.e. the subject) moves to an outer specifier of Spec,TP. In this way, movement to the immediately postverbal position in Kipsigis can be viewed as an instance of mixed A/ \bar{A} -movement, in that it is triggered by both A and \bar{A} features (see e.g. Van Urk 2015; Colley & Privoznov 2020; Scott 2021; Branan & Erlewine 2022; Drummond 2023).

2.5.2 Case marking

Kipsigis displays a typologically rare marked nominative case system (Kouneli 2019; Kouneli & Nie 2021), in which subjects of intransitive and transitive verbs are both marked with nominative case, while nominals with any other grammatical role are unmarked, surfacing in what is called the “oblique” case. The morphological realization of case in Kipsigis is tonal; nominative case marking on nominals generally involves the replacement of the noun’s typical tonal melody with a L-H*-L melody. Some examples are provided in Table 2.15, from Kouneli & Nie (2021:ex. 14).

Oblique		Nominative		Gloss
làakwéet	L.H	làakwèet	L.L	‘child’
ng’óoktá	H.H	ng’òoktà	L.L	‘dog’
sògàróok	L.L.H	sògàròok	L.H.L	‘sugar’
mágáséet	H.H.H	màgásèet	L.H.L	‘skin’
mùgùulèldá	H.HL.L.H	mùgùulèldà	L.H.H.L	‘heart’
kòòkwàatìnwèèk	HL.L.L.HL	kòòkwàatìnwèèk	L.H.H.L	‘villages’

Table 2.15: Nominative L-H*-L tonal melody on nouns

Nominal modifiers like adjectives, possessive pronouns, and demonstratives also have different nominative vs. oblique case forms. In particular, as shown in Kouneli & Nie (2021), the nominative forms of these modifiers are segmentally identical to their oblique forms but have the opposite tonal pattern across-the-board; that is, nominative and oblique modifiers are opposite in every tonal specification. This pattern is shown in (142) with the plural proximal demonstrative suffix *-chu* (§2.3.6), the adjective *mur* ‘dirty’, and the plural suffix *-een*. In the oblique case, *-chu* ‘PROX.PL bears a low tone, while in the nominative case, it bears a high tone. Likewise, the adjective *mur* ‘dirty’ and its plural suffix *-een* bear high and low tones, respectively, in the oblique case, which become low and high tones, respectively, in the nominative case.

- (142) a. Á-kéer-é pèelée(k)-chù múr-èen. OBL: L H.L
 1SG-see-IPFV elephants-PROX.PL.OBL dirty-PL.OBL
 ‘I see these dirty elephants.’ (Kouneli & Nie 2021:ex. 15)
- b. Ø-Rúày pèelèe(k)-chú mùr-éen. NOM: H L.H
 3-run.PL elephants-PROX.PL.NOM dirty-PL.NOM
 ‘These dirty elephants are running.’ (Kouneli & Nie 2021:ex. 16)

This across-the-board switch in tonal polarity is the first of its kind to be described in the linguistic literature and raises theoretical questions about how this type of morphological process can be analyzed (for possible approaches, see Kouneli & Nie 2021; Jolin 2023; Sande To Appear).

All postverbal subjects bear nominative case, regardless of the order of the postverbal constituents. As discussed in §2.5.1, Kipsigis allows for a significant degree of postverbal word order flexibility, although this flexibility does not impact nominative case marking. This pattern can be seen in (143a) - (143b), where the subject *Kibeeet* ‘Kibet’ bears the same H.LH nominative tonal melody whether the sentence displays VSO or VOS word order. However, there are two constructions in Kipsigis—discussed at length in §2.5.5—which allow nominals to surface preverbally. When the subject of the sentence appears in clause-initial position, nominative case marking is lost, as evidenced by the L.HL oblique tonal melody on *Kibeeet* ‘Kibet’ in (143c).

- (143) a. Ø-Chám-è Kíbèet tètá né òo. NOM: H.HL
 3-like-IPFV K.NOM cow REL.SG big.SG
 ‘Kibet likes the big cow.’ (Kouneli 2019:ex. 22a)

- b. Ø-Chám-è tètá né òo **Kibéèt.** NOM: H.HL
 3-like-IPFV cow REL.SG big.SG K.NOM
 ‘Kibet likes the big cow.’ (Kouneli 2019:ex. 22b)
- c. **Kibéèt** kò Ø-chám-è tètá né òo. OBL: L.HL
 K.OBL TOP 3-like-IPFV cow REL.SG big.SG
 ‘Kibet likes the big cow.’ (Kouneli 2019:ex. 22c)

2.5.3 Coordination

In this section, I describe some of the coordination strategies found in Kipsigis. This section is not intended to be an exhaustive discussion of coordination in the language, and there are certainly many more Kipsigis coordinators that I do not discuss here.

2.5.3.1 Conjunction

The conjunction *ak* or *akɔ* can be used to coordinate nouns (144), verbs (145), and clauses (146).

- (144) Meny-e Taputany **ak** Koilong’eet kokweet agεεngε.
 live-IPFV T. and K. village one
 ‘Taputany and Koilong’eet live in the same village.’
- (145) Kɔɔ-Ø-nam **ak** kɔɔ-Ø-tyar mbiireet Kiproono.
 PST2-3-catch and 3.SBJV-throw ball K.
 ‘Kiproono caught and threw the ball.’
- (146) I-pwaat-e piik kole mogoriot Lidya **akɔ** εen iman ko-mogoriot.
 3-think-IPFV people C rich.person L. and in truth 3.SBJV-rich.person
 ‘People think that Lydia is rich and, in truth, she is rich.’

When *ak* or *akɔ* is used to coordinate clauses, the verb in the second clause appears in the subjunctive mood, evidenced by the subjunctive subject marker on *mogoriot* ‘rich person’ in (146).

Kipsigis also has two adversative coordinators that translate to ‘but’ in English: *lakini*, which is a borrowing from Swahili, and *ngandan*. In my data, these words only coordinate two full clauses, as seen in (147) - (148), but further research is necessary to understand if this is the only way that they can be used.

- (147) Ø-Poot-e ng’ookta **lakini** Ø-ru-e pageεt.
 3-bark-IPFV dog but 3-sleep-IPFV cat
 ‘The dog is barking, but the cat is sleeping.’
- (148) Ng’am Cheerop **ngandan** ming’in.
 smart C. but small
 ‘Cherop is smart, but she is young.’

2.5.3.2 Disjunction

The disjunction *anan* can be used to coordinate nouns (149), verbs (150), or clauses (151).

- (149) Koo-a-keer Kibeet **anan** Kiproono.
 PST2-1SG-see K. or K.
 ‘I saw Kibet or Kiproono.’
- (150) Kɪtʃam **anan** mɛɛtʃam kɪtʃabʊt?
 Kɪtʃ-I-cham **anan** ma-I-cham kɪtʃabʊt?
 PST3-2SG-like or NEG-2SG-like book
 ‘Did you like or not like the book?’
- (151) Kɔ-kɔ-wa Kibeet **anan** kɔ-kɔ-it Cheeroono.
 PST1-3.PERF-go.SG K. or PST1-3.PERF-arrive C.
 ‘Kibet left or Cheronno arrived.’

Unlike with the conjunction *ak* or *akɔ*, when the disjunction *anan* is used to coordinate clauses, the verb in the second clause is in the indicative mood and bears the full range of tense and aspect inflection; this can be seen in (151), where the verb in the second disjunct bears the recent past tense morpheme *ka-* and the 3rd person perfect aspect subject marker *kɔ-*.

2.5.4 Questions

2.5.4.1 Polar questions

Polar question formation in Kipsigis involves two morphophonological effects: the addition of the vowel *-i* at the end of the clause and the addition of a super high tone to the final syllable of the clause. When the clause ends in a consonant, addition of *-i* is obligatory and it bears the super high tone (152). However, when the clause ends in a vowel, addition of *-i* is optional; if *-i* is added, then it bears the super high tone (153a), but if *-i* is not added, then the existing clause-final vowel is lengthened and bears the super high tone (153b). In vowel-final clauses, the more common polar question formation strategy is the one seen in (153b), which involves lengthening the clause-final vowel and adding a super high tone to it.

- (152) Koo-i-pel-iis Kipchirchir ɛɛn urerieet-í?
 PST2-3-win-AP K. in race-Q
 ‘Did Kipchirchir win the race?’
- (153) a. Keemutjaan mamaí?
 Kɔ-ii-mut-j-aan mama-í?
 PST1-2SG-bring-APPL-1SG.O mom-Q
 ‘Did you (sg) bring me to mom?’
- b. Kɔɔ-∅-al Kibeet tɛtɛ́áa?
 PST2-3-buy K. cow-Q
 ‘Did Kibet buy a cow?’

Note that the addition of *-í* in (153a) to the clause ending in the word *mama* does not trigger the typical process of vowel coalescence of *a* and *i* to *ee*, as described in §2.2.2. Likewise, addition of the necessarily [+ATR] particle *-í* does not change the ATR specification of the [-ATR] word *mama*, as would be expected given the Kipsigis ATR harmony system (§2.2.2). These patterns suggest that the clause-final polar question particle *-í* is outside the domain that phonological processes apply to.

Besides the addition of *-í* or the super high tone, no structural changes are required for polar question formation (e.g. word order remains the same as in declarative sentences). However, it is possible to add the modal particle *toos* to the beginning of a polar question alongside its canonical polar question marking. The semantic contribution of *toos* is not clear at present; speakers report that there is a slight difference between polar questions with and without *toos* and suggest that the version with *toos* is perhaps “more polite” than the one without it. However, I leave determining the exact interpretation of polar questions with *toos* to future research.

- (154) (Toos) koo-Ø-wa chi sɔgɔl amut-í?
 MOD PST2-3-go.SG person school yesterday-Q
 ‘Did anyone go to school yesterday?’

Biased polar questions are used when the speaker expects a particular answer—either affirmative or negative—to the polar question. Kipsigis uses the clause-final particle *-ís* to form biased polar questions. When the speaker expects an affirmative response to their biased polar question, *-ís* attaches to an affirmative sentence (155). By contrast, when the speaker expects a negative response to their biased polar question, *-ís* attaches to a negated sentence (156).

- (155) Kiprono, Kibet, and Kipchirchir all played Kipchirchir’s favorite game last night. Because Kipchirchir is so good at this game, he always wins. I assume that Kipchirchir won the game, but for confirmation, I ask Kiprono:
 Koo-i-pel-iis Kipchirchir ɛɛn urerieet-ís?
 PST2-3-win-AP K. in game-Q.BIAS
 ‘Kipchirchir won the game, right?’
- (156) Kiprono, Kibet, and Kipchirchir all played Kipchirchir’s least favorite game last night. Because Kipchirchir is terrible at this game, he always loses. I assume that Kipchirchir lost the game, but for confirmation, I ask Kiprono:
 Moo-i-pel-iis Kipchirchir ɛɛn urerieet-ís?
 PST2.NEG-3-win-AP K. in game-Q.BIAS
 ‘Kipchirchir didn’t win the game, right?’

It is an open question whether the particle *-ís* can be decomposed into the standard polar question particle *-í* plus an additional suffix *-s*. However, to my knowledge, there is no other context in the language where one finds the suffix *-s*.

It is impossible to add the modal particle *toos* to a biased polar question with *-ís* (157).

- (157) Yesterday was Tuesday, which is a school day for you. Your mother believes that you went to school yesterday, but she heard a rumor that a group of students skipped and wants to

confirm that you weren't in that group. She asks:

***Toos** koo-ka-i-wa-i sʊgʊl amut-ís?

MOD PST2-?-2SG-go.SG-LP school yesterday-Q.BIAS

Intended: 'You (sg) went to school yesterday, right?'

2.5.4.2 Constituent questions

Constituent questions with the *wh*-words *ng'oo* 'who', *nee* 'what', (*εεn*) *ano* 'where', and *au* 'when' can be formed using two different strategies. It is possible for the *wh*-words to remain postverbal, as seen in the a. examples in (158) - (161), or to surface clause-initially with a form of the relativizer, as seen in the b. examples in (158) - (161).¹⁵ As mentioned in §2.5.1, *wh*-words surface most naturally immediately after the verb.

- (158) a. Koo-i-le kʊʊ-∅-lany ng'oo keetit?
 PST2-2SG-say PST2-3-climb who tree
 'Who did you (sg) say climbed the tree?'
 b. Ng'oo ne koo-∅-le Kibeet ∅-cham-e laakwɛɛt age tʊgʊl?
 who REL.SG PST2-3-say K. 3-like-IPFV child every
 'Who did Kibet say that every child likes?'
- (159) a. Keegeer nee?
 Ka-i-keer nee?
 PST1-2SG-see what
 'What did you (sg) see?'
 b. Nee ne koo-∅-keer?
 what REL.SG PST2-3-see
 'What did he/she see?'
- (160) a. I-pwaat-e Kibeet ko-le koo-a-a-put εεn ano?
 3-think-IPFV K. 3.SBJV-LE PST2-1SG-fall in where
 'Where does Kibet think that I fell?'
 b. Ano ele koo-i-keer-chi Kipchɔmba?
 where REL.OBL PST2-2SG-see-APPL K.
 'Where did you (sg) see Kipchumba?'
- (161) a. Ku-∅-chap au keɛgt Kipchirchir?
 PST3-3-make when cake K.
 'When did Kipchirchir make the cake (long ago)?'
 b. Au ele koo-i-keer Kipchɔmba?
 when REL.OBL PST2-2SG-see K.
 'When did you (sg) see Kipchumba?'

¹⁵For more on the structure and interpretation of the types of sentences in the b. examples, see §2.5.5.

Questions targeting why and how an event happened differ from the constituent questions described above and from each other. *Why*-questions make use of the word *kalryan*, which speakers translate as ‘why’ or ‘what happened.’ In *why*-questions, *kalryan* appears clause-initially and introduces a purpose clause, marked with the prefix *si-* (162). In this way, Kipsigis *why*-questions literally translate to ‘What happened such that X.’

- (162) **Kalryan** *si* *mā-i-am-iisye-i?*
 what.happened PURP NEG-2SG-eat-AP.IPFV-LP
 ‘Why aren’t you (sg) eating?’

Given its position in the clause, its ability to introduce an embedded clause, and the potential tense prefix *ka-*, *kalryan* might appear to be a verb. However, there are two reasons to doubt this analysis. First, there do not seem to be forms of *kalryan* with any other tense prefixes (e.g. *kālryan* or *kūlryan*), which suggests that this form is not a productive verb synchronically. Second, the remainder of the word *kalryan* does not display decomposable verbal morphology.

How-questions, on the other hand, make use of an interesting combination of independently attested morphemes. More specifically, the formation of *how*-questions involves the addition of the andative directional suffix *-ta* to the verb and the use of the locative *wh*-word *ano* ‘where’. This method for forming *how*-questions is highly productive, as evidenced by the range of data examples in (163) - (165).

- (163) **Kōɔɔchaptɛɛ** **ano** *chaiik amut?*
 Kōɔɔ-I-chap-ta-I **ano** *chaiik amut?*
 PST2-2SG-make-AND-LP where chai yesterday
 ‘How did you (sg) make chai yesterday?’
- (164) **Koo-∅-ip-ta** **ano** *Kipchirchir urerieet?*
 PST2-3-win-AND where K. game
 ‘How did Kipchirchir win the game?’
- (165) **Koo-∅-sik-ta** **ano** *Kipchirchir tɛeta-nyim?*
 PST2-3-find-AND where K. cow-3SG.POSS.SG
 ‘How did Kipchirchir find his cow?’

Understanding how this particular combination of directional and locative morphemes generates a *how*-question is an open avenue for future research. Perhaps, for instance, the andative morpheme *-ta* adds some type of semantic “slot” for a directional argument, which can then be filled by an adjunct. In this way, a sentence like (165) would literally mean ‘To where did Kipchirchir find his cow?’, which would simply be an idiomatic way to say ‘how’.

Interestingly, there are a number of verbs whose basic meaning is derived via addition of the andative suffix *-ta*. When *how*-questions are built using these verbs, there is apparent doubling of *-ta*. For instance, the verb ‘buy’ is *al* in Kipsigis (166a), while the verb ‘sell’ is *alda* (166b), which is simply *al* plus the andative suffix *-ta* with regular phonological processes applied (§2.2.1).

- (166) a. Kaa-**al** tɛɛta-nyʊʊn.
 PST1.1SG-buy cow-1SG.POSS.SG
 ‘I bought my cow (recently).’
 b. Kaa-**al-ta** tɛɛta-nyʊʊn.
 PST1.1SG-buy-AND cow-1SG.POSS.SG
 ‘I sold my cow (recently).’

The formation of a *how*-question with *al* ‘sell’ involves the standard procedure: the andative suffix *-ta* is added to the verb, and the locative *wh*-word *ano* is used (167a). By contrast, the formation of a *how*-question with *alda* ‘buy’ involves apparent doubling of the andative suffix *-ta* with an epenthetic linker vowel *-i* (167b).¹⁶

- (167) a. Ka-**al-ta** **ano** Kiproono tɛɛta-nyɪɪn?
 PST1-3-buy-AND where K. cow-3SG.POSS.SG
 ‘How did Kiproono buy his cow (recently)?’
 b. Kɔɔ-**al-ta-i-ta** **ano** Kiproono tɛɛta-nyɪɪn?
 PST2-buy-AND-EP-IT where K. cow-3SG.POSS.SG
 ‘How did Kiproono sell his cow?’

In this way, it is possible that the *how*-question in (167b) involves two instances of the andative suffix *-ta* or that the suffix *-ta* that appears in *how*-questions is simply homophonous with the standard andative suffix in Kipsigis.

2.5.5 Preverbal constituents

While Kipsigis is generally verb-initial (§2.5.1), there are two types of elements that allow material to surface before the verb: the relativizers *ne*, *che*, or *ele* (168) and the particle *ko* (169).

- (168) a. In response to *Kooikoochi ng’oo kitabʊʊt?* ‘Who did he give the book to?’:
 Kaamɛɛt-aap Kiproono **ne** koo-i-koo-chi kitabʊʊt.
 mother-of K. REL.SG PST2-3-give-APPL book
 ‘It’s Kiproono’s mother who he gave the book to.’
 b. In response to *Nee ne koochape mʊrɛnɪk?* ‘What were the men making?’:
 Ngeendeek **che** koo-∅-chap-e mʊrɛnɪk.
 beans REL.PL PST2-3-make-IPFV men
 ‘It’s beans that the men were making.’
 c. In response to *Mwawan ele koowa Kiproono!* ‘Tell me where Kiproono went!’:
 Sʊgʊl **ele** koo-∅-wa Kiproono.
 school REL.OBL PST2-go.SG K.
 ‘It’s school where Kiproono went.’

¹⁶See Kouneli 2022 for arguments that *i* is the default epenthetic vowel in Kipsigis.

- (169) In response to *Mwawan agəbə Cheerop!* ‘Tell me about Cherop!’:
 Cheerop **ko** kəɔ-Ø-al amut maarɪndet-aap pəɔl somok ɛɛn ndoonyo.
 C. TOP PST-3-buy yesterday dress-of hundred three in market
 ‘Cherop bought a dress for 300 shillings at the market yesterday.’

This section describes the structure and interpretation of these constructions. While they appear similar on the surface—at least in terms of their effects on word order—I show that constructions with the relativizers *ne*, *che*, or *ele* are biclausal pseudoclefts, while constructions with *ko* are monoclausal and involve constituent movement to a clause-initial topic position.

2.5.5.1 Pseudoclefts with the relativizers *ne*, *che*, or *ele*

Many different types of constituents can appear in clause-initial position alongside a form of the relativizer. Nominals can surface here, regardless of whether they are core arguments like subjects (168a) and objects (168b), or obliques (168c). Just as number and the core argument vs. oblique status of the head of a Kipsigis relative clause dictates the form of the relativizer (§2.3.5), these same facts about the clause-initial constituent determine the form of the relativizer here; singular core arguments in clause-initial position trigger *ne*, plural core arguments in clause-initial position trigger *che*, and obliques in clause-initial position trigger *ele*.

In addition to nominal obliques like *səɔɔl* ‘school’ in (168c), PPs (170) and temporal adverbs (171) can appear clause-initially alongside the relativizer *ele*. While it is possible for the entire PP to surface in clause-initial position (170a), it is also possible for the preposition *ɛɛn* to be incorporated into the verb, in which case only the object of the preposition appears clause-initially (170b).

- (170) In response to *Kəɔal ɛɛn ano Cheerop maarɪndet amut?* ‘Where did Cherop buy a dress yesterday?’:
- a. ɛɛn ndoonyo **ele** kəɔ-Ø-al maarɪndet.
 in market REL.OBL PST2-3-buy dress
 ‘It’s in the market where she bought a dress.’
 - b. Ndoonyo **ele** kəɔ-Ø-al-ɛɛn maarɪndet.
 market REL.OBL PST2-3-buy-in dress
 ‘It’s in the market where she bought a dress.’
- (171) In response to *Kəɔal av Cheerop maarɪndetaap pəɔl somok ɛɛn ndoonyo?* ‘When did Cherop buy a dress for 300 shillings in the market?’:
- Amut **ele** kəɔ-Ø-al Cheerop maarɪndet-aap pəɔl somok ɛɛn ndoonyo.
 yesterday REL.OBL PST2-3-buy C. dress-of hundred three in market
 ‘It was yesterday when Cherop bought a dress for 300 shillings in the market.’

Unlike temporal adverbs, manner adverbs are ungrammatical in clause-initial position with any form of the relativizer (172). In this way, clause-initial position with a form of the relativizer appears restricted to constituents with nominal properties, where nouns, nominal obliques, PPs, and temporal adverbs are sufficiently nominal to surface in this position, while manner adverbs

are not. In fact, this cutoff in terms of what counts as sufficiently nominal parallels that reported in Bossi & Diercks (2019) for the immediately postverbal position.

- (172) In response to *Kaminda ano laagook baandεεk?* ‘How did the children plant the maize?’:
 *Komyε {**ele** / **ne** / **che**} ka-∅-min.
 well REL.OBL REL.SG REL.PL PST1-3-plant
 Intended: ‘They planted it well (recently).’

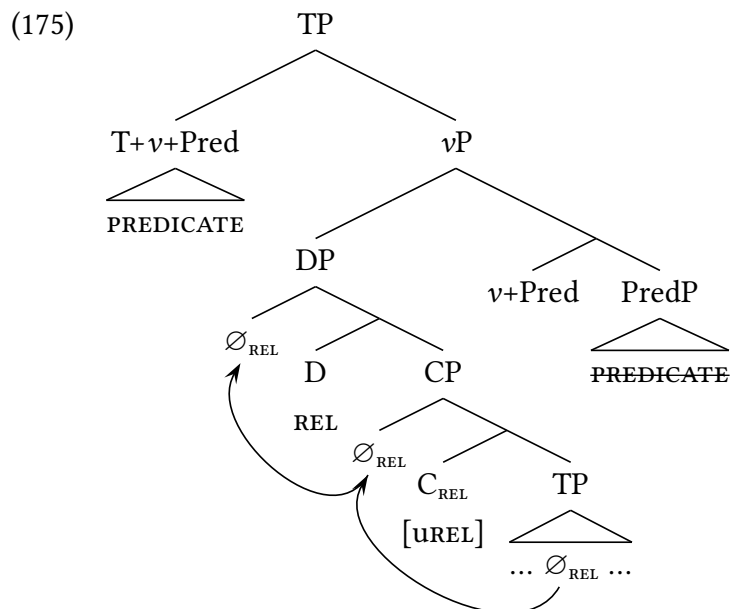
When a constituent appears in clause-initial position in this construction, it receives a focus interpretation. For instance, sentences like those in (168) are appropriate in response to a *wh*-question that targets the clause-initial constituent. These types of sentences are also felicitous in contrastive focus contexts like (173), where the speaker wants to clarify who Kibet hugged.

- (173) Aacha, ma-∅-taraach Kibeet Naansi. Anεε **ne** ka-∅-taraach Kibeet.
 no NEG-3-hug K. N. 1SG REL.SG PST1-3-hug K.
 ‘No, Kibet didn’t hug Nancy. It’s me who Kibet hugged (recently).’

By contrast, this construction is ruled out in contexts that force a topic interpretation for the clause-initial constituent. One such context is found in (174), where the prompt *Mwawan agɔbɔ Cheerop!* ‘Tell me about Cherop!’ establishes Cherop as an aboutness topic (Frey 2004). In this context, the aboutness topic *Cheerop* cannot appear clause-initially with the relativizer *ne*.

- (174) In response to *Mwawan agɔbɔ Cheerop!* ‘Tell me about Cherop!’:
 #Cheerop **ne** kɔɔ-∅-al amut maarɪndɛt-aap pɔgɔl somok εɛn ndoonyo.
 C. REL.SG PST2-3-buy yesterday dress-of hundred three in market
 ‘It’s Cherop who bought a dress in the market yesterday for 300 shillings.’

Turning to the structure of these types of sentences, I suggest that they are biclausal pseudo-clefts (Potsdam & Polinsky 2011), where the clause-initial constituent is a non-verbal predicate and the relativizer introduces a headless relative clause with the relative clause structure introduced in §2.3.5. This pseudo-cleft structure is schematized in (175). The ∅ symbol represents the null nominal head of the relative clause that moves from its base position within the relative clause to Spec,DP above the relativizer in D. Following Bossi & Diercks (2019), I assume that non-verbal predicates in Kipsigis undergo head movement rather than phrasal movement, although nothing crucial in (175) hinges on this distinction; the general pseudocleft structure in (175) is also compatible with an analysis in which the entire PredP moves to a specifier position above the position of the headless relative clause subject.



(177) provides a schematized version of the pseudocleft in (176)—repeated from (173)—to help map the structure in (175) to the data examples seen in this section.

(176) Anεε ne ka-∅-taraach Kibeet.
 1SG REL.SG PST1-3-hug K.
 ‘It’s me who hugged Kibet.’

(177) [Pred Anεε] [RC_{SUBJ} ∅ ne katarach Kibeet].

Evidence for this pseudocleft structure comes from five sources. First, the clause-initial constituent cannot originate within an island for syntactic movement. Focus constructions with a form of the relativizer are sensitive to adjunct (178a), relative clause (178b), and coordination islands (178c).

- (178) a. *Ngeendeek **che** koo-∅-wa Kiproono kotoomwa kɔɔ-∅-chap Chεεpkɔεch.
 beans REL.PL PST2-3-go.SG K. before PST2-3-make C.
 Intended: ‘It’s the beans that Kiproono left before Chepkoech made.’
- b. *Kiproono **ne** koo-∅-keer Chεεpkɔεch ng’ookta ne kɔɔ-∅-sɔs.
 K. REL.SG PST2-3-see C. dog REL.SG PST2-3-bite
 Intended: ‘It’s Kiproono who Chepkoech saw the dog that bit.’
- c. *Pageet **ne** koo-∅-keer Kibeet ng’ookta (ak).
 cat REL.SG PST2-3-see K. dog and
 Intended: ‘It’s a cat that Kiproono saw a dog and.’

In a structure like (175), this type of island sensitivity is predicted because the null head of the relative clause would start out within the island and would need to move from this position during relative clause formation, which is predicted to be impossible.

Second, focus constructions with a form of the relativizer can display multiple instances of tense marking (179a). This possibility is predicted given the structure in (175) because there are two separate instances of T. Yet note that if there are two tense markers present, they must match (179b), which I assume has a semantic explanation. It is possible, though, to have tense marking in the relative clause but not in the matrix clause (168).

- (179) a. Kii anεε ne kiimutyin Kibeet.
Kii anεε ne **kii-i-mut-i-in** Kibeet.
 PST3 1SG REL.SG PST3-3-bring-APPL-2SG.DO K.
 ‘(Long ago) it was me who Kibet brought you (sg) to.’
- b. ***Kii** anεε ne **koo-i-mut-i-in** Kibeet.
 PST3 1SG REL.SG PST2-3-bring-APPL-2SG.DO K.
 Intended: ‘(Long ago) it was me who Kibet brought you (sg) to.’

The third piece of evidence in support of the analysis in (175) comes from patterns of case marking (§2.5.2). In particular, the relativizer in focus constructions always shows nominative case, which is marked via a high tone on the relativizer (180), which is argued to be a realization of D (§2.3.5; Kouneli 2019). I take this case marking pattern to suggest that the entire headless relative clause DP is the subject of the matrix predicate, whose case marking is realized on the relativizer in D. The pseudocleft structure in (175) predicts this behavior, since the headless relative clause introduced by the relativizer is in Spec,νP, where subjects are canonically base-generated.

- (180) a. Ng’òò né ∅-chám-è làakwèet ágè tòngól?
 who REL.SG.NOM 3-like-IPFV child every
 ‘Who does every child love?’
- b. [_{Pred} Ng’òò] [_{RC_{SUBJ}} ∅ né chámè làakwèet ágè tòngól].
- c. *Ng’òò nè ∅-chám-è làakwèet ágè tòngól?
 who REL.SG.OBL 3-like-IPFV child every
 Intended: ‘Who does every child love?’

Fourth, headless relative clauses and predicate fronting are independently attested in the language; in this way, there is no special machinery required to derive a structure like (175). In particular, relative clauses that look just like the constituent with *ne*, *che*, or *ele* in focus constructions exist independently in the language, as seen in (181) and discussed in §2.3.5. Likewise, Kipsigis is generally predicate-initial, as discussed in §2.5.1. Given the existence of headless relative clauses and predicate fronting in the language, there is no novel machinery required to generate the pseudocleft structure in (175).

- (181) a. In response to *Kooitaraach ng’ookta aion?* ‘Which dog did you hug?’:
 Koo-a-taraach **ne** ∅-cham-e Kibeet.
 PST-1SG-hug REL.SG 3-like-IPFV K.
 ‘I hugged the one that Kibet loves.’
- b. [_{TP} Kooataraach [_{RC_{OBJ}} ∅ ne chame Kibeet]].

The fifth piece of evidence in support of this pseudocleft analysis comes from the fact that it is impossible to have object marking on the verb in the relative clause that tracks the features of the clause-initial constituent, even when this object marking is obligatory in basic monoclausal sentences. Such a pattern suggests that the clause-initial constituent is never syntactically the object of the verb—a fact that is captured given the pseudocleft structure in (175). For instance, as discussed in §2.4.6, local person (i.e. first and second person) objects are obligatorily marked on the verb. However, this same type of object marking is impossible when the notional local person object surfaces clause-initially (182). This pattern is expected given a biclausal pseudocleft structure because the clause-initial constituent is never the object of the verb in the relative clause; instead, it is a non-verbal predicate that exists outside of the relative clause.¹⁷

- (182) a. **Anɛɛ** ne koo-∅-keer Kiproono.
 1SG REL.SG PST2-3-see K.
 ‘It’s me who Kiproono saw.’
- b. ***Anɛɛ** ne koo-∅-keer-**an** Kiproono.
 1SG REL.SG PST2-3-see-1SG.O K.
 Intended: ‘It’s me who Kiproono saw.’

In fact, these focus constructions obviate the restriction discussed in §2.4.8; when either notional local person object in a ditransitive appears clause-initially with a form of the relativizer, the *local > local restriction disappears (183) - (184). The disappearance of this restriction falls out of a pseudocleft structure because the clause-initial constituent is never actually within the relative clause to trigger object marking.

- (183) Obviating *1 > 2
- a. Inyee ne kooimutwan Kibeet.
Inyee ne koo-i-mut-u-**an** Kibeet.
 2SG REL.SG PST-3-bring-APPL-1SG.IO K.
 ‘It’s you (sg) who Kibeet brought to me.’
- b. **Anɛɛ** ne kooimutyin Kibeet.
Anɛɛ ne koo-i-mut-i-**in** Kibeet.
 1SG REL.SG PST-3-bring-APPL-2SG.DO K.
 ‘It’s me who Kibeet brought you (sg) to.’
- (184) Obviating *2 > 1

¹⁷Exactly how to derive the lack of object marking that tracks the clause-initial constituent in pseudoclefts depends on the analysis of object marking. If object marking is agreement, its disappearance in a pseudocleft rests on the assumption that the null nominal head of the relative clause lacks person features that correspond to those of the clause-initial constituent. If object marking is cliticization (i.e. morphological incorporation of a pronoun; Harizanov 2014), its disappearance in a pseudocleft is predicted because this pronoun is never within the relative clause.

- a. Anɛɛ ne kooimutuun Kibeet.
Anɛɛ ne koo-i-mut-u-**in** Kibeet.
 1SG REL.SG PST-3-bring-APPL-2SG.IO K.
 ‘It’s me who Kibeet brought to you (sg).’
- b. Inyee ne kooimutyan Kibeet.
Inyee ne koo-i-mut-i-**an** Kibeet.
 2SG REL.SG PST-3-bring-APPL-1SG.DO K.
 ‘It’s you (sg) who Kibeet brought me to.’

In previous work (Bossi 2021), I argued that the obviation of the *local > local restriction in these focus constructions supported a morphological—rather than syntactic—account of the restriction. However, this work treated these focus constructions as monoclausal, involving focus movement of the clause-initial constituent. Given this structure, local > local configurations with and without object fronting have a shared syntax at the relevant point in the sentence’s derivation, and so the ungrammaticality of local > local ditransitives without focus fronting cannot be due to syntactic malformation. Yet given the updated understanding of the syntax of these focus constructions presented here, the argument in Bossi (2021) no longer holds. This is because local > local ditransitive sentences without focus have a fundamentally different syntax than biclausal pseudoclefts, in which one notional object surfaces clause-initially as a non-verbal predicate.

Returning to evidence for the biclausal pseudocleft structure in (175), an apparent challenge comes from patterns of reconstruction. In particular, the clause-initial constituent can be interpreted in a position within the relative clause with respect to binding and *wh*-scope. Starting with binding, (185) shows a baseline Principle C effect in Kipsigis; the null pronominal subject cannot be co-referential with the R-expression *Kiplang’at* ‘Kiplangat’ in the object, presumably because this configuration would lead to illicit binding of the R-expression.

- (185) Koo- $\emptyset_{i/j}$ -keer pichaart-aap **Kiplang’at_j**.
 PST2-3-see picture-of K.
 ‘He_{i/j} saw a picture of Kiplangat_j.’

When the clause-initial constituent in a focus construction contains an R-expression, it still cannot be co-referential with the subject of the relative clause (186), which suggests that the clause-initial constituent must be interpreted below the subject in the relative clause. If the clause-initial constituent were in such a position at some point, the unavailability of the co-referential reading in (186) would be explained as in (185). Yet if the clause-initial constituent were never in the relative clause as in (175), it is unclear why the co-referential reading would be ruled out here.

- (186) Pichaart-aap **Kiplang’at_j** ne koo- $\emptyset_{i/j}$ -keer.
 picture-of Kiplang’at REL.SG PST2-3-see
 ‘It’s a picture of Kiplang’at_j that he_{i/j} saw.’

In addition, when quantified constituents occur clause-initially in focus constructions, they can scope in their surface position or below the subject of the relative clause. In (187), the quantifier *ng’oo* ‘who’ appears in clause-initial position alongside a relative clause whose subject is the

quantified expression *laakwæet agɛ tɔgɔl* ‘every child’. Crucially, both the surface and inverse scope readings are available. While the surface scope reading is expected given the structure in (175), it is less clear how to derive the inverse scope reading; the clause-initial *wh*-word is base-generated in PredP below the entire relative clause, though even in this position, it is not obviously within the scope of the subject of the relative clause.

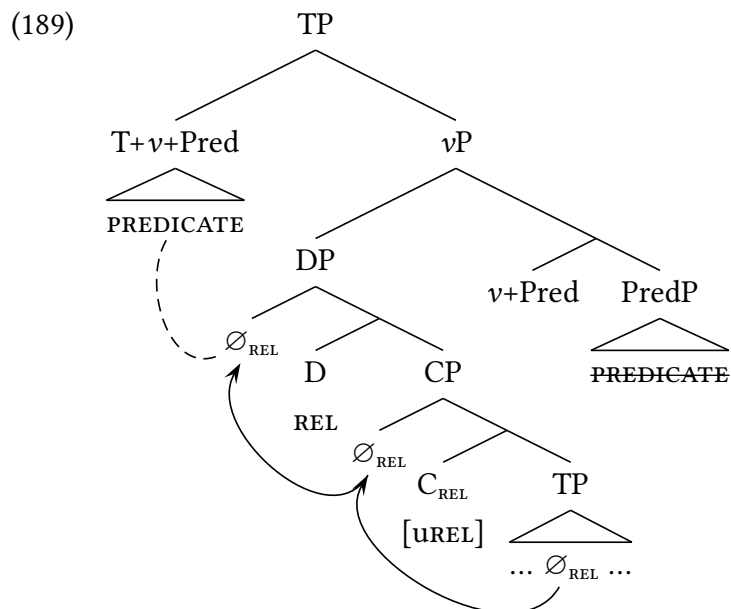
- (187) Ng’oo ne Ø-cham-e laakwæet agɛ tɔgɔl?
 who REL.SG 3-like-IPFV child every
 ‘Who does every child love?’
 ✓ Answer: Kibet. wh > ∀
 ✓ Answer: Chepkoech loves Nancy, Kiplangat loves Linus, Kiprotich loves Nick. ∀ > wh

Furthermore, the availability of the surface and inverse scope readings in (187) cannot be due to quantifier raising of the subject of the relative clause to a position above the *wh*-word *ng’oo* ‘who’ because embedded quantifiers cannot typically scope out of their finite clause. (188) illustrates this point; the $\forall >$ two reading is unavailable, as it would require the embedded quantified expression *chepta agɛ tɔgɔl* ‘every girl’ to undergo quantifier raising to a position above the matrix quantified expression *ng’eeitiik oeeng’u* ‘two boys’.

- (188) Koo-Ø-le ng’eeitiik oeeng’-u ka-Ø-kal Kibeet chepta agɛ tɔgɔl.
 PST2-3-say boys two-NOM PST1-3-mess.with K. girl every
 ‘Two boys said that Kibet messed with every girl.’
 ✓ Kiprotich and Kiprono said that Kibet messed with every girl. two > ∀
 # Kiprotich and Linus said that Kibet messed with Cheptoo, Kipkorir and Victor said that
 Kibet messed with Sharon. *∀ > two

These reconstruction effects cast some doubt on the biclausal pseudocleft structure in (175). However, pseudoclefts often show unexpected connectivity effects across languages, though there is no consensus on the source of this behavior (see e.g. Heycock & Kroch 1999). Given my earlier arguments in favor of a pseudocleft structure, I maintain this structure but acknowledge the challenge that the reconstruction effects pose.

One possible way to integrate these sets of facts is to allow the features of the clause-initial predicate to be inherited by the null head of the relative clause, as schematized in (189). In this way, the clause-initial constituent is never actually within the relative clause but is still able to be interpreted within it. Such a configuration could potentially generate the appearance of reconstruction effects without requiring actual syntactic reconstruction.



In fact, a possible piece of evidence for the feature inheritance analysis in (189) comes from subject marking in these focus constructions. As discussed previously, the clause-initial constituent cannot trigger object marking on the verb within the relative clause. However, subject marking shows a different pattern; the clause-initial constituent *can* optionally trigger subject marking on the verb within the relative clause (190); here, the verb *put* ‘fall’ can show 1SG subject agreement (190a) or can take default third person subject agreement (190b).

- (190) a. **Anεε** ne koo-**aa**-put^j-i keeriingeet.
 1SG REL.SG PST-1SG-fall-APPL hole
 ‘It’s me who fell in a hole.’
- b. **Anεε** ne koo-i-put^j-i keeriingeet.
 1SG REL.SG PST-3-fall-APPL hole
 ‘It’s me who fell in a hole.’

While I do not develop a full argument in support of this approach here, I suggest that this difference stems from the fact that subject marking in Kipsigis is agreement, while object marking is cliticization, as is common cross-linguistically (see e.g. Woolford 2008; Preminger 2009; Nevins 2011; Kramer 2014; Anagnostopoulou 2016; Yuan 2021). Subject agreement on the verb in the relative clause only requires that the subject ϕ -features be present within the relative clause, which can be achieved through the type of feature inheritance in (189). Yet object cliticization requires the object pronoun to be syntactically present within the relative clause, which feature inheritance alone cannot achieve.

2.5.5.2 Topic fronting with *ko*

As with focus constructions, many different types of constituents can appear in clause-initial position with *ko*. Nominals, including subjects (191a) and objects (191b), can surface in this position,

as can some PPs (192) and adverbs with nominal properties (193).

- (191) a. Kiproono **ko** koo-Ø-wa sɔgɔl.
 K. TOP PST2-3-go.SG school
 ‘As for Kiproono, he went to school.’
 b. Susweek **ko** Ø-am-e tuuga.
 grass TOP 3-eat-IPFV cows
 ‘Cows eat grass.’
- (192) Eɛn ndoonyo **ko** kɔɔ-Ø-al Cheerop amut maarɪndɛt-aap pɔgɔl somok.
 in market TOP PST2-3-buy C. yesterday dress-of hundred three
 ‘In the market, Cherop bought a dress for 300 shillings yesterday.’
- (193) Amut **ko** kɔɔ-Ø-al Cheerop maarɪndɛt-aap pɔgɔl somok ɛɛn ndoonyo.
 yesterday TOP PST2-3-buy C. dress-of hundred three in market
 ‘Yesterday, Cherop bought a dress for 300 shillings at the market.’

Regarding sentences like (192), Driemel & Kouneli (2022b:6) observe variation across speakers and across prepositions as to whether PPs can be in this position; in particular, Kouneli (p.c.) notes that most speakers allow PPs with *ɛɛn* ‘in’ to be in this position, while PPs with *ak* ‘with’ are consistently ungrammatical here. While I only have data examples with the preposition *ɛɛn* ‘in’, these PPs are consistently accepted in clause-initial position with *ko* across the speakers who I work with. By contrast, oblique nouns (194) and manner adverbs (195) are unable to appear in this position. In this way, the types of constituents that can surface clause-initially with *ko* are slightly more restricted than those that can be clause-initial with a form of the relativizer.

- (194) *Sɔgɔl **ko** koo-Ø-wa Kiproono.
 school TOP PST2-3-go.SG K.
 Intended: ‘Kiproono went to school.’
- (195) *Mùutyà **kó** Ø-kèet-é Kíbèèt.
 slowly TOP 3-drive-IPFV K.
 Intended: ‘Slowly, Kibet drives.’ (Driemel & Kouneli 2022b:ex. 15)

When a constituent appears in clause-initial position with *ko*, it receives a topic interpretation (Driemel & Kouneli 2022b). Here, I summarize a few key observations from Driemel & Kouneli (2022b), but I refer readers to their work for a full account. When the context establishes particular constituents as contrastive topics, as with *Kibeet* ‘Kibet’ and *Cheebeet* ‘Chebet’ in (196), these constituents surface very naturally in clause-initial position with *ko*. The diagnostic in (196) comes from Buring (2003), who argues that contrastive topics are answers to sub-questions (e.g. what did Kibet eat?) that get at larger discourse questions (e.g. who ate what?).

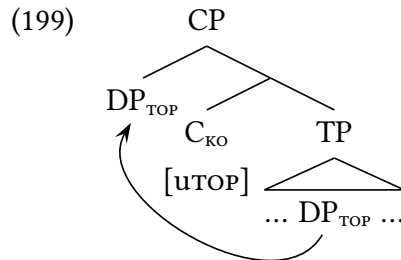
- (196) We were at an event with Kibet, Chebet, and many other people attending, and multiple dishes were available. We want to ask who ate what. What did Kibet eat? What did Chebet eat? (Driemel & Kouneli 2022b:ex. 18)

- a. Kibèèt **kó** kà-∅-ám ng'ééndéek.
 K. TOP PST1-3-eat beans
 'Kibet ate beans.'
- b. Chèebèèt **kó** kà-∅-ám pèendá.
 C. TOP PST1-3-eat meat
 'Chebet ate meat.'

On the other hand, focused elements like *wh*-words (197) and expressions with *mεεkεen* 'only' (198) cannot appear in clause-initial position with *ko*. This pattern suggests that this position is a dedicated topic position, rather than a more general position for constituents with any sort of information structural effect.

- (197) *Ng'áà **kó** kà-∅-ám pèendá?
 who TOP PST1-3-eat meat
 Intended: 'Who ate meat?' (Driemel & Kouneli 2022b:ex. 34)
- (198) *Kibèèt *ínéekéen* **kó** kà-∅-sóòmàn kitàbóót.
 K. only TOP PST1-3-read book
 Intended: 'Only Kibet read a book.' (Driemel & Kouneli 2022b:ex. 35)

I propose that constructions with *ko* involve constituent movement to a clause-initial topic position, as outlined in (199).



Evidence for the structure in (199) comes from four sources. First, topic constructions with *ko* are sensitive to adjunct (200a), relative clause (200b), and coordination islands (200c). This behavior falls out of the structure in (199) because the topical constituent moves from its base position to clause-initial position; this type of movement would not be possible if the topical constituent originated within an island.

- (200) a. *Ng'ookta **ko** koo-∅-raari Cheerop kan kà-∅-wan paageet.
 dog TOP PST-3-laugh Cheerop when PST1-3-chase cat
 Intended: 'Cheerop laughed when the dog chased the cat.'
- b. *Paageet **ko** kà-∅-tarach Cheerop ng'ookta ne kà-∅-wan.
 cat TOP PST-3-hug Cheerop dog REL.SG PST1-3-chase
 Intended: 'Cheerop hugged the dog that chased the cat.'
- c. *Paageet **ko** kà-∅-tarach Cheerop ng'ookta ak.
 cat TOP PST-3-hug Cheerop dog and
 Intended: 'Cheerop hugged the dog and the cat.'

Next, unlike in focus constructions with a form of the relativizer, constructions with *ko* only allow one instance of tense marking (201). This behavior follows from the monoclausal structure in (199), but contrasts with the type of biclausal pseudocleft structure proposed for Kipsigis focus constructions in §2.5.5.1.

- (201) We're talking about all the places that different people were born. Who was born where? Where was Kibet born? Where was Chepkoech born?
- a. Chεεpkɔɛch ko kii-ki-sik-chiin Keericho.
C. TOP PST3-1PL-give.birth-APPL.IPFV Kericho
'Chepkoech was born in Kericho (long ago).'
- b. *Kii Chεεpkɔɛch ko kii-ki-sik-chiin Keericho.
PST3 C. TOP PST3-1PL-give.birth-APPL.IPFV Kericho
Intended: 'Chepkoech was born in Kericho (long ago).'

Third, object marking that tracks the features of the clause-initial constituent is obligatory on the verb in topic constructions with *ko*. As described in §2.4.6, local person objects must be marked on the verb when present. This same type of object marking continues to be obligatory when the notional local person object surfaces clause-initially with *ko* (202). This pattern can be captured on the analysis in (199) because the topical constituent begins the derivation as the object of the verb, where it triggers object marking.

- (202) a. Anεε ko koo-∅-taraach-an Cheepta.
1SG TOP PST2-3-hug-1SG.O C.
'Cheptoo hugged me.'
- b. *Anεε ko koo-∅-taraach Cheepta.
1SG TOP PST2-3-hug C.
Intended: 'Cheptoo hugged me.'

Fourth, topic constructions with *ko* display reconstruction for Principle C, which suggests that it is the topical constituent—rather than any other element in the sentence—that undergoes movement. In (203b), the reading in which *Linus* 'Linus' is co-referential with the subject of the verb *keer* 'see' is ruled out. Given the structure in (199), this behavior arises because the clause-initial constituent begins the derivation in canonical object position below the null subject pronoun; in this position, the null subject pronoun would bind the R-expression, violating Principle C.

- (203) Linus is looking through a big box of pictures. There's a picture of Chebet, one of Kiplangat, one of Lydia, even one of himself. Some of the pictures were stuck together, though, so Linus saw some of them, but didn't see others. I'm explaining which pictures he saw and which pictures he didn't see. I say:
- a. Pichaart-aap Cheebet ko koo-∅-keer.
picture-of C. TOP PST2-3-see
'As the the picture of Chebet, he saw it.'

- b. Lakini pichaart-aap Lmas_j ko ma-∅_{i/*j}-keer.
 but picture-of L. TOP NEG-3-see
 ‘But as for the picture of Linus_j, he_{i/*j} didn’t see it.’

2.5.6 Clausal subordination

In Kipsigis, there are three main strategies for clausal subordination. These include using a subjunctive subordinate clause (204), embedding an indicative clause with no overt complementizer, and embedding an indicative clause with the element *kole* (206). To this point in the dissertation, I have glossed *kole* as C; however, it is actually bi-morphemic, containing the prefix *ko-* ‘3.SBJV’ and *le*, which is discussed in more detail in §2.5.6.3.

- (204) ∅-Mach-e ko-wa Nairobi Koilong’eet.
 3-want-IPFV 3.SBJV-go.SG Nairobi K.
 ‘Koilong’eet wants to go to Nairobi.’
- (205) ∅-Par-e Kiproono koo-a-keer Chepkœech.
 3-think-IPFV K. PST2-1SG-see C.
 ‘Kiproono is under the impression that I saw Chepkœech.’
- (206) I-ngen Cheebet ko-le ∅-ru-e Cheeruyot.
 3-know C. 3.SBJV-C 3-sleep-IPFV C.
 ‘Cheebet knows that Cheruyot is sleeping.’

Verbs that take clausal complements display some degree of flexibility in the embedding strategy that they use—typically with a semantic effect. While it is not the case that any verb can embed a clause using any of the three strategies above, many embedding verbs allow both of the strategies in (204) and (206). In these cases, there is a semantic difference between the two types of constructions. With matrix speech verbs, use of the indicative strategy with *kole* leads to an embedded proposition reading (207a), while use of the subjunctive strategy leads to an embedded command reading (207b), where the matrix attitude holder is instructing that the embedded action be completed.

- (207) a. Ka-∅-chaam Kibeet ko-le ka-∅-chap kimnyeet.
 PST1-3-whisper K. 3.SBJV-C PST1-3-make ugali
 ‘Kibet whispered that he made ugali (recently).’ (Driemel & Kouneli 2022a:ex. 56a)
- b. Ka-∅-chaam-u-an Kibeet a-chap kimnyeet.
 PST1-3-whisper-VEN-1SG.O K. 1SG.SBJV-make ugali
 ‘Kibet whispered to me to make ugali (recently).’ (Driemel & Kouneli 2022a:ex. 56b)

In a similar vein, with perception verbs, use of the indicative strategy with *kole* generates an embedded proposition reading (208a), while use of the subjunctive strategy leads to a control-like reading, where the embedded subject is necessarily co-referential with the matrix subject (208b).

- (208) a. Koo-i-**ruaatit** paageet **ko-le** **ka-∅-nam** mɔriyaat.
 PST2-3-dream cat 3.SBJV-C PST1-3-catch mouse
 ‘The cat dreamed that it caught a mouse.’
- b. Koo-i-**ruaatit** paageet **ko-nam-e** mɔriyaat.
 PST2-3-dream cat 3.SBJV-catch-IPFV mouse
 ‘The cat dreamed of catching a mouse.’

Against this backdrop, in the following subsections, I describe each of the strategies in (204) - (206) in turn, then conclude with a brief discussion of embedded questions.

2.5.6.1 Subjunctive embedded clauses

First, I consider the subjunctive subordination strategy. As noted in Driemel & Kouneli (2022a), there are two types of subjunctive in Kipsigis, which differ in the length of the first person singular subject marker and occur in different contexts. The first person singular subject marker for the Type I subjunctive is *aa-* with a long vowel (§2.4.5). This type of subjunctive appears in conditional clauses (209) and in some temporal adjunct clauses (210), but is generally less common in the language.

- (209) Aa-paypay-iit-u kot **aa**-sich karit.
 1SG-happy-VBLZ-IPFV if 1SG.SBJV-receive car
 ‘I will be happy if I get a car.’
- (210) Kooimutyān Chεεpkɔech Kibeet koon aaputi.
 Koo-i-mut-i-an Chεεpkɔech Kibeet koon **aa**-put-i.
 PST2-3-bring-APPL-1SG.O C. K. when 1SG.SBJV-fall-IPFV
 ‘Chepkoech was bringing me to Kibet when I fell.’

On the other hand, the first person singular subject marker in the Type II subjunctive is *a-* with a short vowel (§2.4.5) and occurs with volitional predicates (211), in purpose clauses (212), and after modal verbs (213; see also §2.5.7). This type of subjunctive is more common in the language than the Type I subjunctive.

- (211) A-mach-e **a**-tun kaaneetiindet.
 1SG-want-IPFV 1SG.SBJV-marry teacher
 ‘I want to marry a teacher.’
- (212) I-mu-an ii-le si **a**-labat-í?
 2SG-scare-1SG.O 2SG.SBJV-LE PURP 1SG.SBJV-run-Q
 ‘Are you scaring me so that I run?’
- (213) Nyaal-u **a**-laach εlmetit.
 MOD-IPFV 1SG.SBJV-wear helmet
 ‘I should wear a helmet.’

In subjunctive clauses, it appears that nominative case marking on subjects is preserved (§2.5.2), as evidenced by the H.HL nominative tonal melody on *Kíbéèt* in (214). However, I leave a full exploration of case assignment in subjunctive clauses for future work.

- (214) Á-mách-é kò-ál Kíbéèt tètá.
 1SG-want-IPFV 3.SBJV-buy K.NOM cow
 ‘I want Kibet to buy a cow.’

2.5.6.2 Indicative embedded clauses with no overt complementizer

The second embedding strategy found in Kipsigis involves embedding an indicative clause with no overt complementizer. This type of clausal embedding is quite restricted in the language and, to my knowledge, only occurs with the speech verb *le* ‘say’ (215) and the biased belief verb *par*, as seen above in (205).¹⁸ In examples like (205) and (215), it is clear that the verb in the embedded clause is in the indicative mood because only verbs in the indicative mood can inflect for tense; in (205), the embedded verb *keer* ‘see’ bears the tense prefix *ka-* ‘PST1’, while in (215), the embedded verb *al* ‘buy’ bears the tense prefix *kɔɔ-* ‘PST2’.

- (215) Koo-Ø-le Cheeroono εεn tuugeet ne mii parak kɔɔ-Ø-al Kibeet tɛɛta.
 PST2-3-say C. in voice REL.SG COP loud PST2-3-buy K. cow
 ‘Cherono said in a loud voice that Kibet bought a cow.’

2.5.6.3 Indicative embedded clauses with the element *kole*

The third embedding strategy—embedding an indicative clause with the element *kole*—is highly productive in the language. A wide range of verbs use this embedding strategy, including *ngen* ‘know’ (206), *pwaat* ‘think’ (216), *kas* ‘hear’ (217), *ruaatit* ‘dream’ (218), *mwa* ‘tell’ (219), among many others.

- (216) I-pwaat-e Cheeroono ko-le kɔɔ-Ø-al Kibeet tɛɛta.
 3-think-IPFV C. 3.SBJV-C PST2-3-buy K. cow
 ‘Cherono thinks that Kibet bought a cow.’
- (217) Kɪɪ-a-kas-εεn-ɪɪn ko-le kɪɪ-Ø-al Kibeet tɛɛta.
 PST3-1SG-hear-in-2SG.O 3.SBJV-C PST3-3-buy K. cow
 ‘I heard from you (long ago) that Kibet bought a cow.’
- (218) Koo-aa-ruaatit ko-le koo-Ø-tyen Kibeet.
 PST2-1SG-dream 3.SBJV-C PST2-3-sing/dance K.
 ‘I dreamed that Kibet sang/danced.’
- (219) Kaa-mwa-chi Cheebeet ko-le Ø-ru-e Cheeruyot.
 PST1.1SG-tell-APPL C. 3.SBJV-C 3-sleep-IPFV C.
 ‘I told Chebet (recently) that Cheruyot is sleeping.’

¹⁸For more on the structure and interpretation of constructions with *par*, see Chapter 4.

These verbs come from a range of lexical classes and include factive verbs like *ngen* ‘know’ and non-factive verbs like *pwaat* ‘think’ and *mwa* ‘tell’.

As mentioned at the beginning of this section, here I include additional discussion of the element *kole*, seen above in (216) - (219). As is common across languages, what appears to be the verb meaning ‘say’—*le* in Kipsigis (215)—can also be used alongside another verb in an attitude report (220).

- (220) Aa-ngen *({ko-le / aa-le}) kɔɔ-∅-al Kibeet tɛɛta.
 1SG-know 3.SBJV-LE 1SG.SBJV-LE PST-3-buy K. cow
 ‘I know that Kibet bought a cow.’

While Kipsigis is highly understudied in general, *kole* has received some attention in the theoretical literature because, in its complementation use, *le* shows a typologically unique pattern of upward-oriented agreement. In particular, it bears subjunctive subject marking that is default third person or that tracks the matrix argument construed as the information source of the attitude report—both of which are illustrated in (220). In addition to the subject marking in (220), *le* can bear extra verbal morphology like imperfective aspect (221a), as well as the applicative suffix and object markers that also track matrix arguments (221b).

- (221) a. Ka-a-mwa-e aa-leelen ka-∅-chɔɔr Kibeet rabɪɪk.
 PST1-1SG-tell-IPFV 1SG.SBJV-LE-IPFV PST1-3-steal K. money
 ‘I was saying that Kibet stole the money.’
 (Driemel & Kouneli 2022a:ex. 16)
- b. Kii-i-mwa-w-aan ii-leen-j-aan kɪ-∅-al Kibeet tɛɛta.
 PST3-2SG-tell-APPL-1SG.O 2SG.SBJV-LE-APPL-1SG PST3-3-buy K. cow
 ‘You told me that Kibet bought a cow.’

In light of these subject and object marking patterns, Diercks & Rao (2019) treat Kipsigis *kole* as a unique upward-agreeing complementizer (hence, the glossing of *le* as C to this point). While upward-oriented subject marking on complementizers is relatively common across Bantu languages, Kipsigis is (to my knowledge) the only language proposed to display upward-oriented object marking, as seen in (221b). However, in more recent work, Driemel & Kouneli (2022a) use the morphological patterns in (220) - (221) to motivate the claim that *le*—even in its complementation use—is uniformly the lexical verb ‘say’. This uniformly verbal analysis captures the full range of morphological patterns seen in (220) - (221) via standard processes of verbal inflection.

However, as observed in Bossi (2023), there are systematic differences across instances of complementation with *le* that pose challenges for the uniformly verbal analysis in Driemel & Kouneli (2022a). First, the extensive inflection on *le* in (221) is only possible when it co-occurs with a matrix speech predicate. While all instances of *le* allow prefixal marking, inflected forms of *le* like those in (221) with imperfective aspect and suffixal marking are ungrammatical with matrix predicates like *keer* ‘see’ (222) and *par* ‘show’ (223), which do not involve speech. When co-occurring alongside these matrix predicates, *le* can only bear prefixal marking, even if the matrix verb shows imperfective aspect (222) or object marking (223).

- (222) I notice that you're dancing, but I don't say anything.
 A-keer-e {aa-le / ko-le / *aa-leelen / *ko-leelen} i-tyen-i.
 1SG-see-IPFV 1SG.SBJV-LE 3.SBJV-LE 1SG.SBJV-LE-IPFV 3.SBJV-LE-IPFV 2SG-dance-IPFV
 'I see that you're dancing.'
- (223) We're hiding under the bed in a game of hide-and-see. We can't speak—otherwise, the
 seeker will find us—but I point to a pair of shoes under the bed.
 Ka-a-par-u-iin {aa-le / ko-le / *aa-leen-j-iin /
 PST1-1SG-show-APPL-2SG.O 1SG.SBJV-LE 3.SBJV-LE 1SG.SBJV-LE-APPL-2SG.O
 *ko-leen-j-iin} koo-mii kwoosiek kitanda arit.
 3.SBJV-LE-APPL-2SG.O PST2-COP shoes bed under
 'I showed you that there were shoes under the bed.'

Second, in its complementation use, *le* can only be modified by adverbs or PPs when it bears extensive verbal inflection (e.g. suffixal marking) and co-occurs with a matrix speech predicate. As seen in (224), the PP *εen arageεneet* 'with surprise' can modify an inflected form of *le* in its complementation use.

- (224) Koo-i-mwa-w-aan ii-leen-j-aan εen arageεneet koo-∅-al Kibeet tεeta.
 PST2-2SG-tell-APPL-1SG.O 2SG.SBJV-LE-APPL-1SG.O in surprise PST2-3-buy K. cow
 'You told me with surprise that Kibet bought a cow.'

However, this same type of PP modification is impossible when *le* co-occurs with a matrix predicate like *ngen* 'know' (225), which does not involve speech and is incompatible with imperfective or object-marked forms of *le*. Note that this ungrammaticality is due to the placement of the PP modifier and not the meaning of the sentence (226).

- (225) *Koo-a-nai aa-le εen arageεneet koo-∅-choo Kibeet prasmuk.
 PST2-1SG-know 1SG.SBJV-LE in surprise PST2-3-steal K. potatoes
 Intended: 'I knew with surprise that Kibet stole the potatoes.'
- (226) Koo-a-nai εen arageεneet aa-le koo-∅-choo Kibeet prasmuk.
 PST2-1SG-know in surprise 1SG.SBJV-LE PST2-3-steal K. potatoes
 'I knew with surprise that Kibet stole the potatoes.'

Given this split in *le*'s behavior, in Bossi (2023), I propose that *le* represents two distinct syntactic atoms: in some cases, it is the lexical verb meaning 'say', but in others, it is a true complementizer devoid of speech semantics and only linked to the verb 'say' diachronically. As an attitude predicate (215) and in complementation structures with speech semantics where *le* can bear extensive verbal inflection (221) and take PP modifiers (224), it is the lexical verb meaning 'say', following Driemel & Kouneli (2022a). On the other hand, when *le* occurs in complementation structures without speech semantics where extensive verbal inflection (222) - (223) and PP modifiers (225) are impossible, it is a semantically-bleached complementizer. On this analysis, *le* is restricted to only prefixal agreement when it is a complementizer. Given this restriction, existing accounts of upward-oriented complementizer agreement can extend to capture the Kipsigis

pattern without requiring Agree to probe upwards (e.g. Diercks 2013; Carstens 2016). For more details on this analysis, see Bossi (2023).

2.5.6.4 Embedded questions

Both polar questions and constituent questions can be embedded in Kipsigis. Embedded polar questions are introduced by the complementizer *koto* (227) or *ngot* (228). These two elements seem to be in free variation with no clear semantic difference between embedded questions with *koto* and those with *ngot*.

- (227) Koomache Kiproono konai koto kooiwee sʊgʊl amut.
 Koo-∅-mach-e Kiproono ko-nai **koto** koo-i-wa-i sʊgʊl amut.
 PST2-3-want-IPFV K. 3.SBJV-know if PST2-2SG-go.SG-LP school yesterday
 ‘Kiproono wanted to know if you (sg) went to school yesterday.’
- (228) Maa-ngen **ngot** koo-i-pel-iis Kipchirchir εen urerieet.
 NEG.1SG-know if PST2-3-win-AP K. in game
 ‘I don’t know if Kipchirchir won the game.’

Unlike with matrix polar questions (§2.5.4.1), there is no clause-final epenthetic vowel or super high tone in embedded polar questions (229).

- (229) *Maa-ngen ngot koo-i-pel-iis Kipchirchir εen urerieet-**í**.
 NEG.1SG-know if PST2-3-win-AP K. in game-Q
 Intended: ‘I don’t know if Kipchirchir won the game.’

Embedded constituent questions use whichever embedding strategy the matrix verb requires and display standard constituent question syntax (§2.5.4.2). For instance, the matrix verb *ngen* ‘know’ embeds an indicative clause with the element *kole* when the embedded clause is a declarative (206). This same pattern is observed when the embedded clause is a question (230); a form of *kole* is obligatory here, and the embedded question is in the indicative mood and shows the typical word order found in Kipsigis constituent questions.

- (230) Aa-ngen *(**aa-le**) ∅-ru-e **ng’oo**.
 1SG-know 1SG.SBJV-LE 3-sleep-IPFV who
 ‘I know who’s sleeping.’

(231) - (234) offer additional examples of embedded constituent questions in Kipsigis. Note that word orders in which the *wh*-word is immediately postverbal, as well as those where it is clause-initial with a form of the relativizer are both possible, even in embedded questions. Following the discussion in §2.5.5.1, I assume that questions like the one in (231) contain an embedded pseudocleft.

- (231) Aa-ngen aa-le **ng’oo** ne koo-∅-keer Kiproono.
 1SG-know 1SG.SBJV-LE who REL.SG PST2-3-see K.
 ‘I know who Kiproono saw.’

- (232) Ii-ngen ii-le Ø-am-e **nee** Kipchirchir.
 2SG-know 2SG.SBJV-LE 3-eat-IPFV what K.
 ‘You (sg) know what Kipchirchir is eating.’
- (233) Aa-ngen aa-le koo-Ø-ip-ta **ano** Kipchirchir urerieet.
 1SG-know 1SG.SBJV-LE PST2-3-win-IT where K. game
 ‘I know how Kipchirchir won the game.’
- (234) Oo-ngen oo-le koo-Ø-chap **au** Kipchirchir kεεgɪt.
 2PL-know 2PL.SBJV-LE PST2-3-make when K. cake
 ‘You (pl) know when Kipchirchir made the cake.’

2.5.7 Modal verbs

Kipsigis has a large inventory of modal verbs, which are summarized in Table 2.16 and classified in terms of modal flavor and force. Modal flavor describes the basis that the speaker is making their claim on—for instance, what the speaker knows about the world, or what the law or other authority says. Modal force describes the strength of the modal claim: whether the world *must* be a particular way or simply *could* be a particular way. Some Kipsigis modal verbs are borrowed from Swahili, which is notated in Table 2.16 using italics.

Modal flavor	Modal force	
	Necessity	Possibility
Epistemic	<i>lazima</i> , nyaalu, <i>raisi</i>	<i>raisi</i> , toot
Circumstantial	yaache	toot
Deontic	<i>lazima</i> , nyaalu	toot

Table 2.16: Modal verbs in Kipsigis

Syntactically, Kipsigis modal verbs differ from other verbs because they do not inflect for tense (235), and they do not display subject agreement (236).

- (235) ***Koo-raisi** ko kartit amut.
 PST2-MOD 3.SBJV cold yesterday
 Intended: ‘It might have been cold yesterday.’
- (236) Eɛn Kenya ko mwa-e ng’otuutiik kole **lazima** ɪ-laach εlmɛtɪt yan
 in K. TOP say-IPFV laws C MOD 2SG.SBJV-wear helmet when
 ii-keet-e pikipiki.
 2SG.SBJV-ride-IPFV motorcycle
 ‘In Kenya, the law states that you (sg) must wear a helmet when you (sg) ride a motorcycle.’

Yet like other verbs, they are compatible with negation (237) and display aspect marking in some cases (238). However, this aspect marking is invariant and does not appear on every modal verb;

for instance, *yaache* and *nyaalu* plausibly contain the imperfective aspect suffixes *-e* and *-u*, respectively, while *toot* does not contain any aspectual morphology.

- (237) a. **Ma-raisi ko** kartit ra.
 NEG-MOD 3.SBJV cold today
 ‘It might not be cold today.’
 b. **Ma-lazima ko** kartit ra.
 NEG-MOD 3.SBJV cold today
 ‘It’s not necessarily cold today.’
- (238) **Yaach-e a-riony.**
 MOD-IPFV 1SG.SBJV-sneeze
 ‘I have to sneeze.’

Finally, Kipsigis modal verbs embed Type II subjunctive clauses, as evidenced by the subjunctive agreement prefixes in the clauses embedded under the modal verbs in (236) - (238).

The rest of this section provides examples of each of the modal verbs in Table 2.16. First, the sentences in (239) - (240) show the use of *lazima*, *nyalu*, and *raisi* to indicate epistemic necessity. Here, the contexts establish that, according to what the speaker knows, it must be the case that the world is a particular way: in (239), that the ball *must* be in Box 3, and in (240), that Kipchirchir *must* be at Chepchirchir’s house.

- (239) The math teacher tells you that the ball is in one of three boxes. You look into Box 1 and Box 2 and discover that it’s not in either of those boxes. [epistemic necessity]
- a. **Lazima ko-miit-een** mbireet pogisit nɔmba somok.
 MOD 3.SBJV-COP-in ball box number three
 ‘The ball must be in Box 3.’
- b. **Nyaal-u ko-miit-een** mbiret pogisit nɔmba somok.
 MOD-IPFV 3.SBJV-COP-in ball box number three
 ‘The ball must be in Box 3.’
- c. **Raisi ko-miit-een** mbiret pogisit nɔmba somok.
 MOD 3.SBJV-COP-in ball box number three
 ‘The ball must be in Box 3.’
- (240) Kipchirchir has coffee at Chepchirchir’s house everyday. Even if he’s sick, he doesn’t miss a day! It’s not obligatory for Kipchirchir; he just goes for coffee there all the time. It’s coffee time now, so: [epistemic necessity]
- a. **Lazima ko-mii** Kipchirchir kaat-aap Chepchirchir.
 MOD 3.SBJV-COP K. house-of C.
 ‘Kipchirchir must be at Chepchirchir’s house.’
- b. **Nyaal-u ko-mii** Kipchirchir kaat-aap Chepchirchir.
 MOD-IPFV 3.SBJV-COP K. house-of C.
 ‘Kipchirchir must be at Chepchirchir’s house.’

- c. **Raisi** ko-mii Kipchirchir kaat-aap Chepchirchir.
 MOD 3.SBJV-COP K. house-of C.
 ‘Kipchirchir must be at Chepchirchir’s house.’

While *lazıma*, *nyaalu*, and *raisi* are interchangeable in most epistemic necessity contexts, there are cases where *nyaalu* is infelicitous, even though *lazıma* and *raisi* are acceptable. In (241), for instance, the doctor concludes that their patient’s headache must be due to stress, since no tests are giving positive results. However, speakers find *nyaalu* inappropriate here because the doctor’s conclusion is not sufficiently motivated; as they put it, *nyaalu* “comes with the expectation of surety” and the context in (241) does not establish enough certainty.

- (241) You have a headache that won’t go away, so you go to the doctor. All the tests show negative. So the doctor concludes that: [epistemic necessity]
- a. **Lazıma** ko-yab-u kabuatuutik.
 MOD 3.SBJV-COME-VEN stress
 ‘It must be from stress.’
- b. #**Nyaal-u** ko-yab-u kabuatuutik.
 MOD-IPFV 3.SBJV-COME-VEN stress
 ‘It must be from stress.’
- c. **Raisi** ko-yab-u kabuatuutik.
 MOD 3.SBJV-COME-VEN stress
 ‘It must be from stress.’

Epistemic possibility—rather than necessity—is found in (242). Here, according to what the speaker knows, it is possible that Cherono will go to Nairobi, but it is not certain that this will happen. In this type of context, both *toot* and *raisi* are acceptable.

- (242) Cherono’s parents told her that she isn’t allowed to go to see her friend in Nairobi because it’s too far away. You heard that Cherono is leaving Bomet next week, but you don’t know where she’s going. Cherono is a daring type of person who usually does things that she isn’t permitted to do. You think: [epistemic possibility]
- a. **Toot** ko-wa Cheerono Nairobi.
 MOD 3.SBJV-go.SG C. N.
 ‘Cherono might go to Nairobi.’
- b. **Raisi** ko-wa Cheerono Nairobi.
 MOD 3.SBJV-go.SG C. N.
 ‘Cherono might go to Nairobi.’

Additionally, an inflected form of the verb *le* ‘say’ can take on a modal function in epistemic possibility contexts. While the modal verb *raisi* is acceptable in (243), the first person present imperfective form of the verb *le* ‘say’ is also felicitous. As with standard cases of embedding under *le*, *le* agrees with the subject, inflects for aspect, and takes an indicative complement clause.

- (243) You're playing hide-and-seek with Kipkoech. He's allowed to hide in the kitchen, living room, or bedroom. You haven't started looking for him yet, but you know that he might be in any one of those three rooms. [epistemic possibility]
- a. **Raisi** ko-miit-een Kɪpkɔɛch chigeet.
 MOD 3.SBJV-COP-in K. kitchen
 'Kipkoech might be in the kitchen.'
- b. **A-leen** miit-een Kɪpkɔɛch chigeet.
 1SG-say.IPFV COP-in K. kitchen
 'Kipkoech might be in the kitchen.'
 (Lit: 'I'm saying that Kipkoech is in the kitchen.')

It is also worth highlighting that the modal verb *raisi* displays what is called “force flexibility” (Matthewson et al. 2007; Rullmann et al. 2008; Davis et al. 2009): it is compatible with epistemic necessity and possibility claims, as seen in (239) - (243). A further example of this modal force flexibility is provided in (244). Here, the modal statement with *raisi* is compatible with the speaker knowing that it *must* have been Ed who arrived (epistemic necessity) or that it *might* have been Ed who arrived (epistemic possibility).

- (244) **Raisi** Ed ne ka-∅-it.
 MOD E. REL.SG PST1-3-arrive
 'It might/must be Ed who just arrived.'
- a. ✓ You have one housemate named Ed. He's the only other person with keys to your apartment. You hear someone open the door using a set of keys, so you conclude that it must be your housemate Ed getting home. [epistemic necessity]
- b. ✓ You have two housemates: Ed and Kyalo. They're the only other people with keys to your apartment. You hear someone open the door using a set of keys, so you conclude that it might be Ed getting home. [epistemic possibility]

This type of modal force flexibility is thought to be rare cross-linguistically and is most widely documented in indigenous languages of the Americas (Matthewson et al. 2007; Rullmann et al. 2008; Davis et al. 2009 for St'át'imcets, Peterson 2010; Matthewson 2016 for Gitksan, Bochnak 2015 for Washo, Deal 2011 for Nez Perce, Menzies 2013 for Nsyilxcen, Jeretič 2021 for Ecuadorian Siona). However, recent work by Newkirk (2022) describes and analyzes force flexible epistemic modals in the Bantu language Kinande. Alongside the Kipsigis facts here, this suggests that modal force flexibility is perhaps more widespread than assumed.

Circumstantial modals are used when physical circumstances make it such that the world must or might be a particular way. (245) illustrates circumstantial necessity: because the speaker has been stuck on the bus, they are now in the position of needing to use the bathroom. Circumstantial necessity is typically conveyed in Kipsigis using the modal verb *yaache* (245a). However, it is also possible to use a first person present imperfective form of the verb *mach* 'want' in these contexts (245b). As with standard cases of embedding under *mach*, *mach* agrees with the subject, inflects for aspect, and takes a subjunctive complement clause.

(245) You're on the bus to Nairobi. You haven't had a chance to go to the toilet for 6 hours, and your bladder is full. You say: [circumstantial necessity]

- a. **Yaach-e** a-yai peek nia.
 MOD-IPFV 1SG.SBJV-make water so.much
 'I have to pee so badly.'
- b. **A-mach-e** aa-yai peek nia.
 1SG-want-IPFV 1SG.SBJV-make water so.much
 'I have to pee so badly.'
 (Lit: 'I want to pee so badly.')

(246) illustrates circumstantial possibility. Here, the physical environment makes it possible for sunflowers to grow in the area. As with epistemic possibility contexts, these situations use the modal verb *toot*.

(246) Claire visited Bomet. She noticed that the climate and many of the plants are similar to places she visited in Tanzania; the temperature is the same, the rainfall is the same, the types of rocks and the soil are the same. But when she looked around, she didn't find any sunflowers anywhere, which she saw all over Tanzania. But because the temperature, rainfall, and soil are the same, she thinks that: [circumstantial possibility]

Toot ko-rut maʷwɛɛk ɛɛn yu.
 MOD 3.SBJV-grow flowers in here
 'Sunflowers could grow here.'

Finally, deontic modals are used to express what is required or allowed according to the law or some other authority. The modal verbs *nyaalu* and *lazima* are used to express deontic necessity, as seen in (247). Here, the context establishes that, according to the rules of the hospital, visitors must leave by a particular time.

(247) You're going to visit your friend in the hospital. When you enter the hospital, you stop at the information desk to inquire what room your friend is in. But the woman at the information desk tells you that you can't visit your friend now because it's already 8pm! She says, "I'm sorry, the hospital regulations say that..": [deontic necessity]

- a. **Nyaal-u** ko-ba tayeek ye ka-∅-it sart taman ak aeeng'.
 MOD-IPFV 3.SBJV-go.PL visitors when PST1-3-arrive time ten and two
 'Visitors must leave by 6pm.'
- b. **Lazima** ko-ba tayeek ye ka-∅-it sart taman ak oeeng'.
 MOD 3.SBJV-go.PL visitors when PST1-3-arrive time ten and two
 'Visitors must leave by 6pm.'

Deontic possibility, on the other hand, is seen in (248) - (249) with the modal verb *toot*. In these contexts, the rules of the hotel or the authority of the speaker on who stays with them makes it possible for the world to be a particular way. In this way, *toot* is an underspecified possibility modal that is compatible with the three modal flavors in Table 2.16.

- (248) We're at a hotel that only allows guests over 18 years old. Kiplangat is 24 years old. He doesn't have to stay at the hotel, but he can if he wants to. [deontic possibility]
Toot ko-pur Kiplang'at.
 MOD 3.SBJV-stay K.
 'Kiplangat may stay.'
- (249) You're making plans for tomorrow night to get together with your friend and potentially stay at their house. Your friend says: [deontic possibility]
Toot ii-pur lang'atʊʊnɪ anan (**toot**) ii-wa-i kaa.
 MOD 2SG.SBJV-stay overnight or MOD 2SG.SBJV-go.SG-LP home
 'You (sg) can stay overnight or you (sg) can go home.'

2.6 Summary

In this chapter, I have outlined several key properties of Kipsigis grammar with a focus on the morphosyntax and semantics of the language. While there are many aspects of Kipsigis grammar that warrant further exploration and description, it is my hope that this grammar sketch lays the foundation for future research in these areas. In addition, the information provided in this chapter offers important background and context for the remaining chapters of this dissertation, which address epistemic modality in two different syntactic domains in Kipsigis—namely, in nominals and in biased attitude reports.

Chapter 3

Higher order ignorance in Kipsigis epistemic indefinites

3.1 Introduction

Epistemic indefinites are indefinite pronouns or determiners that convey speaker ignorance with respect to the witness to the indefinite. For instance, the Spanish sentence in (250a) expresses an existential claim: $\exists x[x$ is a student in the linguistics department & María married $x]$. Use of *algún* additionally indicates that the speaker does not know which linguistics student María married—in contrast to the basic Spanish indefinite *un* (250b), which triggers no such inference.

- (250) *Spanish* (Alonso-Ovalle & Menéndez-Benito 2010:ex. 1, 3)
- a. María se casó con **algún** estudiante del departamento de lingüística.
M. SE married with ALGUN student of.the department of linguistics
'María married a linguistics student.'
 - b. María se casó con **un** estudiante del departamento de lingüística.
M. SE married with UN student of.the department of linguistics
'María married a linguistics student.'

A wide range of epistemic indefinites across languages are discussed in the literature, including: German *irgendein* (Kratzer & Shimoyama 2002), Italian *un qualsiasi* (Aloni & van Rooij 2004; Chierchia 2006) and *un qualche* (Zamparelli 2007), French *quelque* and *un quelconque* (Jayez & Tovená 2006, 2007), the Russian *to* series (Geist & Onea 2007; Kagan 2011), Romanian *vreun* (Farkas 2002; Fălăuş 2014), the Japanese *-ka* series (Alonso-Ovalle & Shimoyama 2014), the Czech *-si* series (Šimík 2015), the Tiwa *-pha* and *-khi* series (Dawson 2018, 2020), and Akan *bi* (Owusu 2019). While these epistemic indefinites all convey some form of ignorance, their empirical profiles differ on a number of dimensions and a range of analyses have been put forth to capture these unique behaviors.

One such parameter of cross-linguistic variation concerns the type of ignorance effects triggered by these different epistemic indefinites. First order ignorance, like that seen in (250a) with

Spanish *algún*, conveys that the speaker does not know which *individual* witnesses the indefinite. Higher order ignorance, on the other hand, conveys that the speaker is ignorant about some relevant *property* of the witness to the indefinite, even if they know exactly which individual it is. An example of this type of higher order ignorance can be seen with the *-khi* indefinite in (251) from Tiwa (Tibeto-Burman; India).

- (251) *Tiwa* (Dawson 2018:ex. 36)
 Ang **shar-khí** India-ne PM-go lak mán-a lí-do.
 1SG who-KHI India-GEN PM-ACC meet-INF go-IPFV
 ‘I’m going to meet some Indian Prime Minister.’

In contexts where (251) is appropriate, there is no question on the speaker’s part about who the Indian Prime Minister is; they know that he is Narendra Modi. However, to felicitously utter (251), the speaker must be ignorant about some contextually relevant property of this individual, ranging from his hair color to what he is wearing to something else entirely.

Dawson (2020) suggests that these different types of ignorance correlate with other empirical properties of the indefinites, like their scope profiles and, by extension, the semantic analyses that they warrant. In particular, she correlates first order ignorance with scopal flexibility and domain-widening semantics, and higher order ignorance with indefinites that display exceptional wide scope and so warrant a choice functional analysis. Here, however, I draw on original field data to show that epistemic indefinites in Kipsigis can convey higher order ignorance but do not show exceptional wide scope. I also show that Kipsigis epistemic indefinites have properties that pose challenges for both domain widening and choice functional analyses of such indefinites. This constellation of facts calls into question the proposed link between ignorance type and scope from Dawson (2020).

Instead, I offer a new account for Kipsigis, according to which the epistemic indefinite introduces basic existential quantification and only triggers ignorance effects due to competition with another type of Kipsigis nominal, which is a choice functional indefinite that conveys some amount of speaker knowledge. In doing so, I propose a re-framing of the correlation between ignorance type and scope in Dawson (2020): namely that, if an epistemic indefinite displays exceptional wide scope, then pragmatically derived ignorance must be higher order. Additionally, I suggest that epistemic effects—both ignorance and knowledge—might offer an additional tool for diagnosing whether there is existential closure of a choice function variable, thus adding to a discussion of this issue in Kratzer (1998) and Matthewson (1999).

The remainder of the chapter is structured as follows. §3.2 overviews the Kipsigis nominal inventory. Then, in §3.3, I characterize the ignorance effects seen with Kipsigis epistemic indefinites. In §3.4, I provide evidence that these ignorance effects are pragmatically derived. §3.5 describes the semantic properties of Kipsigis epistemic indefinites and compares the predictions of existing accounts to the Kipsigis pattern. §3.6 explores the interpretation of Kipsigis nouns with a so-called “secondary suffix”, which I argue to be the relevant competitor to epistemic indefinites in the language. I propose that these forms introduce a free choice function variable that is contextually resolved. Against this backdrop, I offer a new analysis for Kipsigis epistemic indefinites in §3.7, according to which they are basic existential quantifiers and ignorance effects

arise pragmatically due to competition with these choice functional Kipsigis nouns. In §3.8, I discuss the cross-linguistic typology of epistemic indefinites and highlight some key parameters of variation in their behavior and analysis. §3.9 concludes.

3.2 The Kipsigis nominal inventory

To understand how epistemic indefinites fit into the Kipsigis nominal inventory, it is important to recall some key properties of Kipsigis and its nominal inventory from Chapter 2 §2.3.1 - 2.3.3. Kipsigis has been described as a bare noun language (Kouneli 2019) because nouns can have both indefinite (252) and definite interpretations (253) without the need for articles.¹ In the following mini-discourse, the nouns *kaaneetiindet* ‘teacher’, *laakwæet* ‘child’, and *ng’ookta* ‘dog’ in (252) are all indefinites that introduce new discourse referents. In (253), these same nouns act as anaphoric definites, referring back to the referents introduced in the previous sentence.

- (252) Ø-Sooman-chiin **kaaneetiindet laakwæet** kɪtabʊʊt agɔbɔ **ng’ookta**.
 3-read-APPL.IPFV teacher child book about dog
 ‘A teacher is reading a book about a dog to a child.’
- (253) Ø-Tep-eeen **laakwæet kaaneetiindet** kole tyan oo **ng’ookta**.
 3-ask-INST child teacher C how.much big dog
 ‘The child asks the teacher how big the dog is.’

Despite lacking overt articles, Kipsigis nouns are morphologically complex and fall into three number-based noun classes, as described in Chapter 2 §2.3.1 from Kouneli (2019, 2021). Kipsigis nouns contain a root followed by a number suffix (depending on the noun’s class and plurality), then a thematic vowel, then a secondary suffix (254). Nouns do not necessarily contain all of these morphemes—in particular, not all nouns contain a thematic vowel or a number morpheme—but when they co-occur, they surface in the order listed above and seen in (254).

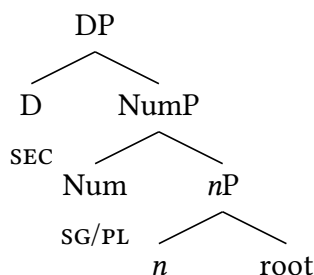
- (254) sigis-yaan-ta-it → sigisyáat
 sock-SG-TH-SEC
 ‘sock’
 (Kouneli 2021:ex. 5c)

In light of this morphological complexity, Kouneli (2019, 2021) proposes that Kipsigis nouns have the structure in (255). Number morphology like singulative *-yaan* in (254) is realized in Num, while the secondary suffix is realized in D.

¹This is a slight simplification of the empirical pattern; noun interpretation in Kipsigis is restricted in a way reminiscent of the pattern seen in other bare noun languages (e.g. Dayal 2004; Deal & Nee 2018), though it is not exactly the same. In particular, existential interpretations are more widely available in Kipsigis than in other bare noun languages. See §3.6 and Chapter 2 §2.3.2 for more discussion.

(255) Kipsigis noun structure

(Kouneli 2021:24)



In this way, while nouns like those in (252) - (254) have been described as bare nouns, I refer to them as “nouns with a secondary suffix” here, since they are not syntactically bare in the canonical sense, as first discussed in Chapter 2 §2.3.1.

In addition to these noun forms, recall from Chapter 2 §2.3.3 that the suffix *-yan* can replace a noun’s secondary suffix to form a dedicated indefinite, as in the b. examples in (256) - (257). While speakers translate nouns with a secondary suffix as ‘the N’ or ‘a N’ depending on the context, forms with *-yan* are consistently translated as ‘some N’, which hints at their function as indefinites that signal some kind of speaker ignorance.

- | | | | | |
|-------|---|---|-----------------------------------|---|
| (256) | a. kar- I -it → kàrít
car-TH-SEC
‘the/a car’ | → | kàrít
(Kouneli 2019:ex. 5c) | b. kar- I - yan
car-TH-YAN
‘some car’ |
| (257) | a. laak-wa-it → làakwéet
child-TH-SEC
‘the/a child’ | → | làakwéet
(Kouneli 2019:ex. 5b) | b. laak-wa- yan
child-TH-YAN
‘some child’ |

Before turning to the interpretation of *-yan* forms, it is important to recall that not all nouns can take the *-yan* suffix, as described in Chapter 2 §2.3.3. First, only singular nouns have *-yan* forms.² All nouns with *-yan* are interpreted as singular, and it is impossible for *-yan* to replace the secondary suffix on a plural noun. This restriction can be seen in (258), where *-yan* cannot surface alongside the plural morpheme *-oy*; instead, it must surface after the thematic vowel in the singular form, as seen in (257b).

- | | | | | |
|-------|---|---|-----------------------------------|--|
| (258) | a. laak-oy-ik → lààgóók
child-PL-SEC
‘children’ | → | lààgóók
(Kouneli 2019:ex. 12b) | b. *laak-oy- yan
child-PL-YAN
Intended: ‘some children’ |
|-------|---|---|-----------------------------------|--|

Second, the set of nouns that *-yan* attaches to is lexically restricted (see (53) in Chapter 2 §2.3.3 for a list of all *-yan* forms that I have found in Kipsigis so far). As a result, the diagnostics

²See §3.7.1 for some discussion of the source of this restriction. In particular, I suggest that indefinite *-yan* arose historically from singulative *-yaan*, which aligns nicely with the fact that *-yan* attaches only to singular nouns.

described in this chapter cannot be applied to every nominal in the language; the examples here make use of a relatively small set of nouns, but the reported patterns are found with other *-yan* forms. Importantly, as shown in Chapter 2 §2.3.3, the addition of *-yan* applies to both native Kipsigis words as well as borrowings, which speaks to its synchronic productivity. Likewise, *-yan* forms crosscut Kipsigis noun classes, applying to nouns from all number-based noun classes, as seen in Chapter 2 §2.3.3. In this way, *-yan*'s distribution suggests that it is synchronically productive and acts independently of the Kipsigis noun class system.

3.3 Characterizing the ignorance effects triggered by *-yan*

With this background in mind, this section justifies the claim that *-yan* forms are epistemic indefinites that signal speaker ignorance—both first order and higher order ignorance. First, it is infelicitous for a speaker to follow a *-yan* form with explicit identification of the witness to the indefinite (259); in using the *-yan* form, the speaker conveys ignorance.

- (259) Kibet and Chepkoech are playing hide-and-seeK—Kibet is the seeker and Chepkoech could be hiding in any room in the house. Kibet says:
 Ø-Uny-e-kεε Cheεpkœεch εεn **room-ɪ-yan**. #Miit-een chigeet.
 3-hide-IPFV-REFL C. in room-TH-YAN COP-in kitchen
 ‘Chepkoech is hiding in some room. #She’s in the kitchen.’
 (contexted adapted from Alonso-Ovalle & Menéndez-Benito 2010)

In this respect, Kipsigis *-yan* forms pattern with other epistemic indefinites across languages, like Spanish *algún* (260) and Tiwa *-khí* (261).

- (260) *Spanish* (Alonso-Ovalle & Menéndez-Benito 2010:ex. 2)
 María se casó con **algún** estudiante del departamento de lingüística: #en concreto
 M. SE married with ALGUN student of.the department of linguistics namely
 con Pedro.
 with P.
 ‘María married a linguistics student, #namely Pedro.’
- (261) *Tiwa* (Dawson 2018:ex. 2)
 Shar-**khí** phi-dom. #Pe-do Mukton.
 who-KHI come-PST 3SG-TOP M.
 ‘Someone came. #Namely, Mukton.’

Second, it is infelicitous to ask for identification of the witness to the indefinite in subsequent discourse after use of a *-yan* form. It is unnatural for a speaker to ask (262b) as a truly information-seeking question; if they do ask this question, speakers report that they should expect the answer *maangen* ‘I don’t know’.

- (262) Kibet and Chepkoech are playing hide-and-seeK—Kibet is the seeker and Chepkoech could be hiding in any room in the house. Kibet says a. and his interlocutor asks b.

- a. Ø-Uny-e-kεε Chepkœech εen rœom-i-yan.
 3-hide-IPFV-REFL C. in room-TH-YAN
 ‘Chepkœech is hiding in some room.’
- b. # Aion?
 which
 ‘Which one?’
 (context adapted from Alonso-Ovalle & Menéndez-Benito 2010)

Here *-yan* forms contrast with Kipsigis nouns with a secondary suffix; in these contexts, interlocutors can ask about the witness to the indefinite in subsequent discourse without any infelicity (263b).

- (263) Kibet and Chepkœech are playing hide-and-seek—Kibeet is the seeker and Chepkœech could be hiding in any room in the house. Kibet says a. and his interlocutor asks b.
- a. Ø-Uny-e-kεε Chepkœech εen rœom-i-it.
 3-hide-IPFV-REFL C. in room-TH-SEC
 ‘Chepkœech is hiding in a room.’
- b. Aion?
 which
 ‘Which one?’

As before, this pattern parallels that found with other epistemic indefinites; (264) - (265) provide examples of the question-and-answer diagnostic for Spanish *algún* and *un*, respectively.

- (264) *Spanish* (Alonso-Ovalle & Menéndez-Benito 2010: ex. 8)
- a. Juan tiene que estar en **alguna** habitación de la casa.
 J. has to be in ALGUN room of the house
 ‘Juan must be in a room of the house.’
- b. #¿En cuál?
 in which
 ‘In which one?’
- (265) *Spanish* (Alonso-Ovalle & Menéndez-Benito 2010: ex. 10)
- a. Juan tiene que estar en **una** habitación de la casa.
 J. has to be in UNA room of the house
 ‘Juan must be in a room of the house.’
- b. ¿En cuál?
 in which
 ‘In which one?’

Last, *-yan* forms are unnatural when it is assumed that the speaker should not be ignorant about the witness to the indefinite or their salient properties. This effect is particularly clear

in sentences with past tense verbs like *tun* ‘marry’ and first person subjects, given the assumption that the speaker should be quite familiar with their spouse. In these cases, *-yan* forms are infelicitous, since they convey ignorance (266).

- (266) # Kii-a-tun **kaaneet-iin-da-yan.**
 PST3-1SG-marry teacher-SG-TH-YAN
 ‘I married some teacher.’

However, changing the person value of the subject (267) or swapping the *-yan* form for a noun with a secondary suffix (268) renders the sentence felicitous. This is because it is perfectly reasonable for the speaker to be ignorant about someone else’s spouse, as in (267), or there simply are no more ignorance effects because there is no *-yan* form, as in (268).

- (267) Kii-∅-tun Kiproono **kaaneet-iin-da-yan.**
 PST3-3-marry K. teacher-SG-TH-YAN
 ‘Kiproono married some teacher.’
- (268) Kii-a-tun **kaaneet-iin-da-it** (kaaneetiindet).
 PST3-1SG-marry teacher-SG-TH-SEC
 ‘I married a teacher.’

In this way, Kipsigis *-yan* forms can be diagnosed as epistemic indefinites that signal speaker ignorance. Notably, though, these forms can convey both first order and higher order ignorance, meaning that a speaker can use a *-yan* form when they are ignorant about either: 1) the individual who witnesses the existential claim (269a), or 2) some salient property of the witness to the existential claim (269b). Use of the *-yan* form is only ruled out when the speaker can identify the witness and knows its salient properties (269c).

- (269) Kibet and Chepkoech are playing hide-and-seek—Kibet is the seeker and Chepkoech is hiding. Kibet says to his babysitter:
 ∅-Uny-e-kεε Cheεpkœch εεn **rɔsm-i-yan.**
 3-hide-IPFV-REFL C. in room-TH-YAN
 ‘Chepkoech is hiding in some room.’
 (context adapted from Alonso-Ovalle & Menéndez-Benito 2010)
- a. ✓ Kibet knows that Chepkoech is in the house, but he doesn’t know which room she’s in, so he can’t find her.
 - b. ✓ Kibet knows that Chepkoech is in the living room, but he doesn’t know where in the house the living room is, so he can’t find her.
 - c. # Kibet cheated, so he knows that Chepkoech is in the living room. He also knows where in the house the living room is, so he can find her.

Another example illustrating this higher order ignorance is found in (270). In this context, the speaker knows who witnesses the indefinite *choorwayan* ‘friend’ but is ignorant about what they are wearing, which is a key property that would enable them to find their friend in the crowded restaurant.

- (270) I'm looking for my good friend in a crowded restaurant. I know exactly which friend I'm looking for, but I don't know what they're wearing today, so it's hard to find them in the restaurant. I say:
- a. A-cheeng'-e **choor-wa-yan...**
 1SG-search-IPFV friend-TH-YAN
 'I'm looking for a friend...'
- b. lakini toma-nyoor-u ngamøn maq-ngen kiy ne i-laach-e.
 but not.yet.1SG-find-IPFV because NEG.1SG-know thing REL.SG 3-wear-IPFV
 'but I can't find them because I don't know what they're wearing.'
 (context adapted from Dawson 2018)

In conveying both first order and higher order ignorance, Kipsigis *-yan* forms pattern with epistemic indefinites like Tiwa *-khi* indefinites, but they differ from other epistemic indefinites like those with Spanish *algún*. In particular, Tiwa *-khi* indefinites are compatible with higher order ignorance; in (271), the speaker knows that the Indian Prime Minister is Narendra Modi but is ignorant about at least one other salient property of this individual. On the other hand, indefinites with Spanish *algún* cannot generally convey higher order ignorance; in (272), the indefinite with *algún* is infelicitous, even though the speaker is ignorant about at least one property of the unknown professor dancing on the table.

- (271) *Tiwa* (Dawson 2018:ex. 36)
 Ang **shar-khi** India-ne PM-go lak mán-a lí-do.
 1SG who-KHI India-GEN PM-ACC meet-INF go-IPFV
 'I'm going to meet some Indian Prime Minister.'
- (272) *Spanish* (Alonso-Ovalle & Menéndez-Benito 2017:ex. 4b)
 L. and P. are visiting the Math department. They don't know anything about the people working there, and they haven't seen any of them before. They suddenly see an individual, who can be inferred to be a professor, frantically dancing on his desk.
 # ¡Mira! ¡**Algún** profesor está bailando encima de la mesa!
 look ALGUN professor is dancing on of the table
 'Look! Some professor is dancing on the table!'

Taken together, the data presented in this section show that Kipsigis *-yan* forms are epistemic indefinites that can convey both first order and higher order ignorance. The following section addresses whether these ignorance effects are lexicalized in the *-yan* suffix itself or whether they arise via pragmatic reasoning and competition with other Kipsigis nominals.

3.4 Ignorance effects are pragmatically derived

Existing analyses of epistemic indefinites fall into two main classes: those that lexicalize ignorance in the meaning of the epistemic indefinite (e.g. Giannakidou & Quer 2011; Aloni & Port 2015) and those that derive it pragmatically via competition between the epistemic indefinite

and another type of nominal in the language (e.g. Kratzer & Shimoyama 2002; Alonso-Ovalle & Menéndez-Benito 2010, 2013, 2017; Dawson 2018, 2020). In this section, I show that the ignorance effects associated with *-yan* forms are pragmatically derived, since they show the hallmarks of conversational implicature; in particular, they are reinforceable, cancellable with sufficient contextual support, and disappear in downward-entailing contexts.

First, speakers often follow a sentence containing a *-yan* form with an explicit statement of ignorance, which indicates that the ignorance component is not part of the asserted content of the *-yan* form. If ignorance were lexically encoded in the *-yan* suffix, the type of reinforcement seen in (273) - (274) would be redundant, since it would simply repeat part of the asserted content of the *-yan* form. By contrast, speakers note that, while the type of reinforcement in (273) - (274) is not necessary to convey ignorance, it is not at all redundant and, therefore, does not repeating part of the asserted content of *-yan*.

(273) Kɔɔ-Ø-al Kibeet kar-**I-yan** ngandan maɑ-nɡen kole aion.
 PST-3-buy K. car-TH-YAN but NEG.1SG-know C which
 ‘Kibet bought some car, but I don’t know which one.’

(274) Miit-een Kibeet ak Cheɛpkɔɛch **rɔɔm-I-yan** ngandan maɑ-nɡen kole aion.
 COP-in K. and C. room-TH-YAN but NEG.1SG-know C which
 ‘Kibet and Chepkoech are in some room, but I don’t know which one.’

Second, the ignorance effects are cancellable, as long as there is sufficient contextual support for this cancellation. As seen previously in (259), a speaker cannot follow use of a *-yan* form with explicit identification of the witness without any established reason to do so. Instead, there must be motivation for cancellation established in the discourse, in which case it is possible. In (275), for instance, the speaker is explicitly withholding information to avoid helping Kibet cheat in the game; in this context, cancellation of the ignorance effects is possible.

(275) Kibet and Chepkoech are playing hide-and-seek. Kibet is trying to cheat and get information from me, but I won’t help him. I say:

a. Ø-Uny-e-kɛɛ Cheɛpkɔɛch ɛɛn **rɔɔm-I-yan**.
 3-hide-IPFV-REFL C. in room-TH-YAN
 ‘Chepkoech is hiding in some room.’

b. Aɑ-nɡen aale rɔɔm-I-it aion ngandan maɑ-mwa-uun.
 1SG-know C room-TH-SEC which but NEG.1SG-say-2SG.IO
 ‘I know which one, but I won’t tell you.’

This behavior suggests a pragmatic derivation of ignorance effects in Kipsigis because, if ignorance were lexicalized, cancellation would lead to a contradiction, rather than simply constituting a marked discourse move that requires contextual justification.

In this way, cancellation of the ignorance effects triggered by *-yan* is possible, although it requires contextual support. Aloni & Port (2015) cite this sort of requirement as evidence against a pragmatic derivation of these kinds of ignorance effects, since without sufficient context, it can seem as though these ignorance effects are not cancellable at all. Yet in general, cancellation of

an implicature constitutes a marked discourse move that is only appropriate in certain discourse configurations (see e.g. Mayol & Castroviejo 2013; Rett 2020). As a result, the fact that cancellation of *-yan*'s ignorance effects must be sufficiently motivated is not particularly unusual. In fact, epistemic indefinites with pragmatically derived ignorance effects in other languages show a similar type of requirement. For instance, Tiwa *-khi* indefinites cannot be cancelled by simply adding 'I know who' after the statement with *-khi*; instead, the context must provide some motivation for withholding information, as seen in (276).

- (276) *Tiwa* (Dawson 2018:ex. 8)
 There is a man who is constantly bothering the speaker, which includes always asking her invasive questions about her recent activities.
Pakhál-khí lí-dom. Ang si-w pakhál, thêbo nága sóng os-ya.
 when-KHI go-PST 1SG know-NEUT when but 2SG.DAT tell AUX-NEG
 'I went sometime. I know when, but I won't tell you.'

In this way, the Kipsigis pattern aligns with the broader cross-linguistic picture, in that cancellation of the ignorance effects triggered by *-yan* is possible, but constitutes a marked—although not contradictory—discourse move.

Third, the ignorance effects typically triggered by *-yan* forms disappear in downward-entailing contexts; when the epistemic indefinite scopes under a conditional operator (277) or negation (278), there is no epistemic effect. In these sentences, which contain *-yan* forms, there is no epistemic effect, and the *-yan* form is best translated with 'any' instead. In fact, as Dawson (2018:353) points out, it is quite difficult to even imagine what ignorance effects would look like when the indefinite is in a downward-entailing context (e.g. 'it is not the case that Kibet swept any room, but/and I don't know which room he didn't sweep').

- (277) Chepkoech and Kibet are playing hide-and-seek—Kibet is the seeker and Chepkoech is hiding. I say:
 Kot ko-uny-e-kεε Chεεpkœech εen **rœm-i-yan** ko-nyoor-u Kibeet.
 if 3.SBJV-hide-IPFV-REFL C. in room-TH-YAN 3.SBJV-find-VEN K.
 'If Chepkoech is hiding in any room, Kibet will find her.'
 ✓ Whichever room Chepkoech is hiding in, Kibet's sure to find her. if > ∃
- (278) You overhear your mom complaining and ask your sibling why she's so upset. Your sibling replies:
 Ma-i-buch Kibeet **rœm-i-yan**.
 NEG-3-sweep K. room-TH-YAN
 'Kibet didn't sweep any room'
 ✓ Kibet didn't sweep any room. ¬ > ∃

This behavior speaks against an analysis in which ignorance is lexically encoded in Kipsigis, since this type of ignorance would not be predicted to disappear in downward-entailing contexts, as it is part of the form's asserted content.³

³An exception to this claim is found in Aloni & Port (2015), where they are able to capture the disappearance of

In light of the data presented here, I conclude that the ignorance effects triggered by *-yan* forms are pragmatically derived via competition with other nominals in the language. However, this observation raises two important questions. First, what other type of nominal competes with Kipsigis *-yan* forms to derive the observed ignorance effects? And second, what are the semantics of *-yan* forms and their competitor that give rise to the precise type of ignorance effects observed in §3.3? In the following section, I describe the distribution of *-yan* in terms of its possible domains of quantification and its scope, and then in §3.6, I compare this distribution with that of nouns with a secondary suffix in Kipsigis. In doing so, I show that *-yan* forms are in competition with these nouns. Based on their empirical properties, I suggest that *-yan* introduces basic existential quantification, while nouns with a secondary suffix introduce free choice function variables that are resolved via a contextually supplied assignment function. In §3.7, I flesh out these semantic proposals and walk through the pragmatic reasoning that derives ignorance effects.

3.5 Semantic behavior of *-yan* forms

Existing pragmatic analyses of epistemic indefinites fall into two main classes: those that rely on domain widening semantics for the epistemic indefinite (Kratzer & Shimoyama 2002; Chierchia 2006; Alonso-Ovalle & Menéndez-Benito 2010, 2013, 2017; Dawson 2018, 2020) and those that propose choice functional semantics for the epistemic indefinite (Yanovich 2005; Dawson 2018, 2020). Both of these types of accounts derive ignorance effects via competition between the epistemic indefinite and other nominals in the language (i.e. plain indefinites).

These different analyses have also been invoked to capture the different kinds of ignorance effects triggered by epistemic indefinites; in particular, Dawson (2018) links domain widening semantics to first order ignorance, and choice functions to higher order ignorance. Notably, these analyses also make different predictions about other behaviors of the epistemic indefinites beyond the type of ignorance conveyed. This section outlines these predictions and shows that Kipsigis *-yan* forms do not pattern neatly with either type of analysis, setting the stage for my novel account in §3.7.

3.5.1 Domain widening analyses

Domain widening analyses of epistemic indefinites (Kratzer & Shimoyama 2002; Chierchia 2006; Alonso-Ovalle & Menéndez-Benito 2010, 2013, 2017; Dawson 2018) impose requirements on the

ignorance effects within the scope of negation, even though they argue that these effects are lexically encoded. On their account, ignorance effects arise when the indefinite requires the witness to be identified in some non-standard way (e.g. via description rather than naming). However, when an existential occurs in the scope of negation, no question of identification arises; instead, they assume that indefinites can also trigger domain widening in these contexts. In this way, indefinites serve a fundamentally different role under negation—widening the domain rather than shifting how an individual is to be identified—which captures the disappearance of ignorance effects. Yet this type of analysis raises questions about the connection between the two functions that epistemic indefinites can have and does not fully capture the Kipsigis pattern, since the ignorance effects seen with *-yan* forms show multiple hallmarks of conversational implicature.

domain that the indefinite quantifies over, arguing that it must be expanded in some way. For instance, Kratzer & Shimoyama (2002) claim that German *irgendein* shows Free Choice effects, meaning that any individual that satisfies the indefinite’s restrictor must be a possible witness for *irgendein* to be felicitous. These effects can be seen in (279), where *irgendein* is appropriate if all doctors are possible marriage options (279a) but infelicitous if only a subset of doctors are possible options (279b).⁴ In light of this pattern, Kratzer & Shimoyama argue that *irgendein* is a maximal domain widener.

- (279) *German* (Kratzer & Shimoyama 2002:ex. 9)
 Mary muss **irgend-einen** Arzt heiraten.
 M. has.to IRGEND-a doctor marry
 ‘Mary has to marry a doctor.’
- a. ✓ There are lots of doctors in the world. Mary has to marry a doctor, and any doctor is a permitted option.
- b. # There are lots of doctors in the world. Mary has to marry one of two doctors—Dr. Heintz or Dr. Dietz—and those are the only permitted options for her.

It is also possible for epistemic indefinites to place weaker requirements on their domain of quantification. For example, Alonso-Ovalle & Menéndez-Benito (2010) argue that Spanish *algún* imposes an anti-singleton constraint on its domain of quantification: it must contain at least two individuals, though it need not be maximal (280b). Alonso-Ovalle & Menéndez-Benito dub this weaker epistemic effect “modal variation” as opposed to Free Choice.

- (280) *Spanish* (Alonso-Ovalle & Menéndez-Benito 2010:ex. 14)
 Juan tiene que estar en **alguna** habitación de la casa.
 J. has to be in ALGUN room of the house
 ‘Juan must be in a room of the house.’
- a. ✓ Maria, Juan, and Pedro are playing hide-and-seek in their country house. Juan is hiding. Maria and Pedro haven’t started looking for Juan yet. Pedro believes that Juan isn’t hiding in the garden or in the barn: he’s sure that Juan’s inside the house. But as far as Pedro knows, Juan could be in **any room in the house**.
- b. ✓ Maria, Juan, and Pedro are playing hide-and-seek in their country house. Juan is hiding. Maria and Pedro haven’t started looking for Juan yet. Pedro believes that Juan isn’t hiding in the garden or in the barn: he’s sure that Juan’s inside the house. Furthermore, Pedro’s sure that Juan **isn’t in the bathroom or in the kitchen**. As far as he knows, Juan could be in any of the other rooms in the house.

These domain widening analyses make it straightforward to encode both Free choice effects and cases of modal variation—depending on how much the domain is widened—and they naturally derive first order ignorance. To see how this works, consider the denotation of Spanish

⁴This is a slight simplification of the empirical picture; German *irgendein* can trigger weaker epistemic effects in some contexts (e.g. under epistemic rather than deontic modals). For more discussion, see Aloni & Port (2015).

algún in (281). At its core, *algún* makes an existential claim. However, it comes with the presupposition that the indefinite's domain of quantification contains two or more individuals; this is formalized in (281) as an anti-singleton requirement on f , a subset selection function that picks out a subset of the individuals denoted by *algún*'s restrictor.

$$(281) \quad \llbracket \text{algún} \rrbracket = \lambda f_{\langle et, et \rangle} . \lambda P_{\langle e, t \rangle} . \lambda Q_{\langle e, t \rangle} : \text{anti-singleton}(f) . \exists x [f(P)(x) \ \& \ Q(x)]$$

(Alonso-Ovalle & Menéndez-Benito 2010:ex. 54)

The anti-singleton constraint in (281) is only satisfied when the subset of individuals with property P picked out by f contains more than one member.

On these accounts, ignorance effects arise as a quantity implicature due to a speaker's choice to use a domain widening indefinite over a plain indefinite (i.e. existential quantification without an anti-singleton presupposition). Plain indefinites are compatible with singleton domains of quantification, in a way that domain widening ones are not. Consequently, existential quantification with a plain indefinite can make a stronger claim: there exists some individual x in the singleton set $\{x\}$ with properties P and Q . This contrasts with the necessarily weaker claim made with a domain widening indefinite: there exists some individual x in the set $\{x, y, \dots\}$ with properties P and Q . When these forms are in competition, use of the domain widening indefinite leads to the inference that the speaker does not know which individual in the necessarily non-singleton domain of quantification witnesses the existential claim; if the speaker knew this information, then they would use the plain indefinite, which allows for a singleton domain of quantification.

In this way, deriving first order ignorance is relatively straightforward on these accounts. However, they rule out ignorance that is *exclusively* higher order. If one only considers individuals in the world of evaluation (following Alonso-Ovalle & Menéndez-Benito 2010), the anti-singleton constraint in (281) makes it impossible to derive higher order ignorance; the speaker cannot use the domain widening indefinite if they know the witness to the indefinite but not one or more of their salient properties. However, if one considers individuals across possible worlds, the non-singleton set that the indefinite ranges over can consist of individuals whose identities remain the same across possible worlds, but whose other salient properties differ. This makes space for higher order ignorance.⁵ However, such an analysis also makes space for first order ignorance, since it is also possible for the intensionalized non-singleton set to consist of individuals whose identities and salient properties vary across possible worlds. In this way, higher order ignorance is possible with a novel understanding of domain widening semantics, but there must be first order ignorance too.

These analyses also predict that domain widening epistemic indefinites should be incompatible with restrictors with singleton extensions (e.g. superlatives, inherently singular denoting nouns). This prediction is welcome for Spanish *algún*, which cannot co-occur with a superlative restrictor (282).

⁵In fact, see Bossi (2022) for an earlier analysis of Kipsigis *-yan* forms that takes this approach.

- (282) *Spanish* (Alonso-Ovalle & Menéndez-Benito 2010:ex. 47)
 # Juan compró **algún** libro que resultó ser el más caro de la librería.
 J. bought ALGUN book that happened to.be the most expensive in the bookstore
 ‘Juan bought a book that happened to be the most expensive one in the store.’

However, this prediction is incorrect for Kipsigis, where *-yan* can take a necessarily singleton, superlative restrictor (283) - (284).⁶ In these examples, the *-yan* form triggers higher order ignorance—in (283) about the car’s make or model, and in (284) about where in the house the relevant room is. In both of these cases, the speaker is missing some crucial piece of information that prevents them from fully identifying the witness to the indefinite.

- (283) I attended the auction where Linus bought the most expensive black car. I saw the car, but I didn’t learn any more information about it like its make or model. I say:
 Kɔɔ-∅-al Linas kar-**r-yan** ne tʊʊi ne koo-kali ɛɛn tʊgʊl ɛɛn ɔkshɛn.
 PST-3-buy L. car-TH-YAN REL.SG black REL.SG PST-expensive in all in auction
 ‘Linus bought the most expensive black car at the auction.’
- (284) Kibet and Chepkoech are playing hide-and-seek—Kibet is the seeker and Chepkoech is hiding. Kibet was told that Chepkoech is in the biggest room in the house. He knows that the living room is typically biggest, so he thinks that Chepkoech must be there. But Kibet has never been here before, so he doesn’t know where the living room is! He says:
 ∅-Uny-e-kɛɛ Cheɛpkɔɛch ɛɛn **rʊʊm-r-yan** ne oo ɛɛn tʊgʊl.
 3-hide-IPFV-REFL C. in room-TH-YAN REL.SG big in all
 ‘Chepkoech is hiding in some room that is the biggest of all.’

In this way, a key prediction of domain widening analyses—that epistemic indefinites of this type should be incompatible with singleton domains of quantification—is not upheld in Kipsigis, though it is in German and Spanish.

3.5.2 Choice functional analyses

Unlike domain widening accounts, choice functional analyses involve existential quantification over choice functions (CFs) rather than over individuals directly (Reinhart 1997; Winter 1997; Kratzer 1998; Matthewson 1999; Yanovich 2005; Dawson 2018, 2020). CFs are functions that take in a set and return a member of that set. For instance, Dawson (2018) claims that Tiwa *-khi* indefinites introduce a CF that ranges over the property denoted by their restrictor (285). A Tiwa sentence with a *-khi* indefinite like that in (286a) has the LF in (286b).

- (285) $\llbracket \text{wh-khi} \rrbracket = \lambda P.f(P)$, where f is a CF (Dawson 2018:ex. 29)

⁶Kipsigis does not have a single, dedicated morphological strategy for creating superlatives. Instead, speakers can use a paraphrase like those in (283) - (284), which literally translates to ‘N that is ADJ in all’. They can also use an ‘exceed’-type comparative, as in ‘N that is ADJ exceeds all’. Although I do not provide a detailed semantic analysis of these constructions here, I assume that they are similar enough to a morphological superlative to illustrate the point that *-yan* allows a singleton domain of quantification. Furthermore, both of these constructions represent very common ways to express superlative meaning across languages (Bobaljik 2012).

(286) *Tiwa* (Dawson 2018:ex. 37)

- a. **Shar-khí** phi-dom.
 who-KHI come-PST
 ‘Someone came.’
- b. $\exists f[\text{CH}(f) \ \& \ \text{came}(f(\text{human}))]$

When a speaker chooses to use this type of higher order quantification rather than quantification over individuals, interlocutors reason that the speaker must be ignorant about how the witness is to be selected or, more technically, what function uniquely picks out the correct individual from the set that the CF ranges over. In this way, the choice functional account derives first order and higher order ignorance in the same way; in cases of first order ignorance, the relevant, unknown way of selecting a witness is via its name or identity, while in cases of higher order ignorance, the relevant, unknown way of selecting a witness involves some other property.

Importantly, because the choice functional indefinite places no requirements on the domain that it ranges over, it should be compatible with a singleton restrictor. This prediction meshes nicely with the Kipsigis facts in (283) - (284) and is upheld in *Tiwa* as well, seen in (287) with the inherently singular denoting noun *India-ne PM* ‘Indian Prime Minister’.

(287) *Tiwa* (Dawson 2018:ex. 36)

- Ang **shar-khí** India-ne PM-go lak mán-a lí-do.
 1SG who-KHI India-GEN PM-ACC meet-INF go-IPFV
 ‘I’m going to meet some Indian Prime Minister.’

However, many implementations of CFs predict that they should take exceptional wide scope—either because the CF variable is contextually resolved (Kratzer 1998)⁷ or because it is existentially closed wide (Matthewson 1999). *Tiwa -khi* indefinites show this predicted exceptional wide scope; forms with *-khi* must take wide scope, even out of islands, as seen in (288) with a conditional antecedent.

(288) *Tiwa* (Dawson 2020:151)

- Chidí **shar-khí** sister-go lak mán-a phi-gai-do, Saldi khúp khâdu-gam.
 if who-KHI sister-ACC meet-INF come-COND-TOP S. very happy-MOD
 ‘If Saldi meets some nun, she would be very happy.’
- a. # Saldi loves nuns. She always wants to talk to any nun she sees. She’s even thinking of becoming one. Every time she sees any nun, she feels very happy. *if > \exists
- b. ✓ There are several nuns that live in a convent nearby, and Saldi doesn’t like any of them, except for one. She always likes to talk with that one sister. \exists > if
 $\exists f[\text{CH}(f) \ \& \ \text{meet}(\text{Saldi})(f(\text{nun})) \rightarrow \text{happy}(\text{Saldi})]$

Kipsigis *-yan* forms, however, can generally scope below or above operators like negation (289), universal quantifiers (290), modals (291), and attitude verbs (292).

⁷On this approach, the indefinite is actually referential and so not a scope-taking element at all. Discussion of scope is, then, somewhat imprecise here.

- (289) You overhear your mom complaining and ask your sibling why she's so upset. Your sibling replies:
 Ma-i-buch Kibeet **rɔɔm-i-yan**.
 NEG-3-sweep K. room-TH-YAN
 'Kibet didn't sweep any/some room'
 ✓ Kibet didn't sweep any room. $\neg > \exists$
 ✓ Kibet forgot to sweep a particular, unknown room. $\exists > \neg$
- (290) Ø-Uny-e-kɛɛ laakwɛɛt aɣɛ tɔgɔl ɛɛn **rɔɔm-i-yan**.
 3-hide-IPFV-REFL child every in room-TH-YAN
 'Every child is hiding in a/some room.'
 ✓ Every child is hiding a different room. $\forall > \exists$
 ✓ Every child is hiding in the same, unknown room. $\exists > \forall$
- (291) Nyaal-u ko-buuch Kibeet **rɔɔm-i-yan**.
 MOD-IPFV 3.SBJV-sweep K. room-TH-YAN
 'Kibet has to sweep any/some room.'
 ✓ Kibet has to sweep any room. $\square > \exists$
 ✓ There is a particular, unknown room that Kibet has to sweep. $\exists > \square$
- (292) Ø-Mach-e ko-uny-e-kɛɛ Cɛɛpkɔɛch ɛɛn **rɔɔm-i-yan**.
 3-want-IPFV 3.SBJV-hide-IPFV-REFL C. in room-TH-YAN
 'Chepkoech wants to hide in any/some room.'
 ✓ Chepkoech wants to hide in any room. want $> \exists$
 ✓ Chepkoech wants to hide in a particular, unknown room. $\exists > \text{want}$

In fact, on some accounts, the exceptional wide scope seen with choice functional indefinites is predicted to hold in additional contexts where indefinite scope is often otherwise restricted. For instance, many indefinites like English *a* must scope below downward-entailing operators that bind into their restrictor (Brasoveanu & Farkas 2011). This restriction is known as the “Binder-Roof Constraint” and can be seen in (293) for English *a*. While *a* is usually scopally flexible, when the downward-entailing operator *no one* binds a pronoun in the restrictor of the indefinite, it can no longer scope above negation (293b).

- (293) No one_i sent a letter that they_i wrote.
 a. ✓ Tim, Emily, Kyle, and Alex each wrote 2 letters, but no one sent any of their own letters. no one $> \exists$
 b. # Tim, Emily, Kyle, and Alex each wrote 2 letters. They all sent one of their letters, but each person didn't send the other one of their letters. $*\exists > \text{no one}$

Once again, this predicted exceptional wide scope is welcome for Tiwa *-khi*, which shows the reverse of the English pattern (294); here, only the wide scope interpretation is available for the *-khi* indefinite, even though the downward-entailing operator *sharbo* ‘no one’ binds a pronoun in the indefinite’s restrictor.

- (294) *Tiwa* (Dawson 2020:173)
 Sharbo_i [pakhâ-khí [_{RC} othê_i pre la-wa] khugrí] -gô marê ton-ya-m.
 nobody which-KHI REFL.GEN buy AUX-NMLZ dog -ACC kill AUX-NEG-PST
 ‘Nobody_i killed a dog that he_i bought.’
- a. # Each person bought several dogs. A rabies outbreak meant that all the dogs had to be killed. However, each person refused to kill their own dogs. *no one > ∃
- b. ✓ Each person bought several dogs. Because of a rabies outbreak, each person killed all of their dogs, except for one. ∃ > no one
 ∃*f*[CH(*f*) & ¬∃*x*[*x* killed *f*(dog that *x* bought)]]

Yet, as previously, this is a bad prediction for Kipsigis *-yan*, which patterns with English *a* in obeying the Binder-Roof Constraint (295); here, only the narrow scope interpretation of the indefinite is possible.

- (295) Ma-∅-mach-e chi_i kɔ-al-da [marında-yan [_{RC} ne koo-∅_i-nɔp-e]].
 NEG-3-want-IPFV person 3.SBJV-buy-IT dress-YAN REL.SG PST-3-sew-IPFV
 ‘No one_i wants to sell any dress that they_i sewed.’
- a. ✓ Chepkoech, Cheronno, and Cheptoo all made many dresses. Each woman was planning to sell all the dresses that she made, but eventually each woman decided that she didn’t want to sell any of her own products. no one > ∃
- b. # Chepkoech, Cheronno, and Cheptoo all made many dresses. Each woman was planning to sell all the dresses that she made, but eventually each woman decided to keep one for herself. *∃ > no one

It is worth noting, however, that not all analyses of CFs make these scopal predictions. Scopal flexibility is possible if existential closure of the CF is permitted at various points in the structure (Reinhart 1997; Winter 1997; Yanovich 2005). Furthermore, the disappearance of this scopal flexibility in Binder-Roof Constraint contexts like (293) - (295) can be captured through stipulations that require existential closure of the CF below operators that bind into the choice functional indefinite’s restrictor (Chierchia 2001; Schwarz 2001, 2011). Nevertheless, analyzing English *a* or Kipsigis *-yan* as CFs introduces complications into the analysis—especially when all the predictions of a choice functional account are upheld in Tiwa. For instance, as Dawson (2020) highlights, there is no clear, independent motivation for the type of stipulation proposed by Chierchia (2001) and Schwarz (2001, 2011), given that existential closure does not affect binding relationships. Furthermore, this type of restriction would have to be language-specific, given that it seems to hold in English and Kipsigis but not in Tiwa. This leads Dawson (2020) to conclude that data like Tiwa (294) are the true hallmark of choice functional indefinites. If this is so, Kipsigis *-yan* forms do not behave as expected for CFs.

Taken together, the data in this section show that the pattern seen with Kipsigis *-yan* does not align with those seen in previous work, either in work supporting domain widening approaches or in work supporting choice functional ones. Table 3.1 summarizes this state of affairs.

	Singleton domain?	Scopal flexibility?
Domain widening	no	yes
Choice functions	yes	no (modulo stipulations)
Kipsigis <i>-yan</i>	yes	yes
Spanish <i>algún</i>	no	yes
Tiwa <i>-khí</i>	yes	no

Table 3.1: Kipsigis *-yan* vs. predictions of existing analyses and documented patterns

3.6 Nouns with a secondary suffix in Kipsigis

With this distribution of *-yan* forms in mind, here I outline the distribution of nouns with a secondary suffix in Kipsigis and provide an analysis of them. In doing so, I show that *-yan* forms and these other nouns are in competition in all of the contexts where use of the *-yan* form triggers ignorance effects. This pattern suggests that competition between *-yan* forms and nouns with a secondary suffix is the relevant instance of competition that derives the ignorance effects seen in §3.3. With the empirical properties of nouns with a secondary suffix established, I propose that the secondary suffix introduces a free choice function variable that is contextually resolved. This sets the stage for §3.7, where I provide a semantic analysis of *-yan* and outline how competition between *-yan* forms and nouns with a secondary suffix gives rise to the precise type of ignorance seen in Kipsigis.

3.6.1 Distribution of nouns with a secondary suffix

As mentioned in §3.2 and in Chapter 2 §2.3.1, Kipsigis is generally classified as a bare noun language, although nouns in the language have a complex internal structure and contain a so-called “secondary suffix”, which differentiates them from canonical bare nouns. Here, I focus on the properties of singular nouns with a secondary suffix, since *-yan* forms are always singular, but see Chapter 2 §2.3.2 for discussion of the interpretation of mass and plural nouns, which also contain a secondary suffix but show a different interpretive pattern. Nouns with a secondary suffix in Kipsigis freely introduce new discourse referents, as seen in (296) - (297). These indefinite uses do not require the numeral *ageenge* ‘one’, which makes for a marked contrast with other languages that have been prominently studied in the bare noun literature, including Hindi and Russian (see e.g. Dayal 2004).

- (296) There are 6 dogs outside. I hear lots of barking and see you come inside looking disheveled. I ask what happened. You say:
 Ka-∅-sʊs-an ng’ook-ta.
 PST1-3-bite-1SG.O dog-SEC
 ‘A dog bit me.’

(297) There's a big beehive outside with lots of bees. You come inside holding your arm. I ask what happened. You say:

Ka-∅-uut-an segeem-yaan-ta-it (segeemyaat).
 PST1-3-sting-1SG.O bee-SG-TH-SEC
 'A bee stung me.'

It is also important to note that nouns with a secondary suffix are not necessarily any standard kind of definite description; there is no previous mention of dogs in (296) or bees in (297) to license an anaphoric definite, and there is no unique dog or bee in context to license a uniqueness definite.

This indefinite use of these nouns is very common in the language. Stories often start with nouns with a secondary suffix as in (298) - (299), both taken from the beginning of different Kipsigis texts. While these nouns can surface in existential constructions with the copula *mii*—e.g. *kiimii moset* 'there was a monkey' in (298)—they do not require this type of structure; the nouns *kɛtɪt* 'tree', *oineet* 'river', *kokweet* 'village', and *cheeplanget* 'leopard' all contain secondary suffixes and introduce new discourse referents without any kind of existential syntax.

(298) Kii-mii **moset** ne kii-∅-am-e logoeek εn **kɛtɪt** parak ne kii-mii
 PST3-COP monkey REL.SG PST3-3-eat-IPFV fruits in tree top REL.SG PST3-COP
 taband-aap **oineet**.
 bank-of river

'There once was a monkey that ate fruits on top of a tree that was on a river bank.'

(299) Kii-∅-leen-j-eech obule kii-mii **cheepkeleliot** ne kii-ki-kuren Tabutany
 PST3-3-say-IPFV-APPL-1PL.IO uncle PST3-COP young.lady REL.SG PST3-1PL-call T.
 ne kii-∅-meny-e **kokweet** ne kii-negit **cheeplanget**.
 REL.SG PST3-3-live-IPFV village REL.SG PST3-near leopard
 'The uncle told us that there was a young lady called Tabutany who lived in a village that was close to a leopard.'

Based on this sort of data, I conclude that nouns with a secondary suffix in Kipsigis are a type of indefinite in the simple sense that they may introduce new discourse referents (see e.g. Dawson & Jenks 2021). These indefinites must take wide scope and are most natural when the speaker is able to identify the noun's referent. This pattern is shown in (300) with negation, where the noun *rɔvmɪt* 'room' must scope above negation; while the narrow scope reading is ruled out, the noun is felicitous in the wide scope context, in which Kibet did not sweep a particular room that is known to the speaker.

(300) Ma-i-buch Kibeet **rɔvmɪt-i-it**.
 NEG-3-sweep K. room-TH-SEC
 'Kibet didn't sweep a certain room.'

- a. # Kibet didn't do any of his chores—he didn't sweep a single room! *¬ > ∃
 b. ✓ Kibet did some of his chores but not all of them. There are 5 rooms that he needs to sweep. Today, he swept 4 of them, but he didn't sweep 1: he skipped the kitchen because it's the messiest room in the house and too much work to sweep. ∃ > ¬

The sentence in (300) is naturally followed by (301), which provides more information about which room Kibet did not sweep.

- (301) Ma-i-buch chiigeet amɔn mur nɪa.
 NEG-3-sweep kitchen because dirty very
 ‘He didn’t sweep the kitchen because it’s very dirty.’

On the other hand, to express the narrow scope reading, speakers use the expression *rɔɔmɪt agɛ tɔgɔl* ‘any room at all’ (302).

- (302) Ma-i-buch Kibeet rɔɔm-ɪ-ɪt agɛ tɔgɔl.
 NEG-3-sweep K. room-TH-SEC any all
 ‘Kibet didn’t sweep any room at all.’

An identical pattern arises with other scope-taking elements like modals (303) and the verb *mach* ‘want’ (304). In these examples, the noun *rɔɔmɪt* ‘room’ is infelicitous in the narrow scope context, but is natural in the wide scope context, particularly if the speaker is able to identify the noun’s referent. As before, the speaker is able to follow their utterance with another sentence that provides more information about the noun’s referent; these continuations are provided in the c. examples.

- (303) Nyaal-u ko-buch Kibeet rɔɔm-ɪ-ɪt.
 MOD-IPFV 3.SBJV-sweep K. room-TH-SEC
 ‘Kibet must sweep a certain room.’
- a. # There are 5 rooms in the house. For his chores, Kibet has to sweep any one of those rooms. It could be the kitchen or the bathroom or the living room. *□ > ∃
- b. ✓ There are 5 rooms in the house. For his chores, Kibet has to sweep the kitchen—the messiest room in the house. Can be followed by (303c). ∃ > □
- c. Nyaal-u ko-buch chiigeet amɔn mur nɪa.
 MOD-IPFV 3.SBJV-sweep kitchen because dirty very.
 ‘He has to sweep the kitchen because it’s very dirty.’
- (304) Ø-Mach-e ko-uny-e-kɛɛ Chɛpkɔɛch ɛɛn rɔɔm-ɪ-ɪt.
 3-want-IPFV 3.SBJV-hide-IPFV-REFL C. in room-TH-SEC
 ‘Chepkoech wants to hide in a certain room.’
- a. # Kibet and Chepkoech are playing hide-and-seek inside the house and outside in the yard. Because it’s cold, Chepkoech wants to hide inside in a room, but she’s happy to hide in any room. *want > ∃
- b. ✓ Kibet and Chepkoech are playing hide-and-seek inside the house and outside in the yard. Because it’s cold, Chepkoech wants to hide inside in the kitchen because it’s the warmest room in the house. Can be followed by (304c). ∃ > want

- c. Ø-Mach-e ko-uny-e-kεε εεn chiigeet ngamɔn purgei ko-siir εεn
 3-want-IPFV 3.SBJV-hide-IPFV-REFL in kitchen because warm 3.SBJV-exceed in
 tɔgɔl.
 all
 ‘She wants to hide in the kitchen because it’s the warmest.’

Across all of these contexts, speaker comments suggest that the nouns with a secondary suffix are receiving specific, referential interpretations. For instance, in response to (303), speakers suggest that “you’re singling out a room” and that an interlocutor might ask the speaker which particular room Kibet has to sweep. Likewise, speakers comment that the sentence in (304) indicates that “there’s a certain room in mind, [that Chepkoech] has a preference for a particular room.”

In this way, the empirical generalization is that nouns with a secondary suffix in Kipsigis display consistent wide scope and are most natural when the speaker can identify the noun’s referent. Against this backdrop, recall from §3.5 that *-yan* forms display scopal flexibility, meaning that they are acceptable in a wider range of contexts than nouns with a secondary suffix. However, ignorance effects arise with *-yan* forms in precisely the contexts where both types of nouns are felicitous. Even though *-yan* forms can take narrow scope with respect to other operators like negation, modals, and attitude verbs, no ignorance effects arise in these contexts. This fact is reflected in speaker translations of *-yan* forms in these contexts as ‘any N’ rather than ‘some N.’ In fact, as noted in §3.4 and Dawson (2018:353), it is difficult to imagine what ignorance effects would look like in these contexts (e.g. ‘Kibet must sweep any room, but/and I don’t know which room he must sweep’). Table 3.2 summarizes this pattern, highlighting that ignorance effects arise in precisely the contexts where nouns with a secondary suffix and *-yan* forms are in competition. Given this distribution, I conclude that competition between these two types of nominals gives rise to the ignorance effects seen with *-yan*.

	Nouns with secondary suffix	<i>-yan</i> forms	Ignorance effects?
$\neg > \exists$	#	✓	No
$\exists > \neg$	✓	✓	Yes
want > \exists	#	✓	No
$\exists > \text{want}$	✓	✓	Yes
$\square > \exists$	#	✓	No
$\exists > \square$	✓	✓	Yes

Table 3.2: Distribution of noun types and ignorance effects in Kipsigis

3.6.2 Semantics of nouns with a secondary suffix

To capture the fact that nouns with a secondary suffix in Kipsigis take obligatory wide scope, I propose that the secondary suffix in singular nouns introduces a choice function variable. Fol-

lowing Winter (1997), I suggest that the choice function variable introduced by the secondary suffix combines directly with the property denoted by its restrictor. The proposed denotation of the secondary suffix is shown in (305).

(305) $\llbracket \text{-SEC} \rrbracket = \lambda P_{\langle e,t \rangle}.f(P)$, where f is a Choice Function

For instance, in the sentence in (306), the secondary suffix within the singular noun *ng'ookta* 'dog' introduces a choice function variable, and the sentence asserts that the individual picked out from $\{x : \text{dog}(x)\}$ by the choice function bit the speaker.

(306) There are 6 dogs outside. I hear lots of barking and see you come inside looking disheveled.
I ask what happened. You say:
Ka-Ø-sʊs-an **ng'ook-ta.**
PST1-3-bite-1SG.O dog-SEC
'A dog bit me.'
[CH(f) & bite(speaker)(f (dog))]

This view, in which the choice function variable combines directly with a property, contrasts with one in which the element that introduces the choice function variable combines with an indeterminate pronoun (i.e. a set of alternatives) to form an indefinite (e.g. Yanovich 2005; Cable 2010; Dawson 2020; among others). This difference is necessitated by the empirical fact that the secondary suffix attaches directly to an ordinary N projection in Kipsigis, rather than to a *wh*-word or another alternative-denoting element.

On a choice functional approach, there are two ways to derive obligatory wide scope: either the choice function variable is existentially closed above other relevant operators (Reinhart 1997; Winter 1997; Matthewson 1999) or it is left free and resolved via a contextually supplied assignment function (Kratzer 1998; Renans 2018). On the latter approach, the indefinite is not actually a scope-taking element at all; it is referential and does not interact scopally with other operators. In addition, the value of the choice function does not need to be contextually supplied or known to all interlocutors—it is sufficient for the speaker to have a particular witness or choice function in mind (Kratzer 1998 following Fodor & Sag 1982).⁸

My description of nouns with a secondary suffix as specific, referential indefinites in §3.6.1 points toward a free choice function variable analysis. Direct evidence for this analysis comes from data like that in (307). This diagnostic rests on the idea that, in order for a free choice function variable to be felicitous, the speaker must have a particular individual or choice function in mind when uttering their sentence. As Matthewson (1999) points out, it is difficult to test what the speaker needs to have in mind when speaking, though the context and mini-discourse in

⁸It is an open question why choice function variables behave differently in this respect than other variables (e.g. pronouns). While it is possible for the value of a free choice function variable to be known only to the speaker, the same is not true for pronouns, which are only felicitous if their intended reference is known to all interlocutors. While it is reasonable to want all contextually-resolved variables to behave similarly in this respect, explaining why this difference exists is a challenge for all analyses that involve free choice function variables and lies outside the scope of this chapter.

(307) seek to establish one particular book as the speaker's main focus. If Kipsigis speakers deem Cherono's statement in (307a) inappropriate, this suggests that the particular book that she has in mind features into the truth conditions of the sentence. If, on the other hand, they deem it appropriate, this suggests that the particular book that she has in mind does not feature into the sentence's truth conditions.

(307) Cherono's parents give her money to go to the store and buy a copy of the book *Half of a Yellow Sun* because she has to write a book report on it. She goes to the store and asks the salesperson for a copy of the book. The salesperson gives her a book in a bag, and Cherono pays for it with her parents' money. When she gets home, she tells her parents:

- a. Kaa-al **kɪtabɔɔt.**
 PST1.1SG-buy book
 'I bought a certain book.'

When uttering (307a), Cherono thinks that the book she bought was *Half of a Yellow Sun*. But when she opens the bag, she finds out that the salesperson made a mistake and she really bought a book by Ngũgĩ wa Thiong'o, not *Half of a Yellow Sun*.

In light of this discovery, was Cherono's statement in (307a) wrong? Speaker's comment: "Yes, once she noticed the mistake [i.e. the book mix-up], then she was wrong."
 (context adapted from Matthewson 1999:ex. 95)

In response to this context and mini-discourse, Kipsigis speakers report that, at the moment when Cherono uttered (307a), her statement was appropriate because she was still under the impression that she bought *Half of a Yellow Sun*. However, upon realizing that the salesperson made a mistake and sold her a different book, speakers comment that Cherono's statement in (307a) is inappropriate. This pattern contrasts with the one reported for St'át'incets by Matthewson (1999) who notes that speakers deem the St'át'incets version of the sentence in (307a) as true just in case any book satisfies the conditions given by the rest of the sentence; in St'át'incets—unlike in Kipsigis—it does not affect the truth conditions of the sentence which particular book satisfies those relevant conditions. In this way, the empirical pattern reported here for Kipsigis suggests that the particular individual that the speaker has in mind when using a noun with a secondary suffix impacts the truth conditions of the sentence; such a pattern is neatly captured on an analysis in which the secondary suffix introduces a free choice function variable that is contextually resolved.

Interestingly, one last piece of evidence for a choice functional analysis of Kipsigis nouns with a secondary suffix comes from a case in which they do *not* take obligatory wide scope. In particular, these nouns are able to take apparent narrow scope under universal quantification, but not under other kinds of operator. In sentences like (308), the noun *rɔvmɪt* 'room' is appropriate in a context where all the children are hiding alone in different rooms (i.e. $\exists > \forall$), as well as a context where all the children are hiding together in the same room (i.e. $\forall > \exists$).⁹

⁹For simplicity, I treat the prepositional phrase in sentences with the verb *uny* 'hide' as an argument of the verb, but this decision has no real bearing on the semantic analysis.

2019). Slightly more detailed semantic description of the secondary suffix is found in Creider & Creider (1989), who claim: “The best way of characterizing the use of these two forms is to say that the primary form is used when the existence of the item referred to is not assumed by the speaker, while the secondary form is used in those situations where the speaker does assume the existence of the item referred to” (169).

While these characterizations do not provide direct evidence for the choice functional analysis here, the reported intuitions fall out of my account relatively straightforwardly. Singular nouns that contain a secondary suffix are specific, referential indefinites that occupy a space somewhere between definite and indefinite interpretations, at least in Kipsigis. On the one hand, these forms are referential—just like definite descriptions—but they are still indefinite, in that they can be used to introduce new discourse referents and do not require uniqueness or anaphoricity to be licensed. Furthermore, the definite-forming operator ι is a particular choice function: one that picks out the unique individual in the context with the property denoted by the set that the definite determiner combines with. In this way, drawing the line between definite and indefinite can be challenging against a choice functional semantic backdrop; this difficulty perhaps explains the varied descriptions of the meaning of secondary noun forms across the literature on Kalenjin languages. An interesting question for future research is whether a Kipsigis-style choice functional analysis can be applied more broadly in languages where the alternation between primary and secondary forms continues.

3.7 An analysis of *-yan*

3.7.1 Semantics of *-yan* forms

Against this backdrop and given the semantic patterns in §3.5, I propose that *-yan* introduces basic existential quantification over individuals, with the denotation in (310).

$$(310) \quad \llbracket \text{-yan} \rrbracket = \lambda P_{\langle e,t \rangle} . \lambda Q_{\langle e,t \rangle} . \exists x [P(x) \ \& \ Q(x)]$$

Evidence that *-yan* introduces this sort of existential quantification comes from two sources. First, *-yan* forms display constrained scopal flexibility; they can generally scope below or above other operators in a sentence, although they cannot outscope an operator that binds into their restrictor (§3.5.2). In addition, *-yan* forms seem to obey scope islands, as seen in (311) where the existential quantification contributed by *-yan* cannot scope out of a conditional antecedent.

- (311) Kot ko-tun Nansi **kaaneet-iin-da-yan**, ko-paipai-iit-u.
 if 3.SBJV-marry N. teacher-SG-TH-YAN 3.SBJV-happy-VBLZ-IPFV
 ‘If Nancy marries a teacher, she’ll be happy.’
- | | |
|--|---------|
| a. ✓ Nancy wants to marry any teacher. | if > ∃ |
| b. # Nancy wants to marry a particular, unknown teacher. | *∃ > if |

Following Dawson (2020), I take this behavior to speak against a choice functional analysis and therefore in favor of ordinary quantification over individuals. Second, *-yan* places no restrictions

3.7.2 Pragmatic competition

As proposed in §3.6.2 and §3.7.1, nouns with a secondary suffix in Kipsigis are necessarily specific and referential, while *-yan* forms involve existential quantification over individuals. In this way, the *-yan* form is less informative than the noun with a secondary suffix in contexts where both are licensed; while the noun with a secondary suffix refers to a certain individual that is somehow identifiable to the speaker, the *-yan* form simply asserts that there is some individual with the properties denoted by *-yan*'s restrictor and scope. This state of affairs sets the stage for a conversational implicature to arise—in particular, a quantity implicature.

I suggest that a speaker's choice to existentially quantify—rather than to refer—implicates that the speaker is unable to identify the noun's referent in a way that is sufficiently relevant to their conversational goals. More technically, this means that the speaker does not know which value of the choice function variable both picks out the correct individual from the set denoted by its restrictor and is a relevant means of identifying this individual, given the speaker's goals. Use of the existential quantifier, then, implicates speaker ignorance—both first order and higher order ignorance. In this section, I walk through this pragmatic reasoning in more detail to show how this kind of competition gives rise to the precise type of ignorance effects illustrated in §3.3.

Consider again the sentences in (314). When the speaker says (314a) with the *-yan* form, they could also say (314b) with the noun with a secondary suffix. The version of the sentence with the *-yan* form expresses that there is some individual x such that x is a room and Chepkoech is hiding in x . By contrast, the version of the sentence in (314b) expresses that there is a choice function whose value is known to the speaker and that Chepkoech is hiding in the room picked out by that choice function.

(314) Kibet and Chepkoech are playing hide-and-seek—Kibet is the seeker and Chepkoech is hiding. Kibet says to his babysitter:

- a. Ø-U_{ny}-e-kεε Chepkɔɛch εεn rɔɔm-I-yan.
 3-hide-IPFV-REFL C. in room-TH-YAN
 'Chepkoech is hiding in some room.'
 $\exists x[\text{room}(x) \ \& \ \text{hide}(x)(C)]$
- b. Ø-U_{ny}-e-kεε Chepkɔɛch εεn rɔɔm-I-it.
 3-hide-IPFV-REFL C. in room-TH-SEC
 'Chepkoech is hiding in a certain room.'
 $[\text{CH}(f) \ \& \ \text{hide}(f(\text{room}))(C)]$

In this way, the sentence in (314b) makes a stronger claim than the one in (314a): it refers to a particular room, rather than simply existentially quantifying over rooms. In a context like the one in (314)—where both noun types are licensed—a speaker's choice to use a less informative *-yan* form over a more informative noun with a secondary suffix implicates that the speaker does not know how to resolve the choice function variable introduced by the secondary suffix; in other words, the speaker does not know which property will uniquely select the appropriate individual from the set that the choice function ranges over. Thus, use of a *-yan* form over a noun with a secondary suffix generates ignorance effects as a quantity implicature.

To see this analysis in action, consider the following examples of first order (315a) and higher order ignorance (315b)—repeated from (269).

(315) Kibet and Chepkoech are playing hide-and-seek—Kibet is the seeker and Chepkoech is hiding. Kibet says to his babysitter:

∅-Uny-e-kεε Chεεpkœech εεn **room-I-yan**.

3-hide-IPFV-REFL C. in ROOM-TH-YAN

‘Chepkoech is hiding in some room.’

(context adapted from Alonso-Ovalle & Menéndez-Benito 2010)

- a. ✓ Kibet knows that Chepkoech is in the house, but he doesn’t know which room she’s in, so he can’t find her.
- b. ✓ Kibet knows that Chepkoech is in the living room, but he doesn’t know where in the house the living room is, so he can’t find her.

Deriving first order ignorance is straightforward given the current semantic set-up. In a context like (315a), the speaker is ignorant about the identity of the room where Chepkoech is hiding; for instance, they know neither the name nor the location of this room. As a result, the speaker is unable to provide any value for the choice function variable that would uniquely select the appropriate room from the set of all rooms. This inability licenses use of the less informative *-yan* form, and ignorance effects arise pragmatically as a result.

Deriving higher order ignorance is slightly more complicated and requires enriching the semantics in §3.6.2. In a context like (315b), the speaker *is* able to identify a value for the choice function variable that uniquely selects the appropriate room from the set of all rooms; they know that the *living-room* property can be plugged in for the choice function variable and return the appropriate room. In this way, the speaker should be justified in using a noun with a secondary suffix here—contrary to fact. Crucially, though, in the context in (315b), the speaker is ignorant about the location of this room, which is essential to achieving their goal of finding Chepkoech. As a result, I suggest that not just *any* property that uniquely selects an individual from the set denoted by its restrictor is a viable value for the choice function variable introduced by the secondary suffix; instead, this property must enable the speaker to identify the unique referent in a way that is sufficiently relevant to their goals, as stated in (316), which is an updated version of (305).

(316) $\llbracket \text{-SEC} \rrbracket = \lambda P_{\langle e,t \rangle}.f(P)$, where f is a Choice Function that enables the speaker to identify the referent in a way that is relevant to their goals

In the hide-and-seek context in (315b), this means that, in order to refer, the speaker must be able to identify a choice function that picks out one particular room *and* that helps them physically locate the person who is hiding. In this way, the property that the speaker plugs in for the choice function variable must enable them to achieve their goals.

With this complication in place, the derivation of higher order ignorance proceeds parallel to that of first order ignorance. In a context like (315b), the speaker is ignorant about the location of the room where Chepkoech is hiding, even though they know the name of this room. While

identifying a room by name is essential in some contexts, in a game of hide-and-seek, knowing how to locate a room is far more important. As a result, the speaker is unable to provide a sufficiently relevant value for the choice function variable that would uniquely select the appropriate room from the set of all rooms. As previously, this inability licenses use of the less informative *-yan* form and leads to the generation of ignorance effects.

On the current analysis, then, first order and higher order ignorance are derived in nearly identical ways. First order ignorance arises when the speaker uses a *-yan* form because they are not able to supply *any* choice function that uniquely picks out the correct individual from the set that the choice function ranges over. Higher order ignorance, on the other hand, arises when the speaker uses a *-yan* form because they are not able to supply a choice function that uniquely picks out the correct individual in a way that is *sufficiently relevant* to their goals.

To conclude, it is worth highlighting that the ignorance effects associated with Kipsigis *-yan* forms arise not because *-yan* contributes any special semantics; instead, ignorance effects are the result of a speaker's choice *not* to use an indefinite that signals some degree of speaker knowledge (i.e. a noun with a secondary suffix). This situation parallels the one in Russian following Geist & Onea (2007). On their account, Russian *to* indefinites convey ignorance much like the kind found with Kipsigis *-yan* forms (see Kagan 2011 for thorough empirical description). Ignorance effects arise because *to* indefinites are in competition with Russian *koe* indefinites, which convey speaker knowledge. While the semantic details of their account differ from mine, the central intuition that ignorance effects can arise when semantically basic forms are in competition with those that convey some degree of speaker knowledge is exactly what I propose here. With this cross-linguistic observation in mind, the following section discusses aspects of the typology of epistemic indefinites across languages and key parameters of variation that contribute to their interpretation.

3.8 Existential quantification, choice functions, and cross-linguistic variation

As mentioned at the start of this chapter, epistemic indefinites across languages differ along several empirical dimensions. In this chapter, I have focused primarily on two properties of epistemic indefinites: 1) the type of ignorance effects that they trigger and 2) their scopal behaviors. In light of the attested cross-linguistic variation along these dimensions, it is reasonable to wonder whether these empirical properties correlate with each other; in fact, Dawson (2020) makes exactly this point, suggesting that first order ignorance correlates with scopal flexibility of the epistemic indefinite, while higher order ignorance correlates with exceptional wide scope. However, the novel Kipsigis pattern reported here speaks against this correlation, since Kipsigis *-yan* forms are scopally flexible but can convey higher order ignorance.

This decoupling of the type of ignorance effect and the scopal behavior of the epistemic indefinite raises the possibility of a new typology, as outlined in Table 3.3. In this system, all combinations of ignorance type and scopal behavior are, in principle, possible. Spanish *algún* conveys

purely first order ignorance and displays scopal flexibility, while Tiwa *-khí* represents the other end of the spectrum, conveying higher order ignorance and taking exceptional wide scope.¹¹ Kipsigis *-yan* fills in the lower left cell in the typology, conveying higher order ignorance while being scopally flexible.

	Scopal flexibility	Exceptional wide scope
(Purely) First order ignorance	Spanish <i>algún</i>	???
Higher order ignorance	Kipsigis <i>-yan</i>	Tiwa <i>-khí</i>

Table 3.3: Typology of epistemic indefinites and their properties

However, the remaining cell in the typology, which represents an epistemic indefinite that conveys purely first order ignorance but takes exceptional wide scope, raises an interesting puzzle. On the one hand, it is an empirical question whether there are indefinites that show this type of scope behavior and that trigger only first order ignorance. To my knowledge, no such epistemic indefinite has been reported in the literature. Yet on the other hand, it is a theoretical question whether such a system could exist.

To test this idea, let's consider the option space. In order to display exceptional wide scope and trigger ignorance effects, I assume that an epistemic indefinite in the upper right cell of the typology in Table 3.3 would introduce an existentially closed choice function variable. The choice functional nature of this indefinite would account for its scopal behavior, while the existential closure of this choice function variable would lead to ignorance effects rather than knowledge effects of the sort seen in Kipsigis. If this existentially closed choice functional indefinite were in competition with a plain indefinite (i.e. a basic existential quantifier), one would expect to see higher order ignorance with this form. This is precisely the pattern seen with Tiwa *-khí*. If this existentially closed choice functional indefinite were in competition with a free choice functional indefinite, one would, once again, expect to see higher order ignorance with this form. This is because, just like in Kipsigis, the free choice functional indefinite would signal some degree of speaker knowledge about how the referent is to be selected from a set; as a result, use of the existentially closed choice functional indefinite would implicate that the speaker lacks this knowledge about how the witness is to be identified.

However, it is possible to complicate the denotation of the existentially closed choice functional indefinite, for instance by adding some kind of domain restriction. If an anti-singleton constraint of the sort proposed in Alonso-Ovalle & Menéndez-Benito (2010) were added to an

¹¹The classification of Spanish *algún* as an epistemic indefinite that conveys purely first order ignorance is, perhaps, questionable; while *algún* is often used when the individual who witnesses the existential claim is unknown, it is possible to use *algún* when the speaker can identify this individual but not via a stable property (see Alonso-Ovalle & Menéndez-Benito 2013). However, the epistemic effects triggered by *algún* are still markedly different from those seen with Kipsigis *-yan* and Tiwa *-khí*; for instance, *algún* is not appropriate in the type of hide-and-seek context in (269b). This distinction motivates the classification in Table 3.3.

existentially closed choice functional indefinite, as with the hypothetical indefinite in (317), what kind of ignorance effects would be predicted?

(317) $\llbracket \text{indefinite}' \rrbracket = \lambda f_{\langle et, et \rangle} . \lambda P_{\langle e, t \rangle} : \text{anti-singleton}(f) . f(P)$, where f is a Choice Function

As above, all instances of competition involving this hypothetical indefinite would lead to higher order ignorance effects. The anti-singleton constraint in (317) simply requires the set that the choice function ranges over to contain more than one individual; however, the pragmatic calculations that actually derive ignorance effects remain the same. Use of this hypothetical indefinite instead of a plain indefinite implicates ignorance about how the witness is to be selected (from a necessarily non-singleton set), as with Tiwa *-khi*. Likewise, use of this hypothetical indefinite instead of a free choice functional indefinite implicates that the speaker lacks knowledge about how the referent is to be selected from a set. While there are presumably other ways to modify the denotation of a choice functional indefinite, the discussion here suggests that existing analytical tools in the literature on epistemic indefinites cannot derive purely first order ignorance with these forms. In this way, these observations lead to a reframing of the implication in Dawson (2020), which proposes that if an epistemic indefinite triggers higher order ignorance, then it must show exceptional wide scope. Instead, I suggest that if an epistemic indefinite displays exceptional wide scope, pragmatically derived ignorance must be higher order.

The discussion in the previous paragraphs also highlights another important point of cross-linguistic variation: whether or not there is existential closure of the choice function variable introduced by a choice functional indefinite. Side-by-side comparison of the Kipsigis and Tiwa systems makes this point particularly salient. In Tiwa, the epistemic indefinite introduces an existentially closed choice function variable, while the competitor is a plain indefinite. In Kipsigis, on the other hand, the epistemic indefinite is a plain indefinite (i.e. existential quantifier), while the competitor introduces a choice function variable that remains free. Table 3.4 summarizes these systems. Note also that the Kipsigis system is somewhat unique cross-linguistically in that the epistemic indefinite is not associated with any kind of special semantics: the heavy lifting in terms of deriving ignorance effects is done by the competitor, which signals some degree of speaker knowledge. While this type of system has been proposed for Russian *to* and *koe* indefinites (Geist & Onea 2007), most analyses of epistemic indefinites cross-linguistically propose an enriched semantics for the epistemic indefinite rather than the competitor.

	Epistemic indefinite	Competitor
Tiwa	\exists -closed CF variable	\exists
Kipsigis	\exists	free CF variable

Table 3.4: Comparison of the Tiwa and Kipsigis systems

In this way, the Tiwa and Kipsigis systems differ semantically only in the presence vs. absence of existential closure of the choice function variable. However, this difference generates two inverse systems. In light of this comparison, I suggest that the Kipsigis system provides further

support for the idea in Dawson (2020) that existential closure is essential for deriving ignorance effects with choice functional indefinites. The Kipsigis data in this chapter show that, in the absence of existential closure, choice functional indefinites in competition with plain indefinites do not trigger ignorance effects; in fact, the opposite happens. In this way, whether or not a choice functional indefinite triggers ignorance effects might offer an additional window into the presence or absence of existential closure. While existential closure of the choice function variable is independently diagnosable (see e.g. the discussion surrounding (307) in §3.7.1; but cf. Kratzer 2003), it is not always straightforward (Matthewson 1999), and the current cross-linguistic picture suggests that epistemic effects—including both ignorance and knowledge effects—might provide another diagnostic for addressing this question.

3.9 Conclusion

This chapter documents and analyzes a novel case of higher order ignorance with Kipsigis epistemic indefinites. While higher order ignorance is less widely discussed in the literature on epistemic indefinites than first order ignorance, the Kipsigis facts add to a growing body of work showing that such effects are attested cross-linguistically and can co-occur with a diverse array of other behaviors (e.g. domain size and scope).

More specifically, I show that Kipsigis *-yan* forms can be used to express ignorance about the individual who witnesses an existential claim (first order ignorance) or ignorance about the contextually salient properties of this individual witness (higher order ignorance). Dawson (2020) links these different types of ignorance effects to different scopal behaviors of the indefinites and, by extension, to different semantic analyses of them; in particular, she ties first order ignorance to scopal flexibility and domain widening semantics, and higher order ignorance to exceptional wide scope and choice functional semantics. However, as detailed in §3.8, the Kipsigis pattern challenges this correlation: while *-yan* forms are compatible with higher order ignorance, they display constrained scopal flexibility that is not straightforwardly predicted on any account of choice functional indefinites.

In light of this, I argue that Kipsigis *-yan* expresses basic existential quantification and only triggers ignorance effects as a result of competition with Kipsigis nouns with a secondary suffix, which introduce free choice function variables that are resolved via a contextually supplied assignment function. These nouns are referential and suggest that the speaker is able to identify the noun's referent in a contextually relevant way. When a speaker chooses to existentially quantify with a *-yan* form—instead of referring with a noun with a secondary suffix—it implicates that the speaker is unable to identify the noun's referent in this contextually relevant way. In cases of first order ignorance, the speaker is unable to provide any value for the choice function variable, while in cases of higher order ignorance, the speaker is unable to provide a contextually relevant value, even if they are able to characterize the referent in another way. This analysis, then, captures the full range of *-yan*'s behavior—from the content of its epistemic effects to its scope.

Chapter 4

Negative bias, reminding, and pragmatic reasoning in Kipsigis belief reports

4.1 Introduction

Semantically, belief reports of the form $x V_{att} p$ report the internal state of the belief holder. However, these expressions also serve a range of pragmatic functions beyond their semantic meaning and so offer a glimpse into the ways that speakers and listeners mutually construct the Common Ground (CG, i.e. the set of propositions that interlocutors agree to treat as true for the sake of a conversation; Stalnaker 1974; Karttunen 1974; Lewis 1979) and achieve their conversational goals. In fact, recent work on child language acquisition highlights the salience of the pragmatic uses of belief reports, showing that children take advantage of the pragmatically enriched meaning of these expressions when acquiring the literal semantic meaning of attitude verbs (Hacquard & Lidz 2018).

Given the importance of the pragmatic uses of attitude reports, this chapter explores one such function—namely, how belief reports contribute to CG management (i.e. how the content of the CG should develop over the course of a conversation; Krifka 2008). Much work has explored how belief reports serve as devices for indicating the status of the reported belief with respect to the CG. Work in this vein has considered factive verbs, which presuppose p (e.g. Kiparsky & Kiparsky 1970; Stalnaker 1974) and, more recently, negatively biased belief verbs, which suggest that p is false or unlikely and so should not be present in the input or output CG (see e.g. Tagalog *akala*, Kierstead 2013; Taiwanese Southern Min *lih-tsun*, Hsiao 2017; Spanish *creerse*, Anvari et al. 2019; Mandarin *yǐwéi*, Glass 2022).

This chapter adds to the growing body of work on biased belief reports by describing and analyzing the biased belief verb *par* in Kipsigis, which has two seemingly contradictory functions. Like negatively biased belief verbs, it can be used to suggest that the reported belief is false or unlikely; in (318b), the use of *par* instead of the neutral belief verb *pwaat* (318a) strongly implies that the speaker is not actually sick.¹

¹Third person indicative subject marking is either null or *i-* as determined by the conjugation class of the verb (i.e.

- (318) a. I-**pwaat**-e kaameε-nyʊʊn kole αα-mnyon-i.
 3-think-IPFV mother-1SG C 1SG-be.sick-IPFV
 ‘My mother **thinks** that I’m sick.’
- b. ∅-**Par**-e kaameε-nyʊʊn αα-mnyon-i.
 3-think-IPFV mother-1SG.POSS 1SG-be.sick-IPFV
 ‘My mother **is under the impression** that I’m sick.’

However, unlike any other currently documented negatively biased belief verb, *par* serves a different function in the present tense with a first person belief holder. Here, *par* has a reminding function: the speaker uses *par* in (319) to indicate that their interlocutor should already know the reported proposition.

- (319) I arrive home and see a guest. I don’t know who the guest is, so I ask my mother who they are. She replies:
 A-**par**-e abuleyaanit.
 1SG-think-IPFV uncle
 ‘**Remember**, it’s your uncle.’ (Lit: ‘I think that it’s your uncle.’)

While this type of reminding function is found with other CG management elements (e.g. the German discourse particles *ja* and *doch*), it has not, to my knowledge been reported for any negatively biased belief verbs. In this way, Kipsigis *par* is typologically unique in that it combines two familiar, yet seemingly contradictory, phenomena.

Against this backdrop, this chapter offers a unified semantic and pragmatic analysis that accounts for both of *par*’s negative bias and reminding functions. I claim that *par* contributes, in addition to its basic belief semantics, an instruction for CG management: *p* must not be *added* to the CG, though the CG may already contain *p*. In formalizing this instruction, I propose a definedness condition that relies on existing pre- and postsuppositional machinery and directly incorporates insights from analyses of both negatively biased belief verbs and reminding discourse particles. Together with context-sensitive pragmatic reasoning, this CG management instruction explains the case of a verb that can be used both to suggest that *p* is false (318b) and to remind the addressee that *p* is true (319).

This analysis of *par* contributes to the growing body of work on negatively biased belief verbs and offers a new perspective on the modeling of different types of CG management. More specifically, the Kipsigis data show that lexical items must be able to impose more complex requirements on the CG, beyond simply checking for membership of $(\neg)p$ in the CG *either* before or after utterance, as is standard in analyses of CG management (e.g. Repp 2013; Grosz 2016; Anvari et al. 2019; Glass 2022). Constructions with Kipsigis *par* are felicitous in a wide range of contexts, including when the input and output CG contain $\neg p$, say nothing with respect to *p*, or contain *p*. In this way, reference to membership of a particular proposition in one of either

Class I or Class II). Many verbs alternate between the two classes with a semantic effect—specifically, anticausative vs. causative interpretation—while many others appear exclusively in one class with no semantic generalization predicting which class a verb is in. In my data, *pwaat* is always in Class II, while *par* is always in Class I. See Chapter 2 §2.4, as well as Towett (1979) and Creider & Creider (1989) for more discussion of Kalenjin verb classes.

the input or output CG is insufficient to capture the full range of effects seen with *par*; instead, a more complex definedness condition that makes reference to both CG_{in} and CG_{out} is required to model the intuition that *par* requires *p* to not be added to the CG. In this way, my analysis of Kipsigis *par* lends further support to the relatively recent idea of postsupposition alongside presupposition (Lauer 2009; Brasoveanu 2013; Glass 2022).

The chapter is structured as follows. In §4.2, I describe the data, with a focus on *par*'s different interpretive effects with non-first person vs. first person belief holders. §4.3 situates the Kipsigis pattern in the typology of CG management elements, including negatively biased belief verbs and reminding particles. In §4.4, I provide an analysis of *par* that spans the semantics-pragmatics interface and attributes its interpretive effects to a CG management instruction that *par* provides: *p* must not be added to the CG. §4.5 concludes.

4.2 The Kipsigis pattern

The Kipsigis verbs *par* and *pwaat* both mean ‘think’, though use of *par* triggers additional inferences about *p*'s status in relation to the CG. Because *par* is a belief verb, its effect on the discourse depends on who the belief holder is—in particular, whether they are the speaker or someone else. Intuitively, this difference arises because individuals can stand in different relationships with their own beliefs vs. those of other people. For this reason, I describe sentences with non-first person belief holders in §4.2.1, then consider sentences with first person belief holders in §4.2.2.

4.2.1 Interpretation with non-first person belief holders

With non-first person belief holders, use of *par* indicates that the speaker knows the reported belief to be false, is biased against the reported belief whether or not its truth is known, or finds the belief holder unreliable. First, if the speaker knows that *p* is false, statements with *pwaat* are strongly dispreferred in comparison to their counterparts with *par*, and statements with factive *ngen* ‘know’ are infelicitous, as illustrated in (320) - (321).² I use #? to indicate strong dispreference—rather than infelicity (#)—as speakers can choose to use *pwaat* if they wish to appear neutral, regardless of whether they actually *are* neutral with respect to *p*.³

(320) I'm healthy, but my mother thinks that I'm sick because I fooled her to skip school. I say:

²To collect the judgements reported here, I first asked Kipsigis speakers if a particular sentence was appropriate in a given context. Then, once I had identified the viable alternatives in this way, I presented speakers with the same context and all viable alternatives to identify their (dis)preferences. This methodology enabled me to make a three-way distinction between felicitous, dispreferred, and infelicitous.

³As seen in (320) - (321), *pwaat* and *par* involve different syntactic complementation strategies: *pwaat* must co-occur with *kole*, while *par* cannot. §4.4.4 discusses this pattern, where I address the analysis of Kipsigis complementation in Driemel & Kouneli (2022a) and show that the semantics of *kole* cannot be responsible for the interpretive differences between *pwaat* and *par*.

- a. #? I-**pwaat**-e kaameε-nyʊʊn kole aa-mnyon-i.
 3-think-IPFV mother-1SG.POSS C 1SG-be.sick-IPFV
 ‘My mother **thinks** that I’m sick.’
- b. ∅-**Par**-e kaameε-nyʊʊn aa-mnyon-i.
 3-think-IPFV mother-1SG.POSS 1SG-be.sick-IPFV
 ‘My mother **is under the impression** that I’m sick.’
- c. # I-**ngen** kaameε-nyʊʊn kole aa-mnyon-i.
 3-know mother-1SG.POSS C 1SG-be.sick-IPFV
 ‘My mother **knows** that I’m sick.’
 (context adapted from Glass 2022:ex. 4)

(321) Cheptoo’s parents think that she’s a very good child who doesn’t drink or go to parties. But they’re wrong—she actually does do these things! I say:

- a. #? I-**pwaat**-e siikiik-aap Cheepta kole ma-∅-ye maiweek.
 3-think-IPFV parents-of C. C NEG-3-drink alcohol
 ‘Cheptoo’s parents **think** that she doesn’t drink alcohol.’
- b. ∅-**Par**-e siikiik-aap Cheepta ma-∅-ye maiweek.
 3-think-IPFV parents-of C. NEG-3-drink alcohol
 ‘Cheptoo’s parents **are under the impression** that she doesn’t drink alcohol.’
- c. # I-**ngen** siikiik-aap Cheepta kole ma-∅-ye maiweek.
 3-know parents-of C. C NEG-3-drink alcohol
 ‘Cheptoo’s parents **know** that she doesn’t drink alcohol.’

The contexts in (320) - (321) establish that the speaker knows the reported beliefs to be false. Consequently, the sentences with *par* in the b. examples are preferred, since they highlight that the reported beliefs are incorrect. Consultants report that the a. examples are sensible but less appropriate in these contexts because they “give the impression that [the belief holder] could be right or wrong”, even though it is already known that they are mistaken. Likewise, belief reports with factive *ngen* ‘know’ are incoherent, since they would require *p* to be true, as seen in the c. examples.

Second, use of *par* indicates that the speaker is biased against *p*, whether or not the truth of the reported belief is known. In (322) - (323), A does not know if Arap Ruto has arrived—evidenced by their initial question. B responds with information about Arap Bett’s belief state, using either *pwaat* or *par*. Only in the case of a belief report with *pwaat* (322) is it appropriate for A to assume that Arap Bett is correct and that Arap Ruto has arrived; this assumption licenses A’s response in (322). When B reports Arap Bett’s belief using *par* (323), it is inappropriate for A to accept this belief as true.

- (322) A: Ka-ko-it Arap Ruto-i?
 PST1-3.PERF-arrive son.of R.-Q
 ‘Has Arap Ruto arrived?’

- B: I-**pwaat**-e Arap Beet kole ka-ko-it.
 3-think-IPFV son.of B. C PST1-3.PERF-arrive
 ‘Arap Bett **thinks** that he has.’
- A: Nen aa-wa ki-kat-kεε any.
 then 1SG-go 1PL-greet-REFL now
 ‘Then I’ll go greet him now.’ (Lit: ‘Then I’ll go, we’ll greet each other now.’)
 (context adapted from Glass 2022:Appendix ex. 13)
- (323) A: Ka-ko-it Arap Ruto-i?
 PST1-3.PERF-arrive son.of R.-Q
 ‘Has Arap Ruto arrived?’
- B: ∅-**Par**-e Arap Beet ka-ko-it.
 3-think-IPFV son.of B. PST1-3.PERF-arrive
 ‘Arap Bett **is under the impression** that he has.’
- A: # Nen aa-wa ki-kat-kεε any.
 then 1SG-go 1PL-greet-REFL now
 ‘Then I’ll go greet him now.’ (Lit: ‘Then I’ll go, we’ll greet each other now.’)

The difference between these cases of speaker bias and those in which the reported belief is false is particularly clear in future-oriented sentences, where the truth of *p* cannot be known in the present moment. In (324), for example, a journalist is reporting a political candidate’s belief about the upcoming election.

- (324) We turn on a Kalenjin TV station and see an impartial journalist of unknown political affiliation discussing the upcoming election. Talking about a viable candidate named Jessica, the journalist says:
- a. I-**pwaat**-e Jεsika kole ∅-sindan-iisyε εεn lewenisiet.
 3-think-AP.IPFV J. C 3-win-IPFV in election
 ‘Jessica **thinks** that she will win the election.’
- b. # ∅-**Par**-e Jεsika ∅-sindan-iisyε εεn lewenisiet.
 3-think-AP.IPFV J. 3-win-IPFV in election
 ‘Jessica **is under the impression** that she will win the election.’
 (context adapted from Glass 2022:ex. 5)

(324b) is inappropriate for an impartial journalist to say because it suggests that the journalist doubts that Jessica will win the election. Speakers comment that this sentence comes across as though the journalist is mocking the candidate who wrongly believes that she will win. However, by introducing explicit bias into the context, speakers’ felicity judgements flip. If the journalist is instead a biased political pundit as in (325), sentences with *par* like (325b) are ideal.

- (325) We turn on a Kalenjin TV station and see a biased political pundit discussing the upcoming election. Talking about a viable candidate named Jessica who he believes will lose the election, the pundit says:

- a. #? I-**pwaat**-e Jɛsɪka kole Ø-sindan-iisye ɛɛn lewenisiet.
 3-think-AP.IPFV J. C 3-win-IPFV in election
 ‘Jessica **thinks** that she will win the election.’
- b. Ø-**Par**-e Jɛsɪka Ø-sindan-iisye ɛɛn lewenisiet.
 3-think-IPFV J. 3-win-AP.IPFV in election
 ‘Jessica **is under the impression** that she will win the election.’
 (context adapted from Glass 2022:ex. 5)

An identical pattern of speaker bias is observed with second person belief holders. An impartial journalist cannot felicitously report their addressee’s future-oriented belief using *par*, as evidenced by the infelicity of (326b).

(326) During an interview, a presidential candidate says *Atinye komong’ unet ne oo kole kipelisieɪ ɛɛn lewenisiet* ‘I have a lot of faith that we will win the election.’ An impartial journalist replies:

- a. Ii-**pwaat**-e kole i-pel-iisye-i ɛɛn lewenisiet.
 2SG-think-AP.IPFV-LP C 2SG-win-IPFV-LP in election
 ‘(So) you **think** that you’ll win the election.’
- b. # I-**par**-e i-pel-iisye-i ɛɛn lewenisiet.
 2SG-think-AP.IPFV-LP 2SG-win-IPFV-LP in election
 ‘(So) you **are under the impression** that you’ll win the election.’
 (context adapted from Glass 2022:ex. 5)

The parallelism between (324) with a third person belief holder and (326) with a second person belief holder shows that this speaker bias is not tied to a third person belief holder. Instead, this interpretive effect arises with any non-first person belief holder. This also includes plural non-first person belief holders, as seen in (327) with the third person plural belief holder *piik* ‘people’ and in (328) with a second person plural belief holder.

- (327) Ø-**Par**-e piik leeluwaat ng’ook-i.
 3-think-IPFV people fox dog-PROX
 ‘People **are under the impression** that this dog is a fox.’
- (328) You’re definitely not coming home tomorrow, but your family seems to think that you are. You say:
 Toos o-**par**-e α-nyoon-e kaa kaaroon-i?
 MOD 2PL-think-IPFV 1SG-come-IPFV home tomorrow-Q
 ‘Are you (pl) **under the impression** that I’m coming home tomorrow?’

Third, speakers use *par* to indicate that they find the belief holder unreliable, even if they do not know the truth of the reported belief themselves. For instance, in an out-of-the-blue context where the speaker has no reason to doubt the belief holder’s reliability, statements with *pwaat* are perfectly natural (329a), while those with *par* are infelicitous (329b).

(329) We walk up to some people at a party and hear them talking about who has and hasn't arrived. We have no idea if Arap Ruto is here, nor any idea why Arap Bett has the beliefs that he does. I say:

- a. Maa-ngen koto ka-ko-it Arap Ruto anan tomo lakini i-**pwaat**-e
 NEG.1SG-know if PST1-3.PERF-arrive son.of R. or not.yet but 3-think-IPFV
 Arap Beet kole ka-ko-it.
 son.of B. C PST1-3.PERF-arrive
 'I don't know if Arap Ruto has arrived yet, but Arap Bett **thinks** that he has.'
- b. # Maa-ngen koto ka-ko-it Arap Ruto anan tomo lakini \emptyset -**par**-e
 NEG.1SG-know if PST1-3.PERF-arrive son.of R. or not.yet but 3-think-IPFV
 Arap Beet ka-ko-it.
 son.of B. PST1-3.PERF-arrive
 'I don't know if Arap Ruto has arrived yet, but Arap Bett **is under the impression**
 that he has.'
 (context adapted from Glass 2022:Appendix ex. 12)

Consultants comment that (329b) is inappropriate because the speaker has no information that would allow them to judge Bett's belief, as use of *par* suggests. However, when the context establishes that the belief holder is unreliable in some way as in (330), *par* becomes natural.

(330) Arap Bett is drunk and is acting confused. I have no idea if Arap Ruto is here or not, but I have reason to doubt Arap Bett's reliability. I say:

- Maa-ngen koto ka-ko-it Arap Ruto anan tomo lakini \emptyset -**par**-e Arap
 NEG.1SG-know if PST1-3.PERF-arrive son.of R. or not.yet but 3-think-IPFV son.of
 Beet ka-ko-it.
 B. PST1-3.PERF-arrive
 'I don't know if Arap Ruto has arrived yet, but Arap Bett **is under the impression** that he
 has.'
 (context adapted from Glass 2022:Appendix ex. 12)

In this way, even when the sentence explicitly states that the speaker does not know whether or not *p* is true, *par* statements are appropriate when the belief holder is deemed unreliable.

4.2.2 Interpretation with first person belief holders

With first person belief holders, *par* has different interpretive effects depending on the tense of the belief verb. Intuitively, this distinction arises because individuals can stand in different relationships with their past beliefs vs. their present ones. People typically assume that their present beliefs are true—otherwise, they would not believe them. However, it is entirely possible for individuals to recognize their past beliefs as false, if their epistemic state has changed in some relevant way. In the past tense, then, speakers use *par* with a first person belief holder when they thought that the reported belief was true, but have since learned that it is false (331) - (332).

- (331) I went to school because I thought there was a meeting, but it was actually cancelled. When I get home, my mom asks why I went to school, so I respond:
- a. #? Kaa-**pwaat**-e kole mii tuiyeet ra.
 PST1.1SG-think-IPFV C COP meeting today
 ‘I **thought** there was a meeting today.’
- b. Kaa-**par**-e mii tuiyeet ra.
 PST1.1SG-think-IPFV COP meeting today
 ‘I **was under the impression** that there was a meeting today.’
 (context adapted from Glass 2022:ex. 11)
- (332) I left my car with a friend while I was out of town because I thought they would use it. But when I get back, I learn that they actually don’t know how to drive! I say:
- a. #? Kaa-**pwaat**-e kole ii-ngen ii-ket karnt.
 PST1.1SG-think-IPFV C 2SG-know 2SG.SBJV-drive car
 ‘I **thought** you knew how to drive a car.’
- b. Kaa-**par**-e ii-ngen ii-ket karnt.
 PST1.1SG-think-IPFV 2SG-know 2SG.SBJV-drive car
 ‘I **was under the impression** that you knew how to drive a car.’

The pattern seen here parallels that in (320) - (321) with third person belief holders; *par* suggests that *p* is false, which is only possible in the past tense with first person belief holders.

Yet in the present tense, speakers use *par* with a first person belief holder for a very different purpose. In sentences like the a. examples in (333) - (335), *par* serves a reminding function. Speakers report that, in these types of sentences, *par* suggests that the addressee should already know *p* or, in the words of one consultant, that “[the reported belief] is already a foregone conclusion”. *Par* serves not to weaken the strength of these statements, but rather to highlight the fact that *p* should be common knowledge—for reasons ranging from shared cultural norms (333) to prior knowledge (334) to physical context (335).

- (333) Church meetings are always loud and take place every Sunday, which we both know. It’s Sunday morning, and we hear lots of noise. You ask me what it is, and I respond:
- a. A-**par**-e mii tuiyeet ra.
 1SG-think-IPFV COP meeting today
 ‘**Of course**, there’s a meeting today.’ (Lit: ‘I think that there’s a meeting today.’)
- b. # Aa-**pwaat**-e aa-le mii tuiyeet ra.
 1SG-think-IPFV 1SG.SBJV-C COP meeting today
 ‘I **think** there’s a meeting today.’
- c. #? Mii tuiyeet ra.
 COP meeting today
 ‘There’s a meeting today.’
- d. # Aa-**ngen** aa-le mii tuiyeet ra.
 1SG-know 1SG.SBJV-C COP meeting today
 ‘I **know** there’s a meeting today.’

- (334) Nick and Sharon live in Oakland, which I should know because I've been to their house before. When they invite me over for dinner, I ask them what city they live in, thinking that it's Berkeley. They want to remind me where they live by saying:
- a. A-**par**-e ki-meny-e Oakland.
1SG-think-IPFV 1PL-live-IPFV Oakland
'We live in Oakland, **as you know**.' (Lit: 'I think that we live in Oakland.')
 - b. # Aa-**pwaat**-e aa-le ki-meny-e Oakland.
1SG-think-IPFV 1SG.SBJV-C 1PL-live-IPFV Oakland
'I **think** we live in Oakland.'
 - c. #? Ki-meny-e Oakland.
1PL-live-IPFV Oakland
'We live in Oakland.'
 - d. # Aa-**ngen** aa-le ki-meny-e Oakland.
1SG-know 1SG.SBJV-C 1PL-live-IPFV Oakland
'I **know** we live in Oakland.'
- (335) I see a friend heading outside in a t-shirt and shorts, even though it's the middle of winter. I want to remind them that it's way too cold for that kind of outfit! I say:
- a. A-**par**-e kartit ra.
1SG-think-IPFV cold today
'**Hang on**, it's cold today.' (Lit: 'I think that it's cold today.')
 - b. # Aa-**pwaat**-e aa-le kartit ra.
1SG-think-IPFV 1SG.SBJV-C cold today
'I **think** it's cold today.'
 - c. #? Kartit ra.
cold today
'It's cold today.'
 - d. # Aa-**ngen** aa-le kartit ra.
1SG-know 1SG.SBJV-C cold today
'I **know** it's cold today.'

In uttering these statements, the speaker is not indicating doubt on their part about *p*; in each of these examples, the speaker is certain of the truth of the reported belief and justified in having this belief. For this reason, the equivalent statements with the neutral belief verb *pwaat* are infelicitous, as seen in the b. examples. Instead, speakers use the *par* statement or the bare proposition, which is possible though strong dispreferred, as it does not explicitly serve a reminding function. For instance, speakers suggest that the c. examples are more polite, since they do not highlight that the addressee is forgetting something, but are less suited to the context, which explicitly calls for reminding. Finally, while the equivalent attitude reports with *ngen* 'know' are more coherent in (333) - (335) than in the negative bias cases, even here speakers prefer the bare proposition *p* over the *ngen* statement—presumably because stating *p* commits the

speaker to knowing *p*, making the attitude report with *ngen* unnecessary (see Simons 2007:1048 for such an analysis of English *I know p*).⁴

Further evidence of *par*'s reminding function is found in (336), where the *par* statement is infelicitous when the speaker cannot assume the reported belief to be shared knowledge with their addressee. The sentence in (336a) with *par* is infelicitous because the addressee cannot be assumed to already know that they should speak to Arap Ruto; rather, getting advice from the advisor about what the addressee should do is the goal of the conversation.

(336) Arap Ruto is in charge of the linguistics department. I'm planning to go to Kenya and need permission to do so. I talk to my advisor, and she says:

- a. # A-**par**-e yaach-e ii-ng'olaal-chi Arap Ruto.
 1SG-think-IPFV MOD-IPFV 2SG.SBJV-speak-APPL son.of R.
 'As you know, you should speak to Arap Ruto.'
- b. Yaach-e ii-ng'olaal-chi Arap Ruto.
 MOD-IPFV 2SG.SBJV-speak-APPL son.of R.
 'You should speak to Arap Ruto.'

Consultants note that (336a) is odd for an advisor to use because it comes across as if they are reprimanding their addressee for not already speaking to Arap Ruto or not knowing to do so.

Notably, this reminding function is restricted to first person belief holders. Although the context in (337) sets the stage for a reminding use of *par*, this reading is unavailable with third person *par*. Speakers note that *par* in (337a) implies that the speaker believes the doctor to be wrong, which is inappropriate given their expertise. In this way, *par* has only a negative bias reading with a third person belief holder. To get at a reminding function, speakers embed the doctor's belief—reported with the neutral belief verb *pwaat*—under 1SG *par* (337b).

(337) You go to the doctor because you're coughing and have a sore throat. The doctor thinks that you have Covid, but didn't run a test because he ran out. He sends you home, but tells you to isolate and follow the Covid guidelines. I see you leaving the house and want to remind you about your diagnosis. I say:

- a. # Ø-**Par**-e dakitaari i-tiny-e koroona.
 3-think-IPFV doctor 2SG-have-IPFV Covid
 '**Remember**, according to the doctor, you have Covid.'
- b. A-**par**-e i-pwaat-e dakitaari kole i-tiny-e koroona.
 1SG-think-IPFV 3-think-IPFV doctor C 2SG-have-IPFV Covid
 '**Remember**, the doctor thinks that you have Covid.'

The same pattern is seen with second person belief holders. Use of 2SG *par* in a reminding scenario like (338) is infelicitous, as it implies that the speaker doubts the addressee's belief. Instead,

⁴This claim, of course, raises the question of why anyone would say *I know p*, when they could just say *p*. Following Simons (2007), I assume that *I know p* is preferred over *p* in some contexts given other discourse considerations like focus (e.g. on the attitude holder or *know*); this analysis aligns with the fact that the *ngen* statements in the d. examples in (333) - (335) are made significantly better when *ngen* 'know' receives stress.

the report of the addressee's belief must be embedded under 1SG *par* (338b). In this way, *par*'s reminding function is restricted to instances with a first person belief holder in the present tense.⁵

(338) This morning, you told me that you were feeling sick and weren't going to school today. But when I see you a bit later, you're getting dressed and putting things in your backpack, as if you're going to school. I want to remind you what you told me earlier, so I say:

- a. # I-**par**-e ii-mnyoon-i.
 2SG-think-IPFV 2SG-be.sick-IPFV
 'Wait a minute, according to you, you're sick.'
- b. A-**par**-e ii-pwaat-e kole ii-mnyoon-i.
 1SG-think-IPFV 2SG-think-IPFV C 2SG-be.sick-IPFV
 'Wait a minute, you think that you're sick.'

To summarize, the specific interpretive effects triggered by *par* depend on the person of the belief holder, as well as the tense of the belief verb, as outlined in (339). With non-first person belief holders, use of *par* indicates that *p* is false, the speaker is biased against *p*, or *x* is unreliable. With first person belief holders, *par*'s effects differ across tenses. In the past tense, *par* suggests that *p* is false, as with non-first person belief holders. Yet in the present tense, *par* serves to remind listeners that they should already know *p*. In this way, *par* serves two seemingly contradictory purposes: to suggest that *p* is or may be false and to remind the addressee that *p* is true.

- | | | |
|-------|--------------------------------------|--------------------------------------|
| (339) | a. Non-first person <i>par p</i> | b. First person past <i>par p</i> |
| | • <i>p</i> is false | • <i>p</i> is false |
| | • Speaker is biased against <i>p</i> | c. First person present <i>par p</i> |
| | • <i>x</i> is unreliable | • Reminds the addressee of <i>p</i> |

4.3 Negatively biased belief and reminding across languages

The negative bias and reminding functions seen with Kipsigis *par* are attested independently across a range of languages. However, *par* is unique in using a single lexical item to accomplish both of these seemingly distinct functions. In this section, I sketch the typology of negatively

⁵I have found speaker variation with respect to the availability of a reminding function with a first person plural belief holder. For one speaker, use of 1PL *par* to serve a reminding function in (1) is natural, since the speaker is responding on behalf of themselves and their partner, while another speaker disprefers (1).

- (1) Nick and Sharon live in Oakland, which I should know because I've been to their house before. When they invite me over for dinner, I ask them what city they live in, thinking that it's Berkeley. They want to remind me where they live by saying:
 Ki-**par**-e ki-meny-e Oakland.
 1PL-think-IPFV 1PL-live-IPFV Oakland
 'We live in Oakland, as you know.'

biased belief verbs, as well as a sample of elements with reminding functions, to show how the Kipsigis pattern fits into the larger, cross-linguistic picture.

Negatively biased belief verbs are reported in several languages, including Tagalog (*akala*; Kierstead 2013), Taiwanese Southern Min (*liah-tsun*; Hsiao 2017), Spanish (*creerse*; Anvari et al. 2019), and Mandarin (*yǐwéi*; Glass 2022). With non-first person belief holders, these verbs show similar behaviors to Kipsigis *par*. For instance, both Spanish *creerse* and Mandarin *yǐwéi* strongly suggest that the reported belief is false or unlikely. The Spanish sentence in (340) implies that it is not actually raining, while the Mandarin sentence in (341) is preferred when the speaker is known to be healthy (i.e. *p* is known to be false). These data parallel the pattern seen with non-first person Kipsigis *par* in §4.2.1.^{6,7}

- (340) *Spanish* (Anvari et al. 2019:ex. 1)
 Juan **se cree** que está lloviendo.
 J. REFL believes that it.is raining
 ‘Juan **incorrectly believes** that it is raining.’
- (341) *Mandarin* (Glass 2022:ex. 4)
 I tell you that I’ve been faking an illness to skip school. I say:
 Māma **yǐwéi** wǒ bìng le.
 mother YIWEI I sick ASP
 ‘Mother **is under the impression** that I’m sick.’

However, *par*’s reminding function with a first person belief holder is unique in the space of previously described negatively biased belief verbs. Other such verbs that have been described in sufficient detail in the literature have no reminding function and are often incompatible with a first person belief holder. For instance, Spanish *creerse* is infelicitous with a first person belief holder (342), while first person Mandarin *yǐwéi* is necessarily interpreted as past tense (343a) or—with enough contextual support—very hedged (343b). Crucially, these verbs lack any sort of reminding function, which differentiates the pattern in (342) - (343) from the Kipsigis one in §4.2.2.

⁶Other languages also have different ways of encoding a similar type of negative bias. For instance, German Konjunktiv I is a reportative subjunctive that serves a function similar to Kipsigis *par* in non-first person cases; the sentence in (1) suggests that *p* is false or unlikely. Here I highlight that negatively biased belief verbs are not the only way to derive negative bias in belief reports, but leave for future research the question of whether German Konjunktiv I warrants a similar, CG management-style analysis (see also Glass 2022:25).

- (1) *German*
 Er denkt, sie sei zuhause.
 he thinks she is.SBJV at.home
 ‘He thinks that she is at home.’

⁷‘Incorrectly’ in the translation of (340) is my addition to highlight the negative bias that arises with Spanish *creerse* in this context. This addition is based on the description around the data example in Anvari et al. (2019), but it is not included in their translation.

- (342) *Spanish* (Anvari et al. 2019:ex. 13)
 # **Me creo** que está lloviendo.
 REFL I.believe that it.is raining
 NOT: ‘Remember, it’s raining.’
- (343) *Mandarin* (Glass 2022:ex. 11-12)
 a. Wǒ yǐwéi jīntiān yǒu ge jiǎngzuò.
 I YIWEI today have CL talk
 ‘I **thought** there was a talk today.’
 NOT: ‘Remember, there’s a talk today.’
 b. Wǒ gèrén yǐwéi nǐ yīnggāi zhèyàng zuò.
 I personally YIWEI you should this.way do
 ‘Personally, I **would think** that you should do this.’
 NOT: ‘You should do this, **as you know**.’

Although it is previously unreported for a negatively biased belief verb, *par*’s reminding function is familiar from discourse particles like German *ja* and *doch*, which instruct the addressee to retrieve from the CG a proposition that is not currently being considered (Döring 2016). For instance, in sentences like the follow-up in (344b), use of *ja* instructs the addressee to retrieve a forgotten proposition from the CG—namely, that Maria has already said that she would like to do the job. Although the addressee has perhaps forgotten this fact, use of *ja* suggests that they should already know *p*.⁸

- (344) *German* (Döring 2016:ex. 39)
 a. Ich würde Maria als Sprecherin vorschlagen.
 I would M. as speaker recommend
 ‘I would recommend Maria as the speaker.’
 b. Sie hat **ja** gesagt, sie würde die Aufgabe gern übernehmen.
 she has JA said she would the task like to.take.over
 ‘**Remember**, she said she would like to take over this task.’
 where $p = \{w : \text{Maria said that she would like to take over this task in } w\}$

In this way, the two functions served by Kipsigis *par* are familiar on their own; a number of verbs—including Spanish *creerse* and Mandarin *yǐwéi*—accomplish only negative bias, while the German discourse particles *ja* and *doch* serve only a reminding function. The novelty of the Kipsigis pattern lies in its combination of these two phenomena in a single lexical item. This unique combination of familiar elements has ramifications for the analysis of Kipsigis *par*. As detailed in the following section, analyses of negatively biased belief verbs typically assume that the speaker is committed to *p* being false or unlikely, while analyses of reminding elements like *ja* and *doch* assume the opposite—that the speaker is committed to *p* being true. In short, existing accounts

⁸‘Remember’ in the translation of (344b) is my addition to highlight the reminding inference that arises with German *ja*. This addition is based on the description that Döring provides in the text surrounding the example, but it is not found in the original translation.

that capture either negative bias or reminding are unable to capture the other function. Against this backdrop, in the following section, I draw inspiration from these two bodies of literature to provide a unified analysis of Kipsigis *par* that accounts for its full range of effects.

4.4 Analyzing Kipsigis *par*

My analysis of Kipsigis *par* makes use of both semantic and pragmatic tools. Semantically, in addition to its standard belief semantics, *par* contains an instruction for CG management: *p* is not to be added to the CG. Pragmatically, then, speakers reason about why this must be the case, which implicates that the reported belief is false or unlikely, the belief holder is unreliable, or the reported belief is already in the CG. This bipartite analysis accounts for the full range of interpretive effects associated with *par*, while requiring minimal semantic differences between *par* and *pwaat*. Then, to conclude this section, I consider two other possible analyses, which seek to explain *par*'s negative bias and reminding functions through different mechanisms, and I show that these alternatives are unable to capture the full range of *par*'s effects.

4.4.1 Semantics of *par*

I assume a framework in which utterances are updates to the CG (Karttunen 1974; Stalnaker 1978). For instance, the asserted content of (345)—what the speaker directly proposes to add to the CG—is the proposition that the speaker's mother has a particular belief (i.e. *x thinks p*). Given assumptions about cooperation in discourse (Grice 1989), listeners typically assume that the speaker of (345) is being truthful and accept this proposition by default. This proposition enters the CG, leading to CG update of the form in (346); the output CG (CG_{out}) contains the input CG (CG_{in}) plus the asserted proposition.

(345) I-*pwaat*-e kaamεε-nyʊʊn kole aa-mnyon-i.
 3-think-IPFV mother-1SG.POSS C 1SG-be.sick-IPFV
 'My mother **thinks** that I'm sick.'

(346) CG_{out} = CG_{in} ∪ {*x thinks p*}

However, depending on how interlocutors evaluate the likelihood of the reported belief and the reliability of the belief holder, belief reports with neutral belief verbs like *pwaat* allow other propositions to enter the CG as well. For instance, when (345) is uttered in a context like (347a), *p* is reasonable and *x* is reliable, so *p* can be added to the CG alongside *x thinks p*. This type of CG update is schematized in (347b).⁹

⁹While the inference from *x thinks p* to *p* is not part of the asserted content of a belief report, listeners often take *x*'s belief in *p* as evidence for *p* itself—especially if *p* seems reasonable and *x* is viewed as reliable (see e.g. the evidential uses of attitude reports in Simons 2007). In this way, interlocutors' acceptance of *p* as true after utterance of a belief report relies on pragmatic reasoning about the belief itself and the belief holder, though this inference is not straightforwardly classified as presupposition or conversational implicature.

- (347) a. My mother is a doctor who is really good at spotting signs of illness before people start feeling sick. My mother took my temperature and felt my lymph nodes and thinks that I'm sick. I'm explaining to you why I can't come to your party today and say:
 b. $CG_{out} = CG_{in} \cup \{x \text{ thinks } p, p\}$
- (348) a. My mother is a hypochondriac who always thinks that I'm sick. I went to bed early last night because I was up late the night before, but now my mother is convinced that I'm sick. I'm complaining to you and say:
 b. $CG_{out} = CG_{in} \cup \{x \text{ thinks } p, \neg p\}$

On the other hand, when (345) is uttered in a context like (348a) where p is unlikely and x is unreliable, the belief report can serve to add $\neg p$ to the CG. In these cases, CG update takes the form in (348b).

In this way, neutral belief verbs like *pwaat* allow for a wide range of CG update possibilities. As a result, I suggest that *pwaat* contributes only standard belief semantics (349): in all the worlds compatible with x 's beliefs, p holds. The fact that *pwaat* allows for a wide range of CG update possibilities aligns with speaker intuitions that use of *pwaat* "leaves the issue open" as to whether or not the reported belief is true, and so general purpose reasoning steps in.

$$(349) \quad \llbracket pwaat \rrbracket = \llbracket think \rrbracket = \lambda p \lambda x \lambda w. \forall w' \in \text{Dox}_{x,w} [p(w') = 1]$$

By contrast, I argue that *par* statements are more restricted in the type of CG update that they allow. The meaning of *par* contains two parts. Its at-issue content provides the same standard belief semantics seen in (349). Yet its not-at-issue content provides a CG management instruction, stating that p must not be *added* to the CG. This instruction rules out the type of CG update seen with *pwaat* in (347b). I formalize this instruction as a definedness condition requiring that p either must not be in CG_{out} or must be in CG_{in} . A denotation for *par* is in (350), though other formalizations of both the at-issue and not-at-issue content are compatible with my analysis; the key takeaway is that *par* comes with a bipartite definedness condition that restricts the shape of the CG before *or* after utterance.¹⁰

$$(350) \quad \llbracket par \rrbracket = \llbracket think \rrbracket = \lambda p \lambda x \lambda w. \forall w' \in \text{Dox}_{x,w} [p(w') = 1]$$

defined only when

- a. $p \notin CG_{out}$, or
- b. $p \in CG_{in}$

(351) Paraphrase of (350a) - (350b): Do not add p to the CG.

This definedness condition ensures that p is not added to the CG in the following way. If p is not in CG_{in} , then it cannot be added to the CG upon acceptance of the *par* statement, since (350a) mandates that p must not be in CG_{out} . On the other hand, if p is already in CG_{in} , *par* continues to be defined (350b), though p cannot be non-trivially added to the CG, since it is already there. Because p is in the input CG, it must also be in the output CG—even without any contribution

¹⁰For the rest of the chapter, I use $\llbracket think \rrbracket$ as shorthand for $\lambda p \lambda x \lambda w. \forall w' \in \text{Dox}_{x,w} [p(w') = 1]$ or an equivalent.

from the *par* statement—given that CG_{out} is built additively upon CG_{in} .¹¹ In this way, the condition in (350) captures the intuition that p must not be added to the CG when a speaker makes a *par* statement, at least not in any non-trivial way. It is worth pointing out, though, that the split between negative bias and reminding seen in the data persists here in the analysis; although these effects are conceptually unified in the paraphrase in (351), the formal implementation in (350) involves a bipartite condition.

4.4.1.1 Evidence for not-at-issue content

Evidence that *par*'s definedness condition is not-at-issue content comes from its behavior in projection contexts; in particular, the bias seen with *par* persists in yes-no questions and antecedents of conditionals.¹² For instance, with yes-no questions, *par* continues to be infelicitous when the reported belief is true, as seen in (352).

(352) You told your family three months ago that you'd be home tomorrow. You're checking to make sure that they remember. You say:

- a. Toos oo-**pwaat**-i kole α-nyoon-e kaa kaaroon-i?
 MOD 2PL-think-IPFV C 1SG-come-IPFV home tomorrow-Q
 'Do you (pl) **think** that I'm coming home tomorrow?'
- b. # Toos o-**par**-e α-nyoon-e kaa kaaroon-i?
 MOD 2PL-think-IPFV 1SG-come-IPFV home tomorrow-Q
 'Are you (pl) **under the impression** that I'm coming home tomorrow?'

Par is only appropriate if the speaker is not coming, but their family thinks that they are (353).

(353) You're definitely not coming home tomorrow, but your family seems to think that you are. You say:

- a. #? Toos oo-**pwaat**-i kole α-nyoon-e kaa kaaroon-i?
 MOD 2PL-think-IPFV C 1SG-come-IPFV home tomorrow-Q
 'Do you (pl) **think** that I'm coming home tomorrow?'
- b. Toos o-**par**-e α-nyoon-e kaa kaaroon-i?
 MOD 2PL-think-IPFV 1SG-come-IPFV home tomorrow-Q
 'Are you (pl) **under the impression** that I'm coming home tomorrow?'

¹¹I assume that the inability of *par* statements to remove a proposition from the CG has a separate explanation that I do not build into *par*'s definedness condition, since this type of CG revision is a non-canonical discourse move that often requires a special mechanism (Farkas & Bruce 2010; Bledin & Rawlins 2016); for instance, Mandarin *yǐwéi* is defined whenever $p \notin CG_{out}$, yet it cannot be used to remove p from the CG (Glass 2022).

¹²It is impossible to test whether *par*'s interpretive effects persist under negation, since this verb cannot be negated while retaining its 'think' meaning. *Par* also means 'kill', though these constructions involve standard transitive syntax rather than clausal embedding. When *par* is negated, it no longer means 'think' and instead only means 'kill'. Interestingly, Mandarin *yǐwéi* also cannot be easily negated (Glass 2022:ex. 18), which suggests that a resistance to negation is perhaps a more widespread property of these sorts of biased belief verbs.

Par's interpretive effects also project from the antecedent of a conditional. In (354), the context establishes that the reported belief is false and, consequently, the *par* statement is preferred over the alternative with *pwaat*—even though the attitude verb is in a conditional antecedent.

- (354) We know that there's no test tomorrow because we're going on a field trip. But Nancy is paranoid and might think that there's going to be a surprise test on the field trip. I say:
- a. #? Kot ko-**pwaat**-e Nancy kole mii tiemuutik kaaroon, ko-sooman-i
 if 3.SBJV-think-IPFV N. C COP test tomorrow 3.SBJV-study-IPFV
 nguuni.
 now
 'If Nancy **thinks** there's a test tomorrow, she's studying now.'
 - b. Kot ko-**par**-e Nancy mii tiemuutik kaaroon, ko-sooman-i nguuni.
 if 3.SBJV-think-IPFV N. COP test tomorrow 3.SBJV-study-IPFV now
 'If Nancy **is under the impression** that there's a test tomorrow, she's studying now.'
 (context adapted from Glass 2022:ex. 16)

4.4.1.2 The distribution of *par* and ruling out existing analyses

As seen in §4.2, *par* statements are felicitous in a wide range of contexts. It is crucial, then, that the definedness condition associated with *par* be compatible with all of these contexts. Here, I show that this is the case for the denotation of *par* in (350). Empirically speaking, *par* statements are licensed when the input and output CG contain $\neg p$ (320) - (321) or say nothing with respect to p (322) - (326); these are the negative bias cases from §4.2. According to (350), *par* is licensed here because the first disjunct of its definedness condition is met (i.e. $p \notin \text{CG}_{\text{out}}$). *Par* statements are also licensed when the input and output CG contain p (333) - (336); these are the reminding cases from §4.2. Here, I follow work on discourse particles with similar reminding functions, which argues that they are felicitous when p is already in the CG, even if an interlocutor is forgetting this fact (see e.g. Repp 2013; Döring 2016; Grosz 2016).¹³ According to (350), *par* is also licensed here because the second disjunct of its definedness condition is met (i.e. $p \in \text{CG}_{\text{in}}$).

In this way, the semantics proposed in (350) render *par* felicitous across a wide range of input and output CGs: those that contain $\neg p$, say nothing with respect to p , or contain p . This distribution fully captures the empirical picture described in §4.2. Table 4.1 summarizes this state of affairs. The leftmost column describes the input CGs that allow a *par* statement, given the definedness condition in (350). The \emptyset symbol represents that CG_{in} says nothing with respect to p ; it does not indicate that CG_{in} is necessarily empty. The rightmost column describes the output CGs that are possible after utterance of a *par* statement, also given the definedness condition in (350).

¹³It is possible for a proposition to be in the CG without all interlocutors realizing it, even while maintaining a Stalnakerian view of the CG as the set of propositions that interlocutors have agreed to treat as true (Stalnaker 1974). Crucially, there is no requirement that all interlocutors are currently attending to all of these propositions, making it possible for someone to have accepted a proposition p as true, but then to forget that they accepted p .

CG _{in}	<i>par</i>	CG _{out}
$\{\neg p\}$	\rightarrow	$\{x \textit{ thinks } p, \neg p\}$
\emptyset	\rightarrow	$\{x \textit{ thinks } p\}$
$\{p\}$	\rightarrow	$\{x \textit{ thinks } p, p\}$

Table 4.1: Input and output CGs compatible with *par* statements (with respect to *p*)

In light of this distribution, *par*'s not-at-issue content cannot be modeled as a condition on just one of either the input or the output CG, as there is no unified description that captures the full range of contexts where *par* is appropriate. In the framework adopted here, presupposition can be modeled as a restriction on the input CG (Stalnaker 1974); for a presupposition to be met, the CG must look a particular way *before* utterance of the trigger. For instance, to capture the idea that *x knows p* presupposes *p*, a prominent analysis of English *know* is that it is defined if and only if the input CG already contains *p* (e.g. Kiparsky & Kiparsky 1970; Stalnaker 1974); that is, $\llbracket \textit{know} \rrbracket$ is defined only when $p \in \text{CG}_{\text{in}}$. To capture *know*'s factive presupposition, then, it is sufficient for the definedness condition to refer only to CG_{in} .

Along similar lines, lexical items can also place restrictions on the CG after an utterance—a phenomenon known as postsupposition (Lauer 2009; Brasoveanu 2013). In contrast to presupposition, the CG must look a particular way *after* utterance of the trigger for a postsupposition to be met. Glass (2022) uses postsupposition to analyze the negatively biased belief verb *yǐwéi* in Mandarin, proposing that *yǐwéi* requires the output CG to be compatible with $\neg p$. This account can be formalized as in (355); to capture *yǐwéi*'s postsupposition, it is sufficient for the definedness condition to refer only to CG_{out} .¹⁴

$$(355) \quad \llbracket \textit{yǐwéi} \rrbracket = \llbracket \textit{think} \rrbracket \\ \text{defined only when } p \notin \text{CG}_{\text{out}}$$

Both pre- and postsuppositional analyses have been proposed to account for negatively biased belief verbs across languages. Glass (2022) offers the postsuppositional analysis above for Mandarin *yǐwéi*, while Anvari et al. (2019) offer a presuppositional analysis of the negatively bi-

¹⁴In fact, the postsupposition proposed for Mandarin *yǐwéi* in Glass (2022) has parallels in work on CG management outside the realm of attitude reports; specifically, the postsupposition that $p \notin \text{CG}_{\text{out}}$ is nearly identical to the CG management content proposed for the FALSUM operator in Repp (2013) and Romero (2015) (1).

- (1) $\llbracket \textit{FALSUM} \rrbracket$: (Romero 2015:ex. 25)
- a. at-issue content: $\lambda p_{\langle s, t \rangle}. \neg p$
 - b. CG management content: $\lambda p_{\langle s, t \rangle}. \lambda w_s. \forall w' \in \text{Epi}_x(w) [\forall w'' \in \text{Conv}_x(w') [p \notin \text{CG}_{w''}]]$
where $\text{Epi}_x(w)$ is the attitude holder's epistemic alternatives and $\text{Conv}_x(w)$ is the set of worlds compatible with the attitude holder's conversational goals

Although Repp and Romero do not call the content in (1b) postsupposition—in particular, $p \notin \text{CG}_{w''}$ —it enforces the same restriction on the shape of the output CG as Glass's analysis of *yǐwéi*.

ased belief verb *creerse* in Spanish, which they argue is contrafactive. They claim that, in order for *creerse*'s presupposition to be met, the input CG must contain $\neg p$.

- (356) $\llbracket \text{creerse} \rrbracket = \llbracket \text{think} \rrbracket$
 defined only when $\neg p \in \text{CG}_{\text{in}}$

Coupled with a syntactic account of neg-raising, this analysis captures the fact that *creerse* typically implies that its complement is false, but suggests that it is true when embedded under negation.

However, it is impossible to extend either type of pre- or postsuppositional analysis directly to Kipsigis *par*, since neither the input nor output CG of a *par* statement necessarily contains a particular proposition, as highlighted in Table 4.1. Instead, I have proposed that *par* imposes a weaker requirement: that updating CG_{in} with $x \text{ par } p$ cannot involve the non-trivial addition of p to the CG. To formalize this intuition, I have relied on a bipartite definedness condition that uses familiar pre- and postsuppositional machinery. The definedness condition in (350) makes reference to the output CG—like postsupposition—and the input CG—like presupposition; in fact, (350a) is the same requirement proposed for Mandarin *yǐwéi* in Glass (2022), while (350b) is the definedness condition often associated with factive predicates like English *know*. In this way, the individual building blocks of *par*'s meaning are familiar. The novelty of the Kipsigis pattern stems from the fact that a single lexical item can be used in either type of context—with the overall effect of preventing non-trivial update of the CG with p . In this way, the unique constellation of properties found with Kipsigis *par* falls out of a new combination of familiar mechanisms.

4.4.2 Pragmatics of *par*

Notably, nothing in *par*'s semantic meaning directly gives rise to its specific interpretive effects, which are summarized from §4.2 in (357).

- (357) Use of *par* instead of *pwaat* or the bare proposition p indicates:
- | | |
|---|---|
| a. that p is false, | c. that x is unreliable, or |
| b. that the speaker is biased against p , | d. that the addressee should know p . |

Par's not-at-issue content simply states when $\llbracket \text{par} \rrbracket$ is defined; it does not make any claims about whether or not the reported belief is true, the belief holder is reliable, etc. Instead, these effects arise via context-sensitive pragmatic reasoning in response to a speaker's choice to use *par* instead of a neutral alternative like *pwaat* (in negative bias cases) or the bare proposition p (in reminding cases). By using *par*, a speaker prevents p from being non-trivially added to the CG. This choice triggers context-sensitive pragmatic reasoning about *why* the reported belief must not be added to the CG—for instance, because it is false or because it is already in the CG. These different types of reasoning take place depending on who the attitude holder is and what the tense of belief verb is and lead to the range of inferences in (357).

To see this pragmatic analysis in action, the following subsections provide sample derivations for each of the effects in (357). In all of these examples, throughout the discourse, listeners assume that the speaker is being cooperative (Grice 1989).

4.4.2.1 p is false

Consider (358)—repeated from (320b)—which strongly suggests that the speaker is not actually sick.

- (358) I'm healthy, but my mother thinks that I'm sick because I fooled her to skip school. I say:
 Ø-**Par**-e kaamεε-nyʊʊn αɑ-mnyon-i.
 3-think-IPFV mother-1SG 1SG-be.sick-IPFV
 'My mother is **under the impression** that I'm sick.'
 (context adapted from Glass 2022:ex. 4)

Upon hearing the utterance in (358), listeners reason as outlined in (359).

- (359) Upon hearing (358), listeners reason:
- For $\llbracket(358)\rrbracket$ to be defined, either $p \notin CG_{out}$ or $p \in CG_{in}$.
 - The speaker intends for their utterance to have a defined truth value.
 - Given (359a) and the fact that p is not in the input CG, the speaker must intend for p to not be in CG_{out} .
 - The speaker knows whether or not they are sick and would not object to this information being added to CG_{out} if it were true.
 - Because the possibility of adding p to CG_{out} is ruled out (359c), p must be false.

4.4.2.2 Speaker is biased against p

Second, the utterance in (360)—repeated from (325b)—suggests that the speaker is biased against p , even though the truth of p is currently unknown.

- (360) We turn on a Kalenjin TV station and see a biased political pundit discussing the upcoming election. Talking about a viable candidate named Jessica who he believes will lose the election, the pundit says:
 Ø-**Par**-e Jesika Ø-sindan-iisye εεn lewenisiet.
 3-think-IPFV J. 3-win-AP.IPFV in election
 'Jessica is **under the impression** that she will win the election.'
 (context adapted from Glass 2022:ex. 5)

Upon hearing the utterance in (360), listeners reason as outlined in (361).

- (361) Upon hearing (360), listeners reason:
- For $\llbracket(360)\rrbracket$ to be defined, either $p \notin CG_{out}$ or $p \in CG_{in}$.
 - The speaker intends for their utterance to have a defined truth value.
 - Given (361a) and the fact that p is not in the input CG, the speaker must intend for p to not be in CG_{out} .
 - The speaker cannot know if Jessica will win the election, but they would not object to this information being added to CG_{out} if they considered it likely.

- e. Because the possibility of adding p to CG_{out} is ruled out (361c), the speaker must view p as unlikely to be true.

The reasoning here parallels that described in §4.4.2.1, though it differs in that the speaker cannot be assumed to know if the reported belief is true—contrary to (359d). Given (361d), listeners instead conclude that the speaker intends for p to not be added to CG_{out} because they are biased against it in some way. In the case of (360), this bias stems from the speaker’s personal beliefs about the upcoming election, though the same reasoning applies across different sources of bias, as shown in the following subsection.

4.4.2.3 x is unreliable

Third, the *par* statement in (362)—repeated from (330)—suggests that p is unlikely to be true because the belief holder is unreliable, even though the speaker is explicitly ignorant about p .

- (362) Arap Bett is drunk and is acting confused. I have no idea if Arap Ruto is here or not, but I have reason to doubt Arap Bett’s reliability. I say:

Maa-ngen koto ka-ko-it Arap Ruto anan tomo lakini Ø-**par**-e Arap
 NEG.1SG-know if PST1-3.PERF-arrive son.of R. or not.yet but 3-think-IPFV son.of
 Beet ka-ko-it.

B. PST1-3.PERF-arrive

‘I don’t know if Arap Ruto has arrived yet, but Arap Bett is **under the impression** that he has.’

(context adapted from Glass 2022:Appendix ex. 12)

Upon hearing the utterance in (362), listeners reason as outlined in (363).

- (363) Upon hearing (362), listeners reason:
- For $\llbracket(362)\rrbracket$ to be defined, either $p \notin CG_{out}$ or $p \in CG_{in}$.
 - The speaker intends for their utterance to have a defined truth value.
 - Given (363a) and the fact that p is not in the input CG, the speaker must intend for p to not be in CG_{out} .
 - The speaker does not know if Arap Ruto has arrived, but they would not object to this information being added to CG_{out} if they considered there to be sufficient evidence for p .
 - Because the possibility of adding p to CG_{out} is ruled out (363c), the speaker must view p as unlikely to be true.

Note that the pragmatic reasoning outlined in (363) is nearly identical to that in (361); the only difference between (363) and (361) is the reason for the speaker’s belief that p is unlikely or insufficiently evidenced. Here, the speaker is biased against the reported belief because they deem the belief holder unreliable, whereas in (361), the speaker is biased against the reported belief due to their own opinions about the upcoming election. In this way, I discuss speaker

bias against p and unreliable attitude holders as two separate empirical phenomena, although the pragmatic reasoning that underlies these two patterns is actually quite similar.¹⁵

The pragmatic reasoning outlined here and in §4.4.2.2 also correctly predicts that speakers must have a reason for using *par* when they do not know if p is true. As noted in §4.2, when the input CG says nothing with respect to p , *par* statements are infelicitous in out-of-the-blue contexts. For instance, (360) is inappropriate for an impartial journalist with no known political affiliation to report and only becomes acceptable when used by a biased political pundit. Likewise, *par* statements are natural when the belief holder is deemed unreliable (362), though these same statements are infelicitous in out-of-the-blue contexts (329b). This pattern is expected given the pragmatic reasoning described here because it hinges on how listeners reason about a speaker's choice to block addition of p to the CG. When listeners are unable to come up with a justification for the speaker's choice to use *par*, as is the case in out-of-the-blue contexts, the pragmatic reasoning described here fails and, consequently, *par* is infelicitous.

4.4.2.4 Addressee should know p

Fourth, in the present tense with a first person belief holder, *par* serves a reminding function; (364)—repeated from (335)—suggests that the addressee should already know that it is cold out.

- (364) I see a friend heading outside in a t-shirt and shorts, even though it's the middle of winter. I want to remind them that it's way too cold for that kind of outfit! I say:
 A-**par**-e kartit ra.
 1SG-think-IPFV cold today
 'Hang on, it's cold today.' (Lit: 'I think that it's cold today.')

The pragmatic reasoning here differs from that seen previously because of the assumptions that listeners make about a speaker's currently held beliefs. More specifically, upon hearing the utterance in (364), listeners reason as outlined in (365).

- (365) Upon hearing (364), listeners reason:
- a. For $\llbracket(364)\rrbracket$ to be defined, either $p \notin CG_{\text{out}}$ or $p \in CG_{\text{in}}$.
 - b. The speaker intends for their utterance to have a defined truth value.
 - c. The speaker believes p to be true; in fact, this is the asserted content of (364).
 - d. Because the speaker believes p to be true, they would not object to it being in CG_{out} .
 - e. Therefore, it must be the case that p is in CG_{in} according to the speaker.

¹⁵Another possible reason to use *par* has to do with the type of evidence for a reported belief. When a speaker considers a belief holder's evidence to be weak, *par* statements are felicitous. This use is mainly attested for one speaker, which is why I do not include it in the main text, but the effect lines up with the facts reported here, which hold across speakers. However, determining what exactly constitutes weak evidence is a question for future research, since it does not align neatly with evidential scales like those proposed by Willett (1988). For instance, if a speaker considers indirect evidence or hearsay to be sound, *par* cannot be used, while direct evidence that the speaker deems faulty in some way can license *par*. I leave exploring the link between evidence source and *par* to future work.

- f. Given this, interlocutors infer that the speaker is trying to remind them of p , since the speaker believes that they are not attending to this point.

As before, *par*'s definedness condition must be met (365a). Yet here, listeners reason differently about which disjunct of the definedness condition applies. Because the speaker believes p —this is the asserted content of (364)—they are presumably comfortable with p being in CG_{out} (365d), since adding true information to the CG is the goal of a conversation. As a result, listeners conclude that p must be in the input CG. Once listeners draw this conclusion, a reminding inference arises naturally, just as with German *ja* and *doch* on many accounts (e.g. Repp 2013; Grosz 2016); even though an interlocutor might not remember p , the speaker of (364) views p as common knowledge, giving rise to a reminding inference. In this way, *par*'s reminding function falls out of 1) how listeners reason about the relationship that speakers have with their present beliefs and 2) how that relationship interacts with the definedness condition attached to *par*.

In this way, the denotation of *par* that I propose lays the groundwork for deriving both negative bias and reminding inferences, given its bipartite definedness condition. However, an alternative approach encodes only negative bias in the belief verb's denotation—adopting a semantics like (355) or (356)—and rules in reminding uses via a pragmatic rescue strategy. The intuition here is that *par*'s negative bias function is default; only when this use is blocked—as with first person belief holders in the present tense—does the reminding function become available. However, if this rescue strategy were based on general pragmatic principles, it should also be available with Mandarin *yǐwéi* or Spanish *creerse*, even though these forms have no such function (§4.3). By contrast, on my account, the variation between Kipsigis vs. Spanish and Mandarin is built into the meaning of *par*. In my opinion, this is a more natural way to capture cross-linguistic variation, instead of proposing that pragmatic principles variably hold across languages.

4.4.2.5 Restricting *par*'s reminding function

Recall from §4.2.2 that *par*'s reminding function is limited to cases with first person belief holders in the present tense. How can this restriction be captured? One tempting possibility is through competition with factive *ngen* 'know' with a non-first person belief holder.¹⁶ Much like English *know*, Kipsigis *ngen* is defined only when $p \in CG_{in}$. By contrast, with non-first person belief holders, *par* is ambiguous between two different possibilities: $p \notin CG_{out}$ or $p \in CG_{in}$. A speaker's choice to use the weaker, ambiguous option *par* over the stronger, unambiguous option *ngen* would implicate that *ngen*'s presupposition is not met, giving rise to negative bias but not reminding with non-first person belief holders.

However, this analysis falls short in contexts where *ngen* is not licensed but reminding uses with a non-first person belief holder remain unavailable. Such a context is found in (366), repeated from (337). Here, the *ngen* statement is infelicitous because the doctor lacks the proper justification required for their belief to be reported as knowledge. A reminding use of third person *par* continues to be unavailable—even without competition with *ngen*—which shows that this competition cannot be solely responsible for the restriction.

¹⁶Many thanks to Judith Tonhauser for pointing out this line of analysis.

(366) You go to the doctor because you're coughing and have a sore throat. The doctor thinks that you have Covid, but didn't run a test because he ran out. He sends you home, but tells you to isolate and follow the Covid guidelines. I see you leaving the house and want to remind you about your diagnosis. I say:

- a. # \emptyset -Par-e dakitaari i-tiny-e koroona.
 3-think-IPFV doctor 2SG-have-IPFV Covid
 'Remember, according to the doctor, you have Covid.'
- b. # I-ngen dakitaari kole i-tiny-e koroona.
 3-know doctor C 2SG-have-IPFV Covid
 'The doctor knows that you have Covid.'

Instead, I suggest that the restriction on *par*'s reminding use arises due to a preference for some kinds of pragmatic reasoning over others. *Par* statements are defined in two different environments: when p is not in CG_{out} or when p is already in CG_{in} . Negative bias and reminding inferences, then, arise pragmatically based on how listeners reason about which disjunct of *par*'s definedness condition applies in a given context. When listeners reason that a speaker used *par* because $p \notin CG_{out}$, a negative bias inference arises (367a). On the other hand, when listeners reason that a speaker used *par* because $p \in CG_{in}$, a reminding inference arises (367b).

(367) Pragmatic reasoning pathways

- a. $\llbracket Par \rrbracket$ is defined because $p \notin CG_{out}$. The speaker must want to block the inference from x *thinks* p to p because p is false, unlikely, or poorly evidenced.
 (negative bias reasoning)
- b. $\llbracket Par \rrbracket$ is defined because $p \in CG_{in}$. The speaker must want to remind an interlocutor of p , since, in the speaker's view, they are not attending to this point.
 (reminding reasoning)

I propose that the type of pragmatic reasoning in (367a) is a default, preferred over the one in (367b) whenever possible. With first person belief holders, only (367b) is possible, since interlocutors know that the speaker believes p and so would not object to p being part of CG_{out} . Yet with non-first person belief holders, both pragmatic reasoning pathways are, in principle, possible because interlocutors do not know whether the speaker believes p . In such cases, the Kipsigis data show that (367a) takes precedence over (367b).

It is possible that this preference is grounded in conversational principles. I assume that the primary goal of conversation is to shrink the context set by adding truthful propositions to the CG, so that interlocutors can determine which possible world they are in. To achieve this goal, it is essential that truthful information be added to the CG *and* that false or unjustified information be kept out of the CG. The reasoning around *par* in (367a) does exactly this: it prevents the addition of a false or unjustified proposition to the CG. By contrast, the reasoning around *par* in (367b) is useful, but does not serve the primary goal of conversation as directly; while it is helpful to remind interlocutors what the CG already looks like, it is arguably not as essential to the development of a conversation as avoiding false CG update. At present, I leave this preference for one line of pragmatic reasoning over another as a stipulation—perhaps grounded in

conversational principles—since there is no clear alternative that captures the restriction. However, understanding how to integrate such a preference into a model of definedness conditions and pragmatic reasoning is a question for further work.

4.4.2.6 Evidence for pragmatic inferences

To conclude this section, I provide evidence that the inferences triggered by *par* are derived pragmatically. Support for this analysis comes from the fact that these effects can be reinforced without redundancy and are cancellable—two hallmarks of pragmatic inferences (Hirschberg 1985; Grice 1989; Potts 2014). For instance, in contexts where *p* is known to be false, speakers can reinforce the falsity of *p* without redundancy. These continuations are particularly natural if the speaker provides additional explanation for the false belief.

- (368) We all know that I'm perfectly healthy. But my mother thinks that I'm sick because I fooled her to skip school.
 Ø-**Par**-e kaamεε-nyʊʊn aa-mnyon-i lakini maa-mnyon-i. Kaa-ng'al-e
 3-think-IPFV mother-my 1SG-be.sick-IPFV but NEG.1SG-be.sick-IPFV PST1.1SG-lie-IPFV
 si maa-wa-i sʊgʊl.
 PURP NEG.1SG-go-LP school
 'My mother is **under the impression** that I'm sick, but I'm not sick. I was lying to not go to school.'
 (context adapted from Glass 2022:ex. 4)

Par's other interpretive effects, like the unreliability of *x*, are also reinforceable, as seen in (369).

- (369) Arap Bett is very drunk and is acting confused. I don't know if Arap Ruto is here or not, but I have reason to doubt Bett's reliability.
 Ø-**Par**-e Arap Bett ka-ko-it Arap Ruto lakini maa-pwaat-e kole
 3-think-IPFV son.of B. PST1-3.PERF-arrive son.of R. but NEG.1SG-think-IPFV C
 ka-ko-it. Ø-Pook-iit-i Arap Bett. Maa-yon-i che Ø-mwa-e.
 PST1-3.PERF-arrive 3-drunk-VBLZ-IPFV son.of B. NEG.1SG-believe-IPFV REL.PL 3-say-IPFV
 'Arap Bett is **under the impression** that Arap Ruto has arrived, but I don't think that he has. Arap Bett is drunk. I don't believe what he says.'
 (context adapted from Glass 2022:Appendix ex. 12)

The fact that these effects can be reinforced without redundancy suggests that they are not part of the asserted content of *par*, but instead arise via pragmatic reasoning.

Par's interpretive effects are also cancellable, though cancellation requires more contextual support than reinforcement, parallel to the pattern seen with Kipsigis *-yan* forms in Chapter 3 §3.4. Continuations like that in (370) are not contradictory, but they are marked discourse moves that typically require the use of veracity emphasizees like *εen iman* 'in truth'. In particular, the speaker's use of *par* in the first half of (370a) suggests that *p* should not be added to the CG simply because many people believe it—even though it ultimately ends up being true. The follow-up in

(370b) then explains the unexpected reason why p is true, at which point it can be added to the CG. In this way, cancellation after use of *par* must serve some rhetorical purpose, intentionally setting up the expectation that p is false in order to subvert this expectation in a surprising or interesting way.

(370) My friend Lydia invented a famous app, and people think she made millions from it. Actually, although my friend never made any money from her app, she inherited money from her parents. I say:

- a. \emptyset -**Par**-e piik mogoriot Lidya, ako $\epsilon\epsilon$ n iman ko-mogoriot...
 3-think-IPFV people rich.person L. and in truth 3.SBJV-rich.person
 ‘People are under the impression that Lydia’s rich, and she actually is...’
- b. Lakini ma-mogoriot kiin ko-al-ta ap. Kii- \emptyset -koo-chi siigiik-chik
 but NEG-rich.person when 3.SBJV-buy-IT app PST3-3-give-APPL parents-3.POSS
 rabimik.
 money
 ‘But she’s not rich from selling the app. Her parents gave her the money.’
 (context adapted from Glass 2022:ex. 9)

Interestingly, negatively biased belief verbs in other languages show a similar pattern in terms of cancellation difficulty. The interpretive effects of Spanish *creerse* cannot be cancelled (Anvari et al. 2019:ex. 10), while those of Mandarin *yǐwéi* are more difficult to cancel than to reinforce (Glass 2022:ex. 27-28). While it is outside the scope of this chapter to offer a full explanation as to why this cancellation difficulty persists cross-linguistically, it is worth noting that this behavior is shared across different biased belief verbs with different empirical profiles in other aspects of the grammar. Nevertheless, the behaviors discussed in this section provide evidence that the specific inferences triggered by use of *par* are pragmatically derived.

4.4.3 The grammatical status of first person present *par*

Because *par*’s reminding function is so different from its negative bias use and is restricted to first person belief holders, it is worth considering the possibility that 1SG present *apare* is a distinct lexical item from the other forms of *par*. For instance, perhaps *apare* is a discourse particle akin to German *ja* or *doch* that is only related to the negative bias use of *par* historically. Yet even if this were the case, the analysis of *par* would have to be compatible with a semantic change pathway that derives a reminding function from a negative bias one, since *apare* is linked to *par* at least morphophonologically. Existing analyses of negative bias do not allow for this possibility. As outlined in §4.4.1, existing pre- and postsuppositional accounts require $\neg p$ to be in CG_{in} (Anvari et al. 2019) or for CG_{out} to be compatible with $\neg p$ (Glass 2022). These requirements preclude the development of *par*’s reminding function, since it would not be licensed with a first person belief holder in the present tense in the first place. By contrast, the analysis here provides a synchronic derivation for *par*’s reminding function (if it is a verb) or lays the groundwork for the diachronic development of this function (if it is a discourse particle).

Additionally, there is evidence to suggest that 1SG present *apare* is synchronically a verb. First, it contains decomposable verbal morphology, including 1SG subject marking *a-* and imperfective aspect *-e*, and it must surface clause-initially, which is the only grammatical position for the verb in Kipsigis (Bossi & Diercks 2019).¹⁷ Second, Kipsigis does not generally have discourse particles; to my knowledge, there are no such particles in the language. While this is not proof that they do not exist, it makes for a marked contrast with languages like German, which have a large inventory of these particles. Given this, if *apare* were a discourse particle, it would be a member of a small—if not singleton—class of elements. Third, *apare* can occur with verbal intensifiers like *ime*. This element can surface in many postverbal positions, even quite distant from the intensified verb (371). Crucially, *ime* can combine with *apare* on its reminding use, as seen in (372); here, even though *ime* is not adjacent to *apare*, it highlights the fact that the addressee should already know *p*, as suggested by speaker comments and ‘clearly’ in the translation.

(371) Ma-Ø-mach-e {ime} ko-wa sʊgʊl {ime} Kiproono {ime}.
 NEG-3-want-IPFV INE 3.SBJV-go.SG school INE K. INE
 ‘Kiproono really doesn’t want to go to school.’

(372) We’re walking through the garden and see animal tracks. The steps are clearly those of a cow: they’re the shape of cow hooves and spaced out like cow hooves. Yet I ask you what animal it was. You reply:
 A-**par**-e tɛɛta ime.
 1SG-think-IPFV cow INE
 ‘As you should really know, this is a cow.’ (Lit: ‘I definitely think that this is a cow.’)

Together, these facts suggest that *apare* is a verb. However, even if this conclusion turns out to be incorrect and *apare* is a discourse particle, the analysis here provides a better foundation for deriving its reminding function than other analyses of negatively biased belief verbs.

4.4.4 Against a syntactic alternative

Although this chapter focuses on the interpretive differences between belief reports with *pwaat* and *par*, these two constructions also involve slightly different syntactic complementation strategies. In this section, I describe the syntactic differences between *pwaat* and *par* statements and show that they cannot be responsible for the interpretive differences described in §4.2. As seen throughout the chapter, complementation with *pwaat* requires the element *kole* (373a), while *par* cannot occur with *kole* (373b).

(373) a. I-**pwaat**-e kaameɛ-nyʊʊn *(kole) aa-mnyon-i.
 3-think-IPFV mother-1SG.POSS C 1SG-be.sick-IPFV
 ‘My mother **thinks** that I’m sick.’

¹⁷While this section is concerned specifically with the grammatical status of 1SG present *apare*, evidence that *par* is generally a verb comes from the fact that it takes verbal morphology, including tense, aspect, and indicative and subjunctive subject marking (in the environments where one would expect these moods to appear in the language). Illustrative data examples can be seen throughout the chapter; for instance, indicative vs. subjunctive subject marking on *par* is found in (320) - (321) vs. (354b).

- b. Ø-**Par**-e kaameε-nyʊʊn (*kole) aa-mnyon-i.
 3-think-IPFV mother-1SG.POSS C 1SG-be.sick-IPFV
 ‘My mother is **under the impression** that I’m sick.’

Some other complement-taking verbs that require *kole* include: *ngen* ‘know’, *mwa* ‘say’, *ruaatit* ‘dream’, and *nereech* ‘be angry’. In this way, verbs that use the *kole* embedding strategy come from a variety of lexical classes and include factive and non-factive verbs.

Diercks & Rao (2019) treat *kole* as a complementizer; hence its glossing to this point in the chapter. However, recent work by Driemel & Kouneli (2022a) argues that it is actually the lexical verb *le* ‘say’ with the third person subjunctive marker prefix *ko-*. For this reason, I refer to this element as “complementizer-like *le*” here. Evidence for this analysis comes from the unique agreement behaviors of complementizer-like *le*; in particular, it shows prefixal agreement with whichever matrix argument is the logophoric center of the belief report. In this way, it often agrees with the matrix subject, though it can also agree with a matrix object when it qualifies as the source of the information reported in the embedded clause (374).¹⁸

- (374) Ka-I-kas-ε:n Kiplàngàt {kò-lé / ì:lé} kà-Ø-tfó:r Kíbê:t rabɪ:nɪk.
 PST1-2SG-hear-APPL K. 3.SBJV-LE 2SG.SBJV-LE PST1-3-steal K. money
 ‘You heard from Kiplangat that Kibet stole the money.’ (Driemel & Kouneli 2022a:ex. 35)

Driemel & Kouneli argue that these agreement behaviors arise because *le* is an embedded lexical verb meaning ‘say’ that agrees with its subject—a null logophoric *pro* that is co-referential with the matrix argument controlling agreement on *le*. On this analysis, sentences like (374) actually involve two instances of embedding: the matrix verb *kas* ‘hear’ embeds a subjunctive TP containing *le* and a logophoric *pro* subject, which then embeds an indicative CP containing the embedded verb *tfɔ:r* ‘steal’. (375) schematizes this state of affairs; the most crucial observation is that sentences like (374) are actually *triclausal* rather than *biclausal*.

- (375) [_{CP} matrix verb_{ind} [_{TP} logophoric *pro* ... *le*_{subjv} [_{CP} embedded verb_{ind}]]]

Driemel & Kouneli extend their morphosyntactic analysis of *le* to its semantics, suggesting that *le* is uniformly a speech verb, and so embedded clauses headed by *le* are sets of contentful saying events. On this analysis, then, there are two differences between statements with *pwaat* vs. *par*: 1) *pwaat* statements have speech semantics that are absent in *par* statements, and 2) *pwaat* statements contain a subjunctive TP that is absent in *par* statements. In the remainder of this section, I consider these differences and show that there is reason to doubt both claims and, regardless, that they cannot be responsible for the interpretive effects described here.

First, as detailed in Bossi (2023) and mentioned in Chapter 2 §2.5.6, there are reasons to think that complementizer-like *le* does not always contribute speech semantics. For instance, complementizer-like *le* is obligatory in attitude reports that do not semantically involve speech,

¹⁸Complementizer-like *le* can also display other morphology indicative of its verbal status, including: suffixal agreement with indirect objects of speech verbs, the applicative morpheme *-chi*, and the reflexive particle *-kεε*. See Driemel & Kouneli (2022a) for more discussion of these syntactic patterns.

as seen in (376) where speech is explicitly denied and in (377) where the attitude holder is an inanimate object incapable of speech.

(376) Koo-aa-ngen *({ko-le / aa-le}) koo-miit-een Kibeet, lakini ma-a-mwa kiy.
 PST2-1SG-know 3.SBJV-LE 1SG.SBJV-LE PST2-COP-in K. but NEG-1SG-say thing
 ‘I knew that Kibet was there, but I didn’t say anything.’

(377) Koo-∅-keer kamera *(ko-le) koo-∅-it Kibeet.
 PST-3-see camera 3.SBJV-LE PST-3-arrive K.
 ‘The camera saw that Kibet arrived.’

If complementizer-like *le* necessarily contributes speech semantics, the data in (376) - (377) are unexpected, since *le* should presumably introduce speech here too. Moreover, even if one adopts the analysis in Driemel & Kouneli (2022a) in which *le* is uniformly a speech verb, speech semantics alone would not derive the interpretive effects described here; there is no reason why the absence of speech semantics in *par* statements would suggest that the reported belief is false or that the belief holder is unreliable, for instance.

In light of semantic patterns like those in (376) - (377)—as well as other syntactic facts about Kipsigis complementation—Bossi (2023) concludes that, in belief reports, *le* is a semantically-bleached complementizer that embeds an indicative clause. In this way, there is no actual difference in mood selection between belief reports with *pwaat* and *par*; both belief verbs embed indicative clauses. In Kipsigis, subjunctive subject marking is in complementary distribution with overt tense marking. Therefore, evidence that *par* embeds an indicative clause comes from the fact that the full range of tense distinctions persists in these embedded clauses. (378) illustrates this point with the three past tenses found in Kipsigis: recent, default, and distant past.

(378) We know that no one saw Chepkoech {earlier today / yesterday / last year}, but Kiproono’s confused and mistakenly thinks that I saw her at these various times. I say:

- a. ∅-Par-e Kiproono kaa-keer Cheεpkεech.
 3-think-IPFV K. PST1.1SG-see C.
 ‘Kiproono is **under the impression** that I saw Chepkoech (recently).’
- b. ∅-Par-e Kiproono koo-a-keer Cheεpkεech.
 3-think-IPFV K. PST2-1SG-see C.
 ‘Kiproono is **under the impression** that I saw Chepkoech.’
- c. ∅-Par-e Kiproono kii-a-keer Cheεpkεech.
 3-think-IPFV K. PST3-1SG-see C.
 ‘Kiproono is **under the impression** that I saw Chepkoech (long ago).’

Given that *pwaat* and *par* both embed indicative clauses, it is unlikely that the interpretive effects associated with *par* arise due to the mood of the embedded clause.

Furthermore, even if one follows Driemel & Kouneli (2022a) in treating complementizer-like *le* as a subjunctive speech verb, the Kipsigis pattern is unexpected because it is the neutral belief report—not the biased one—that contains a subjunctive embedded clause. Mood selection often correlates with the level of certainty indicated by an attitude verb. In particular, the subjunctive

is often used cross-linguistically with verbs of doubting (Siegel 2009), in which case one might expect *par* to select for this mood. However, on Driemel & Kouneli’s view, the Kipsigis pattern is precisely the opposite; *par* selects for an indicative clause, while *pwaat* selects for a subjunctive clause with *le*, which then embeds an indicative clause. This pattern suggests again that *par*’s interpretive effects are not due to the mood of the embedded clause, since selection of the indicative should not contribute negative bias, nor should the absence of the subjunctive. Taken together, these facts suggest that the syntactic differences between *pwaat* and *par* statements are not responsible for the interpretive effects described here.

4.5 Conclusion

This chapter describes and analyzes the biased belief verb *par* in Kipsigis, which serves two seemingly contradictory functions: with a non-first person belief holder, *par* suggests that the reported belief is false or unlikely, while with a first person belief holder in the present tense, *par* reminds the addressee that the reported belief is true. While these functions are familiar on their own (see e.g. other negatively biased belief verbs like Spanish *creerse* and Mandarin *yǐwéi*, and reminding discourse particles like German *ja* and *doch*), Kipsigis *par* is unique in that a single lexical item can serve both functions depending on the context.

Thinking about the cross-linguistic picture, the literature on biased belief reports is quite small, and the Kipsigis pattern is perhaps only surprising from the current typological perspective. Biased belief is a concept that is grammaticalized in more languages than the theoretical literature currently addresses. For instance, other languages are reported to have biased belief verbs, including Caquinte *ji* (Arawakan; O’Hagan 2021), Lobi *ká jma* (Gur; Jarvis p.c.), and Niuean *piko* (Polynesian; Levin & Massam 1985), which is shown in (379).

- (379) Niuean (Levin & Massam 1985:ex. 11b)
Piko e mangafaoa haaku ne fano a koe ki Sāmoa.
 think ABS family my PST go ABS you to Samoa
 ‘My family believed (mistakenly) that you were going to Samoa.’

Whether or not these predicates pattern with Kipsigis *par* is an open question, since the current documentation does not provide the relevant data. However, going forward, it is possible that research will show that the Kipsigis pattern is not actually that unique; perhaps the co-occurrence of negative bias and reminding functions is a hallmark of a particular class of biased belief verbs, as the current analysis makes possible. In addressing this question, it is crucial that researchers consider these biased belief verbs in a range of contexts that vary (at least) in terms of the tense and subject of the belief verb.

On a more theoretical note, to capture the different uses of Kipsigis *par*, I propose that it comes with a not-at-issue instruction for CG management: that *p* not be added to the CG. In light of this CG management instruction, context-sensitive pragmatic reasoning derives the specific interpretive effects seen with *par*. In formalizing this analysis, I show that *par* cannot be modeled using a filter on just one of either the input or output CG, as is standard in analyses of CG management

(e.g. Repp 2013; Grosz 2016; Anvari et al. 2019; Glass 2022). Instead, *par* requires a more complex definedness condition that imposes requirements on the shape of both the input and output CG, following work on pre- and postsupposition.

The utility of postsupposition in modeling Kipsigis *par* provides further evidence in support of this theoretical machinery. While presupposition is well established in the literature, postsupposition is a relatively new idea, which blurs the boundary between semantics and pragmatics. The postsuppositional component of *par*'s definedness condition (i.e. $p \notin \text{CG}_{\text{out}}$) is semantic, yet it controls not only the direct, *semantic* contributions of the utterance, but also the *pragmatic* outputs. Following Glass (2022), I assume that this behavior is characteristic of postsuppositional belief verbs, yet this observation underscores the way in which postsupposition crosses the boundary between semantics and pragmatics. Exploring other applications of postsupposition and its ability to block pragmatic inferences is an interesting area for future research, as the empirical profile of postsupposition continues to become clear.

Chapter 5

Conclusion

This dissertation has examined two different ways to make epistemic modal claims across syntactic categories in Kipsigis—one in the nominal domain and another in the verbal domain. In both of these case studies, the observed epistemic effects arise not due to quantification over possible worlds, but rather are derived using other independently attested machinery like pragmatic competition between nominals or instructions for Common Ground management. In this chapter, I provide a brief summary of the main empirical and theoretical contributions of this work and discuss areas for future research—both within Kipsigis and across languages.

5.1 Empirical and theoretical contributions

5.1.1 Epistemic indefinites

In Chapter 3, I described the properties of Kipsigis *-yan* forms, which are epistemic indefinites that can signal both first order ignorance about the individual witness to the indefinite and higher order ignorance about one or more of the salient properties of that individual. By comparing these *-yan* forms with other epistemic indefinites across languages, I showed that they have a unique constellation of properties; in particular, they are compatible with singleton domains of quantification, unlike domain widening epistemic indefinites (e.g. German *irgendein*, Kratzer & Shimoyama 2002; Spanish *algún*, Alonso-Ovalle & Menéndez-Benito 2010, 2017), and they display scopal flexibility, unlike Tiwa *-khi* indefinites, which Dawson (2018, 2020) argues to be choice functional.

I also showed that the ignorance effects triggered by *-yan* forms are pragmatically derived via competition with Kipsigis singular nouns containing a secondary suffix, which motivated my discussion of the semantics of these forms. As proposed in Kouneli (2021), these Kipsigis nouns have a complex internal structure including a determiner that is called the secondary suffix in the Kalenjin literature. As seen in Chapter 3, these singular nouns necessarily take wide scope and signal speaker knowledge about how the noun's referent is to be identified. In light of these patterns, I suggested that the secondary suffix in Kipsigis singular nouns introduces a free choice

function variable that is resolved via a contextually supplied assignment function, (380).

(380) $\llbracket \text{-SEC} \rrbracket = \lambda P_{\langle e,t \rangle}.f(P)$, where f is a Choice Function that enables the speaker to identify the referent in a way that is relevant to their goals

Then, to capture the full range of patterns seen with Kipsigis *-yan*, I proposed that it introduces basic existential quantification over individuals, (381).

(381) $\llbracket \text{-yan} \rrbracket = \lambda P_{\langle e,t \rangle}.\lambda Q_{\langle e,t \rangle}.\exists x[P(x) \ \& \ Q(x)]$

Following work on Russian *to* and *koe* indefinites by Geist & Onea (2007), I claimed that, when Kipsigis *-yan* forms and singular nouns with a secondary suffix are in competition, use of the existentially quantified *-yan* form rather than the specific, referential singular noun leads to ignorance effects via pragmatic reasoning. More specifically, by choosing to use the *-yan* form rather than the singular noun with a secondary suffix, the speaker indicates that they are unable to provide a value for the choice function variable introduced by the secondary suffix. This inability licenses use of the less informative *-yan* form, and ignorance effects arise pragmatically as a result.

The broader theoretical takeaways from this case study are two-fold. First, the novel Kipsigis pattern calls into question the correlation between ignorance type and scope proposed in Dawson (2020). Dawson suggests that if an epistemic indefinite triggers higher order ignorance, then it must show exceptional wide scope. However, the Kipsigis pattern necessitates a reframing of this correlation; instead, in Chapter 3, I proposed that if an epistemic indefinite displays exceptional wide scope, pragmatically derived higher order ignorance must be possible. Second, the Kipsigis data suggest that the presence or absence of existential closure of choice function variables can be diagnosed using the type of test in Matthewson (1999) and that epistemic effects associated with choice functional indefinites might be another diagnostic for existential closure. More specifically, in Chapter 3, I claimed that existentially closed choice function variables—in a particular nominal inventory—lead to ignorance effects (as with Tiwa *-khi* indefinites), while free choice function variables lead to speaker knowledge effects (as with Kipsigis singular nouns with a secondary suffix).

5.1.2 Biased belief reports

In Chapter 4, I analyzed the biased belief verb *par* ‘think’ in Kipsigis, which has two seemingly contradictory functions. With non-first person belief holders, use of *par* suggests that the speaker views the reported belief as false or unlikely. However, with a first person belief holder in the present tense, *par* serves a reminding function, indicating that the addressee should already know the reported belief to be true. These negative bias and reminding functions are attested independently across languages; for instance, other negatively biased belief verbs like Spanish *creerse* (Anvari et al. 2019) and Mandarin *yǐwéi* (Glass 2022) only have a negative bias function, while discourse particles like German *ja* and *doch* only have a reminding function (Döring 2016; Grosz 2016). The novelty of the Kipsigis pattern lies in the way that it combines two familiar, yet seemingly contradictory, phenomena.

To capture the full range of *par*'s effects, in Chapter 4, I proposed that, in addition to its basic belief semantics, *par* contributes an instruction for Common Ground management (Krifka 2008): that the reported belief is not to be *added* to the Common Ground. This instruction is formalized as in (382) using a disjunctive definedness condition that makes reference to both the input and the output Common Ground. In contrast to existing analyses of negative bias and reminding functions (e.g. Repp 2013; Grosz 2016; Anvari et al. 2019; Glass 2022), I showed that the Kipsigis pattern cannot be modeled as a constraint on one of either the input or the output Common Ground: both concepts must be referenced to account for all of *par*'s functions.

(382) $\llbracket \text{par} \rrbracket = \llbracket \text{think} \rrbracket = \lambda p \lambda x \lambda w. \forall w' \in \text{Dox}_{x,w} [p(w') = 1]$
defined only when

- a. $p \notin \text{CG}_{\text{out}}$, or
- b. $p \in \text{CG}_{\text{in}}$

(383) Paraphrase of (382a) - (382b): Do not add p to the CG.

Given the semantics in (382), I proposed that listeners reason about why the reported belief must not be added to the Common Ground, which gives rise to negative bias or reminding inferences. In all cases, I assumed that speakers intend for their utterance to have a defined truth value; for this to happen, one of the conditions in (382a) - (382b) must be met. With a non-first person belief holder, listeners determine that the condition in (382a) applies and reason the reported belief cannot be added to the output Common Ground because the speaker knows it to be false or because they are biased against it in some way. By contrast, with a first person belief holder, listeners assume that the speaker considers their belief to be true—otherwise, they would not believe it. This assumption sparks a different kind of pragmatic reasoning. Since the speaker presumably wants to add true information to the Common Ground, listeners rule out the possibility that the condition in (382a) applies. Instead, they conclude that (382b) applies and that the reported belief is already in the Common Ground, even if they are forgetting this fact. This reasoning gives rise to a reminding inference.

This case study sheds new light on the pragmatic functions of belief reports and highlights the need for Common Ground management instructions of the sort in (382). As discussed in Chapter 4, belief reports serve a range of pragmatic functions beyond their semantic meaning; in fact, recent work on child language acquisition highlights the salience of the pragmatic uses of belief reports, showing that children take advantage of the pragmatically enriched meaning of these expressions when acquiring the literal meaning of attitude verbs (Hacquard & Lidz 2018). This case study expands the current landscape regarding the pragmatic uses of belief reports, by offering a detailed description of a typologically unique belief verb from an understudied language. Furthermore, as noted above and in Chapter 4, in order to account for the full Kipsigis pattern, the definedness condition attached to *par* must refer to both the input and output Common Ground. In this way, the Kipsigis data provide additional empirical support for the relatively new theoretical concept of postsupposition, which is a counterpart of presupposition that constrains the shape of the output Common Ground, rather than the input Common Ground (Lauer 2009; Brasoveanu 2013; Glass 2022).

5.2 Future directions

This dissertation has laid the foundation for further research within Kipsigis, in Kalenjin languages more broadly, and across other languages. Many open questions were discussed throughout the dissertation, but in the following paragraphs, I detail a handful of particularly interesting avenues for further exploration.

Within Kipsigis, many open questions remain, including ones related to other ways of expressing epistemic modality in the language. For instance, as discussed in Chapter 2, the Kipsigis modal verb *raisi* is compatible with epistemic possibility and necessity claims, and so displays so-called “modal force flexibility”. This sort of flexibility is theoretically interesting, since possibility and necessity modals are thought to involve fundamentally different types of quantification over possible worlds (i.e. existential vs. universal quantification). While force flexible modals have been described in a number of languages (St’át’imcets, Matthewson et al. 2007; Rullmann et al. 2008; Davis et al. 2009, Gitksan, Peterson 2010; Matthewson 2016, Washo, Bochnak 2015, Nez Perce: Deal 2011, Nsyilxcen, Menzies 2013, Ecuadorian Siona, Jeretič 2021, Kinande, Newkirk 2022), modal force flexibility is generally associated with languages of the Americas and with small modal inventories. However, Kipsigis is an African language with a rich inventory of modal verbs and, in this way, the Kipsigis pattern is somewhat unexpected given current assumptions about modal force flexibility. Understanding the extent to which other empirical parallels of modal force flexibility hold in Kipsigis is an exciting area for future work.

Turning to the cross-Kalenjin picture, future research is needed to understand how the patterns reported in this dissertation for Kipsigis do or do not hold in other Kalenjin languages. This question is particularly relevant in the context of the meaning of the secondary suffix proposed in Chapter 3 and summarized above. As noted in Chapter 3, the complex nominal structure proposed for Kipsigis by Kouneli (2021) holds across Kalenjin languages—some of which continue to have a productive alternation between primary forms (i.e. nouns without the secondary suffix) and secondary forms (i.e. nouns with the secondary suffix). Within the Kalenjin literature, descriptions of the interpretation of these secondary forms are varied; some researchers label the secondary form a definite without providing semantic justification (Hollis 1909; Zwarts 2001), while others state that it can be a definite or an indefinite (Tucker & Bryan 1964; Creider & Creider 1989; Kouneli 2019). The choice functional analysis of the secondary suffix offered in this dissertation provides a way of reconciling these conflicting descriptions, since choice functions blur the line between definite and indefinite interpretations, but it is an open question to what extent secondary forms in other Kalenjin languages show the behaviors predicted for choice functional indefinites. A typological study of the interpretation of secondary forms across Kalenjin would provide interesting insights about the semantic analysis of these nouns.

Zooming out, there are also a number of open empirical and theoretical questions concerning indefinites and biased belief reports across languages. In particular, in Chapter 3, I followed Dawson (2020) in suggesting that existentially closed choice function variables are associated with ignorance effects, while free choice function variables are associated with knowledge effects. The connection between existential closure of a choice function variable and the type of epistemic effect requires more significant cross-linguistic study. A detailed, typological survey of

the epistemic effects associated with indefinites whose other properties (e.g. scope, specificity, referentiality) warrant a particular kind of choice functional analysis would shed light on this question. If the correlation between existential closure and ignorance type holds across a broader sample of languages, it would lend further support to the idea that epistemic effects provide another diagnostic for the presence or absence of existential closure of choice function variables. Furthermore, in Chapter 3, I followed Kratzer (1998) in assuming that free choice function variables behave differently than other free variables in requiring only speaker knowledge for them to be felicitously resolved. However, it is an important open question *why* choice function variables would behave differently from other variables in this respect.

Finally, Chapter 4 highlighted several areas for further research related to biased belief reports and the theoretical machinery used to analyze them. As noted at the end of Chapter 4, biased belief verbs have only recently begun to garner attention in the literature, and so the typology of these verbs across languages is an open question. While the discussion in Chapter 4 highlighted the ways in which the pattern seen with Kipsigis *par* is unique, it is entirely possible that there are other verbs with a similar empirical profile. Further cross-linguistic work on biased belief reports would provide valuable insight into this possibility. More theoretically, the analysis in Chapter 4 and summarized above relied on the relatively new theoretical mechanism of post-supposition. Beyond biased belief verbs, postsupposition has been used to model the semantic contribution of free relatives with *-ever* (Lauer 2009), as well as modified numerals like *exactly three boys* (Brasoveanu 2013). Going forward, an important area for further research concerns defining the key characteristics of postsupposition across its varied uses, to better understand what other semantic phenomena might warrant a postsuppositional analysis. The discussion of postsupposition in Glass (2022) is a step in this direction, but more work remains to be done.

Bibliography

- Aloni, Maria & Angelika Port. 2015. Epistemic indefinites and methods of identification. In Luis Alonso-Ovalle & Paula Menéndez-Benito (eds.), *Epistemic indefinites: Exploring modality beyond the verbal domain*, 117–140. Oxford: Oxford University Press.
- Aloni, Maria & Robert van Rooij. 2004. Free choice items and alternatives. In *Proceedings of the KNAW academy colloquium: Cognitive foundations of interpretation*, 5–6. Amsterdam: Edita KNAW.
- Alonso-Ovalle, Luis & Paula Menéndez-Benito. 2010. Modal indefinites. *Natural Language Semantics* 18. 1–31.
- Alonso-Ovalle, Luis & Paula Menéndez-Benito. 2013. Two views on epistemic indefinites. *Language and Linguistics Compass* 7(2). 105–122.
- Alonso-Ovalle, Luis & Paula Menéndez-Benito. 2017. Epistemic indefinites: On the content and distribution of the epistemic component. In Anna Arregui, María Luisa Rivero & Andrés Salanova (eds.), *Modality across syntactic categories*, 11–29. Oxford University Press.
- Alonso-Ovalle, Luis & Junko Shimoyama. 2014. Expressing ignorance in the nominal domain: Japanese *Wh-ka*. In Robert Santana-Labarge (ed.), *Proceedings of WCCFL 31*, 11–20. Somerville, MA: Cascadilla Proceedings Project.
- Anagnostopoulou, Elena. 2016. Clitic doubling and object agreement. In Susann Fischer & Mario Navarro (eds.), *Proceedings of the VII Nereus International Workshop*, 11–42. Konstanz: Fachbereich Sprachwissenschaft Universität Konstanz.
- Anvari, Amir, Mora Maldonado & Andrés Soria Ruiz. 2019. The puzzle of *reflexive belief construction* in Spanish. In *Proceedings of Sinn und Bedeutung 23*, vol. 1, 57–74. Universitat Autònoma de Barcelona, Bellaterra.
- Arregui, Ana, María Luisa Rivero & Andrés Salanova (eds.). 2017. *Modality across syntactic categories*. Oxford: Oxford University Press.
- Bakovič, Eric. 2000. *Harmony, dominance, and control*. New Brunswick, NJ: Rutgers University dissertation.
- Baroja, Tomàs, Peter Sikamoy & Daniel Partany. 1989. *Analytical grammar of the Pokot language*. Trieste: EUT Edizioni Università di Trieste.
- Bledin, Justin & Kyle Rawlins. 2016. Epistemic resistance moves. In *Proceedings of SALT 26*, 620–640. Linguistic Society of America.
- Bobaljik, Jonathan. 2012. *Universals in Comparative Morphology: Suppletion, Superlatives, and the*

Structure of Words. MIT Press.

- Bochnak, M. Ryan. 2015. Underspecified modality in Washo. *UBC Working Papers in Linguistics* 3.
- Bochnak, M. Ryan & Lisa Matthewson (eds.). 2015. *Methodologies in semantic fieldwork*. Oxford: Oxford University Press.
- Bossi, Madeline. 2021. *local > local is morphological. In Rachel Soo, Una Y. Chow & Sander Nederveen (eds.), *Proceedings of WCCFL 38*, 74–84. Somerville, MA: Cascadilla Proceedings Project.
- Bossi, Madeline. 2022. Higher order ignorance in Kipsigis epistemic indefinites. Presentation at Sinn und Bedeutung 27.
- Bossi, Madeline. 2023. Two types of ‘say’-complementation in Kipsigis. Presentation at WCCFL 41.
- Bossi, Madeline & Michael Diercks. 2019. V1 in Kipsigis: Head movement and discourse-based scrambling. *Glossa: a journal of general linguistics* 4(1).
- Branan, Kenyon & Michael Erlewine. 2022. ā-probing for the closest DP. *Linguistic Inquiry* 1–27.
- Brasoveanu, Adrian. 2013. Modified numerals as post-suppositions. *Journal of Semantics* 30(2). 155–209.
- Brasoveanu, Adrian & Donka Farkas. 2011. How indefinites choose their scope. *Linguistics and Philosophy* 34. 1–55.
- Bunyi, Grace. 1999. Rethinking the place of African indigenous languages in African education. *International Journal of Educational Development* 19(4). 337–350.
- Büring, Daniel. 2003. On D-trees, beans, and B-accent. *Linguistics and Philosophy* 26. 511–545.
- Cable, Seth. 2010. *The grammar of Q: Q-particles, Wh-movement, and pied piping*. Oxford: Oxford University Press.
- Carstens, Vicki. 2016. Delayed valuation: A reanalysis of goal features, “upward” complementizer agreement, and the mechanics of Case. *Syntax* 19(1). 1–42.
- Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6(4). 339–405.
- Chierchia, Gennaro. 2001. A puzzle about indefinites. In Carlo Cecchetto, Gennaro Chierchia & Maria Teresa Guasti (eds.), *Semantic interfaces: reference, anaphora and aspect*, 51–89. Stanford, California: CSLI.
- Chierchia, Gennaro. 2006. Broaden your views: Implicatures of domain widening and the “logicality” of language. *Linguistic Inquiry* 37(4). 535–590.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Massachusetts: M.I.T. Press.
- Colley, Justin & Dmitry Privoznov. 2020. On the topic of subjects: Composite probes in Khanty. In Mariam Asatryan, Yixiao Song & Ayana Whitmal (eds.), *Proceedings of NELS 50*, 111–124. University of Massachusetts GLSA.
- Collins, James N. 2019. Definiteness determined by syntax: A case study of Tagalog. *Natural Language & Linguistic Theory* 37. 1367–1420.
- Creider, Chet & Jane Creider. 1989. *A grammar of Nandi*. Hamburg: Helmut Buske Verlag.

- Creider, Chet A. 1982. *Studies in Kalenjin nominal tonology*. Dietrich Reimer.
- Davis, Henry, Lisa Matthewson & Hotze Rullmann. 2009. 'out of control' marking as circumstantial modality in St'át'imcets. In Lotte Hogeweg, Helen de Hoop & Andrej Malchukov (eds.), *Cross-linguistic semantics of tense, aspect, and modality*, 205–244. Amsterdam: John Benjamins.
- Dawson, Virginia. 2018. A new kind of epistemic indefinite. In Uli Sauerland & Stephanie Solt (eds.), *Proceedings of Sinn und Bedeutung 22*, vol. 1, 349–366. ZAS.
- Dawson, Virginia. 2020. *Existential quantification in Tiwa: Disjunction and indefinites*. Berkeley, CA: UC Berkeley dissertation.
- Dawson, Virginia & Peter Jenks. 2021. Documenting and describing (in)definiteness. Unpublished manuscript.
- Dayal, Veneeta. 2004. Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy* 27(4). 393–450.
- Dayal, Veneeta & Yağmur Sağ. 2020. Determiners and bare nouns. *Annual Review of Linguistics* 6. 173–194.
- Deal, Amy Rose. 2011. Modals without scales. *Language* 87(3). 559–585.
- Deal, Amy Rose & Julia Nee. 2018. Bare nouns, number, and definiteness in Teotitlán del Valle Zapotec. *Proceedings of Sinn und Bedeutung 21*(1). 317–334.
- Diercks, Michael. 2013. Indirect agree in Lubukusu complementizer agreement. *Natural Language & Linguistic Theory* 31(2). 357–407.
- Diercks, Michael & Meghana Rao. 2019. Upward-oriented complementizer agreement with subjects and objects in Kipsigis. In Emily Clem, Peter Jenks & Hannah Sande (eds.), *Theory and description in African linguistics: Selected papers from the 47th Annual Conference on African Linguistics*, 369–393. Berlin: Language Science Press.
- Dimmendaal, Gerrit. 2000. Number marking and noun categorization in Nilo-Saharan languages. *Anthropological Linguistics* 42(2). 214–261.
- Dimmendaal, Gerrit. 2012. Metrical structures: A neglected property of Nilotic (and other African language families). *Studies in Nilotic Linguistics* 5. 1–26.
- Dimmendaal, Gerrit. 2017. Areal contact in Nilo-Saharan. In Raymond Hackley (ed.), *Cambridge handbook of areal linguistics*, 446–470. Cambridge: Cambridge University Press.
- Dimmendaal, Gerrit & F. K. Erhard Voeltz. 2008. Africa. In Christopher Moseley (ed.), *Encyclopedia of the world's endangered languages*, 579–634. Routledge.
- Döring, Sophia. 2016. *Modal particles, discourse structure, and common ground management: Theoretical and empirical aspects*. Berlin, Germany: Humboldt-Universität zu Berlin dissertation.
- Driemel, Imke & Maria Kouneli. 2022a. C-agree is local subject-verb agreement in Kipsigis. Unpublished manuscript.
- Driemel, Imke & Maria Kouneli. 2022b. A dedicated topic position in Kipsigis. In Henrison Hsieh (ed.), *Proceedings of Triple A 8*, 1–20. NUS ScholarBank.
- Driemel, Imke, Ahmet Bilal Özdemir & Marie-Luise Popp. 2020. Direction towards Person: Canonical inverse and reverse PCC in Adyghe. Unpublished manuscript.
- Drummond, Emily. 2023. *Clause structure and ergativity in Nukuoro*. Berkeley, CA: UC Berkeley

dissertation.

- Eberhard, David M., Gary F. Simons & Charles D. Fennig. 2021. Kipsigis. In *Ethnologue: Languages of the world*, Dallas, TX: SIL International 24th edn.
- Ehret, Christopher. 1971. *Southern Nilotic history: Linguistic approaches to the study of the past*. Evanston, IL: Northwestern University Press.
- Farkas, Donka. 2002. Varieties of indefinites. In Brendan Jackson (ed.), *Proceedings of SALT 12*, 59–83. Ithaca, NY: Cornell University.
- Farkas, Donka & Kim Bruce. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27. 81–118.
- Fink, Teresa. 2005. *Attitudes toward languages in Nairobi*: University of Pittsburgh MA thesis.
- von Fintel, Kai & Anthony S. Gillies. 2008. An opinionated guide to epistemic modality. In T. S. Gendler & J. Hawthorne (eds.), *Oxford studies in epistemology*, vol. 2, 32–62. Oxford: Oxford University Press.
- Fodor, Janet Dean & Ivan Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5. 355–398.
- Franciscar, Kamuren & Bartoo Phylis. 2012. The morphosyntactic differences among Kalenjin dialects: An analysis of Kipsigis, Tugen, and Pokot. *Research on Humanities and Social Sciences* 2(7). 56–65.
- Frey, Werner. 2004. A medial topic position for German. *Linguistische Berichte* 198. 153–190.
- Fălăuș, Anamaria. 2014. (partially) free choice of alternatives. *Linguistics and Philosophy* 37. 121–173.
- Geist, Ljudmila. 2010. Bare singular NPs in argument positions: Restrictions on indefiniteness. *International Review of Pragmatics* 2. 191–227.
- Geist, Ljudmila & Edgar Onea. 2007. Specificity and implicatures. In Maria Aloni, Paul Dekker & Floris Roelofsen (eds.), *Proceedings of the Sixteenth Amsterdam Colloquium*, 109–114. Amsterdam: ILLC.
- Giannakidou, A. & Josep Quer. 2011. Against universal free choice: Free choice and referentially vague indefinites in Greek, Catalan, and Spanish. Unpublished manuscript.
- Glass, Lelia. 2022. The negatively biased Mandarin belief verb *yǐwéi*. *Studia Linguistica*.
- Grice, H. Paul. 1989. Logic and conversation. In *Studies in the way of words*, Cambridge, MA: Harvard University Press.
- Grosz, Patrick. 2016. Information structure and discourse particles. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford Handbook of Information Structure*, 336–358. Oxford University Press.
- Hacquard, Valentine & Jeffrey Lidz. 2018. Children’s attitude problems: Bootstrapping verb meaning from syntax and pragmatics. *Mind & Language* 1–24.
- Hall, Beatrice, R. M. R. Hall, Martin Pam, Amy Myers, Stephen A. Antell & Godfrey K. Cheron. 1974. African vowel harmony systems from the vantage point of Kalenjin. *Africa und Übersee* 57. 241–267.
- Halle, Morris & Jean-Roger Vergnaud. 1981. Harmony processes. In Klein Wolfgang & Willem J. M. Levelt (eds.), *Crossing the boundaries in linguistics: Studies presented to Manfred Bierwisch*,

- 1–22. Dordrecht: Springer.
- Harizanov, Boris. 2014. Clitic doubling at the syntax-morphophonology interface: A-movement and morphological merger in Bulgarian. *Natural Language and Linguistic Theory* 32(4). 1033–1088.
- Heycock, Caroline & Anthony Kroch. 1999. Pseudocleft connectedness: Implications for the LF interface level. *Linguistic Inquiry* 30(3). 365–397.
- Hirschberg, Julia. 1985. *A theory of scalar implicature*. Philadelphia, PA: University of Pennsylvania dissertation.
- Hollis, Alfred Claud. 1909. *Nandi: Their language and folklore*. Negro Universities Press.
- Hsiao, Pei-Yi K. 2017. On counterfactual attitudes: A case study of Taiwanese Southern Min. *Lingua Sinica* 3(4).
- Jayez, Jacques & Lucia Tovenà. 2006. Epistemic determiners. *Journal of Semantics* 23(3). 217–250.
- Jayez, Jacques & Lucia Tovenà. 2007. Evidentiality and determination. In Atle Grønn (ed.), *Proceedings of Sinn und Bedeutung 12*, Oslo: ILOS.
- Jeretič, Paloma. 2021. *Neg-raising modals and scaleless implicatures*. New York: New York University dissertation.
- Jerono, Prisca. 2012. *Tugen word order: A Minimalist perspective*. European Scientific Institute.
- Jerono, Prisca, S. Hillary, C. Andrew & O.N.J.A. 2012. A unified orthography for kalenjin languages of Kenya. In Casas, South Africa.
- Jolin, Armel. 2023. A non-polar, concatenative account of Kipsigis across-the-board tonal polarity. Talk at Old World Conference of Phonology.
- Kagan, Olga. 2011. On speaker identifiability. *Journal of Slavic Linguistics* 19(1). 47–84.
- Karttunen, Lauri. 1974. Presupposition and linguistic context. *Theoretical Linguistics* 1(1-3). 181–194.
- Kierstead, Gregory. 2013. Shifted indexicals and conventional implicature: Tagalog *akala* ‘falsely believe’. Talk presented at SALT 23, University of California, Santa Cruz.
- Kiparsky, Paul & Carol Kiparsky. 1970. Fact. In Manfred Bierwisch & Karl E. Heidolph (eds.), *Progress in linguistics*, 143–173. The Hague and Paris: Mouton.
- KNBS. 2009. Kenya population and housing census.
- KNBS. 2019. Kenya population and housing census.
- Kouneli, Maria. 2019. *The syntax of number and modification: An investigation of the Kipsigis DP*. New York, NY: New York University dissertation.
- Kouneli, Maria. 2021. Number-based noun classification: The view from Kipsigis. *Natural Language & Linguistic Theory* 39. 1195–1251.
- Kouneli, Maria. 2022. Inflectional classes in Kipsigis. *Glossa* 47(1).
- Kouneli, Maria & Yining Nie. 2021. Across-the-board tonal polarity in Kipsigis: Implications for the morphology-phonology interface. *Language*.
- Kramer, Ruth. 2014. Clitic doubling or object agreement: The view from Amharic. *Natural Language and Linguistic Theory*.
- Kratzer, Angelika. 1998. Scope or pseudoscope? Are there wide-scope indefinites? In Susan

- Rothstein (ed.), *Events and grammar*, 163–196. Dordrecht: Kluwer Academic Publishers.
- Kratzer, Angelika. 2003. A note on Choice Functions in context. Unpublished manuscript.
- Kratzer, Angelika & Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. In Yukio Otsu (ed.), *Proceedings of the Tokyo conference on psycholinguistics*, vol. 3, 1–25. Tokyo: Hituzi Syobo.
- Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3-4). 243–276.
- Lauer, Sven. 2009. Free relatives with -ever: Meaning and use. Unpublished manuscript.
- Levin, Juliette & Diane Massam. 1985. Surface ergativity: Case/Theta relations reexamined. In *Proceedings of NELS 15*, vol. 15 1, 286–301. GLSA UMass Amherst.
- Lewis, David. 1979. Attitudes de dicto and de se. *Philosophical Review* 88. 513–543.
- Little, Carol Rose. 2020. *Mutual dependencies of nominal and clausal syntax in Ch’ol*. Ithaca, NY: Cornell University dissertation.
- Local, John & Ken Lodge. 2004. Some auditory and acoustic observations on the phonetics of [ATR] harmony in a speaker of a dialect of Kalenjin. *Journal of the International Phonetic Association* 34(1). 1–16.
- Lodge, Ken. 1995. Kalenjin phonology and morphology: A further exemplification of underspecification and non-destructive phonology. *Lingua* 96(1). 29–43.
- Matthewson, Lisa. 1999. On the interpretation of wide-scope indefinites. *Natural Language Semantics* 7(1). 79–134.
- Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70. 369–415.
- Matthewson, Lisa. 2016. Modality. In *Cambridge handbook of formal semantics*, 525–559. Cambridge University Press.
- Matthewson, Lisa, Henry Davis & Hotze Rullmann. 2007. Evidentials as epistemic modals: Evidence from st’át’imcets. *Linguistic Variation Yearbook* 7. 201–245.
- Mayol, Laia & Elena Castroviejo. 2013. How to cancel an implicature. *Journal of Pragmatics* 50(1). 84–104.
- Menzies, Stacey. 2013. *Nsyilxcen modality: Semantic analysis of epistemic modality*. Vancouver: University of British Columbia MA thesis.
- Moroney, Mary. 2021. Updating the typology of definiteness: Evidence from bare nouns in Shan. *Glossa: a journal of general linguistics*.
- Muthwii, Margaret Jepkirui. 2004. Language of instruction: A qualitative analysis of the perceptions of parents, pupils and teachers among the Kalenjin in Kenya. *Language Culture, and Curriculum* 17(1). 15–32.
- Nevins, Andrew. 2010. *Locality in vowel harmony*, vol. 55 Linguistic Inquiry Monographs. Cambridge, MA: MIT Press.
- Nevins, Andrew. 2011. Multiple agree with clitics: person complementarity vs. omnivorous number. *Natural Language and Linguistic Theory* 29(4). 939–971.
- Newkirk, Lydia. 2022. *Be flexible, but not too flexible: Limited variable-force modals in Kinande*

- and the typology of modal force*. New Brunswick, NJ: Rutgers University dissertation.
- O'Hagan, Zachary. 2021. Morphosyntax and semantics of psych-predicates in Caquinte. *Cadernos de Etnolingüística* 1–38.
- Owusu, Augustina. 2019. A closer look at the Akan determiner *bi*: An epistemic indefinite analysis. In Samson Lotven, Silvina Bongiovanni, Phillip Weirich, Robert Botne & Samuel Gyasi Obeng (eds.), *African linguistics across the disciplines: Selected papers from the 48th Annual Conference on African Linguistics*, 257–280. Berlin: Language Science Press.
- Peterson, Tyler. 2010. *Epistemic modality and evidentiality in Gitksan at the semantics-pragmatics interface*: University of British Columbia dissertation.
- Potsdam, Eric & Maria Polinsky. 2011. Questions and word order in Polynesia. In C. Moyse-Faurie & Joachim Sabel (eds.), *Topics in Oceanic morphosyntax*, 121–153. Berlin: Mouton de Gruyter.
- Potts, Christopher. 2014. Presupposition and implicature. In Shalom Lappin & Chris Fox (eds.), *Handbook of contemporary semantic theory*, 168–202. Wiley Blackwell 2nd edn.
- Preminger, Omer. 2009. Breaking agreements: Distinguishing agreement and clitic doubling by their failures. *Linguistic Inquiry* 40(4). 618–666.
- Pylkkänen, Liina. 2008. *Introducing arguments* (Linguistic Inquiry Monographs 49). Cambridge, MA; London, England: MIT Press.
- Reinhart, Tanya. 1997. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20(4). 335–397.
- Renans, Agata. 2018. Two types of choice-functional indefinites: Evidence from Ga (Kwa). *Topoi* 37. 405–415.
- Repp, Sophie. 2013. Common Ground management: Modal particles, illocutionary negation and *Verum*. In Daniel Gutzmann & Hans-Martin Gärtner (eds.), *Expressives and beyond: Explorations of conventional non-truth-conditional meaning*, Oxford: Oxford University Press.
- Republic of Kenya. 2010. *The Constitution of Kenya*.
- Rett, Jessica. 2020. Manner implicatures and how to spot them. *International Review of Pragmatics* 12. 44–79.
- Rolle, Nicholas. 2018. *Grammatical tone: Typology and theory*. Berkeley, CA: UC Berkeley dissertation.
- Romero, Maribel. 2015. High negation in subjunctive conditionals and polar questions. In Eva Csipak & Hedde Zeijlstra (eds.), *Proceedings of Sinn und Bedeutung 19*, University of Göttingen.
- Rottland, Franz. 1982. *Die südnilotischen Sprachen*, vol. 7. Berlin: D. Reimer.
- Rullmann, Hotze, Lisa Matthewson & Henry Davis. 2008. Modals as distributive indefinites. *Natural Language Semantics* 16. 317–357.
- Saltin, Bengt, Henrik Larsen, Nicolàs Terrados, Jens Bangsbo, Thomas Bak, Chang Keun Kim, Jan Svedenhag & C. J. Rolf. 1995. Aerobic exercise capacity at sea level and at altitude in Kenyan boys, junior and senior runners compared with Scandinavian runners. *Scandinavian journal of medicine & science in sports* 5(4). 209–221.
- Sande, Hannah. To Appear. Is grammatical tone item-based or process-based? *Phonology*.
- Sauerland, Uli. 1998. *On the making and meaning of chains*. Cambridge, Massachusetts: MIT

dissertation.

- Sauerland, Uli. 2000. Two structures for English restrictive relative clauses. In *Proceedings of the Nanzan GLOW*, 351–366. Nagoya: Nanzan University.
- Sauerland, Uli. 2003. Unpronounced heads in relative clauses. In *The interfaces: Deriving and interpreting omitted structures*, 205–226. Amsterdam: John Benjamins.
- Schwarz, Bernhard. 2001. Two kinds of long-distance indefinites. In R. van Rooy & M. Stokhof (eds.), *Proceedings of the 13th Amsterdam Colloquium*, 192–197. Amsterdam: ICCL.
- Schwarz, Bernhard. 2011. Long distance indefinites and choice functions. *Language and Linguistics Compass* 5(12). 880–897.
- Schwarz, Florian. 2013. Two kinds of definites cross-linguistically. *Language and Linguistics Compass* 7(10). 534–559.
- Scott, Tessa. 2021. Formalizing two types of mixed A/ \bar{A} movement. Unpublished manuscript.
- Siegel, Laura. 2009. Mood selection in Romance and Balkan. *Lingua* 119. 1859–1882.
- Simons, Mandy. 2007. Observations on embedding verbs, evidentiality, and presupposition. *Lingua* 117(6). 1034–1056.
- Stalnaker, Robert. 1974. Pragmatic presuppositions. In Milton Munitz & Peter Unger (eds.), *Semantics and philosophy*, 197–213. New York, NY: New York University Press.
- Stalnaker, Robert. 1978. Assertion. In Peter Cole (ed.), *Syntax and semantics*, vol. 9, 315–322. New York: Academic Press.
- Toweett, Taaita. 1975. *Kalenjin nouns and their classification with notes on phonology and a noun list appendix*. Nairobi: University of Nairobi MA thesis.
- Toweett, Taaita. 1979. *A study of Kalenjin linguistics*. Kenya Literature Bureau.
- Tucker, Archibald N. & Margaret A. Bryan. 1964. Noun classification in Kalenjin: Nandi Kipsigis. *African Language Studies* 5; 6. 192–247; 117–187.
- Van Urk, Coppe. 2015. *A uniform syntax for phrasal movement: A Dinka Bor case study*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Šimík, Radek. 2015. Epistemic indefinites under epistemic modals in Czech. In G. Zybatow, P. Biskup, M. Guhl, C. Hurtig, O. Mueller-Reichau & M. Yastrebova (eds.), *Slavic grammar from a formal perspective: Proceedings of the 10th Anniversary Conference on Formal Description of Slavic Languages*, 425–442. Frankfurt: Peter Lang.
- Willett, Thomas. 1988. A cross-linguistic survey of the grammaticization of evidentiality. *Studies in Language* 12. 51–97.
- Winter, Yoad. 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20. 399–467.
- Woolford, Ellen. 2008. Active-stative agreement in Lakota: Person and number alignment and portmanteau formation. Unpublished manuscript.
- Yanovich, Igor. 2005. Choice-functional series of indefinite pronouns and Hamblin semantics. In Effi Georgala & Jonathan Howell (eds.), *Proceedings of SALT XV*, 309–326. Ithaca, NY.
- Yuan, Michelle. 2021. Diagnosing object agreement vs. clitic doubling: An Inuit case study. *Linguistic Inquiry* 52(1). 153–179.

- Zamparelli, Roberto. 2007. On singular existential quantifiers in Italian. In Ileana Comorovski & Klaus Heusinger (eds.), *Existence: Semantics and Syntax*, 293–328. Dordrecht: Springer.
- Zwarts, Joost. 2001. Number in Endo-Marakwet. In *Advances in Nilo-Saharan linguistics: Proceedings of the 8th Nilo-Saharan linguistics colloquium*, 281–294. University of Hamburg.
- Zwarts, Joost. 2004. The phonology of Endo: A southern Nilotic language of Kenya. In *LINCOM Studies in African Linguistics*, vol. 59, LINCOM GmbH.