Elements of (in)definiteness and binding: A Mayan perspective

Justin Royer

Department of Linguistics
McGill University
July 2022

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

À ma mère, Johanne Pelletier, pour tout.

Abstract

This thesis explores topics pertaining to the syntax and semantics of nominal expressions, with a focus on definite, indefinite, demonstrative, and pronominal elements. The data are drawn from original work on Mayan languages, especially Chuj, an underdocumented language predominantly spoken in Guatemala and Mexico.

The first part zooms in on the elements that play a role in the syntactic composition of the extended nominal domain, and the semantic and pragmatic contributions that result from combining these elements together. By showcasing great complexity within the extended nominal domain, I argue that Chuj is particularly illuminating for topics which have been at the core of debates in the syntax and semantics of DPs, such as the encoding of definiteness versus indefiniteness, the internal syntax and semantics of demonstratives, the nature of pronouns, and the ways in which the contextual domain of nominal expressions is implicitly or explicitly delimited. A recurring theme of the thesis is that, by virtue of being radically decompositional, Chuj often challenges pre-existing assumptions about the primitivity of certain linguistic expressions. Instead, the Chuj data align with an increasing number of works that argue that traditional notions, such as definiteness, come in different guises (e.g., Schwarz 2009; Arkoh and Matthewson 2013; Jenks 2018; Jenks and Konate to appear), or that these different notions arise as a result of a decomposition of functional heads within the nominal domain (e.g., Déchaine and Wiltschko 2002; Leu 2008; Simonenko 2014; Coppock and Beaver 2015; Hanink 2018; Ahn 2019).

The second part of the thesis zooms out of the internal syntax of nominal expressions and into the distribution of covalued nominals within sentences. This part also provides data on Ch'ol, another Mayan language. I show that while Ch'ol behaves entirely as expected given the Binding Conditions (Reinhart 1983, Chomsky 1986), Chuj appears to consistently tolerate violations of Condition C, privileging linear precedence as the determining factor in the distribution of R-expressions and pronouns. The Chuj data thus initially seem to cast doubt on a long tradition to treat the Binding Conditions as universal. I argue that the difference between Chuj and Ch'ol can be largely explained if, contrary to Ch'ol, Chuj exhibits "high-absolutive" syntax, independently proposed to account for a number of syntactic phenomena in a subset of Mayan languages (Coon, Mateo Pedro, and Preminger 2014; Coon, Baier, and Levin 2021). High-absolutive syntax creates configurations in which the internal argument asymmetrically c-commands the external argument, bleeding otherwise expected binding relations from the external argument into the internal argument. The violations of Condition C in Chuj are thus only apparent. The outcome is that despite initial evidence to doubt the universality of the Binding Conditions, a universalist approach can be maintained.

Abrégé

Cette thèse présente une étude sur la syntaxe et la sémantique des expressions nominales, avec un accent sur les éléments définis, indéfinis, démonstratifs et pronominaux. Les données proviennent de travaux originaux sur les langues mayas, en particulier le chuj, une langue sous-documentée parlée principalement au Guatemala et au Mexique.

La première partie concerne les éléments qui forment la syntaxe du domaine nominal étendu, ainsi que les contributions sémantiques et pragmatiques qui résultent de leurs combinaisons. En mettant de l'avant la grande complexité du domaine nominal chuj, je soutiens que l'analyse de cette langue apporte un éclairage nouveau sur certains sujets au cœur de la linguistique théorique, tels que la différence entre les expressions définies et indéfinies, la composition interne des démonstratifs, la nature des pronoms, et la restriction du domaine contextuel. Un thème récurrent dans cette étude est qu'en démontrant une décomposition fine, le chuj remet en question certaines suppositions préexistantes sur les composantes de base des expressions linguistiques. En effet, les données du chuj concordent avec un nombre croissant de travaux qui soutiennent que certaines notions communes, comme la définitude, peuvent se présenter sous différentes formes (p. ex. Schwarz 2009; Arkoh et Matthewson 2013; Jenks 2018; Jenks et Konate à paraître), ou encore que ces différentes notions nécessitent une décomposition des éléments fonctionnels formant le domaine nominal (p. ex. Déchaine et Wiltschko 2002; Leu 2008; Simonenko 2014; Coppock et Beaver 2015; Hanink 2018; Ahn 2019).

La deuxième partie porte sur la distribution des expressions nominales à l'intérieur de phrases complètes. Cette partie analyse également des données du ch'ol, une autre langue maya. Je montre que, bien que le ch'ol se comporte comme prévu conformément aux Principes de la Théorie du Liage (PTL) (Reinhart 1983, Chomsky 1986), le chuj semble tolérer des violations du Principe C, en privilégiant souvent la préséance linéaire comme facteur déterminant pour la distribution des expressions nominales. À première vue, les données du chuj semblent donc mettre en doute une tradition qui consiste à traiter les PTL comme universels. Je soutiens que la différence entre le chuj et le ch'ol s'explique en grande partie du fait que le chuj, contrairement au ch'ol, présente une syntaxe aux absolutifs élevés, qui a été proposée indépendamment pour expliquer un certain nombre de phénomènes morphosyntaxiques dans un sous-ensemble des langues mayas (Coon et al. 2014; Coon et al. 2021). Ce type de syntaxe crée des phrases dans lesquelles l'argument interne c-commande l'argument externe, brisant certaines relations de liage autrement attendues de l'argument externe envers l'argument interne. Les violations du Principe C ne sont donc qu'apparentes. Par conséquent, malgré les doutes initiaux soulevés par le chuj sur l'universalité des PTL, une approche universaliste peut être maintenue.

Resumen

La presente tesis explora temas relacionados con la sintaxis y semántica de expresiones nominales, centrándose en los elementos definidos, indefinidos, demostrativos y pronominales. Los datos provienen de trabajo original sobre lenguas mayas, especialmente el Chuj, una lengua que se habla principalmente en Guatemala y México.

La primera parte se centra en los elementos que componen la sintáctica del dominio nominal extendido, y en las contribuciones semánticas resultantes al combinar estos elementos. Al mostrar una gran complejidad, el dominio nominal del Chuj es esclarecedor para temas que han sido el centro de debates en la sintaxis y la semántica nominal, como la codificación de la definitud y de la indefinitud, la sintaxis y semántica interna de los demostrativos, la naturaleza de los pronombres y las formas en que el dominio contextual de las expresiones nominales está implícitamente o explícitamente delimitado. Un tema recurrente de esta tesis es que, en virtud de mostrar una descomposición nominal radical, el Chuj a menudo desafía suposiciones preexistentes sobre la primitividad de ciertas expresiones lingüísticas. Por lo tanto, los datos se alinean con un número creciente de trabajos que argumentan que algunas nociones básicas, como la definitud, se presentan en diferentes formas (Schwarz 2009; Arkoh y Matthewson 2013; Jenks 2018; Jenks y Konate en prensa), o que diferentes nociones semánticas surgen de una descomposición de elementos funcionales dentro de la frase nominal (Déchaine y Wiltschko 2002; Leu 2008; Simonenko 2014; Coppock y Beaver 2015; Hanink 2018; Ahn 2019).

La segunda parte se aleja de la sintaxis interna de las expresiones nominales y se enfoca en la distribución de los nominales coreferenciales dentro de oraciones. Esta parte también proporciona datos del Ch'ol, otra lengua maya. Muestro que, mientras que el Ch'ol se comporta totalmente como se espera dados los Principios de Ligamiento (PL) (Reinhart 1983, Chomsky 1986), el Chuj parece tolerar violaciones del Principio C, privilegiando frecuentemente la precedencia linear como factor determinante en la distribución de los nominales. Los datos del Chuj parecen, por lo tanto, poner en duda una larga tradición de tratar a los PL como universales. Propongo que la diferencia entre el Chuj y el Ch'ol puede explicarse en gran medida si solo el Chuj exhibe una sintaxis absolutiva-alta, propuesta de forma independiente para explicar una serie de fenómenos morfosintácticos en un subconjunto de lenguas mayas (Coon et al. 2014; Coon et al. 2021). La sintaxis absolutiva-alta crea configuraciones en las que el objeto manda-c ('c-command') al sujeto, desviando las relaciones de ligamiento sintáctico esperadas desde el objeto hacia el sujeto. Entonces, las violaciones del Principio C en Chuj son simplemente aparentes. El desenlace es que, a pesar de pruebas iniciales para dudar de la universalidad de los PL, no solo se puede mantener su universalidad, sino que el Chuj la apoya.

Acknowledgements

This thesis sees light after more than eight years of studies, life, and absolute happiness in Montreal. A mountain of irreplaceable individuals have contributed to this journey, both within and outside academia. I'm well aware of the privilege I have had to be surrounded by such an extraordinary network of collaborators, mentors, teachers, colleagues, friends, and family. The following list is by no means exhaustive.

First, I would like to extend my deepest gratitude to the Chuj speakers, consultants, friends, and collaborators whose generosity and linguistic insights are reflected throughout the entirety of this thesis. *¡Yuj wal yos t'ayex hemasanil!* La primera persona a quien tengo que agradecer es sin duda a Matal (Magdalena) Torres. He tenido el enorme placer de conocer a Matal desde hace más de seis años, y ha sido la colaboradora más amable, sabia y paciente que jamás hubiera podido imaginar. Agradezco no solo su disposición a compartir su maravillosa lengua y cultura conmigo, sino también su amistad, por la que guardo un inmenso cariño. ¡Gracias Magda! Tres personas adicionales que han contribuido de manera significativa con los datos del chuj abordados en esta tesis son Matin (Mateo) Pablo, Elsa Velasco y Yun Torres. Les agradezco infinitamente por haber colaborado conmigo en muchas ocasiones a lo largo de los últimos seis años. También agradezco a Elsa y a Yun por su constante generosidad y hospitalidad al siempre acogerme en su hogar. Por su amistad y voluntad de compartir su idioma conmigo en los últimos años, también agradezco a Reinalda Domingo, Pedro (Petul) Felipe Gómez, Xuwan Gómez, Diego Neper López García, Estela Pablo, Santiago Pablo, Alejandra Torres, Ana Torres, Agenor Torres País, Lena Torres, Mach'ol Torres, Mekel Torres, Miguel Torres, Petul Torres, Petul (Tigo) Torres País, Rogelio Torres, Xun Torres Velasquez, Xuwan Torres, Ana Velasco, y Heb'in Velasco. También quiero agradecer a Gilda Ixem Cum y a su familia por su amistad y por haberme ofrecido la bienvenida en su hogar en el encantador municipio de Patzún.

I am also greatly indebted to the most qualified team of advisors I could have dreamed of having: Jessica Coon, Luis Alonso-Ovalle and Aron Hirsch. I have no doubt that the content of this thesis and many of its proposals fall in great part as an immediate outcome of their guidance and wisdom.

Jessica, more than anyone, has contributed to my academic development. Her broad and profound knowledge of linguistics, her commitment to the field, her humanity, and the confidence she instills in her students, are a mere sample of the reasons that make her an exemplary advisor. From providing advice on putting together handouts, to planning a fieldwork trip to Guatemala, to discussing deep theoretical topics in syntax, Jessica has

been present in my journey as a grad student from day one to the end. She is an all-round inspiration, both within and outside linguistics. Thank you for everything Jessica!

Luis has also been an essential source of guidance during my studies. His enthusiasm for linguistics and his admirable positivity are contagious. He also has the *magical* ability to make excessively difficult topics in semantics seem totally accessible and understandable, which is one of the many qualities that have made him an excellent mentor. I am thankful for all the work we have conducted together throughout the years, and keep fond memories of difficult—and fun!—puzzle solving sessions about *komon* and *yalnhej*! Thank you Luis for your constant confidence in me and for your mentorship.

I was lucky to coincide with Aron during his postdoc at McGill from 2017–2019. After a semester of meetings related to a seminar he was co-teaching with Luis, I asked him whether he'd like to co-supervise my first evaluation paper. I have not regretted it. Aron has been a key mentor for me and has taught me much of what I know about the topics discussed in this thesis. A master of narratives and storytelling, he has also taught me the art of explaining one's work as clearly and efficiently as possible. I am very grateful for all the generous and thought-provoking meetings together—which were so entertaining they sometimes lasted half a day(!)—and for his overall presence and advising throughout my studies. Thank you Aron!

I have also had the privilege to receive excellent mentorship from faculty members outside of my thesis committee, which have greatly influenced my path as a student. I especially thank Michael Wagner for his guidance and wisdom on my second evaluation paper at McGill and for his rigorous feedback as internal examiner on the thesis. At Concordia, I thank Alan Bale for co-supervising my honours thesis, and Dana Isac for supervising me on the very first serious research project I undertook in linguistics.

I would also like to express my gratitude to Lisa Matthewson, who I was fortunate to have as external examiner on this thesis. Lisa's comments and questions were immensely clear and insightful, and have contributed substantially to this work. Thank you Lisa!

I have loved every moment of my experience at the Linguistics Department at McGill. I have benefited from conversations and courses provided by many faculty members, particularly Meghan Clayards, Brendan Gillon, Heather Goad, Martina Martinović, Bernhard Schwarz, Junko Shimoyama, and Lisa Travis. I also benefited from conversations and friendships from graduate students within the department throughout the years, particularly Jason Borga, Amy Bruno, Francesco Gentile, William Johnston, Masashi Harada, Henrison Hsieh, Willie Myers, Mathieu Paillé, Jonny Palucci, Clint Parker, Liz Smeets, Michaela Socolof, James Tanner, and Connie Ting. I am also very grateful to Giuliana Panetta and Andria De Luca for being amazing administrators, and for their constant support, laughs, and casual conversations over the past five years.

My first three years of doctoral studies at McGill were funded by the Social Sciences and Humanities Research Council of Canada via a Joseph-Armand Bombardier doctoral scholarship. I am very grateful for the opportunities provided by this funding. I also thank Mitacs and the Mexico Secretariat of Public Education, as well as McGill University, for funding fieldwork trips to Guatemala and Mexico.

Before starting at McGill, I was trained in linguistics at Concordia University. My extremely positive experience at Concordia, and the mentors I met during my time there, were what incited me to pursue graduate studies in the first place. Particularly, I thank

Dana Isac and Charles Reiss for their guidance and friendship throughout the years, and for their devotion to the training of future linguists and to the field of linguistics more generally. I also thank Alan Bale, Laura Grestenberger, Madelyn Kissock, and Grace Kuo.

I have benefited from conversations and meetings with many members of the linguistics community. These include Maayan Abenina-Adar, Judith Aissen, Scott Ander-Bois, Nico Baier, Ryan Bennett, Bronwyn Bjorkman, Telma Can Pixabaj, Ivano Caponigro, Julien Carrier, Emily Clem, Lauren Clemens, Richard Compton, Silvio Cruschina, Emily Elfner, Samuel Herrera-Castro, Nick Hopkins, Peter Jenks, Tom Leu, Yoann Léveillé, B'alam Mateo-Toledo, Lisa Matthewson, Keir Moulton, Heather Newell, Jaime Pérez-González, Omer Preminger, Rodrigo Ranero, Betsy Ritter, Andrés Salanova, Tessa Scott, Sigwan Thivierge, Mireille Tremblay, Hugo Hector Vázquez López, Morelia Vázquez Martínez, and Roberto Zavala Maldonado. I also thank my awesome linguistics colleagues and friends, Dan Brodkin, Cristina Buenrostro, Elizabeth Carolan, Paulina Elias, Aurore Gonzalez, Robert Henderson, Carol-Rose Little, Pedro Mateo Pedro, Mary Moroney, Willie Myers, and Matal Torres for collaborating with me on a number of different projects throughout the years.

Mes ami.e.s ont été une source nécessaire d'énergie tout au long de cette aventure. Merci entre autres à Bianca, María, Covadonga, Esteban, Vanessa, Mathieu, Karine, Maya, Miriam, Renée et Jean. A special thanks to Stella and her cat Buddy, whose feline company certainly made dissertation writing easier.

Enfin et surtout : je suis très conscient que cette thèse voit le jour aujourd'hui, en grande partie, grâce à ma famille. Ma reconnaissance envers vous est infinie. Merci avant tout à Camilo pour son appui indispensable, sa patience et son amour ¡Te amo Milo! Merci à mi suegrita Meridalba, qui a toujours été à l'affût des mes progrès académiques, y también a mis cuñados queridos Joaquín y Marianly. Merci à mon frère Alexandre, ma belle-sœur Marie-Ève et aux adorables Liam et Léa. Merci également à Daniel et à mes grands-parents Rosemary (thanks Grandma!), Fernand, Denise et Thérèse. Pour finir, je remercie de tout cœur mes parents, Alain et Johanne, pour votre soutien permanent et pour votre amour inconditionnel. Je vous aime.

List of abbreviations

*	ungrammatical construction
#	infelicitous construction
??	degraded construction
✓	acceptable/felicitous
✗	acceptable/infelicitous
( )	optional segment
*( )	non-optional segment
1	first person
2	second person
3	third person
‖	prosodic break
A	"set A" (ergative/possessive) markers
ADV	verb stem internal adverb
AP	antipassive suffix
AF	agent focus suffix or prefix
AG	agentive suffix
B	"set B" (absolutive) markers
CAUS	causative
CLF	noun classifier
CLF.NUM	numeral classifier
COMP	complementizer
CON	discourse connective marker
DEIX	deictic particle
DEP	dependent clause marker
DET	determiner
DFLT	default

DIR	directional light verb
DISC	discourse marker
DS	directional suffix
DTV	derived transitive status suffix
EMPH	emphasis
EXT	existential predicate
FOC	focus marking
HUMAN	human plural marker
INDF	indefinite
IPFV	imperfective aspect
IRR	irrealis clitic
IV	intransitive status suffix
NEG	negation
NML	nominal suffix
NMLZ	nominalization
P	plural person in agreement markers
PASS	passive
PFV	perfective
PL	plural marker
PREP	preposition
PROG	progressive aspect
PRON	pronoun
REP	reportative
S	singular agreement marking
SG	singular
SS	status suffix
STAT	stative suffix
SUF	unanalyzed suffix
TAM	tense-aspect-mood marking
TOP	topic marker
TV	transitive verb

Table of Contents

Chapter 1

Introduction

This thesis explores topics pertaining to the syntax and semantics of nominal expressions, with a focus on definite (1), indefinite (2), demonstrative (3), and pronominal (4) elements.

(1) Definite determiners
Magda read **the** book.

(2) Indefinite quantifiers
Magda read **a** book.

(3) Demonstratives
Magda read **this** book.

(4) Pronouns
Magda read **it**.

It is commonly assumed that the bolded functional items in (1)–(3) are in complementary distribution. From the perspective of English, this assumption is supported by the fact that none of these items can co-occur within the same nominal phrase:

(5) a. *Magda read **this the** book.
b. *Magda read **a the** book.
c. *Magda read **this a** book.
d. *Magda read **a this the** book.

As has long been noted in work on the syntax and semantics of the extended nominal domain (see e.g., Alexiadou et al. 2007; Giusti 2015; Leu 2015), however, this assumption is undermined when we expand our empirical focus to languages beyond English. Examples from Spanish and Greek, for example, teach us that co-occurrences between definite determiners and demonstratives are possible in at least some languages:

(6) Spanish
El libro **este** fue publicado en 1990.
DET book DEM was published in 1990
'This book was published in 1990.' (Brugè 1996: 28)

(7) Greek
Dhen perimena **afti** **tin** antidhrasi.
not expected-1SG DEM DEF reaction
'I didn't expect this reaction.' (Alexiadou et al. 2007: 108)

The Spanish and Greek examples show us that definite determiners and demonstratives, at least in these languages, are not in complementary distribution. Contrary to what the English datapoint in (5-a) initially suggests, then, definite determiners and demonstratives are not necessarily incompatible with each other.

Relatedly, Arsenijević (2018) recently shows that demonstratives can co-occur with overt indefinite determiners in Serbo-Croatian, as shown in (8).

(8) Serbo-Croatian (Arsenijević 2018: 179)
Sa **tim** **nekim** ambicijam su ušli u Evroligu.
with DEM.INST some.INST ambitions AUX.PL entered in Euro.league
'With those ambitions, they entered the Euro league.'

Once again, we find that combinations of functional items within the nominal domain, of the kind that are illicit in English (5-c), seem to be possible in other languages.

Chuj, a Mayan language which serves as the main empirical focus of this thesis, provides an extreme case of a language that tolerates combinations of different functional items within the nominal domain. Indeed, like other closely-related Mayan languages (Craig 1977, Zavala 1992, Mateo Toledo 2017), it wears on its sleeve unusually rich nom-

inal syntax (Hopkins 1967, Maxwell 1987), with co-occurrences of functional items that would be completely unexpected from the perspective of languages like English. As we will see, this rich morphology presents strong support for a radical decomposition of functional items typically understood as primitives in the literature, with several elements partaking in the composition of semantic distinctions. The relevant configurations we will focus on are summarized in Table 1.1:

Table 1.1: Possible DP configurations in Chuj

	Configuration	Example	Rough translation
①	CLF + NP	*nok' tz'i'*	'the dog'
②	CLF + NP + DEIX	*nok' tz'i' chi'*	'the/that dog' (anaphoric)
③	CLF	*nok'*	'it'
④	INDF + NP	*jun tz'i'*	'a dog'
⑤	INDF + CLF + NP	*jun nok' tz'i'*	'a (certain) dog'
⑥	INDF + NP + DEIX	*jun tz'i' chi'*	'that dog'
⑦	INDF + CLF + NP + DEIX	*jun nok' tz'i' chi'*	'that dog' (exophoric)

The above table shows that Chuj is much more liberal than English in tolerating combinations of functional items within the nominal domain. In fact, assuming for now that noun classifiers (CLF) should be taken as the Chuj counterpart of a definite article (see line ①), then it appears that Chuj allows all of the ungrammatical English combinations of functional items shown in example (5). Compare for instance, (5-a) with ②, (5-b) with ⑤, (5-c) with ⑥, and (5-d) with ⑦.

Against the empirical picture presented in Table 1.1, the first part of this thesis will thus seek to provide answers to the following two questions:

Q1 What morphosyntactic elements are involved in the composition of nominal expressions used to pick out or quantify over entities?

Q2 What semantic and pragmatic distinctions are expressed by different combinations of functional items within the nominal domain?

In Chapters 3–5, I provide a compositional analysis of each of the combinations of items in Table 1.1, showing that different combinations of functional elements lead to subtle se-

mantic distinctions, with relevance to Q1 and Q2. As we will see, the resulting meanings that arise from combining different functional items within the nominal domain provide crucial insight into topics which have been at the core of debates in linguistic theory, such as the encoding of definiteness versus indefiniteness, the internal syntax and semantics of demonstratives, the nature of pronouns, and the ways in which the contextual domain of nominal expressions is implicitly or explicitly delimited. A recurring theme throughout the thesis will be that, by virtue of being radically decompositional, Chuj often challenges pre-existing assumptions about the primitivity of certain linguistic expressions, such as definite determiners and demonstratives. Instead, the Chuj data align with an increasing number of works arguing for a decomposition of certain functional elements within the extended nominal domain (see e.g., Déchaine and Wiltschko 2002, Leu 2008, 2015, Coppock and Beaver 2015, Arsenijević 2018, Hanink 2018, Ahn 2019, Hsu and Syed 2020), including many that have traditionally been viewed as simplex expressions.

The second part of this thesis will zoom out of the internal syntax of nominal expressions and into the distribution of bound and coreferential expressions within sentences. Specifically, by taking a close look at patterns of nominal binding and coreference in Chuj and Ch'ol (another Mayan language), Chapter 6 will seek to provide answers to the following general question:

Q3 What conditions are imposed on co-referential or bound nominal expressions within sentences?

Once again, I will argue that Mayan languages provide a unique perspective towards answering this question. The reason is that Chuj, like some of its other close relatives (Craig 1977, Hoekstra 1989, Aissen 2000), diverges from Ch'ol in exhibiting unusual patterns of nominal binding and coreference. In fact, while Ch'ol seems to generally abide by the Binding Conditions, Chuj seems to entirely ignore them. Instead, only linear precedence seems to dictate the placement of R-expressions and pronouns in Chuj, often to the apparent violation of the Binding Conditions. Given that the Binding Conditions have tradi-

tionally been regarded as universal (Grimshaw and Rosen 1990, Grodzinsky and Reinhart 1993, Reuland 2010, 2011, 2016), comparable patterns in Popti', a close relative of Chuj, led previous authors to conclude that the Binding Conditions (as conceived in work like Reinhart 1983 and Chomsky 1986) could not be universal, since they do not dictate the distribution of nominals in this language (Hoekstra 1989, Aissen 2000). My goal will be to argue that this conclusion was reached too quickly: once we understand the syntax of the relevant sentences in these languages, which crucially diverge from Ch'ol, we see that the binding violations are only illusory, because there are no structural relations between the relevant expressions. The general lesson will be that despite initial evidence to doubt the universality of the Binding Conditions, a universalist approach can not only be maintained, but is supported by the Chuj data.

The rest of this thesis is structured as follows. Chapter 2 provides basic information about the demographics of Chuj, as well as the morphosyntactic pieces involved in the Chuj extended nominal domain, focusing on the components of central interest in this dissertation: (i) noun classifiers, (ii) indefinite determiners, (iii) deictic particles, and (iv) possessive constructions. This chapter will also provide information about the methodology used for the collection and interpretation of data, which includes both data extracted from corpora and data obtained via direct elicitation.

In Chapters 3–5, I then turn to the first two questions posed above (i.e., *What morphosyntactic elements are involved in the composition of nominal expressions used to pick out or quantify over entities?*; and *What semantic and pragmatic distinctions are expressed by different combinations of functional items within the nominal domain?*). In attempting to account for (i) the morphosyntactic constitution of the expressions in Table 1.1 and (ii) how they compositionally interact in order to deliver the subtle semantic distinctions to be described, it will be necessary to provide an analysis in steps, with adequate revisions made upon exposure to new data.

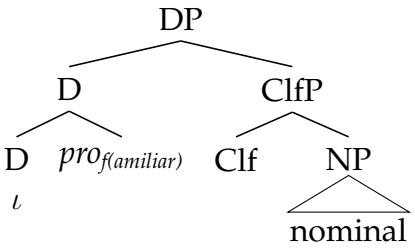
Chapter 3 will break the ice by providing the foundations for an analysis of definiteness in Chuj (lines ①–③ of Table 1.1). Adopting a Situation Semantics (Kratzer 1989, Elbourne 2005, Schwarz 2009), I will argue that Chuj provides key support for recent theories that distinguish between weak and strong definite descriptions (e.g., Schwarz 2009, 2013; Arkoh and Matthewson 2013; Jenks 2015b, 2018; Hanink 2018, Ahn 2019). A set of morphemes called "noun classifiers" contributes a uniqueness presupposition, composing directly with nominals to form weak definites. To form strong definites, I show that two pieces are required: (i) the noun classifier, which again contributes the same uniqueness presupposition, and (ii) extra morphology that contributes an anaphoricity presupposition. Chuj strong definites thus provide explicit evidence for a decompositional account of weak and strong definites, as also advocated by Hanink (2018) and Ahn (2019). I then extend this analysis to third person pronouns, which are realized in Chuj with bare classifiers, and which I propose come in two guises depending on their use. On the one hand, based on previous work (Postal 1966, Cooper 1979, Heim 1990), I argue that classifier pronouns can sometimes be E-type pronouns: weak definite determiners which combine with a covert index-introducing predicate. In such cases, classifier pronouns represent a strong definite description. On the other hand, I argue based on diagnostics established by Bi and Jenks (2019) that Chuj classifier pronouns sometimes arise as a result of NP ellipsis (Elbourne 2001, 2005). In such cases, classifier pronouns reflect a weak definite description.

Chapter 4 then expands the empirical focus to the expression of indefiniteness in Chuj and the role that noun classifiers play in indefinite expressions. I show that noun classifiers can also combine with indefinite determiners, focusing on lines ④–⑤ of Table 1.1. This is puzzling given the analysis proposed in Chapter 3, where noun classifiers are treated as weak definite articles. All else being equal, it is not clear why and how a definite determiner should be able to co-occur with an indefinite determiner. The goal of Chapter 4 is thus to provide a solution to this puzzle, all while maintaining the core pro-

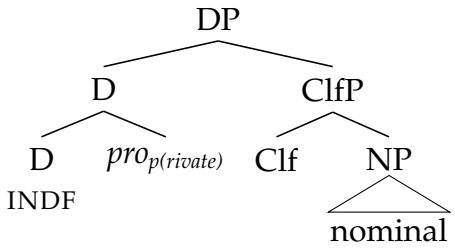
posal about noun classifiers in Chapter 3, namely that they impose existence and uniqueness presuppositions on the interpretation of nominal expressions. Empirically, I show that noun classifiers do make a semantic contribution to the interpretation of the indefinite. Building on previous work on closely related Mayan languages (Craig 1977, 1986b, Ramsay 1985, Trechsel 1995, Zavala 2000), I show that when noun classifiers combine with indefinite determiners, the result is a specific indefinite (with a wide-scope interpretation). I then argue that this can be explained if indefinite-classifier combinations are interpreted as singleton indefinites (Schwarzschild 2002). In short, I propose that when co-occurring with indefinite determiners, the uniqueness component of the noun classifier indicates that the domain of the relevant indefinite expression is delimited to a singleton set. To account for this new data, parts of the analysis from Chapter 3 will have to be refined. A first proposal will come from the semantic type of the noun classifier, which, building on Coppock and Beaver 2015, I argue denotes a partial identity function over predicates. A second key proposal will come from the kind of value assigned to the situation argument of definite versus indefinite expressions, responsible for contextual domain restriction (Schwarz 2009, 2012). Specifically, building on Beaver and von Fintel (2013) and Arsenijević (2018), I will argue for a distinction between two kinds of indexed situation variables, in accordance with the kinds of conditions they impose on the discourse participants' ability to retrieve the intended contextual domain restriction associated with an expression. In the case of definite descriptions, the DP is headed by the null morpheme ι, which *must* combine with a situation pronoun whose value is known or is *familiar* to both the speaker and hearer. Indefinite expressions, on the other hand, combine with a situation pronoun whose value *does not necessarily need* to be retrievable by all discourse participants (Schwarzschild 2002), one whose values can be *private*. The general idea is schematized below:[1]

[1]As will be made clear as the thesis unfolds, a "familiar situation variable" is one whose value has to be known to both the speaker and addressee(s). I do not take this to mean that any referent in the situation is necessarily itself identifiable with an index in the variable assignment (as will be the case with strong definite descriptions).

(9) Weak definite description



(10) Indefinite+classifier DP



The crucial point will be that when the noun classifier occurs within an indefinite expression, its presupposition is triggered relative to a private situation pronoun, and the effect of the noun classifier's presuppositional component in such cases is not a weak definite description, but a singleton indefinite.

In Chapter 5, I then show that the refined entries for noun classifiers in Chapter 4 can be extended to account for the necessity of noun classifiers in definite descriptions, if weak definiteness is further decomposed with an additional ι in the head of DP, as shown in (9). This ι will crucially combine with a *familiar situation pronoun*, which will result in the non-assertive component of the noun classifier projecting as a classic pragmatic presupposition. I also show that the theory of definiteness and indefiniteness proposed in this thesis can be straightforwardly extended to explain why deictic particles can combine with both definite and indefinite nominal expressions alike, as shown in the configurations on lines ②, ⑥, and ⑦ of Table 4.1. In particular, the entry for deictic particles provided in Chapter 3 can be minimally revised so as to account for the fact that they may compose with both definite and indefinite "bases", all while keeping to the core proposal that they introduce indices, which when combined with a definite base results in a strong definite description.

Chapter 6 turns to another and final puzzle by zooming out of the internal syntax of nominal expressions and into the distribution of covalued nominal expressions within sentences. In addition to Chuj, this chapter will also provide extensive data from Ch'ol, another Mayan language. I show that while Ch'ol behaves entirely as expected with regards to the binding conditions, Chuj appears to consistently tolerate violations of Con-

dition C, often privileging linear precedence as the determining factor in the distribution of R-expressions and pronouns. The Chuj data therefore initially seem to cast doubt on a long tradition to treat the binding conditions as universal (e.g., Grodzinsky and Reinhart 1993, Reuland 2010, 2011). I argue that the difference between Chuj and Ch'ol can be largely explained if, contrary to Ch'ol, Chuj exhibits 'high-absolutive' syntax, independently proposed to account for a number of morphosyntactic phenomena in a subset of Mayan languages (Coon et al. 2014, Coon et al. 2021). High-absolutive syntax creates configurations in which the internal argument asymmetrically c-commands the external argument, bleeding otherwise expected binding relations from the external argument into the internal argument. The violations of Condition C in Chuj are thus only apparent. I further argue (i) that linear precedence effects in Chuj are a reflex of a more general anti-cataphora constraint on free nominals, which can also be shown to apply to Ch'ol, and (ii) that there are corners of Chuj where the binding conditions do apply, and that in such cases linear precedence is irrelevant for the distribution of covalued nominals. This means that the binding conditions are active in Chuj, even though idiosyncratic syntactic properties of the language often render their application impossible. Therefore, despite Chuj initially leading us to believe that the Binding Conditions are not universal, a universalist approach can not only be maintained, but is supported by the data.

Finally, chapter 7 concludes with a summary of the proposals put forth in this thesis, as well as a brief discussion of the cross-linguistic consequences that follow from these proposals.

Chapter 2

Chuj, a Mayan language

> *T'alaj chi'jun, mato ayb'aj olelolonelejxi. Mato ayb'aj oleyalxi. Hataton sb'iwixi kan juntzanh kik'ti tik. Hi, ay jun wab'nak tob' ichatik yaji: max cham laj juntzanh yistorya komam kicham tik.* (Pedro Domingo Gómez, txt CP200715)
>
> 'Maybe this is why at some point you might talk about this again. Or maybe you'll tell someone again. And this is how these stories of ours will live on. Yes, there's a saying that goes: "these stories of our ancestors don't die."'

The principle empirical data discussed in this dissertation are drawn from original fieldwork on Chuj, a Mayan language. The goal of this chapter is to provide preliminary background information on the demographics and grammar of Chuj, as well as on the methodology used in fieldwork and elicitation. Emphasis will be put on the nominal domain, the main topic of study in this work.

In section 2.1, I provide basic demographic information about Chuj and other Mayan languages. In section 2.2, I establish the fieldwork methodology used for the collection of linguistic data. Section 2.3 then provides information about the orthographic conventions used in this work. I then turn to a discussion of basic grammatical properties of Chuj, with the verbal domain discussed in section 2.4 and the nominal domain discussed in section 2.5. In discussing the nominal domain, I will focus on the grammatical items of particu-

lar relevance for this thesis, including (i) noun classifiers, (ii) indefinite determiners, (iii) adnominal demonstrative constructions, and (iv) possessive constructions.

2.1 Demographics, vitality, and previous work

2.1.1 The Mayan language family

The Mayan language family consists of about 30 languages, primarily spoken in Guatemala, Mexico, Belize and Honduras; see Figure 2.1 on the next page. Mayan languages are generally classified according to four major branches, two of which are sub-grouped into two further secondary branches. The linguistic classification of Mayan languages, as proposed by Kaufman (1974, 1976), is provided in Table 2.1.

Table 2.1: The Mayan Language Family (based on Kaufman 1976 and Law 2014)

	Primary branch	Secondary Branch	Languages
PROTO MAYAN	Yukatekan		Itzaj (Itza'), Lacandon (Lakantun), Mopan, Yukatek (Maya)
	Huastecan		Chicomuceltec (Kabil) Huastec (Teenek)
	Western	Ch'olan-Tseltalan	Ch'ol, Ch'olti', Chontal (Yokot'an), Ch'orti', Tseltal, Tsotsil
		Q'anjob'alan	**Chuj**, Akatek, Mocho' Popti', Q'anjob'al, Tojol-ab'al
	Eastern	K'ichean	Achi, Kaqchikel, K'iche' Poqomam, Poqomchi', Q'eqchi' Sakapultek, Sipakapense Tz'utujil, Uspantek
		Mamean	Awakatek, Chalchitek Ixil, Mam, Tektitek (Teko)

As noted by Law (2014), while the classification in Table 2.1 is overall widely accepted, there are still ongoing debates about the correct classification of a subset of the Mayan

languages, most notably with regards to the classification of Huastecan languages and Tojol-ab'al (see e.g., Campbell 1977, Robertson 1977, Schumann 1981, Robertson 1992, Hopkins 2006, Gómez Cruz 2017).[1] For overviews of Mayan languages, see England 1988, 1994, 2001; Bennett et al. 2016; Aissen et al. 2017.

Figure 2.1: Current-day Mayan-speaking area (Law 2014, p. 25)
This is a derivative of "Present geographic distribution of Mayan languages in Mexico and Central America" by noahedits, used under CC BY-SA 4.0.

[1]The classification of Tojol-ab'al has been subject to debate. Recent work, however, has argued that it shows signs of a 'mixed' language, showing not only several similarities with Chuj, but also with Tseltal (Law 2011, 2014, Gómez Cruz 2017).

2.1.2 Demographics

This thesis will focus on Chuj (pronounced [ʧux]), a language that belongs to the Western branch of Mayan languages and is generally classified as a member of the Q'anjob'alan sub-family, along with Akatek, Mocho', Popti' (formerly known as Jakaltek), Q'anjob'al, and Tojol-ab'al (though see footnote 1 on Tojol-ab'al). Chuj speakers refer to their language as *koti'*. As of 2009, Chuj was spoken by an estimated 70,000 speakers in Guatemala and Mexico (Piedrasanta 2009, Buenrostro 2013b), with the vast majority of speakers living in Guatemala. Chuj is also spoken by significant diaspora communities across North America, with estimates of speakers being in the thousands across several states in the United States of America (Maxwell 1993, Hopkins 2021, Kaplan 2021). Figure 2.2 shows the location of Chuj communities in Guatemala and Mexico.

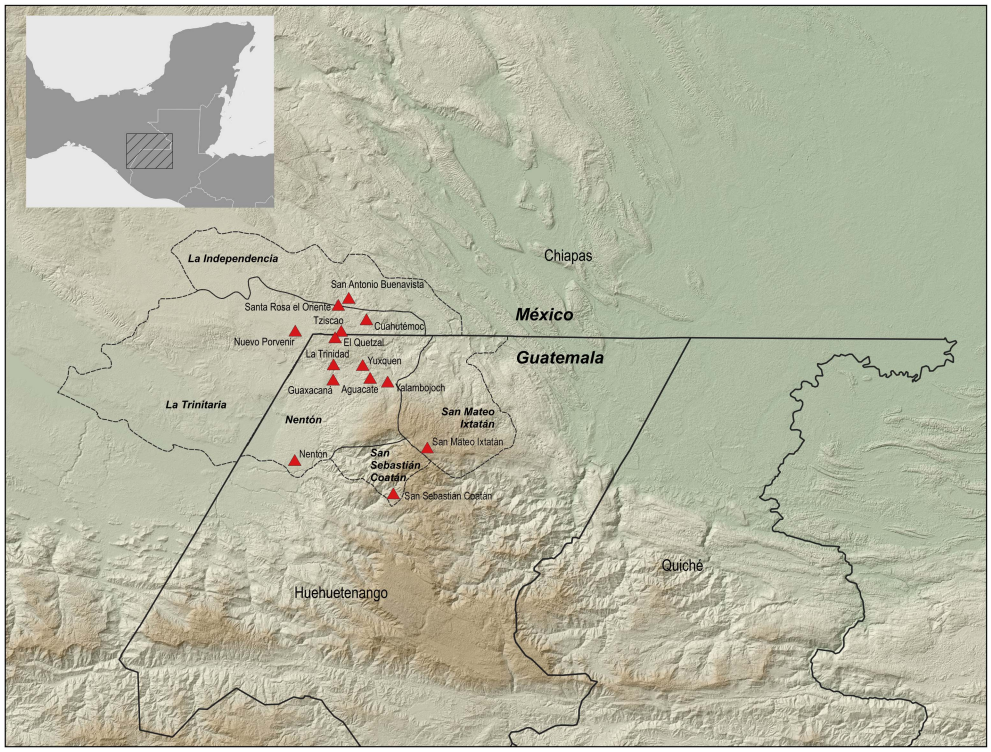


Figure 2.2: Map of Chuj speaking area
Created by and used with permission from Gerardo Jiménez Delgado, Laboratorio de Análisis Espacial y Digital, IIA–UNAM

As seen in Figure 2.2, Chuj communities in Guatemala are located in three municipalities of the Department of Huehuetenango: San Mateo Ixtatán, San Benito Nentón, and San Sebastián Coatán. Many of these communities are located in the *Sierra de los Cuchumatanes* (indicated with brown shading), which hosts some of the highest mountains in Central America, with drastic changes in the climate and surrounding ecosystem from one community to the other. As Hopkins (2021) notes about the cold climate of San Mateo Ixtatán:

> Given the climate, it is no surprise that the name of the language derives from a prominent feature of households, the sweatbath. "Chuj" is a word that is ultimately of Mamean origin, but is used in local Spanish for the low structures that sit at the sides of houses, used for ordinary bathing as well as curing ceremonies. I once asked a man why they didn't bathe in the rivers, and he looked at me astonished and said, "Good Lord! Do you know how cold that water is?" (Hopkins 2021: 3)

Two main dialects of Chuj are generally recognized (Maxwell 1981, Domingo Pascual 2007, García Pablo and Domingo Pascual 2007, Buenrostro 2013b): the San Mateo Ixtatán and San Sebastián Coatán dialects. In Guatemala, the first dialect is spoken in the municipalities of San Mateo Ixtatán and San Benito Nentón, while the second dialect is primarily spoken in the municipality of San Sebastián Coatán. Since communities from both dialects of Chuj emigrated to Mexico during the Guatemala Civil War, both dialects are spoken in Mexico. Unless otherwise indicated, the data discussed in this thesis were obtained in work with speakers of the San Mateo Ixtatán dialect.

Chuj communities in Mexico are now located primarily in Chiapas in two municipalities: La Trinitaria and La Independencia. While there were Chuj communities located in Mexico before the war, many Chuj communities now located in Mexico were formed as a result of large displacements of Maya speakers in the Mexican states bordering Guatemala (Quezada and O'Dogherty 1986, Buenrostro 2001, Limón 2009, Córdova Hernández 2014, López and Herrera 2019). In 2015, the Instituto Nacional de Lenguas Indígenas estimated an approximate 2890 Chuj speakers located in Mexico.

2.1.3 Vitality

According to the language vitality and endangerment scale of UNESCO, Chuj is considered "vulnerable" (Moseley 2010). However, the overall vitality of Chuj varies from Guatemala to Mexico. The language has a high degree of vitality in Guatemala, as reflected by the large number of monolingual speakers, particularly women and young children (Buenrostro 2013b). In most Chuj communities in Huehuetenango, Chuj is often the main language spoken at home and continues to be transmitted to younger generations, with many children learning Spanish only upon exposure to the elementary school system. Moreover, the *Smakb'enal Sti' Chonhab' Chuj / Comunidad Lingüística Chuj*—a subbranch of the governing body of Mayan languages of Guatemala (the *Academia de Lenguas Mayas de Guatemala*)—remains active in Chuj municipalities of Huehuetenango, providing a number of services, including free access to L1 and L2 Chuj courses.

In Mexico, the vitality of the Chuj language is significantly lower than in Guatemala, and it is considered endangered by Buenrostro (2013b) and Córdova Hernández (2014). Buenrostro (2013b) attributes the loss of Chuj to (i) the country's linguistic politics, which at one point prohibited the use of Chuj and other Indigenous languages, and (ii) the devastating effects of the Guatemalan refugee crisis in Mexico, which caused many Mexican-born Chuj speakers to stop speaking their language in fear of being labelled as refugees.[2] Based on Fishman's (1991, 2001) Graded Intergenerational Disruption Scale (GIDS) for language endangerment, Córdova Hernández (2014) further notes that while approximately 50% of communities have reached stage 6 of the GIDS (Chuj is only transmitted within households and communities and not within external institutions), another 50% of communities have reached stage 7, which critically involves the absence of intergenerational transmission. That said, Buenrostro (2013b) mentions the existence of certain Mexican communities where a process of language revitalization has been observed during the last decade (see also Córdova Hernández 2014 for relevant discussion).

[2]For details on the Guatemala refugee crisis in Mexico, see Quezada and O'Dogherty (1986)

In sum, Chuj remains a vulnerable language, and some communities have experienced a decrease in the number of native speakers. Nonetheless, its vitality in the Department of Huehuetenango, as well as revitalization efforts in Mexican communities, hopefully indicate a promising and healthy future for the language.

2.1.4 Previous work on Chuj

The earliest published work on grammatical properties of Chuj dates back to Nicholas Hopkins' 1967 doctoral dissertation entitled *The Chuj Language* as well as work by Williams and Williams (1966). There have since been two more doctoral theses (Maxwell 1981, Buenrostro 2013b), as well as a number of articles and undergraduate theses (Hopkins 1970, 1972, 1980a, 1980b, 2012a, 2012b; Williams and Williams 1971; Maxwell 1976, 1978, 1983, 1987, 1990, 1993, 1995; Schumann 1981; Buenrostro et al. 1989, Buenrostro 1992, 1994, 1995a, 1995b, 1996, 1998a, 1998b, 2000, 2001, 2002, 2004, 2005, 2007, 2009, 2010a, 2010b, 2011, 2013a, 2014, 2015a, 2015b, 2017, 2018; Medina and Buenrostro 2003; Duncan 2007; Hou 2011; Bielig 2015; Carolan 2015; Coon 2016b, 2018, 2019; Kotek and Erlewine 2016, 2019; Coon and Carolan 2017; Gómez Cruz 2017; Royer 2017, 2019, 2020, 2021a, 2021b, 2021c, 2022; to appeara; Henderson et al. 2018; Elias 2019; Royer and Alonso-Ovalle 2019; Alonso-Ovalle and Royer 2020, 2021, 2022; Coon and Royer 2020; Buenrostro and Royer 2022; Little, Moroney, and Royer 2022a; Royer et al. 2022; Vázquez Hernández 2022; Brodkin and Royer to appear; Felipe Gómez 2022, López García 2022). There are also two grammars produced by the *Academia de Lenguas Mayas de Guatemala* (Domingo Pascual 2007; García Pablo and Domingo Pascual 2007); three dictionaries (Felipe Diego and Gaspar Juan 1998; Comunidad Lingüística Chuj 2003; Hopkins 2012a); a recently published book of Chuj texts (Hopkins 2021); and a forthcoming collection of Chuj articles edited by Cristina Buenrostro (Buenrostro, to appear). Finally, two collections of narratives, field notes, and recordings, are archived on the *Archive of the Indigenous Languages of Latin America* (AILLA; ailla.utexas.org) (Hopkins 2018; Mateo Pedro and Coon 2018).

2.2 Fieldwork and methodology

Unless otherwise specified, all Chuj data discussed in this thesis were obtained via one of two sources: (i) targeted and theoretically-informed elicitation tasks conducted by myself and (ii) the corpus of transcribed narratives and interviews made available by Mateo Pedro and Coon (2018) on AIILA, the latter of which are identified with "txt" in examples.

Data from targeted elicitations were obtained through several years of collaborative work with Chuj speakers in Montreal, Mexico and Guatemala. Principle consultants on projects have been Matal Torres, Matin Pablo, Elsa Vázquez and Yun Torres. The former two collaborators reside in Montreal, and immigrated to Canada as adults in the 1990s as an immediate consequence of the Guatemala Civil War and resulting Maya Genocide. The latter two collaborators reside in Yuxquen, a small town of approximately 250 inhabitants located in the municipality of San Benito Nentón in Guatemala. A picture of the town of Yuxquen, my principle area of research in Guatemala, is shown in Figure 2.3.

Figure 2.3: Town of Yuxquen, San Benito Nentón, Huehuetenango

In collecting data on Chuj, I employed a number of contextualized elicitation tasks, as prescribed in existing resources on linguistic fieldwork methodology, including Payne 1997, Matthewson 2004, Bowern 2008, Krifka 2011, and Bochnak and Matthewson 2015, 2020. The three principle elicitation tasks were the following:

(11) Main methods used during elicitation

a. Requests for well-formedness judgements of sentences uttered in Chuj.
b. Translation tasks from Spanish to Chuj, and from Chuj to Spanish.
c. Requests for acceptability judgements of well-formed Chuj utterances.

The first task in (11-a), also sometimes called a grammaticality judgement task (see e.g., Matthewson 2004), allows one to probe into whether a particular string of words constitute a morphosyntactically well-formed sentence in the language. Without a provided context, if the consultant responds that the utterance is well-formed, then one can conclude that it is a grammatical sentence in the language. If the answer is that it is not well-formed, the researcher can merely conclude that the sentence is not *necessarily* grammatical, since a negative judgement could be provided if the speaker is unable to think of a proper context legitimizing the use of the sentence.

Much of the data discussed in the current work concerns an investigation of *meaning*, which involves determining the truth conditions or pragmatic felicity conditions of full sentences (Matthewson 2004). As emphasized in much work in semantic fieldwork (see e.g., Matthewson 2004, AnderBois and Henderson 2015, Bohnemeyer 2015, and Bochnak and Matthewson 2020), a crucial component of research in semantics leans on establishing an appropriate state-of-affairs for which a targeted utterance—one which is already known to be morphosyntactically well-formed—can be judged for acceptability. As discussed in detail in Matthewson 2004 and Bochnak and Matthewson 2020, if a consultant judges a syntactically well-formed utterance as acceptable in the provided context, then it is determined that the sentence is both true and felicitous in the context. If they provide a judgement of unacceptable, then one can conclude either that the sentence is false

or unacceptable for another pragmatic reason.[3] Accordingly, the researcher and consultant must agree on a clear *context* for which a full sentence will receive an acceptability judgement. In conducting acceptability judgement tasks (11-c), I therefore systematically accompanied the sentence with an appropriate context. Translation tasks were also often accompanied with context. Contexts were provided to the speakers either orally, with pictures, or by recreating an appropriate physical situation. Orally-provided contexts were always provided in the contact language, Spanish. Translations from Spanish to English in examples are my own.

Both positive and negative acceptability judgements were taken into consideration to investigate the semantic or pragmatic contribution of different expressions. When judgements were unclear or more fuzzy, speakers were sometimes asked to judge the appropriateness of a sentence, given a context, on a scale of 1 to 5—1 being less acceptable and 5 being most acceptable. Comments from speakers about particular expressions or contexts were taken into consideration whenever relevant. A sample judgement task, with English as the object language, is provided below.

(12) Example judgement task to investigate the contribution of 'the'
Context: In your town, there are three priests. In a conversation, you tell another resident of your town the following sentence:
Target sentence [in object language]: *I spoke with **the** priest yesterday.*
Question: Is this sentence acceptable given the context?

Also note that in judgements or translation tasks from Chuj to Spanish, consultants were routinely invited to repeat the targeted Chuj sentence. This allowed me to make sure that (i) the speaker had heard the targeted sentence well, (ii) they had not generously

[3] A subtype of acceptability judgement task is a truth value judgement task (Matthewson 2004), used when it can be pre-established that a morphosyntactically well-formed sentence does not contain any presupposition failures or is pragmatically well-formed in the context. Because a large part of this thesis is concerned with presuppositions, and since it has been established that there is not a clear relationship between presupposition failure and truth-value intuitions (see e.g, von Fintel 2004), I have mostly opted to ask for acceptability judgements rather than truth value judgements in this thesis.

corrected an ungrammatical or unacceptable sentence, and (iii) my transcription of the targeted sentence did not contain any mistakes.

Whenever appropriate, I also employed resources that were specifically established to investigate topics of study in this work. These include two questionnaires on definiteness (Jenks 2015a, Vázquez-Rojas Maldonado et al. 2017), as well as Wilkins's (2018) demonstrative questionnaire. I also employed one elicited production task, which involved consultants narrating Mercer Mayer's (1969) illustration book *Frog, Where Are You?*.

2.3 Chuj orthographic conventions

The examples presented here and throughout this work are written using the Chuj standard orthography, which is based on the alphabet created by the *Academia de Lenguas Mayas de Guatemala* in 1987 and used in the most recent prescriptive and descriptive grammars of Chuj (Domingo Pascual 2007, García Pablo and Domingo Pascual 2007). Table 2.2 shows the Chuj practical orthography on the first line, with IPA correspondences on the second line. Materials taken from works adopting different orthographic conventions (e.g., Hopkins 1967, Maxwell 1981, Hopkins 2012b) have been updated to reflect the standard orthography.

Table 2.2: Chuj orthography (<>) - IPA correspondence ([])

<	a	b'	ch	ch'	e	h	i	j	k	k'	l	m	n	nh	o	p	r	s	t	t'	tz	tz'	u	w	x	y	'	>
[	a	ɓ	ʧ	ʧ'	e	h/∅	i	x	k	k'	l	m	n	ŋ	o	p	r	s	t	t'	ʦ	ʦ'	u	w/v/b	ʃ	j	ʔ	]

A few notes on the orthographic conventions in Table 2.2 are in order. First, note that the apostrophe <'> corresponds to two distinct phonetic realizations: it represents (i) an ejective or implosive when it appears immediately following a consonant, and (ii) a glottal stop consonant when it appears following a vowel or between two vowels <V'V>. Glottal stops before word-initial vowels are not represented in the orthography (see (13) below).

Second, as indicated in the above table, the phoneme represented as <w> varies in phonetic realization between a bilabial approximant [w], a labiodental fricative [v], and a voiced bilabial stop [b] (distinct from the implosive stop <b'>). The distribution of [w], [v], [b] largely overlaps (even alternating for the same words and sentences for the same speakers). Following other works (e.g. García Pablo and Domingo Pascual 2007; Royer et al. 2022), I assume that the three phonetic realizations are in free variation, and represent them all with the letter <w>.

Finally, a third important note concerns the letter <h>. In addition to marking glottal fricatives, this letter can also indicate the absence of an otherwise present word-initial glottal stop (García Pablo and Domingo Pascual 2007, Hopkins 2012a, Lesure 2016). This convention is not adopted in all works (see e.g. Buenrostro 2013b). The presence or absence of <h> leads to minimal pairs, such as the following:

(13) *ch'anh ab'* → [ʧ'aŋ ʔaɓ] 'the hammock'

(14) *ch'anh hab'* → [ʧ'aŋ aɓ] 'your hammock'

2.4 Verbal and clausal structure

2.4.1 Verbal predicates

As is the case in other Mayan languages (England 2001, Grinevald and Peake 2012, Coon 2016a, Aissen et al. 2017), Chuj is a head-marking, ergative-absolutive language. Fully inflected verbal predicates exhibit the templates in (15-a) and (16-a) below, shown with representative examples. As shown, the verbal stem typically consists of a Tense-Aspect-Mood (TAM) marker, person/number morphology, a root, possibly one or more derivational suffixes (glossed simply as "VOICE" in (15-a) and (16-a)), and finally, what is known in Mayan literature as a "status suffix". As discussed in Royer 2020, 2022, status suffixes

in Chuj only surface overtly at the edge of certain prosodic domains and are otherwise dropped (see also Henderson 2012 on K'iche' and Aissen 1992 more generally).

(15) a. **Transitive verb template in Chuj**
TAM – Set B (ABS) – Set A (ERG) – ROOT – VOICE – SS

b. *Ixachwila'.*
Ix - ach - w - il - a'.
PFV - B2S - A1S - see - TV
'I saw you.'

c. *Ixonheyanhtej.*
Ix - onh - ey - anh - t - ej.
PFV - B1P - A2P - cure - CAUS - DTV
'Y'all cured us.'

(16) a. **Intransitive verb template in Chuj**
TAM – Set B (ABS) – ROOT – VOICE – SS

b. *Ixonhb'ati.*
Ix - onh - b'at - i.
PFV - B1P - go - IV
'We went.'

c. *Ixexanhtaji.*
Ix - ex - anh - t - aj - i.
PFV - B2P - cure - CAUS - PASS - IV
'Y'all were cured.'

As shown above, the two sets of person/number markers, known in Mayanist literature as "Set A" and "Set B" (Bricker 1977; Larsen and Norman 1979; Robertson 1980), cross-reference arguments on the predicate, revealing an ergative-absolutive alignment pattern. Transitive stems appear with Set A (ergative) marking cross-referencing the subject (15-a), and Set B (absolutive) morphemes cross-reference the object. Intransitive subjects are marked with Set B (16-a), with the exception of the progressive aspect, which shows an apparent split in alignment, discussed in Coon and Carolan 2017 and Coon and Royer 2020. Set A and Set B morphemes are provided in Table 2.3. As in other Mayan languages, Set A forms have both pre-consonantal and prevocalic allomorphs and are also used to cross-reference possessors on nominals (see §2.5.4). Notice that third person

singular Set B has no phonological exponent, and will therefore not be represented in glosses. Finally, also note that the orthographic <h> on person/number markers is only used when the relevant markers appear at the beginning of a phonological word, and not when they appear internal to words, as is the case in the examples in (15) and (16) above.

Table 2.3: Chuj Set A and Set B morphemes

	Set A (ergative/possessive)		**Set B** (absolutive)
	___C	___V	
1S	*(h)in-*	*w-*	*(h)in*
2S	*(h)a-*	*h-*	*(h)ach*
3S	*s-*	*y-*	Ø
1P	*ko-*	*k-*	*(h)onh*
2P	*(h)e-*	*hey-*	*(h)ex*
3P	*s-*	*y-*	Ø

Note that in the rest of this thesis, I will sometimes simplify glossing of derivational morphology within complex words, whenever the morphological break is not crucial to the point at hand.

2.4.2 Clausal syntax and word order

Like other Mayan languages, Chuj exhibits verb-initial word order in discourse-neutral contexts (see e.g. England 1991, Aissen 1992, Clemens and Coon 2018). The basic word order in San Mateo Chuj is VOS, though, as discussed in Chapter 6 of this thesis, VSO is sometimes permitted in special circumstances. A basic VOS sentence is provided below:

(17) *Ixschi' nok' mis nok' tz'i'.*
Ix-s-chi' [$_{OBJ}$ nok' mis] [$_{SUBJ}$ nok' tz'i'].
PFV-A3-bite CLF cat CLF dog
'The dog bit the cat.' (Chuj)

While VOS is the basic word order, subjects and objects frequently appear in a preverbal position to mark topic and focus, a common pattern across Mayan languages (Aissen 1992, Clemens and Coon 2018, Royer 2022). In Chuj, topics and foci exhibit different prop-

erties. Third person topics require overt pronominal resumption (see §2.5.1 on pronouns) and show prosodic contours characteristic of intonational phrases, including a prosodic break before the rest of the clause—indicated with "‖" in (18) (see Royer 2022):

(18) **Subject topic in Chuj**

Ha nok' tz'i' ixschi' nok' mis nok'.

[$_{\text{TOP}}$ Ha nok' tz'i'] ‖ ix-s-chi' [$_{\text{OBJ}}$ nok' mis] [$_{\text{SUBJ}}$ **nok'**].

TOP CLF dog PFV-A3-bite CLF cat PRON

'As for the dog, it bit the cat.'

Foci, on the other hand, do not trigger resumption and do not show the characteristic prosodic features of intonational phrases. As discussed in more detail in Chapter 6 of this thesis, focused transitive subjects trigger the obligatory presence of the Agent Focus suffix *-an* on the verb stem, as shown in (19) (see e.g. Aissen 2017 and references therein). Note that I indicate focused expressions with small capital letters.

(19) **Agent focus in Chuj**

Ha nok' tz'i' ixchi'an nok' mis.

[$_{\text{TOP}}$ Ha nok' tz'i'] ix-chi'-**an** [$_{\text{OBJ}}$ nok' mis].

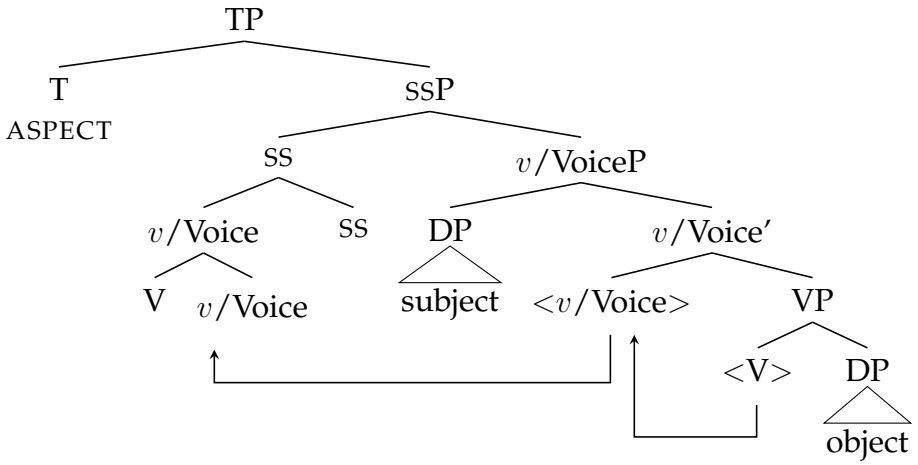
FOC CLF dog PFV-bite-AF CLF cat

'THE DOG bit the cat.' (Chuj)

The derivation of word order across different Mayan languages remains a controversial topic (Clemens and Coon 2018, Little 2020b). In this thesis, I follow Clemens and Coon (2018) in assuming that verb-initial word order (at least in Chuj) is derived from a base-generated SVO order with subsequent syntactic and post-syntactic operations deriving VOS word order (though see Aissen 1992, Coon 2010a, Douglas, Ranero, and Sheehan 2017 and Little 2020b for alternative proposals). Following Clemens and Coon (2018), I assume that the verb undergoes head movement to a projection hosting status suffix morphology, and label it accordingly as "SSP". I also assume, following Coon (2019) on Chuj, that the external argument of transitive verbs is merged in the specifier of a single bundled v/Voice head, triggering Set A on the verb stem. Set B, on the other hand, can have

different sources (v/Voice or T) depending on the construction (see Coon, Baier, and Levin 2021, as well as Chapter 6 for more details). A structure for the transitive verb is provided below. I assume, following Aissen 1992, that tense-aspect markers are instantiated in the head of TP.

(20) Verb-initial word order in Mayan (Clemens and Coon 2018)



Following Clemens and Coon 2018 and Coon et al. 2021, I further assume that the order of post-verbal arguments can be affected by a variety of factors, including object raising over the subject or phonological weight of arguments. These two options, as well as the derivation of word order more generally, will be discussed in more detail in Chapter 6.

2.5 The extended nominal domain

Chuj, much like other Mayan languages from the Q'anjob'alan sub-branch, is particularly relevant for questions pertaining to the syntax and semantics of the extended nominal domain, because it exhibits exceptionally rich nominal morphology, providing evidence for complex nominal structure (Maxwell 1987). The elements of the Chuj extended nominal phrase are schematized below, where 'top-to-bottom' indicates 'left-to-right' order. The bolded items indicate the morphology that will play an important role in this dissertation.

(21) *Extended nominal phrase in Chuj*

Focus/topic marking
Determiners/quantifiers
Numerals
Numeral classifiers
Human plural marker
Noun classifiers
Set A possessive mophology
Adjectives
Noun
Adjectives and relative clauses
Possessors
Deictic particles

While these items rarely all co-occur together within the same phrase, large nominal constituents can be found in narratives, as in (22) and (23). Furthermore, speakers have no trouble producing and judging large nominal constituents in controlled elicitation settings, including phrases containing all of the pieces in (21). This is exemplified in (24) (see also discussion in Maxwell 1987).

(22) *Howanh heb' winh ajb'ulej chi' ixchamxi.*

[Ho-wanh heb' winh aj-b'ulej chi'] ix-cham-x-i.
five-NUM.CLF PL.HUM CLF.MASC AG-b'ulej DEIX PFV-die-ADV.now-IV
'These five *B'ulejers* are now dead.' (txt, CP280715)

(23) *Ha jun wakwanh winakil tz'och junjun k'u t'ay atz'am.*

[Ha jun wak-wanh winak-il] tz'-och junjun k'u t'ay atz'am.
FOC INDF six-NUM.CLF man-NML IPFV-out each day PREP salt
'Six men would go out each day to the salt mines.' (txt, CX200715)

(24) *Ha juntzanh oxwanh heb' winh skotak unin junk'olal ix Malin chi' ixpax heb'.*

[Ha juntzanh ox-wanh heb' winh s-kotak unin junk'olal ix
TOP some three-N.CLF PL NOUN.CLF A3-small child happy N.CLF
Malin chi'] ix-pax heb'.
Malin DEIX PFV-return-IV PL.PRON
'These three small happy children of Malin, they came back.'

A notable aspect of Q'anjob'alan languages is their extensive system of nominal classification, described at length in Day 1973; Craig 1977, 1986b, 1990; Ramsay 1985, Zavala 1992,

2000; and Hopkins 2012b. As an example, consider the morphemes that classify the noun *unin* 'child' in (24). First, *-wanh* is one of several numeral classifiers, which in Chuj occur with Mayan-based numerals above 'one', and signals that the noun is animate. Second, *heb'* is a plural marker that only appears with human-denoting nominals (in San Mateo Chuj). Finally, the noun classifier *winh* indicates that the noun is male.[4] Crucially, I assume that these items belong to different syntactic projections, as evidenced by the fact that they (i) make reference to different classes, and (ii) may sometimes co-occur. Of these three grammatical items, I will only focus on *noun* classifiers, which play a critical role in the formation of definite expressions in Chuj. For more information about numeral classifiers in Chuj and other Mayan languages, see Hopkins 1970 and Little, Moroney, and Royer 2022b.

In addition to the grammatical items in (21), a number of derivational morphemes can be found on nominal stems. This is the case in (22), with the agentivizing prefix *aj-*. Another example is the nominal affix *-Vl* in (23) (the vowel of which is partially harmonic with the last vowel on the nominal). This suffix is extremely common in Chuj, and can be found on nominals for a number of reasons, such as marking inalienable nouns and so-called abstractive nominal expressions that do not denote individuals (see Royer et al. 2022 and Felipe Gómez 2022 for discussion of this affix). See Hopkins 1967 and Maxwell 1981 more generally for extensive lists of derivational morphology in Chuj, including those involved in deriving nominals.

[4] Zavala (2000), on the related language Akatek, describes a fourth category of classificatory morphology: so-called "sortal classifiers" like in (i). These classifiers, also attested by the dozens in Chuj (Hopkins 1970, 2012a), are derived from an open class of positional roots and convey a wide range of meanings associated with the inherent shape or position of the entities described by the nominal predicate. Such sortal classifiers are also found in other Mayan languages (Berlin 1968; Arcos López 2009; Bale et al. 2019).

(i) kaa-eb' **xoyan** ixim paat.
two-NUM.CLF SORT.NUM.CLF CLF tortilla
'two tortillas.' (Akatek; Zavala 2000: 125)

While these morphemes are also found in Chuj, they frequently take the position of the numeral classifier, which to my knowledge is also the case in other Q'anjob'alan languages. See Little, Moroney, and Royer 2022b for further discussion of these items in Chuj.

The main items of study in Chapters 3–5 of this thesis will be (i) noun classifiers, (ii) indefinite determiners, and (iii) deictic particles. These items are of particular interest for the questions posed in the introduction, since they appear to play a critical role in the composition of different nominal expressions in Chuj. As indicated in Table 2.4, these items interact to deliver different kinds of subtle semantic distinctions.

Table 2.4: Singular DP configurations in Chuj

	Configuration	Example	Rough translation
①	CLF + NP	*nok' tz'i'*	'the dog'
②	CLF + NP + DEIX	*nok' tz'i' chi'*	'the/that dog' (anaphoric)
③	CLF	*nok'*	'it'
④	INDF + NP	*jun tz'i'*	'a dog'
⑤	INDF + CLF + NP	*jun nok' tz'i'*	'a (certain) dog'
⑥	INDF + NP + DEIX	*jun tz'i' chi'*	'that dog'
⑦	INDF + CLF + NP + DEIX	*jun nok' tz'i' chi'*	'that dog' (exophoric)

Before turning to an analysis of the syntax and semantics of the configurations in Table 2.4, I provide a brief description of each of the items in this table. Specifically, §2.5.1 provides a description of noun classifiers, §2.5.2 a description of indefinite determiners, and §2.5.3 a description of deictic particles and the formation of adnominal demonstratives in Chuj. In §2.5.4, I also provide a short description of possessive constructions in Chuj, which will play an important role in Chapter 6 of this thesis. Note that I will mostly only focus on *singular* interpretations of nouns in the rest of this dissertation, leaving plural interpretations for future research.

2.5.1 Noun classifiers

The Mayan languages spoken in the Cuchumatanes highlands of Guatemala, including both Q'anjob'alan and Mamean languages, are unique within the family in having developed a system of noun classifiers not found in Proto-Mayan. According to Hopkins (2012b), this system of noun classifiers arose as a diffusion from contact with Chiapanec, an Otomanguean language formerly spoken in the region. These items are typologi-

cally rare, being attested in only a limited set of language families, and are unlike more commonly studied *numeral* classifiers inasmuch as their presence or absence in a nominal expression is not connected to numerals or quantificational expressions (Aikhenvald 2000, Grinevald 2000). Noun classifiers have received a good deal of attention in work on Mayan languages from the Q'anjob'alan sub-branch (Craig 1977, 1986b, 1990; Ramsay 1985, Buenrostro et al. 1989, Zavala 1992, 2000, Hopkins 2012b, Buenrostro 2017; see also England 2017 and Lemon 2019 on Mamean languages). The dialect of Chuj under study in this work has sixteen such classifiers, provided in Table 2.5. All noun classifiers closely resemble a noun in the language, a fact that Hopkins (2012b) attributes to the recent development of the noun classifier system. For instance, *ix*, the classifier for female entities, is homophonous with the noun *ix* 'woman', and *nok'*, the classifier for animals, is homophonous with *nok'* 'animal'.

Table 2.5: *Chuj noun classifiers* (see also Hopkins 2012b)

CLF	**Introduces**	**Example**	
ix	female individual/proper name	*ix chichim*	'the elder (f.)'
winh	male individual	*winh icham*	'the elder (m.)'
w(inh)aj	masculine proper names	*waj Matin*	'Mateo'
nok'	animals & derived products	*nok' nholob'*	'the egg'
te'	wood & related entities	*te' k'atzitz*	'the log'
anh	plants & related entities	*anh paj'ich*	'the tomato'
k'en	stone/metal & related entities	*k'en tumin*	'the money'
lum	earth & related entities	*lum yaxlu'um*	'the mountain'
ch'anh	vines & related entities	*ch'anh hu'um*	'the paper'
ixim	corn & related entities	*ixim wa'il*	'the tortilla'
atz'am	salt & related entities	*atz'am atz'am*	'the salt'
ha'	liquids	*ha' melem*	'the river'
k'ak/k'apak	cloth(es)	*k'ak nip*	'the huipil'
k'inal	rain	*k'inal nhab'*	'the rain'
naj/ni'o'	young (male) individual/proper name	*ni nene*	'the (m.) baby'
uch/utni	young (female) individual/proper name	*uch nene*	'the (f.) baby'

As made clear in the above table, the choice of the noun classifier depends on physical or social properties of the noun. Examples of many of the noun classifiers in the table are provided below. The examples were all extracted from the corpus of Chuj narratives and

interviews from Mateo Pedro and Coon (2018). Note that while I indicate for purposes of illustration class membership in a subset of examples in this chapter, the remaining chapters will gloss all noun classifiers uniformly as "CLF" to diminish the size of glosses.

(25) **Classifier for female individuals / proper names (*ix*)** (txt, CD300715)

Ha'onhxo pax ixkok'ayb'ej heb' ix komagrona.

Ha'-onh-xo pax ix-ko-k'ayb'ej heb' **ix** komagrona.
TOP-B1P-ADV also PFV-A1P-teach PL CLF.FEM midwife
'We were the ones who taught the midwives.'

(26) **Classifier for male individuals (*winh*)** (txt, CP200715)

Ichok ta ay jun ilya', tonse yujchi' ay winh lesalwum.

Ichok ta ay jun ilya', tonse yujchi' ay **winh** lesal-wum.
Like if EXT INDF illness, then why EXT CLF.MASC payer-AG
'Now if there's an illness, that's why the "rezador" (the one who prays) exists.'

(27) **Classifier for male proper names (*waj/winhaj*)** (txt, CJ210715)

Ixcham heb' winh anima', tzyal waj Petul tik.

Ix-cham heb' winh anima', tz-y-al **waj** Petul tik.
PFV-die PL CLF.MASC person IPFV-A3-say CLF.MASC.PROPER Petul DEIX
"'The men died", says Petul.'

(28) **Classifier for animal entities (*nok'*)** (txt, CP260715)

Ixlajw el nok' komiston yoj nok' choj.

Ix-lajw el **nok'** ko-miston yoj **nok'** choj.
PFV-finish DIR.out CLF.ANIMAL A1P-cat by CLF.ANIMAL cougar
'Our cat died because of the cougar.'

(29) **Classifier for wooden entities (*te'*)** (txt, CM300715)

Tzinchonh pax te' niwak son.

Tz-in-chonh pax **te'** niwak son.
IPFV-A1S-sell also CLF.WOOD big marimba
'I also sell the big *marimbas*.'

(30) **Classifier for plant entities (*anh*)** (txt, CP010815)

Tzk'e' te' china t'a yib'an schamnak, ma anh tzoyol, ma anh keneya.

Tz-k'e **te'** china t'a yib'an s-cham-nak, ma **anh** tzoyol,
IPFV-climb CLF.WOOD orange PREP over A3-die-PFV, or CLF.PLANT chayote
ma **anh** keneya.
or CLF.PLANT banana
'One puts the orange over the deceased relative, or the chayote, or the banana.'

(31) **Classifier for stone entities (*k'en*)** (txt, CD300715)

Max hinyam laj k'en machit, k'en ch'akab', k'en asadon.

Max hin-yam laj **k'en** machit, **k'en** ch'akab', **k'en**
NEG.IPFV A1S-grab NEG CLF.STONE machete, CLF.STONE axe, CLF.STONE
asadon.
hoe
'I don't use the machete, or the axe, or the hoe.'

(32) **Classifier for earth entities (*lum*)** (txt, CP090815)

Ichachi' tz'aj kob'at t'a lum desyerto chi'.

Icha-chi' tz'-aj ko-b'at t'a **lum** desyerto chi'.
like-DEIX IPFV-be A1P-go PREP CLF.EARTH desert DEIX
'This is how we'd get to the desert.'

(33) **Classifier for paper entities (*ch'anh*)** (txt, CP280715)

Yelta ch'anh titulo prob'isyonal chi' t'ayonh.

Y-elta **ch'anh** titulo prob'isyonal chi' t'ay-onh.
A3-come.out CLF.PAPER certificate temporary DEIX PREP-A1P
'The temporary certificate came to us.'

(34) **Classifier for corn entities (*ixim*)** (txt, CP010815)

Tzyak'an tratar heb' ix to tzk'e'ta ixim ojlem.

Tz-y-ak'-an tratar heb' ix to tz-k'e'ta **ixim** ojlem.
IPFV-A3-give-DEP try PL CLF.PRON COMP IPFV-take.out CLF.CORN tamal
'They (fem.) try to bring up the tamales.'

(35) **Classifier for salt entities (*atz'am*)** (txt, CM210715)

Olonhb'atxok kik' atz'am katz'am.

Ol-onh-b'at-x-ok k-ik' **atz'am** k-atz'am.
PROSP-B1P-go-ADV.now-IRR A1P-bring CLF.SALT A1P-salt
'Now we're going to bring our salt.'

(36) **Classifier for water entities (*ha'*)** (txt, CD300715)

Tzsb'o'an kot ha' k'ak'ala' winh hinmam.

Tz-s-b'o'-an kot **ha'** k'ak'ala' winh hin-mam.
IPFV-A3-make-SUF DIR.arrive CLF.WATER hot.water CLF.MASC A1S-father
'My father prepared the hot water.'

(37) **Classifier for cloth entities (*k'apak/k'ak*)** (txt, CP260715)

Ichaton tik yik'ti'al k'apak nip.

Icha-ton tik y-ik'ti'al **k'apak** nip.
like-COMP DEIX A3-story CLF.CLOTH huipil.
'This is the story of the *huipil*.'

In all of the above examples, we see that the presence of a noun classifier before a noun gives rise to a definite translation. Such uses, which have long been noticed in the literature on Q'anjob'alan languages (Craig 1986b, 1990; Buenrostro et al. 1989; Zavala 1992; García Pablo and Domingo Pascual 2007; Buenrostro 2017), will be the focus of Chapter 3. However, as shown in Table 2.4 from page 28, and as also noted in previous work on other Q'anjob'alan languages, this is not the only use of noun classifiers. For one, noun classifiers can combine with other functional items, including indefinite determiners (discussed in §2.5.2 and Chapter 4) and deictic particles (discussed in §2.5.3, as well as Chapters 3 and 5).

Second, noun classifiers function more generally as third person anaphoric pronouns, appearing without an overt noun to refer back to previously introduced referents. I will refer to such uses of noun classifiers as "classifier pronouns", and, for convenience, will gloss these as "CLF.PRON" in examples. In (38-b), notice that classifier pronouns must be used to refer back to the two referents introduced in the prior sentence (38-a).

(38) a. *Ixsman jun onh ix ix.*
Ix-s-man [$_{OBJ}$ jun onh] [$_{SUBJ}$ ix ix].
PFV-A3-buy INDF avocado CLF woman
'The woman bought an avocado.'

b. *Ixslo'an te' ix.*
Ix-s-lo'-an [$_{OBJ}$ **te'**] [$_{SUBJ}$ **ix**].
PFV-A3-eat-CON CLF.PRON CLF.PRON
'And then she ate it.'

Within the Mayan language family, the requirement for overt non-emphatic third person pronouns is an innovation. Most Mayan languages are robustly *pro*-drop (Coon 2016a), a fact that will also be discussed in detail in Chapter 6. In Chapter 6, I will further describe other pronominal expressions in the language, including pronominal expressions that must be employed instead of classifier pronouns when two nominal expressions are covalued within the same sentence.

It is important to note that the choice of the noun classifier can sometimes have repercussions for how a nominal expression gets interpreted. This is the case in the examples in (39) and (40). In (39), the choice of the classifier affects whether the nouns *unin* 'child' and *doktor* 'doctor' are interpreted as female or male. In (40), the choice of the classifier discriminates between two types of tomatoes: those which grow on plants (*anh paj'ich* 'tomato') and those which grow on trees (*te' paj'ich* 'tomarillo'):

(39) Alternation between masculine/feminine nominal expressions

a. *Ixyanhtej winh unin ix doktor.*
Ix-y-anhtej **winh** unin **ix** doktor.
PFV-A3-cure CLF.MALE child CLF.FEMALE doctor
'The (female) doctor cured the boy.'

b. *Ixyanhtej ix unin winh doktor.*
Ix-y-anhtej **ix** unin **winh** doktor.
PFV-A3-cure CLF.FEMALE child CLF.MALE doctor
'The (male) doctor cured the girl.

(40) Alternation between plant/tree nominal expressions

a. *Ixinlo' anh paj'ich.*
Ix-in-lo' **anh** paj'ich.
PFV-A1S-eat.sweet CLF.PLANT tomato
'I ate the tomato (standard tomato; *solanum lycopersicum*).'

b. *Ixinlo' te' paj'ich.*
Ix-in-lo' **te'** paj'ich.
PFV-A1S-eat.sweet CLF.WOOD tomato
'I ate the tomarillo (tree tomato; *cyphomandra betacea*).'

Crucially, however, this is only true of a subset of nouns. Other nouns are more rigid with regards to the classifier they admit, even when the classifier appears to contradict the material with which the referent of the noun is made, suggesting that noun classifiers have undergone some level of "semantic bleaching" in the process of grammaticalizing as functional items (Craig 1986b, Zavala 2000). Consider the case of *pat* 'house'. In Huehuetenango, it is common for houses to be made of either wood, stone, or mud, which all else being equal, could take three different kinds of classifiers (*te'* (wood), *k'en* (stone), or *lum* (earth/soil)). Nonetheless, *pat* can only combine with the classifier for wooden entities, *te'* (41-a). As shown in (41-b), this is so even when it is immediately asserted that the house in question is not made of wood.

(41) a. *te' pat*
{**te'/*k'en/*lum**} pat
CLF.WOOD/CLF.STONE/CLF.EARTH house
'the house'

b. *¡Tewach' kilan te' pat chi'! Ke'ennhej ix'aj te'.*
¡Te'-wach' k-il-an **te'** pat chi'! K'e'en-nhej ix'-aj
INTS-good A1P-see-DEP CLF.WOOD house DEIX stone-only PFV-be
te'.
CLF.PRON.WOOD
'The house looks great! It's made entirely of stone.'

Moreover, there is a sense in which there are "default classifiers" for each noun. For instance, when a table is entirely made of plastic, for which there is no corresponding classifier in the dialect of Chuj under study (though see Hopkins 2012b on a possible

classifier for plastic items), the classifier for wooden objects, and not the one for metal objects, must be used. This is so even though it is otherwise possible to use the classifier for metal objects with tables that are made of metal.

(42) **Context:** Talking about a plastic table we saw on the Ikea website:
Maj hinman laj te' mexa, yojto kawcho te'.
Maj hin-man laj {**te'**/#**k'en**} mexa, yojto kawcho {**te'**/***k'e'en**}.
NEG.PFV A1S-buy NEG CLF.WOOD/STONE table, because plastic CLF.PRON
'I didn't buy the table, because it's made of plastic.'

Finally, note that not all nouns have a corresponding noun classifier. Nouns that typically do not take a classifier are summarized below.

(43) **Nouns that typically do not take a classifier** (based on Zavala 2000 for Akatek)
a. Abstract nouns (e.g., friendship, religion, party)
b. Nominalizations (e.g., buying, sleeping, dancing)
c. Temporal expressions (e.g., year, month)
d. Some locative nouns (e.g., market, temple, place names)
e. Body parts (with the exception of bones and teeth)
f. Nouns that do not fall in a classificatory category (non-binary person, the word for food in general).

Examples of nouns that never combine with a noun classifier are provided in (44)–(46). In (44), we see that the nouns *relijyonh* 'religion' and *k'inh* 'party' do not have corresponding classifiers. The same goes for the noun *iglesya* 'church' in (45) and non-binary or mixed-gender uses of the word *doktor* 'doctor' in (46). In such cases, definite nominals appear without a noun classifier.

(44) *Haxo ixjawi relijyonh tik t'a kokal jun yo, kasi ixyak' ewitar k'inh heb' winh.*
Haxo ix-jaw-i **relijyonh** tik t'a ko-kal jun yo, kasi
and.when PFV-arrive-IV religion DEIX PREP A1P-middle DISC DISC, almost
ix-y-ak' ewitar **k'inh** heb' winh.
PFV-A3-do avoid party PL CLF.PRON
'And when the religion arrived among us, you know, then they almost avoided the parties.' (txt, CP010815)

(45) *Mantalaj iglesya tik.*

Man-ta-laj **iglesya** t'atik.
NEG-still-NEG church here
'The church wasn't here yet.' (txt, CP2201715)

(46) *Yowalil ay impormasyonh tzsk'an heb' doktor.*

Yowalil ay impormasyonh tz-s-k'an heb' **doktor**.
obligatory EXT information IPFV-A3-ask PL doctor
'There's information that the doctors must ask.' (txt, CP300715)

Since a major goal of this thesis will be to provide syntactic and semantic generalizations about the distribution of noun classifiers, I will focus exclusively on nominals that have an attributed classifier.

2.5.2 Indefinite expressions

The previous literature on Q'anjob'alan languages generally describes two indefinite determiners (see e.g. Craig 1986b on Popti'; Zavala 1992 on Akatek; Mateo Toledo 2017 on Q'anjob'al; and García Pablo and Domingo Pascual 2007 and Buenrostro 2022 on Chuj). The morpheme *jun*, homophonic with the numeral 'one', is used as a singular indefinite determiner, whereas the plural indefinite determiner *juntzanh* 'some' is generally used as a plural indefinite determiner.[5] In addition to the plural indefinite *juntzanh*, *jun* can also be reduplicated to deliver a plural indefinite determiner.[6] Examples of these determiners within existential clauses are provided in (47). As shown in (47-a), the singular indefinite determiner is required to obtain a singular indefinite interpretation. As for plural indefinites (47-b)–(47-c), neither *juntzanh* or *junjun* are required to obtain a plural indefinite interpretation. While I will exclusively focus on singular nominal expressions in Chapters 3-6 of this thesis, it remains to be understood what semantic effects arise when plural

[5]Cognate forms of *jun* as a singular indefinite determiner are widespread across the Mayan language family. See for instance Henderson 2014 on Kaqchikel, Can Pixabaj 2017 on K'iche', England 2017 on Mam, and Curiel Ramírez del Prado 2017 on Tojol-ab'al.

[6]Interestingly, reduplication of *jun* in Chuj can also deliver a distributive universal quantifier, which best translates as 'each'. The existential/universal ambiguity of *junjun* is also observed in other Mayan languages, such as Kaqchikel (see Henderson 2014: 47-49).

indefinite determiners are omitted, and what kinds of semantic distinctions are expressed by the use of *juntzanh* versus *junjun*.

(47) a. *Ay jun tz'i' t'atik.*
Ay [#**(jun)** tz'i'] t'a-t'ik.
EXT INDF dog PREP-here
'There's a dog here.'

b. *Ay juntzanh tz'i' t'atik. / Ay tzi'i' t'atik.*
Ay [**(juntzanh)** tz'i'] t'atik.
EXT some dog here
'There are (some) dogs here.'

c. *Ay junjun tz'i' t'atik.*
Ay [(**junjun**) tz'i'] t'atik.
EXT some dog here
'There are (some) dogs here.'

Singular nominal expressions in thematic positions of finite verbs likewise require the presence of the singular indefinite determiner to give rise to an indefinite interpretation. Examples taken from narratives are provided below. The indicated acceptability judgements have been verified with speakers with relevant contexts.

(48) a. *Elanhchamel tzb'o'il jun nip.*
Elanhchamel tz-b'o'-il #**(jun)** nip.
quickly S-make-NMLZ INDF huipil
'They finish a huipil quickly.' (txt, CM201115)

b. *Haxo winh ak'an jun wakax.*
Ha-xo winh ak'-an #**(jun)** wakax.
FOC-and CLF.PRON give-AF INDF cow
'And HE donated a cow.' (txt, CJ2201715)

c. *Haxo winh tzmanan jun kantela.*
Ha-xo winh tz-man-an #**(jun)** kantela.
FOC-and CLF.PRON IPFV-buy-AF INDF candle
'And he buys a candle.' (txt, CJ240715)

The determiner *jun* behaves like an indefinite quantifier in other languages in its ability to take scope under modal or negative operators. The sentence in (49) can either con-

vey that the speaker did not see any priest at all (50-a), the low-scope interpretation, or that there is a priest such that they did not see them (50-b), the wide-scope interpretation.[7]

(49) *Maj wil laj jun pale.*
Maj w-il laj **jun** pale.
NEG.PFV A1S-see NEG INDF priest
'I didn't see a priest.'

(50) Felicitous contexts for (49)
a. ✓ $\neg\exists x$[PRIEST(x) $\wedge$ SEE(x)(SPEAKER)]
Context: I wanted to see a priest, but never saw one. I tell you (49).
b. ✓ $\exists x$[PRIEST(x) $\wedge \neg$SEE(x)(SPEAKER)]
Context: I saw several priests, but there's one particular priest I didn't see. I tell you (49).

Note that low-scope interpretations like (50-a) can be forced when the singular indefinite combines with the irrealis clitic *=ok*, resulting in an interpretation akin to indefinite uses of *any* in English. The semantic and syntactic status of *=ok* remains to be understood. In Chuj, it can encliticize to a wide range of grammatical items whenever these items appear under the scope of irrealis mood markers (see §5.2.2).[8]

(51) *Maj wil laj junok pale.*
Maj w-il laj **jun=ok** pale.
NEG.PFV A1S-see NEG INDF-IRR priest
'I didn't see any priest (not even one).'

Some speakers show a preference for the use of *jun=ok* in certain low-scoping environments, including negative sentences and polar questions (see Buenrostro 2022, Buenrostro and Royer 2022), an observation that has also been made for Q'anjob'al (Bervoets 2014, Becker 2021). We will return to a discussion of *jun=ok* indefinites in Chapter 5 (§5.2.2),

[7] As will be discussed at length in Chapter 4, the presence of a noun classifier in the indefinite DP in (49) would force a wide scope interpretation, and would therefore only be compatible with the context in (50-b) and not (50-a).

[8] *=ok* can encliticize to a number of items from different lexical categories in Chuj, including verbal predicates in certain aspects, non-verbal predicates, numerals, quantifiers, directional particles, and a number of adverbs, whenever these items appear under the scope of an "irrealis" operator (Buenrostro 2015a). Other Mayan languages share the same cognate morpheme (see e.g. Polian 2007 and Vinogradov 2012).

where I show that there seems to be some amount of linguistic variation within the Q'anjob'alan language family with regards to the distribution of *jun=ok*.

As discussed at length in Chapter 4, indefinite determiners can co-occur with noun classifiers in Chuj. This is the case in the two examples in (52), taken from narratives (see §4.1.2 for evidence that such constructions do not involve a partitive syntax).

(52) **Indefinite+classifier DPs**

a. *Ayek' jun winh pale tzkuchan Pegre Mundo.*
Ay-ek' **jun winh** pale tz-kuchan Pegre Mundo.
EXT-DIR.pass INDF CLF.MALE priest IPFV-called Father Mundo
'There was a priest called Father Mundo.' (txt, CJ2201715)

b. *Ay jun k'en nhak'e'en hab' b'aj ayem heb' winh.*
Ay **jun k'en** nhak'e'en hab' b'aj ay-em heb' winh.
EXT INDF CLF.STONE cave REP where EXT-DIR.down PL CLF.PRON
'Apparently there's a cave where they'd be.' (txt, CP2601715)

Indefinite determiners can also combine with classifier pronouns to derive indefinite pronouns. This can be observed in the examples below. Notice how *jun* can combine with a classifier pronoun in examples (53-b)–(53-d).[9]

[9] As also noted in Ramsay 1985 (footnote 3) on Popti', indefinite pronouns sometimes appear to combine with a classifier pronoun to introduce a new referent in discourse. This only ever seems to happen when the head noun would be *winak* 'man'.

(i) *Ay jun winh tzkuchan Kun Chab'in.*
Ay jun winh s-kuchan Kun Chab'in.
EXT INDF CLF IPFV-called Kun Chab'in
'There a man called Kun Chab'in.' (txt, CP280715)

(53) **Co-occurrences of *jun* with classifier pronouns** (txt, CP200715)

a. *Ay juntzanh heb' winh witzak'lik, sayb'il yuj heb' winh t'atik.*
Ay juntzanh heb' winh **witzak'lik**, say-b'-il yuj heb'
EXT INDF.PL PL CLF earthlords look.for-DERIV-NMLZ by PL
winh t'atik.
CLF.PRON here:
'There were some earthlords that were searched for here:'[10]

b. *jun winh b'e'al Onh tik,*
jun winh b'e'al Onh tik,
INDF CLF.PRON trail Aguacate DEIX
'one on the trail to Aguacate,'

c. *jun winh b'e'al Sambram,*
jun winh be'al Sambram,
INDF CLF.PRON trail San Francisco
'one on the trail to San Francisco,'

d. *jun winh b'e'al t'a Tzalantaj tik.*
jun winh b'e'al t'a Tzalantaj tik.
INDF CLF.PRON trail PREP Tzalantaj DEIX
'and one on the way to the trail to Tzalantaj.'

A major contribution of Chapter 4 will be to pinpoint the semantic contribution of nominal expressions that bear both an indefinite determiner and a classifier, versus indefinite expressions that appear without a noun classifier. As mentioned above, I will only focus on singular expressions, leaving an investigation of plural nominal expressions for the future. However, it is important to emphasize that, except for nouns that do not denote humans (see (56) below), nouns may also sometimes combine alone with a noun classifier to deliver a plural indefinite. In fact, in such cases, an ambiguity arises between a (singular/plural) definite interpretation and (plural) indefinite interpretation. This is shown in (54) with the three possible translations.

[10]See Hopkins (2012b) on the term *witzak'lik*, which refers to spiritual beings living in the mountains of the Cuchumatanes who are believed to protect the Chuj people. The term *witzak'lik*, sometimes written *witz ak'lik*, literally translates as 'mountain plain'. It belongs to a class of *difrasismos* in Chuj, a common metaphorical device across Mesoamerica which involves the compounding of two lexical items in order to create metaphorical meaning (see e.g. Montes de Oca Vega 2004).

(54) *Ixkil nok' tz'i' t'achi'.*
Ix-k-il [**nok'** tz'i'] t'achi'.
PFV-A1P-see CLF dog there
'We saw {the dog / the dogs / dogs / *a dog} over there.'

The nominal expression *nok' tz'i'* in (54) can yield either a definite description, in either the singular or plural, or a plural indefinite. It cannot yield a singular indefinite, since, as mentioned above, singular indefinites require an overt indefinite determiner.[11] Finally, note that bare nouns can alternatively appear without a noun classifier, in which case only a plural indefinite interpretation arises:

(55) *Ixkil tz'i' t'achi'.*
Ix-k-il [tz'i'] t'achi'.
PFV-A1P-see dog there
Can only mean: 'We saw dogs over there.'

The singular/plural ambiguity does not arise with human-denoting nouns. Singular human-denoting nouns require a classifier when definite, as shown in (56). For plural interpretations, human denoting nouns require the presence of the plural marker *heb'* whenever a noun classifier is present. Moreover, the presence of *heb'* goes hand-in-hand with the presence of a noun classifier. That is, whenever a plural noun is classifiable, *heb'* can only appear if a classifier is present, and vice-versa. Expressions in thematic positions that have the form *heb'* + classifier + noun, however, remain ambiguous between an indefinite and definite interpretation, as indicated in the translation in (56-b).

[11] This ambiguity only arises in the San Mateo Ixtatán dialect of Chuj—the one under study in this work—and not the San Sebastián Coatán dialect, since the San Sebastián Coatán dialect shows overt plural marking on all nouns (human, animate, and inanimate; see e.g. García Pablo and Domingo Pascual 2007).

(56) a. *Ixkil winh ichamwinak t'achi'.*
Ix-k-il [**winh** ichamwinak] t'achi'
PFV-A1P-see CLF elder.man there
Can only mean: 'We saw the elder man over there.'

b. *Ixkil heb' winh ichamwinak t'achi'.*
Ix-k-il [**heb' winh** ichamwinak] t'achi'.
PFV-A1P-see PL CLF elder.man there
'We saw {the elder men/elder men} over there.'

The possible interpretations of Chuj nominals without overt indefinite determiners are summarized below.

Table 2.6: Possible interpretations of nominals without overt indefinite determiners

	human					non-human			
	DEF		INDF			DEF		INDF	
	SG	PL	SG	PL		SG	PL	SG	PL
① N	✗	✗	✗	✓		✗	✗	✗	✓
② CLF+N	✓	✗	✗	✗		✓	✓	✗	✓
③ *heb'*+CLF+N	✗	✓	✗	✓		✗	✗	✗	✗

Table 2.6 shows several important properties. First, bare nouns in Chuj, shown in row ①, can only give rise to bare plural interpretations. Second, row ② shows that there is never ambiguity between *singular definites* and *singular indefinites*. This is shown more specifically with grey shading. Again, the reason for this is that singular indefinites require the presence of an overt indefinite determiner. Finally, there is sometimes a three-way ambiguity between *singular/plural* definites and *plural* indefinites. This is the case for non-human nominal expressions, where we see that CLF+N constituents can give rise to either (i) a singular definite description, (ii) a plural definite description, or (iii) a bare plural indefinite. However, as also shown on the same row, human denoting nominals that exhibit a CLF+N template are never ambiguous: they can only convey a singular definite description, because they require the presence of the overt plural marker *heb'* (see row ③). This fact is further reflected in the ungrammatical existential sentence in (57) below, with the nominal expression *winh munlajwum* 'the worker':

(57) *Ay [**winh** munlajwum] t'a te' hin-pat.
EXT CLF worker PREP CLF A1S-house
Literal intended translation: 'There's the worker in my house.'

Example (57) shows that human denoting classifier + noun configurations are ungrammatical in existential sentences (see also discussion in Buenrostro and Royer 2022). The ungrammaticality of (57) is entirely expected from a cross-linguistic perspective: definite nominals tend to be ungrammatical in existential constructions (see e.g. Milsark 1974, Enç 1991, McNally 2016). Given that human denoting classifier + nominal expressions necessarily give rise to a singular definite interpretation, it makes sense that such nominal expressions are not admitted in existential constructions.

In sum, for a subset of (non-human denoting) nominals, there is some overlap between the expression of singular/plural definiteness and plural indefiniteness in Chuj. As already mentioned, the rest of this work will focus on the syntactic and semantic distribution of singular nominal expressions, for which definite/indefinite ambiguities never arise. I leave an investigation of plurality in Chuj for future work.

2.5.3 Deictic particles and adnominal demonstratives

Demonstrative constructions will play an important role in Chapters 3 and 5 of this dissertation. As discussed in Buenrostro and Royer 2022, to form an adnominal demonstrative construction in Chuj, it is necessary to combine one of two deictic particles, *tik* or *chi'* (uniformly glossed "DEIX" below), with a syntactic base showing the characteristic properties of either a definite or indefinite nominal expression. The two options are shown below:

(58) Singular adnominal demonstrative with definite base

a. Form: [*(CLF) N DEIX]

b. *Wojtak winh winak tik.*
W-ojtak [**winh** winak **tik**].
A1S-know CLF man DEIX
'I know this man.'

c. *Wojtak winh winak chi'.*
W-ojtak [**winh** winak **chi'**].
A1S-know CLF man DEIX
'I know that man.'

(59) Singular adnominal demonstrative with indefinite base

a. Form: [*(*jun*) (CLF) N DEIX]

b. *Wojtak jun winak tik. / Wojtak jun winh winak tik.*
W-ojtak [**jun** (winh) winak **tik**].
A1S-know INDF CLF man DEIX
'I know this man.'

c. *Wojtak jun winak chi'. / Wojtak jun winh winak chi'.*
W-ojtak [**jun** (winh) winak **chi'**].
A1S-know INDF CLF man DEIX
'I know that man.'

As hinted in the above translations, the choice between *tik* versus *chi'* is determined by distance and visibility with respect to a deictic centre, which in Chuj is always the speaker. While *tik* is used for referents that are visibly close to the speaker, *chi'* is used with those that are visibly far from the speaker or not visible to them at all. In anaphoric uses of adnominal demonstratives, which do not necessarily involve a referent in the immediate situation, the distal particle is usually preferred over the proximal particle (see Buenrostro and Royer 2022 for details). As also shown in the above examples, noun classifiers can optionally combine with *jun* to form adnominal demonstratives with indefinite bases. This parallels the apparent syntactic optionality of noun classifiers in indefinite phrases, which was described in the preceding section.

From the perspective of the existing literature on adnominal demonstratives (e.g. Diessel 1999, Roberts 2002, Wolter 2006, Elbourne 2008), the formation of demonstrative con-

structions in Chuj shows interesting properties. For one, adnominal demonstratives do not seem to form a "primitive". That is, the deictic particles must always combine with other functional items within the extended nominal domain, either with a noun classifier in the case of (58) or with an indefinite determiner in the case of (59). Bare demonstratives are not allowed on (classifiable) nouns:

(60) **Context:** The speaker is pointing at one man that is close and visible to them, and says:
Wojtak winak tik.
W-ojtak [winak tik].
A1S-know man DEIX
'I know this man.'

As I will argue in the coming chapters, this suggests that "demonstratives" do not form a primitive category in Chuj, but are rather derived compositionally (as proposed in recent work on other languages, including Hanink 2018, 2021 and Ahn 2019).

Second, taking into consideration the fact that demonstratives have been traditionally analyzed as a subtype of definite article, an analysis that is maintained in much recent work on demonstratives (Roberts 2002, Wolter 2006, Elbourne 2008, 2013, Ahn 2019), it is somewhat surprising that adnominal demonstratives in Chuj sometimes combine with indefinite determiners (see also Can Pixabaj 2017 on this possibility in K'iche' and Arsenijević 2018, or example (8) above, for this possibility in Serbo-Croatian).[12] As a matter of fact, the ability for *jun* to occur in demonstrative phrases could be taken to cast doubt on the proposal that *jun* is an indefinite determiner in the first place. Indeed, comparable data in Akatek, a close relative of Chuj, led Zavala (2000: 118) to propose that the cognate of *jun* in Akatek "functions as a *singulative* (a marker which restricts the reference to a single entity) and not as an indefinite marker".

[12]However, it should be stressed that, while they have received considerably less attention, "indefinite" uses of demonstratives in English have long been documented (e.g. *there's this guy I know who speaks dozens of languages*). See for example Prince 1981, Maclaran 1982, Gundel et al. 1993, Matthewson 1998, 2008, 2013, Ionin 2006, Arsenijević 2018, as well as Doran and Ward 2019 more generally on the heterogeneity of demonstrative uses.

Nonetheless, there is good independent reason to believe that deictic particles in Chuj *do* have the ability to combine with indefinite expressions. For one, note that both the plural indefinite determiners *juntzanh* and *junjun* may likewise combine with deictic particles to derive plural adnominal demonstratives:

(61) Plural admoninal demonstratives

a. *Ixinman juntzanh chinhb'il tik.*
Ix-in-man [**juntzanh** chinhb'il **tik**].
PFV-A1S-buy INDF.PL guitar DEIX
'I bought these guitars.

b. *Ixinman junjun chinhb'il tik.*
Ix-in-man [**junjun** chinhb'il **tik**].
PFV-A1S-buy INDF.PL guitar DEIX
'I bought these guitars.'

Second, *tik* and *chi'* may combine with other expressions that otherwise yield existential meaning, such as *jab'/jak* 'few/little' and *jaye'* 'a good amount of' (also the interrogative word meaning 'how much'). While more work is needed to understand the semantic contribution of such demonstrative constructions, relevant examples are provided below. While the (a) examples show indefinite expressions in their use without a deictic particle, the (b) examples show their use in combination with a deictic particle.

(62) a. *Ay jab' smunlajel winh t'a K'axep chi'.*
Ay [**jab'** s-munlajel] winh t'a K'axep chi'.
EXT few A3-work CLF.PRON PREP K'axep DEIX
'He had little work in K'axep.' (txt, CX200715)

b. *K'ilab' yaj jab' iglesya chi'.*
K'ilab' y-aj [**jab'** iglesya **chi'**].
clay A3-be few church DEIX
'That church was made of mud.' (txt, CP220715)

(63) a. *Ixinman jaye' pat.*
Ix-in-man [**jay-e'** pat].
PFV-A1S-buy how.much-NUM.CLF house]
'I bought a good amount of houses.'

b. *Ixinman jaye' pat chi'.*
Ix-in-man [**jay-e'** pat **chi'**].
PFV-A1S-buy how.much-NUM.CLF house DEIX
'I bought that (good) amount of houses.'

Third, as discussed in Buenrostro and Royer 2022, the semantic and pragmatic conditions on the use of demonstratives within indefinite versus definite expressions diverge in ways that tally with their respective bases. For instance, demonstratives with indefinite bases in Chuj are perfectly acceptable as first mentions of referents in discourse, a prototypical property of indefinites (Heim 1982), and in such cases are admitted in existential constructions. The following example, taken from the first sentence of a story in Hopkins 2021 (recorded in 1965), illustrates this well (glosses and felicity judgement added for illustration).

(64) *Komo ha t'ay pekatax, ay jun winh icham chi'...*
Komo ha t'ay pekatax, ay [#**(jun)** winh icham chi']...
DISC FOC PREP before EXT INDF CLF man DEIX
'Once upon a time, there was an old man...' (Hopkins 2021: 45)

In (64), the referent *jun winh icham chi'* 'this old man', is clearly introduced for the first time in the story. Yet, it can appear with a deictic particle. As indicated by Chuj consultants, omitting the indefinite in this case is not possible.

The sentence in (64) also shows that adnominal demonstratives that combine with indefinite determiners are perfectly acceptable in existential sentences, which, as already mentioned, tend to only admit indefinite expressions across languages (Milsark 1974, Enç 1991, McNally 2016). This practice is quite common in Chuj narratives. Another example from a more recent narrative is provided below:

(65) *Ay jun lugar chi' tzkuchan "Pat B'ojoch" t'a peka'.*
Ay [**jun** lugar **chi'**] tz-kuchan "Pat B'ojoch" t'a peka'.
EXT INDF place DEIX IPFV-called Pat B'ojoch PREP before
'There was this place called *Pat B'ojoch* before.' (txt, CP1901715)

As hinted in (64), using a demonstrative with a definite base in existential constructions is simply ungrammatical. This is shown in (66). This kind of example fits with the general unacceptability of human denoting classifier + noun expressions in existential sentences (see (57) above).

(66) *Ay [**winh** ichamwinak **tik**] t'atik.
EXT CLF elder.man DEIX here
Intended: 'There's this old man here.' (Buenrostro and Royer 2022)

Another particularity of demonstratives with indefinite bases in Chuj, which distinguishes them from demonstratives with definite bases, is that they are overwhelmingly used in cases of *exophora*. "Exophora" describes uses of demonstratives in which the referent is present in the immediate situation and the speaker is pointing at the referent with some form of gesture (see e.g. Diessel 1999, Ahn and Davidson 2018, Ahn 2019). Such uses are also sometimes referred to as "deictic uses" in the formal semantics literature (Heim and Kratzer 1998 and Doran and Ward 2019). In Chuj, when the speaker brings their addressee's attention towards a referent for the first time—most often with a gesture—consultants indicate a clear preference for the demonstrative with an indefinite base:[13]

(67) **Context:** There's only one ant on the shoulder of the addressee, but the addressee hasn't seen the ant. The speaker points to the ant and tells the addressee:

a. *¡Hilnab' jun sanich tik t'a hajenhjab'!*
¡H-ilnab' [**jun** sanich **tik**] t'a ha-jenhjab'!
A2S-look INDF ant DEIX PREP A2S-shoulder
'Look at this ant on your shoulder!'

b. # ¡H-ilnab' [**nok'** sanich **tik**] t'a ha-jenhjab'!
A2S-look CLF ant DEIX PREP A2S-shoulder

[13]The contexts to test exophoric deixis were adopted from Wilkins's (2018) demonstrative questionnaire.

Royer (2019) and Buenrostro and Royer (2022) show that demonstratives with definite bases like (277-b), on the other hand, tend to be used with referents that have already been mentioned in the discourse, and are thus anaphoric (though they can still sometimes be used in exophoric contexts). Such uses are naturally much more common in narrative texts, which do not always involve referents present in the immediate discourse.

In sum, the observations (i) that deictic particles are in general possible with indefinite quantifiers in Chuj, (ii) that they can appear in existential sentences in co-occurrence with indefinite quantifiers, and (iii) that they have different uses depending on whether they combine with a definite base or indefinite base, all support the view that deictic particles in Chuj are compatible with both definite and indefinite bases. A major theme of Chapter 3 will be devoted to this topic, where I show that deictic particles are generally required with anaphoric definite descriptions in Chuj. I will also show that demonstratives with definite bases lead to maximality presuppositions that are not attested with demonstratives that combine with indefinite determiners (see specifically §3.3.2.1).

In Chapter 4, I then return to the use of deictic particles in indefinite expressions, providing a compositional analysis of the pieces involved in the extended nominal domain, which allows for deictic particles to combine with both definite and indefinite expressions.

2.5.4 Possessive constructions

Possessors will play an important role in the discussion of Chapter 6, where I provide an analysis of patterns of nominal binding and coreference in Chuj. I therefore provide brief background on nominal expressions with overt possessors here.

As in other Mayan languages (England 2001, Coon 2016a, Aissen et al. 2017), Chuj possessors trigger Set A (ergative/possessive) morphology as a prefix on the possessee. Notice, for instance, the identical third person Set A prefix on the verb stem and possessees in (68) (see Table 2.3 above for the list of Set A prefixes).

(68) *Hat'a tzsb'o' winh smam winh winak chi' yet' winh smam ix ix chi'.*
Ha-t'a tz-**s**-b'o' [winh **s**-mam [$_{POSS}$ winh winak chi']] yet'
FOC-to.there IPFV-A3-make CLF A3-father CLF man DEIX with
[winh **s**-mam [$_{POSS}$ ix ix chi']].
CLF A3-father CLF woman DEIX
'They put the man's father and the woman's father there.' (txt, CM250715)

As seen in (68), possessors appear on the right of possessees in Chuj.[14] This fact will play an important role in motivating the analysis of binding and coreference in Chapter 6.

Note that possessive marking in Chuj has a more flexible distribution than in English. Like other languages (e.g., Italian, Russian; see Schoorlemmer 1998; Haspelmath 1999), possessors in Chuj can co-occur with indefinite determiners. This is shown in (69).

(69) Possessed indefinites in Chuj

a. *Ha ix Malin, olsman ix jun slibro ni Xun.*
Ha ix Malin, ol-s-man ix [**jun** **s**-libro ni Xun].
TOP CLF Malin PROSP-A3-buy CLF.PRON INDF A3-book CLF Xun
'As for Malin, she'll buy a book of Xun's.'

b. *Ixlajwi jun hinkurso.*
Ix-lajw-i [**jun** **hin**-kurso].
PFV-FINISH-IV INDF A1S-course
'I finished a course of mine.' (txt, CJ240715)

Also as in other languages (e.g. Russian; Freeze 1992), possessed nominals in Chuj appear in existential sentences to convey *have*-predication (see also Elias 2019).

(70) *Ay nok' swakax heb' winh.*
Ay nok' **s**-wakax heb' winh.
EXT CLF A3-cow PL CLF.PRON
'They (masc.) own cows.' (Lit: 'There exists their cows') (txt, CP2801715)

Finally, it is important to note that noun classifiers are always optional in the presence of possessive marking, a fact that is also described by Craig (1977) for Popti'.[15] This is so

[14] Except in the case of *wh*-possessors, which must appear before the possessee, a case of "pied-piping with inversion" (see e.g. Aissen 1996; Coon 2009, Little 2020a).

[15] Optional definite marking in possessed nominals is attested in other languages, such as Brazilian Portuguese (Haspelmath 1999).

even when the nominal expression is arguably interpreted as a definite description (71-a), where noun classifiers are otherwise required (71-b):

(71) a. **Context:** The speaker only has one dog and their addressee knows this:
Ixwil hintz'i' / Ixwil nok' hintz'i'.
Ix-w-il (**nok'**) hin-tz'i'.
PFV-A1S-see CLF A2S-dog
'I saw my dog.'

b. **Context:** The speaker and addressee are staying at a hostel, and have both been advised that there is a dog in the hostel. After just spotting the dog, the speaker tells their addressee:
Ixwil nok' tz'i'.
Ix-w-il *(**nok'**) tz'i'.
PFV-A1S-see CLF dog
'I saw the dog.'

Another example, taken from a narrative, is provided below. In (72-a), we see the possessed nominal *hinman icham* 'my grandfather' with a noun classifier. About one minute later in the conversation, the speaker mentions their grandfather again. As shown in (72-b), they do not use a noun classifier on that second occasion.

(72) Optionality of classifier with possessive morphology (txt, CP190715)

a. *Xun Heb'in sb'i' winh hinmam icham chi'.*
Xun Heb'in s-b'i [**winh hin**-mam icham chi'].
Xun Heb'in A3-name CLF A1S-father elder DEIX
'My grandfather's name was Xun Heb'in'

b. *Ichachi' ay yemnak kot hinmam icham chi' t'a jun B'ojoch tik.*
Icha-chi' aj y-em-nak kot [**hin**-mam icham chi'] t'a
so-DEIX be A3-descend-PFV.STAT DIR.arrive A1S-father elder DEIX PREP
jun B'ojoch tik.
INDF B'ojoch DEIX
'This is how my grandfather had arrived here in *B'ojoch*.'

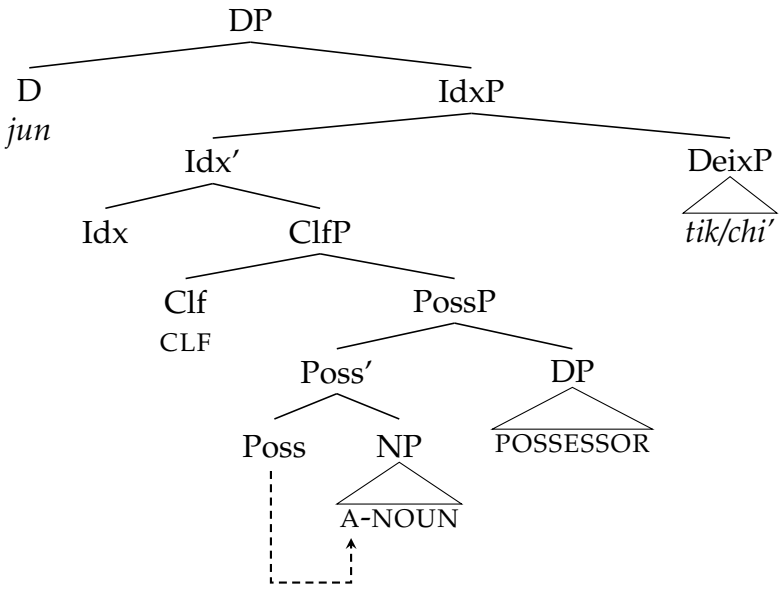
Since possessive marking is compatible with both definite and indefinite nominal expressions, I will not attempt to provide a semantic composition of definiteness and indefiniteness with and without possessive marking. Chapters 3–5 of this thesis will therefore

only focus on definite and indefinite expressions that are unpossessed, with possession only becoming relevant in Chapter 6.

2.5.5 Assumptions about nominal syntax and word order

Before moving on, I establish some assumptions about the syntax of the extended nominal domain in Chuj. While the syntax of the clausal domain has received considerable attention in Mayanist literature, the syntax of the nominal domain remains relatively understudied. Since the specific syntactic details of the extended nominal domain are not relevant for the current purposes, I will assume a simplified syntax throughout this thesis with right-branching specifiers for items that are found at the right of the noun. A schema is provided in (73) for the items of relevance in this thesis. I also assume, following Coon (2017b) on Ch'ol, that the source for Set A possessor agreement on the noun is a functional projection above the NP, which hosts the possessor. This parallels assignment of Set A morphology in the verbal domain, which is also argued to arise as a consequence of an Agree relation between the head of a functional projection, v/Voice, and the external argument located in the specifier of that projection (Coon 2017b).

(73) Syntax of DPs in Chuj



Three notes about the assumed syntactic structure in (73) are in order. First, the above syntax assumes, following numerous authors, the existence of functional projections above the head nominal (e.g., Abney 1987, Szabolcsi 1987, Ritter 1991, Alexiadou et al. 2007, Giusti 2015, Jenks and Konate to appear). Under this view, noun phrases are not the maximal projection in the extended nominal domain (see Bruening 2009, 2020 and Salzmann 2020 for discussion and potential criticisms of this approach). Whether or not this is right is not crucial for the proposals put forth in this chapter. However, I submit that the rich array of functional items found in extended noun phrases in Chuj (see (21) above) serves to motivate this treatment of the extended nominal domain.

A second assumption has to do with the syntactic position of possessors and deictic particles. It is widely assumed that comparable items across languages sit in specifier positions (Jackendoff 1977, Szabolcsi 1983, 1987, Speas and Fukui 1986, Abney 1987, Alexiadou et al. 2007, Giusti 2015, Coon 2017b, Hanink 2018, Ahn 2019, Hsu and Syed 2020, Jenks and Konate to appear), a proposal I adopt in (73). Specifically, following Szabolcsi (1987)—and Coon (2017b) on Mayan—I assume that possessors merge relatively low in the extended nominal domain in a projection I represent as "PossP".[16] As for demonstratives, I follow Simonenko (2014) and Hanink (2018, 2021) in assuming that they sit in a projection labeled "IdxP" below the head of the extended nominal projection (though see Jenks and Konate to appear for criticisms of this position).

Finally, notice that possessors and deictic particles are found in *right*-side specifiers in (73). It should be stressed, however, that right-branching specifiers are represented only as a matter of convenience. While right-specifier analyses of clausal word order in Mayan exist (Aissen 1992 and Little 2020b), the correct analysis of word order within extended nominal projections in Chuj—and more generally across other Mayan languages—is still open for debate. Moreover, there are ways to maintain a DP syntax that does not rely on the presence of right-branching specifiers, as was the case for the clausal domain schema-

[16] In Chapter 6, I will suggest that possessors can also sometimes be re-merged in the specifier of DP.

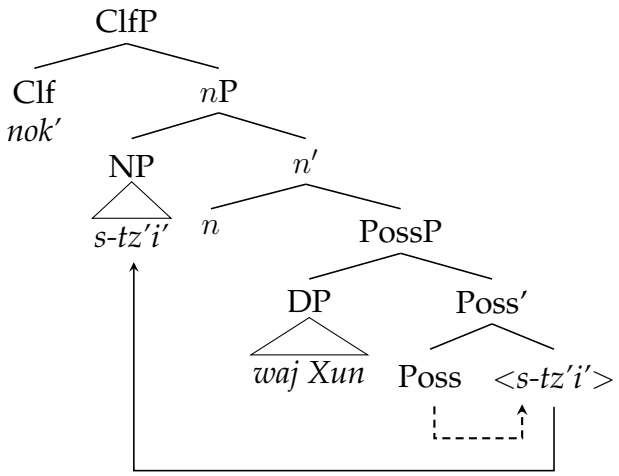
tized in (20) above. For instance, building on Coon (2009, 2017b) on Ch'ol, material found on the right of nominals could be derived via steps of leftward phrasal movement to intermediate functional projections. This is shown for a nominal expression with a possessor in (74) and one with both a possessor and a deictic particle in (75). Such derivations might be taken to be conceptually advantageous, since they do not violate previously proposed constraints on word order within the DP (see e.g. Cinque 2005, but also Abels and Neeleman 2012). I leave an in-depth analysis of the derivation of word order in the nominal domain for future work.

(74) Phrasal movement of NP to "spec, nP"

a. *nok' stz'i' waj Xun*

nok' s-tz'i' waj Xun
CLF A3-dog CLF Xun
'Xun's dog'

b.

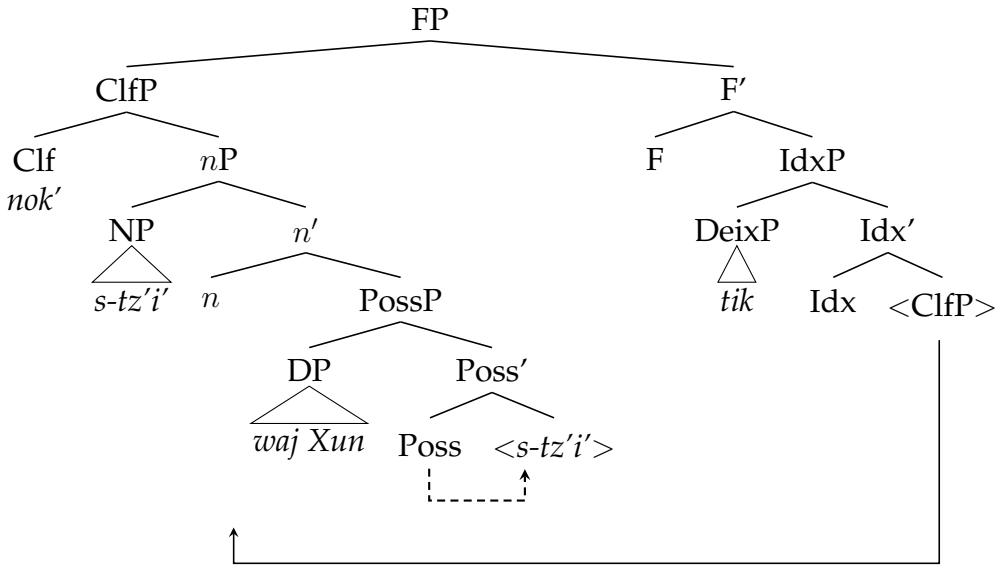


(75) Phrasal movement of ClfP to higher functional projection (labeled "FP")

a. *nok' stz'i' waj Xun tik*

nok' s-tz'i' [waj Xun] tik
CLF A3-dog CLF Xun DEIX
'this dog of Xun's'

b. *Phrasal movement of ClfP to "FP"*



2.6 Summary

This chapter has provided a description of the demographics and grammar of Chuj, with a focus on the extended nominal domain, which we saw exhibits a rich array of functional items. Emphasis was put on the items that will play a central role in the rest of this thesis, namely those that play a role in the composition of definite and indefinite expressions: (i) noun classifiers and third person classifier pronouns, (ii) indefinite determiners, and (iii) adnominal demonstratives. As foreshadowed throughout this chapter, the rest of this thesis will focus almost exclusively on singular expressions. Table 2.7, repeated from above, summarizes the possible combination of these morphemes within nominal expressions. The goals of Chapters 3–5 will be to provide a syntactic and semantic analysis of the pieces in this table.

Chapter 3 will provide the foundations for an analysis of definite expressions in Chuj, and will therefore focus on the configurations on the top half of Table 2.7, those on lines

Table 2.7: Singular DP configurations in Chuj

	Configuration	Example	Rough translation
①	CLF + NP	*nok' tz'i'*	'the dog'
②	CLF + NP + DEIX	*nok' tz'i' chi'*	'the/that dog' (anaphoric)
③	CLF	*nok'*	'it'
④	INDF + NP	*jun tz'i'*	'a dog'
⑤	INDF + CLF + NP	*jun nok' tz'i'*	'a (certain) dog'
⑥	INDF + NP + DEIX	*jun tz'i' chi'*	'that dog'
⑦	INDF + CLF + NP + DEIX	*jun nok' tz'i' chi'*	'that dog' (exophoric)

①–③. Chapters 4 and 5 will then turn to the configurations in the second half of the table, those which play a role in the composition of indefinite expressions. While I attempt to account for (i) the semantics of the items from Table 2.4 and (ii) how they compositionally interact in order to deliver the subtle semantic distinctions to be described, some of the proposals put forth in earlier stages of the thesis will have to be refined in later parts of the thesis, upon exposure to new empirical findings. Proposed entries for lexical items that are not final will be marked accordingly. By providing an analysis of definiteness and indefiniteness in Chuj, these chapters will thus provide answers to the first two questions posed in the introduction, namely: (i) What morphosyntactic elements are involved in the composition of nominal expressions?; and (ii) What semantic and pragmatic distinctions are expressed by different combinations of functional items within the nominal?

The general structure of these chapters is as follows. Chapter 3 will provide a first approach at the semantics of noun classifiers, including an analysis of when they appear alone with a noun, when they combine with deictic particles, and when they appear alone as pronouns (configurations ①–③). In Chapter 4, I then build on and refine the proposals in Chapter 3 to account for the ability for noun classifiers to co-occur with indefinite determiners (configurations ④–⑤). While the overall semantic contribution of noun classifiers will remain intact, the entries from Chapter 3 will be refined so as to account for their ability to co-occur with indefinites. I then show in Chapter 5 that the refined entries for noun classifiers can be extended to account for the necessity of noun classifiers in definite

descriptions. I also show that the theory of definiteness and indefiniteness proposed in this thesis can be straightforwardly extended to explain why deictic particles can combine with both definite and indefinite nominal expressions.

Finally, Chapter 6 will be concerned with the third question posed at the outset of this dissertation, namely: What conditions are imposed on bound and coreferential nominal expressions within sentences? In section 2.5.4 of the current chapter, I provided brief background on possessive constructions. As we will see, possessive constructions will provide us with key insight into the distribution of covalued nominals in Chuj.

Chapter 3

The composition of definiteness in Chuj

Noun classifiers form a typologically rare class of grammatical item, attested in only a limited set of language families, including in languages belonging to the Q'anjob'alan sub-branch of Mayan (Aikhenvald 2000; Grinevald 2000). Though, as mentioned in Chapter 2, these items have received considerable attention in the Mayanist literature (see e.g. Craig 1977, 1986b, 1986a and Ramsay 1985 on Popti'; Hopkins 1967, Maxwell 1981, Buenrostro et al. 1989 and Royer 2017 on Chuj; Zavala 2000 on Akatek; Mateo Toledo 2017 on Q'anjob'al; and Hopkins 2012b on the Q'anjob'alan languages more generally), they have received little study in formal semantics. As noun classifiers play a quintessential role in the composition of nominal expressions in Chuj, one of the general goals of this thesis will thus be to fill this gap, by taking a close look at their syntactic and semantic distribution. More locally, the goal of this chapter is to provide a preliminary analysis of noun classifiers, based on what is probably their most canonical use in the language: their obligatory appearance in *definite descriptions*, as in (76).

(76) *Ixwil nok' tz'i'.*
Ix-w-il [**nok'** tz'i'].
PFV-A1S-see CLF dog
'I saw the dog(s).'

More specifically, this chapter will show that an investigation of noun classifiers in Chuj can inform us on the underlying syntax and semantics of different kinds of definite descriptions (Schwarz 2009, 2013, 2019), and how these different kinds of expressions connect to pronouns, if pronouns are to be understood as concealed definite descriptions (Postal 1966, Evans 1977, Cooper 1979, Déchaine and Wiltschko 2002, Elbourne 2005).

As already noted in Chapter 2, Chuj's noun classifiers exhibit a wide distribution, appearing in a variety of syntactic and semantic environments, playing what appears to be a central role in the composition of the DP. Table 3.1, repeated from Table 2.4, summarizes the syntactic environments in which they appear.

Table 3.1: Possible DP configurations in Chuj

	Configuration	Example	Rough translation
①	CLF + NP	*nok' tz'i'*	'the dog'
②	CLF + NP + DEIX	*nok' tz'i' chi'*	'the/that dog' (anaphoric)
③	CLF	*nok'*	'it'
④	INDF + NP	*jun tz'i'*	'a dog'
⑤	INDF + CLF + NP	*jun nok' tz'i'*	'a (certain) dog'
⑥	INDF + NP + DEIX	*jun tz'i' chi'*	'that dog'
⑦	INDF + CLF + NP + DEIX	*jun nok' tz'i' chi'*	'that dog' (exophoric)

Given their wide distribution, and, as we will see, the subtle semantic distinctions that emerge from combining the different items in Table 3.1, it will be necessary to provide an analysis of noun classifiers in steps. Accordingly, this chapter will exclusively focus on the configurations in ①–③, namely when classifiers appear alone with nouns ①, when they co-occur with a deictic particle ②, and when they appear alone as pronouns ③. Building on observations from previous work (Buenrostro et al. 1989; García Pablo and Domingo Pascual 2007, Buenrostro 2017), I provide a preliminary analysis of noun classifiers as *weak* definite determiners (Schwarz 2009): these items essentially impose uniqueness and existence conditions on the extension of the NP. I will also argue that the distribution of Chuj noun classifiers offers important insight into the growing literature that establishes a distinction between two kinds of definite descriptions, namely

the distinction between *weak* and *strong* definites (Schwarz 2009, Aguilar-Guevara et al. 2019). In particular, Chuj provides evidence that strong definiteness can be compositionally derived from weak definiteness (as in e.g. Hanink 2018, Ahn 2019), rather than being lexically encoded in separate determiners, as proposed in previous work (Schwarz 2009; Arkoh and Matthewson 2013; Jenks 2018; Jenks and Konate to appear).

In a nutshell, I will argue that noun classifiers occur in the configurations in ①–③ of Table 2.4 because these configurations all require a uniqueness presupposition, contributed by the noun classifier. As I will show, the uniqueness-based approach to noun classifiers straightforwardly accounts for their appearance alone with nouns ①, which result in weak definite descriptions. To create *strong* definites, which further contribute a familiarity (or anaphoricity) presupposition (Schwarz 2009), I show that noun classifiers must combine with additional morphology. In particular, the familiarity presupposition, formalized with an index interpreted relative to a contextually-determined assignment function, is triggered by the presence of a deictic particle ②. Finally, if noun classifiers are uniformly weak definite determiners, a question arises as to why they can be used alone as pronouns ③, which in most cases are used anaphorically. I argue that *anaphoric* third person classifier pronouns in Chuj are essentially E-type pronouns (Cooper 1979, Heim 1990): weak definite classifiers that combine with a covert index-introducing predicate. As such, classifier pronouns are just an alternative form of strong definites, with the familiarity presupposition being introduced covertly. The proposed semantic outputs for each of the different configurations in ①-③ are summarized in Table 3.2:

Table 3.2: Classifier configurations and semantic output

①	CLF + NP	weak definite
②	CLF + NP + DEIX	strong definite
③	CLF + $[\lambda x.\ x = g(i)]$	strong definite (= anaphoric pronoun)

This chapter only focuses on the configurations in rows ①–③ of Table 3.1. Therefore, the analysis of noun classifiers and of other items presented in this chapter is only pre-

liminary, and will have to be refined given new configurations considered in the next chapters, namely those in rows ④–⑦ of Table 3.1. Nonetheless, the core semantic proposal of noun classifiers will remain the same throughout the thesis: noun classifiers contribute the signature property of weak definite descriptions, a non-assertive uniqueness condition.

The rest of this chapter is structured as follows. In section 3.1, I provide brief background on previous accounts of noun classifiers in Q'anjob'alan languages, as well as a discussion of the assumptions related to the theoretical framework (situation semantics) I adopt in the rest of this thesis. In section 3.2, I provide information about the distinction between *weak* and *strong* definite determiners established in Schwarz 2009 and argue that Chuj noun classifiers have the semantics of weak definite articles. In section 3.3, I then argue that strong definites are built compositionally in Chuj, and provide a formal analysis of this composition. In section 3.4, I account for pronominal uses of noun classifiers. Section 3.5 concludes.

3.1 Previous work and theoretical assumptions

Here, I briefly discuss in section 3.1.1 previous pragmatic/semantic notions with which noun classifiers have been described, arguing that these notions are either too vague or make wrong predictions. Section 3.1.2 then presents the entry for noun classifiers motivated in the rest of this chapter—one in which the noun classifiers denote weak definiteness—and lays out some important theoretical assumptions related to the framework adopted in this work.

3.1.1 Previous work on noun classifiers

The wide distribution of noun classifiers in Chuj—as attested in Table 3.1 above—naturally led previous researchers to offer more general, and therefore less principled, accounts of

their distribution. In particular, apart from definiteness (Buenrostro et al. 1989, García Pablo and Domingo Pascual 2007), which is the notion I will rely on in this work, there are two pragmatic/semantic notions with which the semantic conditions on the use of noun classifiers have been described.

One notion that has been used is *referentiality*. Indeed, both Craig (1986b) and Zavala (2000), working respectively on Popti' and Akatek (both closely-related to Chuj), have argued that noun classifiers are markers of *referential* nominal expressions. While the association of noun classifiers to referentiality seems largely correct on a descriptive level, and concurs with many of the findings presented in this dissertation (which show that noun classifiers tend to largely appear with DPs that refer to specific entities in the world), the term "referentiality" is not sufficiently defined such as to make clear predictions about the exact distribution of noun classifiers. Moreover, as will be discussed in this chapter, it is not entirely accurate to state that classifiers mark referentiality, since, as discussed below, they are felicitous in covarying uses of definites, which do not refer to particular entities in the world.

A second notion with which noun classifiers have been identified is "pragmatic importance in discourse", which again, both Craig (1986b) and Zavala (2000), building on Ramsay (1985), resort to. Following Givón (1981, 1985), Craig (1986b: 271) defines pragmatic importance in terms of nominal referents which are most likely to be picked up again in the following discourse. Again, this cannot be the sole contribution of noun classifiers. Consider, for instance, the following narrative sequence:

(77) a. *Ixinxit' ek' t'a te' spat waj Xun.*
Ix-in-xit' ek' t'a te' s-pat waj Xun.
PFV-B1S-go DIR.pass PREP CLF A3-house CLF Xun
'I went to Xun's house.'

b. *Haxo, ixinjakan te' pwerta.*
Haxo, ix-in-jakan [#**(te')** pwerta].
then PFV-A1S-open CLF door .
'Then, I opened the door.'

c. *¡Ha waj Xun, tzuyan ek' winh t'a ssat piso!*
Ha waj Xun, tzuy-an ek' winh t'a s-sat piso!
TOP CLF Xun, lie-STAT DIR.pass CLF PREP A3-face floor
'Xun was lying (unconscious) on the floor!'

An account that treats noun classifiers as markers of important participants in discourse predicts that their presence should sometimes be, if not always be, optional. In the narrative sequence in (77), the speaker is telling the addressee that Xun, a man that they know, was lying unconscious on the floor. The noun *pwerta* 'door' is not an important participant in the discourse, and would likely not be mentioned again in following discourse, yet the presence of a classifier is enforced.

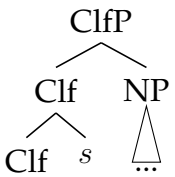
We therefore need a more principled account of noun classifiers. In the rest of this thesis, I will argue that—despite their appearance within indefinite expressions—the right semantic notion to describe the distribution of noun classifiers is *weak definiteness*.

3.1.2 Definiteness and situation semantics

In this chapter, I propose an analysis of noun classifiers in Chuj, which can account for the observations put forth in previous work on noun classifiers (including Craig (1986b) and Zavala's (2000)), all while explaining the presence of noun classifiers in contexts not predicted in previous accounts, including non-referential uses. The main proposal will be that noun classifiers instantiate weak definite determiners, in the sense of Schwarz 2009, 2013. For Schwarz, weak definites are "Fregean" (Frege 1892), in that they encode a presupposition that there is a unique satisfier of the NP taken as an argument by the determiner (see also Strawson 1950, Heim 1991, Elbourne 2005, 2013). Following Percus (2000), Schwarz (2009, 2012), and Elbourne (2013), I further assume that nominal expressions involve a syntactically-represented but unpronounced *situation* pronoun (Barwise and Perry 1983, Kratzer 1989, 2019), which I take to instantiate particulars of type s (Kratzer 1989), and which serves to restrict the domain with which the nominal expression is interpreted (78). As argued at length in the aforementioned literature, the

possibility for the situation pronoun to be bound by a quantifier (over situations) is instrumental in deriving covarying interpretations of definite descriptions. For now, I locate the situation argument as a sister to the head of the phrase containing the classifier, which I will label "Clf" in this dissertation. Note, though, that in Chapter 4, the location of the situation pronoun and resulting order of composition will be revisited, so as to account for new data. In particular, s will be sister to a head above the ClfP.

(78) Syntactically-represented situation pronouns



The preliminary denotation for noun classifiers is provided in (79).[1] As suggested for the weak definite determiner in Schwarz 2009, the noun classifier takes two arguments, a situation pronoun and an NP property, and returns the unique satisfier of that NP in the situation (note again that the order of composition with respect to the situation argument will be modified in Chapter 4). If there is no unique satisfier of the NP in the situation, the uniqueness presupposition in (79) is not met and the output is undefined.[2]

(79) Denotation of noun classifiers (to be modified in chapter 4)
$[\![\text{CLF}]\!] = \lambda s.\lambda P_{\langle e,\langle s,t\rangle\rangle}\colon |\{y\colon P(y)(s)\}| = 1.\ \iota x[P(x)(s)]$

Before moving on, two notes about the entry of the noun classifier in (79) are in order.

First, as is the case in most accounts of domain restriction via contextually-supplied variables (Cooper 1975; Barwise and Perry 1983; Westerståhl 1984; von Fintel 1994; Percus 2000; Büring 2004; Wolter 2006; Keshet 2008; Schwarz 2009), it is a question how exactly

[1] Like previous work (e.g. Sharvy 1980, Link 1983), I assume that the uniqueness presupposition is just a sub-case of a more general maximality presupposition on definite descriptions, and I leave aside the discussion of plural definite descriptions.

[2] I use set notation in (79) to maintain consistency with the updated classifier denotation to be introduced in chapter 3. That said, (79) is semantically equivalent to the more common weak definite notation in (i); see also (86-a) below.

(i) $[\![\text{CLF}]\!] = \lambda s.\lambda P_{\langle e,\langle s,t\rangle\rangle}\colon \exists!x[P(x)(s)].\ \iota x[P(x)(s)]$

the contextual variable s in (79) gets its value, and what kinds of values it can receive. I will assume, following Schwarz (2009, 2012), that situation variables can be either free, bound by a quantifier over situations, or bound by a syntactically-represented topic situation (see Schwarz 2012 and Kratzer 2019 for a detailed discussion of these different options). For free situation variables, I follow Schwarz (2009) in assuming that these get their value from a contextually supplied assignment function. Crucially, a free situation variable cannot simply take as its value any situation without constraint, otherwise a (sub)situation could always be found, such that it includes just one satisfier of the NP, and the presupposition of the weak definite could always be met or accommodated. This issue will be at the forefront of the discussion in Chapter 4. For now, though, the central assumption I am making is that the situation pronouns with which definite descriptions are evaluated must pick out a large enough situation so as to include all salient satisfiers of the NP. As for bound situation pronouns, I will simply be assuming the system in Schwarz (2009: §3.2.2), and will only briefly discuss quantificationally bound situation variables when it becomes relevant (§3.3.2). See Schwarz 2009, Chapter 4, for discussion of topic situations, how the situation pronoun of the determiner could be bound by a topic situation, and how topic situations could be derived from questions.

Second, note that the denotation in (79) ignores the fact that noun classifiers vary depending on the noun they introduce. I assume that this is no different than the fact that the French definite articles *le* and *la* vary according to gender. Chuj is just an extreme case, as it has sixteen versions of the same definite article. Though I leave open the issue of how the choice of the classifier is determined, one possibility is that the features associated with different classifiers are introduced in the syntax as presuppositional modifiers that denote *partial* identity functions. This is similar to the presuppositional analyses of ϕ-features in e.g. Cooper 1983, Heim 1990, and Heim 2008. Another option would be to encode a 'class' presupposition in the entry of each classifier. For instance, the classifier for animals *nok'* could encode a presupposition to the effect that the nominal expression

denotes an entity or set of entities that are a subset of a specified class of "animal entities" or "entities derived from animals":

(80) Denotation of animal noun classifier (to be modified in chapter 4)
$[\![nok']\!] = \lambda s.\lambda P_{\langle e,\langle s,t\rangle\rangle}$: $|\{y\colon P(y)(s)\}| = 1 \wedge \{y\colon P(y)(s)\} \subseteq \{z\colon \textsc{animal}(z)\}.\ \iota x[P(x)(s)]$

In the rest of this thesis, I will generally leave out the class information associated with the different noun classifiers, focusing instead on the syntactic and semantic contributions that seem to be shared by all of these items.

The rest of this chapter is divided as follows. After providing background on the distinction between weak and strong definites, section 3.2 provides evidence that noun classifiers are required in weak definite descriptions. I then argue in section 3.3 that strong definites are derived compositionally in Chuj, by combining weak definite classifier phrases with additional morphology. Section 3.4 then accounts for pronominal uses of noun classifiers.

3.2 Chuj classifiers and weak definite descriptions

Here, I first provide general background on Schwarz's (2009) proposed contrast between weak and strong definites in section 3.2.1, and then provide evidence that Chuj classifier–noun constructions pattern like weak definites in section 3.2.2.

3.2.1 Background: Two kinds of definites across languages

Though there are many approaches to the semantics of definiteness, two families of accounts stand out. On the one hand, some accounts posit that definite determiners introduce a uniqueness (or maximality) presupposition (e.g. Frege 1892; Russell 1905; Strawson 1950; Hawkins 1978, Heim 1991, Heim and Kratzer 1998, Elbourne 2005, 2013, Coppock and Beaver 2015). On the other hand, some accounts posit that definite determiners encode a presupposition that the speaker and addressee are *familiar* with the referent of

the nominal expression (Christophersen 1939; Kamp 1981; Heim 1982, Chierchia 1995). There are also hybrid accounts, which incorporate aspects of both views (e.g. Farkas 2002; Roberts 2003).[3]

More recently, based on observations dating back to Ebert 1971 that some languages overtly distinguish between different kinds of definite articles, Schwarz (2009) proposes that there are two kinds of definite determiners crosslinguistically: weak definites, which encode only uniqueness, and strong definites, which encode familiarity.[4] The overt contrast between weak and strong definites is observed in German in the ability for different article forms to contract with prepositions. Weak definite forms of articles occur in environments where the referent of the nominal expression is unique in the context, but where it has been neither previously mentioned in discourse nor exophorically/deictically identified. Example (81) illustrates this, with the key feature to notice being that the weak article phonologically contracts with the preposition *von*:

(81) *Weak definite article in German* (Schwarz, 2009, 42)
Der Empfang wurde **vom** / #**von dem** Bürgermeister eröffnet.
the reception was by.the$_{weak}$ / by the$_{strong}$ mayor open
'The reception was opened by the mayor.'

Strong definites, on the other hand, are required when the referent of the nominal expression is present in prior discourse as well as when the referent is exophorically identified. In that case, contraction with the preposition is not possible, as illustrated in (82):

[3]I have oversimplified the range of theories on definite descriptions. For example, while most of the uniqueness-based accounts of definite descriptions assume that they also introduce an existence presupposition, Coppock and Beaver (2015) recently argue that the English definite article only presupposes uniqueness, and not existence. Moreover, not all uniqueness-based theories of definite descriptions encode uniqueness as a presupposition. Non-presuppositional accounts include Russell 1905, Donnellan 1966 and Neale 1990, who argue that definite determiners *assert* uniqueness. See Elbourne (2013), chapter 1, for an overview.

[4]This reinforces the view more generally that "definiteness" is not necessarily a primitive semantic or pragmatic category, as suggested in much work prior to Schwarz 2009, such as Hawkins 1978, Gundel et al. 1993, Himmelmann 1997 and Lyons 1999; see also Becker 2021.

(82) *Strong definite article in German* (Schwarz, 2009, 23)

Hans hat einen Schriftsteller und **einen Politiker** interviewt. Er hat
Hans has a writer and a politician interviewed. He has
#**vom** / **von dem** **Politiker** keine interessanten Antworten
from.the$_{weak}$ / from the$_{strong}$ politician no interesting answers
bekommen.
gotten

'Hans interviewed a writer and politician. He didn't get any interesting answers from the politician.'

As Schwarz shows, the above two examples form only a subset of environments in which weak and strong definites are observed. In sections 3.2.2 and 3.3, I discuss a broader range of environments in which both kinds of definites arise. As we will see, Chuj consistently marks the distinction characterized by Schwarz.

While the weak/strong definite contrast in German is only perceivable when a determiner appears adjacent to a preposition, we will see that it is perceivable throughout all definite environments in Chuj. This is also the case in other languages that have been reported to exhibit a contrast between weak and strong definites. For instance, Arkoh and Matthewson (2013) argue that while weak definites are realized as bare nouns in Akan (Kwa, Niger-Congo), strong definites require the 'familiar' determiner *nʊ́*. An example illustrating this use of *nʊ́* is provided in (83).

(83) *Narrative segment in Akan* (Arkoh and Matthewson 2013: (13))

a. Mʊ̀-tɔ́-ɔ̀ èkùtú.
1SG.SUBJ-buy-PAST orange
'I bought an orange.'

b. Èkùtú **n**ʊ́ yɛ̀ dɛ̀w pápá.
orange FAM be nice good
'The orange is/was really tasty.'

In (83-b), the referent of *èkùtú* 'orange' has already been introduced in the previous sentence (83-a), and is therefore familiar. The use of *nʊ́* in (83-b) is enforced.

Akan weak definites, on the other hand, do not tolerate the presence of *nó*. According to Arkoh and Matthewson (2013), the sentence in (84) is odd given the context they provide, because the referent of *bànkyí* is not familiar to the hearer.

(84) Akan – Context provided by Arkoh and Matthewson (2013: p. 9):
"Esi visits her friend Ama and in conversation, Ama utters [this sentence]. [...] Esi has no prior knowledge of the said cassava".
?? Ésì fá bànkyí **n**ó áà ó-gú kɛ̀ntsɛ́n mù nó bɔ̀rà.
Esi take cassava FAM REL it-pour basket in FAM come
'Esi, bring the cassava that is in the basket.'

In recent work, Jenks (2018) highlights similar facts in Mandarin: while weak definites are realized as bare nouns in this language, strong definites obligatorily appear with a demonstrative. As we will see, this is even more similar to Chuj, which requires the use of deictic particles with strong definites. For example, consider the following narrative segment, adapted from Jenks 2018:

(85) *Narrative segment in Mandarin*
a. Jiaoshi li zuo-zhe yi ge nansheng he yi ge nüsheng.
classroom inside sit-PROG one CLF boy and one CLF girl
'There are a boy and a girl sitting in the classroom.'
b. Wo zutian yudao #(**na ge**) **nansheng**.
I yesterday meet that CLF boy
'I met the boy yesterday.'

As shown above, definites that have been previously introduced in discourse in Mandarin require the presence of a demonstrative. This is contrary to weak definites, which according to Jenks, must surface as bare nouns (see, for instance, the absence of a demonstrative with the noun *jiaoshi* 'classroom').

Schwarz (2009), Arkoh and Matthewson (2013) and Jenks (2018) all provide an account of the weak/strong definite distinction by assuming that weak and strong definite articles are separate lexical items. In particular, they argue that strong definites have the same core semantics as weak definites, with the minimal addition that strong definites take an extra index-introducing argument. The denotations for weak and strong definite

determiners, modelled in situation semantics (Barwise and Perry 1983, Kratzer 1989), are reproduced below from Schwarz 2009:

(86) a. *Weak definite article* (adapted from Schwarz 2009)
$\lambda s_r.\lambda P$: $\exists!x[P(x)(s_r)].\iota x[P(x)(s_r)]$

b. *Strong definite article*
$\lambda s_r.\lambda P.\lambda y$: $\exists!x[P(x)(s_r) \wedge \boldsymbol{x = y}]$. $\iota x[P(x)(s_r) \wedge \boldsymbol{x = y}]$

In the above denotations, both the weak definite (86-a) and the strong definite (86-b) trigger uniqueness presuppositions within a particular situation. The crucial difference lies in the fact that the strong definite takes an extra index argument (λy), which has the effect of introducing a familiarity (or anaphoricity) condition.[5] Assuming that the index argument is saturated by a covert variable, whose value will be determined by the assignment function, the denotation of strong definites will only be defined if the satisfier of the NP is identical to the value for an index in the assignment function, and thus anaphorically or exophorically identifiable to the speaker and hearer.[6]

Importantly, Arkoh and Matthewson and Jenks share the assumption in Schwarz 2009 that the distinction between weak and strong definites is realized by separate lexical items. While weak definites are derived via a covert determiner in Akan and Mandarin, strong definites independently encode both a uniqueness and familiarity presupposition.

[5] In Hanink 2018 and Jenks 2018, the index argument is introduced in the denotation of the strong definite article as a predicate (type $\langle e,t\rangle$). Evidence for this comes from the fact that the index argument can sometimes be realized by overt predicates the languages i(see Jenks 2018, section 4.4.). Though this is not crucial for the current discussion, note that in the refined analysis of deictic particles in Chapter 5, I will also adopt a predicative denotation for deictic particles. The facts discussed in this subsection, however, will remain intact in the refined analysis.

[6] This is in assuming that entities in the range of the assignment function must be identifiable to both the speaker *and* hearer. This indeed seems to follow from Heim and Kratzer's (1998) Appropriateness Condition in (i) below (see also Beaver and von Fintel 2013). The intuition is the following: the only way for the hearer to be able to determine the truth conditions of an LF, and therefore the meaning of a sentence, is for the hearer to be able to identify the value for every free variable in a LF. While I will propose that this assumption is correct for definite descriptions, I will question its validity in the case of indefinites and other cases of quantifier domain restriction in Chapter 4.

(i) The Appropriateness Condition (Heim and Kratzer 1998: 243)
A context C is *appropriate* for an LF ϕ only if C determines a variable assignment g_C whose domain includes every index which has a free occurrence in ϕ.

In addition to Arkoh and Matthewson 2013 and Jenks 2018 on Akan and Mandarin, Ebert (1971) and Schwarz's (2009) observations on the crosslinguistic nature of definiteness have led to a large body of work, with a great deal of support that languages across various families distinguish weak from strong definites (see e.g., Simonenko 2014 on Austro-Bavarian German; Jenks 2015b on Thai; Cho 2016 on Korean; Ingason 2016 on Icelandic; Simpson 2017 on the Jinyun variety of Chinese; Hanink 2018 on Washo; Cisnero 2019 on Cuevas Mixtec; Irani 2019 on American Sign Language; Schwarz 2019 on various languages; Šereikaitė 2019 on Lithuanian; Little 2020b on Ch'ol; and Jenks and Konate to appear on Marka-Daring). In the next sections, I contribute to this view of definiteness with additional empirical support from Chuj, showing that this language also overtly marks this distinction. However, I show that Chuj strong definites are transparently decomposed, with a weak definite, namely the classifier + noun phrase, as their core. As such, Chuj shows overt evidence that strong definites can be derived compositionally, contrasting with the theories developed in Schwarz 2009, Arkoh and Matthewson 2013 and Jenks 2018, where weak and strong definites are hardwired as separate lexical items. In providing a decompositional account, my proposal aligns with a recent proposal by Hanink (2018, 2021), who also provides a decompositional account of this distinction in German and Washo. I will, however, argue for a different compositional route to strong definiteness. This will have important consequences for the resulting interpretation, namely whether the uniqueness presupposition of strong definites is evaluated relative to the intersection of the NP predicate with the index argument, as in Hanink 2018 and other work, or only with respect to the NP predicate itself. In arguing for the latter option, the current proposal ultimately opens up the possibility for variation in the interpretive properties of strong definites across languages.

3.2.2 Weak definites require classifiers in Chuj

Based on crosslinguistic evidence, Schwarz (2009, 2013, 2019) argues that the different uses of definite determiners in (87) all involve "weak definites". As we will see below, all of these subtypes of definites in Chuj get realized by combining a classifier with a noun, suggesting that noun classifiers pattern like weak definite articles.[7]

(87) **Subtypes of weak definites**
1. "Immediate" situation uses of definites
2. "Larger/global" situation uses of definites
3. Kind-denoting definites
4. Situation-dependent covarying uses of definites

The terms "immediate" and "larger situation uses" are due to Hawkins (1978), who argues for a uniqueness-based approach to definite determiners. Briefly, immediate-situation uses occur when a speaker makes reference to a unique entity present in the immediate context (e.g. *the table* if the speaker is in a kitchen). Larger situation uses, on the other hand, occur when a speaker makes reference to a unique entity in a larger context (e.g. *the president* if the speaker is in Guatemala and is referring to the current president of Guatemala).

In Chuj, both immediate and larger situation uses of definite articles require the presence of a noun classifier, as expected if classifiers encode weak definiteness. Examples of immediate and larger situation uses are provided in (88) and (89):

(88) **Context**: There's one book. The speaker asks you to move it.
Ak' em ch'anh libro t'achi'.
Ak' em [#**(ch'anh)** libro] t'achi'.
put DIR.down CLF book there
'Put the book over there.' (Immediate situation use)

[7] Schwarz (2009) also includes "part-whole bridging definites" (see Clark 1975, Hawkins 1978) under the category of weak definites: examples like: *The computer is broken, because **the keyboard** has a problem*. This subtype of definite in Chuj takes obligatory possessive marking (*the keyboard* must be formally possessed by *the computer*). As discussed in Chapter 2, noun classifiers are never obligatory with possessed nominals.

(89) Larger situation uses

a. **Context**: At a presidential ceremony in Guatemala.

Ixk'och ix Presidente.

Ix-k'och [#**(ix)** Presidente].
PFV-arrive CLF Presidente
'The president arrived.'

b. *Tz'elta wakch'ub' winh alkal.*

Tz'-elta wak-ch'ub' [#(**winh**) alkal].
IPFV-go.out six-clay.jar CLF mayor
'And six saltwater jars were for the mayor.' (txt, CX200715)

Importantly, if there is no unique satisfier of the NP predicates in (88) and (89), a classifier–noun construction cannot be used. Consider, for instance, (90):

(90) **Context**: There are two books. The speaker asks you to move <u>one</u> of the two.

Ak' em [**ch'anh** libro] t'achi'.
put DIR.down CLF book there

The third use identified in (87) is the use of definite articles to refer to kinds, a relatively common pattern across languages (see e.g. Chierchia 1998). As illustrated in (91), Chuj classifiers are required in such cases:

(91) **Context**: Talking about which animals, in general, are dangerous.

Te'ay smay nok' ajawchan.

Te'-ay s-may #**(nok')** ajawchan.
INTS-EXT A3-danger CLF rattle.snake
'The rattlesnake is / Rattlesnakes are very dangerous.'

In the above example, *nok' ajawchan* 'the rattlesnake' does not refer to a particular rattlesnake, but to rattlesnakes in general. Again, the necessity for the classifier to combine with kind-denoting predicates is expected if classifiers are weak definite determiners.

Another relevant example is provided in (92), this time extracted from a narrative. In this sentence, the speaker is enumerating the clothes traditionally worn by men in San Mateo Ixtatán. As seen in (92-b), all of the nouns co-occur with a noun classifier.

(92) Kind-denoting expressions in narrative (txt, CM201115)

a. *Hatik kok'apak, ha'onh mero ajchonhab' honh. Ichok honh winak:*
Ha-tik ko-k'apak, ha'-onh mero aj-chonhab' honh. Ichok honh winak:
FOC-DEIX A1P-clothes, FOC-B1P very AG-village B1P, like A1P man
'This is our clothes, for the true villagers. For us men:

b. *k'apak wex, nok' lopil, k'ak kamix, k'apak chak payu'.*
k'apak wex, **nok'** lopil, **k'ak** kamix, **k'apak** chak payu'...
CLF pant CLF capixay CLF shirt CLF red handkerchief
'the pants, the capixay, the shirt, the red hankerchief...'

Related to the kind-denoting cases observed above is the use of definite determiners in English examples like (93-a), discussed in more detail in Carlson and Sussman 2005, Carlson et al. 2006, and Aguilar-Guevara 2014. These are notable for allowing sloppy interpretations under ellipsis. Contrast (93-a) with (93-b):

(93) Sloppy interpretation of weak definites in English (Carlson and Sussman 2005)

a. Fred went to *the store*, and Alice did, too. (OK as different stores)
b. Fred went to *the desk*, and Alice did, too. (must be the same desk)

Aguilar-Guevara (2014) argues that the sloppy interpretation in (93-a) is made possible given the ability for definite descriptions to refer to kinds. Similarly, Schwarz (2009, 2014) argues that sloppy uses of definite determiners should be subsumed under the category of weak definite articles.

With a subset of nominals, including *bus* 'bus', Chuj shows similar sloppy interpretations of bare classifier + noun phrases under ellipsis.

(94) **Context**: The speaker grabbed a different bus than Xun.
Ixinyam k'en bus, yet' pax waj Xun.
Ix-in-yam [#**(k'en)** bus], y-et' pax waj Xun.
PFV-A1S-grab CLF bus , A3-with also CLF Xun
'I grabbed the bus, and so did Xun.'

The context in (94) forces an interpretation in which the speaker and Xun grabbed a different bus, leading to the sloppy interpretation of *k'en bus* 'the bus'. The noun classifier *k'en*

is obligatory with the nominal *bus* in this case. These kinds of data are entirely expected if Chuj bare classifier + noun phrases form weak definite descriptions.

Finally, Schwarz (2009) argues that *weak* definites can sometimes have "covarying" uses, crucially when *not* preceded by an antecedent. This use of the weak definite can also be observed in Chuj, as seen in (95).

(95) *Juntakel ixek' waj Xun t'a jun chonhab', ixlolon winh yet' winh alkal.*

Juntakel ix-ek' waj Xun t'a jun chonhab' ix-lolon winh yet' [
every.time PFV-pass CLF Xun PREP INDF town, PFV-talk CLF with
#**(winh)** alkal].
CLF mayor
'Whenever Xun visited a town, he spoke with the mayor.'

Under the most salient interpretation of (95), the (weak) definite description *winh alkal* 'the mayor' covaries with respect to each town *Xun* visited. That is, Xun spoke with the unique mayor of each town. As argued in detail in Schwarz 2009 (sections 3.2.2.3 and 4.3) and Jenks 2018, a situation semantics entry for the weak definite articles, like the one in (79), can capture such examples. Essentially, the situation pronoun of the definite article can be bound by a quantifier over situations, such that the uniqueness presupposition is relativized to the situation variable that the universal quantifies over. This yields an interpretation paraphrasable as "in every situation *s*, Matin met the unique mayor in *s*", with the uniqueness presupposition projecting universally.

Importantly, if the uniqueness presupposition is not met in each situation, then the use of a nominal expression with a classifier is considered infelicitous:

(96) **Context**: Many towns that Xun visited had several marimba players.
Juntakel ixek' waj Xun t'a jun chonhab', ixlolon winh yet' winh sonum.

Juntakel ix-ek' waj Xun t'a jun chonhab', ix-lolon winh yet'
every.time PFV-pass CLF Xun PREP INDF town, PFV-talk CLF with
[**winh** sonum].
CLF marimba.player
Means: 'Whenever Xun visited a town, he spoke with *the* marimba player.'

It is not possible to convey (96) with a classifier phrase in the provided context. Since some towns have more than one marimba player, it is not the case that there are unique marimba players for every town, and the classifier is therefore not allowed.

In sum, the data seen in this section suggest that noun classifiers are required in constructions that are prototypical of weak definite descriptions across languages. Classifier + noun phrases trigger the presupposition, given a certain contextually-defined situation, that there is exactly one satisfier of the NP in that situation. In the next section, we will see that strong definites, despite also requiring a noun classifier, are differentiated by their requirement for additional morphology. I will argue that this is because strong definite descriptions are compositionally derived from a weak definite base in Chuj.

3.3 Decomposing strong definites in Chuj

In this section, I first show that strong definite descriptions in Chuj require the addition of a deictic particle (§3.3.1). I then provide a decompositional account of strong definite descriptions in Chuj (§3.3.2).

3.3.1 Strong definites in Chuj

In Chuj, strong definites are overtly distinguished from weak definites in requiring morphology in addition to the classifier. In particular, they must appear with one of the two deictic particles discussed in section 2.5.3, either *tik* or *chi'*, both of which occupy a post-nominal position:[8]

[8]When referring to *tik* and *chi'*, I will avoid the term "demonstrative", and use the term "deictic particle" instead. There are two reasons for this terminological choice. First, these particles are not semantically free, insofar as they require co-occurrence with a functional item (including a noun classifier or indefinite determiner; see Chapters 2 and 5) in order to derive the equivalent of a demonstrative construction in English. The second reason is that the same particles combine with a number of functional expressions in the language in order to convey different meanings, such as prepositions and temporal adverbs.

(97) *nok' tz'i' tik*
nok' tz'i' **tik**
CLF dog DEIX.PRX
'this dog'

(98) *nok' tz'i' chi'*
nok' tz'i' **chi'**
CLF dog DEIX
'this dog'

Demonstrative constructions were described in some detail in section 2.5.3, and will be discussed in greater length in Chapter 5. However, for the time being, recall that *tik* is the proximal particle used when the intended referent is close to the speaker, whereas *chi'* is the distal particle, used when the intended referent is further away from the speaker or not visible to them (see Maxwell 1981 and Buenrostro and Royer 2022). Both particles may also be used for exophoric reference (generally accompanied with a gesture). Since anaphoric uses of definite articles are almost exclusively realized with the distal particle *chi'* (with some exceptions; see Buenrostro and Royer 2022), the strong definites discussed in this chapter will only bear the distal particle.

Schwarz (2009) lists the cases in (99) as environments requiring a strong definite.

(99) **Subtypes of strong definites**[9]
1. Anaphoric uses of definites.
2. Covarying anaphoric definites (e.g. donkey sentences).
3. Producer-product bridging uses of definites.

The example in (100) shows an anaphoric use of a Chuj strong definite. As shown in the possible continuation of (100-a) in (100-b), anaphoric uses of full nominal expressions (i.e. not pronouns, see §3.4) must bear both a noun classifier and a deictic particle:

[9]Based on German data, Schwarz (2009) also lists DPs that take restrictive relative clauses as an environment that licenses strong forms of definite articles (though see Wiltschko 2013 and Simonenko 2014 for potential complications). However, not all languages require strong forms with restrictive relative clauses (e.g. Mandarin), and given the semantics of definite articles provided in Schwarz 2009, it is unclear why the strong form should even be required. In Chuj, though DPs that take restrictive relative clauses can appear with demonstratives, this does not seem to ever be obligatory.

(100) a. *Ay jun tz'i' yet' jun miston tachi'.*
Ay [jun **tz'i'**] yet' jun miston t'achi'.
EXT one dog with one cat there.
'There's a dog$_1$ and a cat there.'

b. *Saksak nok' tz'i' chi'.*
Saksak [**nok'** tz'i' #**(chi')**].
white CLF dog DEIX
'The dog$_1$ is white.'

In (100-b), the noun classifier must obligatorily co-occur with the deictic particle *chi'*, since the referent of the nominal has already been introduced in the previous sentence.[10]

Another example extracted from a narrative is provided below. As can be observed, the weak definite expression *winh pale'* 'the priest' is introduced for the first time in the narrative in (101-b). The following mentions of this referent in (101-c) and (101-d) must then appear with a deictic particle (which speakers confirm is obligatory in this context):

(101) Narrative sequence with weak and strong definites (txt, CP220715)

a. *Haxo junxo lugar b'ajt'il syak' b'aptisar sb'a heb' hato t'a Jakaltenango.*
Haxo junxo lugar b'ajt'il s-y-ak' b'aptisar s-b'a heb' hato
and other place where IPFV-A3-give baptism A3-REFL PL.PRON COMP
t'a Jakaltenango.
PREP Jacaltenango
'And another place they'd go get baptized was in Jacaltenango.

b. *Yojto hat'a tzjaw winh pale'...*
Yojto ha-t'a tz-jaw [**winh** pale']
because FOC-PREP IPFV-come CLF priest
'Because the priest there would come.'

c. *... Hat'a ayek' winh pale' chi'...*
Ha-t'a ay-ek' [**winh** pale' **chi'**]
FOC-PREP EXT-DIR CLF priest DEIX
'This priest was over there...'

[10] It is unclear how long the anaphoric form of the definite article is obligatory in discourse, a fact that is also discussed in Ebert 1971 and Schwarz 2009, 2019. For example, the anaphoric form of the definite article seems to be obligatory with nominals that co-refer with a nominal in immediately preceding sentences. However, once a referent becomes "central" to the narrative, the weak form of the article might become appropriate. Since this is an issue that extends to all existing theories on the distinction between weak and strong definites, I leave it for future work.

d. *... Hanhej t'a tzjaw winh pale' chi'...*
Ha-nhej t'a tz-jaw [**winh** pale' **chi'**].
FOC-only PREP IPFV-come CLF priest DEIX
'And only over there would that priest ever come.'

In addition to anaphoric uses, as seen in (100) and (101) above, it is widely agreed that strong forms of definite articles are also required in covarying anaphoric uses of full definite descriptions, such as donkey anaphora (Schwarz 2009, 2013, 2019; Jenks 2018). Contrary to the covarying use of weak definites observed in the previous section (95), covarying anaphoric uses are preceded by an overt antecedent in the sentence in which they are found. This is the case in donkey sentences, where the entity denoted by *the donkey* co-varies based on its owner.

(102) Every man who owns [a donkey]$_1$ loves [the$_{strong}$ donkey]$_1$.

Now consider a similar donkey sentence in Chuj (103). As can be observed, covarying anaphoric uses of definites require the presence of both a classifier and a deictic particle. That is, under a covarying reading in which every person hunted a different bird, the deictic particle cannot be felicitously omitted; in fact, omission of the deictic particle leads to an interpretation in which every person hunted the same bird.

(103) Donkey sentence in Chuj
Masanil heb' anima' ixilan junjun much, ixsmak'cham nok' much chi' heb'.
Masanil heb' anima' ix-il-an junjun much, ix-s-mak'-cham [**nok'** much
all PL person PFV-see-AF INDF.DIST bird, PFV-A3-hit-die CLF bird
#**(chi')**] heb'.
DEIX PL
'Every person that saw a bird, hunted that bird.'

Strong definites are also argued to arise with a subtype of "bridging definite" (Clark 1975), also known as "associative anaphora" (Hawkins 1978) or "inferrables" (Prince 1981). The kind of bridging definite that requires strong definites is the "producer-product bridging definite". An English example is provided below.

(104) John bought a book yesterday. **The author** is French. (Schwarz 2009)

In the above example, *the author* picks out the author of the book that was introduced in the previous sentence. As discussed at length in Schwarz 2009, such definites require the strong article form. Consider now the following bridging definite in Chuj:

(105) Producer-product bridging definite
Ixwatej jun libro. Tewach' ix tz'ib'um chi'.
Ix-w-awt-ej jun libro. Te-wach' [**ix** tz'ib'um #**(chi')**].
PFV-A1S-read-DTV one book INTS-good CLF writer DEIX
'I read a book. The author is really good.'[11]

As demonstrated in (105), producer-product bridging definites in Chuj require the presence of both the classifier and deictic particle, as expected if classifier–noun–deictic sequences form strong definites.

As seen in the final three rows of Table 2.5 from Chapter 2, classifiers are also used with proper names. A relevant example is provided below:

(106) *Ixwil waj Kixtup.*
Ix-w-il **waj** Kixtup.
PFV-A1S-see CLF Kixtup
'I saw Kixtup.'

In such cases, I assume that classifier pronouns still contribute weak definiteness (see Elbourne 2005 and Fara 2015 for similar accounts of proper names, based on Burge 1973, and for evidence against the "direct referential" view of proper names in Kaplan 1989). Under this view, uniqueness is encoded with a covert definite article in English (Elbourne 2005), but with an overt definite article in Chuj. Interestingly, proper names can co-occur with deictic particles in Chuj. Though I have decided to leave aside the question of how proper names can be understood in this thesis, an issue I hope to explore in future work,

[11] According to the three consultants I have been able to ask, there is another way to convey this utterance. One could alternatively prefix the nominal expression *ix tz'ib'um* with Set A (possessive) marking, such that it is formally possessed by *the book*. Without possessive marking, however, the deictic particle is required.

a preliminary look at corpora reveals that discourse anaphoric uses of proper names tend to behave just like other strong definites in co-occurring with deictic particles:

(107) *Haxo waj Kuxin Yakchin chi', pu'an chakb'ok junelxo.*
Haxo [**waj** Kuxin Yakchin **chi'**], pu'-an chak-b'-ok junelxo.
and CLF Kuxin Yakchin DEIX , blow-DEP red-INCOH-IRR again
'And Kuxin Yakchin, he blew again so as to make it red again.' (txt, CJ210715)

In sum, we have seen that Chuj deictic particles play a crucial role, together with noun classifiers, in deriving strong definites in the language. While weak definite environments involve a classifier, strong definite environments require both a classifier and a deictic particle. Before moving on, it is worth further highlighting the clear similarities between Chuj and Mandarin: both derive strong definites with the same morphology used with complex demonstrative expressions. Crucially, however, the Chuj data suggest a departure from previous accounts of strong definites. Recall from above that in Jenks 2018, there are two separate definite articles. One is ι, a null definite determiner with the semantics of the weak definite. The other is the deictic particle, which incorporates the semantics of ι but adds an index argument. The Chuj data seem to indicate that strong definites themselves should in fact be decomposed in two pieces, an observation which I account for in the next subsection.

3.3.2 Building strong definites from weak definites

I propose that strong definites in Chuj are derived compositionally via two ingredients: (i) noun classifiers, which trigger a uniqueness presupposition, and (ii) deictic particles, which introduce an index that essentially imposes a familiarity condition. The account builds on Schwarz 2009, 2013, 2019; Arkoh and Matthewson 2013; Jenks 2018; Hanink 2018, 2021; and Ahn 2019, but departs from these authors in two respects. First, while Schwarz, Arkoh and Matthewson, and Jenks attribute the distinction between weak and strong definites to a lexical ambiguity, I argue with Hanink and Ahn that the distinction

is achieved compositionally. Second, the proposal differs from all previous accounts with regards to the resulting presupposition of strong definites: while uniqueness is evaluated with respect to the intersection of the NP predicate with the index-introducing argument in previous accounts, I propose that it is only evaluated with respect to the NP predicate in Chuj, and provide support for this choice in section 3.3.2.1 (see also chapter 5).

As already discussed in section 3.1.2, I provide the preliminary denotation in (79) for noun classifiers, repeated below for convenience (recall that I am ignoring the 'classificatory' presupposition of the classifier). This entry can account for all instances of weak definites seen in section 3.2.2, where a classifier appears alone with a nominal.[12]

(108) Denotation of noun classifiers (to be modified in Chapter 4)
$[\![\text{CLF}]\!] = \lambda s.\lambda P_{\langle e,\langle s,t\rangle\rangle}$: $|\{y\colon P(y)(s)\}| = 1.\ \iota x[P(x)(s)]$

In words, noun classifiers take as their first argument a situation pronoun, whose value determines the situation or (sub)situation in which the uniqueness presupposition triggered by the classifier will get evaluated, and then combines with a predicate to yield an argument of type e, namely the unique satisfier of the NP in the situation. While the denotation in (108) will have to be modified in the next chapter to accommodate new data, the crucial point, which will remain intact throughout the rest of this dissertation, is that noun classifiers introduce a uniqueness presupposition.

I propose that the role of deictic particles, then, is to contribute the familiarity (or anaphoricity) presupposition of strong definites. An entry for the particle *chi'* is provided in (109). The deictic particle denotes a partial identity function of type $\langle e, e\rangle$. In the presupposition, the deictic particle makes use of an index interpreted relative to a contextually provided assignment function. Under the assumption that the assigned values for indices in the assignment function must generally be available to all discourse participants in order for the truth conditions of a sentence be calculated (Heim and Kratzer

[12]Recall that I assumed that differences among the choice of the classifier are located in further non-assertive conditions encoded within the lexical entry of the classifier (see e.g. (80) above).

1998), the presence of a deictic particle in a nominal expression effectively triggers the familiarity presupposition associated with strong definites.

(109) Denotation of deictic particle *chi'* (to be refined in Chapter 5)
$[\![chi'_i]\!]^{g,c} = \lambda x\colon x = g(i).\ x$

I further propose that the proximal deictic particle, *tik*, to be discussed in more detail in Chapter 5, encodes a stronger presuppositional variant of (109), with the added condition that the intended referent be relatively close to the speaker:

(110) Denotation of deictic particle *tik* (to be refined in Chapter 5)
$[\![tik_i]\!]^{g,c} = \lambda x\colon x = g(i) \wedge \textsc{close.to.speaker.of.}c(x).\ x$

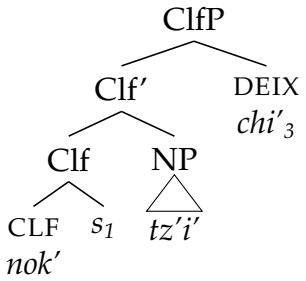
The difference between (109) and (110) can explain the 'default' nature of *chi'* versus *tik*. Since *chi'* is under-specified for speaker proximity, its sole contribution is to introduce an index, and can be used when the intended referent is either far away from the speaker, not present at all in the utterance context, or even in contexts where there is no intended referent (when the index is bound by a quantifier). Again, and as noted for both entries in (109) and (110), the denotation of these items will have to be slightly modified in Chapter 5 to accommodate other syntactic environments in which the deictic particles may be found. The crucial proposal, however, will remain the same: deictic particles introduce a familiarity presupposition as a product of identification of an entity in the range of the assignment function.

To illustrate how strong definites are derived in Chuj, consider the structure and composition for the strong definite expressions in (111-a). As shown in (111-b), I assume that the noun classifier occupies the head of a classifier phrase (as opposed to the head of "DP"), a proposal I defend in Chapter 4. More crucially for our current purposes, I propose that the noun first combines with the classifier, and that the deictic particle is then combined the classifier–noun constituent. I also assume that Chuj deictic particles are located in a right-side specifier in a position structurally above the noun classifier (see

Alexiadou et al. 2007 on demonstratives occupying specifier positions). See Chapter 5 for further discussion.

(111) a. nok' tz'i' chi'
CLF dog DEIX
'this/the dog'

b.



(112) ⟦Clf'⟧(⟦chi'_3⟧^{g})
P: $|\{y\text{: DOG}(y)(s_1)\}| = 1 \wedge \underline{\iota y[\text{DOG}(y)(s_1)] = g(3)}$
A: $\iota x[\text{DOG}(x)(s_1)]$

In this derivation, the classifier *nok'* first introduces a uniqueness presupposition, requiring that there be exactly one salient dog in s_1. If this presupposition is met, the classifier returns this entity. The second step is for the deictic particle to compose with the classifier–noun constituent. Given (109), the deictic particle bears an index, which must be in the domain of the variable assignment, and presupposes that its entity argument be identical to the value of this index (i.e. the 'familiarity' presupposition). I propose that for the relevant "dog" in (111-a) to be in the range of the assignment function, it must have either already been introduced in discourse, or be exophorically/deictically identifiable to both the hearer and speaker. The condition thus captures exophoric as well as anaphoric uses of demonstratives. If the familiarity presupposition (underlined in (112)) is met—namely if the relevant dog is picked out by the index 3 in the variable assignment—then the particle *chi'* composes with the unique salient dog in the situation, returning it back as the referent of the nominal expression. The overall result is a strong definite, realized compositionally by combining the weak definite semantics of the noun classifier in (108) with the semantics of the deictic particle in (109).

As discussed in detail by Schwarz (2009) and Jenks (2018), who both build on insights from dynamic approaches to donkey anaphora (Kamp 1981; Heim 1982; Groenendijk

and Stokhof 1991; Chierchia 1995), the index introduced by strong definite articles can provide us with a desirable semantics for anaphoric covarying readings of strong definites (e.g. donkey anaphora), assuming that the index of the strong definite can be bound by an (unselective) universal quantifier (see e.g. Jenks 2018, section 4.4 for discussion). However, as also acknowledged by both authors, it is not clear why exactly a strong definite is *required* in donkey sentences, since the situation variable of the weak definite could also be bound by a universal quantifier over situations, giving rise to the right semantic output (see Elbourne 2005). Though this is an issue that requires further work, I follow Jenks (2018) in assuming that the use of the strong definite is enforced because of a pragmatic pressure to realize and bind indices whenever possible. In particular, see Jenks's *Index!* principle. Jenks views this principle as "a specific instance of *Maximize Presupposition!*" (see also Ahn 2019, §2.6.3 for relevant discussion).

Relatedly, also notice that the decompositional analysis just provided reveals an entailment relation between Chuj 'weak' and 'strong' definites: uniqueness is still presupposed with strong definites (see (112)), and therefore 'strong definiteness' entails 'weak definiteness'.[13] That is, when the classifier appears with a noun by itself, it triggers a uniqueness presupposition, and when a deictic particle is added, the presupposition of the classifier survives and the deictic particle adds an additional familiarity presupposition. Assuming that the two constructions are 'competitors', then the obligatory presence of deictic particles with strong definites in Chuj could be understood as a classic instance of *Maximize Presupposition!* (Heim 1991, Percus 2006, Sauerland 2008, Schlenker 2012). That is, with or without the deictic particle, the assertive content of the DP in (111-a) is identical for any given context. Assuming as standard that *Maximize Presupposition* compares LFs with identical assertive content (see also §5.1.2.2 below), and favours those which presuppose the most information, it follows that the presence of the deictic particle should always be required in strong definite descriptions.

[13] This observation is due to an anonymous reviewer who provided feedback in Royer (to appeara), the article related to this chapter.

The next three subsections are divided as follows. I first discuss in section 3.3.2.1 a prediction regarding the scope of the quantifier introducing the uniqueness presupposition that follows from the decompositional account of noun classifiers just put forth, and which contrasts with the analysis provided in Schwarz 2009, Arkoh and Matthewson 2013, Jenks 2018 and Hanink 2018, and I show that at least in Chuj, this prediction is borne out. I then discuss in section 3.3.2.2 an apparent exception to the appearance of deictic particles with strong definites, namely when strong definites appear inside a topicalized DP. Finally, section 3.3.2.3 addresses the fact that the proposed denotation for weak definite classifiers also encodes an existence presupposition, which Coppock and Beaver (2015) recently contest in relation to the definite article in English. I argue that the issues discussed by Coppock and Beaver do not straightforwardly extend to Chuj.

3.3.2.1 The scope of uniqueness and deixis

The decompositional account of strong definites just proposed departs from the analysis of strong definites in many previous works in one crucial respect (though see Ahn 2019, §3.2.1, for a decomposition similar to the current account). Recall that for these proposals, the index plays a role in the content of the uniqueness presupposition. Consider, again, Schwarz's entry for the strong definite in (113). A uniqueness presupposition is triggered for the intersection of the NP predicate with the index-introducing argument (where the index comes from a covert variable that saturates the third argument). The relevant segment is underlined for convenience.

(113) $[\![\text{the}_{strong}]\!] = \lambda s_r.\lambda P.\lambda y$: $\underline{\exists!x[P(x)(s_r) \wedge x = y]}$. $\iota x[P(x)(s_r) \wedge x = y]$

Within the presupposition of the strong definite article (underlined), the quantifier enforcing uniqueness ($\exists!$) takes scope over the index (y). This has important consequences for the content of the uniqueness presupposition: it will be satisfied when there is exactly one entity which is both a satisfier of the NP and identical to the index. This means that

one could utter a strong definite description even if there is more than one salient satisfier of the NP predicate in the situation, since at most one entity will ever be identical to the index.

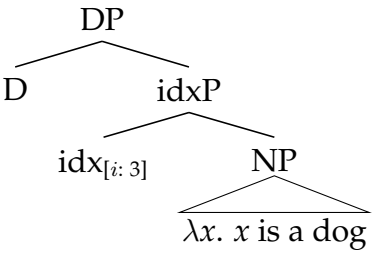
Hanink's (2018, 2021) decompositional account of strong definites in German and Washo makes the same prediction. For Hanink, the index-introducing argument, which she proposes denotes a property, first combines with the NP via Predicate Modification (Heim and Kratzer 1998). The uniqueness presupposition is subsequently evaluated with respect to the result of this combination. Lexical entries and a relevant decomposition for the strong definite DP *the dog* are provided in (114).

(114) Lexical entries and decomposition in Hanink 2018, 2021

a. $[\![\text{idx}]\!]^g = \lambda y.\ y = g(i)$

b. $[\![\text{the}_{weak}]\!] = \lambda \text{P: } \exists!x[\text{P}(x)].\ \iota x[\text{P}(x)]$

c.



d. $[\![\text{DP}]\!]^g = \iota x[x \text{ is a dog} \wedge x = g(3)]$, defined only if $\underline{\exists!x[x \text{ is a dog} \wedge x = g(3)]}$.

As seen in the underlined part of (114-d), the uniqueness presupposition of the weak definite article is again evaluated with respect to the intersection of the NP with the indexical property. Since $g(3)$ will only ever pick out a single entity, the uniqueness presupposition can be met even if there is more than one dog in the context.

The decompositional account I have proposed is slightly different. If the index is introduced outside of the uniqueness trigger, then the familiarity presupposition will be added on top of the presupposition that there is a unique satisfier of the NP in the situation, and so uniqueness in the situation should still hold. The presupposition in (112) is repeated below for convenience:

(115) Presuppositions resulting from composition of *nok' tz'i' chi'* (111-a):
P: $|\{x\text{: DOG}(x)(s_1)\}| = 1 \wedge \underline{\iota y[\text{DOG}(y)(s_1)] = g(3)}$

This presupposition imposes the condition that there be a unique satisfier of the NP in s_1, a dog in this case, and that the unique dog of s_1 be identical to an entity in the range of the assignment function, namely $g(3)$. Therefore, contrary to (113) and (114), the condition that there be one satisfier of the NP in the situation is maintained.

We might expect the result in (115) to have consequences for the felicity conditions of classifier–noun–deixis constructions, especially when used *exophorically*. That is, it might be infelicitous to utter (111-a) if there is more than one dog.[14] Though exophoric uses of demonstrative phrases will be discussed in further detail in Chapter 5, a preliminary look suggests that this prediction is, at least partially, borne out (see also Alonso-Ovalle and Royer 2021).

There are at least two ways demonstrative phrases can be used in exophoric contexts in Chuj:

(116) *¡Yam nok' tz'i' chi'!*
¡Yam [*nok'* tz'i' **chi'**]'!
grab CLF dog DEIX
'Grab that dog!'

(117) *¡Yam jun tz'i' chi'!*
¡Yam [*jun* tz'i' **chi'**]!
grab one dog DEIX
'Grab that dog!'[15]

While the noun and the demonstrative co-occur with a classifier in (116), they co-occur with the numeral *jun* 'one' in (117). While it is acceptable to utter both of these sentences in a context where there is only one dog, the speakers I have consulted indicate a clear preference for (117) if the context contains more than one dog (118-b). That is, an impera-

[14] As Schwarz (2009: §2.2.2.3) notes, strong definite articles tend to also be used *exophorically*, in addition to their *anaphoric* use (see Diessel 1999 on the distinction between exophoric and anaphoric (or endophoric) uses of demonstratives). Jenks (2018) further argues that demonstratives in Mandarin *are* the strong definite article. I therefore assume that exophoric demonstrative phrases in Chuj are a kind of familiar definite article that make use of an index in their denotation (for similar proposals of demonstratives, see e.g. King 2001, Roberts 2002, Wolter 2006, Elbourne 2008, Patel-Grosz and Grosz 2017, Hanink 2018, and Jenks 2018).

[15] Under the current proposal, where deictic particles denote partial identity functions of type $\langle e, e\rangle$, it is not immediately clear how the demonstrative composes with *jun tz'i'* here. In Chapter 5, I provide a modified entry of deictic particles and noun classifiers which can explain why deictic particles can combine with both noun classifier phrases and *jun* phrases.

tive like (116) is judged perfectly acceptable in a setting like (118-a), but less so in a setting like (118-b). The sentence in (117) without a noun classifier, on the other hand, is judged equally felicitous in both settings in (118).[16]

(118) Scenarios for (116) and (117)

a. One dog is in front of you, and it's trying to steal your food. Pointing at that dog, you ask your child to grab it. (116) = ✓| (117) = ✓

b. There are several dogs around you; one of them is trying to steal your food. Pointing at it, you ask your child to grab that one dog. (116) = ?? | (117) = ✓

The fact that (116) is dispreferred by speakers in a context where there is more than one dog supports the decomposition proposed in this chapter: classifiers impose a uniqueness presupposition on top of which the demonstrative adds a familiarity condition. We should therefore expect to see the effects of the uniqueness presupposition in classifier–noun–demonstrative constructions when there is more than one satisfier of the NP, as seems to be the case.

It should be noted, though, that the judgments are not categorical, and that there is considerable speaker variability. In particular, of the three speakers I have been able to consult on this datapoint, one judged (116) as infelicitous in (118-b), whereas two judged it as more or less acceptable. Crucially though, all indicated a clear preference for (117) in this setting. It is important to note that this kind of variability is not entirely unexpected, given a situation semantics approach to definiteness. That is, the uniqueness presupposition is evaluated with respect to the situation picked out by the situation variable—i.e., the situation variable does the domain restriction (Schwarz 2009). If the situation variable is set, for instance, as the entire utterance situation, a failure of the uniqueness presupposition is expected. On the other hand, speakers and hearers may sometimes be willing to admit a more 'minimal' value for the situation variable (e.g., one in which only the dog that is being pointed at is considered, and the other dogs are discarded). In that case,

[16]It also remains to be understood why (117) is not blocked in (118-a), assuming Maximize Presupposition! (Heim 1991). I leave this puzzle aside for future work.

it would be possible for the uniqueness presupposition to hold even if there are several dogs in the larger utterance context. This is not any different from English speakers' ability to use a weak definite description in a situation where uniqueness does not necessarily hold, as in *pass me the book*, if there is only one salient or relevant book in a library with hundreds of books. The essential point for our concerns, however, is that if the classifier takes scope below the deictic particle, as proposed here, then we should sometimes perceive the effects of the uniqueness presupposition, as is observed in the degradedness of (116) in a context with more than one dog.

In sum, though this subsection has presented some amount of evidence that situational uniqueness must hold for strong definites with deictic particles in Chuj, it does not have to be the case that all strong definites across languages are so construed. In fact, in section 3.4, I will claim that contrary to strong definites with overt NPs, anaphoric uses of pronouns involve an indexical argument that applies in the scope of the uniqueness trigger, yielding a result equivalent to the denotations for strong definites provided in previous work. If the current analysis is on the right track, the overall result, then, is that we might expect to find different flavours of strong definiteness from language to language.

3.3.2.2 Strong definites and topichood

There is an exception to the generalization that demonstratives are needed for strong definites: when a Chuj DP is topicalized, the demonstrative is optional. Chuj topics tend to appear at the left periphery (with the topic marker *ha*), and they obligatorily corefer with a resumptive pronoun in the main clause (Bielig 2015, Royer 2021b). This is shown in (119-b), which could naturally follow the utterance in (100-a), repeated in (119-a).

(119) a. *Ay jun tz'i' yet' jun mis t'atik.*
Ay [jun **tz'i'**] yet' jun mis t'atik.
EXT one dog with one cat here.
'There's a dog$_i$ and a cat here.'

b. *Ha nok' tz'i' chi', saksak nok' / Ha nok' tz'i', saksak nok'.*
[***(Ha)** nok' tz'i' **(chi')**], saksak **nok'**.
TOP CLF dog DEIX white CLF.PRON
'The dog$_i$ is white.'

I tentatively propose that topicalized projections involve a topic head, overtly realized as *ha*, that introduces a presupposition requiring the referent of the DP to be discourse-old, and that this circumvents the need for an additional deictic particle. Topicalized constituents are cross-linguistically associated with discourse-old referents (see e.g. Prince 1992, von Fintel 1994, and Aissen 1992 on Mayan). If only constituents whose referent is discourse-old can be topics in Chuj, then it follows that topicalized constituents will always be anaphoric, even without a demonstrative. Interestingly, Mandarin features the same exception with strong definites—demonstratives are optional with topicalized DPs (Jenks 2018, §5.3)—suggesting that this may be a general property of strong definites across languages.[17] I leave a detailed analysis for future work.

3.3.2.3 Do definites presuppose existence?

In recent work, Coppock and Beaver (2015) argue that the English definite determiner does not encode an existence presupposition, a presupposition it has commonly been associated with since at least Frege 1892. They offer a denotation along the lines of (120):

(120) $[\![\textit{the}]\!] = \lambda P\colon |P| \leq 1.\ \lambda x.\ P(x)$

Under this denotation, the definite determiner combines with NP predicates to yield a predicative meaning—*the dog* denotes the predicate of being a dog, defined only if there

[17] In fact, a similar hypothesis is put forth by Jenks (2018), who suggests that topicalized DPs in Mandarin do not need to be indexed, because they are made salient by the Question Under Discussion (Roberts 1996; Büring 2003; Schwarz 2009).

is one or *less than one* dog in the context. In other words, Coppock and Beaver take predicative uses of definite articles, as in (121), to be their most basic use. This is in opposition to the denotation proposed here, as well as in Schwarz 2009 and subsequent work, where the definite determiner is understood to (i) yield an entity (rather than a predicate), and (ii) trigger an existence presupposition.

(121) a. Scott is not [the only author of Waverley].
b. John considers this woman [the queen of the world]. (Coppock and Beaver 2015)

As Coppock and Beaver note, the absence of an existence presupposition is especially supported in examples like (121-a). For them, *only author of Waverley* denotes a predicate which holds of an entity if and only if that entity and no other is an author of Waverley. But under its most salient interpretation, (121-a) conveys that Scott is one of at least two authors of Waverley, in which case there is no satisfier of *only author of Waverley*. Since the sentence is felicitous, they conclude that *the* should not presuppose existence. To account for argumental definites, Coppock and Beaver propose two type-shifts based on Partee 1986 (IOTA and EX), which together type-shift *the*-predicates to type *e* arguments (IOTA) or existential quantifiers (EX). The idea is that when *the*-predicates combine with IOTA, IOTA adds an existence presupposition and the result is a run-of-the-mill definite description presupposing uniqueness and existence ($|P|$=1); this is what Coppock and Beaver describe as the *determinate* interpretation. On the other hand, EX denotes an existential quantifier, and so *the*-predicates combined with EX merely presuppose weak uniqueness without existence ($|P|\leq 1$); this is what they describe as the *indeterminate* interpretation. Indeterminate definites are then taken to account for the most salient readings of the sentences in (121).

It is not clear, however, that this analysis of the definite article can be fully extended to Chuj classifiers. One reason is that bare Chuj classifier–noun constructions are cate-

gorically banned from surfacing as predicates (this has also been noted in Craig 1986b on Popti' and Zavala 1992, 2000 on Akatek). This is shown in (122).

(122) *Alkal waj Xun.*
(***winh**) Alkal waj Xun.
CLF mayor CLF Xun
'Xun is (the) mayor.'

Moreover, the morpheme usually used to convey the meaning of *only* in Chuj, *nhej*, is not compatible with predicative nominals (regardless of the presence of a classifier), as opposed to English (121-a).[18] This means that it is impossible to test utterances like English (121-a) in Chuj, and therefore it is impossible to verify the key evidence presented in Coppock and Beaver 2015 against an analysis of definite determiners as presupposing existence.

(123) **Context**: *The village we are in has more than one mayor.*
*Mok-nhej-laj alkal waj Xun.
NEG-only-NEG mayor CLF Xun
Intended: 'Xun is not the only mayor.'

In sum, the evidence for a predicative analysis of Chuj noun classifiers, without the existence presupposition usually attributed to definite determiner, seems to be lacking. One reason is that bare classifier–noun configurations cannot be used predicatively. Another reason is that the examples that could diagnose the lack of existence presuppositions with definite articles are ineffable in Chuj. This is not to say, however, that a predicative analysis of noun classifiers along the lines of Coppock and Beaver's account of the definite article in English could not in principle be adapted to account for the distribution of noun classifiers in Chuj. In fact, in Chapter 3, a refined predicative entry for noun classifiers, albeit one which keeps with the existence presupposition, will be proposed.

[18] To convey the meaning of (121-a), speakers use a construction along the lines of English *it's not just Scott who is an author of Waverly*.

3.3.2.4 Summary

I have proposed a decompositional account of strong definites in Chuj. While noun classifiers introduce a uniqueness presupposition, deictic particles contribute a familiarity presupposition, namely that the entity output by the weak definite classifier is in the range of the assignment function. In the next section, we turn to an apparent issue for this account: the fact that classifiers can appear alone, and crucially without deictic particles, as anaphoric pronouns. I provide a solution, which essentially proposes a view of pronouns as concealed definite descriptions (Postal 1966, Evans 1977, Cooper 1979, Heim 1990, Elbourne 2005, 2013, among many others). Building on Cooper 1983 and Heim 1990, I assume that classifier uses of anaphoric pronouns are definite determiners that combine with a null predicative variable, which also serves to introduce an index. In that sense, classifier pronouns are conceived of as just another kind of strong definite. However, they also differ from the strong definites with deictic particles discussed in this section insofar as the index argument is introduced below the classifier, yielding a strong definite with the same scopal properties as the ones in work like Schwarz 2009, Arkoh and Matthewson 2013, Jenks 2018, and Hanink 2018, 2021.

3.4 Decomposing pronouns

Most Mayan languages are robustly *pro*-drop (Coon 2016a; Aissen et al. 2017). However, Q'anjob'alan languages are an exception, since noun classifiers serve as third person pronouns (henceforth "classifier pronouns"), and under most circumstances cannot be dropped (see Chapter 6 for the conditions allowing for classifier drop). Consider the following example:

(124) *Ay jun tz'i' t'achi'. Lan sway nok'.*
Ay [jun tz'i'] t'achi'. Lan s-way [*(**nok'**)].
EXT one dog there. PROG A3-sleep CLF
'There's a dog$_i$ there. It$_i$ is sleeping.'

In the first sentence of (124), an indefinite *jun tz'i'* 'a dog' introduces a new referent into the discourse. In the second, the use of the classifier *nok'* alone is sufficient to refer back to the dog that was introduced in the previous sentence.

The piece of data in (124) is somewhat surprising given the proposal from the previous section that strong definites in Chuj can be decomposed. That is, if classifiers only introduce a uniqueness presupposition when combined with an overt noun (and not a familiarity presupposition), then why can they surface alone as pronouns when the pronominal expression is anaphoric to a previously introduced referent? Even more surprising is the fact that the classifier pronoun in (124) cannot in general co-occur with a deictic particle, even when used anaphorically:[19]

(125) *Ay jun tz'i' t'achi'. Lan sway nok'.*
Ay [jun tz'i'] t'achi'. Lan s-way [nok' **(#chi')**].
EXT one dog there PROG A3-sleep CLF DEIX
'There's a dog$_i$ there. It's sleeping.'

This starkly contrasts with anaphoric uses of classifiers with overt nominals, which as shown in examples like (100), *require* the presence of a deictic particle.

Another important observation concerns the use of classifier pronouns in donkey sentences, which again do not allow a deictic particle:

(126) *Masanil heb' anima' ixilan junjun much, ixsmak'cham nok' heb'.*
Masanil heb' anima' ix-il-an junjun much, ix-s-mak'-cham [**nok'** (#chi')]
all PL person PFV-see-AF one bird, PFV-A3-hit-die CLF DEIX
heb'.
PL
'Each person that saw a bird killed it.'

[19]Though classifier pronouns generally appear without a deictic particle, there are special circumstances under which they can optionally appear with one—for instance, when they appear in a topic position:

(i) *Ha winh chi', techanh stelwi winh.*
Ha winh **chi'**, te-chanh s-telwi winh.
TOP CLF.PRON DEIX INTS-tall A3-fall CLF.PRON
'As for him, he's very tall.'

While I leave this issue for future work, it could be that deictic particles become exceptionally appropriate when an anaphoric nominal expression is focused or topicalized.

Again, the absence and unacceptability of a deictic particle in (126) is surprising given the fact that anaphoric uses of definite descriptions in donkey sentences with overt nominals usually require one (see (103) above).[20]

If we want to maintain the semantics of noun classifiers as weak definite determiners, as proposed in (79), two questions must be addressed: (i) why can noun classifiers arise without an overt noun as pronominal expressions that are anaphoric to a previously introduced referent?; and (ii) how is anaphoricity encoded, if not with a deictic particle?[21] In what follows, I address these questions.

3.4.1 Pronouns as definite descriptions

Since Postal 1966, many syntactic and semantic analyses of pronouns, or at least a subtype of what has been referred to as pronouns, posit that they are actually definite determiners with null or elided NPs (e.g. Cooper 1979; Abney 1987; Heim 1990; Ritter 1995; Déchaine and Wiltschko 2002; Elbourne 2005; Arkoh and Matthewson 2013; Clem 2017; Patel-Grosz and Grosz 2017; Bi and Jenks 2019). There are many reasons to support this view. For one, pronominal elements and determiners often look alike (German examples are from Elbourne 2001):

[20] Matthewson (2008) describes a reminiscent—though slightly distinct—pattern in St'át'imcets: while pronouns can be used in donkey sentences, full DPs, which she demonstrates lack a familiarity presupposition (in St'át'imcets), cannot.

[21] As pointed out to me by Lisa Matthewson (p.c.), the second question (how is anaphoricity encoded, if not with a deictic particle?) presupposes that anaphoricity *must* be encoded in Chuj. However, anaphoricity could simply not be encoded at all (see Matthewson 2008 for such an analysis of pronouns in St'át'imcets). If this were the case, it could be that there is simply no anaphoric pronoun in Chuj, and that the ability for classifier pronouns to pick out anaphoric referents is tolerated on the basis of there being no competing anaphoric form. Under this kind of proposal, however, it would remain unclear why classifier pronouns in Chuj cannot co-occur with a deictic particles in sentences like (125) and (126), which are *required* with anaphoric definites with overt nouns. An important question to address if we want to maintain a uniform analysis of noun classifiers is thus: why can classifier pronouns be anaphoric to previously introduced referents, but weak definites (CLF+N) cannot?

(127) French

a. Je vois **la** femme.
I see the woman
'I see the woman.'

b. Je **la** vois.
I her see
'I see her.'

(128) German

a. Hans sieht **den** Mann.
Hans sees the man
'Hans sees the man.'

b. Hans sieht **den**.
Hans sees him
'Hans sees him.'

Furthermore, it has long been observed that pronouns tend to share more with determiners than they do with nouns in their distribution (Postal 1966; Abney 1987). A classic example comes from first and second person pronouns in English, as well as third person pronouns in some dialects of English, which pattern like determiners, and unlike nouns, in accepting an overt noun (Postal, 1966):

(129) we (linguists), you (people), you (lucky guy), them (artists)...

Finally, pronouns and definite determiners often show similar effects, notably in cases of *donkey anaphora* (Heim 1990, Elbourne 2005) (also compare the Chuj examples in (103) and (126) above):

(130) Every person who owns a donkey loves {it/the donkey}.

At least two types of accounts have been proposed to explain the similarity between pronouns and definite descriptions. On the one hand, Elbourne (2013) proposes that the only difference between full DPs and pronouns is NP-deletion. In other words, *the* and pronouns such as *it*, *she*, and *he* exhibit identical semantics (ignoring gender features). The contrast between articles and pronouns lies solely in the phonology: while *the* appears before overt NPs, pronominal forms appear before elided NP complements:

(131) a. [the [NP]]
b. [it [~~NP~~]]

(132) a. [the [dog]]
b. [it [~~dog~~]]

Another strategy has been to assume that pronouns are definite determiners that combine with special unpronounced morphology, and which must critically involve an index

interpreted relative to the assignment function (see e.g. Cooper 1979, Heim 1990, Elbourne 2001, 2005). For such theories, pronouns in English are also considered as morphophonological variants of the definite article:

(133) a. [the [NP]]
b. [it [λx. x = g(i)]]

(134) a. [the [dog]]
b. [it [λx. x = g(i)]]

Interestingly, the Chuj data just seen appear to favour the second of these two accounts (though see §3.4.2 shortly for data supporting the first account). Recall from (125) that classifier pronouns do not generally co-occur with deictic particles, which obligatorily appear with strong definites in Chuj. All else being equal, an NP-deletion account of pronouns (e.g. Elbourne 2013) would therefore predict that anaphoric classifier pronouns always appear with deictic particles. That is, if pronominal uses of classifiers were identical to determiner uses of classifiers, except for deletion of the NP in the phonology, then we would expect that both would require a deictic particle when used anaphorically. However, as already seen in (125), this prediction is not borne out. An analysis with a covert index, on the other hand, offers a straightforward account of the absence of a deictic particle with anaphoric pronouns. Under such accounts, weak definite articles (when functioning as pronouns) combine with a null variable, which introduces an index. This means that adding an index-introducing deictic particle would have no further effect—it would render the demonstrative's contribution trivial. The absence of the demonstrative with classifier pronouns could then be explained given general structural economy constraints on the addition of redundant structure, in line with Cardinaletti and Starke 1999, Schlenker 2005a, Patel-Grosz and Grosz 2017.

I therefore propose that anaphoric uses of classifier pronouns involve a null predicative variable, provided in (135), whose sole contribution is to introduce an index. Since this index can presumably be bound, it is possible to account for the use of classifier pronouns in donkey sentences (see (126)).

(135) $[\![pro_i]\!]^g = \lambda x.\ x = g(i)$

Classifier pronouns are thus E-type pronouns in essence, and denote the unique entity identical to a contextually-determined entity in the range of the assignment function:

(136) $[\![[[\text{CLF } s_1] \ pro_3]]\!]^g$ is defined only if $\exists! x[x = g(3) \text{ in } s_1]$.
When defined = $\iota x[x = g(3) \text{ in } s_1]$

As suggested to me by an anonymous reviewer on a related work, there may be a second empirical reason to favour an E-type approach to classifier pronouns like the one illustrated in (136). If only NP-deletion were at issue, we might expect pronouns to always trigger a uniqueness presupposition, and we would expect sentences like (137) to be infelicitous, since there is clearly no unique elder woman in the context in (137). This prediction is not borne out; (137) is judged felicitous by speakers.

(137) **Context**: Everyone in the village attended a meeting. There are several female elders in the village.
Ay jun b'ek'anh, ay tas sk'anb'ej jun ix chichim t'a skal heb' ix chichimtak chi'. Ixk'e' wa'an ix...
Ay jun b'ek'anh, ay tas s-k'an-b'-ej jun ix chichim t'a
EXT one moment EXT what A3-ask-SUF-DTV one CLF female.elder PREP
s-kal heb' ix chichim-tak chi'. Ix-k'e' wa'an **ix** [...]
A3-among PL CLF female.elder-PL DEIX PFV-rise stand-STAT CLF
'At one point, one of the female elders asked a question. She stood up [...]'

An E-type approach, on the other hand, does not make this prediction. Since there will always be at most one entity that is identical to any given index, the uniqueness presupposition in (136) can be met, even if there are several female elders in the situation. This means that the use of an (E-type) classifier pronoun in Chuj should be possible in sentences similar to (136), as is the case.

In sum, we now have answers to the questions set out at the end of the previous subsection: (i) why can noun classifiers be used alone as anaphoric pronouns?; and (ii) how is familiarity introduced, if not with a deictic particle?

Regarding (i), I showed that it was possible to keep with the weak definite semantics of classifiers in (79) if classifiers combine with a null predicative variable, as independently

proposed for E-type pronouns in Cooper 1979 and Heim 1990. This theory of pronoun formation relies on the widely-held assumption that pronouns are concealed definite descriptions, an assumption that is especially compelling for Chuj, seeing as pronouns and determiner uses of classifiers exhibit no allomorphic variation (unlike determiners and pronouns in, say, English).

Regarding (ii), I argued following a number of works on E-type pronouns that in their use as third person pronouns, classifiers can combine with a null index-introducing variable, thereby bleeding the need for an independent index-introducing deictic particle (possibly due to structural economy constraints). However, I proposed that with classifier pronouns, the familiarity presupposition gets introduced below the uniqueness trigger, revealing a denotation for the strong definite that is slightly different to the one that results from the composition of classifier–noun–deictic particle constructions, where the familiarity presupposition is evaluated on top of the uniqueness presupposition (compare (136) with (115)). This denotation for anaphoric classifier pronouns can therefore be seen as an alternative compositional path to strong definiteness in Chuj, which aligns more closely with the proposed denotations for the strong definite article in previous work.

Finally, the proposal has implications for theories of pronouns that view them as (weak) definite descriptions with elided NPs (e.g. Elbourne 2013). That is, I showed that this view of anaphoric pronouns would make a wrong prediction for sentences like (137) in Chuj, and that anaphoric pronouns were better understood as determiners which combine with covert index-introducing predicates.

3.4.2 Are there weak definite pronouns?

I have just proposed that, as weak definite determiners, noun classifiers can combine with a covert index-introducing predicate to yield an E-type pronoun. This accounts for most pronominal cases of classifiers, since classifier pronouns tend to be used anaphorically. However, given that noun classifiers have the semantics of weak definite articles, it is

interesting to consider whether they could also be used non-anaphorically, or in other words as "weak definite pronouns". In this subsection, I show that classifier pronouns can sometimes behave as weak definites, and propose that it is only in such cases that Chuj pronouns are truly definite determiners with elided NPs, as proposed more generally for pronouns in Elbourne 2013.

The idea that the pronominal system of a language might be influenced by its determiner system is not new. This hypothesis is put forth by Matthewson (2008), who states that "perhaps in general, the semantics of third-person pronouns in a language L is based on the semantics of determiners *in L*". More recently, Bi and Jenks (2019), building on work by Patel-Grosz and Grosz (2017) on German and Clem (2017) on Tswefap, explicitly argue that a language's pronominal inventory should be isomorphic to its determiners, proposing the following generalization:

(138) **Determiner-pronoun parallelism** (Bi and Jenks 2019, (6))
Whatever distinction a language makes in its determiner system will be mirrored in its pronominal system.

To support this generalization, Bi and Jenks (2019) argue that Mandarin, which recall from section 3.2.1 marks the distinction between weak and strong definites, also marks it in its pronominal system. As summarized in the table below, while "weak definite pronouns" are entirely covert and combine with ι, "strong definite pronouns" tend to require a demonstrative.[22] Note that Bi and Jenks follow Elbourne (2005) in assuming that pronouns involve elided NPs.

Table 3.3: Determiners/pronouns configurations in Mandarin (Bi and Jenks 2019)

	determiner	pronoun
weak definite	ι + NP	ι + ~~NP~~
strong definite	NP + DEIX	~~NP~~ + DEIX

[22]Note that Bi and Jenks (2019) point to some complications. Namely, there are apparent instances of strong definite pronouns that do not require a demonstrative. In such cases, much like what I proposed in section 3.4.1 for Chuj, Bi and Jenks assume the presence of a null index.

Bi and Jenks establish a number of tests to contrast weak definite from strong definite pronouns. One proposed environment for weak definite pronouns is anaphora to indefinites under the scope of negation within a conditional or disjunction, or so-called "bathroom sentences" (Roberts 1989, due to Barbara Partee):

(139) Either the building does not have [a bathroom]$_i$, or it$_i$ is in a funny place.

Since the most salient (and perhaps only) interpretation of (139) is one in which the indefinite *a bathroom* appears under the scope of negation, there is no entity that satisfies the property of being *a bathroom* in the first conjunct of (139). This means that there is no individual in the discourse that can get picked up by the assignment function, and so *it* must be a weak definite in (139). As corroborated by Bi and Jenks (examples (12), (13)), demonstrative pronouns are expectedly infelicitous in Mandarin "bathroom sentences", and a null pronoun must instead be used.

Against this backdrop, consider the Chuj sentence in (140), which shows that noun classifiers can appear as pronouns in "bathroom sentences":

(140) *Malaj stumin waj Xun, o max chax laj k'en yoj winh.*
Malaj [s-tumin] waj Xun, o max chax laj **k'en** y-oj winh.
NEG.EXT A3-money CLF Xun, o NEG find NEG CLF A3-by CLF
'Xun either has no money$_i$ or he can't find it$_i$.'

As seen above, the classifier pronoun *k'en* can be used even though it has no antecedent in the discourse. This suggests that classifier pronouns can be weak definites, a welcome result if classifiers encode weak definiteness.

Bi and Jenks also show that weak (null) pronouns in Mandarin are forced in cases of *situation-dependent covariation* or so-called "president sentences" (Evans 1977). Consider the following example:

(141) *T'a Yuxquen, ha anima' te'xajan ix alkal, haxo t'a Gracias, malaj mach xajanan winh.*
T'a Yuxquen, ha anima' te'-xajan [ix alkal], haxo t'a Gracias,
PREP Yuxquen, TOP people INTS-like CLF mayor and PREP Gracias,
malaj mach xajan-an [**winh**].
NEG.EXT WHO like-CON CLF
'In Huxk'e'en, people like the (female) mayor, but in Gracias, no one likes him (i.e. the male mayor).'

In (141), the use of the pronoun *winh* has again no clear antecedent (i.e. the unique (female) mayor of Huxk'e'en is not also the unique (male) mayor of Gracias). Since Chuj classifier pronouns are allowed in such sentences, we are again led to conclude that classifier pronouns can sometimes track weak definites.

The examples in (140) and (141) ultimately suggest that there must be more than one type of classifier pronoun in Chuj. Classifier pronouns cannot always involve a null index-introducing predicate, as was proposed for classifier pronouns in section 3.4.1, since in the weak definite uses of pronouns in (140) and (141), the assignment function cannot supply a value for the index that would be required by strong definite pronouns. Therefore, I propose that weak uses of classifier pronouns instantiate cases of definite determiners with elided NPs in Chuj. As such, while strong uses of classifier pronouns in Chuj involve a classifier with a null predicate that introduces an index, weak uses involve an elided NP:

(142) (At least) two kinds of pronouns in Chuj:
a. CLF + $g(i)$ = strong pronoun
b. CLF + ~~NP~~ = weak pronoun

It should be acknowledged that if configurations like (142-b) are sometimes possible for 'weak' pronouns, it is mysterious why [CLF + NP + DEIX] configurations are not also generally possible to form 'strong' pronouns (see (125) above). I tentatively propose that the preference for (142-a) results from structural economy constraints, as proposed for similar phenomena in Cardinaletti and Starke 1999, Schlenker 2005a, Katzir 2011, and Patel-Grosz and Grosz 2017. Concretely, since [CLF + *g(i)*] is structurally less complex

than [CLF + NP + DEIX], the former is favoured. Note, though, that classifier pronouns do sometimes exceptionally co-occur with deictic particles (see footnote 19), most commonly when topicalized or focused. Though I have decided to set aside this observation for future work, it could very well be that the structural economy constraint can sometimes be lifted for pragmatic reasons, as suggested in footnote 19 from this chapter.

To summarize, I have extended the generalization proposed by Bi and Jenks in (138) to Chuj. Though I have argued that the generalization is formally correct for Chuj—since there are two kinds of pronouns—the distinction between weak and strong definites is not *overtly* reflected in Chuj's pronominal system. The conclusion that emerges is that languages that overtly distinguish weak and strong definites in their determiner system will not necessarily *overtly* make this distinction in their pronominal system.

3.5 Taking stock and some cross-linguistic implications

In this chapter, I have proposed a decompositional account of definiteness and pronoun formation in Chuj. At the heart of all of the constructions we observed were noun classifiers. I provided a preliminary analysis of noun classifiers as functional items with weak definite semantics. Essentially, noun classifiers trigger the presupposition that there is a unique satisfier of the NP in a situation, a proposal that will remain constant throughout the rest of this thesis. I then argued that strong definites (including anaphoric pronouns) are derived compositionally, by combining the weak definite semantics of noun classifiers with additional overt (or covert) morphemes signalling familiarity. Overall, while weak definites are always realized by combining a classifier with an NP, there are at least three strategies to obtain strong definiteness, summarized in table 3.4.

As discussed in section 3.4, the account has implications for theories of pronoun formation. Based on previous work on the distinction between weak definite pronouns and strong definite pronouns (Patel-Grosz and Grosz 2017, Clem 2017, and Bi and Jenks 2019),

Table 3.4: Classifier configurations

weak definite	CLF + NP	
	CLF + ~~NP~~	(= weak definite pronoun)
strong definite	CLF + NP + DEIX	
	TOP + CLF + NP	
	CLF + $[\lambda x.\ x = g(i)]$	(= anaphoric pronoun)

I argued that there are at least two kinds of pronominal constructions formed with noun classifiers in Chuj, which together reflect the distinction between weak and strong definites. I proposed that while anaphoric pronouns combine with covert index-introducing predicates to form E-type pronouns, weak definite uses of classifier pronouns involve NP ellipsis, and thus lack an index.

Finally, I suggested that there may be crosslinguistic variation in how the index responsible for introducing the familiarity presupposition with strong definites is evaluated with respect to the uniqueness trigger. Specifically, the index is introduced at a wide scope position above the uniqueness trigger in classifier–noun–deictic constructions, but below the uniqueness trigger with anaphoric pronouns. This could be a general point of cross-linguistic variation, and so "strong definites" might be expected to differ slightly in their presuppositions from language to language.

One typological question that results from the proposal in this chapter concerns the extent to which strong definites are crosslinguistically decomposable. As already discussed in section 3.3, the compositional nature of strong definites observed for Chuj is not straightforwardly captured in previous proposals, including the recent typology of definiteness marking in Jenks 2018, reproduced below (see also Moroney 2021).

Table 3.5: *Typology of definiteness marking* (Jenks, 2018)

	Bipartite	Marked anaphoric	Generally marked	Marked unique
Unique	Def_{weak}	Ø	Def	Def_{weak}
Anaphoric	Def_{strong}	Def_{strong}	Def	Ø
Languages	German, Lakhota	Mandarin, Akan, Wu	Cantonese, English	(unattested)

In this typology, *bipartite* languages are languages which overtly and distinctively mark the contrast between weak and strong definites; *marked anaphoric* languages are languages which only overtly mark strong definites, but not weak definites; *generally marked* languages are languages which overtly mark definiteness, but do not make a distinction between weak and strong definites; and *marked unique* languages would correspond to the other logical but unattested possibility: languages that mark weak definites, but not strong definites.

Crucially, under this typology of definiteness marking, weak and strong definite determiners are conceived of as separate lexical items. This seems correct for some languages, including Fering (Ebert 1971), which marks the weak definite article, *a*, distinctively from the strong definite article, *di* (see also discussion in Jenks and Konate to appear):

(143) Fering

a. Ik skal deel tu **a** / ***di** kuupmaan.
I must down to the$_{weak}$ / the$_{strong}$ grocer
'I have to go down to the grocer.'

b. Oki hee an hingst keeft. ***A** / **Di** hingst haaltet.
Oki has a horse bought the$_{weak}$ / the$_{strong}$ horse limps
'Oki has bought a horse. The horse limps.' (Ebert 1971: 161)

At first glance, Chuj appears to pattern with Fering in behaving like a bipartite language, since it overtly and distinctively marks the contrast between weak and strong definites. However, the distribution of weak and strong definites in Chuj, at least as proposed in this chapter, points towards another type of language: one which marks the distinction compositionally, as argued in section 3.3.2.

Taking this observation one step further, the distribution of Chuj definites opens up the possibility that the distinction between weak and strong definites is also compositional in other languages, as independently proposed in recent work by Hanink (2018, 2021) and Ahn (2019). In this respect, notice that the difference between Mandarin weak and strong definites, analyzed as lexically ambiguous in Jenks 2018, would require minimal modification to extend the current compositional account: ι could derive the unique-

ness presupposition for both weak and strong definites, and the Mandarin demonstrative's sole contribution, then, would be to introduce a familiarity presupposition.

There is, moreover, a typological reason to favour a decompositional analysis. As highlighted in Jenks 2018 and in table 3.5, there is a gap in the typology of definite determiners: no language only marks weak definites. While lexical-ambiguity theories do not straightforwardly predict this gap, decompositional accounts do.[23] That is, languages which only have definite determiners that trigger uniqueness presuppositions will always come out as "generally marked", since weak definiteness is just one piece in the composition of strong definites.

While the current chapter has focused on the role of noun classifiers in definite descriptions, and proposed a preliminary analysis of the composition of definiteness in Chuj, an issue of central importance still remains to be addressed. The issue is that noun classifiers can co-occur with indefinite determiners. While this fact will at first glance seem to compromise the current account (why should definites co-occur with indefinites?), I will argue in the next two chapters, based on the observation that indefinite + classifier expressions yield specific indefinites, that a version of the weak definite semantics of noun classifiers proposed in this chapter should nevertheless be maintained. An important revision to the current account will be that weak definite descriptions must be further decomposed into two pieces. The additional step of decomposition will lead us to a view of definiteness in Chuj similar to the one advocated by Coppock and Beaver (2015), where the noun classifier denotes an identity function over predicates. However, the basic tenets of the current proposal—that noun classifiers impose uniqueness and existence conditions—will remain intact.

[23] Jenks (2018) suggests that this gap can be explained if definite articles always grammaticalize from demonstratives, following Greenberg (1978). There is no principled reason to believe, however, that this is the only way definite determiners should grammaticalize. In fact, Chuj contradicts this claim: noun classifiers grammaticalized from nouns (Hopkins 2012b).

Chapter 4

Singleton indefinites and private situations

The previous chapter laid out the empirical and theoretical foundations for an analysis of definiteness in Chuj. Whereas weak definite descriptions were argued to result from the obligatory combination of a noun classifier with an overt noun, strong definite descriptions were argued to require a weak definite base, with additional index-introducing morphology. A natural move, provided only with the data from Chapter 3, was to treat noun classifiers as weak definite determiners. The analysis provided a way to capture the syntax and semantics of the first three possible DP configurations in Table 3.1, repeated below for convenience.

Table 4.1: Possible DP configurations in Chuj

	Configuration	Example	Rough translation
①	CLF + NP	*nok' tz'i'*	'the dog'
②	CLF + NP + DEIX	*nok' tz'i' chi'*	'the/that dog' (anaphoric)
③	CLF	*nok'*	'it'
④	INDF + NP	*jun tz'i'*	'a dog'
⑤	INDF + CLF + NP	*jun nok' tz'i'*	'a (certain) dog'
⑥	INDF + NP + DEIX	*jun tz'i' chi'*	'that dog'
⑦	INDF + CLF + NP + DEIX	*jun nok' tz'i' chi'*	'that dog'

When we expand our empirical focus to a broader range of Chuj nominal expressions, however, the proposal that noun classifiers are weak definite determiners becomes heavily challenged. This is because noun classifiers, our alleged weak definite determiners, can also combine with indefinite determiners. Indeed, considering line ⑤ of the above table, we face a serious empirical challenge for the current treatment of noun classifiers, statable as the "too-many-determiner" puzzle in (144):

(144) **The "too-many-determiner" puzzle**
If noun classifiers are definite determiners, why can they sometimes co-occur with indefinite determiners?

Crucially, Table 4.1 also shows that indefinite expressions do not *need* to contain noun classifiers. Indefinite expressions can be well-formed only by combining an indefinite determiner with an NP, as in ④. This is in clear contrast with definite descriptions, which *require* a noun classifier. A piece of the puzzle will thus be to understand what contribution noun classifiers have when they combine with indefinite determiners.

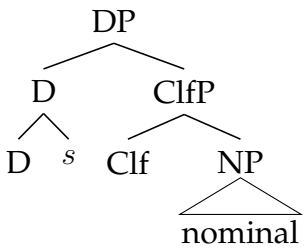
The goal of this chapter will be to provide a solution to (144), all while maintaining the core proposal about noun classifiers from Chapter 3, namely that they impose existence and uniqueness presuppositions on the interpretation of nominal expressions.[1] Building on previous work on closely-related Mayan languages (Craig 1986b, Ramsay 1985, Trechsel 1995, Zavala 2000), I will show that when noun classifiers combine with indefinite determiners, the result is a specific indefinite. I argue that this can be explained if configurations like ⑤ are interpreted as singleton indefinites (Schwarzschild 2002). In short, when co-occurring with indefinite determiners, I propose that the uniqueness condition of noun classifiers is responsible for delimiting the domain of those indefinites to a singleton set.

[1]Whenever I use the term "presupposition" alone, I mean the "semantic presupposition" associated with a lexical entry. I will use the terms "(semantic) presuppositions", "conditions", and "non-assertive content" interchangeably. As I will argue in section 5.1.2.2, the semantic presupposition of noun classifiers does not always result in a pragmatic presupposition.

In an attempt to account for (144), parts of the analysis from Chapter 3 will have to be refined throughout the next two chapters. Two key refinements in our analysis of noun classifiers will come from (i) the syntactic locus of situation pronouns within nominal expressions, and (ii) the kinds of values this situation variable may be assigned.

Regarding (i), I will follow von Fintel (1994), Büring (2004), and Schwarz (2009, 2012) (among others) in proposing that syntactically-represented domain restriction variables within extended nominal projections—here *situation pronouns*—are exclusively instantiated within complex DP heads, as exemplified in (145).

(145) DP structure in Chuj

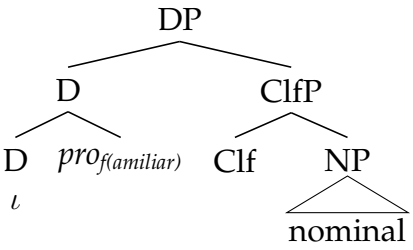


In (145), the situation argument is located in a complex DP head, above the projection of the noun classifier (it is no longer taken as the first argument of classifier, as was the case in Chapter 3). This will have as an effect that the situation within which the uniqueness condition of the noun classifier will get evaluated will only be saturated upon merger of a DP head, which I will assume is null in the case of definite descriptions, but overt in the case of indefinite expressions.

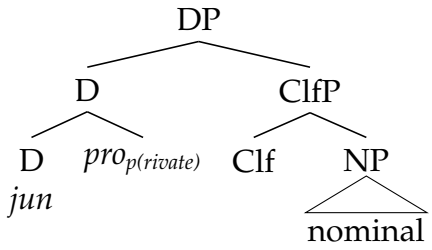
This is where refinement (ii) comes into play. Building on Beaver and von Fintel (2013) and Arsenijević (2018), I will argue for a distinction between two kinds of indexed situation variables, in accordance with the kinds of conditions they impose on the discourse participants' ability to retrieve the intended contextual domain restriction associated with an expression. In the case of definite descriptions, I argue that the DP is headed by the null morpheme, ι, which *must* combine with a situation pronoun whose value is known to both the speaker and their audience. In the case of singleton indefinites, on the other

hand, the indefinite determiner (overt in Chuj) combines with a situation pronoun whose value *does not necessarily need* to be retrievable by all discourse participants (Schwarzschild 2002, Kratzer 2003, Matthewson 2008). The general idea is schematized below:

(146) Weak definite description



(147) Indefinite+classifier DP



I argue that the difference between (146) and (147) has crucial effects for the evaluation of the uniqueness condition encoded in the lexical entry of noun classifiers. Noun classifiers impose a uniqueness condition, but the relative breadth of this condition is affected by the kind of situation considered. If the value for the situation pronoun is known to all discourse participants, the result is a classic uniqueness presupposition. If it is not, the result is a singleton indefinite.

A major theoretical proposal of the next two chapters will thus be that our ontology of situation pronouns must be expanded in the following way: not only do natural language determiners provide us with the means to restrict the evaluation domain of expressions to situations whose value is common knowledge to all discourse participants (pro_f), it also provides us with the means to do so with respect to situations that are assumed by the speaker to be unknown to a subset of discourse participants (pro_p). And while some determiners seem to require that situation variables receive common-knowledge values (e.g., definites), others seem to show no such requirement (e.g., indefinites). As I will argue, the proposal that we need private domain restriction variables is not necessarily novel. Once one embraces a variable approach to domain restriction (Barwise and Perry 1983; Westerståhl 1984; von Fintel 1994; Percus 2000; Wolter 2006; Keshet 2008; Schwarz

2009), as I will do here, it is essentially just a reformulation of what Schwarzschild (2002) termed the "Privacy Principle", couched in a variable approach to domain restriction:

(148) **The Privacy Principle** (Schwarzschild 2002, 307)
It is possible for a felicitous utterance to contain a restricted quantifier even though members of the audience are incapable of delimiting the extension of the (implicit) restriction without somehow making reference to the utterance itself.

The proposal is also related to a number of recent works on the kinds of assignments that should be considered appropriate for the proper interpretation of indexed variables within Logical Forms (LFs) (Beaver and von Fintel 2013, Silk 2016, King 2018). The Chuj data, and the Privacy Principle, essentially require us to relax the conditions imposed on the interpretation of LFs that contain free indexed variables, such that the assigned values for these free variables do not always have to be identifiable to all discourse participants. This is contrary to standard assumptions about assignment functions in which the interpretability of LFs is assumed to be contingent on the discourse participants' ability to *agree* on the value for free variables (see e.g., Heim and Kratzer 1998).

The chapter is structured as follows. Sections 4.1 and 4.2 will first spell out the empirical facts surrounding the configurations in lines ④ and ⑤ of Table 4.1 by stating the "too-many-determiner" puzzle in further detail, and providing evidence that indefinite expressions of the type found on line ⑤ give rise to "specific" interpretations. The central goal of section 4.3 will then be to provide an analysis of noun classifiers in co-occurrence with indefinites.

In adapting the proposals of Chapter 3 to the new empirical findings from this chapter, the semantic types and entries of the noun classifier and deictic particles will inevitably have to be revisited. This will be the goal of Chapter 5. In particular, in section 5.1, we return to the contribution of noun classifiers within definite descriptions, where I argue that the revised entries for noun classifiers proposed in section 4.3 can be extended to the Chuj data if weak definites are further decomposed so as to exhibit the structure schematized

in (146) above. The contribution of noun classifiers to meaning, however, will remain very similar to the one proposed in Chapter 3. Section 5.2 then turns to a discussion of the predictions made by the refined entry for noun classifiers, which are borne out, as well as a discussion of potential linguistic variation in the denotation of noun classifiers within the Mayan language family. Finally, section 5.3 will turn to deictic particles, which following observations from Chapters 2 and 3, I argue are compatible with both definite and indefinite bases alike (see lines ②, ⑥, and ⑦ of Table 4.1). I will show that the entry from Chapter 3 can be minimally revised to account for the fact that deictic particles may combine with both kinds of bases, all while keeping to the core proposal that deictic particles introduce indices, which in the case of definite demonstratives result in a strong definite description.

4.1 The too-many-determiner puzzle

If we adopt a uniqueness-based approach to noun classifiers, as proposed in Chapter 3, then examples like those in (149) and (150) initially appear to be problematic. The issue is that noun classifiers, our alleged weak definite determiners, can co-occur with indefinite determiners (see §2.5.2 for relevant background on indefinite determiners in Chuj):

(149) *Ay jun te' kajonh niwakil*
Ay [jun **te'** kajonh] niwakil.
EXT INDF CLF box big
'There's a big box (where rocks and sand are carried).' (txt, CJ220715)

(150) a. *Ewi schonh jun nok' xumpil ix ix.*
Ewi s-chonh [jun **nok'** xumpil] ix ix.
yesterday A3-sell INDF CLF hat CLF woman
'Yesterday the woman sold a hat.'

b. *Ewi sman nok' xumpil jun ix ix.*
Ewi s-man nok' xumpil [jun **ix** ix].
yesterday A3-buy CLF hat INDF CLF woman
'Yesterday a woman bought the hat.'

Repeating the "too-many-determiner" puzzle from example (144), a proper analysis of noun classifiers will thus need to provide an answer to the question in (151):

(151) **The "too-many-determiner" puzzle**
If noun classifiers are definite determiners, why do they sometimes co-occur with indefinite determiners?

What is more, noun classifiers are never actually needed in indefinite expressions. The sentences in (149) and (150) would be all equally well-formed in the absence of the bold-faced noun classifiers. Therefore, part of the "too-many-determiner" puzzle is to understand what contribution noun classifiers have when they do combine with indefinite determiners.

In the rest of this section, I first expand on some of the issues that would arise if we kept to the specific entry for noun classifiers proposed in Chapter 3 (§4.1.1). I then rule out a hypothetical partitive analysis of indefinite+classifier constructions (§4.1.2).

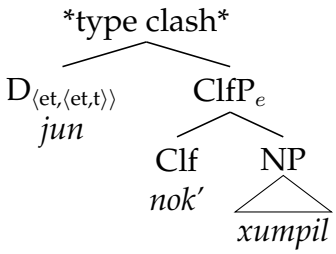
4.1.1 The puzzle

Given the entry for noun classifiers in Chapter 3, repeated below for convenience, the possibility of indefinite+classifier co-occurrences raises at least two important issues.

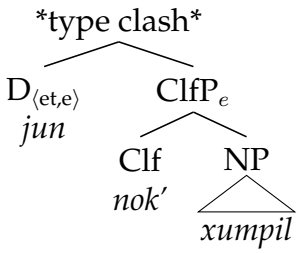
(152) Denotation of noun classifiers (repeated from (79))
$[\![\text{CLF}]\!] = \lambda s.\lambda P_{\langle e,\langle s,t\rangle\rangle}\colon |\{y\colon P(y)(s)\}| = 1.\ \iota x[P(x)(s)]$

First, assuming standard theories of indefinites, we find ourselves with clear type mismatches. In static semantics, indefinite determiners have been standardly construed in at least one of two ways: (i) as quantifiers (Russell 1905, Montague 1973, Barwise and Cooper 1981) or (ii) as choice functions (Reinhart 1997, Winter 1997, Kratzer 1998, Matthewson 1999). These two options, with a syntax that assumes that the noun classifier is generated in a ClfP below DP, are represented in the schemas below:

(153) Indefinites as generalized existential quantifiers



(154) Indefinites as choice functions



In both cases, the entry provided for noun classifiers in Chapter 3 leads to a type clash. Indeed, Chapter 3 followed standard approaches to definite determiners in situation semantics in proposing that classifier + noun constituents result in type-*e* denotations (Schwarz 2009, Arkoh and Matthewson 2013, Jenks 2018). This, as shown in the above examples, does not yield an adequate type for either approach to indefinites. One could of course try to maintain the semantic entry of the classifier from Chapter 3, and resort to type-shifting mechanisms in order to solve the compositional issues, a possibility explored in Royer 2019. For instance, applying Partee's (1986) *Ident*-shift to the classifier + noun constituent would return a predicate, allowing for composition to proceed in both cases. However, as will become clear in the discussion immediately below, this option is undesirable. The reason is that, under such a theory, the effects of the triggered (uniqueness) presupposition would be the same in the case of the indefinite and the definite, which is not the case (see example (156) below).

A second issue, and a more dramatic one, is that definites and indefinites should *a priori* not be compatible with each other. Under virtually all theories of definiteness and indefiniteness, the two are generally regarded as polar opposites of each other, or

at least as presuppositional competitors (see e.g. Abbott 2004, Heim 2011). From this standpoint, it would seem to follow that weak definite classifiers should not be compatible with indefinite determiners, contrary to fact. As recently discussed by Dayal (2019), there are generally two dimensions with which definites and indefinites have been contrasted (in addition to their semantic type), which in turn align with particular assumptions about the correct formalization of definiteness in terms of uniqueness versus familiarity. One classic account maintains that indefinites minimally contrast with definites on a uniqueness/non-uniqueness basis. Under this view, while definites assert or trigger the presupposition that the NP predicate has a unique satisfier in the context, indefinites do not (Frege 1892, Russell 1905, Strawson 1950, Hawkins 1978, a.o.). A second line of account views the difference in terms of familiarity. Under this view, while definites trigger a familiarity presupposition, indefinites do not, and the latter are therefore generally associated with a novelty condition (Christophersen 1939, Karttunen 1976, Heim 1982, a.o.). To the extent that these conditions should be encoded in the denotation of definite versus indefinite determiners, we might therefore expect the two to be incompatible.

Crucially, indefinites with noun classifiers in Chuj comply with both of the above-mentioned conditions on indefinites (i.e., non-uniqueness and non-familiarity). The following example, taken from a narrative, provides positive evidence in favour of this point.

(155) **Contextualization**: This sentence was uttered in an interview in which the speaker is recounting her experience in her local school system. The speaker is remembering the different teachers who instructed her throughout her experience:
Ay jun ix maestra skuchan Sofia.
Ay #(**jun**) **ix** **maestra** s-kuchan Sofia.
EXT INDF CLF teacher A3-called Sofía
'There was a teacher called Sofía.' (txt, CJ240715)

Given the context in which this sentence was uttered, the indefinite expression *jun ix maestra* 'a teacher' clearly complies with both prototypical diagnostics for indefinites: (i)

the context provides us with a situation in which there is clearly not a unique teacher (i.e., non-uniqueness holds), and (ii) the speaker is recounting her experience for the first time, has not yet talked about this teacher in the interview, and the interviewer is not familiar with said teacher (i.e., non-familiarity holds). As shown by the felicity judgement in the example, the presence of the indefinite determiner is necessary in this case.

More importantly, the felicity of the sentence in (155) in the described context is unexpected given the proposal from the previous chapter that classifiers contribute the presupposition that there is a unique satisfier of the NP property in a situation. Clearly, in co-occurrence with indefinite determiners, noun classifiers do not trigger a uniqueness presupposition (at least not with respect to a salient situation known to both the speaker and addressee). This is in striking opposition to DPs without an overt indefinite determiner, which *do* trigger such a presupposition. This difference can be further highlighted by the minimal contrast between the sentence in (156), with an indefinite DP, and the sentence in (157), with a definite DP:

(156) **Context**: There are five priests in Yuxquen and the speaker and addressee know it. The speaker and addressee live and currently are in Yuxquen. Out of the blue, the speaker says:

Ixinlolon yet' jun pale ewi / Ixinlolon yet' jun winh pale tikneik.

Ix-in-lolon yet' [#(jun) (**winh**) pale] tikneik.
PFV-B1S-speak with INDF CLF priest today
'I spoke with a priest today.'

(157) **Context**: There is just one priest in Yuxquen and the speaker and addressee know it. The speaker and addressee both live and currently are in Yuxquen. Out of the blue, the speaker says:

Ixinlolon yet' winh pale tikneik.

Ix-in-lolon yet' [(#jun) #(**winh**) pale] tikneik.
PFV-B1S-speak with INDF CLF priest today
'I spoke with the priest today.'

The context in (156) is such that there is *more than one relevant priest* in the relevant situation surrounding the utterance. As indicated, *jun* must appear in this sentence, and while the

presence of the noun classifier is not required, its presence is appropriate. In contrast, the context in (157) is such that there is *only one priest* in the relevant situation. Accordingly, the noun classifier becomes obligatory, and the presence of an indefinite determiner in this context is judged infelicitous.

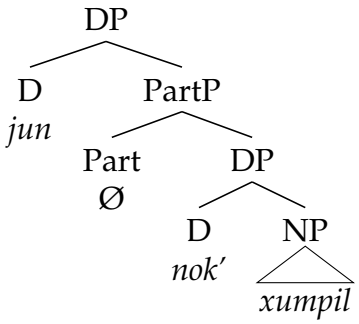
In sum, we have seen that indefinite + classifier co-occurrences behave differently from definite descriptions on at least two levels: (i) classifiers are required with definites, but not with indefinites, and (ii) bare classifier phrases trigger a presupposition to the effect that there exists a unique satisfier of the NP, while indefinite + classifier phrases do not seem to. Before summarizing this section, a brief note on partitive structures in Chuj is in order.

4.1.2 Ruling out a partitive analysis

Before concluding this section, it is important to establish that indefinite + classifier expressions are not partitive. If they were, then these data would not actually be problematic; the relevant DP in (158-a), for instance, would simply translate as *one of the hats*. And if this were the case, the bracketed DP in (158-a) could simply exhibit a partitive syntax like (158-b), with two separate DP layers.

(158) a. *Ewi schonh jun nok' xumpil ix ix.*
Ewi s-chonh [jun **nok'** xumpil] ix ix.
yesterday A3-sell INDF CLF hat CLF woman
'Yesterday the woman sold a hat.'

b. Hypothetical partitive structure for *jun nok' xumpil*



In addition to consistent non-partitive translations provided by Chuj collaborators, as well as controlled judgement tasks with contexts where a partitive is not supported, a partitive analysis of these data can be ruled out for at least two more reasons. First, across languages, partitives have been shown to be disallowed in existential constructions like the ones in (149), repeated in (159) (see e.g. Milsark 1974; Enç 1991). If a partitive construction were actually involved in the example in (159), then the exceptional grammaticality of this construction would be left unexplained.

(159) *Ay jun te' kajonh niwakil*
Ay [jun **te'** kajonh] niwakil.
EXT INDF CLF box big
'There's a big box (where rocks and sand are carried).' (txt, CJ220715)

A second reason to disfavour a partitive analysis of indefinite + classifier phrases is that partitives can sometimes be overtly distinguished from non-partitives in Chuj. Compare, for example, (160) with (150-b) above. The only difference is in the presence of the plural marker *heb'* within the DP:

(160) *Ewi sman nok' xumpil jun heb' ix ix.*
Ewi s-man nok' xumpil [jun **heb' ix** ix].
yesterday A3-buy CLF hat INDF PL CLF woman
'Yesterday one of the women bought the hat.'

The presence of plural marking within partitive DPs in Chuj is entirely expected given previous work on partitivity: partitives tend to only select pluralities (see e.g. de Hoop 1997 and Chierchia 1998). It is important to note, however, that only human nouns receive overt plural marking in San Mateo Ixtatán Chuj. This means that the distinction between partitive and non-partitive DPs is often opaque. This is the case, for instance, with the DP *jun nok' xumpil* in (158-a), which is ambiguous between a non-partitive reading ('a hat') and a partitive one ('one of the hats'). The data discussed in the rest of this chapter has been controlled for this ambiguity.

4.1.2.1 Summary

The optional co-occurrence of noun classifiers with indefinite determiners raises compositional challenges for the analysis of noun classifiers provided in Chapter 3. In particular, two relevant observations stand out:

(161) Relevant empirical differences between definite and indefinite DPs in Chuj:
1. While definite DPs require a noun classifier, indefinite DPs can take a classifier, but do not require one.
2. While bare classifier DPs induce existence and uniqueness presuppositions, indefinite DPs with classifiers do not seem to.

Provided with this state of affairs, we have two options regarding the semantics of noun classifiers. One option would be to reject the proposal from Chapter 3, and link the presuppositions seen in definite classifier DPs to an external vacuous morpheme or type-shifting operation (see e.g. Chierchia 1998). A second option would be to maintain a version of the weak definite analysis of noun classifiers from Chapter 3, and explain why the semantic presuppositions of the noun classifier are apparently absent when the noun classifier co-occurs with an indefinite.

In the next section, I argue for the latter option, based on a third empirical observation about indefinite + classifier phrases. Building on previous work on Q'anjob'alan languages (Ramsay 1985, Craig 1986b, Trechsel 1995, Zavala 2000), I show that the inclusion of a noun classifier in an indefinite DP is not without semantic consequence: it forces a *specific* interpretation of that indefinite. This observation will lead to the proposal in section 4.3 that the semantic presupposition of the noun classifier in combination with an indefinite is to reduce the restrictor of that indefinite to a singleton set, which will account for the specific indefinite properties established (Schwarzschild 2002). The result will be a unified entry for the noun classifier, which keeps to its non-assertive uniqueness and existence contribution defended in Chapter 3.

4.2 Classifier-indefinites as specific indefinites

In this section, I show that indefinites containing classifiers exhibit properties that other indefinites do not exhibit. In particular, they (i) appear to take obligatory wide scope over operators, and (ii) they show "epistemic specificity", both of which have been described as canonical properties of "specific indefiniteness" in the literature (see e.g. Farkas 2002 and Farkas and Brasoveanu 2021).[2] I first provide a brief summary of what is meant by wide-scope and epistemic interpretations of indefinites, and then provide evidence that indefinite DPs with noun classifiers consistently lead to such interpretations.

4.2.1 Specific indefinites

The exceptional scope-taking abilities of indefinites have received considerable investigation in semantic theory, first propelled by Fodor and Sag's (1982) influential paper, and followed by a number of works and important insights in the 1990s and early 2000s (Farkas 1981, 2002; Abusch 1994; Cresti 1995; Reinhart 1997; Winter 1997; Kratzer 1998; 2003; Matthewson 1999; Chierchia 2001; Schwarz 2001; von Heusinger 2002; Portner and Yabushita 2001; Portner 2002; Schwarzschild 2002; and Breheny 2003). While different semantic analyses have been proposed, these authors all agree on the fact that indefinites are less limited than universal quantifiers in their ability to be interpreted outside the scope of well-established syntactic islands (Ross 1967). An example (162-a) with its pertinent wide-scope interpretation (162-b) is provided below:

[2]Note, however, that specificity is a controversial notion that has been used to describe many different linguistic phenomena. In this chapter, I only concentrate on so-called scopal and epistemic specificity. Also note that epistemically specific indefinites and wide-scope indefinites are lumped together under the term "referential indefinites" in Fodor and Sag 1982. Since I ultimately keep to a quantificational analysis of indefinites, I will avoid this term in this thesis. For overviews on specificity distinctions, see von Heusinger 2011, 2019 and Farkas and Brasoveanu 2021.

(162) a. If a friend of mine from Texas had died in the fire, I would have inherited a fortune.

b. $\exists x[x$ is a friend of mine from Texas $\wedge$
$[x$ had died in fire $\rightarrow$ I would have inherited a fortune$]]$

Sentences like (162-a) led Fodor and Sag (1982) to propose, building on previous work (e.g. Chastain 1975, Wilson 1978), that indefinites are ambiguous between a quantificational and referential interpretation. This provided a solution to the scope puzzle: the apparent exceptional scope-taking abilities of indefinites were not due to the fact that they differed from other quantificational expressions in being able to scope out of syntactic islands, but rather due to the fact that they could also be interpreted referentially (contrary to other quantifiers). Under a referential interpretation, indefinites essentially refer to a particular entity, and therefore give the illusion of a wide-scope reading.

Following the work of Fodor and Sag, additional scope-taking abilities of indefinites were subsequently identified, casting doubt on their proposed ambiguity. In particular, many authors showed that indefinites could also receive "intermediate-scope" interpretations (Farkas 1981, King 1988, Abusch 1994), a possibility not predicted by the referential / quantificational ambiguity. A relevant example adapted from Schwarzschild (2002: 295) is provided below.

(163) Everyone$_1$ had read most of the reviews about a movie that happened to be their$_1$ favourite.

As Schwarzschild mentions, a very natural interpretation of (163) can be paraphrased with (164). Under this interpretation, the indefinite appears to scope between *everyone* and *most*:

(164) For each person, there was a movie that was their favourite and they had read most of the reviews that were written about it.

The existence of intermediate scope interpretations is problematic for the treatment of indefinites in sentences like (162-a) as purely referential. In the intermediate-scope inter-

pretation paraphrased with (164), the value of the indefinite covaries with each person, and thus does not refer to a particular entity in the world. Such observations led a number of scholars to explore alternative accounts of the exceptional scope-taking abilities of indefinites, including choice-functional approaches (e.g. Reinhart 1997, Winter 1997, Kratzer 1998, Matthewson 1999, von Heusinger 2002, Dawson 2020) and singleton domain restriction approaches (Portner and Yabushita 2001, Portner 2002, Schwarzschild 2002, Breheny 2003).[3]

It is common for work on specific indefinites to establish a descriptive contrast between "scopal specificity", as seen above, and "epistemic specificity" (von Heusinger 2011, 2019). The latter type of specificity has been traditionally contrasted with wide scope indefinites given that epistemic specificity cannot be explained in terms of scope. However, in most of the above-cited work, wide-scope indefinites and "epistemically-specific indefinites" receive the same analysis—in Fodor and Sag 1982, for instance, they both fall under the purview of referential indefinites. An example of such an indefinite in English, due to Fodor and Sag (1982, 355), is provided below:

(165) A student in the syntax class cheated on the final exam.

As Fodor and Sag (1982) note, the above indefinite can receive two interpretations. Under the specific interpretation, the speaker has a particular cheating student in mind. Under the non-specific interpretation, the speaker merely knows that one of the students cheated, but does not know which student that is.

As noted by von Heusinger (2019), much work on specific indefinites has followed the path of Fodor and Sag, as I will do below, in reducing wide-scope indefinites and epistemic specificity to a unitary phenomenon (see also Farkas and Brasoveanu 2021 for a similar proposal). In the next subsection, I go through examples that demonstrate that co-occurrences of noun classifiers with indefinites always lead to specific indefinite inter-

[3]See also Endriss 2009, Brasoveanu and Farkas 2011, and Charlow 2014, 2020 for yet other accounts of the exceptional scope-taking abilities of indefinites.

pretations, including both in the context of wide scope indefinites and epistemic specificity.

4.2.2 Specific indefinites in Chuj

In this section, I show that when a noun classifier co-occurs with an indefinite in Chuj, it forces a specific interpretation of that indefinite. The claim that classifier-indefinite co-occurrences result in "marked" interpretations of indefinites broadly related to "reference" is not entirely novel. As already mentioned in Chapter 3, previous work on Q'anjob'alan languages, starting with Craig's (1977) thorough investigation of Popti' (formerly known as Jakaltek), connected the use of noun classifiers with referential semantics (see also Ramsay 1985, Trechsel 1995, and Zavala 2000). For instance, Craig (1986b: 273) proposed that:

> In their noun adjunct function [as opposed to their pronominal function] the classifiers mark referential NPs which are thematically important. This thematic importance is the feature shared by the marked indefinite and the definite NPs.

Likewise, on another closely-related Mayan language called Akatek, Zavala (2000: 140) notes that:

> Noun classifiers are used to explicitly mark third-person nominals as individuated, referential and thematically important items in discourse. In contrast, non-individuated and non-referential nominals as well as nominals which refer to participants of backgrounded sections of discourse are not tagged with noun classifiers.

In other words, both authors concur on the fact that when indefinite determiners co-occur with noun classifiers, semantic or pragmatic effects broadly related to reference arise. Here, I will argue that these effects can be understood if noun classifiers force specific interpretations of indefinites.

Before considering examples of indefinites that co-occur with noun classifiers, it is useful to first consider the example in (166), where the indefinite surfaces without a classifier. This example, adapted from Matthewson (1999), features an operator with which the indefinite could interact with scope. Assuming a restrictor analysis of conditionals (Lewis 1975; Kratzer 1986), the specific interpretation of the indefinite arises when the existential is interpreted outside the scope of the covert universal modal. The non-specific interpretation arises when the existential is interpreted within the restriction of that universal modal. The scenarios in (167) provide appropriate contexts for these two possibilities.

(166) **Context:** Malin is organizing a party in the village.
Tejunk'olal ix Malin tato tz-jaw jun icham.
Te-junk'o'olal ix Malin [tato tz-jaw [jun icham]].
INTS-happy CLF Malin if IPFV-come INDF elder
'Malin will be happy if an elder comes (to the party).'

(167) Scenarios for (166)
a. There is just one elder called Xun, such that if Xun comes to the party, Malin will be happy. wide scope/specific = ✓
b. Malin will be happy if at least one elder comes to the party, but it doesn't matter who. narrow scope/non-specific = ✓

The felicity of both scenarios in (166) indicates that in the absence of a noun classifier, an existential in Chuj can receive both a wide or narrow scope interpretation. In that sense, bare indefinite-noun sequences in Chuj behave just like regular indefinites in languages like English in admitting both specific (wide scope) and non-specific (narrow scope) interpretations (i.e., the English translation is also fine in both scenarios).

When an indefinite quantifier co-occurs with a noun classifier, on the other hand, the result is different: the indefinite must be interpreted as if it were taking wide scope over other operators, including out of syntactic islands. This can be observed in the example in (168), which minimally contrasts with (166) in that the indefinite co-occurs with the classifier for male entities, *winh*.

(168) **Context**: Malin is organizing a party in the village

Tejunk'olal ix Malin tato tzjaw jun winh icham.

Te-junk'o'olal ix Malin [tato tz-jaw [jun **winh** icham]].
INTS-happy CLF Malin if IPFV-come INDF CLF elder
'Malin will be happy if an elder comes (to the party).'

(169) Scenarios for (168)

a. There is just one elder, Xun, such that if Xun comes to the party, Malin will be happy. specific = ✓

b. Malin will be happy if at least one elder comes to the party, but it doesn't matter who. non-specific = ✗

As demonstrated by the (in)felicity of (168) in (169), the co-occurrence of a noun classifier with an indefinite forces a specific interpretation of the indefinite. In other words, the only possible interpretation of (168) is one in which the indefinite DP *jun winh icham* 'an elder' takes wide scope over the universal modal that the antecedent of the conditional restricts.

Another example of the wide scope behaviour of noun classifiers is presented in (170) and (171). In this case, the existential could potentially be interpreted either outside or within the scope of the universal quantifier *junjun* 'each'.

(170) *Junjun kinhib'al, tzmunlaj jun ix ix t'a chonh.*

Junjun kinhib'al, tz-munlaj [jun **ix** ix] t'a chonh.
each morning IPFV-work INDF CLF woman PREP store
'Each morning, a woman works in the store.'

(171) Scenarios for (170)

a. Each morning, the same woman works in the store. specific = ✓

b. Each morning, only one woman works in the store, but this can vary (e.g. Malin works Mondays, Xuwan Tuesdays, etc.) non-specific = ✗

In (170), the only possible interpretation of the indefinite DP with the classifier, *jun ix ix* 'a woman', is one in which the existential appears to take wide scope over the universal quantifier *junjun* 'each'. In other words, the identity of the woman working in the store cannot co-vary with the days of the week; it has to be the same woman. Though not

explicitly illustrated here, the wide scope interpretation is no longer forced when the indefinite appears without a classifier, and accordingly, such sentences are compatible with the scenarios in (171).

The two examples just provided show that noun classifiers force a specific interpretation of indefinites introduced with the singular quantifier *jun*, but note that noun classifiers force specific interpretations of *all* indefinite quantifiers in the language. For instance, when a noun classifier co-occurs with the plural indefinite quantifier *juntzanh*, the result is also a specific indefinite:

(172) *Hingana tzinman juntzanh ch'anh libro.*
Hin-gana tz-in-man [juntzanh **ch'anh** libro].
A1S-desire IPFV-A1S-buy some CLF book
'I want to buy some books.'

(173) Scenarios for (172)

a. There are books I want to buy, and I know exactly which ones. specific = ✓

b. I want to buy books, because I feel like I don't read enough. However, I don't know which ones I'll buy yet. non-specific = ✗

In the above example, the classifier forces an interpretation in which the speaker already knows exactly which books they want to buy, as shown by the felicitous and infelicitous scenarios in (173). Again, the only possible interpretation of (172), where a classifier co-occurs with the plural indefinite determiner, is one in which the indefinite takes wide scope over the modal *want*. In the rest of this chapter, we focus exclusively on singular indefinites.

So far, all of the Chuj examples that were provided in this subsection involved alternative interpretations of indefinites that could be potentially framed in terms of scope: indefinite+classifier DPs force wide-scope interpretations. It can also be shown, however, that the co-occurrence of a noun classifier with an indefinite leads to epistemically specific interpretations of indefinites. Consider the following example:

(174) *Hijan ixkochi' nok' kaxlan, ixik' b'at nok' jun nok' tz'i'.*
Hijan ix-ko-chi' nok' kaxlan, ix-ik' b'at nok' [jun **nok'** tz'i'].
almost PFV-A1P-eat CLF chicken PFV-bring DIR.go CLF INDF CLF dog
'We were going to eat the chicken, but a dog stole it.'

(175) Scenarios for (174)

a. A particular dog known to the speaker, say Fido, stole the chicken. specific = ✓

b. There are traces of a dog in the speaker's house, e.g. paw prints, that lead the speaker to think that a dog stole the chicken. non-specific = ✗

In (174), the indefinite DP *jun tz'i'* 'a dog' in the second conjunct can appear with or without a noun classifier. Chuj collaborators agree that to felicitously use the noun classifier, the speaker must, in some broad sense, have a particular dog in mind. This is demonstrated by the felicitous and infelicitous scenarios in (174). Importantly, note that for the speaker to "have a particular dog in mind", they must not necessarily need to be able to *identify* the dog.[4] This utterance is felicitous if the speaker is reporting facts that they heard from someone else about an event in which a particular dog called Fido stole the chicken, without them necessarily being able to identify Fido.

4.2.3 Summary

In sum, when appearing with a noun classifier, Chuj indefinite expressions must be interpreted as specific. This includes cases where an indefinite appears to take obligatory wide scope over an operator, and cases of so-called epistemic specificity. Note, however, that these labels are only descriptive. In the following section, I treat both kinds of specific indefinites as the same phenomenon (following many authors; see e.g., von Heusinger 2019). I propose an analysis of these constructions in which the noun classifier essentially serves to restrict the domain of the indefinite quantifier to a singleton set.

[4] I assume that if the speaker knows a particular dog stole the food, then they can create a mental representation for this dog, even though they are not able to physically identify the dog in question. This mental representation for this dog can in turn be a part of the situation which the speaker intends as the value for the situation pronoun.

4.3 Indefinites and singleton domain restriction

The previous section showed that when a noun classifier co-occurs with an indefinite, a specific indefinite interpretation arises. Given these new data, the goal of this section will be to revise the entry for noun classifiers such that: (i) we can explain the specific indefinite facts from section 4.2, and (ii) we can have a unified account of noun classifiers, which also accounts for their obligatory participation in weak definite descriptions.

The guiding strategy to unify the contribution of noun classifiers will come from a key observation from Schwarzschild (2002) about a fundamental difference between specific indefinites and definites. As he notes:

> There are familiarity conditions on the use of definites which do not apply to indefinites. It would be odd for me to assert out of the blue ***the aluminium toothbrush*** *is in a museum in New Hampshire*, despite the fact that there is a unique aluminium toothbrush. However, I could, out of the blue, speak of there being ***an aluminium toothbrush*** *in New Hampshire* and I would, in this case, be using a singleton indefinite, a complete one in fact. Fodor and Sag's *a friend of mine* is also singleton, albeit incomplete, and since it is likewise indefinite there is no requirement that the 'referent' be familiar to all discourse participants. This freedom appears to allow the content of the contextual supplementation to be less transparent to the hearer in a way that would be impossible with a definite. What we have in effect is an incomplete indefinite description, where the completion is asymmetrically available to the speaker but not to the hearer. (Schwarzschild 2002: 292)

Along the same lines, Dayal (2019: 59) recently argues that while "definites denote individuals who are identifiable by discourse participants, the referent of a specific indefinite is typically not identifiable to the hearer". Following a number of authors who argue that the locus of domain restriction is the determiner (von Fintel 1994; Büring 2004; Gillon 2006; Schwarz 2009, 2012), I will argue that the 'familiarity' distinction observed in the use of definite versus indefinite expressions and pinpointed by Schwarzschild and Dayal

boils down to a selectional constraint on the "kind" of situation variable that such expressions take. Roughly, I propose that while definite determiners constrain the possible value of their situation pronoun to one that is identifiable to all discourse participants, indefinite determiners combine with one whose value is not necessarily familiar to all discourse participants, at least from the speaker's perspective. In other words, natural language provides us with the means to interpret expressions not only with respect to "familiar situations", but also "private situations":

(176) Situation variables in DPs come in two varieties:

a. familiar situation variables (pro_{f_i}), which typically combine with definite determiners, are accessible to all discourse participants.

b. private situation variables (pro_{p_i}), which typically combine with quantificational determiners, are potentially not accessible to all discourse participants.

Keeping with an approach in which noun classifiers encode a semantic presupposition to the effect that there is a unique satisfier of the NP, the key modification in providing a unified semantic treatment of noun classifiers will lie in how and when the situation argument of the classifier gets saturated. When it gets saturated by a familiar situation pronoun, I argue that the semantic presuppositions of the classifier result in classic pragmatic presuppositions. When the classifier's situation argument is saturated by a private situation pronoun, on the other hand, its semantic presuppositions do not result in pragmatic presuppositions, and instead, the result is a singleton indefinite. As I will discuss in Chapter 5, the resulting theoretical picture is thus one in which semantic presuppositions or the non-assertive component associated with lexical items are not always mapped to pragmatic presuppositions (see e.g., Matthewson 2006 and Tonhauser et al. 2013 for related claims).

In section 4.3.1, I first provide more background on the singleton approach to specific indefinites. Sections 4.3.2 and 4.3.3 then spell out the claim in (176) in more detail, and provide evidence, building on a number of previous works (Partee 1973, Beaver and von

Fintel 2013, Silk 2016), for independent motivation in discriminating between at least two kinds of variables in our semantic model. Section 4.3.4 then shows how singleton indefinites can be compositionally derived from combining an indefinite with a noun classifier. Section 4.3.5 then summarizes, and provides a brief discussion of remaining questions and of a potential alternative account.

4.3.1 Specific indefinites as singleton indefinites

Schwarzschild (2002) argues that specific indefinites should derive from implicit domain restriction of an indefinite quantifier to a singleton (see also von Fintel 1999, Portner and Yabushita 2001, Portner 2002, Schwarzschild 2002, Breheny 2003 for essentially the same proposal). Under this view, specific indefinites arise not because they take wide scope relative to an operator, nor because they are referential, but because there is an explicit or implicit domain restrictor that delimits the extension of the quantifier's restrictor to a singleton set (see also Collins 2018). Schwarzschild exemplifies this proposal with elided syntactic constituents, showing how the quantificational domain could be implicitly restricted to denote a singleton set:

(177) If a relative of John <that I have in mind> dies, he will inherit a fortune.

This kind of approach can explain key properties of specific indefinites across languages. For instance, it is compatible with Farkas and Brasoveanu's (2021) proposal that the unifying feature of specific indefinites across languages is *stability of reference*, versus *variability of reference*. Essentially, specificity markers tend to impose a constraint that limits variation of values for relevant variables across a set of contextual assignments to a unitary entity. Singleton domain restriction does just this: if the existential quantifier can only quantify over a single individual, then stability of reference follows naturally.[5]

[5]Note that in Schwarzschild 2002, there may still be covarying uses of "singleton indefinites". This happens when the restrictor contains a bound variable. That is, for Schwarzschild (2002: 295), "a 'singleton indefinite' is an indefinite whose restrictor has a singleton extension, *relative to each relevant assignment of values to any bound variables in the restrictor*". See § 4.3.5.1 for similar observations in Chuj.

At the heart of singleton restriction approaches to specific indefinites lies a crucial assumption about the general availability of implicit domain restriction, at least with certain quantifiers. As discussed at the outset of this chapter, Schwarzschild titles this restriction The Privacy Principle (repeated below for convenience):

(178) **The Privacy Principle** (Schwarzschild 2002, 307)
It is possible for a felicitous utterance to contain a restricted quantifier even though members of the audience are incapable of delimiting the extension of the (implicit) restriction without somehow making reference to the utterance itself.

Whatever one's approach to domain restriction (and more than one approach could potentially co-exist; Collins 2018), the Privacy Principle describes a necessary empirical generalization about the basic workings of quantifier domain restriction. If implicit domain restriction results from covert constituents, as in (177), then the Privacy Principle implies that covert syntactic constituents must sometimes be licensed in the absence of linguistic antecedents (in defiance of basic ellipsis principles; Hankamer and Sag 1976). On the other hand, if domain restriction is implemented through covert contextual variables in the syntax (Cooper 1975; Barwise and Perry 1983; Westerståhl 1984; von Fintel 1994; Percus 2000; Martí 2003; Büring 2004; Wolter 2006; Keshet 2008; Schwarz 2009, 2012), then contextual variables must sometimes be assigned values that are not identifiable to the hearer. In a situation semantics, where situation variables contribute domain restriction, the Privacy Principle forces us to conclude that the latter route is sometimes possible.

In sum, if we want to provide a singleton indefinite analysis of specific indefinite expressions, like indefinite + classifier expressions in Chuj, we minimally need the following two ingredients:

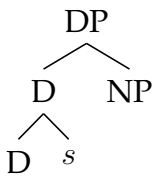
(179) Ingredients to derive specific indefinites as singleton indefinites.
a. A *singleton domain* shifting operation must be imposed on the indefinite;
b. The output of this domain shifting operation can be *unidentifiable* to a subset of discourse participants (i.e., the Privacy Principle holds).

In section 4.3.4, I propose that noun classifiers are essentially the overt realization of the ingredient in (179-a): they restrict the quantificational domain of the indefinite to a singleton set. First, however, the next two subsections (i) spell out the need for ingredient (179-b) in more detail (§4.3.2) and (ii) provide a way to formally capture the Privacy Principle in the current framework (§4.3.3).

4.3.2 Two kinds of free variables

One of the oft-mentioned benefits of situation semantics is that it offers a unified account of a number of seemingly disparate phenomena, one of which is contextual domain restriction effects (Kratzer 2021). As in other accounts of quantifier domain restriction (Cooper 1975, Westerståhl 1984, von Fintel 1994, Martí 2003), proponents of situation semantics assume that syntactically-represented variables contribute information about the context within which the truth conditions of an utterance, or of expressions within this utterance, are determined. Schwarz (2012) provides detailed argumentation that, within DPs, the syntactic locus for this variable is D^0 (see also Büring 2004):

(180) Situation arguments as sister of D^0:

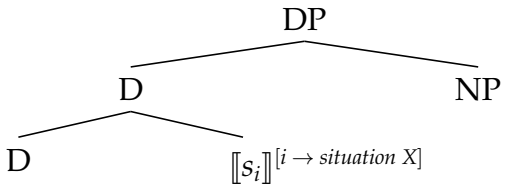


As also argued in Schwarz 2009, 2012, the value for this situation within DPs may be obtained in one of at least three ways: (i) it could be bound by the topic situation (with which the entire sentence gets evaluated), (ii) it could be bound by a quantifier over situations, or (iii) it could be free, and accordingly, receive its value from a contextually-determined assignment function.[6]

[6] As shown in von Fintel (1994), contextual variables restricting quantifiers behave like other more commonly accepted variables in showing deictic, discourse anaphoric, and bound readings. This is shown in the following examples adapted from Beaver and von Fintel 2019, slide 57:

Of particular interest to us here is what kinds of restrictions apply to free situation variables. Starting with textbook basics, a free indexed variable in the sister of D^0, as in (181), will have to be mapped to a particular situation in the range of the assignment function, via Heim and Kratzer's (1998) *pronouns and traces rule* (182).

(181) Free situation variable



(182) Pronouns and traces rule (Heim and Kratzer 1998: 241)
If α is a pronoun or a trace, g is a variable assignment, and $i \in \text{dom}(g)$,
then $[\![\alpha]\!]^g = g(i)$

Given the Privacy Principle, which implies that the value for situation variables can sometimes not be identifiable to all discourse participants, a fundamental question is what it takes for Logical Forms (LFs) with free variables to be interpretable. On that question, Heim and Kratzer (1998: 243) provide the following relevant discussion:

> If you utter a sentence like (I) *She is taller than she*, then your utterance is felicitous only if the utterance situation provides a value for the two occurrences of the pronouns "she". Given that referring pronouns bear indices at LF, (I) has some representation such as (II) *She*$_1$ *is taller than she*$_2$ and we can think of an utterance situation as fixing a certain partial function from indices to individuals. An appropriate utterance situation for LF (II) is one that fixes values for the indices 1 and 2.

In the paragraphs following the above quote, Heim and Kratzer (1998: 243) go on to propose the *Appropriateness Condition*:

(i) a. Can **everybody** hear me? (deictic/exophoric reading)
b. Yesterday I met **a group of students**. **Most** were working hard. (anaphoric reading)
c. Whenever I met **a group of students**, **most** were working hard. (bound reading)

(183) **The Appropriateness Condition**
A context C is *appropriate* for an LF ϕ only if C determines a variable assignment g_C whose domain includes every index which has a free occurrence in ϕ.

The *Appropriateness Condition* imposes a pragmatic constraint on the *domain* of the assignment function: all free occurrences of an indexed variable must be in the domain of the assignment function. The condition does not specify whether the assignments for these indices map to values that must be identifiable to all discourse participants. Nonetheless, the above quote from Heim and Kratzer does lead us to the initial impression that all discourse participants should be able to identify the mappings (from indices to linguistic expressions) set by the assignment function (see also discussion in Beaver and von Fintel 2013, 2019). Indeed, it is inappropriate to utter (184) if the addressee(s) cannot identify what the two occurrences of *she/her* refer to.

(184) She_1 is taller than her_2

Likewise, it is inappropriate to use a definite or deictic DP if the addressee does not know the context within which these expressions are to be evaluated:

(185) **Context:** We're having dinner. We never talked about whales together. I tell you:
a. # Hey, by the way, [**the s_i whale**] escaped!
Defined iff the value for $g(i)$ is familiar to all discourse participants.
If defined, $\text{ESCAPED}(x)(\iota x[\text{WHALE}(x)(s_i)])$
b. # Hey, by the way, **that whale** escaped!

In a situation semantics, the problem with (185) could well be recast as the addressee(s) failure to identify the value for the bolded expressions' free domain restricting situation variables. After all, how could the addressee(s) judge whether the uniqueness presupposition of the definite article (185-a) is met if they have no knowledge of the context within which the uniqueness presupposition is meant to be interpreted?

While these data point to a theory in which the values for free variables must be known to all discourse participants, there is also evidence that free variables can sometimes be

assigned values that are not identifiable to all discourse participants, as shown by Partee (1973) and recently highlighted in Beaver and von Fintel 2013, 2019, Silk 2016, Arsenijević 2018 and King 2018. Partee (1973) provides the following insightful example:

(186) They$_7$ haven't installed my telephone yet. (Partee 1973: 603)

The pronoun in (186), which clearly is not bound by a quantifier over situations, "seems to be referring to whoever it is that's supposed to install the telephone" (Partee 1973: 603). Here, a free variable pronoun thus appears to be assigned a value that is not known to the discourse participants.

Arguing that tense must be treated with a variable semantics, a now widely held view of the semantics of tense, Partee further provides the following example:

(187) John went to a private school. (Partee 1973: 603)

As Partee argues, the actual value for the past tense in (187) does not need to be known in order to adequately judge the truth conditions of the sentence. That John went to private school at some *underspecified* point in the past is all that matters.

In sum, free variables do not always map to values that are known to all of the discourse participants. If domain restriction is executed through contextual variables, this is a necessary implication of Schwarzschild's Privacy Principle. At the same time, it does seem like a subset of expressions do actually require that the value for the domain restriction variable be known or familiar. As discussed above, this is the case with third person pronouns (184) and definite descriptions (185). We therefore arrive at the following empirical generalization about situation variables:

(188) **Situation (non)familiarity**
Some DPs (e.g., definites) require that the value for their contextual/situation variable be known to all discourse participants; other DPs (e.g., indefinites) show no such requirements.

Beaver and von Fintel (2013) note a number of possible solutions to explain the generalization set by *Situation (non)familiarity*.[7] As foreshadowed earlier in this section, one possibility they mention is to differentiate between two kinds of variables, those which must be "familiar" to all discourse participants, and those which do not have to be (which they label "secret"). The next section briefly sketches a way to implement this distinction in our semantic model.

4.3.3 Private situations

Building on the empirical observations from the previous section, as well as on Beaver and von Fintel (2013) who suggest that values for free variables could sometimes be secret, I propose that natural language provides us with the means to restrict the evaluation domain of expressions to two kinds of situations: (i) those whose value is common knowledge for all discourse participants ($s_{familiar}$) and (ii) those whose value is potentially unknown to a subset of discourse participants ($s_{private}$). To formalize such a system, we could add an additional variable type to our pronouns and traces rules, namely "pro_{p_i}":

(189) Traces and pronouns rule (modified from Heim and Kratzer 1998)
If α is a pro_{f_i}, pro_{p_i}, or a trace, g is a variable assignment, $i \in \text{dom}(g)$, then
$[\![\alpha]\!]^g = g(i)$

In order to account for the felicity conditions on the use of pro_f versus pro_p, it is possible to revise Heim and Kratzer's *Appropriateness Condition* on free variables, such that it establishes separate appropriateness conditions on the use of each kind of free variable:

(190) **(Revised) Appropriateness Condition**
A context C is *appropriate* for an LF ϕ only if C determines a variable assignment g_C whose domain includes every index i which has a free occurrence in ϕ, and:
a. for every pro_{f_i}, $g_C(i)$ must be known to all discourse participants; and
b. for every pro_{p_i}, $g_C(i)$ does not have to be known to all discourse participants.

[7] The discussion is embedded more generally under a discussion of "strong contextual felicity" constraints (Tonhauser et al. 2013).

Finally, I propose that whether or not a familiar or private situation pronoun must be used is hardwired in the syntactic selection requirements of different determiners. Specifically, while definite determiners c-select familiar situation variables, indefinite determiners (and potentially other quantificational expressions) c-select private situation variables:

(191) a. $\text{D}_{\text{DEF}} \rightarrow pro_f$
b. $\text{D}_{\text{INDF}} \rightarrow pro_p$

To facilitate the reading of entries in the coming discussion, I will sometimes indicate, on the right-hand side of semantic entries, whether the lexical item c-selects a familiar situation pronoun or private situation pronoun in the syntax. Illustrative examples are provided below for *the* and *a* in English. In the semantics, however, both kinds of determiners select for the same kind of primitive semantic object: a situation variable of type s.

(192) $[\![the]\!] = \lambda s.\lambda P_{\langle e, st\rangle}\colon \exists!x[P(x)(s)].\ \iota x[P(x)(s)]$ $[the \rightarrow pro_f]$

(193) $[\![a]\!] = \lambda s.\lambda P_{\langle e,\langle s,t\rangle\rangle}.\lambda Q_{\langle e,\langle s,t\rangle\rangle}.\lambda s'.\ \exists x[P(x)(s) \wedge Q(x)(s')]$ $[a \rightarrow pro_p]$

In sum, based on the discussion from section 4.3.2, I argued that our semantic model must distinguish between two kinds of situation variables, *familiar* and *private* situation variables. The necessity for "private variables" is further supported by recent works on domain restriction, which reach a similar conclusion (Beaver and von Fintel 2013, 2019; Silk 2016, Arsenijević 2018). With these assumptions in place, we can move to the analysis of indefinite + classifier phrases as singleton indefinites.

4.3.4 Deriving singleton indefinites in Chuj

Having established our assumptions about the existence of "private situations", we can finally turn to the ingredient of central interest in this chapter: noun classifiers. The goal of this section is to show that the data on specific indefinites from section 4.2 can be ex-

plained by maintaining the core semantic proposals from Chapter 3—that noun classifiers encode existence and uniqueness conditions (i.e., "semantic presuppositions"). While the core proposals from Chapter 3 will be maintained, the semantic type of noun classifiers will have to be altered. In section 5.1, I will show how the revised entries can be adapted to account for the necessity of noun classifiers in definite descriptions. Section 5.2 will also present a number of predictions made by the final analysis of noun classifiers.

Building on Elbourne's entry for indefinite quantifiers in situation semantics, I define indefinites as existential quantifiers that select a situation argument as their first argument, as in (194).

(194) Indefinite determiner *jun* (building on Elbourne 2005, 2013)
$[\![jun]\!] = \lambda s.\ \lambda P_{\langle e,\langle s,t\rangle\rangle}.\lambda Q_{\langle e,\langle s,t\rangle\rangle}.\lambda s'.\ \exists x[P(x)(s) \wedge Q(x)(s')]$ $[jun \rightarrow pro_p]$

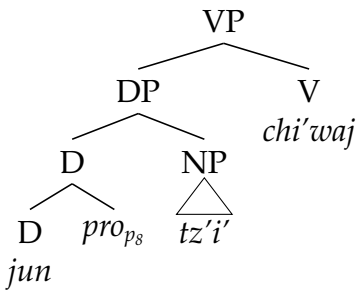
Notice the situation pronoun of the restrictor of the quantifier (the NP) is valued by a different situation pronoun than that of the scope of the quantifier (VP), which will presumably be valued by the topic situation once the remaining situation argument in (195-c) is saturated.

The entry in (194) follows a number of recent works that locates contextual restriction variables in D^0 (see e.g., von Fintel 1994, Büring 2004, Gillon 2006, Schwarz 2009, 2012).[8] Since the indefinite determiner syntactically selects a private situation pronoun as its first argument, the addressee(s) will not take for granted that they are familiar with the value for the situation argument. In a DP without classifiers, combining an NP with an indefinite determiner will simply yield a generalized quantifier whose domain is established by the value for the private situation. The syntax and semantics of a relevant indefinite DP is schematized below (see §2.4.2 above for assumptions about how verb-initial word order would be derived in (195)):

[8] Etxeberria and Giannakidou (2019) provide a similar view, but argue that weak quantifiers, including indefinite determiners, do not denote existential quantifiers.

(195) a. *Ixchi'waj jun tz'i'.*
Ix-chi'waj jun tz'i'.
PFV-bark INDF dog
'A dog barked.'

b.



c. $[\![\text{VP}]\!]^g = \lambda s'.\ \exists x[\text{DOG}(x)([\![pro_{p_8}]\!]^g) \wedge \text{BARK}(x)(s')]$

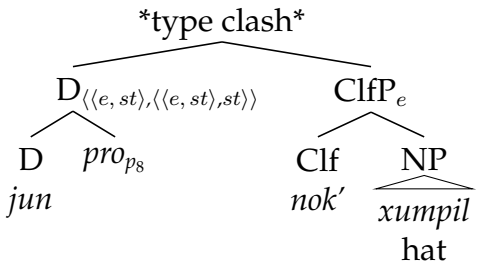
As per the *Revised Appropriateness Condition* in (190), notice that when uttering (199-a), it is not necessary for the addressee(s) to know the value assigned to the situation pronoun of the indefinite. One could felicitously utter this sentence with a particular situation with several dogs in mind even if the addressee(s) are unable to determine the value for the situation pronoun.

Turning to noun classifiers, recall the following entry from Chapter 2:

(196) Denotation of noun classifiers (repeated from (79))
$[\![\text{CLF}]\!] = \lambda s.\lambda P_{\langle e,\langle s,t\rangle\rangle}\colon |\{y\colon P(y)(s)\}| = 1.\ \iota x[P(x)(s)]$

In (196), the classifier first takes a situation argument, returns a function which takes an NP, and gives back the unique entity that satisfies the NP in a relevant situation. As already foreshadowed at the beginning of this chapter, this entry is compositionally incompatible with the entry provided for indefinites. Combining the classifier with an existential quantifier would lead to a semantic type clash:

(197) Indefinites incompatible with (196)



Therefore, if we want to keep to a quantifier analysis of indefinites, as singleton approaches to specific indefinites have it (Schwarzschild 2002), then we must modify the entry of the classifier in such a way as to allow it to compose with indefinites. I propose the final entry in (198), which keeps to a uniqueness-based approach to noun classifiers. Again, for simplicity, I leave the "classifying" presupposition out of the denotation of the noun classifier, but see (281) below for a full-fledged entry of the noun classifier for animal entities with its class presupposition.

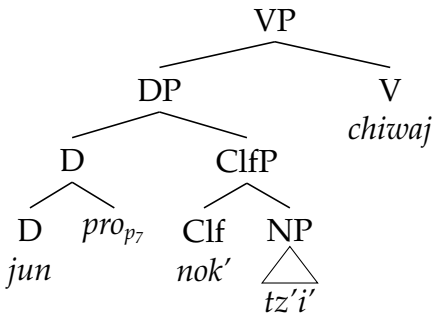
(198) Classifier denotation (modified, and final)
$[\![\text{CLF}]\!] = \lambda P_{\langle e,\langle s,t\rangle\rangle}.\lambda x.\ \lambda s\colon |\{y\colon P(y)(s)\}| = 1.\ P(x)(s)$

The noun classifier in (198) now takes a property and returns the same property with the additional presupposition that there is a unique satisfier of the property in the situation with which the DP gets interpreted. Also notice that the entry in (198) is reminiscent of the predicative entry for *the* proposed by Coppock and Beaver (2015), discussed in Chapter 3, except that the noun classifier still invokes existence and uniqueness in its non-assertive component.

There is a critical distinction between the new entry in (198) and the old entry in (196): the classifier no longer takes a situation variable as its first argument. Instead, its situation argument will be valued upon merger of a determiner. While we return to how the situation pronoun gets valued in the case of definite descriptions in section 5.1, the crucial point here is that when the classifier combines with an indefinite determiner, its situation argument will get valued by the *private situation pronoun* selected by that indefinite:

(199) a. *Ixchi'waj jun nok' tz'i'.*
Ix-chi'waj jun **nok'** tz'i'.
PFV-bark INDF CLF dog
'A dog barked.'

b.



c. $[\![\text{DP}]\!]^g = \lambda Q_{\langle e,\langle s,t\rangle\rangle}.\lambda s'\colon |\{y\colon \text{DOG}(y)(\underline{[\![pro_{p_7}]\!]^g})\}| = 1.\ \exists x[\text{DOG}(x)(\underline{[\![pro_{p_7}]\!]^g}) \wedge Q(x)(s')]$

d. $[\![\text{VP}]\!]^g = \lambda s'\colon |\{y\colon \text{DOG}(y)(\underline{[\![pro_{p_7}]\!]^g})\}| = 1.\ \exists x[\text{DOG}(x)(\underline{[\![pro_{p_7}]\!]^g}) \wedge \text{BARK}(x)(s')]$

The result in (199-c) is a DP that denotes a singleton indefinite. That is, the classifier encodes a presupposition that there be a unique satisfier of DOG in a situation that is private to the speaker. This situation could be restricted to the point where there is only one dog, even though there is more than one salient dog in the more salient situation shared by the discourse participants.

In other words, we now have a solution to the too-many-determiner problem identified at the beginning of this chapter: noun classifiers can co-occur with indefinite determiners because (i) they are instantiated below D^0 in the extended nominal domain, and (ii) their composition with an NP as identity functions over properties yields a semantic object which can be fed to a quantifier. We can also now explain why the uniqueness condition of noun classifiers does not seem to behave as a run-of-the-mill uniqueness presupposition in combination with indefinite determiners: if it is taken for granted by the discourse participants that the situation variable's value is possibly unknown to the addressee(s), then they cannot conclude that the Common Ground entails that there is only one satisfier of the NP. Doing so would require them to agree on the value for the situation, which is not the case with private situation pronouns. Hence, this theory implies

that the non-assertive component in the denotation of noun classifiers does not always surface as a pragmatic presupposition, a topic I discuss in further detail in section 5.1.2.2 below.

Repeating a relevant example from above, the DP in (200) will end up denoting, roughly, the unique priest that the speaker has in mind without necessarily presupposing that there is a unique priest in the context.

(200) **Context**: There are five priests in Yuxquen and the speaker and addressee know it. The speaker and addressee live and currently are in Yuxquen. Out of the blue, the speaker says:

Ix-in-lolon yet' jun winh pale ewi.
Hin-lolon yet' [#(jun) **winh** pale] ewi.
B1S-speak with INDF CLF priest yesterday
'I spoke with a priest yesterday.'

(201) $[\![(200)]\!]^g = \lambda s\colon |\{y\colon \text{DOG}(y)(\underline{[\![pro_{p_7}]\!]^g})\}| = 1.\ \exists x[\text{PRIEST}(x)(\underline{[\![pro_{p_7}]\!]^g}) \wedge \text{SPOKE.WITH}(x)(\text{SPEAKER})(s)]$

The value assigned to pro_{p_7} could simply be the minimal situation involving the speaker's conversation with the priest, one which is only epistemically accessible to the speaker. Such a situation may well have contained only one priest, and there is no reason why the value for this situation would need to be identifiable to the addressee(s). Also notice that the new entry for noun classifiers still involves an existence condition: if the set of priests in a given situation has a cardinality of one, then the existence of priests in that situation follows. However, just like the uniqueness condition, the existence condition will not behave as a classic presupposition when combined with a private situation pronoun (again, see §5.1.2.2 for further discussion). We therefore successfully derive the specific indefinite observations made in section 4.2, all while maintaining the proposal from Chapter 3 that the non-assertive component of noun classifiers involves uniqueness and existence conditions on the extension of the NP.

4.3.5 Summary and interim discussion

In this chapter, I proposed that DPs containing noun classifiers restrict the domain of a quantifier to a singleton set. The proposal builds on independent ingredients from the literature, including domain restriction-based approaches to the exceptional scope of indefinites (von Fintel 1999; Schwarzschild 2002; Breheny 2003), as well as approaches that assume that domain restriction in the nominal domain is achieved via syntactically-represented contextual variables located within complex DP heads (e.g., von Fintel 1994, Büring 2004, Schwarz 2009, 2012). A leading empirical observation in deriving the contribution of noun classifiers in indefinite expressions has been Schwarzschild's (2002) Privacy Principle, which implies that the contextual domain of certain quantificational expressions can be privately restricted. This led us to the proposal that indefinite determiners combine with "private" situation variables. When such variables saturate the situation argument of the noun classifier, I argued that the result is a singleton indefinite, thereby deriving the specific indefinite facts observed in section 4.2. We therefore have a solution to the "too-many-determiner" puzzle introduced at the outset of this chapter. While, at first glance, classifiers appear to trigger completely disparate presuppositions when they appear with definite and indefinite expressions, I proposed that the overall non-assertive contribution of noun classifiers can be unified, if indefinites introduce different requirements on the value for the situation argument.

This chapter almost exclusively focused on co-occurrences of indefinite determiners with noun classifiers. However, we also saw that indefinite DPs without classifiers can give rise to specific interpretations; see for instance (166) above. A rationale for the classifier's optionality with singleton indefinites, but not in definite descriptions, will be provided in section 5.1.2. However, I note here that a singleton approach to indefinites is compatible with indefinites receiving specific interpretations, even *without a classifier*. In principle, the situation pronoun that combines with the indefinite determiner could pick

out a situation that contains only one satisfier of its restrictor NP, and therefore a specific indefinite interpretation would arise. Though the classifier introduces a presupposition that *forces* a singleton interpretation, such an interpretation remains *available* even without the classifier. This aligns with data from other languages, such as English, where Schwarzschild (2002) argues that an indefinite determiner can be restricted to a singleton set without an overt cue (like *a* + *certain* or *a* + relative clause).

While the current proposal builds on the proposal in Chapter 3 in encoding a non-assertive uniqueness condition within the denotation of the noun classifier, the new denotation for noun classifiers has been refined in important ways from the one put forth in Chapter 3. Namely, the noun classifier now denotes an identity function over properties and thereby no longer takes a situation variable as its first argument. In the next chapter, we return to definite descriptions. I propose that the same noun classifier denotation can be extended to explain the necessity of noun classifiers in weak definite descriptions, so long as (i) building on Coppock and Beaver's (2015) analysis of definite descriptions in English, we further decompose weak definite descriptions into two morphological pieces, and (ii) the head of DP in weak definite descriptions must combine with a *familiar situation pronoun*. I will also show that minimal modifications may also be made to the denotation of deictic particles, such that we may account for their ability to co-occur with indefinite determiners in Chuj.

Before moving on, however, I provide a brief discussion of an open issue having to do with the availability of binding the situation variable of indefinite phrases (§4.3.5.1). I also briefly discuss choice-functional approaches to wide scope interpretations of indefinites, arguing that a singleton indefinite account of specific indefinites with noun classifiers in Chuj is to be favoured (§4.3.5.2).

4.3.5.1 An open issue: Binding and pro_p

In this section, I have proposed, following previous work on quantifier domain restriction (Westerståhl 1984; von Fintel 1994; Büring 2004; Schwarz 2009), that indefinite determiners combine with a contextual variable. This variable serves to delimit the quantificational domain of the determiner. One prediction the current analysis makes is that the situation argument of the indefinite determiner, much like the one of the definite determiner, should sometimes show signs of being bound by a higher quantifier over situations.

First recall the example from Chapter 3, which involved covariation of a weak definite determiner. The relevant example is repeated below for convenience.

(202) *Juntakel ixek' waj Xun t'a jun chonhab', ixlolon winh yet' winh alkal.*
Juntakel ix-ek' waj Xun t'a jun chonhab' ix-lolon winh yet'
every.time PFV-pass CLF Xun PREP INDF town, PFV-talk CLF with
[#**(winh)** alkal].
CLF mayor
'Whenever Xun visited a town, he spoke with the mayor.'

In (202), the definite description *winh alkal* 'the mayor' covaries with respect to each town *Xun* visited. The proposal in Schwarz 2009 and Chapter 3 is that in such cases, the situation pronoun of the definite article is bound by the situation argument of the quantifier *juntakel* 'every time', such that the uniqueness presupposition is relativized to the situation variable that the universal quantifies over. As mentioned above, this yields an interpretation paraphrasable as "in every situation s, Matin spoke with the unique mayor in s".

Since, under the current system, indefinite determiners also combine with situation variables, we predict that it should be possible for the situation variable of the indefinite determiner to also sometimes be bound by a quantifier over situations. This prediction is borne out: comparable covarying examples involving an indefinite + classifier DP can be shown to arise in Chuj. Consider the example in (203) with its context:

(203) **Context:** Xun is a filmmaker, who is interested in interviewing the eldest man of every town he visits. To describe this, I say:

Juntakel ixek' waj Xun t'a jun chonhab', ixlolon winh yet' jun winh ichamwinak.

Juntakel ix-ek' waj Xun t'a jun chonhab', ix-lolon winh yet'
ever.time PFV-pass CLF Xun PREP INDF town, PFV-speak CLF.PRON with
[jun **winh** ichamwinak].
INDF CLF elder.man

'Whenever Xun visited a town, he spoke with a (certain) elder.'

There is a sense in which the indefinite expression *jun winh ichamwinak* 'an elder' in the above sentence has a specific interpretation. The context leads to an interpretation in which in every town Xun visited, he spoke with a *certain* elder of that town, namely the eldest elder. Therefore, the uniqueness condition of the noun classifier seems to be, yet again, relativized to the situation variable that *juntakel* 'every time' quantifies over. The sentence essentially conveys that "in every situation s, Xun spoke with an elder in s, where (i) s contains only one elder (the uniqueness condition), and (ii) s is not necessarily known to all discourse participants". While I leave a more detailed study of the conditions on bound indefinite expressions for future work, the important point is that the situation variable of indefinite expressions seems to be bindable, as predicted by a variable-based approach to domain restriction.

4.3.5.2 On choice-functional approaches to wide scope indefinites

Since noun classifiers force specific interpretations of indefinites, a natural alternative to the analysis proposed here could have been to consider whether noun classifiers denote choice function variables, an approach that has been adopted by many authors to account for the exceptional scope of indefinites (see e.g. Reinhart 1997; Winter 1997; Kratzer 1998; Matthewson 1999). There are two major views on the choice function analysis of indefinites. One line of account, proposed by Reinhart (1997) and Winter (1997), argues that indefinite determiners may always denote existentially bound choice functions, regardless of the narrow, intermediate, or wide scope reading obtained. This is done by

allowing for the scope of the indefinite to be determined by the site of existential closure over the choice function, which can occur at any point in the structure. A second line of account is proposed by Kratzer (1998) and Matthewson (1999), who argue that indefinite expressions can be both choice-functional or quantificational (Kratzer (1998) argues that in English, they are lexically ambiguous). However, unlike in Reinhart 1997 and Winter 1997, these authors propose that choice functional indefinites ultimately only give rise to widest-scope interpretations.[9]

Extending either of these two analyses to capture the contribution of noun classifiers, however, faces a number of non-trivial complications (in addition to some of the more general conceptual and empirical costs of choice-functional approaches to indefinites; see e.g., Schwarz 2001, 2004, 2011, Endriss 2009, and Heim 2011).

A first general issue is that noun classifiers are clearly not the indefinite determiner proper, but occur morphosyntactically separately from indefinite determiners like *jun* (which without a noun classifier can take both narrow or wide-scope). In choice-functional approaches to indefinites, it is the indefinite determiner that denotes a choice-function. If a choice-functional analysis of noun classifiers were pursued, we would require the stipulation that there are two separate choice-functional expressions or that a choice-functional determiner can appear within the restrictor of a separate quantificational determiner.[10]

A second issue concerns the particular view of choice functions put forth by Reinhart (1997) and Winter (1997), who both argue that existential closure over choice functions can occur *at any site* in a clause. If the classifier were to denote a choice function variable,

[9]By assuming that choice-functional indefinites are not existentially quantified (Kratzer 1998) or that they may only be existentially quantified at the highest level (Matthewson 1999), intermediate readings are not immediately derived in Kratzer 1998 and Matthewson 1999. Therefore, much in the same way as in Schwarzschild 2002, these authors propose that the (pseudo)-intermediate scope effect of indefinites is due to the presence of a (potentially implicit) bound variable.

[10]One could keep to a choice functional analysis of the indefinite determiner, all while proposing that the choice function combines with a predicative ClfP, one which introduces a uniqueness condition. This kind of analysis would be analogous to the analysis of noun classifiers proposed in this thesis, with the added stipulation that existential quantification may be be severed from the semantics of the indefinite determiner. Under this kind of analysis, there would effectively be two ways of deriving a specific indefinite: (i) by having existential closure scope wide or (ii) by restricting the domain of quantification to a singleton set via the noun classifier.

then these accounts would predict that such indefinites should be able to receive narrow scope interpretations. As discussed in section 4.2, this does not seem to be the case.

Finally, a third issue is that an analysis of classifiers as (bound) choice function variables would not straightforwardly extend to the other non-indefinite environments in which noun classifiers occur (obligatorily), as in cases where noun classifiers appear in weak definite descriptions or surface alone as pronouns. It would also not account for the interpretative differences described in the different constructions in which classifiers arise. For instance, if the classifier is just an existentially bound choice function variable, no uniqueness presupposition will be expected when noun classifiers appear without the indefinite determiner *jun*.

In sum, I conclude that attempting to account for the specific indefinite facts observed in Chuj by appealing to a choice function analysis of noun classifiers would be inadequate. Of course, this does not preclude that so-called wide scope indefinites may be derived by assigning *indefinite determiners* a choice functional analysis (see e.g., Matthewson 1999 on the possibility of overt choice functions in St'át'imcets; or footnote 10 on the previous page). However, if the preceding argumentation is correct, the Chuj data suggest that recourse to choice functions cannot be the sole method available to derive the apparent exceptional scope of indefinites.

Chapter 5

A unified theory of the syntax and semantics of Chuj DPs

Having provided refined entries for noun classifiers to account for their occurrence in indefinite expressions in Chapter 4, this chapter revisits the analysis of definite descriptions from Chapter 3. I show that it is possible to extend the previous chapter's proposal on noun classifiers in order to explain the necessity of noun classifiers in definite descriptions if weak definite descriptions are further decomposed into two pieces: (i) noun classifiers, which contribute a uniqueness and existence conditions, and (ii) ι, which heads the DP and is also the locus of a *familiar* situation pronoun. I also provide a refined entry for deictic particles, which accounts for their appearance in both definite and indefinite expressions.

The chapter is structured as follows. In section 5.1, we revisit the analysis of weak definite descriptions. In section 5.2, I then provide a discussion surrounding the resulting analysis of noun classifiers on three levels. First, I show that the final analysis of noun classifiers in Chuj, both within definite and indefinite expressions, makes a number of predictions, which I show are borne out. Second, I provide a discussion of apparent sources of linguistic variation in the denotation of noun classifiers within the Mayan

language family, and argue that this variation could potentially be straightforwardly captured if one adopts the denotation for noun classifiers proposed in this work. And third, I show that the analysis of definites and specific indefinites proposed here has the potential to extend to similar phenomena described in languages of three distinct language families. Finally, section 5.3 returns to the contribution of deictic particles, which following observations from Chapters 2 and 3, I argue are compatible with both definite and indefinite bases alike. I sketch a compositional analysis of deictic particles, showing that it is possible to revise the entries from Chapter 3 in order to account for the fact that they may combine with both kinds of bases. The resulting composition allows to maintain the core proposal that deictic particles introduce indices, which in the case of definite demonstratives result in a strong definite description.

5.1 Definites revisited

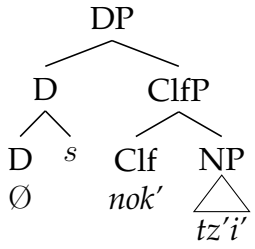
Now that we have an analysis of specific indefinites in classifier+indefinite constructions in Chuj, we must revisit the contribution of classifiers in (weak) definite descriptions. That is, we must understand (i) why we obtain a classic weak definite uniqueness presupposition when a classifier appears in a definite description, and (ii) why classifiers are *required* for definite descriptions (and third person pronouns), but not with indefinite expressions. Section 5.1.1 addresses point (i), while section 5.1.2 addresses point (ii).

5.1.1 Decomposing weak definites further

In the previous chapter, I proposed that noun classifiers are instantiated below DP in a projection labeled ClfP. I further proposed that situation pronouns in the DP are merged as part of a complex D^0 (von Fintel 1994, Büring 2004, Schwarz 2012). In a weak definite description, which requires the combination of a bare noun classifier with an NP, this implies that the head of the DP will be phonologically vacuous:

(204) a. *nok' tz'i'*
CLF dog
'the$_{weak}$ dog'

b.



To account for classifier DPs in Chuj, I propose that such DPs are headed by a null expression, ι, with the following crucial assumption: ι obligatorily selects a *familiar situation pronoun*. Since the noun classifiers' situation argument is only valued upon merger of D^0, this familiar situation pronoun is ultimately fed as the value for the situation argument of the noun classifier. The noun classifier will subsequently impose a uniqueness presupposition relative to a familiar situation, which I argue effectively derives a weak definite description.

The entry for ι is provided in (205), adapted from previous work (see e.g. Arkoh and Matthewson 2013, Coppock and Beaver 2015, Jenks 2018). It first combines with a situation variable, then takes a property, and returns the unique satisfier of that property in the situation. As in previous work, ι also triggers a uniqueness and existence presupposition on the extension of the NP. While I write this in the entry in (205), I will sometimes omit this presupposition in derivations for brevity. Critically, as indicated on the right of the denotation in (205), I assume that like other definite DP heads, ι specifically selects a *familiar* situation variable in the syntax.[1]

(205) $[\![\iota]\!]^g = \lambda s.\lambda P_{\langle e,\langle s,t\rangle\rangle} : \exists!y[P(y)(s)].\ \iota x[P(x)(s)]$ $[\iota \rightarrow pro_f]$

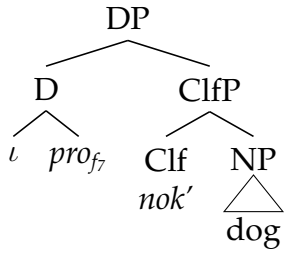
[1]As already indicated in Chapter 3, I assume that the uniqueness presupposition is just a sub-case of a more general maximality presupposition on definite descriptions (see e.g., Sharvy 1980, Link 1983), and leave aside the discussion of plural definite descriptions.

Now, recall that noun classifiers impose the condition that there be a unique satisfier of the NP relative to the situation argument, which is only valued once D^0 is merged. The final entry for noun classifiers proposed in Chapter 4 is repeated below for convenience.

(206) Classifier denotation (final)
$[\![\text{CLF}]\!] = \lambda P_{\langle e,\langle s,t\rangle\rangle}.\lambda x.\ \lambda s\text{: } |\{y\text{: } P(y)(s)\}| = 1.\ P(x)(s)$

Since ι selects a familiar situation pronoun as its first argument, the proposition that there is a unique satisfier of the NP in the situation picked up by the familiar situation pronoun should be part of the Common Ground, and will therefore emerge as a classic pragmatic presupposition (note that here *familiar situations* are contrasted with *familiar referents*, which I assume are signalled via the presence of a deictic particle; see section 5.3 below for further discussion). That is, if the uniqueness condition is imposed on a situation that is assumed to be known to all discourse participants, then it follows that there will only be one satisfier of the relevant NP in the situation. This is schematized below:

(207) a.



b. $[\![\text{ClfP}]\!]^g = \lambda x.\ \lambda s\text{: } |\{y\text{: } \text{DOG}(y)(s)\}| = 1.\ \text{DOG}(x)(s)$
c. $[\![\text{DP}]\!]^g =$
Presupposition: $|y\text{: } \text{DOG}(y)([\![pro_{f_7}]\!]^g)| = 1$
Assertion: $\iota x[\text{DOG}(x)([\![pro_{f_7}]\!]^g)]$

For completeness, notice that the new analysis now explains why, in the minimal pair of (200) from Chapter 4 in (208) without an indefinite determiner, the resulting interpretation is one in which it is presupposed that there is a unique priest in the relevant situation that the discourse participants agree on, and therefore that a run-of-the-mill uniqueness presupposition arises:

(208) **Context**: There's only one priest in Huxk'e'en and the speaker and addressee know it. The speaker and addressee live in Huxk'e'en.
Ixinlolon yet' winh pale.
Ix-in-lolon yet' [#(**winh**) pale].
PFV-A1S-speak with CLF priest
'I spoke with the priest.'

(209) $[\![(208)]\!]^g = \lambda s\colon |\{y\colon \textsc{priest}(y)(s_f)\}| = 1.\ \textsc{spoke.with}(\iota x[\textsc{priest}(x)(\underline{[\![pro_{f_7}]\!]^g})])(\textsc{speaker})(s)$

Under this analysis, weak definite determiners are thus further decomposed in two parts. While traditional Fragean approaches to the definite determiner in Situation Semantics (e.g. Schwarz 2009, Jenks 2018) package (i) domain restriction and (ii) a uniqueness presupposition in the lexical semantics of the definite determiner, the current account severs ingredients (i) from (ii). The reason for this severance arose as a response to the "too-many-determiner" puzzle in section 4.1, which led us to conjecture that the uniqueness and existence conditions of the noun classifier could be introduced lower than D^0 in the extended nominal structure. That is, the fact that the classifier can co-occur with a separate quantifier seems to suggest that the classifier itself is not instantiated in the head of DP, making it natural to analyze weak definites with an external DP head. This is not to say that the definite determiner in English does not hardwire both ingredients in the same lexical item. In fact, perhaps it may, given the impossibility of expressions such as **a the book* (though see Coppock and Beaver 2015 for a similar decomposition of the definite article in English). What the current proposal shows instead is that there may be more than one path to the composition of weak definiteness from one language to another, and that this difference might sometimes correlate with the ability for one language to identify weak definites and specific indefinites with overlapping morphology.

On this note, it is also worth noting that the current system resembles the recent analysis of definiteness provided by Coppock and Beaver (2015) in a number of interesting ways. As discussed in Chapter 3, Coppock and Beaver (2015) argue that definite descriptions are decomposed in two pieces: (i) an ι type-shifter, which triggers a uniqueness and

existence presupposition and (ii) an entry for *the* of type $\langle et,et \rangle$, which introduces a weak uniqueness presupposition (without an existence presupposition). An entry equivalent to theirs in the current notation is repeated below from (120):

(210) $[\![\textit{the}]\!] = \lambda P\colon |P| \leq 1.\ \lambda x.\ P(x)$ (Coppock and Beaver 2015)

The differences between the current analysis and the one proposed by Coppock and Beaver (2015) are (i) the assumption that the classifier and ι involve a situation argument, responsible for domain restriction (following Schwarz 2009 and others) and (ii) the assumption that the noun classifier also presupposes existence, contrary to the entry in (210), which does not require existence (*the*(P) is defined even when $|P| = 0$). In section 5.2.2, I suggest that an entry along the lines of (210) may be appropriate for other Q'anjob'alan languages, providing further reason to believe that there may be more than one compositional path to weak definiteness across languages.

5.1.2 On the necessity of classifiers with definites

While the analysis just presented explains why a uniqueness presupposition arises when the classifier appears alone with a noun in Chuj, we must still understand why noun classifiers are *obligatory* in such constructions. That is, recall that while noun classifiers are not necessary to yield specific indefinites, they are obligatory when occurring in weak definite phrases. A fresh contrast is provided below for illustration.

(211) **Context**: There's a book, namely *The Little Prince*, that you want to buy. You tell your friend:
Hingana tzinman jun ch'anh libro.
Hin-gana tz-in-man [jun **(ch'anh)** libro].
A1S-desire IPFV-A1S-buy INDF CLF book
'I want to buy a book.'

(212) **Context**: You already told your friend a few days ago that there's a particular book that you find interesting. You tell your friend:

Hingana tzinman ch'anh libro.

Hin-gana tz-in-man [ɩ #**(ch'anh)** libro].

A1S-desire IPFV-A1S-buy CLF book

'I want to buy the book.'

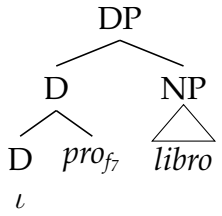
Given the current setup, it is unclear why ɩ could not alternatively be used alone in (212). Compositionally, there is no reason why it should not be able to do so. The result would yield the unique entity that satisfies the nominal property in the situation taken as a first argument by ɩ:

(213) a. # ɩ libro

book

Intended: 'the book'

b.



c. $\iota x[\text{BOOK}(x)([\![pro_{f_7}]\!]^g)]$

Also important to mention is the fact that bare nouns are not generally banned in Chuj. As shown in (214), bare nouns can serve as arguments to yield bare indefinite plurals (see discussion in section 2.5.2):

(214) *Ixinman onh.*

Ix-in-man **onh**.

PFV-A1S-buy avocado

'I bought avocados.'

Therefore, the current analysis does not provide a clear explanation for the necessity of noun classifiers in definite descriptions. In other words, we need to account for the following generalization:

(215) **Generalization about noun classifiers and (in)definiteness**
Noun classifiers are obligatory in definite descriptions, but optional in specific indefinite descriptions.

In the rest of this subsection, I provide two possible ways of understanding the necessity of noun classifiers with definite descriptions. One explanation, discussed in section 5.1.2.1, is based on empirical observations about the placement of noun classifiers in definite DPs containing numerals. The second, discussed in section 5.1.2.2, is based on pragmatic reasoning, and the consequences of the existence of private situation pronouns for theories of Maximize Presupposition. Note that the two explanations are not incompatible with each other, meaning that both could conspire to explain (215).

5.1.2.1 Syntactic relationship between ClfP and DP

One potential source of insight into the need for noun classifiers in definite descriptions comes from an empirical observation about the possible syntactic positions of noun classifiers with DPs that contain *numerals*. The pattern of interest is the following: while noun classifiers typically follow numerals, they can sometimes appear before them. Crucially, however, this displacement can only happen when the DP is interpreted as definite. The two possible DP configurations are summarized below:

(216) Classifier displacement interacts with semantic interpretation

a. No classifier displacement → definite & indefinite interpretations possible
$[_{\text{DP}}\ \emptyset\ [_{\text{NumP}}\ \text{NUMERAL}\ [_{\text{ClfP}}\ \text{CLF}\ [_{\text{NP}}\ \text{NOUN}\ ]]]]$

b. Classifier displacement → only definite interpretation possible
$[_{\text{DP}}\ \text{CLF}\ [_{\text{NumP}}\ \text{NUMERAL}\ [_{\text{ClfP}}\ t_i\ [_{\text{NP}}\ \text{NOUN}\ ]]]]$

Q'anjob'alan languages vary in the extent of constructions in which classifier movement is possible. In Q'anjob'al, for instance, classifier movement seems to be in general available: noun classifiers can appear in two positions within the complex DP, after numerals and

before numerals. While (217-a) is underspecified for definiteness (compatible with both definite and indefinite interpretations), (217-b) only gives rise to a definite interpretation (Pedro Mateo Pedro, p.c.; see also Mateo Toledo 2017). Also note that noun classifiers cannot appear twice within the same extended nominal projection (217-c).

(217) Q'anjob'al

a. *Xchiwaj kak'on no' tx'i'.*
X-chiwaj [ka-k'on **no'** tx'i'].
PFV-bark two-NUM.CLF CLF dog
'Two dogs barked / the two dogs barked.'

b. *Xchiwaj no' kak'on tx'i'.*
X-chiwaj [**no'** ka-k'on tx'i'].
PFV-bark CLF two-NUM.CLF dog
'The two dogs barked / not: 'Two dogs barked.'

c. *Xchiwaj no' kak'on no' tx'i'.*
*X-chiwaj [**no'** ka-k'on **no'** tx'i'].
PFV-bark CLF two-NUM.CLF CLF dog

In the San Mateo Ixtatán dialect of Chuj under study in this work, classifier displacement is also possible in partitive constructions of the type "of the NUM *x*, NUM *x*", which likewise force a definite interpretation of the DP.

(218) Chuj from San Mateo Ixtatán

a. *Ha wakwanh nok' tz'i', ixinb'ik' oxwanh nok'.*
Ha [wak-wanh **nok'** tz'i'], ix-in-b'ik' ox-wanh nok'.
EMPH six-NUM.CLF CLF dog PFV-A1S-wash three-NUM.CLF CLF.PRON
'Of (the) six dogs, I washed three.'

b. *Ha nok' wakanh tz'i', ixinb'ik' oxwanh nok'.*
Ha [**nok'** wak-wanh tz'i'], ix-in-b'ik' ox-wanh nok'.
EMPH N.CLF six-NUM.CLF dog PFV-A1S-wash three-NUM.CLF CLF.PRON
'Of the six dogs, I washed three.'

While the Chuj consultants of the San Mateo Ixtatán dialect do not seem to admit classifier displacement within canonical arguments, as in the Q'anjob'al data in (217), García Pablo and Domingo Pascual (2007) document this possibility in the dialect of San Sebastián Coatán:

(219) Chuj from San Sebastián Coatán (García Pablo and Domingo Pascual 2007: 211)

Ixway nok' chawanh kalnel tz'on.

Ix-way [**nok'** cha-wanh kalnel tz'on].
PFV-sleep CLF two-NUM.CLF sheep skinny
'The two skinny sheep slept.'

As shown in (219), the noun classifier can appear before the numeral, resulting in a definite translation (as provided by the authors). Note that the noun classifier in San Sebastián Coatán Chuj can also appear after numerals, just like in San Mateo Chuj and Q'anjob'al. García Pablo and Domingo Pascual (2007) provide the following examples:

(220) Chuj from San Sebastián Coatán (García Pablo and Domingo Pascual 2007: 159)

a. *oxwanh nok' tz'i'.*

ox-wanh **nok'** tz'i'
three-NUM.CLF CLF dog
'three dogs'

b. *hukwanh nok' okes.*

huk-wanh **nok'** okes.
seven-NUM.CLF CLF coyote
'seven coyotes'

Importantly, neither Q'anjob'al nor the two dialects of Chuj seem to allow classifier displacement over the indefinite determiner, also the numeral 'one'. A Q'anjob'al example is shown below:

(221) Q'anjob'al

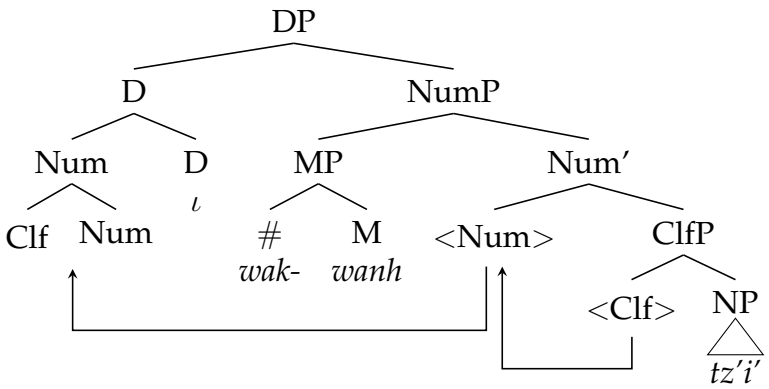
Xchiwaj no' jun tx'i'.

*X-chiwaj [**no'** jun tx'i'].
PFV-bark CLF one dog
Intended: 'The one dog barked.'

Though more work is needed to understand the semantic contribution of classifier displacement within the extended DP, and why certain dialects of Chuj restrict it more than others, these data indicate a tight connection between the head of DP and noun classifiers, *specifically when the DP is interpreted as a definite description*. A potential avenue

to explain the necessity of noun classifiers in definite descriptions could thus be syntactic: perhaps, the heads of definite DPs must enter in a syntactic relationship with the heads of ClfPs, which sometimes results in movement of the classifier to the head of DP. A structure for *nok' wakwanh tz'i'* showing this possibility is provided in (222). Following Bale and Coon (2014), Bale et al. (2019), and Little et al. (2022a), I assume that the numeral and numeral classifier constituent are instantiated in a measure phrase in the specifier of NumP. To derive the order in (218-b), the classifier undergoes successive head movement to the head of the DP (which I assume reconstructs at LF).

(222) Classifier movement in Chuj



This kind of analysis could potentially explain why classifier displacement is not observed in indefinite DPs. If there is a systematic syntactic relationship between definite DPs and ClfPs, but not between indefinite DPs and ClfPs (see the absence of displacement with *jun* in (221) above), then classifier movement should only be observed in the case of definite DPs. We might also expect this syntactic relationship to *always* be necessary, regardless of whether displacement is observed on the surface. If this is the case, then the necessity of noun classifiers in definite DPs could result from the necessity of D^0 to enter in a syntactic relationship with Clf^0.

In sum, while the nature of potential syntactic relationships between definite D^0 and Clf^0, as well as the kind of head movement involved in the above examples, remains to be understood, classifier movement within the extended DP provides us with a potential

line of analysis to explain the necessity of noun classifiers within definite DPs in Chuj and other Q'anjob'alan languages.[2]

5.1.2.2 *Maximize Presupposition* and private situations

A second compatible line of account to explain the necessity of noun classifiers in definite phrases lies in pragmatic reasoning, and more specifically, in how we conceive of effects that fall under the purview of *Maximize Presupposition* (Heim 1991; Percus 2006; Chemla 2008; Sauerland 2008; Singh 2011; Schlenker 2012). Though more work is needed to flesh out the effects of admitting private situations for a theory of pragmatics, the guiding intuition behind this solution would be the following. If a speaker assumes that their addressee is not in a position to identify the content of a semantic presupposition, because this semantic presupposition involves a *private situation*, then the proposition related to this semantic presupposition should not be part of the Common Ground. And since the principle of *Maximize Presupposition* is geared towards the Common Ground, then presuppositions based on private situations should not be targeted by this principle.

The principle of *Maximize Presupposition* was originally put forward by Heim (1991) to make sense of contrasts like the following:

(223) a. The sun is shining.
b. #A sun is shining.

(224) Common Ground: $|\{x\text{: }\textsc{student}(x)\}| = 2$
a. Both students passed the exam.
b. #All students passed the exam.

As Heim (1991) and subsequent work argued, a crucial component for understanding the contrasts in (223) and (224) lies in the fact that both sentences in each example are

[2]This kind of head movement would fall under the purview of *syntactic* head movement in the sense of Harizanov and Gribanova 2019. That is, classifier movement (i) does not form new words, (ii) may not necessarily obey the head movement constraint (depending on the analysis of the extended nominal domain) and (iii) can have interpretive effects. In this case, it would give rise to obligatory definite interpretations. However, as pointed out by Harizanov and Gribanova, the interpretive effects may also result from the head that causes the movement in the first place (a definite D^0, in this case).

"contextually equivalent". The idea is that the set of sentences in each example are semantically equivalent in the context, but the (a)-sentences are favoured because they presuppose more content than the (b)-sentences. A definition of "contextual equivalence", taken from Singh (2011), is provided below:

(225) Contextual equivalence: (Singh 2011: 151)
LFs ϕ and ψ are contextually equivalent with respect to context c iff
$\{w \in c : [\![\phi]\!](w) = 1\} = \{w \in c : [\![\psi]\!](w) = 1\}$

In simpler terms, two propositions are contextually equivalent whenever their assertive content is identical relative to a context, generally understood as the Common Ground (Schlenker 2012), where the Common Ground involves the set of propositions that the speaker and addressee mutually agree to be true.

A standard implementation of *Maximize Presupposition* takes this principle to be evaluated at the level of the entire sentence (Heim 1991; Sauerland 2008; Schlenker 2012). Under this view, *Maximize Presupposition* compares entire propositions that are contextually equivalent, and favours those that presuppose the most content. To illustrate how this all works together, consider again the contextually-equivalent sentences in (223). Since the Common Ground establishes that there is only one sun on Earth, *Maximize Presupposition* predicts that only (223-a) should be felicitous. That is, the definite article competes with the indefinite article for *Maximize Presupposition*, and since the definite article triggers a uniqueness presupposition that is not triggered by the indefinite, (223-a) must be used. Likewise, it is infelicitous to introduce an NP with *all* if the Common Ground entails that there are only two satisfiers of that NP, as shown in (224), despite the fact that a sentence with *all* would convey the same assertive content as the sentence with *both*.

This view, however, cannot be extended to the data at hand. Given the decomposition of weak definiteness provided in this section, a proposition derived from an LF containing a noun classifier presupposes just as much as a proposition without a classifier. This is because ι, as discussed in section 5.1.1, also presupposes uniqueness and existence. A

formulation of *Maximize Presupposition* at the level of propositions can thus not explain the necessity of classifiers within definite DPs.

There are, however, formulations of *Maximize Presupposition* relativized to *lexical alternatives* rather than full propositions (Percus 2006; Coppock and Beaver 2015), which could in fact predict the necessity of noun classifiers in Chuj definite DPs. Percus (2006), for instance, provides the following formulation:

(226) Percus' Maximize Presupposition (Percus 2006)

a. Alternatives are only defined for lexical items. For any lexical item, the alternatives consist of all "presuppositionally stronger" items of the same syntactic category.

b. Do not use ϕ if a member of its Alternative-Family ψ is felicitous and contextually equivalent to ϕ.

Under this definition, *Maximize Presupposition* still technically operates over entire propositions (or LFs), but requires that the LF that contains the presuppositionally strongest *lexical item* among an "Alternative-Family" of LFs be used. The Alternative-Family of an LF can be defined as follows.[3]

(227) **Alternative-Family of an LF** (as defined in Singh 2011: 155)
The *Alternative-Family* of LF ϕ is the set of LFs that can be generated by replacing a lexical item in ϕ with one of its lexical alternatives.

In other words, *Percus' Maximize Presupposition* compares an LF without a presuppositional lexical item with an LF with a presuppositional item (an expression whose entry invokes a presupposition), and states that if these two LFs are contextually-equivalent, the former cannot be used.

As discussed by Percus (2006), this kind of analysis can account for presuppositional competition effects not predicted in more global accounts of *Maximize Presupposition*. Con-

[3]Note that Singh (2011) ultimately does not endorse the theory of Percus (2006). While engaging in this debate falls outside, see Anvari 2018 for further evidence in favour of Percus (2006).

sider, for instance, the sentence in (228) from Anvari (2018), which showcases competition between *all* and *both*.

(228) Example from Anvari 2018: 1
I am critical of {#all, both} of the two mainstream presidential candidates.

As Anvari (2018) notes, accounts of *Maximize Presupposition* based on entire propositions do not predict a competition between *all* and *both* in (228). That is: "[*the two candidates*] already presupposes that there are exactly two candidates, and the same presupposition should, by standard assumptions, be triggered by [*both of the two candidates*] and by [*all of the two candidates*]" (Anvari 2018: 1). On the other hand, if *Maximize Presupposition* is relativized to force LFs with presuppositionally stronger lexical items, as is the case in Percus 2006, then the competition in (228) is predicted. The lexically stronger presuppositional item should be favoured, even if it adds no overall presuppositional content to the overall proposition.

In light of this background, it is interesting to consider two central proposals from this chapter: (i) that there are two kinds of contextual variables—familiar situation pronouns and private situation pronouns—and (ii) that the former are selected by definite determiners (here ι), while the latter are selected by indefinite determiners. This is critical: given that *Maximize Presupposition* depends on the notion of "contextual equivalence" (225), we might expect that the choice of contextual variable will have effects on the application of this pragmatic principle. Here, I would like to propose that this is indeed the case, and that it might serve to explain why noun classifiers are obligatory in definite descriptions, but optional with indefinite expressions.

Consider first what happens when the noun classifier's situation argument is saturated by a familiar situation pronoun, as is the case with definite descriptions. Suppose that noun classifiers compete with a null lexical item heading the ClfP which does not convey uniqueness. If this is the case, Percus' Maximize Presupposition should enforce the use of the classifier in definite descriptions. That is, assuming the lexical alternatives

⟨Ø,CLF⟩, where the classifier is presuppositionally stronger than Ø, we expect that Maximize Presupposition should always favour an LF with a classifier over an LF with Ø.'[4] Building on Coppock and Beaver's (2015) entry for *a* in English, which they likewise assume competes with *the* within a modified version of Percus' Maximize Presupposition, we could define the null competitor as an identity function without presuppositional content, as in (229). Compare this entry with that of the noun classifier (230), repeated from above. The assertive content is the same, the only difference being that noun classifiers encode an additional uniqueness (and existence) condition.

(229) $[\![\text{Ø}]\!] = \lambda P_{\langle e,\langle s,t\rangle\rangle}.\lambda x.\ \lambda s.\ P(x)(s)$

(230) Classifier denotation (repeated from above)
$[\![\text{CLF}]\!] = \lambda P_{\langle e,\langle s,t\rangle\rangle}.\lambda x.\ \lambda s\colon |\{y\colon P(y)(s)\}| = 1.\ P(x)(s)$

When combined with ι, an LF with a noun classifier and an LF with Ø will be contextually equivalent. That is, since ι necessarily combines with a familiar situation pronoun, propositions that differ only with respect to the choice of a classifier over Ø will have the same assertive content relative to the context. Consider, for instance, the following two sentences, with their corresponding LFs (assuming for simplicity existential closure of the situation argument of the verb):

[4] At first glance, this analysis relies on the assumption that a string that is not well-formed (ι+Ø+NOUN) can still be a competitor for *Maximize Presupposition*. However, assuming that a ClfP always projects, strings of the type ι+Ø+NOUN may actually be well-formed with the subset of Chuj nouns which are not classifiable, for instance (see §2.5.1 for discussion of nouns that cannot bear a classifier in Chuj):

(i) *Teb'utan merkado tikneik.*
Te-b'ut-an [ι Ø merkado] tikneik.
INTS-full-STAT market today
'The market is very full today.'

It is also important to note, however, that other presuppositional lexical items have been argued to compete with null morphemes in previous work. This is the case with *too*, which, as Chemla (2008) proposes, competes with a null morpheme (⟨Ø,*too*⟩):

(ii) I had tea and John had tea (too/#Ø). (Chemla 2008: 144)

(231) a. *Ixchi'waj nok tz'i'.*
Ix-chi'waj [nok' tz'i'].
PFV-bark CLF dog
'The dog barked.'

b. LF: [barked [$_{DP}$ ι pro_{f_7} [$_{ClfP}$ **nok'** dog]]]

c. Presupposition = $|y$: DOG$(y)([\![pro_{f_7}]\!]^g)| = 1$

d. Assertion = $\exists s$[PAST$(s) \wedge$ BARK$(\iota x$[DOG$(x)([\![pro_{f_7}]\!]^g)]$

(232) a. # Ix-chi'waj [Ø tz'i'].
PFV-bark Ø dog
Intended: 'The dog barked.'

b. LF: [barked [$_{DP}$ ι pro_f [$_{ClfP}$ **Ø** dog]]]

c. Presupposition = $|y$: DOG$(y)([\![pro_{f_7}]\!]^g)| = 1$

d. Assertion = $\exists s$[PAST$(s) \wedge$ BARK$(\iota x$[DOG$(x)([\![pro_{f_7}]\!]^g)]$

The two sentences in (231-a) and (232-a) have the same assertive content. Both assert that there is a past barking situation s such that the agent of s is the maximal dog in a situation picked up by $g(7)$ familiar to both the speaker and addressee(s). By the definition in (225), the two LFs in (231-b) and (232-b) are also contextually equivalent: all worlds in which (232-a) is true are also going to be worlds in which (234-a) is true. Moreover, as per the definition in (226), the classifier is also a lexical alternative of Ø, and so the LF in (231-b) will be a member of the Alternative-Family of the LF in (232-b). Whenever the presupposition that there exists a unique dog in the situation picked up by $g(7)$ is met (which will always be the case with ι given that ι also presupposes uniqueness), Percus' *Maximize Presupposition* should thus require (231-a) over (232-a), and therefore the presence of the noun classifier should be enforced. Crucially, since $g(7)$ is a familiar situation pronoun, it will be identifiable to both the speaker and their audience.

When the noun classifier combines with an indefinite determiner, on the other hand, recall from section 4.3 that the uniqueness condition will be interpreted relative to a "private" situation. Thus, the semantic presupposition of the noun classifier will merely contribute the information that there is a unique satisfier of the NP in a situation that the

speaker assumes is not identifiable to the addressee. I suggest that for this to be the case, the semantic presupposition of the noun classifier cannot be part of the Common Ground:

(233) **Principle against private situations in the Common Ground**
If a proposition is based on a private situation, that proposition cannot be part of the Common Ground.

The intuition behind the principle in (233) is that for a situation to be private, it must be that propositions based on that private situation are not *mutually agreed upon*. After all, when using an indefinite with a private situation, the speaker is implying that their addressee may not be familiar with the value for the situation pronoun they intend. Now, given that the definition of Contextual Equivalence relies on evaluating propositions with respect to the Common Ground, (233) implies that *Maximize Presupposition* will not apply to propositions based on a private situation. The two sentences in (234) and (235), with and without a classifier, will therefore be equally well-formed:

(234) a. *Ixchi'waj jun nok' tz'i'.*
Ix-chi'waj [jun nok' tz'i'].
PFV-bark INDF CLF dog
'A dog barked.'

b. LF: [barked [$_{DP}$ *jun* pro_{p_7} [$_{ClfP}$ **nok'** dog]]]

c. Non-assertive content = $|y\text{: } \textsc{dog}(y)([\![pro_{p_7}]\!]^g)| = 1$

d. Assertive content = $\exists s[\textsc{past}(s) \wedge \exists x[\textsc{dog}(x)([\![pro_{p_7}]\!]^g) \wedge \textsc{bark}(x)(s)]]$

(235) a. *Ixchi'waj jun tz'i'.*
Ix-chi'waj [jun Ø tz'i'].
PFV-bark INDF Ø dog
'The dog barked.'

b. LF: [barked [$_{DP}$ *jun* pro_{p_7} [$_{ClfP}$ **Ø** dog]]]

c. Non-assertive content = N/A

d. Assertive content = $\exists s[\textsc{past}(s) \wedge \exists x[\textsc{dog}(x)([\![pro_{p_7}]\!]^g) \wedge \textsc{bark}(x)(s)]]$

In sum, if the assumption in (233) is on the right track, the *semantic* presupposition of the classifier will only truly result in a *pragmatic* presupposition (Stalnaker 1974) when it

combines with a definite determiner, but not when it combines with an indefinite determiner. This is because the content of the noun classifiers' semantic presupposition will only appear as a proposition in the Common Ground when it is based on a familiar situation pronoun. I proposed that this explains why noun classifiers are not enforced in specific indefinite DPs. When their situation argument is saturated by a private situation pronoun, they do not compete with Ø for *Maximize Presupposition*, and so (234-a) and (235-a) are both expected to be felicitous sentences. Therefore, the current analysis of definiteness and indefiniteness in Chuj, which crucially relies on the existence of two kinds of situation variables, provides us with the potential means to understand the necessity of noun classifiers with definite descriptions, but not with specific indefinites. I also argued that only a formulation of *Maximize Presupposition* based on lexical alternatives, as in Percus 2006 and Coppock and Beaver 2015, can explain the necessity of noun classifiers in definite descriptions (at least under the current approach).

While the current pragmatic approach could offer an account of the necessity of noun classifiers within definite descriptions, more work is needed to understand the consequences of introducing private situations within a theory of semantics and pragmatics. At first glance, for instance, it seems to complicate the taxonomy of the range of "non-assertive" content that may be encoded in lexical items (see Matthewson 2006 and Tonhauser et al. 2013 for related claims). In particular, *Stalnaker's Bridge*, which requires that semantic presuppositions systematically map to pragmatic presuppositions (see e.g., von Fintel 2008), only arises when the presupposition is based on a familiar situation pronoun. That is, semantic presuppositions associated with the same lexical items surface as pragmatic presuppositions only in a subset of cases.

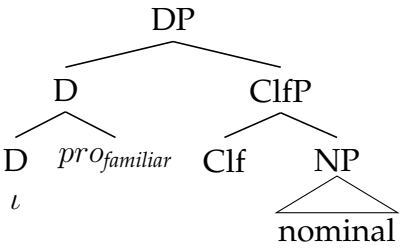
5.1.3 Summary and extension to third person pronouns

This section has extended the proposals from section 4.3 to definite descriptions, by refining the analysis of weak definite descriptions from Chapter 3. While the new analy-

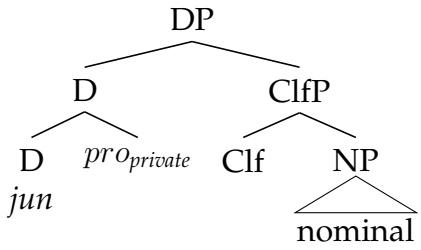
sis keeps to the core proposals from Chapter 3—noun classifiers introduce existence and uniqueness conditions on the extension of the NP—there were two key refinements. First, weak definites are further decomposed with a null DP layer, headed by ι, so as to allow for the syntactic locus of situation pronouns to be higher in the nominal projection, in a parallel fashion to indefinite expressions. Second, I argued that ι, contrary to indefinite quantifiers, c-selects for a familiar situation pronoun. A consequence of this was that the non-assertive component of the classifier only projects as a pragmatic presupposition when combined with ι, but not when combined with an indefinite determiner.

The new analysis provides a unified approach to the contribution of noun classifiers in definite and indefinite expressions. If the DP is headed by ι, the uniqueness condition is interpreted relative to a situation that is familiar to both speaker and addressee(s), and the result is a weak definite description. If the DP is headed by an indefinite determiner, then uniqueness is interpreted relative to a situation that is not necessarily known to all discourse participants, which I argued gives rise to a singleton indefinite. The two options are schematized again below:

(236) Weak definite description



(237) Indefinite+classifier DP



I also provided two—mutually-compatible—ways of understanding why noun classifiers are *obligatory* with definite descriptions, but *optional* with specific indefinite expressions. One of these solutions, elaborated in section 5.1.2.2, capitalizes on the fact that the uniqueness condition of the noun classifier only projects as a pragmatic presupposition when the DP is headed by ι, but not when the DP is headed by *jun*. If this is so, then the

necessity for noun classifiers in definite descriptions can be made to follow as a classic instance of *Maximize Presupposition*, sensitive only to pragmatic presuppositions.

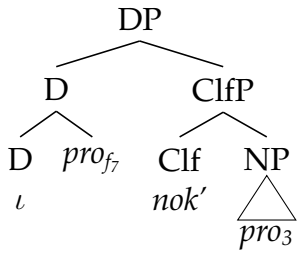
Finally, it should be noted that the analysis proposed in this section naturally extends to the case of third person anaphoric pronouns. In particular, in Chapter 3, I proposed that anaphoric uses of classifier pronouns involve a null predicative variable, repeated in (238). The predicative variable's sole contribution is to introduce an index:

(238) $[\![pro_i]\!]^g = \lambda x.\ x = g(i)$

Combining this predicative variable with the new derivation for weak definite descriptions will result in a meaning equivalent to the one discussed in Chapter 3. The pronoun will essentially pick out the unique entity identical to the value for an index in a contextually provided assignment function g. This is shown in (239) for anaphoric uses of the classifier pronoun *nok'*.

(239) a. *nok'* 'it'

b.



c. $[\![[[\ \iota\ pro_f\ ]\ nok'\ pro_i\ ]\ ]\!]^g$ is defined only if $|y\text{: } [\lambda z.\ z = g(3)](y)| = 1$.
When defined $= \iota x[x = g(3) \text{ in } g(7)]$

I assume that the presence of the noun classifier is required for the same reasons discussed in the previous section for the derivation of other definite descriptions. Moreover, as discussed in Chapter 3, I assume that the uniqueness condition of third person anaphoric pronouns will always be met, since there will always be at most one entity that is identical to the entity picked out by any given index. That is, there will always only be one entity identical to $g(3)$ in the range of the assignment function, guaranteeing that the unique-

ness condition in (239-c) is satisfied. Third person classifier pronouns should therefore be required, as is the case.

5.2 Noun classifiers: Predictions, linguistic variation, and potential extensions

In the preceding sections and chapters, I have shown that noun classifiers play a crucial role in the composition of DP. I provided a compositional analysis that can explain the distribution of noun classifiers in definite and specific indefinite expressions. Noun classifiers essentially introduce a uniqueness and existence condition, whose semantic/pragmatic effect varies depending on what kind of situation pronoun is merged with D^0.

In this section, I first show in section 5.2.1 that the current analysis makes a number of Chuj-internal predictions, largely borne out. Section 5.2.2 then discusses potential evidence for linguistic (micro)variation in the denotation of noun classifiers within the Q'anjob'alan language family, showing that minimal modifications to the current analysis could be straightforwardly implemented to account for this variation. Finally, section 5.2.3 turns to data from three unrelated languages, which could potentially be explained by extending the analysis of Chuj noun classifiers.

Note that some of the examples presented here build off of Craig (1977, 1986b) and Zavala's (1992, 2000) description and analysis of similar data in two Mayan languages closely related to Chuj: Popti' and Akatek. In a similar vein as the current account, Craig and Zavala's goals were to show that noun classifiers induce "referential" interpretations of indefinites (here reinterpreted as "singleton indefinites").

5.2.1 Chuj-internal predictions

5.2.1.1 Evidence of existence presupposition

The final entry for noun classifiers proposed in this thesis preserves the existence and uniqueness conditions of the noun classifier from Chapter 3. Modulo cases where the nominal expression appears under the scope of a quantifier over situations (see Chapter 3), this in turn means that noun classifiers should not be permitted with nominal expressions that do not pick out or quantify over a specific (single) entity in the context. That is, for the presuppositions of the classifier to be met, there must a unique satisfier of the nominal predicate within the situation of evaluation, regardless of whether this situation is 'familiar' or 'private' to all discourse participants. As we already saw in section 4.2, DPs with classifiers indeed lack non-singleton quantificational interpretations of indefinites. Instead, they can be used with indefinite expressions that have a single witness. There is, however, additional evidence for the non-assertive existence component of noun classifiers.

One piece of evidence comes from the inability of noun classifiers to appear as the objects of so-called "incorporated antipassives". Much work on Q'anjob'alan languages have noted that noun classifiers are incompatible with such constructions (Maxwell 1976; Craig 1986b; Zavala 2000; Coon 2019):

(240) *Ixonhchonhwi anh onh.*
Ix-onh-chonh-w-i [(***anh**) onh].
PFV-B1P-sell-AP-IV CLF avocado
'We sold avocados' (lit. we avocado-sold) (adapted from Coon 2019: 57)

Coon (2019) argues that the NP in (240) denotes a property that restricts the denotation of transitive verbs like *chonh* 'to sell' from events of selling to events of avocado-selling. The nominal expression *onh* in (240) neither picks out or quantifies over any particular individual, and no existence or uniqueness inference arises. For one, this sentence could

be felicitously used to describe a situation in which the speakers went avocado-selling, but unfortunately did not manage to sell any avocados. This is clearly incompatible with the existence presupposition of the noun classifier. Moreover, Coon (2019) shows that incorporated nominals are "discourse opaque", in the sense of Farkas and de Swart (2003):

(241) Coon 2019: 57

a. *Ixinmanwi onh.*
Ix-in-man-w-i onh.
PFV-B1S-buy-AP-IV avocado
'I avocado-bought.'

b. *Yaxto anh.*
Yax-to anh.
green-already CLF.PRON
Intended as continuation of (241-a): 'It was ripe.' / 'They were ripe.'

The fact that objects of incorporated antipassives cannot be referred back to in the discourse is entirely expected given previous works that distinguish between purely predicative nominals and those used to pick out or quantify over individuals. Indeed, Doron (1988) treats the availability of anaphoric relations as a key diagnostic in determining a nominal expression's status as "referential", and conversely, the unavailability of anaphoric relations as a diagnostic for the lack of reference.

A similar state of affairs holds of nominals that appear as modifiers inside an extended DP, as also shown by Craig (1986b) and Zavala (2000). Consider the following examples:

(242) a. *Ixjaw winh chonhum ixim wa'il.*
Ix-jaw [winh chonhum (***ixim**) wa'il].
PFV-arrive CLF salesperson CLF tortilla
'The tortilla salesman arrived.'

b. *Teniwan ixim.*
Te-niwan ixim.
PFV-big CLF.PRON
Intended as continuation of (242-a): 'It was very big.'

In (242-a), the use of the classifier with *wa'il* 'tortilla' is ungrammatical. This is expected, since in this particular case, *wa'il* does not pick out or quantify over any particular entity,

but rather serves to restrict the denotation of *chonhum* 'salesperson' to only the salespeople who sell tortillas. The presence of a noun classifier would lead to the inference that there exists a unique tortilla that the relevant salesman sells, which is clearly not the intended meaning of (242-a). Also notice that the nominal expression *wa'il* in (242-a) is discourse opaque (it does not produce an antecedent for anaphora), as shown in (242-b).

Moreover, as also shown by Craig (1986b) and Zavala (2000), noun classifiers in Q'anjob'alan languages are illicit with the predicates of predicational copular sentences, a fact which was already noted in section 3.3.2.3:

(243) a. *Doktor ix ix tik.*
(*__ix__) Doktor ix ix tik.
CLF doctor CLF woman DEIX
'This woman is a doctor.'

b. *Hingana ha waj Xun alkalil.*
Hin-gana ha waj Xun (*__winh__) alkalil.
A1S-want PART CLF Xun CLF mayor
'I want Xun to be mayor.'

While the semantics of copular sentences in Mayan requires further study, it would be unlikely for predicate nominals in predicational copular clauses to be compatible with the existence and uniqueness conditions of noun classifiers. In fact, languages tend to overtly distinguish predicational copular clauses from equative (or specificational) copular clauses, the latter of which instantiate two full-fledged definite DPs (see e.g. Mikkelsen 2005). As expected, Chuj equative copular clauses, which are morphosyntactically differentiated from predicational copular clauses, *do* require noun classifiers:

(244) *Ha ix doktor ha ix ix tik.*
Ha *(__ix__) doktor ha ix ix tik.
PART CLF doctor PART CLF woman DEIX
'This woman is the doctor.'

The utterance in (244) is distinct from predicational copular clauses, which cannot appear with the particle *ha*. As shown, the noun classifier is obligatory in such cases.

In sum, the current analysis of noun classifiers predicts that they should be incompatible with nominals that get used as predicates, which do not pick out or quantify over single entities. As briefly shown here, this prediction is borne out.

5.2.1.2 Incompatibility of co-occurrence with "non-specific" expressions

Given that noun classifiers in indefinites yield singleton indefinites, another empirical prediction that the current analysis makes is that they should be illicit with indefinite expressions, which, for independent reasons might not be compatible with singleton domains. Many such expressions have been identified in the literature, including the direct objects of imperatives and interrogatives (Dayal 2019), and free-choice indefinites (Kratzer and Shimoyama 2002, Alonso-Ovalle and Menéndez-Benito 2010, 2018).

Starting with imperatives, Dayal (2019) recently argues that direct objects of imperatives should be systematically incompatible with singleton indefinites. The reason is essentially pragmatic: if a speaker does not reveal the contextual restriction of the direct object to their hearer, then it is impossible for their hearer to follow through with the order. As she mentions, this reasoning can also account for the infelicity of English imperatives with *a certain*, as in (245).

(245) # Buy a certain book!

As predicted, indefinites in imperatives in Chuj are judged infelicitous in co-occurrence with a noun classifier (246).

(246) *¡Man jun libro!*
¡Man jun (#ch'anh) libro!
buy INDF CLF book
'Buy a book!'[5]

[5]With the right context, this sentence could be felicitously interpreted as 'buy one of the books'. See discussion of partitives in §4.1.2 above.

A similar kind of pragmatic incompatibility can also be shown to arise when noun classifiers appear within interrogative expressions (Dayal 2019). Consider, for instance, the following example, which shows that noun classifiers cannot co-occur with the interrogative word *tas* 'what':

(247) *¿Tas nipal hagana?*
¿Tas **(#anh)** nip-al ha-gana?
WH CLF huipil-NML A2S-desire
'What *huipil* (traditional garment worn by women) do you want?

If we take interrogative words to be existential quantifiers (Karttunen, 1977), then the infelicity of examples like (247) follows from the analysis. This is because restricting the domain of an interrogative word to a singleton set would render the question trivial. That is, questions presuppose that some answer is true, but if the speaker is able to delimit the domain of the interrogative to a singleton, then it amounts to presupposing that one answer is true. In other words, the question denotation will contain just one possible answer and it will be presupposed that this one answer is true, leading to expected (and observed) infelicity.

Related to the *wh*-questions seen above are a class of modal indefinites in Chuj, which can combine with the morpheme *yalnhej* to derive free choice effects (see e.g. Kotek and Erlewine 2019, Royer 2021b, Alonso-Ovalle and Royer 2022). Like Spanish, Chuj features a random-choice modal indefinite, *yalnhej tas*, described and analyzed as a free-choice indefinite in Kotek and Erlewine 2016, Royer 2021b, and Alonso-Ovalle and Royer 2022. Such modal indefinites are not compatible with noun classifiers:

(248) *Yalnhej tas itajil ixinyama'.*
Yalnhej tas **(*anh)** itaj-il ix-in-yam-a'.
FC WH CLF herb-NML PFV-A1S-grab-TV
'I grabbed a random herb.'

The presence of *yalnhej* in (248) leads to a free choice effect: the speaker is conveying that they grabbed some herb indiscriminately. Though the proper semantic characterization of

yalnhej DPs requires more work. it is entirely expected that free choice morphemes should be incompatible with singleton restrictors. For instance, in Alonso-Ovalle and Royer 2022, which builds on the analysis of the Spanish free choice indefinite *uno cualquiera* in Alonso-Ovalle and Menéndez-Benito 2018, the free choice modal effect in (248) would essentially be derived by identifying all of the alternative entities that are *herbs* in the context, and conveying that the decisions of the agent did not discriminate among these different entities. If the domain of the quantifier were restricted down to a singleton, it would follow that no alternative herbs could be considered, and we should therefore expect infelicity, as is the case.

In sum, we have seen that several constructions, which are predicted to be incompatible with singleton domain restriction, ban the presence of noun classifiers. In the next subsection, we turn to another expression which appears to ban singleton domain restriction. However, I show that there seems to be variation within the Q'anjob'alan language family with regards to the acceptability of noun classifiers with this expression, potentially revealing linguistic variation in the denotation of noun classifiers across different Q'anjob'alan languages.

5.2.2 Linguistic variation in noun classifiers: the case of *jun=ok*

As discussed in the previous subsection, noun classifiers are generally incompatible with indefinite expressions that obligatorily convey "non-specificity". Here, I discuss another one of these items, which exhibits a distribution reminiscent of a Negative Polarity Item (NPI) in English. I show that the data from Chuj is by and large consistent with the proposed entry for noun classifiers in this chapter: the collaborators on this project judge co-occurrences of noun classifiers with *jun=ok* as infelicitous—a fact which echoes previous descriptions of the same item in two other Q'anjob'alan languages (Craig 1986b, Zavala 2000). However, I also show that there seems to be variation in the reported speaker judgements, even in Chuj, which I take to be indicative of potential linguistic variation in

the denotation of noun classifiers. I then show that minimal modifications to the current analysis of noun classifiers could be straightforwardly implemented to account for this variation, which I view as a welcome conceptual advantage of the current analysis.

Mayan languages of the Q'anjob'alan sub-branch feature an irrealis clitic, *=ok* in Chuj, described as an NPI in Royer (2019) and as a "non-specific" indefinite in recent work (Buenrostro 2022). While more work on the semantics of *jun=ok* is ultimately needed, it is interesting to see what happens when *jun=ok* co-occurs with noun classifiers. If *jun=ok* is an NPI (which have been regarded in some work as "domain wideners"; Kadmon and Landman 1993), or forces non-specific interpretations of indefinites, then we might expect noun classifiers to be incompatible with this expression.

Like weak NPIs (e.g., *any* in English), *jun=ok* is generally licensed in downward entailing environments, as in (249-a)–(249-c), but can also sometimes appear in upward entailing environments, in which case it gives rise to free choice effects. The sentence in (249-d), for instance, roughly conveys that the speaker does not care what flower they find.

(249) a. *Maj wil laj junok pale.*
Maj w-il laj **jun=ok** pale.
NEG.PFV A1S-see NEG INDF-IRR priest
'I didn't see any priest (not even one).'

b. *¿Ixhil yek' junok tz'i'?*
¿Ix-h-il y-ek' jun=**ok** tz'i'?
PFV-A2S-see A3-DIR.pass INDF=IRR dog
'Did you see any dog?'

c. *Tato tzhil junok anima', ¡tzhal t'ayin!*
Tato tz-h-il jun=**ok** anima', '!tz-h-al t'ay-in!
if IPFV-A2S-see INDF=IRR person, IPFV-A2S-say PREP-B1S
'If you see a person, tell me.'

d. *Lan hinsayan junok xumak.*
Lan hin-sayan jun=**ok** xumak.
PROG A1S-look.for INDF=IRR flower
'I'm looking for any flower (no matter which).'

As already discussed in Chapter 2, the presence of *=ok* tends to lead to an interpretation of the indefinite as if it took low-scope below other operators in the clause. For instance, the utterance in (249-a), repeated from (51) above, triggers the interpretation in (250-a) made available by the context indicated below, and does not allow the reverse scope order in (250-b) with its context (the reverse order is only possible with *jun* without *=ok*):

(250) Felicitous and infelicitous contexts for (249-a)

a. ✓ $\neg\exists x$[PRIEST(x) $\wedge$ SEE(x)(SPEAKER)]
Context: I wanted to see a priest, but never saw one. I tell you (51).

b. ✗ $\exists x$[PRIEST(x) $\wedge$ $\neg$SEE(x)(SPEAKER)]
Context: I saw several priests, but there's one particular priest I didn't see. I tell you (51).

Since *jun=ok* forces low scope interpretation of indefinites, noun classifiers should be incompatible with *jun=ok*. The Chuj collaborators on this project do indeed judge co-occurrence of *jun=ok* with noun classifiers as infelicitous. Two examples are provided below:[6]

(251) a. *Maj chax laj junok libro wu'uj.*
Maj chax laj jun=ok (#**ch'anh**) libro w-u'uj.
NEG.PFV find NEG INDF=IRR CLF book A3-by
'I didn't find any book(s).'

b. *¿Ixhil yek' junok tz'i'?*
¿Ix-h-il y-ek' jun=ok (#**nok'**) tz'i'?
PFV-A2S-see A3-DIR.pass INDF=IRR CLF dog
'Did you see any dog(s)?'

This same restriction is identified by Craig (1986b) and Zavala (2000) for the cognate of *jun=ok* in Popti' and Akatek. For instance, Craig notes that (p. 268-69): "when the objects of non-implicative verbs are non-referential, they are marked with the irrealis suffix *=oj* (which is subject to vowel harmony) and do not carry classifiers" (see Zavala (2000), p.

[6]Note that given the right context, (251-a) and (251-b) could receive felicitous partitive interpretations, e.g. 'I didn't find any of the books'. See §4.1.2.

138 for the relevant passage on Akatek). Craig provides the example in (252), crucially without a noun classifier:

(252) X'oq' heb' ix say-a' **hun-uj** munlab'al.
started PL CLF.PRON look.for-TV INDF-IRR pot
'They started looking for a pot.' (Popti': Craig 1986b: 269)

While the abovementioned incompatibility of noun classifiers with *jun=ok* in Chuj and cognates of *jun=ok* in Popti' and Akatek is entirely expected given the denotation assigned to noun classifiers in this thesis, it is interesting to note a point of potential linguistic variation within Chuj and another Q'anjob'alan language with regards to co-occurrences of noun classifiers with *jun=ok*. Specifically, there is reason to believe that noun classifiers are compatible with the cognate of *jun=ok* in Q'anjob'al, a fourth Q'anjob'alan language, as well as for some speakers of Chuj.

If *jun=ok* does indeed force a low-scope interpretation of the indefinite for these speakers, this suggests that the semantic presuppositions associated with noun classifiers may perhaps vary across the sub-family. Here, I show that the refined analysis of noun classifiers proposed in this chapter and in Chapter 4 could be straightforwardly modified to accommodate such variation, which I take as a conceptual benefit of the current account.[7]

[7] Another point of potential variation requiring further work may concern the actual distribution of *jun=ok*. As mentioned in Chapter 2, some speakers show a preference for the use of *jun=ok* in certain low-scoping environments, such as negative sentences (251-a) and polar questions (251-b) (see Buenrostro 2022, Buenrostro and Royer 2022). This led Buenrostro (2022) to hypothesize that *jun=ok* is a non-specific indefinite determiner, and therefore that the use of *jun* without *=ok* implies specificity. However, non-specific interpretations of *jun* without *=ok* are consistently judged possible by the Chuj speakers I have consulted. This was already shown in several examples above, including the conditional sentence in (166). As shown in (i), the equivalent of (249-d) with *jun* in a non-specific context is also felicitous in the absence of *=ok*. Of course, there may well be variation from one Mayan language to another regarding the necessity of irrealis marking on indefinites in irrealis environments. As discussed in Becker 2021, the necessity of *=ok* in Q'anjob'al seems to be much stronger than in other Mayan languages, such as Ch'ol, where she notes that the cognate irrealis marker "is not used systematically in nonspecific irrealis contexts" (p. 285).

(i) **Context:** To decorate my house, I like to have flowers. But I'm not difficult, any flower is fine. I go outside to look for a flower, and someone asks me what I'm doing. I tell them:
Lan hinsayan jun xumak.
Lan hin-sayan **jun** xumak.
PROG A1S-look.for INDF flower
'I'm looking for a flower.'

Again, the main area of variation lies in judgements about the acceptability of *jun=ok* with noun classifiers. While the consultants on this project tend to reject (non-partitive) co-occurrences of *jun=ok* with noun classifiers—a result which was also initially noted for Popti' and Akatek (Craig 1986b, Zavala 2000), there appears to be linguistic variation in the acceptability of *jun=ok*+classifier co-occurrences, both internally to Chuj and in Q'anjob'al. For example, Buenrostro (2022) recently reports uses of noun classifiers in Chuj with *jun=ok* indefinites, such as the following (where the narrow scope reading is indicated in the translation):

(253) *Tzinnib'ej tzinman junok te' pat.*
Tz-in-nib'-ej tz-in-man **jun=ok te'** pat.
IPFV-A1S-want-DTV IPFV-A1S-buy INDF=IRR CLF house
'I want to buy a house (and I don't know which).' (Buenrostro 2022: 15)

I have also been able to identify cases of *jun=ok* co-occurring with indefinites in Chuj narratives. However, the number of co-occurrences remains relatively low. In 17 transcriptions ranging from 30 minutes to an hour in the corpus of Mateo Pedro and Coon (2018), I only identified about a dozen co-occurrences of *jun* with noun classifiers (there were 186 instantiations of *jun=ok*). What is more, most *jun=ok*+classifier co-occurrences identified in narratives are amenable to a partitive analysis, in which case the presence of a classifier is expected. That is, *jun=ok* co-occurs with a *definite* DP in such cases. I provide two relevant examples below.

Moreover, the licensing requirements for *=ok* in Chuj seem to be tied to the presence of irrealis mood operators, rather than non-specificity of the DP. For instance, epistemic non-specificity cannot be indicated with *jun=ok* in episodic sentences (which I assume are not irrealis given that they describe completed events):

(ii) **Context:** There are traces of a dog in the speaker's house, e.g. paw prints, that lead the speaker to think that a dog stole the chicken. However, since there are dozens of dogs in the village, it is unclear what dog stole the chicken.
Hijan ixkochi' nok' kaxlan, ixik' b'at nok' jun tz'i'.
Hijan ix-ko-chi' nok' kaxlan, ix-ik' b'at nok' [**jun**(#=ok) tz'i'].
almost PFV-A1P-eat CLF chicken PFV-bring DIR.go CLF INDF=IRR dog
'We were going to eat the chicken, but a dog stole it.'

In (254), the classified noun *jun=ok te' te' chi'* was used in a conversation after the speaker had been talking about the different kinds of wooden sticks available to build music instruments. Given this context, and the fact that the nominal expression co-occurs with the deictic particle *chi'*, it is reasonable to think that this DP involves a partitive syntax.

(254) *Mato ay juok te' te' chi' espesyal t'ay, tz'aj kalani.*
Mato ay [jun=**ok** **te'** te' **chi'**] espesyal t'ay, tz'aj k-alan-i
or-COMP EXT INDF=IR CLF stick DEIX special PREP, how A1P-say-IV
'But there's one of these sticks that's special, you know.' (txt, CM300715)

In (255), we see that the plural marker *heb'* appears after *jun=ok*, likewise indicating a partitive construction (see §4.1.2 above):

(255) *[...] mato ay junok heb' winh kaxlanh winak tzjawi.*
[...] ma-to ay [jun=**ok** **heb' winh** kaxlanh winak] tz-jaw-i'.
or-COMP EXT INDF=IRR PL CLF ladino man IPFV-come-IV
'[...] or if there's one of the ladinos who comes.' (txt, CJ230715)

I was, however, able to identify a few cases of *jun=ok* co-occurring with classifiers, which at least at first glance, do not seem to be amenable to a partitive analysis. This is the case in (256), which the consultants on this project find infelicitous:

(256) *Pero lasmia ichok honh hayik wan slolon junok winh ichamwinak, max kak' och laj kochikin t'a.*
Pero lasmia ich=ok honh hayik wan s-lolon **jun=ok** **winh** ichamwinak,
but sadly like=IRR B1P when PROG A3-speak INDF=IRR CLF elder.man,
max k-ak' och laj ko-chikin t'a.
NEG.IPFV A1P-give DIR.out NEG A1P-ear PREP
'But sadly when an elder would speak, we didn't pay attention.' (txt, CM210715)

Relatedly, similar co-occurrences of *jun=ok* with classifiers are reported for Q'anjob'al (Bervoets 2014; Becker 2021), where a noun classifier can optionally co-occur with the cognate indefinite *jun=oq*:

(257) Q'anjob'al (Bervoets 2014)

a. Ch-y-och-ej naq Pedro [**jun=oq** **ix** ix] q'ajab' b'ay naq.
IPFV-A3-want-DTV CLF Pedro INDF=IRR CLF woman to CLF
'Pedro wanted a woman (any woman) to talk to him.'

b. Aqwal ch-in man [**jun=oq** no' hin chej].
should IPFV-A1S buy INDF=IRR CLF A1S horse
'I should buy a horse (any horse).'

In sum, the collaborators on this project generally reject *jun=ok*+classifier co-occurrences, a fact which is replicated for closely-related languages like Popti' and Akatek (Craig 1986b, Zavala 2000). This is expected, since, all else being equal, *jun=ok* should not be compatible with a uniqueness-based approach to noun classifiers. However, we also saw that cognate *jun=ok*-classifier co-occurrences seem possible in Q'anjob'al (Bervoets 2014, Becker 2021), and for other speakers of Chuj (Buenrostro 2022).

It is perhaps not surprising to find some amount of linguistic variation in the syntactic and semantic distribution of noun classifiers.[8] According to Hopkins (2012b), noun classifiers in the subset of Mayan languages that have them are a relatively recent innovation. There are also reported shifts in the use of noun classifiers in other Q'anjob'alan languages (Grinevald 2016), and rich dialectal microvariation among Chuj speakers and speakers of other Mayan languages (Maxwell 1981; García Pablo and Domingo Pascual 2007; England 1990; Lemon 2019). It is therefore perfectly conceivable that speakers internalize different semantic entries for noun classifiers. Under the current account, formalizing this shift would be quite simple. For instance, the non-assertive content of the classifier could be minimally manipulated in many ways, with foreseeable effects for the distribution of noun classifiers. Some hypothetical entries are provided below:

[8] In fact, in Mayan languages of the Mamean sub-branch, where dialectal variation has received more attention (England 1990; England 2017), noun classifiers pattern in a number of different ways from one dialect to the other (Lemon 2019).

(258) Hypothetical variation in the denotation of *nok'* (classifier for animals)

a. **Uniqueness + existence + class (current proposal)**
$\lambda P_{\langle e,\langle s,t\rangle\rangle}.\lambda x.\ \lambda s\colon |\{y\colon P(y)(s)\}| = 1 \wedge \{y\colon P(y)(s)\}\subseteq\{z\colon \textsc{animal}(z)\}.\ P(x)(s)$

b. **Existence + class**
$\lambda P_{\langle e,\langle s,t\rangle\rangle}.\lambda x.\ \lambda s\colon |\{y\colon P(y)(s)\}| \neq 0 \wedge \{y\colon P(y)(s)\}\subseteq\{z\colon \textsc{animal}(z)\}.\ P(x)(s)$

c. **Uniqueness + class (à la Coppock and Beaver 2015)**
$\lambda P_{\langle e,\langle s,t\rangle\rangle}.\lambda x.\ \lambda s\colon |\{y\colon P(y)(s)\}| \leq 1 \wedge \{y\colon P(y)(s)\}\subseteq\{z\colon \textsc{animal}(z)\}.\ P(x)(s)$

d. **Class**
$\lambda P_{\langle e,\langle s,t\rangle\rangle}.\lambda x.\ \lambda s\colon \{y\colon P(y)(s)\}\subseteq\{z\colon \textsc{animal}(z)\}.\ P(x)(s)$

What we see is that, given a predicative analysis of noun classifiers, it is possible to formulate alternative denotations of noun classifiers with weaker non-assertive content, including formulations in which the classifier requires existence but not uniqueness (258-b), uniqueness but not existence (258-c), or even formulations in which it requires neither uniqueness or existence (258-d). While the alternative denotations of the noun classifiers in (258-b)–(258-d) make different predictions and should be verified, a strength of the current analysis of noun classifiers is that it can potentially be extended to explain slight semantic differences in the use of cognate noun classifiers across the Q'anjob'alan language family.

I leave the semantic characterization of *=ok* in combination with *jun*, and in other environments more generally, as well as an investigation of variation in the semantics of noun classifiers across the Mayan language family, for future work.

5.2.3 Potential cross-linguistic extensions

If the current analysis of definites and specific indefinites in Chuj is on the right track, it leads to the following typological expectation: other languages might pattern with Chuj in exhibiting morphological overlap between weak definiteness and specific indefiniteness. That is, other languages might feature singleton-inducing morphemes (with a semantics similar to that of noun classifiers) that appear in both definite and (specific)

indefinite environments. As it turns out, many languages have been argued to lump together specificity and definiteness (Lyons 1999; Becker 2021). Here, I point to three of these languages, which exhibit a strikingly similar distribution to Chuj, a fact which I claim reinforces the possible empirical scope of the proposal.

A number of languages exhibit configurations potentially amenable to the analysis provided in this work, where a morpheme used to convey weak definiteness can also co-occur with an indefinite determiner in order to give rise to a specific interpretation. One such language is Turkish, which as argued by several authors, marks specificity distinctions with indefinites via differential object marking (see e.g. Enç 1991, Kornfilt 1997. von Heusinger and Kornfilt 2005, 2017, and Hedberg et al. 2009). The examples in (259) show that while Turkish requires accusative marking on direct objects that have definite, specific indefinite, and demonstrative interpretations—indefinites not specified for specificity do not appear with the accusative marker.

(259) Turkish (von Heusinger and Kornfilt 2005: 5)

a. (Ben) bir kitab oku-du-m.
I a book read-PAS-1SG
'I read a book.' (indefinite)

b. (Ben) kitab-**ı** oku-du-m.
I book-ACC read-PAS-1SG
'I read the book.' (definite)

c. (Ben) **bir** kitab-**ı** oku-du-m.
I a book-ACC read-PAS-1SG
'I read a specific book.' (specific indefinite)

d. (Ben) **bu** kitab-**ı** oku-du-m.
I this book-ACC read-PAS-1SG
'I read that book.' (demonstrative)

The Turkish nominal expressions in (259), and their resulting meanings, bear striking similarities with those described for Chuj in the last three chapters. We see that direct object definite descriptions (259-b) take a morpheme, the accusative marker *ı*, which can co-occur with the indefinite determiner, *bir*, to force specific interpretations of the indefinite

(259-c). Moreover, the same accusative marker can also co-occur with a deictic in order to create a demonstrative, as is also the case in Chuj (see Chapter 2–3 and the next section).

Another language which exhibits a similar distribution is Persian, which exhibits the same morphological pieces as Turkish and Chuj. As Hedberg et al. (2009) and von Heusinger and Sadeghpoor (2020) show, definite (260-b) and specific indefinite (260-c) direct objects share a morpheme, *rā*, not found on non-specific indefinites (260-a):

(260) Persian (Hedberg et al. 2009: 5)

a. Emruz **ye** vakil mi-bin-am.
today a/one lawyer DUR-see-1SG
'I am seeing a lawyer today.' (indefinite)

b. Ali ketâb-***(o)** xund
Ali book-RĀ read-3SG
'Ali read the book.' (definite)

c. Emruz **ye** vakil-(i)-**o** mi-bin-am.
today a/one lawyer-I-RĀ DUR-see-1SG
'I am seeing a (particular) lawyer today.' (specific indefinite)

von Heusinger and Sadeghpoor (2020) further show, with context, that wide-scope readings are forced in the presence of *rā*:

(261) Persian; examples adapted form von Heusinger and Sadeghpoor 2020: 126.

a. **Context:** There were three possible books I could buy. I didn't buy any of them.

Man hattā **ye** ketāb ham naxaridam.
I even a book also not.bought.1SG
'I didn't buy any book.' (¬ > ∃)

b. **Context:** There were three possible books I could buy. I bought two of them but not three.

Man **ye** ketāb **ro** naxaridam.
I a book RĀ not.bought.1SG
'There is a book I didn't buy.' (∃ > ¬)

These data stand in striking parallel to the data seen in Chuj and Turkish. Yet again, we see that specific indefinites and definites pattern alike in requiring overt accusative marking. The similarities are summarized in Table 1.2 below.

Table 5.1: Turkish, Persian and Chuj DP configurations

Turkish (direct objects)	Persian	Chuj	Result
bir - NOUN	*ye* - NOUN	*jun* - NOUN	indefinite
NOUN - ACC	NOUN - RĀ	CLF - NOUN	definite
bir - NOUN - ACC	*ye* - NOUN - RĀ	*jun* - CLF - NOUN	specific indefinite

Finally, another language which appears to parallel Chuj, Turkish, and Persian in the distribution of definites and specific indefinites is Makassarese, an Austronesian language spoken in South Sulawesi, Indonesia.[9] Though the status of indefiniteness in Makassarese remains to be further studied, Jukes (2006) notes that the definite article "=*a* can be attached to *se're* 'one' [...] with the meaning of 'a certain X, a particular X"':

(262) Makassarese; examples obtained from Jukes 2006

a. Ku=kanre=i unti=**a**.
1=eat=3 banana=DEF
'I eat the bananas.' (Jukes 2006: 152)

b. Ri **se're=a** kampong amm-antang=i se're=**a** tu kalabini.
PREP one=DEF village MV-stay=3 one=DEF person couple
'In a certain village lived a certain couple' (Zainuddin Hakim 1991: 126)

Again, this is similar to Chuj, Turkish, and Persian, where the same morpheme that participates in definite descriptions appears to combine with an indefinite morpheme to form a specific indefinite.

Other languages which possess morphemes that exhibit overlap in their appearance with definite and specific indefinite expressions have been reported (see e.g. Mosel and Hovdhaugen 1992, Ionin 2006 on Samoan; Hendrick 2005 on Tongan; Givón 1969 on Bemba; Blass 1990 on Sisaala, Matthewson 1999, 2008 on St'át'imcets; Becker 2021 on Ayoreo). While more work is needed to establish what kinds of parallels these languages

[9] I thank Dan Brodkin for pointing me to these data.

exhibit with Chuj, I submit that the potential cross-linguistic extensions observed in this section provide independent support for the line of analysis proposed in this work.

5.3 Demonstratives revisited

Having provided a refined analysis of noun classifiers in Chuj, which allows for a unified treatment of their contribution in the composition of DP, we now revisit the denotation of deictic particles, which as discussed in Chapters 2 and 3 of this thesis, also play an important role in the composition of Chuj nominal expressions. I first summarize relevant information seen in the previous chapters about demonstrative constructions. I then turn to a unified analysis of deictic particles, which closely resembles the analysis provided in Chapter 3, but can account for the different demonstrative uses observed in Table 5.2, repeated from 4.1 for convenience.

Table 5.2: Possible DP configurations in Chuj

	Configuration	Example	Rough translation
①	CLF + NP	*nok' tz'i'*	'the dog'
②	CLF + NP + DEIX	*nok' tz'i' chi'*	'the/that dog' (anaphoric)
③	CLF	*nok'*	'it'
④	INDF + NP	*jun tz'i'*	'a dog'
⑤	INDF + CLF + NP	*jun nok' tz'i'*	'a (certain) dog'
⑥	INDF + NP + DEIX	*jun tz'i' chi'*	'that dog'
⑦	INDF + CLF + NP + DEIX	*jun nok' tz'i' chi'*	'that dog' (exophoric)

5.3.1 The composition of Chuj demonstratives revisited

Deictic particles in Chuj show at least two intriguing properties. First, as discussed at length in Chapter 3, they are generally required to form strong definite descriptions (unless there is overt topic marking). In particular, strong definites with overt nominals require deictic particles to combine with a *bare classifier phrase* (without an indefinite determiner):

(263) Chuj strong definite
[CLF + N + DEIX]

In Chapter 3, I provided a preliminary compositional analysis of deictic particles in combination with weak definite descriptions. I proposed the entries in (264) and (265) for the default and proximal deictic particles respectively, which both denote a partial identity function over individuals (type $\langle e,e \rangle$). In the non-assertive component of these entries, the deictic particle introduced an index, and presupposed that the satisfier of the NP was the value for that index in the range of the assignment function.

(264) Denotation of deictic particle *chi'* (repeated from (109))
$[\![\text{DEIX}_i]\!]^g = \lambda x\text{: } x = g(i).\ x$

(265) Denotation of deictic particle *tik* (repeated from (110))
$[\![tik_i]\!]^{g,c} = \lambda x\text{: } x = g(i) \wedge \text{CLOSE.TO.SPEAKER.OF.}c(x).\ x$

These entries remain compatible with the weak definite semantics assigned to classifier DPs in this chapter. The function could simply be fed the DP headed by ι, which is of the right semantic type for the deictic particle to take as argument (type e).

However, when we consider a wider set of environments in which deictic particles appear, there is a compositional confound, which is where the second intriguing property of deictic particles comes into play. As discussed in Chapter 2 and in Buenrostro and Royer 2022, as well as briefly in section 3.3.2.1, deictic particles can combine with what are arguably both definite and indefinite expressions. That is, singular demonstrative constructions in Chuj can be formed in one of two ways: (i) by combining a deictic particle with a bare classifier phrase, indicative that the DP is interpreted as definite, or (ii) by combining the deictic particle with an indefinite determiner. The two options are represented below:

(266) Demonstratives with definite base (DDB)

a. *nok' tz'i' tik*
nok' tz'i' **tik**
CLF dog DEIX.PRX
'the/this dog'

b. *nok' tz'i' chi'*
nok' tz'i' **chi'**
CLF dog DEIX.DFLT
'the/that dog'

(267) Demonstratives with indefinite bases (DIB)

a. *jun tz'i' tik / jun nok' tz'i' tik*
jun (nok') tz'i' **tik**
INDF CLF dog DEIX.PRX
'this dog'

b. *jun tz'i' chi' / jun nok' tz'i' chi'*
jun (nok') tz'i' **chi'**
INDF CLF dog DEIX.DFLT
'that dog'

From the perspective of existing theories on the formal semantics of demonstratives, it is surprising that deictic particles are compatible with indefinite DPs (though discussion in Matthewson 1998, 2013 on similar data in St'át'imcets and Arsenijević 2018 on Serbo-Croatian). Previous authors working on complex demonstratives have widely assumed that demonstratives are a sub-type of definite article (see e.g. Roberts 2002; Wolter 2006; Elbourne 2008, 2013; Ahn 2019). Nonetheless, as discussed in section 2.5.3, DIBs show a collection of properties that are characteristic of indefiniteness—all of which do *not* arise with DDBs. This includes (i) their ability to introduce new referents in discourse and (ii) their ability to appear in existential sentences, which ban definite descriptions in Chuj (see, for instance, example (57)). Both of these properties are illustrated in the following example, repeated from (64):

(268) *Komo ha t'ay pekatax, ay jun winh icham chi'...*
Komo ha t'ay pekatax, ay [***(jun)** winh icham chi']...
DISC FOC PREP before, EXT INDF CLF man DEIX
'Once upon a time, there was an old man...' (Hopkins 2021: 45)

Furthermore, as briefly discussed in section 3.3.2.1, the two kinds of demonstratives show another property characteristic of their respective composition. While DDBs trigger uniqueness/maximality effects, DIBs do not. Another example illustrating this effect is provided below:

(269) **Context** (recreated physically with consultant): There are four ants on the arm of the addressee. The speaker and the addressee have been talking about these ants. Signalling just one of the ants, the speaker tells the addressee:

a. *¡Ha jun sanich tik, te'wach' kilani!*
Ha **jun** sanich **tik**, te'-wach' k-il-an-i
FOC INDF ant DEIX INTS-good A1P-see-DEP-IV
'This ant looks great!' (indefinite base)

b. *¡Ha nok' sanich tik, te'wach' kilani!*
Ha **nok'** sanich **tik**, te'-wach' k-il-an-i
FOC CLF ant DEIX, INTS-good A1P-see-DEP-IV
Intended: 'This ant looks great!' (definite base)
Consultant comment: This is fine if you're talking about all of the ants.

The example in (269-b) shows that DDBs give rise to maximality effects, not attested with DIBs (269-a).

Given the collection of evidence showing that deictic particles *do* indeed combine with indefinite expressions to derive demonstrative constructions, we must therefore seek a unified account of deictic particles, which can account for their co-occurrence with both definite and indefinite expressions. To arrive at a unified analysis of deictic particles, we face both compositional and semantic challenges. First, we need an entry and syntax for Chuj deictic particles that allow them to semantically compose with both definite and indefinite expressions alike. Second, we need the overall result to deliver an appropriate semantics for both kinds of demonstratives—that is, one which can account for the uniqueness versus non-uniqueness effects observed with DDBs and DIBs, respectively. Here, I show that it is possible to explain both of the facts if we (i) adopt the semantics of definiteness and indefiniteness proposed in the previous sections and (ii) minimally change the denotation of the deictic particles, all while keeping to the core proposal from

Chapter 3, namely that deictic particles introduce an index. In doing so, this section thus provides further evidence for the account of indefiniteness and definiteness proposed in the previous sections of the present chapter.

5.3.2 A unified analysis of deictic particles

As currently stated, the entries for (270) and (271), repeated from above, are clearly not compatible with demonstratives that take an indefinite phrase as their base.

(270) Denotation of deictic particle *chi'* (repeated from Chapter 3)
$[\![chi'_i]\!]^g = \lambda x\text{: } x = \text{g}(i).\ x$

(271) Denotation of deictic particle *tik* (to be refined in Chapter 3)
$[\![tik_i]\!]^{g,c} = \lambda x\text{: } x = \text{g}(i) \wedge \text{CLOSE.TO.SPEAKER.OF.}c(x).\ x$

Under the current system, indefinites are generalized quantifiers, and will therefore denote a function of type $\langle\langle e,st\rangle, st\rangle$, which is not of the right type to compose with these denotations. To remediate this situation, all while keeping with the core proposal that deictic particles introduce indexical meaning, I sketch an analysis of deictic particles in which they denote not identity functions over individuals, but simple properties (of type $\langle e,st\rangle$), which I claim enter the derivation via Predicate Modification. The entries are provided in (272). In these entries, the variable is introduced both in the non-assertive and assertive components of the deictic particle (see Hanink 2018, 2021 for a similar set-up). Notice that the non-assertive component of the deictic particle also requires *existence*: the index must map to an entity in the range of the assignment function, one which is "a part of" the situation with which the DP is interpreted (the part relation is indicated with "$\leq$", see discussion in Kratzer 2021).

(272) Revised denotation for deictic particles

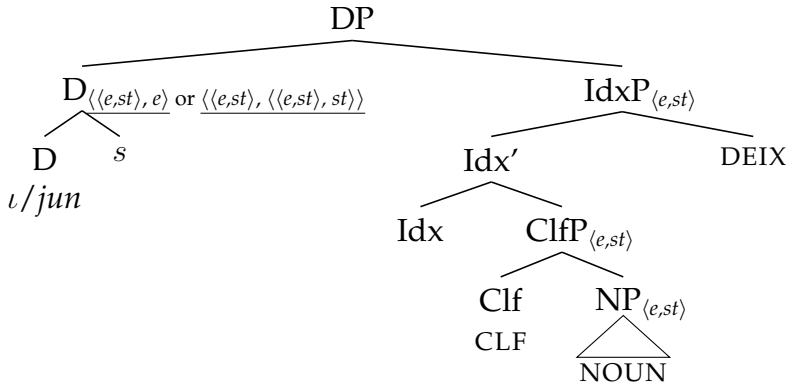
a. $[\![\ chi'_i\ ]\!]^g = \lambda x.\ \lambda s\text{: } \exists y[y \leq s \wedge y = g(i)].\ x = g(i)$

b. $[\![\ tik_i\ ]\!]^g = \lambda x.\ \lambda s\text{: } \exists y[y \leq s \wedge y = g(i) \wedge \text{CLOSE.TO.SPEAKER}(y)(s)].\ x = g(i)$

In (272), notice that the main proposal from Chapter 3 about the semantic contribution of deictic particles is preserved: deictic particles introduce an index and add the information that the nominal predicate is identical to the value for this index in the range of the assignment function.

As for the syntactic position of the deictic particle, I maintain that they sit in the specifier of a position dedicated to indexical material within the DP, which I label "IdxP" following previous work. Following Hanink (2018, 2021) and Ahn (2019), I propose that this projection is specifically located between the NP and DP, where D^0 is headed by ι and indefinite determiners, which both select arguments of type $\langle e,st\rangle$:

(273)

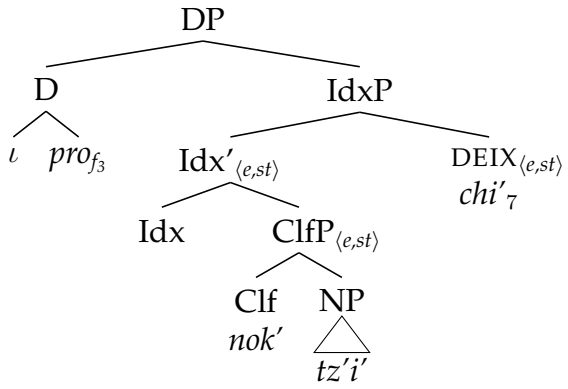


Given this syntax, the revised denotation for deictic particles can now compose with both definite and indefinite expressions alike, delivering meanings which seem to, at least at first glance, concord with the semantic contribution of both types of demonstrative constructions. Specifically, I propose that the deictic particle composes with the output of the classifier phrase via Predicate Modification (Heim and Kratzer 1998), the result of which can subsequently combine either with ι or with an indefinite quantifier, yielding either an entity (e) or a function of generalized quantifier type ($\langle\langle e,st\rangle, st\rangle$).

To provide concrete examples, consider first a derivation for a demonstrative with a definite base:

(274) a. *nok' tz'i' chi'*
nok' tz'i' chi'
CLF dog DEIX
'that dog'

b. Demonstrative with definite base:



c. $[\![\text{IdxP}]\!]^g =$
$\lambda x.\ \lambda s\colon |\{y\colon \text{DOG}(y)(s)\}| = 1 \wedge \exists z[z \leq s \wedge z = g(7)].\ \text{DOG}(x)(s) \wedge x = g(7)$

d. $[\![\text{DP}]\!]^g$ is defined only if
$\underline{|\{y\colon \text{DOG}(y)([\![pro_{f_3}]\!]^g)\}| = 1} \wedge \exists z[z \leq [\![pro_{f_3}]\!]^g \wedge z = g(7)]$
If defined: $\iota x[\text{DOG}(x)([\![pro_{f_3}]\!]^g) \wedge x = g(7)]$

The above derivation captures the uniqueness presupposition of demonstratives with definite bases (underlined), thereby accounting for empirical observations regarding the scope of the uniqueness presupposition discussed in section 3.3.2.1. Because the noun classifier introduces a uniqueness condition relative to a familiar situation, the DP in (274-d) is only defined if there is a unique satisfier of the NP in a situation familiar to all discourse participants. This explains the data seen in (116)–(117) from Chapter 3 and (269) from this section, where it was shown that classifier+nominal+deixis expressions trigger uniqueness presuppositions (or a maximality presupposition in the case of plural expressions). Furthermore, it follows from the logic of section 5.1.2.2 that both the noun classifier and deictic particle should be *obligatory* (assuming that a DP with a deictic particle competes with a DP without a deictic particle). The reason is that both introduce

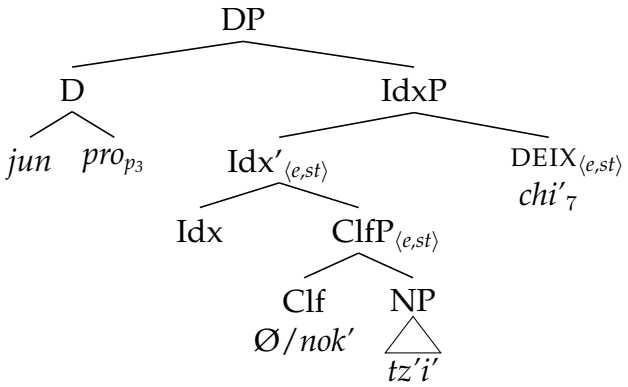
non-assertive meaning, which when combined with a familiar situation pronoun, should be obligatory to encode.

Also notice that since there can be at most one entity valued by a given index in the assignment function, it follows that this index will always map to the same individual as the one picked out in the assertive component. The result is thus the same as in Chapter 3: a decomposed strong definite description, which triggers both (i) a presupposition that there is a unique entity satisfying the NP in the situation and (ii) a presupposition that this unique entity is equivalent to the value for an index in the assignment function. I propose that for the entity to be identified by an index in a situation which is familiar to all discourse participants, it must be the case that the entity is *familiar* to the discourse participants.

Now consider the derivation for a demonstrative with an indefinite base in (275-b). There are two options, which depend on whether a noun classifier surfaces. That is, as was the case with indefinite expressions in the previous sections, I assume that Ø denotes an identity function and is a competitor with noun classifiers for *Maximize Presupposition*; see (229) above.

(275) a. *jun tz'i' chi' / jun nok' tz'i' chi'*
jun (nok') tz'i' chi'
INDF CLF dog DEIX
'that dog'

b. Demonstrative with indefinite base:



c. Option 1 (with Ø)
[[DP]] is defined only if
$\exists y[y \leq [\![pro_{p_3}]\!]^g \wedge y = g(7)]$
If defined: $\lambda Q.\ \lambda s.\ \exists x[[\text{DOG}(x)([\![pro_{p_3}]\!]^g) \wedge x = g(7)] \wedge Q(x)(s)]$

d. Option 2 (with Clf)
[[DP]] is defined only if
$|\{y\text{: DOG}(y)([\![pro_{p_3}]\!]^g)\}| = 1 \wedge \exists z[z \leq [\![pro_{p_3}]\!]^g \wedge z = g(7)]$
If defined: $\lambda Q.\ \lambda s.\ \exists x[[\text{DOG}(x)([\![pro_{p_3}]\!]^g) \wedge x = g(7)] \wedge Q(x)(s)]$

As in the previous sections, I assume that the lack of uniqueness presupposition (in the pragmatic sense) follows from the fact that indefinite determiners select for a private situation pronoun. Extending the analysis of *Maximize Presupposition* in section 5.1.2.2, we can also explain why deictic particles are not obligatory with indefinite expressions: since the presuppositional content of (275-d) involves a private situation pronoun, the proposition described by the presupposition should not be in the Common Ground. Again assuming that deictic particles compete with null morphemes for a lexical version of *Maximize Presupposition* (Percus 2006), the pragmatic principle will thus not enforce the presence of either the noun classifier or the deictic particle.

Before concluding, I highlight two consequences of this theory of deictic particles. A first consequence has to do with the status of the deictic particles as expressions which introduce indices. Like any other such expression, we predict that it should be possible for the index of the indefinite demonstrative to be bound by a higher quantifier. While more work is needed to determine what the constraints on binding indefinite demonstratives are, there is evidence that this prediction is borne out. Consider an "indefinite" version of the donkey sentence with its context:

(276) **Context:** Assume that people who own cows always love most the cow that tends to lead the herd.

Masanil heb' anima' ay swakax, ay jun nok' wakax chi' tzschamk'olej heb' t'a skal nok' mas chi'.

Masanil heb' anima' ay s-wakax, ay jun **nok'** wakax ??**(chi')**
all PL people EXT A3-cow, EXT INDF CLF A3-cow DEIX
tz-s-chamk'ol-ej heb' t'a s-kal nok' mas chi'.
IPFV-A3-love-DTV PRON.PL PREP A3-among CLF.PRON more DEIX

'For every person who owns cows, there's a cow that they love more than the others.'

In (276), the indefinite demonstrative covaries with respect to each cow-owning person. Since each person only has one cow which they prefer, maximality does not hold, and an indefinite demonstrative is used instead. Interestingly, speakers indicate a strong preference for an indefinite demonstrative in this case.

Second, it is important to signal that the current theory of deictic particles implies that *indices* are not what immediately contribute the familiarity presupposition of strong definites. Instead, the "familiarity presupposition" arises as a result of indexing a DP whose domain is delimited by a familiar situation pronoun. The idea is that for the entity to be in the range of the assignment function in a situation picked up by a familiar situation pronoun, then it must be the case that this entity is also familiar to all discourse participants. However, in DIBs, deictic particles, and thus an index, can be introduced even when the referent is not anaphorically identifiable. One way to implement this view in the current system would be to say that deictic particles introduce indices whose value can be felicitously unknown to some of the discourse participants, or to use the previous terminology, indices whose value is "private". This potentially provides us with the means to unify the contribution of "demonstratives" as indexical expressions, which tend to exhibit a wide range of uses across languages (see e.g. Doran and Ward 2019), including uses which involve first mentions of referents in discourse.

While I leave for future work a full-fledged investigation of the occurrence of deictic particles within indefinite expressions, it is clear that indefinite expressions carrying de-

ictic particles contribute additional information compared to their indefinite counterparts without a deictic particle. For instance, as discussed in Chapter 2, indefinite demonstratives are the default form used for exophoric reference (also called "deictic" uses in the literature). In particular, when it is clearly established that the speaker is drawing the attention of their addressees to an entity visible to them, an indefinite demonstrative must be used:

(277) **Context:** There's only one ant on the shoulder of the addressee, but the addressee hasn't seen the ant. The speaker points to the ant and tells the addressee:

a. *¡Hilnab' jun sanich tik t'a hajenhjab'!*
¡H-ilnab' [**jun** sanich **tik**] t'a ha-jenhjab'!
A2S-look INDF ant DEIX PREP A2S-shoulder
'Look at this ant on your shoulder!'

b. # ¡H-ilnab' [**nok'** sanich **tik**] t'a ha-jenhjab'!
A2S-look CLF ant DEIX PREP A2S-shoulder

Moreover, when appearing with nominal expressions that introduce an entity for the first time in discourse, the presence of a deictic particle seems to be tied to what Ionin (2006) describes as *noteworthiness*, which is also the contribution of indefinite uses of the demonstrative *this* in English (see also Prince 1981 and Maclaran 1982). What Ionin (2006) shows is that the felicitous use of *this*-indefinites in English is conditional on the referent exhibiting some noteworthy property, which will often be explicitly mentioned in the immediate discourse or as a modifier of this indefinite. This is shown in (278) and (279)

(278) Examples from Maclaran 1982, p 88

a. He put {✓a/#this} 31 cent stamp on the envelope, so he must want it to go airmail.

b. He put {✓a/✓this} 31 cent stamp on the envelope, and only realized later that it was worth a fortune because it was unperforated.

(279) a. I found this blue apple on my plate!

b. #I found this apple on my plate!

Indefinite uses of *this* are only appropriate if the speaker intends to say something noteworthy about the referent of the indefinite expression. In (278-b), *this* is appropriate be-

cause the stamp in question is worth a fortune. In (279-a), it is appropriate because apples are not normally blue and the apple in question thus exhibits a noteworthy property.

Similar facts seem to hold with indefinite uses of demonstratives in Chuj. Consider, for instance, the sentences following the utterance from Hopkins (2021) in (268). In the sentences following the DIB, the speaker intends to convey a specific property of the referent of *jun winh icham chi'* 'this old man', namely that he has many sons. The rest of the story centres around one of the man's sons, and so this information could be said to be "noteworthy" for the rest of the story.

(280) Narrative sequence from Hopkins 2021: 45.

a. *Komo ha t'ay pekatax, ay jun winh icham chi'.*
Komo ha t'ay pekatax, ay [***(jun)** winh icham chi']
DISC FOC PREP before, EXT INDF CLF man DEIX
'Once upon a time, there was an old man...'

b. *Ay juntzanh yuninal winh.*
Ay juntzanh y-unin-al winh.
EXT PL.INDF A3-child-INAL CLF.PRON
'He had sons.'

c. *Tzijtum yuninal winh chi'.*
Tzijtum y-unin-al winh chi'.
Many A3-son-INAL CLF.PRON DEIX
'He had *many* sons.'

Again, the goal of this section was to sketch a refined analysis of deictic particles, one which can explain why deictic particles can combine with both definite and indefinite bases. While the analysis of strong definites with deictic particles from Chapter 3 essentially remains intact, I leave for future work an exploration of the pragmatic and semantic contribution of deictic particles in combination with indefinite expressions.

5.4 Summary of the proposals

In the previous two chapters and the current, I provided an analysis of the composition of the Chuj DP, with a focus on the pieces that play a role in the composition of defi-

niteness and indefiniteness. These pieces are summarized below. The final entries for each of the functional items in Table 5.3 are also provided below. Finally, I also provide the Revised Appropriateness Condition, which determines the condition on the felicitous use of familiar versus private variables, respectively c-selected by ι (282) and indefinite determiners (283).

Table 5.3: Singular DP configurations in Chuj

	Configuration	Output in Chuj	Rough translation
①	ι + CLF + NP	*nok' tz'i'*	'the dog'
②	ι + CLF + NP + DEIX_i	*nok' tz'i' chi'*	'the/that dog' (anaphoric)
③	ι + CLF + pro_i	*nok'*	'it'
④	ι + CLF + <NP>	*nok'*	'it'
⑤	$\exists$ + NP	*jun tz'i'*	'a dog'
⑥	$\exists$ + CLF + NP	*jun nok' tz'i'*	'a (certain) dog'
⑦	$\exists$ + NP + DEIX_i	*jun tz'i' chi'*	'that dog'
⑧	$\exists$ + CLF + NP + DEIX_i	*jun nok' tz'i' chi'*	'that dog' (exophoric)

(281) **Noun classifier** (with "class" presupposition)
$[\![\text{CLF}]\!] = \lambda P_{\langle e,\langle s,t\rangle\rangle}.\lambda x.\ \lambda s\text{: } |\{y\text{: } P(y)(s)\}| = 1 \wedge \{y\text{: } P(y)(s)\} \subseteq \{z\text{: } \text{CLASS}(z)\}.\ P(x)(s)$

(282) ***Iota***
$[\![\iota]\!]^g = \lambda s.\lambda P_{\langle e,\langle s,t\rangle\rangle}.\iota x[P(x)(s)]$ $[\iota \rightarrow pro_f]$

(283) **Singular indefinite determiner (∃)**
$[\![jun]\!] = \lambda s.\ \lambda P_{\langle e,\langle s,t\rangle\rangle}.\lambda Q_{\langle e,\langle s,t\rangle\rangle}.\lambda s'.\ \exists x[P(x)(s) \wedge Q(x)(s')]$ $[jun \rightarrow pro_p]$

(284) **Deictic particles**

a. $[\![\ chi'_i\ ]\!]^g = \lambda x.\ \lambda s\text{: } \exists y[y \leq s \wedge y = g(i)].\ x = g(i)$

b. $[\![\ tik_i\ ]\!]^g = \lambda x.\ \lambda s\text{: } \exists y[y \leq s \wedge y = g(i) \wedge \text{CLOSE.TO.SPEAKER}(x)(s)].\ x = g(i)$

(285) **Null predicative *pro* in anaphoric pronouns**
$[\![\ pro_i\ ]\!]^g = \lambda x.\ x = g(i)$

(286) **(Revised) Appropriateness Condition**
A context C is *appropriate* for an LF ϕ only if C determines a variable assignment g_C whose domain includes every index i which has a free occurrence in ϕ, and either:

a. for every pro_{f_i} or trace, the value for $g_C(i)$ must be known to all discourse participants; or
b. for every pro_{p_i}, the value for $g_C(i)$ does not have to be known to all discourse participants.

The overall result is a unified analysis of the pieces involved in the composition of definite and indefinite nominal expressions in Chuj, which I have argued can account for the subtle semantic differences that arise as a consequence of different combinations of the functional items in (281)–(285). The core proposals from the last two chapters are summarized below:

(287) Summary of proposals in Chapters 4 and 5

a. Indefinite+classifier combinations yield singleton indefinites, explaining the described specific indefinite interpretations associated with this combination of functional items (§4.2).
b. There are two kinds of situation variables: private situation variables and familiar situation variables (§4.3).
c. The kind of variable c-selected by a lexical item (ι versus *jun* indefinites) affects the breadth of the semantic presuppositions contributed by the noun classifier. When combined with ι, the uniqueness presupposition surfaces as a pragmatic presupposition (§5.1). When combined with *jun*, the result is a singleton indefinite (§4.3).
d. The unified theory of noun classifiers can explain (i) their necessity in definite descriptions and also as third person pronouns and (ii) the fact that they can optionally co-occur with specific interpretations of indefinites (§5.1.2).

e. Deictic particles introduce a syntactically-represented index, and denote the property of being identical to this index (§5.3). While deictic particles are obligatory with strong definites, they are compatible with both definite and indefinite expressions, accounting for definite and indefinite uses of demonstratives in Chuj.

More generally, Chuj provides evidence for a rich decomposition of indefinite and definite nominal expressions, aligning with much recent work advocating for the decomposition of DP in some form or another (e.g. Leu 2008; Simonenko 2014; Coppock and Beaver 2015; Hanink 2018; Ahn 2019). By doing so, Chuj contributes important insight into crosslinguistic theories of the syntax and semantics of definite, indefinite, pronominal, and demonstrative nominal expressions.

Having focused on the internal composition of nominal expressions in the previous three chapters, the next and final chapter in the body of this thesis zooms out of the internal syntax of nominal expressions and into the distribution of covalued nominal expressions within sentences.

Chapter 6

Binding and anti-cataphora in Chuj and Ch'ol

Patterns of nominal binding and coreference have long been a rich area of research in linguistics (e.g., Lees and Klima 1963; Ross 1967; Langacker 1969; Jackendoff 1972; Reinhart 1976, 1983; Chomsky 1981, 1986; Lebeaux 1984; Reuland and Koster 1991; Pollard and Sag 1992; Grodzinsky and Reinhart 1993; Reinhart and Reuland 1993), a tradition that has persisted in recent work in generative syntax (e.g., Hornstein 2001, 2007; Reuland 2001, 2011; Kayne 2002; Zwart 2002; Safir 2004, 2008, 2014; Büring 2005; Schlenker 2005b; Davis 2009; Hicks 2009; Kratzer 2009; Rooryck and vanden Wyngaerd 2011, Drummond et al. 2011; Despić 2013, 2015; Bruening 2014, Bruening 2021; Ahn 2015; Charnavel and Sportiche 2016). While a wide range of proposals have been put forth to capture the distribution of bound and coreferential nominals, one fact stands out: the empirical generalizations about the distribution of covalued nominal expressions within sentences have remained remarkably constant throughout the years; since Chomsky 1981 and Reinhart 1983, most authors still aim to derive a version of the binding conditions in (288).

(288) **The binding conditions**
Condition A – An anaphor must be locally bound
Condition B – A pronoun must be locally free
Condition C – An R-expression must be free

What is more, most of the literature on the binding conditions has assumed a definition of 'binding' that is sensitive to some form of command relation (regardless of whether generalizations about binding are explained trans-derivationally, as in Chomsky 1981, Reinhart 1983 and Bruening 2021, or via a set of principles or derivationally through movement operations or Agree, as in Hornstein 2001 and Reuland 2011). A classic definition is *c-command*, provided in (289), which has been implemented in what is probably the majority of works since Reinhart 1976, 1983.[1]

(289) Classic definition of binding
NP_A binds NP_B iff (i) NP_A c-commands NP_B and (ii) NP_A and NP_B are covalued.

It is striking that the conditions in (288) seem to hold across a great many languages, with little space for linguistic variation. As indicated by Grodzinsky and Reinhart (1993) and Reuland (2010, 2011), this suggests that whatever properties of natural language underly the binding conditions must be universal. This suggestion is further supported by a vast array of experimental studies in language acquisition and with aphasic patients, which all coincide in presenting robust evidence for the innateness of the binding conditions (see e.g. Grimshaw and Rosen 1990, Grodzinsky and Reinhart 1993, Grodzinsky et al. 1993).

In this chapter, I show that Chuj exhibits a number of properties that, at first glance, appear to challenge the view that the binding conditions are universal. In presenting the challenges from Chuj, I will also provide data from Ch'ol, another Mayan language

[1] I use the term "covalue" as a neutral, descriptive term, to mean that two nominals are either (i) in a syntactic binding relation (one nominal c-commands and is covalued with the other) or (ii) are both free (no c-command relation), but covalued (see e.g. Heim 2007 for similar use of this term). Similar terms that have been used are 'codetermination' (Heim 1998) and 'coconstrual' (Safir 2008). I remain agnostic as to whether (i) must involve bound variable anaphora in the semantics (see e.g. discussion in Reinhart 1983, Safir 2008 and Bruening 2021).

from the Western Branch of Mayan languages (see Table 2.1), which differs from Chuj minimally in ways that will be relevant for the described challenges.

The core of the puzzle is as follows: while the binding conditions behave as expected in Ch'ol, they appear to play little role in regulating the distribution of covalued nominal expressions in Chuj, as also reported for the closely-related language Popti' (Craig 1977, Hoekstra 1989; Woolford 1991; Trechsel 1995; Aissen 2000). Most notably, though the Ch'ol and Chuj sentences in (290) and (291) exhibit the same word order on the surface, we will see compelling evidence that the R-expression in Chuj sentences like (291) is not the subject, but the possessor, the subject being instantiated as a covert pronoun. In other words, the right parse for Ch'ol is (292-a), as expected, whereas the right parse for Chuj is, unexpectedly, (292-b).

(290) **Ch'ol**
Tyi ichoño iwakax ajAna.
Tyi i-choñ-o i-wakax aj-Ana.
PFV A3-sell-TV A3-cow CLF-Ana
'Ana$_1$ sold her$_1$ cow.'

(291) **Chuj**
Ixschonh swakax ix Ana.
Ix-s-chonh s-wakax ix Ana.
PFV-A3-sell A3-cow CLF Ana
Lit: 'She$_1$ sold Ana$_1$'s cow.'

(292) a. sold [$_{\text{OBJ}}$ cow [$_{\text{POSS}}$ *pro*$_1$]] [$_{\text{SUBJ}}$ Ana$_1$] = Ch'ol (290)
b. sold [$_{\text{OBJ}}$ cow [$_{\text{POSS}}$ Ana$_1$]] [$_{\text{SUBJ}}$ *pro*$_1$] = Chuj (291)

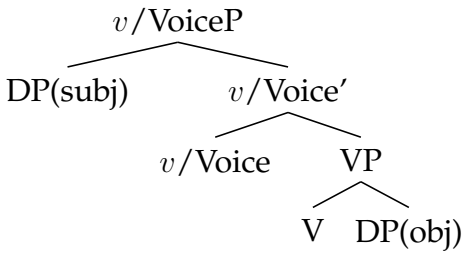
All else being equal, the Chuj configuration in (292-b) looks like a violation of Condition C, per the literal translation under (291). These data are thus puzzling if the binding conditions are universal.

After presenting in section 6.1 the Chuj data that are problematic from the perspective of the binding conditions, the goal of this chapter will be to argue that it is possible to account for these puzzling data without having to deny the universality of the binding conditions. In section 6.2, I argue that in every configuration in which the binding conditions are inoperative in Chuj, there are no c-command relations between the relevant covalued expressions, and therefore no violations of the binding conditions. For instance,

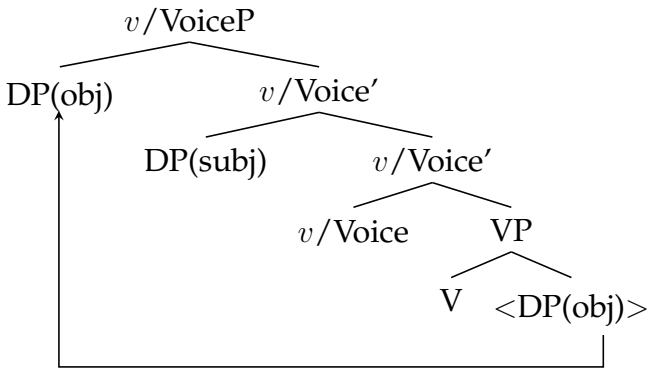
I argue that while the subject c-commands the possessor of the object in the Ch'ol sentence in (292-a), it does not in Chuj (292-b).

Central to the proposal will be the adoption of the 'high-absolutive' approach to the *Ergative Extraction Constraint* (EEC), which describes a general ban on the A'-extraction of ergative subjects in a subset of Mayan languages (see Aissen 2017, table 30.3). In the high-absolutive approach, the EEC arises because the absolutive object systematically raises to a position above the ergative subject (Coon et al. 2014, Assmann et al. 2015; Coon et al. 2021). Proponents of this analysis propose that this is the case in Chuj, but not in Ch'ol.[2]

(293) Low-absolutive syntax (Ch'ol)



(294) High-absolutive syntax (Chuj)



I argue that the binding differences between Chuj and Ch'ol are another, pervasive correlate of the independently proposed need for high-absolutive syntax in Chuj. In particular, object raising bleeds c-command relations between the subject and nominals contained inside the object, which obviates the application of the binding conditions. The binding conditions do not apply in (291), because the two covalued nominals are both free. In

[2]Following Coon (2019), I assume a single bundled v/Voice head in Chuj and Ch'ol.

Ch'ol (290), on the other hand, object raising does not happen. Since the two nominals are in a c-command relation, Condition C applies to prohibit the R-expression from occurring as the possessor.

Developing observations in Royer 2021c and Coon et al. 2021, a survey of several languages indicates that similar facts are repeated across the Mayan language family. Mayan languages that exhibit the EEC, like Chuj, allow an R-expression contained within the object to be covalued with the subject, whereas those that do not exhibit the EEC, show Condition C effects in such data. The resulting typological generalization will therefore provide strong empirical support for high-absolutive approaches to the EEC.

There is an additional layer of puzzle regarding the distribution of covalued nominals in Ch'ol and Chuj. In addition to the apparent inapplicability of Condition C in Chuj, there appears to be a *linear precedence* constraint on the distribution of certain covalued nominals. For instance, the R-expression in the Chuj sentence seen in (291) *obligatorily* appears in the possessor, and not the subject (which in this case is a null pronoun). Based on (i) a discussion of clear patterns of free, nominal covaluation in both Ch'ol and Chuj, and (ii) a detailed discussion of reflexive sentences in Chuj, which I show abide by the binding conditions, I argue in sections 6.3 and 6.4 that the linear precedence constraint only applies to free expressions. Expressions that are bound under c-command, on the other hand, are subject to the binding conditions. More precisely, I show that the two-fold generalization in (295) describes the distribution of covalued nominal expressions in both languages:

(295) **Generalization about covalued expressions in Chuj and Ch'ol**

a. If a nominal is bound under c-command, it is subject to structurally-determined binding conditions like (288) (linear precedence is irrelevant).

b. If a free pronoun is covalued with an R-expression, the R-expression must linearly precede the free pronoun (linear precedence is relevant).

The generalization in (295-a) is completely expected given standard approaches to syntactic binding. Also expected is the fact that bound and free pronouns are treated differently, since most accounts of syntactic binding draw a distinction between the two. Less expected, however, is (295-b): linear precedence plays a central role in determining the realization of free nominals. This is important, since it will allow us to unify the constraints on pronominalization in Ch'ol and Chuj, which at first glance appear to be entirely disparate (see (292-a) vs (292-b)). Both languages are subject to the binding conditions (295-a), and both ban cataphora with free pronouns (295-b). The only difference between the two languages is that Chuj exhibits high-absolutive syntax (294), with the effect that Chuj seems to be more frequently affected by the constraint in (295-b) than Ch'ol is.

Given (295), I turn in section 6.5 to a discussion of how the difference between free and bound nominals could be formalized. Based on movement theories of syntactic binding (e.g. Hornstein 2001; Kayne 2002; Zwart 2002), the main suggestion will be that while constraints pertaining to the binding conditions could be derived entirely in terms of how PF interprets movement chains, constraints on free nominals could result from how PF treats identical, externally merged lexical items within a derivation.

Finally, section 6.6 concludes with a brief discussion of a theoretical consequence posed by the Mayan patterns of nominal covaluation discussed in this chapter. In particular, I argue that the anti-cataphora constraint in (295-b) offers a strong empirical argument for the necessity of indices in syntax. This casts doubt on Inclusiveness (Chomsky 1995, Chomsky 2001), and more generally on recent theories of syntactic binding, which assume that indices cannot form part of Universal Grammar (e.g., Reuland 2001, 2011).

Before we continue, I specify that this chapter concentrates almost exclusively on the conditions that govern the *syntactic distribution* of covalued nominal expressions in Ch'ol and Chuj. Throughout the chapter, I will frequently use the term 'syntactic binding', which I take to be the syntactic conditions that govern the *distribution* of covalued nominals. For the better part of the chapter (with the exception of section 6.5, where it will

become necessary to discuss semantic binding), I remain agnostic as to whether syntactic binding and 'semantic' binding—the latter of which I take to concern the conditions on bound variable anaphora—exhibit a one-to-one correspondence. While much work in formal linguistics assumes a one-to-one correspondence between the two (e.g., Reinhart 1983; Heim and Kratzer 1998; Hornstein 2001, Büring 2005, Heim 2007), recent work has brought into question whether the two can consistently be subsumed under a single phenomenon (see e.g., Safir 2008; Barker 2012; and Bruening 2021 for relevant work).[3] Given the limited understanding of semantic binding in Mayan, I will mostly not engage with this question in this chapter, and will therefore avoid diagnostics from semantic binding (e.g., sloppy identity under ellipsis) as a test for syntactic binding.

6.1 The puzzle: Apparent binding violations in Chuj, but not in Ch'ol

In this section, I discuss data from Chuj which appear to, at least at first glance, challenge the universality of the binding conditions in (288). The issue is that the binding conditions appear to be largely irrelevant for the realization of covalued nominals in this language. Instead, only *linear precedence* seems to matter. I also show that the Chuj data stand in stark contrast with comparable data from Ch'ol, where the binding conditions apply as expected without reference to linear order. I first provide brief relevant background on Chuj and Ch'ol in sections 6.1.1 and 6.1.2, turning to the relevant empirical data in section 6.1.3. Section 6.1.4 then summarizes the findings of this section, and presents the questions to be addressed in the rest of the chapter.

[3]These authors show that the conditions on variable binding and the conditions on the pronunciation of covalued nominals do not always match. For instance, while the possessor of a subject can bind a direct object in English (*Every boy$_1$'s mother loves him$_1$*), the possessor does not have to be realized as an R-expression (*His$_1$ mother loves John$_1$*). Therefore, semantic binding does not necessarily imply syntactic binding. It is unclear to me at this point whether this is also the case in Mayan languages.

6.1.1 Background

While the previous chapters of this thesis have mainly focused on Chuj, this chapter also provides extensive data on Ch'ol, a related language of the Western branch of Mayan languages. Ch'ol is currently spoken by approximately 252,000 speakers in a few states of Southern Mexico, including primarily Chiapas, Tabasco, and Campeche; see Figure 2.1 for a relevant map. The Ch'ol data in this thesis come from original fieldwork conducted with one speaker of the Tila dialect.[4]

The grammatical properties of Ch'ol closely resemble those described for Chuj in Chapter 2, as well as the more general grammatical properties shared within the Mayan language family (England 2001, Coon 2016a, Aissen et al. 2017). Both are head-marking, ergative-absolutive languages, and both exhibit verb-initial word order in discourse-neutral contexts. Like Chuj, Ch'ol also exhibits basic VOS word order, though VSO is sometimes permitted under special circumstances in both languages (Coon 2010b, Clemens and Coon 2018, Little 2020b). Basic VOS sentences are provided below for comparison:

(296) *Ixschi' nok' mis nok' tz'i'.*
Ix-**s**-chi' [$_{OBJ}$ nok' mis] [$_{SUBJ}$ nok' tz'i'].
PFV-A3-bite CLF cat CLF dog
'The dog bit the cat.' (Chuj)

(297) *Tyi ik'uxu mis jiñi ts'i'.*
Tyi **i**-k'ux-u [$_{OBJ}$ mis] [$_{SUBJ}$ jiñi ts'i'].
PFV A3-bite-TV cat DET dog
'The dog bit the/a cat.' (Ch'ol: Coon 2010a, 43)

As in Chuj and other Mayan languages, "Set A" prefixes in Ch'ol are used to cross-reference ergative subjects and possessors, whereas "Set B" morphemes are used for absolutive arguments. Also like in Chuj, there is no overt instantiation of third person Set B morphology, and therefore third person Set B is not represented in glosses. Examples with overt Set B morphemes in both Chuj and Ch'ol are provided below:

[4]For grammatical overviews of Ch'ol, see Coon 2010a, Vázquez Álvarez 2011, and Little 2020b.

(298) *Ixinachela'.*
Ix-**in**-**a**-chel-a'.
PFV-B1S-A2S-hug-TV
'You hugged me.' (Chuj)

(299) *Tyi amek'eyoñ.*
Tyi **a**-mek'-e-**yoñ**.
PFV A2-hug-TV-B1
'You hugged me.' (Ch'ol)

Notice that the Set B morphemes appear in different positions in Chuj and Ch'ol. While Set B morphemes follow tense-aspect marking and precede the root in Chuj (298), they follow the stem in Ch'ol (299).[5] This will be relevant for the discussion of the analysis in section 6.2.

Possessive constructions will play an important role in this chapter. As discussed in Chapter 2, possessors appear post-nominally in Chuj and trigger Set A (ergative/possessive) morphology as a prefix on the possessee. The same is true in Ch'ol. Notice, for instance, the identical Set A prefixes on the Chuj and Ch'ol verbs in (296) and (297).

(300) *smis ix Ana*
s-mis ix Ana
A3-cat CLF Ana
'Ana's cat' (Chuj)

(301) *imis ajAna*
i-mis aj-Ana
A3-cat CLF-Ana
'Ana's cat' (Ch'ol)

Since this chapter is about patterns of nominal covaluation between R-expressions and pronouns, some background on the status of pronouns in Chuj and Ch'ol is in order. Before providing the puzzle in section 6.1.3, we therefore turn to a brief description of pronominals in both languages.

[5] I follow the norms of each language in terms of where and how a word boundary is written internal to the stem. For instance, while perfective marking is written as a separate word in Ch'ol, it is not in Chuj. Note, however, that it is not clear that these orthographic conventions reflect a meaningful difference in the morphosyntax of these languages.

6.1.2 Pronominal expressions in Chuj and Ch'ol

Most Mayan languages are robustly *pro*-drop and lack overt non-emphatic third person pronouns.[6] Ch'ol is no exception. In appropriate contexts, both the subject and object must be dropped:

(302) *Tyi imänä.*
Tyi i-män-ä [$_{OBJ}$ *pro*] [$_{SUBJ}$ *pro*].
PFV A3-buy-TV PRON PRON
'She bought it.' (Ch'ol)

As already discussed in Chapters 2 and 3 of this thesis, however, Chuj diverges considerably from Ch'ol and most other Mayan languages when it comes to pronouns: it is not *pro*-drop and it features non-emphatic third person pronouns. Specifically, like other languages of the Q'anjob'alan sub-branch of Mayan languages (Craig 1986b; Zavala 2000; Hopkins 2012b; Mateo Toledo 2017), Chuj's noun classifiers function as third person non-emphatic pronouns in the language. Also shown in (303-b), under most circumstances, classifier pronouns cannot be dropped:

(303) Chuj
a. *Ixsman jun te' onh winh winak.*
Ix-s-man [$_{OBJ}$ jun **te'** onh] [$_{SUBJ}$ **winh** winak].
PFV-A3-buy INDF CLF avocado CLF man
'The man bought an avocado.'

b. *Ixslo'an te' winh.*
Ix-s-lo'-an [$_{OBJ}$ ***(te')**] [$_{SUBJ}$ ***(winh)**].
PFV-A3-eat.sweet-CON CLF.PRON CLF.PRON
'And then he ate it.' .

There are special circumstances, however, where the use of a classifier pronoun becomes illicit. Consider (304-a); this sentence can only be interpreted with disjoint reference—

[6] Vázquez Álvarez (2011: 153) lists *jiñ* as a third person emphatic pronoun in Ch'ol. As he mentions, overt personal pronouns are generally used emphatically, and tend to appear preverbally in dedicated topic and focus positions (Aissen 1992).

Xun must have spoken with another man's mother. To arrive at a joint-reference reading, a null pronoun (*pro*) must instead be used, as in (304-b).

(304) Chuj

a. *Ixlolon waj Xun yet' ix snun winh.*
Ix-lolon [$_{SUBJ}$ waj Xun] [$_{PP}$ y-et' ix s-nun **winh**].
PFV-speak CLF Xun A3-with CLF A3-mother CLF.PRON
'Xun$_1$ spoke with his$_{2/*1}$ mother.'

b. *Ixlolon waj Xun yet' ix snun.*
Ix-lolon [$_{SUBJ}$ waj Xun] [$_{PP}$ y-et' ix s-nun ***pro***].
PFV-speak CLF Xun A3-with CLF A3-mother PRON
'Xun$_1$ spoke with his$_1$ mother.'

The use of *pro* in Chuj is highly restricted, as also observed for Popti' by Craig (1977). The basic generalization about *pro* can be stated as follows:

(305) *Generalization about 'pro' in Chuj* (adapted from Craig 1977 and Trechsel 1995)
The null pronoun *pro* must be covalued with another nominal expression inside the minimal CP in which it occurs (where relative clauses do not count as separate CPs).

In example (304-b), notice that the condition in (305) is met, since *pro* is covalued with the R-expression *Xun* inside the same minimal CP. An example for which the condition in (305) is not met is provided in (306). Here, the subject of the complement clause is covalued with the subject of the matrix clause. Since the two expressions appear in separate "minimal CPs", the subject of the complement clause cannot be realized as *pro* and a classifier pronoun must instead be used:

(306) *Ixyal winh winak to ixb'at winh.*
Ix-y-al winh winak [$_{CP}$ to ix-b'at {**winh/*pro**}].
PFV-A3-say CLF man COMP PFV-go CLF.PRON/PRON
'The man$_1$ said that he$_{1/2}$ left.' (Chuj)

As indicated in the translation, the sentence in (306) is ambiguous. The classifier pronoun can be covalued with *winh winak* 'the man' or could alternatively refer to another male individual. As predicted by the generalization in (305), the use of *pro* is not allowed.

In order to make progress on our understanding of patterns of nominal covaluation in Chuj, this chapter will mostly only focus on the distribution of *pro*. Based on a variety of data, I will show that sentences involving *pro* often appear to violate Condition C. In fact, *pro*'s distribution seems to be largely governed by the generalization in (307):[7]

(307) *Generalization about covalued expressions in Chuj*
If covalued expressions appear in the same minimal CP, the linearly first must be an R-expression (or an overt classifier pronoun), and the rest are realized as *pro*.

The evidence, which we immediately turn to in the next subsection, will come from various domains. Also note that though the current section focuses specifically on covaluation relations between the external and the internal argument, we will return to covaluation relations with oblique phrases in section 6.2.4, where contrasts between Chuj and Ch'ol will again be observed.

6.1.3 Puzzle: Linear precedence and binding in Chuj and Ch'ol

Here, I show evidence for the generalization in (307): linear precedence plays a fundamental role for the distribution of covalued nominals in Chuj, often in apparent violation

[7] Aissen (2000) formulates the domain of *pro* in the related language Popti' prosodically, as applying to intonational phrases:

(i) Condition on [*pro*] (Aissen 2000: 191)
The anaphor [*pro*] must be co-indexed with a nominal which precedes in within the same intonational phrase.

Aissen further shows that while complement clauses and topics form their own intonational phrase, relative clauses do not. As noted by an anonymous reviewer on related work, the prosodic generalization seems more adequate, since it can explain why relative clauses do not count as separate domains for *pro* (see also Royer 2022 on intonational phrases in Chuj). Note, though, that it is not crucial for the present purposes whether the conditions on *pro* apply at the syntactic or prosodic level. In section 6.5, I suggest that the linear precedence effects should be understood as a more general ban on backwards PF deletion (397). Under this view, the prosodic generalization would be equally (if not more) suitable.

of Condition C. We will also see that this is not the case in Ch'ol, where the binding conditions seem to apply as expected, irrespective of linear order. The evidence comes from data relevant to adverb placement options (§6.1.3.1), coordination (§6.1.3.2), relative clauses (§6.1.3.3), object A'-extraction (§6.1.3.4), and word order alternations (§6.1.3.5).

6.1.3.1 Adverb placement in 'extended reflexive' constructions

Following Aissen (1999), I refer to transitive sentences in which the external argument is covalued with the possessor of the internal argument as "extended reflexives". Examples of such constructions were already presented at the outset of this chapter, in (290) and (291). These examples are repeated below for convenience:

(308) Ch'ol extended reflexive
Tyi ichoño iwakax ajAna.
Tyi i-choñ-o i-wakax aj-Ana.
PFV A3-sell-TV A3-cow CLF-Ana
'Ana$_1$ sold her$_1$ cow.'

(309) Chuj extended reflexive
Ixschonh swakax ix Ana.
Ix-s-chonh s-wakax ix Ana.
PFV-A3-sell A3-cow CLF Ana
Lit: 'She$_1$ sold Ana$_1$'s cow.'

Since Ch'ol and Chuj are VOS languages with postnominal possessors, the right syntactic parse of the sentences in (308) and (309) is not immediately transparent from surface word order. That is, the R-expression *Ana* could be in subject position, as in (310-a). Alternatively, *Ana* could be in the possessor of the object, in which case the subject would be realized as *pro* (310-b).

(310) a. sold [$_{\text{OBJ}}$ cow [$_{\text{POSS}}$ Ø$_1$]] [$_{\text{SUBJ}}$ Ana$_1$] (lit: Ana$_1$ sold her$_1$ cow)
b. sold [$_{\text{OBJ}}$ cow [$_{\text{POSS}}$ Ana$_1$]] [$_{\text{SUBJ}}$ *pro*$_1$] (lit: She$_1$ sold Ana$_1$'s cow)

Assuming that external arguments c-command internal arguments, (310-b) looks like a classic violation of Condition C: a pronoun appears to c-command a covalued R-expression. Nevertheless, I will provide evidence that this is the right parse of the Chuj sentence in (309). At the same time, there is also evidence that the right parse in Ch'ol is the expected, Condition C-complying parse in (310-a). To diagnose the syntactic position of the

R-expression, I will primarily consider evidence from adverb placement options, which differ between Chuj and Ch'ol. We will also see evidence from coordinated sentences and contexts of A'-extraction in sections 6.1.3.2 and 6.1.3.4.

Before showing the relevant test examples, first note that in regular transitive sentences, some Ch'ol and Chuj adverbs can appear either between the object and the subject, as in (311-a) and (312-a), or after the subject, as in (311-b) and (312-b).[8]

(311) Ch'ol

a. *Tyi ichoko tyuñ abi jiñi alob.*
Tyi i-chok-o [$_{OBJ}$ tyuñ] **abi** [$_{SUBJ}$ jiñi alob].
PFV A3-throw-TV stone yesterday DET boy
'The boy threw the stone yesterday.' (Coon 2010a: p. 241)

b. *Tyi ichoko tyuñ jiñi alob abi.*
Tyi i-chok-o [$_{OBJ}$ tyuñ] [$_{SUBJ}$ jiñi alob] **abi** .
PFV A3-throw-TV stone DET boy yesterday
'The boy threw the stone yesterday.'

(312) Chuj

a. *Ixsb'o' tek junelxo waj Xun.*
Ix-s-b'o' [$_{OBJ}$ tek] **junelxo** [$_{SUBJ}$ waj Xun].
PFV-A3-make meal again CLF Xun
'Xun made the meal again.'

b. *Ixsb'o' tek waj Xun junelxo.*
Ix-s-b'o' [$_{OBJ}$ tek] [$_{SUBJ}$ waj Xun] **junelxo**.
PFV-A3-make meal CLF Xun again
'Xun made the meal again.'

Also note that adverbs cannot intervene between the possessor and the possessee in either language, as shown in the following examples:

(313) a. *Tyi k-chok-o i-tyuñ **abi** jiñi alob.
PFV A1-throw-TV A3-stone yesterday DET boy
Intended: 'I threw the boy's stone yesterday.' (Ch'ol)

b. *Ix-in-b'o' s-tek **junelxo** waj Xun.
PFV-A1S-make A3-meal again CLF Xun
Intended: 'I made Xun's meal again.' (Chuj)

[8] In Chuj, not all adverbs can appear between the object and the subject. Adverbs that can appear in that position include *junelxo* 'again', *junelnhej* 'at once', and *masanil* 'completely/all'.

However, when we consider extended reflexive constructions, the adverb placement options change in Chuj, but not in Ch'ol. First consider the following Ch'ol data which differ minimally from the sentences in (311), with the only difference being that the referent of the object is now formally possessed by the referent of the subject, as indicated by the presence of Set A possessive marking on the object:

(314) Ch'ol

a. *Tyi ichoko ityuñ abi jiñi alob.*
Tyi i-chok-o [$_{OBJ}$ **i**-tyuñ] **abi** [$_{SUBJ}$ jiñi alob].
PFV A3-throw-TV A3-stone yesterday DET boy
'The boy$_1$ threw his$_1$ stone yesterday.'

b. *Tyi ichoko ityuñ jiñi alob abi.*
Tyi i-chok-o [$_{OBJ}$ **i**-tyuñ] [$_{SUBJ}$ jiñi alob] **abi** .
PFV A3-throw-TV A3-stone DET boy yesterday
'The boy$_1$ threw his$_1$ stone yesterday.'

The adverb *abi* 'yesterday' can still intervene between *tyuñ* and *jiñi alob*, just like in (311). Since Ch'ol adverbs never intervene between objects and their possessor, as shown in (313-a), the data in (314) suggest that the R-expression *jiñi alob* is the subject, and that the possessor is a null pronoun:

(315) Proposed structure for Ch'ol (314)
threw [$_{OBJ}$ stone [$_{POSS}$ *pro*$_1$]] {yesterday} [$_{SUBJ}$ the boy$_1$] {yesterday}

Crucially, the pattern in (315) is expected given the binding conditions. If subjects c-command objects in Ch'ol, as has been proposed in independent work (see e.g., Clemens and Coon 2018), Condition C should block the R-expression from occurring as the possessor of the object. Also notice that linear precedence seems to be irrelevant in (315), since a pronoun precedes an R-expression with which it is covalued. This is exactly as expected given analyses of the binding conditions.[9]

[9]The fact that *surface* linear precedence is irrelevant in sentences like (314) is also predicted given accounts of the binding conditions that, in part, make reference to precedence, such as Bruening 2014.

Now consider the Chuj sentences in (316), which differ minimally from the Chuj sentences in (312). Strikingly, the adverb *junelxo* 'again', which could appear between *tek* and *waj Xun* in (312-a), can no longer appear between these two words in (316-a). Instead, there is only one option: the adverb must follow the R-expression *Xun*, as in (316-b).

(316) Chuj

a. *Ix-s-b'o' **s**-tek **junelxo** waj Xun.
PFV-A3-make A3-meal again CLF Xun
Intended: 'Xun made his meal again.' (cf. (312-a))

b. *Ixsb'o' stek waj Xun junelxo.*
Ix-s-b'o' **s**-tek waj Xun **junelxo**.
PFV-A3-make A3-meal CLF Xun again
'Xun made his meal again.'

The restriction observed in (316-a) can be explained if the R-expression *Xun* must appear as the possessor instead of the subject, as schematized in (317). That is, if the R-expression is the possessor, then we expect that no adverb should be able to intervene between it and the possessee (recall from (313-b) that Chuj adverbs cannot appear between the possessor and possessee).

(317) *Proposed structure for Chuj* (316-b)
made [$_{\text{OBJ}}$ meal [$_{\text{POSS}}$ Xun$_1$]] {again} [$_{\text{SUBJ}}$ *pro*$_1$] {again}

Without making further assumptions about the syntax of (316-b), the parse in (317) looks like a violation of Condition C. If the subject c-commands the possessor of the object, then we would expect Condition C to ban the R-expression from appearing in possessor position, contrary to what the data suggest. It is also notable that linear precedence appears to be deterministic for the distribution of R-expressions and pronouns in Chuj (as indicated by the impossibility of (312-a)), whereas hierarchical structure determines their distribution in Ch'ol (see grammaticality of (314-a)).

6.1.3.2 Coordination in extended reflexive constructions

Extended reflexive sentences with coordinated objects present another environment showing the presence of Condition C in Ch'ol, and its apparent inapplicability in Chuj. Consider first the following set of sentences from Ch'ol and Chuj:

(318) Ch'ol coordination in extended reflexives

a. *Tyi its'äñä its'i' yik'oty imis ajAna.*

Tyi i-ts'äñ-ä [$_{\&P}$ i-ts'i' [$_{POSS}$ ***pro***] yik'oty i-mis [$_{POSS}$ ***pro***]]
PFV A3-wash-TV A3-dog PRON and A3-cat PRON
[$_{SUBJ}$ **aj-Ana**].
CLF-Ana
'Ana$_1$ washed her$_1$ dog and her$_1$ cat.'

b. *Tyi i-ts'äñ-ä [$_{\&P}$ i-ts'i' [$_{POSS}$ **aj-Ana**] yik'oty i-mis [$_{POSS}$ ***pro***]]
PFV A3-wash-TV A3-dog CLF-Ana and A3-cat PRON
[$_{SUBJ}$ ***pro***].
PRON
Intended: 'Ana$_1$ washed her$_1$ dog and her$_1$ cat.'
Could mean: 'They$_1$ washed Ana$_2$'s dog and cat.'[10]

(319) Chuj coordination in extended reflexives

a. *Ix-s-b'ik [$_{\&P}$ nok' s-tz'i' [$_{POSS}$ ***pro***] yet' nok' s-mis [$_{POSS}$ ***pro***]]
PFV-A3-wash CLF A3-dog PRON and CLF A3-cat PRON
[$_{SUBJ}$ **waj Xun**].
CLF Xun
Intended: 'Xun$_1$ washed his$_1$ dog and his$_1$ cat.'

b. *Ixsb'ik nok' stz'i' waj Xun yet' nok' smis.*

Ix-s-b'ik [$_{\&P}$ nok' s-tz'i' [$_{POSS}$ **waj Xun**] yet' nok' s-mis [$_{POSS}$ ***pro***]]
PFV-A3-wash CLF A3-dog CLF Xun and CLF A3-cat PRON
[$_{SUBJ}$ ***pro***].
PRON
'Xun$_1$ washed his$_1$ dog and his$_1$ cat.'
Lit: 'He$_1$ washed Xun$_1$'s dog and his$_1$ cat.'

[10]The judgements are provided for parses of these sentences with neutral intonation. The intended meaning for this sentence could be judged acceptable if there is a marked prosodic break before *yik'oty imis* 'and her cat'. In such cases, the literal translation would be 'Ana$_1$ washed her$_1$ dog ||, and also her$_1$ car', and so the R-expression in Ch'ol would still need to be in subject position. Moreover, after such a prosodic break, *yik'oty imis* sounds like an afterthought and can appear after adverbs.

Again, Chuj and Ch'ol do not pattern alike. Assuming that transitive subjects c-command objects in both languages, Condition C should require the R-expression to appear in subject position, as is the case in Ch'ol (318). In Chuj (319), however, we observe that the R-expression must appear in the linearly first covalued DP, which in this case is the first possessor of the coordinated extended reflexive object. Therefore, Condition C seems to be once again ignored in Chuj.

6.1.3.3 Relativized objects

Sentences with object relative clauses provide a third area of evidence for the role of linear precedence in determining the order of covalued nominals in Chuj, and for the apparent lack of Condition C effects. Ch'ol examples are not considered here, since objects with full relative clauses in this language obligatorily trigger VSO order, making it impossible to test the relevant sentences (Clemens and Coon 2018; Little 2020b).

Before considering examples with covalued nominals, first consider the sentences in (320-a). Importantly, though speakers indicate a general preference for VSO order in this case, both VOS and VSO are possible when the object of a transitive sentence is relativized:

(320) Chuj

a. *Olyawtej ch'anh libro sman ix ewi ix Ana.*
Ol-y-awtej [$_{OBJ}$ ch'anh libro [$_{RC}$ s-man **ix** ewi]]
PROSP-A3-read CLF book A3-buy CLF.PRON yesterday
[$_{SUBJ}$ **ix** **Ana**].
CLF Ana
'Ana$_1$ will read the book that she$_2$ bought yesterday.' (VOS)

b. *Olyawtej ix Ana ch'anh libro sman ix ewi.*
Ol-y-awtej [$_{SUBJ}$ **ix Ana**]$_1$ [$_{OBJ}$ ch'anh libro [$_{RC}$ s-man **ix**$_2$ ewi]] (VSO)

As foreshadowed in (305), relative clauses in Chuj do not count as a separate CP domain for *pro*. This means that, per the generalization in (307) above which forces the use of *pro* whenever there is covaluation with another nominal expressions within a minimal CP,

the presence of the classifier pronoun *ix* 'her' in the object of the relative clause in (320-a) forces a disjoint reference reading: Ana is not the person who bought the book.

Now consider a minimal counterpart to (320) in which the subject is covalued with the object of the relative clause (i.e., where Ana *is* the person that bought the book). Again, there appear to be two ways of conveying this sentence; (321-a) and (321-b) are judged equally grammatical and semantically equivalent. As discussed in further detail in section 6.1.3.5, the examples in (321) show that re-ordering the subject and object has effects on whether the R-expression appears in the matrix subject or inside the object. Also shown is that the potential VOS configuration in (321-c) is ungrammatical.

(321) Chuj

a. *Olyawtej ch'anh libro sman ix Ana ewi.*
Ol-y-awtej [$_{OBJ}$ ch'anh libro [$_{RC}$ s-man **ix** **Ana** ewi]]
PROSP-A3-read CLF book A3-buy CLF Ana yesterday
[$_{SUBJ}$ ***pro***].
PRON
'Ana$_1$ will read the book that she$_1$ bought yesterday.' (VOS)
Lit: 'She$_1$ will read the book that Ana$_1$ bought yesterday.'

b. *Olyawtej ix Ana ch'anh libro sman ewi.*
Olyawtej [$_{SUBJ}$ **ix Ana**] [$_{OBJ}$ ch'anh libro [$_{RC}$ sman [$_{SUBJ}$ ***pro***] ewi]]. (VOS)
Lit: 'Ana$_1$ will read the book that she$_1$ bought yesterday.'

c. *Olyawtej [$_{OBJ}$ ch'anh libro [$_{RC}$ sman ***pro*** ewi]] [$_{SUBJ}$ **ix Ana**]. (VOS)

Though we discuss VSO sentences like (321-b) in section 6.1.3.5, the crucial point for the current discussion is the possibility of VOS sentences like (321-a), and the impossibility of the minimally different VOS sentence in (321-c). Remarkably, in (321-a), the R-expression can be realized inside the relativized object, in which case the subject is pronominalized. That this is the right syntactic parse can be inferred from surface word order alone. The adverb *ewi* 'yesterday' necessarily modifies the predicate of the relative clause, since the predicate of matrix clause bears future (prospective) aspect. And since *ewi* intervenes between the position of the possessor and the position of the subject (see (321-a) above),

we can infer that the R-expression *Ana* realizes the possessor. In other words, (321-a) literally translates as *She*$_1$ *will read the book that Ana*$_1$ *bought yesterday*, where *she* and *Ana* are covalued. Finally, (321-c) shows us that the R-expression must appear in the linearly first position: (321-a) is the only possible VOS parse, because linear precedence matters.

Again, these data are surprising given the binding conditions. If the subject c-commands the object, then (321-a) appears to violate Condition C.

6.1.3.4 Object A'-extraction

A fourth area of evidence comes from cases of object A'-extraction (Coon et al. (2021) report similar facts in other Mayan languages, see also Craig 1977 on Popti').

As briefly discussed in Chapter 2, across many languages of the Mayan language family—Chuj and Ch'ol included—focused expressions undergo obligatory displacement to a preverbal position (see e.g. Aissen 1992). Relevant examples for both languages are provided in (322) and (323):

(322) *Sa' ta' ijuch'u ajMaría.*
[$_{\text{OBJ}}$ Sa']$_i$ ta' i-juch'-u t_i aj-María.
corn PFV A3-grind-TV CLF-Malin
'María ground CORN.' (Ch'ol, adapted from Little 2020b: (60))

(323) *Ha waj Kixtup ixyil ix Malin.*
[$_{\text{OBJ}}$ Ha waj Kixtup]$_i$ ix-y-il t_i ix Malin.
FOC CLF Kixtup PFV-A3-see CLF Malin
'Malin saw KIXTUP.' (Chuj)

Following Aissen (1992) and subsequent work on Mayan (e.g., Velleman 2014, Coon et al. 2021), I assume that preverbal foci involve movement, specifically A'-movement.

Object A'-extraction can shed further light on the role of linear precedence in Chuj. Consider what happens when the object of an extended reflexive construction is focused in Chuj:

(324) *Ha smam waj Xun ixyila'.*

[$_{OBJ}$ Ha s-mam [$_{POSS}$ **waj Xun**]]$_i$ ix-y-il-a' t_i [$_{SUBJ}$ ***pro***].
FOC A3-father CLF Xun PFV-A3-see-TV PRON

'Xun$_1$ saw HIS$_1$ FATHER.'

Lit: 'He$_1$ saw XUN$_1$'S FATHER.' (Chuj, cf. (316))

In (324), surface word order suffices to show that the R-expression is the possessor and that the subject is null, since the verb now intervenes between the possessor and the subject. Assuming that A'-movement must reconstruct for binding (Chomsky 1995; Fox 1999; Sportiche 2006; Lebeaux 2009; Legate 2014; van Urk 2015), the result again appears to violate Condition C. Also note that the opposite configuration in which the R-expression is realized in subject position is ungrammatical in Chuj:

(325) *[$_{OBJ}$ Ha s-mam [$_{POSS}$ ***pro***]]$_i$ ix-y-il t_i [$_{SUBJ}$ **waj Xun**].
FOC A3-father PFV-A3-see CLF Xun

Intended: 'Xun$_1$ saw HIS$_1$ FATHER.' (Chuj)

The contrast between Chuj and Ch'ol is striking yet again. When the object of an extended reflexive A'-extracts in Ch'ol, the R-expression must be realized in subject position and the possessor must be null (as expected given Condition C):

(326) *Iwakax tyi ichoño ajAna.*

[$_{OBJ}$ I-wakax [$_{POSS}$ *pro*]]$_i$ tyi i-choñ-o t_i [$_{SUBJ}$ **aj-Ana**].
A3-cow PRON PFV A3-sell-TV CLF-Ana

'Ana$_1$ sold HER$_{1/*2}$ COW.' (Ch'ol)

Configurations in which the possessor is overtly realized as an R-expression are impossible under a covalued reading. As discussed in Coon et al. 2021, such configurations lead to obligatory disjoint reference:

(327) *Iwakax ajAna tyi ichoño.*

[$_{OBJ}$ I-wakax [$_{POSS}$ **aj-Ana**]]$_i$ tyi i-choñ-o t_i [$_{SUBJ}$ ***pro***].
A3-cow CLF-Ana PFV A3-sell-TV PRON

'She$_1$ sold ANA$_{2/*1}$'S COW.' (Ch'ol)

For completeness, it can also be shown that in cases of A'-extraction of objects with relative clauses, linear precedence effects are, once again, observed in Chuj. In (328-a), the R-expression must be realized inside the relativized object; (328-b) is judged ungrammatical.

(328) Chuj

a. *Ha ch'anh libro ixsman ix Ana olyawtej.*
[$_{OBJ}$ Ha ch'anh libro [$_{RC}$ ix-s-man **ix** **Ana**]]$_i$ ol-y-awt-ej t_i
FOC CLF book PFV-A3-buy CLF Ana PROSP-A3-read-DTV
[$_{SUBJ}$ ***pro***].
PRON
'Ana$_1$ will read THE BOOK THAT SHE$_1$ BOUGHT.' (VOS)
Lit: She$_1$ will read THE BOOK THAT ANA$_1$ BOUGHT.'

b. *[$_{OBJ}$ Ha ch'anh libro [$_{RC}$ ixsman ***pro***$_1$]]$_i$ olyawtej t_i [$_{SUBJ}$ **ix Ana**]$_1$.
Lit intended: 'Ana$_1$ will read THE BOOK THAT SHE$_1$ BOUGHT.'

Clearly, A'-extracted objects in Chuj do not yield Condition C effects, in contrast with Ch'ol. These data thus provide further indication that precedence plays a fundamental role in the distribution of covalued nominals in Chuj, often to the apparent detriment of Condition C.

6.1.3.5 VSO/VOS alternations

As already noted above, optional VSO/VOS alternations in Chuj can sometimes arise given the right circumstances. The possibility for VSO is often correlated with sentences in which the object is phonologically heavy or the subject is phonologically weak, a fact which is consistent with alternations found in other Mayan languages (see e.g., references in Clemens and Coon 2018). In section 6.1.3.3, for instance, we saw that when the internal argument is modified with a relative clause, VSO word order is exceptionally possible. Such alternations provide a final piece of evidence for the role of linear precedence in Chuj. Considering again the example in (321) above, we can observe that when the subject of the embedded verb precedes the subject of the matrix verb, as in (321-a), it is the subject

of the embedded verb that gets realized as an R-expression. On the other hand, when the subject of the matrix verb precedes the subject of the embedded verb, as in (321-b), it is the subject of the matrix verb that gets realized as an R-expression.

Optional alternations in VOS/VSO word order in Chuj are also exceptionally possible when the subject of a transitive verb is a classifier pronoun. This means that we can also witness the importance of linear precedence in extended reflexive constructions. As we saw in section 6.1.3.1, in VOS extended reflexive constructions, there is evidence that it is the possessor, and not the subject, that gets overtly realized. An example with a classifier pronoun is provided in (329-a) for illustration. In VSO sentences, on the other hand, the subject must get realized as a classifier pronoun, while the possessor is null (329-b):

(329) Chuj VSO/VOS alternations

a. *Ixschonh swakax winh.*
Ix-s-chonh [$_{OBJ}$ s-wakax [$_{POSS}$ **winh**]] [$_{SUBJ}$ *pro*].
PFV-A3-sell A3-cow CLF.PRON PRON
'He$_1$ sold his$_1$ cow.' (VOS)

b. *Ixschonh winh swakax.*
Ix-s-chonh [$_{SUBJ}$ **winh**] [$_{OBJ}$ s-wakax [$_{POSS}$ *pro*]].
PFV-A3-sell CLF.PRON A3-cow PRON
'He$_1$ sold his$_1$ cow.' (VSO)

Word order alternations therefore provide a fourth argument for the role linear precedence in Chuj, and the apparent lack of Condition C violations.

6.1.4 Summary and questions to be addressed

In this section, we have seen evidence from possessors, relative clauses, object A'-extraction, and alternations in VOS/VSO word order that the distribution of nominals in Chuj conforms to the generalization in (307), repeated below.

(307) **Generalization about covalued nominal expressions in Chuj**
If covalued expressions appear in the same minimal CP, the linearly first must be an R-expression (or an overt classifier pronoun), and the rest are realized as *pro*.

From a crosslinguistic perspective, these facts are surprising. Given the data seen so far, Chuj appears to consistently violate Condition C, a pattern that is not expected if the binding conditions are universal (Grodzinsky and Reinhart 1993; Reuland 2010, 2011). The data are even more surprising considering the fact that other Mayan languages, like Ch'ol, *do* generally abide by the binding conditions. The main findings of this section are summarized in Table 6.1.

Table 6.1: Evidence of Condition C in Ch'ol vs. Chuj

	Data diagnostic	Ch'ol Condition C-abiding?	Chuj Condition C-abiding?
§6.1.3.1	Adverbs in ext. reflexives	yes	no
§6.1.3.2	Coordinated ext. reflexives	yes	no
§6.1.3.3	Relativized objects	n/a	no
§6.1.3.4	Object A'-extraction	yes	no
§6.1.3.5	VOS/VSO alternations	n/a	no

We thus find ourselves at a juncture. One option could be to explore the possibility that the binding conditions in (288) are not universal, or that they need to be modified in order to accommodate the linear precedence facts in Chuj. A second option would be to maintain the universality of the binding conditions, but explore the possibility that there is something special about the syntax of Mayan languages like Chuj that reconciles the unexpected data with respect to the binding conditions. In the next sections, I will argue in favour of the second option. In particular, I argue that Mayan languages for which linear precedence seems to play a central role, like Chuj, exhibit a different syntax than languages where only structural relations seem to matter, like Ch'ol. It is this special syntax that ultimately leads to configurations in which the binding conditions are inactive.

More concretely, the central goal of the rest of this chapter will be to answer the following three questions, relevant to the Chuj and Ch'ol data that we observed in this section:

(330) Questions to be addressed

1. Why is Condition C seemingly ignored in Chuj, but not in Ch'ol? §6.2
2. Why does linear precedence play a role in the distribution of covalued nominals in Chuj? §6.3
3. Do the binding conditions ever apply as expected in Chuj? §6.4

Section 6.2 focuses on the first question. The leading proposal will be that in every configuration in which the binding conditions do not appear to apply in Chuj, the relevant covalued expressions are not in a c-command relation. In other words, the Condition C violations are only apparent, since the relevant covalued nominals in Chuj are free. The Chuj data are therefore entirely *consistent* with the binding conditions, and do not serve as evidence against a universalist approach to syntactic binding.

Section 6.3 turns to the second question. We will see that the Chuj linear precedence effects seen in this section are just one symptom of a more general *anti-cataphora* constraint that regulates the distribution of *free nominals* in Mayan. That is, in clear cases of free pronouns, where c-command relations do not hold between covalued nominal expressions, both Chuj and Ch'ol, unlike languages like English, show a general ban on cataphora (backwards pronominalization).

Section 6.4 then turns to the third question, answering it in the affirmative. A crucial conclusion will be that when nominals are in a c-command relation in Chuj, the regular binding conditions *do* apply, and override the anti-cataphora constraint. The overall conclusion will be that Chuj and Ch'ol, despite looking completely disparate at first glance, are subject to the exact same constraints on pronominalization: (i) the binding conditions (for bound nominals) and (ii) a ban on cataphora (for free nominals). The only difference between the two languages is the fact that Chuj exhibits idiosyncratic syntactic proper-

ties, such as high-absolutive syntax, which often obviate the application of the binding conditions and therefore lead to the illusion the binding conditions are inactive.

Finally, it should be reiterated that similar patterns of nominal covaluation have been described by Craig (1977) for Popti', a close-relative of Chuj. Though the Chuj data have not been previously discussed, the Popti' data have received much more attention in the theoretical literature, including in work such as Hoekstra 1989, Woolford 1991, Trechsel 1995 and Aissen 2000. However, the sentences in which surprising binding effects may be observed in Popti' are much more limited, because Popti' exhibits rigid VSO word order. The proposals of the next sections will nonetheless ultimately build on components of the analyses in Craig 1977, Trechsel 1995, and Aissen 2000, as discussed below.

6.2 High-absolutive syntax and syntactic binding

The goal of this section is to show that the surprising Chuj data seen in the previous section are *consistent* with the binding conditions, despite initial appearances. Specifically, I argue that "high-absolutive syntax"—independently proposed to underlie other morphosyntactic phenomena in a subset of Mayan languages (Coon et al. 2014, a.o.)—can explain the surprising lack of Condition C effects in Chuj.

I first provide background on the high-absolutive approach to syntactic ergativity in section 6.2.1. Section 6.2.2 then lays out the proposal. In section 6.2.3, I elaborate on one prediction made by the analysis, which I show is borne out. Finally, in section 6.2.4, I consider additional data from oblique phrases, which are not immediately covered by the high-absolutive approach, and I suggest a tentative account of these data that keeps to the general theme of the proposal.

6.2.1 Syntactic ergativity and object raising in Mayan languages

In addition to the differences in pronominal realization described in section 6.1, Mayan languages like Chuj and Ch'ol differ in their syntax in another, better-known respect: while both languages are morphologically-ergative, only Chuj demonstrates syntactic ergativity in the form of transitive subject extraction asymmetries, also known as the "Ergative Extraction Constraint" (EEC) (Aissen 2017, Coon et al. 2021):

(331) Chuj → EEC

a. *Ixachyil ix ix.*
Ix-ach-y-il ix ix.
PFV-B2S-A3-see CLF woman
'The woman saw you.'

b. *¿Mach$_j$ ix-ach-y-il-a' t_j?
who PFV-B2S-A3-see-TV
Intended: 'Who saw you?'

(332) Ch'ol → no EEC

a. *Tyi ik'eläyety x'ixik.*
Tyi i-k'el-ä-yety x-'ixik.
PFV A3-see-DTV-B2 CLF-woman
'The woman saw you.'

b. *¿Maxki tyi ik'eläyety?*
¿Maxki tyi i-k'el-ä-yety?
who PFV A3-see-DTV-B2
'Who saw you?'

As seen above, only Chuj exhibits the EEC: (331-b) shows that an ergative subject cannot be A'-extracted from a canonical transitive sentence in Chuj. In Ch'ol, on the other hand, ergative subjects can be freely A'-extracted, as seen in (332-b).

Transitive subject extraction asymmetries, such as the one in (331), have received a lot of attention in the Mayanist literature (e.g. Aissen 1999, 2011, 2017; Stiebels 2006; Coon et al. 2014; Preminger 2014; Assmann et al. 2015; Erlewine 2016; Henderson and Coon 2018; Coon et al. 2021, a.o., for work specifically on Mayan). Since space prevents me

from doing justice to the various proposals (though see Deal 2016 and Polinsky 2017 for relevant overviews of related analyses of syntactic ergativity across languages, and Aissen 2017 specifically on Mayan), I only provide details on the type of analysis that is directly relevant for the analysis proposed in this chapter, namely the one first proposed (for Mayan) in Coon et al. 2014.

Building on previous work on syntactic ergativity in other languages (Campana 1992; Bittner and Hale 1996; Aldridge 2004), Coon et al. (2014) propose that whether or not a language exhibits the EEC stems from a deep syntactic parameter in the Mayan language family (see also Assmann et al. 2015; Coon et al. 2021). In 'high-absolutive' languages like Chuj, the absolutive object consistently raises to a position above the ergative subject (333). In low-absolutive languages like Ch'ol, no such raising occurs (334).

(333) High-absolutive language (e.g., Chuj)
$[_{v/VoiceP}$ OBJECT [SUBJECT $[_{VP}$ V <OBJECT>]]]

(334) Low-absolutive languages (e.g., Ch'ol)
$[_{v/VoiceP}$ SUBJECT $[_{VP}$ V OBJECT]]]

Though formalized differently in different works, the main idea is that the high-absolutive syntax in (333) creates an *intervention problem* for the extraction of the ergative subject; this intervention is taken to be at the source of the EEC:

(335) Raising of object in high-absolutive languages blocks subject extraction
$[_{CP}$ __ ... $[_{v/VoiceP}$ OBJECT [SUBJECT $[_{VP}$ V <OBJECT>]]]]

As noted by the proponents of this analysis, there is a crucial correlation between the high/low-absolutive parameter and the position of the absolutive morpheme, first noticed by Tada (1993). In high-absolutive Mayan languages, Set B (absolutive) morphemes tend to be realized between aspect marking and Set A (ergative) morphemes, as in (336-a).

In low-absolutive Mayan languages, Set B morphemes appear as a suffix on the verb stem (336-b).

(336) a. Verb stem in high-absolutive languages
TAM - **Set B** - Set A - verb - suffixes (see (331-a))

b. Verb stem in low-absolutive languages
TAM - Set A - verb - suffixes - **Set B** (see (332-a)

Coon et al. (2014) argue that this correlates with the locus of absolutive licensing. In high-absolutive languages, the source of absolutive is finite T/Infl, whereas it is transitive v/VoiceP in low-absolutive languages.

In order to extract the ergative subject in high-absolutive Mayan languages like Chuj, speakers employ a special construction, known as the "Agent Focus (AF) construction". In this construction, the verb lacks Set A (ergative) agreement and is suffixed with special morphology (*-an* in Chuj).

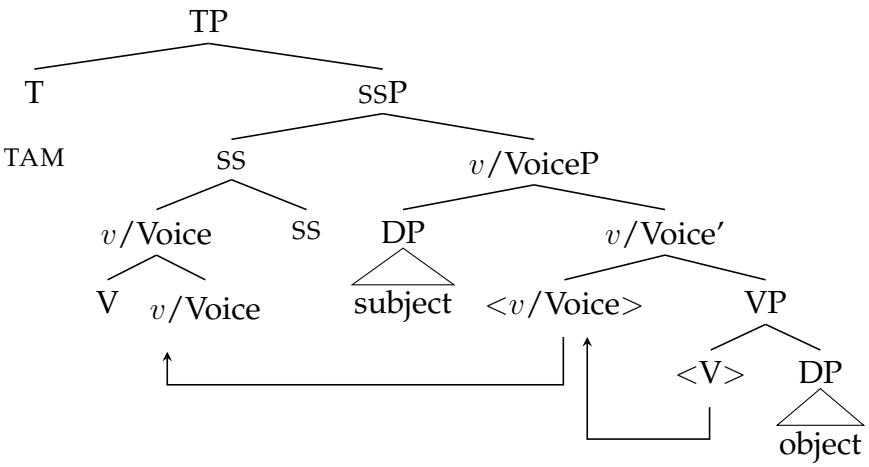
(337) *Ha ix chichim ixachilani.*
Ha ix chichim ix-ach-il-**an**-i.
FOC CLF elder.woman PFV-B2S-see-AF-IV
'THE ELDER saw you.' (Chuj Agent Focus)

In the Agent Focus construction, Coon et al. (2014) propose that the absolutive object does not raise, and that absolutive morphemes exceptionally have a low source (v/Voice, instantiated as *-an* in Chuj). Therefore, the EEC is circumvented and the ergative subject can extract.

Finally, before we move on, a note on word order in Mayan is in order. All else being equal, we might expect objects to systematically precede the subject in high-absolutive sentences (assuming leftward movement over the subject). Though this is indeed generally the case in Chuj, it is not necessarily the case in other high-absolutive Mayan languages: many high-absolutive Mayan languages exhibit VSO word order. Moreover, many low-absolutive Mayan languages, Ch'ol included, exhibit VOS word order.

As discussed in Chapter 2, I follow Clemens and Coon (2018) in assuming that verb-initial word order is derived from a base-generated SVO order via head-movement of the verb to a higher functional projection:

(338) Verb-initial word order via head movement (Clemens and Coon 2018)



Also following Clemens and Coon, I assume that the order of post-verbal arguments is sensitive to the phonological properties of arguments (e.g., phonological weight; see §6.1.3.5). Syntactic trees in examples will therefore not always match the actual word order seen in Chuj and Ch'ol. For other proposals of word order in Mayan, see England 1991, Aissen 1992, Coon 2010b, Douglas et al. 2017, and Little 2020b.

Having established the relevant background, I now show how the high-absolutive approach to the EEC can help us understand the surprising differences in patterns of nominal binding in Chuj and Ch'ol, without having to abandon a universal approach to the binding conditions.

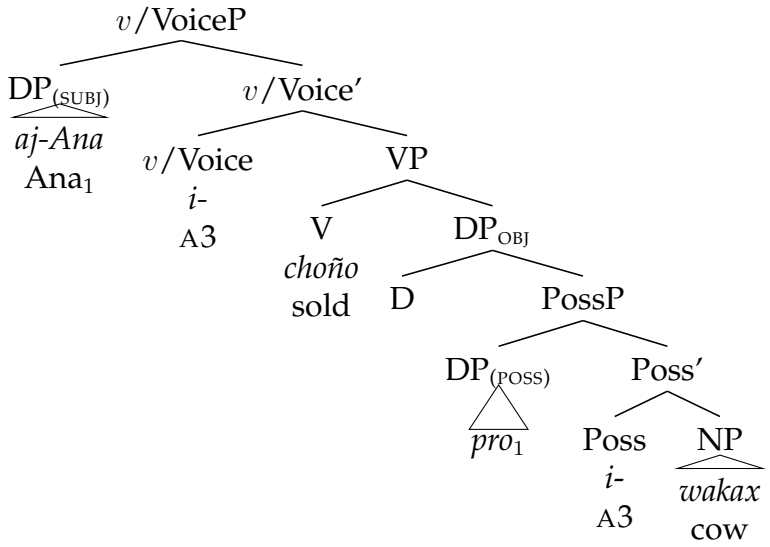
6.2.2 Proposal: High-absolutives bleed Condition C

We start with Ch'ol. Consider the extended reflexive construction in (339-a), with the proposed 'low-absolutive' syntax in (339-b). As in Chapter 2, I assume that the locus of Set A morphemes are functional projections that sit above the VP and NP.

(339) Ch'ol extended reflexive construction and corresponding structure

a. *Tyi ichoño iwakax aj-Ana.*
Tyi i-choñ-o [$_{OBJ}$ i-wakax [$_{POSS}$ *pro*]] [$_{SUBJ}$ aj-Ana].
PFV A3-sell-TV A3-cow *pro* CLF-Ana
'Ana$_1$ sold her$_{1/*2}$ cow.'

b.



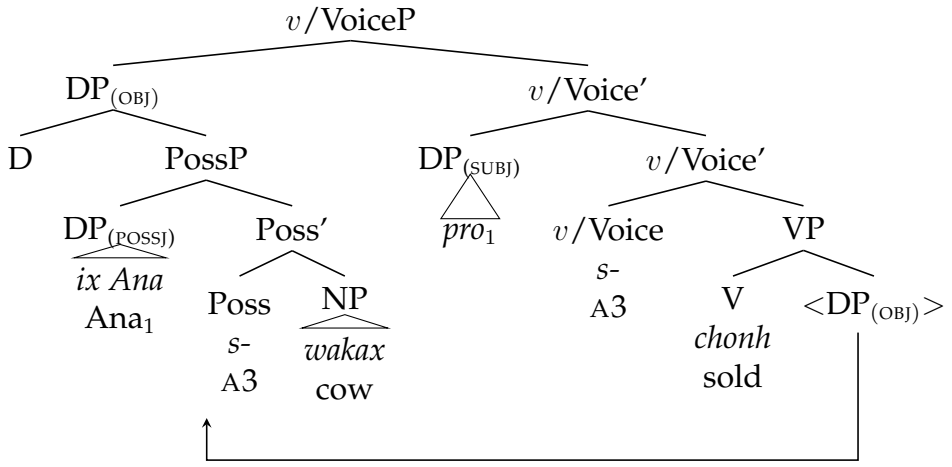
That Ch'ol generally abides by Condition C, as seen in section 6.1, is not surprising given the low-absolutive syntax in (339-b). Note that the subject c-commands and therefore binds the object, so Condition C should apply. Since generating the R-expression in the possessor of the object in (339-b) would clearly lead to a violation of Condition C, we expect that the R-expression can only occur in subject position. In other words, whenever the subject is covalued with an expression contained inside the absolutive object in Ch'ol, we expect that Condition C will be operative and that an R-expression will have to occur in subject position, as we saw is the case in section 6.1.

Now let us consider Chuj. Assuming high-absolutive syntax, with the object raised over the subject, (340-a) has the structure in (340-b). Though not discussed explicitly in previous work on high-asbolutive Mayan languages, this syntax should have a crucial consequence for binding: object raising should bleed c-command relations between subjects and covalued DPs inside the object. This is schematized in (340-b).

(340) Chuj extended reflexive construction and corresponding structure

a. *Ixschonh swakax ix Ana.*
Ix-s-chonh [$_{OBJ}$ s-wakax [$_{POSS}$ ix Ana]] [$_{SUBJ}$ *pro*].
PFV-A3-sell A3-cow CLF Ana *pro*
Lit: 'She$_1$ sold Ana$_1$'s cow.'

b.



Once the object raises to a position above the subject in (340-b), the subject no longer c-commands the possessor within the object. And the same goes for the possessor of the object, which also does not c-command the subject. This is crucial: if neither the subject nor the possessor c-command each other in sentences like (340-a), then nothing should prevent the R-expression from appearing in possessor position. In other words, once we adopt a high-absolutive syntax for Chuj, as independently-motivated in previous work, then we find that the apparent violations of Condition C described in section 6.1 were only illusory.[11] The two covalued expressions in (340-a) are not in a binding relationship; they are both free expressions.

Importantly, I assume that object raising is an instance of A-movement (as proposed by Coon et al. 2021), and furthermore that A-movement does not reconstruct for Condition C. The absence of A-movement reconstruction in (340-b) is an essential component of the proposal. If the object obligatorily reconstructs, then we would expect (340-b)

[11]Note that Trechsel (1995) independently arrives at the conclusion that the absolutive object must asymmetrically c-command the ergative subject in Popti', precisely in order to explain data related to the apparent lack of Condition C effects in this language.

to trigger a Condition C violation. As seen in section 6.1, this clearly is not the case in Chuj. Moreover, the claim that A-movement does not reconstruct for Condition C is not new, and was proposed in many works including Chomsky 1995, Fox 1999; Lasnik 1999, Legate 2014 (see Takahashi 2010 for an overview).[12] Indeed, though A'-movement in English must reconstruct for Condition C, as illustrated by the ungrammaticality of (341), there is evidence that it does not for A-movement (342):

(341) *A'-movement → reconstruction*
*[Which picture of John$_1$]$_i$ did he$_1$ like t_i?

(342) *A-movement → no reconstruction* (Fox 1999: 192)
[Every argument that John$_1$ is a genius]$_i$ seems to him$_1$ t_i to be flawless.

In sum, I argued that the high-absolutive approach to the EEC can explain the fact that Chuj, as opposed to Ch'ol, shows surprising patterns of nominal covaluation. High-absolutive syntax leads to a syntactic configuration in which neither the subject nor expressions contained inside the object c-command each other, thereby explaining the absence of Condition C violations in all of the Chuj sentences seen in section 6.1. As a result, we do not need to deny the universality of the binding conditions in order to explain the distribution of nominal expressions in Chuj, despite initial appearances.

The current proposal makes a number of predictions. For one, notice that high-absolutive syntax leads to configurations in which the absolutive object asymmetrically c-commands the subject. This means that when the absolutive object is covalued with an expression inside the subject, Condition C effects should be perceived. I elaborate on this prediction in section 6.4.1, showing that while the prediction is indeed borne out, the data are complicated by the fact that objects also *precede* subjects in the testable sentences, making it impossible to untangle Condition C effects from linear precedence effects. In the

[12]The crucial point here is that there is no A-movement reconstruction *for Condition C*. I am not claiming that A-movement reconstruction is always impossible, but simply that it is impossible in the relevant examples (see Fox 1999, Boeckx 2001 and Sportiche 2006 for reasons to think that there is sometimes reconstruction for A-movement).

next section, we focus on another—this time typological—prediction made by the current analysis, which upon preliminary investigation, is borne out.

6.2.3 Prediction: Nominal covaluation across Mayan

Since the current analysis relies on the low/high-absolutive parameter, we make the following pan-Mayan typological prediction:

(343) **Typological prediction**

a. Mayan languages that exhibit the EEC (and therefore exhibit high-absolutive syntax) should behave like Chuj in showing no Condition C effects in syntactic contexts comparable to the ones seen in §6.1.

b. Mayan languages that do not exhibit syntactic ergativity (and therefore exhibit low-absolutive syntax) should behave like Ch'ol in complying as expected to Condition C in the syntactic contexts seen in §6.1.

A straightforward way to test this prediction in different Mayan languages is with sentences in which the object of an extended reflexive is extracted. High-absolutive languages should exhibit—or at least allow—the parse in (344-a), whereas low-absolutive languages should systematically exhibit the parse in (344-b):

(344) a. High-absolutive languages (possessor should be overt and subject null):
[$_{\text{OBJ}}$ … [$_{\text{POSS}}$ R-expression$_1$]] verb [$_{\text{SUBJ}}$ *pro*$_1$]

b. Low-absolutive languages (possessor should be null and subject overt):
[$_{\text{OBJ}}$ … [$_{\text{POSS}}$ *pro*$_1$]] verb [$_{\text{SUBJ}}$ R-expression$_1$]

Though I leave an extensive cross-Mayan investigation of this prediction to future work, preliminary investigation suggests that this prediction is borne out, as also discussed in Coon et al. 2021. Of the languages that I could survey, three were high-absolutive and two were low-absolutive (see Coon et al. 2021 for classification of low/high absolutive languages). The three high-absolutive languages all belong to different sub-branches of Mayan languages, and could thus be considered distant relatives within the family (Law 2014).

All three high-absolutive languages, Q'anjob'al, Mam, and Kaqchikel, show the same (lack of) Condition C effects as Chuj: when the object of an extended reflexive is fronted, the R-expression is realized in possessor position, and the subject is null:

(345) Q'anjob'al

a. *A no' swakax naq Xhunik max stxono'.*

[$_{OBJ}$ A no' s-wakax [$_{POSS}$ **naq Xhunik**]] max s-txon-o' [$_{SUBJ}$ ***pro***].
FOC CLF A3-cow CLF Xhunik PFV A3-sell-TV PRON

'Xhunik$_1$ sold HIS$_1$ COW.'

b. *A no' swakax max stxon naq Xhunik.*

[$_{OBJ}$ A no' s-wakax [$_{POSS}$ ***pro***]] max s-txon [$_{SUBJ}$ **naq Xhunik**].
FOC CLF A3-cow PRON PFV A3-sell CLF Xhunik

Cannot mean: 'Xhunik$_1$ sold HIS$_1$ COW.'
Means: 'Xhunik$_1$ sold HIS$_2$ COW.'

(346) Mam

a. *A tchej Xwan o tz'ok tb'yo'n.*

[$_{OBJ}$ A t-chej [$_{POSS}$ **Xwan**]] o tz'-ok t-b'yo-'n [$_{SUBJ}$ ***pro***].
DET A3S-horse Xwan PFV B3S-DIR A3S-hit-DS PRON

'Xwan$_1$ hit HIS$_1$ HORSE.'

b. *A tchej o tz'ok tb'yo'n Xwan.*

[$_{OBJ}$ A t-chej [$_{POSS}$ ***pro***]] o tz'-ok t-b'yo-'n [$_{SUBJ}$ **Xwan**].
DET A3S-horse PRON PFV B3S-DIR A3S-hit-DS Xwan

Cannot mean: 'Xwan$_1$ hit HIS$_1$ HORSE.'
Means: 'Xwan$_1$ hit HER$_2$ HORSE.'

(347) Kaqchikel

a. *Ja ri ruwakx ri xta Ana xuk'ayij.*

[$_{OBJ}$ Ja ri ru-wakx [$_{POSS}$ **ri xta Ana**]] x-u-k'ayi-j [$_{SUBJ}$ ***pro***].
FOC DET A3S-cow DET CLF Ana PFV-A3-sell-DTV PRON

'Ana$_1$ sold HER$_1$ COW.'

b. *Ja ri ruwakx xuk'ayij ri xta Ana.*

[$_{OBJ}$ Ja ri ru-wakx [$_{POSS}$ ***pro***]] x-u-k'ayi-j [$_{SUBJ}$ **ri xta Ana**].
FOC DET A3S-cow PRON PFV-A3-sell-DTV DET CLF Ana

Cannot mean: 'Ana$_1$ sold HER$_1$ COW.'
Means: 'Ana$_1$ sold THEIR$_2$ COW.'

Conversely, the two low-absolutive languages I surveyed, Tojol-ab'al and Tseltal, both behave like Ch'ol: Condition C is active in the relevant sentences, as seen in the examples below. Recall from Chapter 2 that while Tseltal is part the Cholan-Tseltalan subbranch of Mayan languages, Tojol-ab'al is often categorized as a close-relative of Chuj (Q'anjob'alan) (Kaufman 1969, Hopkins 2006).[13]

(348) Tojol-ab'al

a. *Ja' ja swakax xchona ja Jwani'.*

[$_{OBJ}$ Ja' ja s-wakax [$_{POSS}$ ***pro***]] x-chon-a [$_{SUBJ}$ **ja** **Jwan-i'**].
FOC DET A3-cow PRON A3-sell-TV DET Jwan-DET

'Jwan$_1$ sold HIS$_1$ COW.'

b. **Ja' ja swakax ja Jwan xchona.*

[$_{OBJ}$ Ja' ja s-wakax [$_{POSS}$ **ja** **Jwan**]] x-chon-a [$_{SUBJ}$ ***pro***].
FOC DET A3-cow DET Jwan A3-sell-TV PRON

Intended: 'Jwan$_1$ sold HIS$_1$ COW.'[14]

(349) Tseltal

a. *Ja' xwakax la xchon teWane.*

[$_{OBJ}$ Ja' x-wakax [$_{POSS}$ ***pro***]] la x-chon [$_{SUBJ}$ **te** **j-Wan-e**].
FOC A3-cow PRON PFV A3-sell DET CL-Wan-DET

'Wan$_1$ sold HIS$_1$ COW.'

b. *Ja' xwakax jWan la xchon.*

[$_{OBJ}$ Ja' x-wakax [$_{POSS}$ **j-Wan**]] la x-chon [$_{SUBJ}$ ***pro***].
FOC A3-cow CL-Wan PFV A3-sell PRON

Cannot mean: 'Wan$_1$ sold HIS$_1$ COW.'
Means: 'Wan$_1$ sold HIS$_2$ COW.'

Strikingly, the prediction that Condition C effects should be *absent* from the relevant sentences in high-absolutive languages, but *present* in low-absolutive languages, is borne out in a range of languages. Insofar as the prediction is confirmed across the Mayan family,

[13]The classification of Tojol-ab'al has been subject to debate. Recent work, however, has convincingly argued that it shows signs of a 'mixed' language, showing not only several similarities with Chuj, but also with Tseltal (Law 2011, 2014, Gómez Cruz 2017).

[14]In Tojol-ab'al, the non coreferential reading requires an applicative dative construction (Hugo Hector Vázquez López, p.c.).

it constitutes not only strong support for the current analysis, but also more generally support for high-absolutive approaches to syntactic ergativity in Mayan.

6.2.4 Oblique phrases and binding

While this chapter focuses primarily on covaluation relations between external and internal arguments of transitive clauses (and nominals contained within both), important differences between Chuj and Ch'ol in terms of nominal covaluation may also be observed between other types of nominals. Specifically, as I show in this subsection, similar linear precedence effects, surprising from the perspective of classic assumption about the structural position of arguments in the clause, are observed in Chuj, but not in Ch'ol, when the main arguments of verbs are covalued with expressions inside oblique phrases. Since the differences in patterns of nominal covaluation were attributed to object raising in the previous sections, the current analysis does not immediately cover cases of nominal covaluation with oblique phrases. This short subsection thus provides new data concerning variation between oblique phrases in Chuj and Ch'ol, and initial steps toward an account of the contrast between the two languages.

Consider the set of sentences from Chuj (350) and Ch'ol (351), which involve covaluation between the absolutive subject and an expression contained inside a PP. Notice that in (350-a) and (351-a), the PPs are postverbal, whereas in (350-b)-(350-c) and (351-b)-(351-c), the PPs appear in a preverbal position.

(350) Chuj

a. *Ixway waj Xun t'a spat.*
Ix-way [$_{SUBJ}$ **waj Xun**] [$_{PP}$ t'a s-pat [$_{POSS}$ ***pro***]].
PFV-sleep CLF Xun PREP A3-house PRON
'Xun$_1$ slept in his$_1$ house.'

b. *T'a spat waj Xun ixwayi.*
[$_{PP}$ T'a s-pat [$_{POSS}$ **waj Xun**]] ix-way-i [$_{SUBJ}$ ***pro***] .
PREP A3-house CLF Xun PFV-sleep-IV PRON
'Xun$_1$ slept in his$_1$ house.' (Lit: In Xun$_1$'s house he$_1$ slept)

c. *[$_{PP}$ T'a spat [$_{POSS}$ ***pro***]] ixwayi [$_{SUBJ}$ **waj Xun**].
'Xun$_1$ slept in his$_1$ house.' (Lit: In his$_1$ house Xun$_1$ slept)

(351) Ch'ol

a. *Tyi wäyi **ajRosa** tyi yotyoty.*
Tyi wäy-i [$_{SUBJ}$ **aj-Rosa**] [$_{PP}$ tyi y-otyoty [$_{POSS}$ ***pro***]].
PFV sleep-IV CLF-Rosa PREP A3-house *pro*
'Rosa$_1$ slept in her$_1$ house.'

b. *Tyi yotyoty tyi wäyi **ajRosa**.*
[$_{PP}$ Tyi y-otyoty [$_{POSS}$ ***pro***]] tyi wäy-i [$_{SUBJ}$ **aj-Rosa**]
PREP A3-house *pro* PFV sleep-IV CLF-Rosa
'In her$_1$ house Rosa$_1$ slept.'

c. *Tyi yotyoty **ajRosa** tyi wäyi*
[$_{PP}$ Tyi yotyoty [$_{POSS}$ **ajRosa**]] tyi wäyi [$_{SUBJ}$ ***pro***].
'Rosa$_1$ slept in her$_{2/*1}$ house.'

Yet again, what we see is that linear precedence matters in Chuj, while it does not in Ch'ol. When the PP follows the subject, as in (350-a) and (351-a), Chuj and Ch'ol exhibit the same pattern: the subject is obligatorily realized as an R-expression and the covalued expression inside the PP is null. When the PP precedes the subject, however, the two languages diverge. In Chuj the R-expression must be realized inside the fronted PP, as shown in (350-b) and (350-c). In Ch'ol it must be realized in subject position, as shown in (351-b) and (351-c). In other words, we see reconstruction effects for Condition C in Ch'ol, but not in Chuj.

Consider another Chuj example, this time with a transitive sentence. Again, the crucial point to notice is that the linearly first covalued expression is the one to get realized as an R-expression:

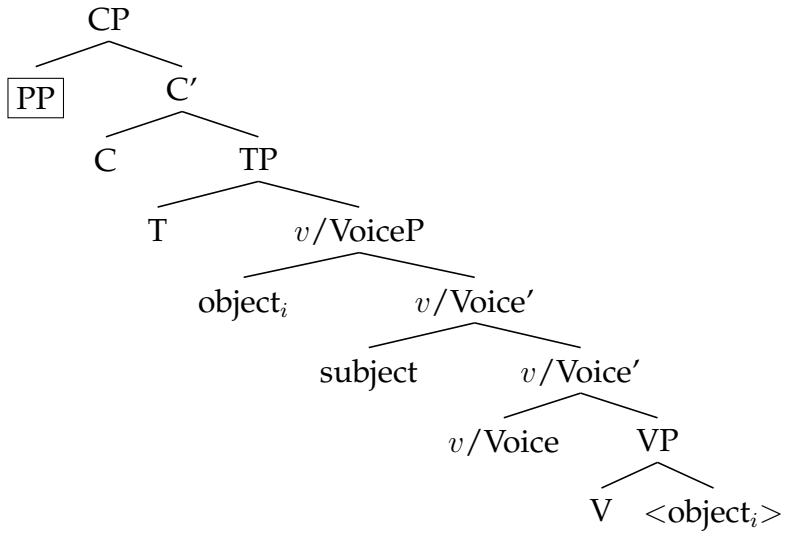
(352) Chuj

a. *Ixyik' b'at k'atzitz waj Xun t'a spat.*
Ix-y-ik' b'at k'atzitz [$_{SUBJ}$ **waj Xun**] [t'a s-pat [$_{POSS}$ ***pro***]].
PFV-A3-carry DIR.go wood CLF Xun PREP A3-house *pro*
'Xun$_1$ carried wood to his$_{1/*2}$ house.'

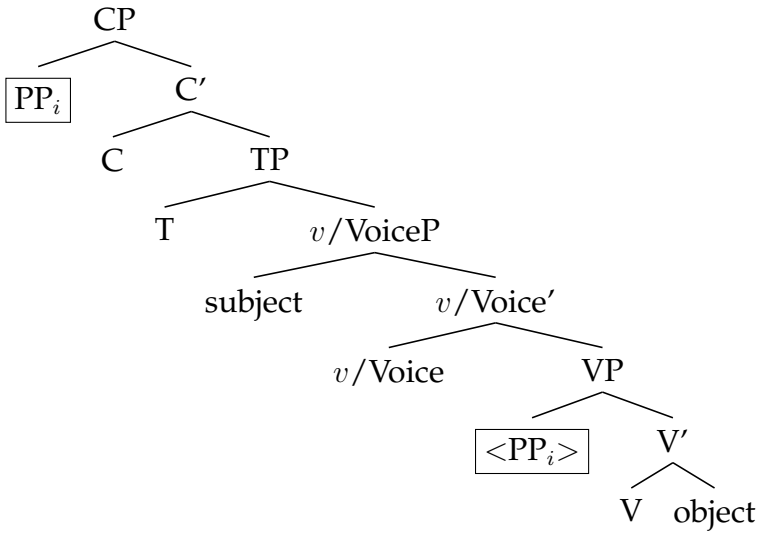
b. *T'a spat waj Xun ixyik' b'at k'atzitz.*
[T'a s-pat [$_{POSS}$ **waj Xun**]] ix-y-ik' b'at k'atzitz [$_{SUBJ}$ ***pro***]
PREP A3-house CLF Xun PFV-A3-carry DIR.go wood *pro*
'To his$_{1/*2}$ house Xun$_1$ carried wood.' (Lit: To Xun$_1$'s house, he$_1$ carried wood)

In the spirit of the the proposal in section 6.2.2, I suggest that the absence of Condition C effects in Chuj can be explained based on the fact that there are no c-command relations between the relevant covalued nominal expressions in (350-b) and (352-b). The main idea goes as follows: while fronted PPs in Chuj are base-generated in their preverbal position above the main arguments, PPs in Ch'ol can be generated low and A'-extracted to a preverbal position. Concretely, I propose that preverbal PPs in Chuj and Ch'ol are derived as follows (high/low-absolutive syntax is also exemplified below):

(353) Base-generated preverbal PP in Chuj



(354) A'-extracted preverbal PP in Ch'ol



As far as Condition C is concerned, the two structures in (353) and (354) make different predictions. In the structure for Chuj in (353), no c-command relations hold between the main arguments and expressions contained inside the PP, and so no Condition C effects are expected. In the structure for Ch'ol in (354), the subject binds inside the PP before A'-movement, which feeds Condition C effects. Though the evidence in favour of the height of oblique phrases is limited, and I leave a more detailed analysis of these data to future

work, I provide two arguments that this analysis is on the right track.[15] One is conceptual and one is empirical.

On the conceptual side, it should be stressed that the high-absolutive approach to the EEC proposed by Coon et al. (2014) and Coon et al. (2021) already *predicts* that adjunct extraction in transitive sentences like (352) would be problematic. In these accounts, the prediction is that nothing besides the absolutive object should be able to extract from a transitive v/VoiceP, because the raised object blocks A'-extraction of all other kinds of constituents (see discussion in §5.3 of Coon et al. 2014, and §3.3 of Coon et al. 2021). At least in a transitive sentence, preverbal adjuncts would therefore need to be base-generated high, a prediction Coon et al. (2014) provide evidence for based on the behaviour of low adverbs in Q'anjob'al.[16] Though the ban on (low) adjunct extraction is only predicted for transitive sentences, it would not be surprising if some high-absolutive languages, such as Chuj, had simply generalized a base-generation strategy for preverbal adjuncts. I suggest that this is indeed the case.

On the empirical side, note that the relative ordering of PPs with respect to subjects in Chuj and Ch'ol is different. Oblique phrases in Chuj can only appear in peripheral positions, as in (355-a) and (355-b). They can never intervene between the verb and the object (355-c), or between the object and the subject (355-d):

[15]Note that even in English, the status of syntactic binding between arguments and PPs is complicated (Reinhart 1976, Pesetsky 1995, Bruening 2014). See, for instance, Bruening 2014 §3.1 and §5.3 for relevant discussion.

[16]In some Mayan languages (but not Q'anjob'alan languages), the extraction of low adjuncts requires a special construction, with v/Voice-like morphology in the main clause (see e.g. Ayres 1983, England 1997, Henderson 2007, Velleman 2014, Can Pixabaj 2015; Mendes and Ranero 2021). It would be interesting to see whether the presence of adjunct extraction morphology has effects on the realization of covalued nominals when oblique phrases are fronted in these languages. I leave this question for future work.

(355) Chuj

a. *Ixsman ixim ixim ix Rosa t'a merkado.*
✓Ix-s-man [$_{OBJ}$ ixim ixim] [$_{SUBJ}$ ix Rosa] [$_{PP}$ **t'a merkado**].
PFV-A3-buy CLF corn CLF Rosa PREP market
'Rosa bought corn at that market.

b. *T'a merkado ixsman ixim ixim ix Rosa.*
✓[$_{PP}$ **T'a merkado**] ix-s-man [$_{OBJ}$ ixim ixim] [$_{SUBJ}$ ix Rosa].

c. *Ix-s-man [$_{PP}$ **t'a merkado**] [$_{OBJ}$ ixim ixim] [$_{SUBJ}$ ix Rosa].

d. *Ix-s-man [$_{OBJ}$ ixim ixim] [$_{PP}$ **t'a merkado**] [$_{SUBJ}$ ix Rosa].

In Ch'ol, on the other hand, the position of PPs is more flexible. PPs readily intervene between objects and subjects, as seen in (356-d):

(356) Ch'ol

a. *Tyi imäñä ixim ajRosa tyi merkadu.*
✓Tyi i-mäñä [$_{OBJ}$ ixim] [$_{SUBJ}$ aj-Rosa] [$_{PP}$ **tyi merkadu**].
PFV A3-buy corn aj-Rosa PREP market
'Rosa bought corn at the market.

b. *Tyi merkadu tyi imäñä ixim ajRosa.*
✓[$_{PP}$ **Tyi merkadu**] tyi i-mäñä [$_{OBJ}$ ixim] [$_{SUBJ}$ aj-Rosa].

c. *Tyi i-mäñä [$_{PP}$ **tyi merkadu**] [$_{OBJ}$ ixim] [$_{SUBJ}$ aj-Rosa].

d. *Tyi imäñä ixim tyi merkadu ajRosa.*
✓Tyi i-mäñä [$_{OBJ}$ ixim] [$_{PP}$ **tyi merkadu**] [$_{SUBJ}$ aj-Rosa].

While VO-PP-S order is possible in Ch'ol, it is not in Chuj. To my knowledge, this is a novel observation that should be tested across other low-absolutive and high-absolutive Mayan languages. Though I remain agnostic about the derivation of word order in this chapter, I take these differences as indicative that PPs occupy different structural positions in both languages. In particular, this is compatible with the proposal that PPs in Chuj are always base-generated above the external and internal arguments, regardless of whether they are preverbal or postverbal.

A reviewer on related work asks about how the high base-generation theory of obliques in Chuj could account for indirect object PPs (such as goal PPs), which by assumption

would need to be base-generated in a lower position in the structure. As discussed in Coon et al. 2014, high-absolutive Mayan languages, contrary to some low-absolutive Mayan languages, do not possess a distinct class of double object constructions. And to my knowledge, it is not clear that any verb inherently requires two objects in Chuj, which raises the question of whether Chuj possesses distransitive verbs in the first place, and therefore genuine indirect object PPs. I thus leave a detailed investigation of these constructions in Chuj for future work.

6.2.5 Summary

This section provided an answer to the first question posed at the end of section 6.1, namely: why do the binding conditions apply as expected in Ch'ol, but not in Chuj? I argued that the apparent violations of Condition C seen in section 6.1 (and §6.2.4) were not actually violations of Condition C. The general conclusion is that in all of the Chuj sentences that appear to obviate Condition C, there are no c-command relations between the relevant covalued nominal expressions, and so the binding conditions do not apply.[17]

Though we now understand why the surprising Chuj sentences seen in section 6.1 do not violate Condition C—they all involve *free* nominals—we have yet to discuss the fact that linear precedence matters for the distribution of free nominals in Chuj. We turn to this question in the next section.

6.3 Anti-cataphora in Chuj and Ch'ol

In the previous section, I argued that high-absolutive syntax often bleeds syntactic binding relations between covalued nominal expressions in Chuj. This in turn leads to con-

[17]Following the majority of works, I have assumed that the right 'command' notion for the binding conditions is *c*-command. As far as I can tell, however, other approaches, such as Bruening's (2014) *precedence* and *phase command*, could also be used to derive the differences between Chuj and Ch'ol in terms of syntactic binding. See Bruening 2014: §5.4 for a discussion of how precedence and phase-command could apply for VOS languages (crucially, only those that do not exhibit high-absolutive syntax).

figurations in which the binding conditions do not dictate the distribution of covalued nominals, because the relevant covalued expressions are *free*. However, we have yet to address the fact that Chuj prohibits cataphora between covalued nominals that are free. That is, configurations such as (357) are impossible in Chuj:

(357) Impossible extended reflexive construction in Chuj
*verb $[_{\text{OBJ}} \ldots [_{\text{POSS}} \textit{pro}_1]]$ $[_{\text{SUBJ}}$ R-expression$_1$]

In this section, I propose that the ungrammaticality of (357) is attributable to a general ban on cataphora for free pronouns, which as we will see, must also be posited to account for clear cases of free pronouns in not only Chuj, but also Ch'ol. In particular, I propose the following generalization:

(358) **Anti-cataphora with free pronouns in Chuj and Ch'ol**
If a pronoun is free and covalued with an overt nominal (R-expression/classifier pronoun), the overt nominal must linearly precede the free pronoun (linear precedence is relevant).

The ban on cataphora in (358) is surprising coming from the perspective of English. In English, cataphora is often taken to be permitted precisely in cases where nominals are free, and so not regulated by the binding conditions (e.g. Ross 1967; Reinhart 1983; Kayne 2002; Bruening 2014). Here are some examples:

(359) a. Those who know her$_1$ adore Zelda$_1$. (Reinhart 1983: (2))
b. Those who know Zelda$_1$ adore her$_1$

(360) a. Her$_1$ mother likes Bernice$_1$'s friends.
b. Bernice$_1$'s mother likes her$_1$ friends. (Bruening 2014: (6a-b))

The position of R-expressions and pronouns in (359) and (360) is flexible. For Reinhart (1983), this is directly attributed to the fact that the two nominals are not in a binding relation; they are both free.

Yet Chuj and Ch'ol differ from English, and the evidence for the constraint in (358) is overwhelming. Even in examples which indisputably involve free pronouns, the effects

of (358) can be noticed. Consider, for instance, the Chuj sentence in (361). With or without high-absolutive syntax, it is difficult to imagine how the two covalued expressions could c-command each other, since one expression is embedded inside the subject and the other is embedded inside the object. (361-b) illustrates that contrary to what we see in the English translations, cataphora remains illicit:

(361) Chuj free pronoun → linear precedence matters

a. *Tzschamk'olej stz'i' ix Ana ix ix ixlolon yet'ok.*
Tz-s-chamk'ol-ej [$_{OBJ}$ s-tz'i' **ix Ana**] [$_{SUBJ}$ ix ix ix-lolon
IPFV-A3-like-DTV A3-dog CLF Ana CLF woman PFV-speak
y-et'ok ***pro***].
A3-with PRON
Lit: 'The woman that spoke with her$_1$ likes Ana$_1$'s dog.'

b. *Tz-s-chamk'ol-ej [$_{OBJ}$ s-tz'i' ***pro***] [$_{SUBJ}$ ix ix ix-lolon y-et'
IPFV-A3-like-DTV A3-dog PRON CLF woman PFV-speak A3-with
ix Ana].
CLF Ana

Strikingly, the preference for linear precedence with free nominals is not limited to Chuj. In examples which clearly involve two free nominals which are covalued, Ch'ol also respects (358):

(362) Ch'ol free pronoun → linear precedence matters[18]

a. *Tyi ipejkä ajRosa jiñi x'ixik ta'bä ik'ele.*
Tyi i-pejk-ä [$_{OBJ}$ **aj-Rosa**] [$_{SUBJ}$ jiñi x-'ixik [$_{RC}$ ta'=bä
PFV A3-speak-DTV CLF-Rosa DET CLF-woman PFV=REL
i-k'el-e ***pro***]].
A3S-see-TV PRON
Lit: 'The woman who saw her$_1$ spoke with Rosa$_1$.'

b. *Tyi i-pejk-ä [$_{OBJ}$ ***pro***$_1$] [$_{SUBJ}$ jiñi x-'ixik [$_{RC}$ ta'=bä
PFV A3-speak-DTV PRON DET CLF-woman PFV=REL
i-k'el-e **aj-Rosa**$_1$]].
A3S-see-TV CLF-Rosa

[18]The sentence in (362-b) could have a grammatical disjoint interpretation: 'the woman who saw him/her$_2$ spoke with Rosa$_1$.' Note though, that (362-a) can also have a disjoint interpretation. Therefore, the preference for (362-a) cannot be due to ambiguity resolution.

In (362-a), an expression contained inside the subject is covalued with the object. Since Ch'ol is a low-absolutive language, this means that neither covalued expression c-commands the other, and by assumption, that the two expressions are not in a binding relation. What we see is that the R-expression must be realized first, in this case as the object. The opposite order, where the R-expression is realized inside the subject **??**, is judged ungrammatical under a joint reference reading.

We can further see that linear precedence matters in Ch'ol sentences like (362) when we consider comparable sentences with fronted subjects.

(363) Ch'ol free pronoun → linear precedence matters

a. *Jiñi x'ixik ta'bä ik'ele ajRosa tyi ipejkä.*
[$_{SUBJ}$ Jiñi x-'ixik [$_{RC}$ ta'=bä i-k'el-e **aj-Rosa**]] tyi
DET CLF-woman PFV=REL A3S-see-TV CLF-Rosa PFV
i-pejk-ä [$_{OBJ}$ ***pro***].
A3-speak-DTV PRON
'THE WOMAN WHO SAW ROSA$_1$ spoke with her$_1$.

b. *[$_{SUBJ}$ Jiñi x-'ixik [$_{RC}$ ta'=bä i-k'el-e ***pro***$_1$]] tyi i-pejk-ä
DET CLF-woman PFV=REL A3S-see-TV PRON PFV A3-speak-DTV
[$_{OBJ}$ **aj-Rosa**$_1$]
CLF-Rosa

In (363), in contrast with (362), the R-expression must now be realized inside the subject, and not as the object. This is because the subject now *precedes* the object. Again, this shows that linear precedence matters for the distribution of free pronouns, even in Ch'ol.

The generalization in (358) extends to other pronominal uses. Recall from section 6.1.1, and from previous chapters of this thesis, that Chuj employs overt pronouns in certain syntactic domains, roughly when two covalued expressions in a sentence are in different CPs (excluding relative clauses). Crucially, when a free overt pronoun is covalued with an R-expression, and no c-command relations hold between the two, the R-expression must *precede* the pronoun:

(364) Chuj

a. *Tato tzb'at waj Xun t'a Xan Matin, tejunk'o'olal olaj winh.*
[$_{CP}$ Tato tz-b'at **waj Xun** t'a Xan Matin], te-junk'o'olal ol-aj
if IPFV-go CLF Xun to San Mateo INTS-happy PROSP-be
winh.
CLF.PRON
'If Xun$_1$ goes to Xan Matin, he$_1$ will be very happy.'

b. *[$_{CP}$ Tato tzb'at **winh** t'a Xan Matin], tejunk'o'olal olaj **waj Xun**.
Intended: 'If he$_1$ goes to Xan Matin, Xun$_1$ will be very happy.'

In (364), the two covalued expressions are clearly not in a c-command relation, since one is contained inside a preverbal clausal adjunct. Yet again, what we see is that linear preference triumphs: the sentence in (364-b) is judged ungrammatical by speakers.

The Ch'ol counterparts of the Chuj sentences in (364) show the same linear precedence effects:

(365) Ch'ol

a. *Mi tyi majli ajJuan tyi San Cristóbal, tyijikña kej yubiñ.*
[$_{CP}$ Mi tyi majl-i **aj-Juan** tyi San Cristóbal], tyijikña kej y-ubiñ
if PFV go-IV CLF-Juan PREP San Cristóbal , happy PROSP A3-feel
pro.
PRON
'If Juan$_1$ goes to San Cristóbal, he$_1$'ll be happy.'

b. ??/*[$_{CP}$ Mi tyi majli ***pro*** tyi San Cristóbal] tyijikñaj kej yubiñ **aj-Juan**.
Intended: 'If he$_1$ goes to San Cristóbal, Juan$_1$'ll be happy.'

The preceding results are important because they show that both Chuj and Ch'ol are subject to the constraint on cataphora in (358), which applies between free R-expressions and covalued pronouns. The only difference between the two languages is the frequency with which anti-cataphora effects may be perceived. That is, since Chuj is language in which c-command relations between the subject and nominals inside the object are consistently disrupted by object raising, the anti-cataphora constraint can also be perceived, as opposed to in Ch'ol, in basic VOS sentences whenever the subject is covalued with nominal inside the object (i.e., all those that looked like violations of Condition C in §6.1).

As it turns out, Chuj and Ch'ol are not alone in imposing linear precedence constraints on the distribution of free nominals. Many researchers have noticed the existence of linear precedence constraints on pronominalization, exactly when the binding conditions do not apply (see e.g. Tai 1973 and Huang 1982 on Mandarin; Huang 1982 on Japanese; Mohanan 1981, 1983 on Malayalam; Chung 1989 on Chamorro; Kazanina and Phillips 2001, Reuland and Avrutin 2004, and Kazanina 2005 on Russian; and Christodoulou 2008 on Greek). For instance, Huang (1982) shows that linear precedence plays a central role for the distribution of (pro)nominals in Mandarin Chinese (see also Lust 1986; Lust et al. 1996, and Su 2020). He provides the following sentences:

(366) Mandarin Chinese (Huang 1982: 388)

a. [[da-le **Zhangsan**$_1$ de] neige ren], dui **ta**$_1$ hen bu keqi.
hit-ASP Zhangsan DE that man to him very not polite
'The man that hit Zhangsan$_1$ was very impolite to him$_1$.'

b. *[[da-le **ta**$_1$ de] neige ren], dui **Zhangsan**$_1$ hen bu keqi.
Intended: 'The man that hit him$_1$ was very impolite to Zhangsan$_1$.'

(367) Mandarin Chinese (Huang 1982: 389)

a. [[wo kanjian **Zhangsan**$_1$ de] shihou], **ta**$_1$ zai dazi.
I see Zhangsan DE time he at type
'When I saw Zhangsan$_1$, he$_1$ was typing.'

b. *[[wo kanjian **ta**$_1$ de] shihou], **Zhangsan**$_1$ zai dazi.
Intended: 'When I saw him$_1$, Zhangsan$_1$ was typing.'

The Mandarin Chinese data parallel those of Chuj and Ch'ol above: these languages impose constraints on cataphora. Notice that all of the (intended) translations of the ungrammatical sentences in Chuj, Ch'ol, and Mandarin examples seen above were grammatical in English. We therefore arrive at the following point of variation about the crosslinguistic distribution of free pronouns: while some languages ban cataphora with free nominals (Chuj, Ch'ol, Mandarin, etc.), others allow it (English, French, etc.).[19]

[19]The extent to which cataphora is tolerated could be gradable across languages. For instance, while cataphora is possible with English possessors (*Their*$_1$ *mother saw Kim*$_1$), my (native-speaker) judgement is that comparable sentences are impossible in French (**Sa*$_1$ *mère a vu Kim*$_1$). In a similar vein, while linear precedence restrictions with free nominals in Mandarin Chinese and Japanese seem widespread and robust, both languages allow cataphora under special circumstances. See discussion on pages 391 and 393 of Huang 1982.

As seen in section 6.1, it cannot be the case that anti-cataphora applies for the distribution of *all* covalued expressions, at least not in Ch'ol. In Ch'ol, which does not exhibit high-absolutive syntax, linear precedence is often irrelevant: when an R-expression c-commands a pronoun with which it is covalued, the binding conditions apply as expected regardless of whether the pronoun precedes the R-expression. This means that the anti-cataphora constraint in (358) can only apply to *free* expressions, and not to those which are *bound* under c-command. In other words, we arrive at the following two-fold generalization about the distribution of covalued nominal expressions in Chuj and Ch'ol, repeated from the introduction:

(368) Generalization about nominal expressions in Chuj and Ch'ol

a. If a nominal is bound, it is subject to structurally-sensitive binding conditions (linear precedence is irrelevant).

b. If a free pronoun is covalued with an R-expression, the R-expression must linearly precede the free pronoun (linear precedence is relevant). (repeated from (358))

The generalization in (368) points us to two clear conclusions. The first is that the grammars of Chuj and Ch'ol treat bound and free nominals in fundamentally different ways. This is not a controversial claim; virtually all theories of nominal covaluation draw a formal distinction between the two (e.g. Reinhart 1983, Grodzinsky and Reinhart 1993, Heim and Kratzer 1998, Hornstein 2001, Reuland 2001, Büring 2005, 2011, Rooryck and vanden Wyngaerd 2011; Bruening 2014). The second conclusion—and perhaps the more surprising one—is that Chuj and Ch'ol impose further restrictions on free expressions than other languages, such as English. As discussed, it is well-known that cataphora is possible with, precisely, *free expressions* in English (e.g. Reinhart 1983; Kayne 2002; Bruen-

ing 2014).[20] This led us to a crosslinguistic point of variation: in some languages, linear precedence regulates the distribution of free nominals; in others, it does not.

Though we have seen evidence for both generalizations in (368) in Ch'ol, we have not yet seen any evidence for (368-b) in Chuj. Section 6.2 only showed that Chuj was *compatible* with the binding conditions, and not that the binding conditions were actually operative in Chuj. The austerity of positive evidence is not entirely unexpected: high-absolutive languages are exceptional in that they rarely exhibit the configurations in which binding relations are expected to hold. This means that most cases of covalued nominals in Chuj involve free nominals, and thus we expect linear order to play a fundamental role.

In the next section, I argue that upon closer inspection, there are corners of Chuj where an R-expression syntactically-binds a pronoun. I show that in such cases the binding conditions apply as expected, and linear precedence is irrelevant. This ultimately supports the generalization in (368), and strongly reinforces the view that the binding conditions may reflect a universal feature of language (Grodzinsky and Reinhart 1993, Reuland 2010, 2011).

6.4 Binding under c-command, even in Chuj

My goal in this section is now to show that, in spite of their vast apparent inapplicability, the binding conditions *are* operative in Chuj. In section 6.4.1, we first consider sentences which, although entirely compatible with the binding conditions, cannot provide conclusive evidence that the binding conditions apply. This is so because of a confound from linear precedence: since the binders also precede their bindees, it is impossible to be certain that the binding conditions, and not linear precedence, are at issue. In section 6.4.2, I

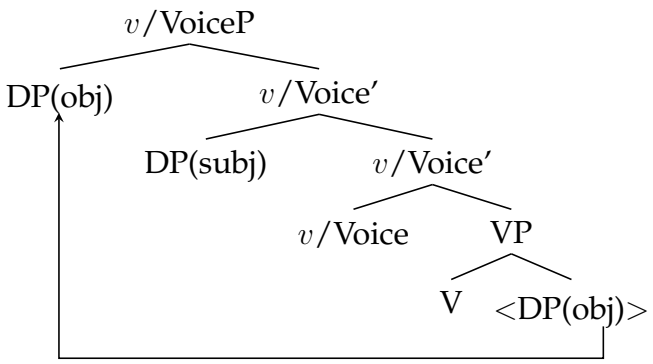
[20] As pointed out to me by Michael Wagner (p.c.), R-expressions can be anaphoric in English only if certain prosodic conditions are met, namely if they are destressed. Williams (1997) calls this process "anaphoric destressing" (see also Bianchi 2009). One way of accounting for the variation between English versus Chuj and Ch'ol with regards to (anti)-cataphora might therefore lie in the (un)availability of mechanisms in the respective languages allowing for anaphoric R-expressions, such as anaphoric destressing. Perhaps, Chuj and Ch'ol simply lack such a mechanism, leading to the overall ungrammaticality of sentences with anaphoric R-expressions.

then argue that reflexive sentences *do* provide conclusive evidence that the binding conditions are active in Chuj, since in such sentences, it can be shown that an R-expression can bind a pronoun which it not only c-commands but also linearly follows.

6.4.1 Absolutives binding into ergatives

In a high-absolutive configuration like (340-b), or (369) below, the ergative subject does not c-command the absolutive object. However, the object as a whole *does* c-command the subject.

(369) High-absolutive syntax (Chuj)



All else being equal, Condition C effects should emerge in Chuj data equivalent to (370), since the object c-commands the subject in the high-absolutive structure (cf. discussion in Trechsel 1995 on Popti'). That is, with the object c-commanding into the subject in Chuj, we predict that the R-expression should be forced to appear as the object in Chuj sentences equivalent to the English sentences in (370) (e.g., the literal Chuj equivalent of (370-b) should have to be *His*$_1$ *mother saw Xun*$_1$).

(370) a. The woman that saw Xun$_1$ scolded him$_1$.
b. Xun$_1$'s mother saw him$_1$.

Let us start with sentences like (370-a) in Chuj. As can be observed, the prediction is borne out. Only the object can be realized as an R-expression in (371).

(371) Chuj

a. *Ixstumej waj Xun ix ix ixilani.*
Ix-s-tum-ej [$_{OBJ}$ waj Xun] [$_{SUBJ}$ ix ix ix-il-an-i *pro*].
PFV-A3-scold-DTV CLF Xun CLF woman PFV-see-AF-IV *pro*
'The woman that saw Xun$_1$ scolded him$_1$.'
Lit: 'The woman that saw him$_1$ scolded Xun$_1$.'

b. *Ixstumej [$_{OBJ}$ *pro*] [$_{SUBJ}$ ix ix ixilan waj Xun].
Lit: 'The woman that saw Xun$_1$ scolded him$_1$.'

Though compatible with Condition C, the data in (371) cannot be taken as strong evidence that Condition C is operative. The reason is simply that Condition C and linear precedence cannot be dissociated: we cannot be certain whether the R-expression must be parsed as the object because of Condition C, or because the object linearly precedes the subject. Moreover, fronting the ergative subject to a preverbal position will not provide us with better insight. For the transitive subject to A'-extract in (371), an Agent Focus construction is needed. As discussed in section 6.2.1, this will result in a 'low-absolutive' configuration, and so no c-command relations will hold between the two covalued nominals.

What about sentences like (370-b), where the absolutive object is covalued with the possessor of the ergative subject? Interestingly, such sentences are ineffable in Chuj, an observation that has already been made for other Mayan languages, including both high- and low-absolutive languages (see e.g. Craig 1977, Aissen 1997, 1999, and Zavala 2007). Examples of ungrammatical sentences are provided below for Chuj:[21]

(372) Chuj

a. *Ix-y-il [$_{OBJ}$ *pro*] [$_{SUBJ}$ ix s-nun [$_{POSS}$ waj Xun]].
PFV-A3-see PRON CLF A3-mother CLF Xun
Intended: Xun$_1$'s mother saw him$_1$ (could mean 'Xun$_1$ saw his$_1$ mother.')

b. *Ix-y-il [$_{OBJ}$ waj Xun$_1$] [$_{SUBJ}$ ix snun [$_{POSS}$ *pro*$_1$]].

[21] Note that the ungrammaticality of the sentences in (372) does not arise from the inability for ergative subjects to bear possessors. Sentences of the type *Xun$_1$'s mother saw him$_2$* are grammatical.

The fact that sentences like (372) are also ineffable in low-absolutive Mayan languages, including Ch'ol (see Zavala 2007: 296 for relevant Ch'ol data), suggests that the ineffability of such sentences is not related to the low/high-absolutive parameter. In fact, Aissen (1997, 1999) and Zavala (2007, 2017) provide an analysis of the ineffability of sentences like (372) based on constraints related to *obviation* in Mayan. As far as the current proposal goes, however, the crucial point is that we cannot use data such as (370-b) to test for effects of the binding conditions.

In sum, though the prediction that the absolutive object binds inside the ergative subject in Chuj is not falsified, sentences like (371) also do not provide evidence for Condition C effects. We must therefore look at other kinds of data to test whether the binding conditions are operative in Chuj and can be dissociated from linear precedence constraints. We turn to this evidence in the next section.

6.4.2 When the binding conditions prevail: The case of reflexives

Here, I argue that, although Chuj in general exhibits high-absolutive syntax, there are exceptional cases where the external argument c-commands the internal argument. In such cases, the binding conditions are operative, and linear precedence is irrelevant.

Let us consider what would need to happen in order for the external argument to c-command the internal argument in a language like Chuj. There are two possibilities. Either (i) the internal argument exceptionally does not raise, or (ii) reconstruction of object raising is exceptionally possible, allowing for the internal argument to be in its base position for binding. Coon et al. (2021: §4.3) independently discuss option (ii) in relation to sentences in which the EEC is exceptionally circumvented in Chuj. Here, I show evidence for option (i) from the perspective of reflexive sentences.[22] In short, I show that reflexive

[22]Option (i) is of course also instantiated in the Agent Focus construction, where the use of the AF v/VoiceP head allows for the internal argument to remain *in situ* (Coon et al. 2014), and thus be asymmetrically c-commanded by the external argument. All cases involving AF are indeed compatible with the proposal that the binding conditions are active in AF constructions: in AF constructions, 'non-pronominal' arguments must be realized in the external argument position, as predicted given the binding conditions.

objects do not undergo object raising. Since high-absolutive syntax exceptionally does not happen, effects of Conditions A and C can be observed in reflexive sentences.

I first provide background on the internal syntax of reflexive anaphors in Mayan in section 6.4.2.1, arguing that they exhibit the same *internal* syntax as possessed nominals, similar to the one proposed for 'extended reflexives' in the previous sections. In section 6.4.2.2, however, I show that the *external* syntax of reflexives and extended reflexive differs, insofar as reflexive sentences do not exhibit high-absolutive syntax. The exceptional syntax of reflexive sentences ultimately allows us to observe Condition C effects from the external argument into the internal argument, despite the former following the latter.

6.4.2.1 The internal syntax of reflexive anaphors

Reflexive anaphors across Mayan pattern with possessed nouns in appearing with Set A (possessive) agreement and serving as the thematic object of transitive verbs (see Ayres 1980, Hou 2013, and Aissen 2017). Examples from Chuj and Ch'ol are provided below:

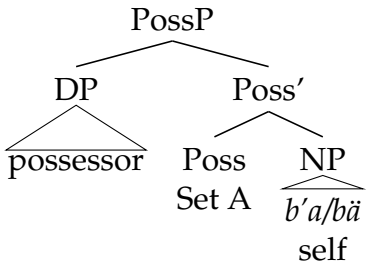
(373) *Ixyil sb'a ix Ana.*
Ix-y-il s-**b'a** ix Ana.
PFV-A3-see A3-self CLF Ana
'Ana saw herself.' (Chuj)

(374) *Tyi ik'ele ibä ajAna.*
Tyi i-k'el-e i-**bä** aj-Ana.
PFV A3-see-TV A3-self CLF-Ana
'Ana saw herself.' (Ch'ol)

Building on Coon 2017a, I propose that reflexive anaphors (in Chuj and Ch'ol) minimally exhibit the internal syntax in (375), with a possessor merged into a PossP specifier above the NP, which also triggers Set A agreement on the noun (see section 2.5.5).

However, since all AF constructions also involve A'-extraction of the external argument to a position that *linearly precedes* the internal argument, it is impossible to be certain whether it is truly the binding conditions, and not a constraint on cataphora, that adjudicates the distribution of nominals in AF constructions.

(375) Internal syntax of reflexive anaphors in Chuj and Ch'ol



Alternatively, one could consider the possibility that reflexive anaphors exhibit a simplex syntax, and do not project possessors, as proposed for the Mayan language Kaqchikel in Burukina 2019.[23] Since projection of a possessor is an essential component of the discussion to come, it is important to provide evidence for the structure in (375).

A first reason to think that anaphors in Mayan involve a complex possessive structure is that anaphors *look like* possessed nominals. In fact, the literature on Mayan languages normally describes anaphors as possessed or relational *nouns* (e.g. Ayres 1980; England 2001). Consider the following paradigms, which show that reflexive anaphors in both Chuj and Ch'ol pattern with regular possessed nouns in being inflected with Set A agreement for person and number:

(376) *Inflected reflexive in Chuj*
- a. hin-b'a 'myself'
- b. ha-b'a 'yourself'
- c. s-b'a 'himself/herself/themselves'
- d. ko-b'a 'ourselves'
- e. he-b'a 'yourselves'

(377) *Possessed noun in Chuj*
- a. hin-tz'i' 'my dog'
- b. ha-tz'i' 'your dog'
- c. s-tz'i' 'his/her/their dog'
- d. ko-tz'i' 'our dog'
- e. he-tz'i' 'y'all's dog'

(378) *Inflected reflexive in Ch'ol*
- a. k-bä 'myself/ourselves'
- b. a-bä 'yourself/yourselves'
- c. i-bä 'himself/herself/themselves'

(379) *Possessed noun in Ch'ol*
- a. k-ts'i' 'my/our dog'
- b. a-ts'i' 'your/y'all's dog'
- c. i-ts'i' 'his/her/their dog'

[23] Extending Labelle's (2008) analysis of French reflexives, Burukina argues that the reflexive counterpart of *b'a/bä* in Kaqchikel (*-i'*) is the overt realization of an agreement relation between a special reflexive v/Voice head and an object variable contained inside an unsaturated VP. For Burukina, the reflexive morpheme *-i'* is therefore of a similar kind as the reflexive clitic *se* in Romance languages like French.

A second reason to believe that reflexive anaphors in Chuj involve genuine possessive morphology, and are not simplex reflexive pronouns, is that this morphology disappears in syntactic environments that generally disallow possessives in Chuj. First consider the transitive nominalization in (380), derived with the suffix *-oj* (see Coon and Carolan 2017 and Coon and Royer 2020). The incorporated objects of such nominalizations can never appear with 'higher-level' nominal structure, including possessive morphology or determiners.

(380) *Ixayamoch moloj kape.* (Chuj; Coon and Royer 2020: 157)
Ix-a-yamoch [mol-oj (***te'**) (***ha**)-kape (***chi'**)].
PFV-A2S-begin gather-NMLZ CLF A2S-coffee DEIX
Intended: 'You started to gather the/your/that coffee.'

At least in some Mayan languages, reflexive anaphors can appear in such nominalizations, in which case they surface unpossessed. This is shown in (381). The fact that *b'a* appears unpossessed in environments where possessive morphology is blocked suggests that regular instantiations of reflexive anaphors involve genuine possessive structure in examples like (376) and (378).

(381) Chuj
a. *Ixb'at winh b'o'oj b'a'il.*
Ix-b'at winh [b'o'-oj b'a'-il].
PFV-go CLF.he make-NMLZ self-NML
'He began to prepare himself.'
b. *Hanhej iloj b'a'il tzyak' winh.*
[Ha-nhej il-oj b'a'-il] tz-y-ak' winh.
FOC-only see-NMLZ self-NML IPFV-A3-do CLF.him
'All he does is to look at himself.'

Also note that in the examples in (381), the reflexive noun *b'a* 'self' is suffixed with *-Vl* (whose vowel is partially harmonic with the vowel of the nominal root, realized as *-il* in (381)). Though the exact distribution of *-Vl* in Chuj requires further work, *-Vl* sometimes

appears on inalienable nominals which exceptionally appear unpossessed (Maxwell 1981, Buenrostro 1996 and Royer et al. to appear).

In sum, the fact that (i) *b'a* can appear unpossessed in environments that disallow possessive morphology and (ii) that it can bear *-Vl* morphology characteristic of nominals in Chuj, provides further evidence that normal instances of reflexive anaphors exhibit a complex internal syntax. In particular, these facts suggest that reflexive anaphors normally project possessive structure (though not in (381-a)), as proposed in (375).

6.4.2.2 Detecting the binding conditions in Chuj

I have just argued that reflexive anaphors project possessors in Chuj and Ch'ol. But what kind of nominal is instantiated in the possessor position? That is, in Chuj sentences equivalent to *Xun saw himself*, is the R-expression in subject position, as in (382-a), or is it instantiated in possessor position, as in (382-b)?

(382) Two logical parses of 'Xun saw himself' in Chuj

a. saw $[_{\text{OBJ}}$ self $[_{\text{POSS}}$ *pro*]] $[_{\text{SUBJ}}$ Xun]

b. saw $[_{\text{OBJ}}$ self $[_{\text{POSS}}$ Xun$_i$]] $[_{\text{SUBJ}}$ *pro*$_i$]

Since Chuj normally exhibits high-absolutive syntax, we might expect—though we will shortly see this is not the case—that reflexive sentences should pattern exactly like extended reflexive sentences from section 6.1.3.1 in showing the pattern in (382-b), where the possessor is instantiated by the R-expression and the subject is a null pronoun. After all, the only surface difference between the reflexive in (383) and the extended reflexive in (384) below appears to be in the choice of the noun, and we know from section 6.1.3.1 that in extended reflexive constructions like (384), it is the possessor that is instantiated by the R-expression, and not the subject. In fact, it was the similarity between (383) and (384) which led Aissen (1997) to name constructions like (384) "extended reflexives".

(383) Chuj reflexive

Ixyil sb'a waj Xun.

Ix-y-il s-**b'a** waj Xun.

PFV-A3-see A3-self CLF Xun

'Xun$_1$ saw himself$_1$.'

(384) Chuj extended reflexive

Ixyil stz'i' waj Xun.

Ix-y-il s-**tz'i'** waj Xun.

PFV-A3-see A3-dog CLF Xun

'Xun$_1$ saw his$_1$ dog.'

Despite their surface similarity, there is reason to think that the R-expression *waj Xun* in the sentences in (383) and (384) occupies different syntactic positions, with the reflexive exhibiting the parse in (382-a). Recall from section 6.1.3.1 that adverbs can normally intervene between the external argument and the internal argument in Chuj, but that this exceptionally does not hold in extended reflexives. These kinds of data were taken as evidence that in such cases, the R-expression is the possessor, and not the subject. A new minimal pair is provided below. Again, the only difference between (385-a) and (385-b) is the presence of Set A (possessive) agreement on the noun *tz'i'* 'dog'.

(385) a. Chuj transitive sentence

Ixyil nok' tz'i' junelxo waj Xun. / Ixyil nok' tz'i' waj Xun junelxo.

Ix-y-il nok' tz'i' {junelxo} waj Xun {junelxo}.

PFV-A3-see CLF dog again CLF Xun again

'Xun saw the dog again.'

b. Chuj extended reflexive construction

Ixyil nok' stz'i' waj Xun junelxo.

Ix-y-il nok' **s**-tz'i' {*junelxo} waj Xun {junelxo}.

PFV-A3-see CLF A3-dog again CLF Xun again

'Xun$_1$ saw his$_1$ dog again.'

Consider now an example with a reflexive object, as in (386). The judgments from speakers are robust: reflexives pattern differently than extended reflexives in allowing identical adverb placement options as regular transitive sentences:

(386) *Ixyil sb'a junelxo waj Xun. / Ixyil sb'a waj Xun junelxo.*

Ix-y-il s-**b'a** {junelxo} waj Xun {junelxo}.

PFV-A3-see A3-self again CLF Xun again

'Xun$_1$ saw himself$_1$ again.' (Chuj; compare with (385-a))

Because adverbs do not intervene internal to possessive phrases (see (313) above), this suggests that the possessor is null and the subject overt in reflexive sentences like (386), as schematized in (387-a), to be contrasted with the extended reflexive syntax for (385-b) in (387-b).

(387) a. Chuj reflexive in (386)
saw [$_{\text{OBJ}}$ self [$_{\text{POSS}}$ *pro*]] {again} [$_{\text{SUBJ}}$ Xun] {again}

b. Chuj extended reflexive in (385-b)
saw [$_{\text{OBJ}}$ dog [$_{\text{POSS}}$ Xun$_i$]] {again} [$_{\text{SUBJ}}$ *pro*$_i$] {again}

Especially striking is the fact that linear precedence, or the anti-cataphora constraint in (358), is irrelevant in deriving reflexive sentences, in stark contrast with all of the examples we saw in section 6.1, where linear precedence played a crucial role. That is, under the assumption that reflexives in Chuj involve possessors, (386) shows us that the linearly second of two covalued DPs gets realized as an R-expression, and therefore that the anti-cataphora constraint from (358) cannot be regulating the distribution of these nominals.

I propose that (387-a), contrary to (387-b), simply exemplifies a Condition C effect: the R-expression must appear in subject position, because the subject c-commands the object (and its possessor). Crucially, for Condition C to be active, it cannot be the case that high-absolutive syntax occurs, otherwise the subject would not c-command inside the reflexive anaphor. We therefore arrive at the following hypothesis:

(388) **Proposal about reflexive sentences in Chuj**
Transitive sentences with reflexive objects do not exhibit high-absolutive syntax.

Given (388), (i) the ergative subject will bind the reflexive object, (ii) Condition C will be operative, and (iii) linear precedence is expected to play no role in determining how the relevant covalued expressions should be realized, as the data in (386) indicate (and as is the case more generally in low-absolutive objects in Ch'ol). That is, the anti-cataphora constraint only applies to free expressions, and since the *pro* in (387-a) is bound by the ergative subject, Condition C will block it from occurring as an R-expression.

As it turns out, there is strong support both cross-linguistically and Chuj-internally for the proposal in (388). First, there is good reason to believe that universal constraints on reflexive anaphors (Condition A) should be incompatible with high-absolutive configurations. Prohibitions on the realization of anaphors in external argument positions have long been observed (Anderson 1976), and Brodkin and Royer (to appear) recently argue more specifically that this ban—one which follows from virtually all modern approaches to Condition A (see §6.5 for one such approach based on Hornstein 2001 and Zwart 2002)—extends to high-absolutive languages (in Mayan and beyond). Therefore, assuming (i) that there is a crosslinguistic ban on ergative anaphors (Brodkin and Royer to appear), (ii) that Condition A requires reflexive anaphors to be locally bound (Reinhart 1983, a.o.), and (iii) that A-movement does not reconstruct for binding (Chomsky:1995mp, a.o.), then it is not surprising that high-absolutive sentences with reflexive anaphors should be ineffable (after object raising, the anaphor will not be locally bound by the external argument). Under such circumstances, we may therefore expect reflexive anaphors to remain low.

Second, a closer look at the sentences containing reflexive anaphors in Chuj reveals robust evidence for the proposal that reflexive objects do not raise above the ergative subject. In fact, reflexive sentences show a constellation of properties, which we turn to now, that are to be expected if the object remains low.

For one, recall that transitive subjects are blocked from extracting in high-absolutive languages, an effect which is claimed to be caused by high-absolutive syntax. As noted in much previous work on the EEC (see e.g. Craig 1977, Mondloch 1981, Ordóñez 1995, Aissen 1999, 2017, Pascual 2007, Coon and Henderson 2011, Hou 2013, Velleman 2014, Coon et al. 2014, Coon et al. 2021), this constraint is exceptionally circumvented with reflexive objects, as shown in (389) (this is the case across Mayan; cf. Aissen 2017, table 30.3):

(389) *¿Mach ixyil sb'a?*
¿Mach ix-y-il s-**b'a**?
who PFV-A3-see A3-self
'Who saw themself?' (Chuj; compare with (331) above)

The absence of EEC effects in (389) could be explained if object raising fails to occur (Ordóñez 1995, Coon et al. 2014): no intervention effect will arise, and so the transitive subject should be free to extract (though see §6.4.3 shortly for further discussion of these and related data). Data like (389) are thus consistent with the proposal that reflexive anaphors remain low in Chuj.

A second piece of evidence for the low position of reflexive anaphors comes from the observation that reflexive objects cannot themselves undergo A'-extraction:

(390) *Ha s-**b'a** ix-y-il waj Xun.
FOC A3-self PFV-A3-see CLF Xun
Intended: 'Xun saw HIMSELF.' (Chuj)

Assuming, following Coon et al. (2014) and Coon et al. (2021), that arguments must first move to the edge of the v/VoiceP phase in order to A'-extract (in both Chuj and Ch'ol), we can potentially make sense of this restriction. Contrary to non-reflexive objects, reflexives can never raise to that position, and can therefore never be A'-extracted.

A third piece of evidence comes from coordination. Reflexive objects cannot coordinate with other DPs:

(391) *Ix-y-il [$_{\&P}$ s-**b'a** yet' ix Malin] winh k'ayb'um.
PFV-A3-see A3-self and A3 Malin CLF teacher
Intended: 'The teacher saw himself and Malin.' (Chuj)

The impossibility of coordinating reflexive anaphors with regular DPs is not surprising given the current proposal. If one conjunct requires a low-absolutive syntax (the reflexive), and the other requires a high-absolutive syntax (the other DP), then it is not surprising that such sentences should result in ineffability.

Word order provides a fourth argument. Recall from section 6.1.3.5 that Chuj exceptionally allows VSO word order when the transitive subject is a classifier pronoun (or when the subject is phonologically-light; see Clemens and Coon 2018). This option is not allowed when the internal argument is a reflexive anaphor, VOS being the only possible order:[24]

(392) Chuj

a. *Ixyil sb'a winh.*
Ix-y-il [$_{OBJ}$ s-**b'a**] [$_{SUBJ}$ winh].
PFV-A3-see A3-self CLF.PRON
'He saw himself.' (VOS)

b. *Ix-y-il [$_{SUBJ}$ winh] [$_{OBJ}$ s**b'a**].

Though I remain agnostic about how postverbal word order is derived, these data again support the existence of an important structural distinction between reflexive and non-reflexive objects.

A final piece of evidence in favour of a structurally-low position for reflexive anaphors comes from other Mayan languages, such as Kaqchikel (a high-absolutive language), which feature overt third person absolutive (Set B) agreement in the plural. As discussed by Burukina (2019), plural reflexive anaphors (393-a), contrary to other plural third person internal arguments (393-b), do not trigger Set B (absolutive) plural agreement (see also Can Pixabaj 2015 for similar claims in K'iche'):

[24]The same word order facts are observed in Mayan languages that otherwise exhibit rigid VSO word order, such as Mam, Popti', and Q'anjob'al. In these languages, reflexives exceptionally trigger *VOS* word order (see e.g. England 1983: (7-43) on Mam; Craig 1977: 217, on Popti'; and Coon et al. 2014: (77) on Q'anjob'al).

(393) Kaqchikel → no Set B agreement with reflexives

a. *Rije xekitz'et ki'.*
Rije x-**(*e)**-ki-tz'ët k-i'.
PRON.3P PFV-B3P-A3P-see A3P-REFL
'They saw themselves.' (Burukina 2019: (2))

b. *Yïn xeintz'ët rje'.*
Yïn x-**e**-in-tz'ët rje'.
I PFV-B3P-A1S-see they
'I saw them.' (Imanishi 2019: (6))

Again, this makes sense if reflexive anaphors remain low in the structure. In Coon et al. 2021, raising of the object leads to an Agree relation with T/Infl, the locus of absolutive morphemes. If reflexive objects never raise in the first place, failure of agreement is expected.

To summarize, a constellation of properties about reflexive anaphors in Chuj provide strong support for the view proposed here that they do not undergo object raising. This in turn leads to the possibility of binding relations from the external argument into the reflexive object.

Though the ultimate takeaway of this section should be that reflexive sentences in Chuj show Condition C effects due to their lack of high-absolutive syntax, it is interesting to consider *why* reflexive sentences should lack high-absolutive syntax. One reason might simply be that high-absolutive syntax, for some reason or other, creates ineffable results for Condition A (see Brodkin and Royer to appear for discussion). Another reason, already proposed in previous work on Mayan (Ordóñez 1995, Coon et al. 2014, Coon et al. 2021), could be that reflexive anaphors in Chuj and other Mayan languages exhibit a reduced syntax, and simply lack the functional structure corresponding to DP. For example, if anaphors are structurally-reduced, then we might expect them not to be targeted by the EPP feature that causes object raising in the first place—the EPP feature could be relativized to probe only for DPs. If this is the correct explanation, some of the data discussed in this subsection could be attributed to anaphors lacking a DP layer. For instance,

the impossibility of extracting reflexive objects (390) could be due to relativization of A'-probes to DPs (in the spirit of Coon et al. 2021), and the impossibility of coordinating reflexive anaphors with regular DPs (391) could be recast as a constraint on coordinating expressions of different categories. Under this type of analysis, structurally-reduced reflexives could still project 'lower' possessors, as proposed for inalienable nouns in recent work (e.g., Alexiadou 2003, Myler 2014, and Tyler 2021). Moreover, though not illustrated in this section, all of the properties discussed for reflexive anaphors in Chuj also hold of Ch'ol reflexives. This could be taken as evidence that reflexive anaphors in general lack functional structure found in other types of transitive objects, even canonical low-absolutive objects in Mayan languages like Ch'ol. Exploring these options further promises to be deeply informative for our understanding of Mayan syntax, and perhaps of Condition A anaphors more generally. For the local purposes, however, the important point is simply that reflexives in Chuj do not exhibit high-absolutive syntax, and thus provide us with the opportunity to diagnose instantiations of the binding conditions.

Before summarizing this section, the next subsection briefly discusses a piece of data that potentially complicates the view that reflexives and extended reflexives exhibit a different syntax. I nonetheless show that, when we consider the larger set of syntactic properties used to diagnose the low position of reflexive objects, it becomes clear that extended reflexives generally pattern with high-absolutive transitives, and unlike reflexive sentences, thus supporting the proposal that reflexive sentences exhibit a different syntax.

6.4.3 Extended reflexives and EEC circumvention

In the previous subsection, I proposed that transitive sentences with reflexive anaphors in Chuj do not exhibit high-absolutive syntax. I argued that a potential correlate of this property lies in the fact that reflexive sentences circumvent the EEC. Again, the key property to notice is that the subject of a reflexive sentence can exceptionally undergo A'-extraction

without Agent Focus morphology, normally necessary for the A'-extraction of transitive subjects in Chuj.

(394) *¿Mach ixyil sb'a?*
¿Mach ix-y-il s-b'a?
who PFV-A3-see A3-self
'Who saw themself?' (Chuj)

I argued that the above data makes sense if reflexive anaphors never undergo object raising, thus providing evidence for their 'low' position. While data like (389) are consistent with the internal argument remaining low, these data may not be a direct argument for the low position of the internal argument. In fact, as pointed out by a reviewer on related work, this logic leads to an apparent complication. As noted in much work on the EEC (Craig 1977, Mondloch 1981, Ordóñez 1995, Aissen 1999, Aissen 2017, Pascual 2007, Coon and Henderson 2011, Hou 2013, Velleman 2014, Coon et al. 2014, Newman 2020, Coon et al. 2021), in a subset of high-absolutive languages, including Chuj, extended reflexive sentences also appear to circumvent the EEC, allowing agent extraction without Agent Focus morphology. Notice, for instance, the lack of Agent Focus in the following well-formed question:

(395) *¿Mach ixyil snun?*
¿Mach ix-y-il ix s-nun?
who PFV-A3-see CLF A3-mother
'Who saw their mother?' (Chuj)

On the surface, the non-compliance of extended reflexives with the EEC looks like an empirical puzzle for the current analysis. In particular, central to my proposal is that Chuj extended reflexives systematically exhibit high-absolutive syntax. This proposal is instrumental in explaining why no Condition C violations arise in Chuj when the possessor of an *extended reflexive* object is an R-expression, and conversely, why Condition C effects *do* arise in Chuj *reflexive* sentences (as well as in comparable extended reflexive sentences in low-absolutive Ch'ol). But if the absolutive object raises above the ergative subject in Chuj

extended reflexives, then it is not clear why the EEC should be circumvented in cases like (395). In fact, the expectation is that extended reflexive objects should *induce* the EEC.

While data like (395) look problematic on the surface, it is important to emphasize that extended reflexive objects otherwise behave like transitive objects with regards to all of the other properties used to diagnose the low position of reflexive objects in the previous subsection. These differences are summarized in Table 2.

Table 6.2: Shared properties of reflexives, ext. reflexives, and transitives in EEC languages

Reflexives	Extended reflexives	Other transitives
cannot be A′-extracted (390)	can be A′-extracted (§6.1.3.4)	can be A′-extracted
cannot be coordinated (391)	can be coordinated (§6.1.3.2)	can be coordinated
rigid VOS (392)	VSO possible (§6.1.3.5)	VSO possible
no Set B agreement (393-a)	Set B agreement (393-b)	Set B agreement
no EEC (389)	no EEC (395) (subset of high-abs lang.)	EEC compliant

The only case in which reflexives and extended reflexives seem to pattern alike is in their apparent ability to circumvent the EEC. Moreover, this does not hold across the Mayan language family as a whole. While the EEC seems to be systematically circumvented in reflexive sentences across the entire language family, only a subset of Mayan languages appear to circumvent the EEC with extended reflexives (see e.g. Hou 2013, Aissen 2017). Furthermore, as seen throughout this chapter, the two constructions behave differently in terms of the binding conditions: while extended reflexives in Chuj do not show Condition C effects, reflexives do.

Given this state of affairs, it is reasonable to assume that extended reflexives do exhibit high-absolutive syntax, and therefore that the non-compliance of the EEC in reflexives versus extended reflexives must receive *different explanations*, a point that is also made by Coon et al. (2021: §4.3). Assuming, as in section 6.4.2.2, that reflexive sentences circumvent the EEC because they do not exhibit high-absolutive syntax, the question becomes how extended reflexives can also circumvent the EEC *while still showing a high-absolutive syntax*. That is, we need a high-absolutive theory of extended reflexive objects (as is the

case with other transitive objects), which can also explain why no Agent Focus is needed in (395). Though more work is needed to flesh out the syntax and semantics of extended reflexive constructions, and its relation to the EEC, one relevant observation is that the EEC is a constraint on the extraction of *ergative subjects*. Since reflexive and extended reflexive constructions involve possessors that are covalued with the ergative subject, it may be a hasty conclusion to assume that what extracts in (395), and potentially also (394), is the ergative subject. All else being equal, extraction of the possessor could also be at issue. I leave exploring this possibility for future work.

6.4.4 Summary

In sum, this section has argued that there are corners of Chuj where the external argument c-commands the internal argument. In such cases, if the external argument is covalued with an expression contained inside the internal argument, the binding conditions prevail, and linear precedence becomes irrelevant. This supports the two-fold generalization in (368): while the distribution of free nominals is constrained by an anti-cataphora restriction, bound nominals are subject to the binding conditions, regardless of whether they precede their binder. Therefore, despite preliminary reasons to doubt the existence of the binding conditions in Chuj—and in turn their universality—we find evidence that they are sometimes operative.

6.5 Interface conditions on bound and free nominals

The Chuj and Ch'ol findings of this chapter have led us to the generalization about the distribution of nominal expressions in (368), repeated below. A proper theory of the Chuj and Ch'ol patterns of nominal covaluation will therefore have to account for this generalization.

(396) Generalization about nominal expressions in Chuj and Ch'ol

a. If a nominal is bound, it is subject to structurally-sensitive binding conditions (linear precedence is irrelevant).

b. If a pronoun is free and covalued with an overt nominal (R-expression/classifier pronoun), the overt nominal must precede the free pronoun (linear precedence is relevant).

It remains to be understood where exactly in the grammar these generalizations apply. Here, I provide some first steps towards answering this question, by suggesting that both generalizations can be viewed as regulated by the syntax-phonology interface. In section 6.5.1, I first argue, following Craig (1977) and Aissen (2000) on Popti', that since the distribution of free pronouns in Chuj and Ch'ol is sensitive to linear order, it is natural that the generalization in (396-b) should be regulated by PF. Adopting the view that patterns of "pronominalization" may sometimes result from ellipsis processes at PF, as proposed in Chapter 3 for Chuj (see e.g. Postal 1966, Craig 1977, Elbourne 2001, 2005), I further argue that the generalization in (396-b) follows form a more general ban on backwards ellipsis, which is consistently judged unacceptable in Chuj.[25] In section 6.5.2, I then conjecture more generally that the generalizations in (396) could boil down to the way the phonology handles externally merged ("free pronouns") versus internally merged covalued nominals ("bound pronouns"). This view of nominal covaluation has the advantage that it allows for a simple account of the difference between bound and free nominals, a difference that moreover relies on independently-needed assumptions about the grammars of Chuj and Ch'ol. Finally, in section 6.4.3, I show how the line of analysis presented in this section could potentially shed light on a well-known puzzle about extended reflexives in a subset of high-absolutive languages: that they seem to circumvent the EEC (Craig 1977, Mondloch 1981, Ordóñez 1995, Aissen 1999, Aissen 2017, Pascual 2007, Coon and Henderson 2011, Hou 2013, Velleman 2014, Coon et al. 2014, Coon et al. 2021).

[25] As discussed in Chapter 3, this is not to say that all of the items typically described as "pronouns" involve ellipsis.

6.5.1 Anti-cataphora as a ban on backwards deletion

Let us start with the anti-cataphora generalization in (368-b)/(396-b). Craig (1977) and Aissen (2000), on the Mayan language Popti', concur in proposing that the ban on cataphora should be viewed as a product of deletion or ellipsis, which I assume applies at PF (see Merchant 2019 for PF approaches to ellipsis). This view relies on the well supported hypothesis that a subset of "pronouns" in Chuj and Ch'ol, or at least the ones discussed in this chapter, result from partial or full ellipsis at PF (see e.g. Postal 1966, Craig 1977, Elbourne 2001, 2005, and Chapter 3). Building on these works, I propose the PF principle in (397), which is simply a more precise formulation of the generalization in (396-b). Notice that the generalization in (397) makes reference to *indices*, which violates *Inclusiveness* (Chomsky 1995, 2001), an issue I return to in the conclusion.

(397) **PF principle against cataphora with free nominals**
If two or more free expressions bear the same index within the same clause, only the linearly first can be realized overtly (as an R-expression/classifier pronoun), and the rest must undergo deletion.

The principle in (397) essentially states that when two co-indexed expressions are not in a syntactic binding relation (neither c-commands the other), the one that comes linearly first is privileged for pronunciation. Crucially, it also takes as a basic assumption that free nominals are externally merged as full nominal expressions in the syntax, and that they undergo ellipsis under identity with a nominal bearing the same index in the same clause. An example for an extended reflexive in Chuj is provided below for illustration (where "< >" indicates elided copies in movement chains and strikeout indicates ellipsis without movement):

(398) Chuj extended reflexive constructions

a. *Ixschonh swakax ix Ana.*
Ix-s-chonh [$_{OBJ}$ s-wakax [$_{POSS}$ ix Ana]] [$_{SUBJ}$ *pro*] .
PFV-A3-sell A3-cow CLF Ana PRON
Lit: 'She$_7$ sold Ana$_7$'s cow.'

b. Numeration: { Ana, Ana, cow, sell, T^0 ... }

c. [sold [$_{OBJ}$ cow [$_{POSS}$ Ana$_7$]]$_i$ [$_{SUBJ}$ ~~Ana$_7$~~] <OBJ>$_i$]

As schematized in (398-c), I propose that the lexical item *Ana*, bearing the index 7, is externally merged twice in the derivation (notice that it appears twice in the numeration). After object raising, both instantiations of the expression *Ana* are free. Therefore, the anti-cataphora principle in (397) applies and the second DP undergoes ellipsis. To be precise, this means that what has been referred to as a null pronominal in Chuj sentences like (398-a) actually corresponds to an elided DP.

The revised constraint in (397) ties anti-cataphora to a ban on backwards ellipsis. We might therefore expect backwards ellipsis to be generally constrained. Preliminary investigation from Chuj suggests that this expectation is borne out, as can be observed in the sentences in (399) to (401), which showcase different kinds of ellipsis (note that backwards ellipsis or backwards cataphora is tolerated in all English translations; my judgements, verified with two speakers):

(399) DP ellipsis in Chuj

a. *Tato tzyal yak'an te' son waj Xun, tzyal pax yak'an ix Malin.*
Tato tz-yal y-ak'-an **te'** **son** waj Xun, tz-yal pax
if IPFV-can A3-give-DEP CLF marimba CLF Xun IPFV-can also
y-ak'-an ~~**te' son**~~ ix Malin.
A3-give-DEP CLF marimba CLF Malin
'If Xun can play the marimba, Malin can play it too.

b. *Tato tzyal yak'an ~~**te' son**~~ waj Xun, tzyal pax yak'an **te' son** ix Malin.
Intended 'If Xun can play it, Malin can play the marimba too.[26]

[26]The English translation sentence does not actually involve DP ellipsis, but a cataphoric pronoun.

(400) Sluicing in Chuj

a. *Ay jun mach ixjawi, pero machekel mach.*
Ay jun mach **ix-jaw-i**, pero ma-chekel mach ~~**ix-jaw-i**~~.
EXT some WHO PFV-arrive-IV, but NEG-know who PFV-arrive-IV
'Someone arrived, but I don't know who ~~arrived~~.

b. *Machekel mach ~~**ixjawi**~~, pero ay junmach **ixjawi**.
Intended: 'I don't know who ~~arrived~~, but someone arrived.'

(401) Stripping in Chuj

a. *Tato tzsman jun yonh waj Xun, ha ix Malin paxi.*
Tato **tz-s-man** **jun y-onh** waj Xun, ha ix Malin paxi
if IPFV-A3-buy one A3-avocado CLF Xun, FOC CLF Malin also
~~**tz-s-man**~~ ~~**jun y-onh**~~ ~~**ix**~~.
IPFV-A3-buy one A3-avoado CLF.PRON
'If Xun buys an avocado, so will Malin ~~buy an avocado~~.

b. *Tato ha waj Xun paxi ~~**tzsman jun yonh**~~, **tzsman jun yonh** ix Malin.
Intended: 'If Xun does ~~buy an avocado~~, Malin will buy an avocado too.'

While forward ellipsis is clearly possible in Chuj, attempts at constructing sentences with backwards ellipsis in this language are consistently judged ungrammatical. These data are revealing: the constraint on cataphora with free nominals (397), which builds on previous work on closely-related Popti' (Craig 1977, Aissen 2000), leans on the assumption that externally-merged nominal expressions that are covalued with another externally-merged nominal expression within the same sentence undergo *ellipsis*. Thus, it is conceivable that the anti-cataphora constraint is part of a larger, language-specific constraint on ellipsis resolution.[27]

6.5.2 Binding as internal merge

If free nominals in Chuj involve the external merger of co-indexed nominals, which are in turn constrained by the PF principle in (397), then what component of grammar enforces the binding conditions for nominals that are in a c-command relation? While the

[27] The connection between ellipsis and free nominals would not be novel: work as early as Ross 1967 proposed to group together patterns of cataphora in English with those involving backwards ellipsis.

results of the chapter are compatible with different approaches to the binding conditions, a natural hypothesis is that the distribution of bound nominals is also regulated by PF. And since the anti-cataphora constraint targets *externally* merged nominals, we can consider whether syntactic binding could be representative of how PF constrains *internally* merged nominals. Under this view, bound nominals would effectively be interpreted as part of movement chains, which, given the Extension Condition (Chomsky 1995), should necessarily be sensitive to c-command.

Several authors have proposed that syntactic binding should involve some form of internal merge (Hornstein 2001, 2007; Kayne 2002; Zwart 2002, Boeckx et al. 2007; Rodrigues 2010; Sauerland 2013; Charnavel and Sportiche 2021). For instance, Hornstein (2001) and Zwart (2002) propose that Condition A is the reflex of local movement from a position within a complex anaphor to the external argument position, a theory that relies on the assumption that movement to a thematic position is possible (Hornstein 1999, 2001). The sensitivity of Condition A to c-command then follows from the Extension Condition (Chomsky 1995), which essentially ensures that the moved constituent will c-command its copy. In such accounts, movement of a DP into a non-thematic position is preferred over external merger of a co-indexed expression (see Hornstein 2007: p. 51, for discussion; see also Shima 2000 for relevant discussion).[28] As illustration, we can take a Chuj reflexive sentence as in (402). Recall that reflexives exceptionally do not exhibit high-absolutive syntax, and so the ergative subject c-commands the reflexive object. In light of Hornstein 2001 and Zwart 2002, we can derive this sentence by moving the possessor of the reflexive NP *b'a* 'self' in (402-a) to the external argument position. Assuming a copy theory of movement (Chomsky 1993; Bobaljik 2002), this will give rise to the structure in (402-c) (see Newman 2020 p. 40 for a related proposal on Q'anjob'al reflexive and extended reflexive constructions).

[28] An economy principle such as 'merge-over-move' must still be preferred for nominals in thematic positions to avoid over-generating sentences like Kim_i *loves* $<Kim_i>$ to mean 'Kim loves themself'.

(402) Chuj reflexive

a. *Ixyil sb'a waj Xun.*
Ix-y-il [s-**b'a**] waj Xun.
PFV-A3-see A3-self CLF Xun
'Xun_1 saw himself$_1$.'

b. Numeration: { Xun, b'a, see, T^0 ... }

c. [saw [$_{\text{OBJ}}$ self [$_{\text{POSS}}$ <Xun>]] [$_{\text{SUBJ}}$ Xun]]

In (402), notice that the numeration, contrary to the one provided for the extended reflexive in (398), contains only one instance of the lexical item *Xun*. This expression is first externally merged in the possessor position of the internal argument, headed by the nominal *b'a* 'self', which triggers Set A morphology (*s-*) on the noun. The same lexical item, *Xun*, is then internally merged in the thematic position of the external argument. Crucially, the phonological component will treat the lower copy as any other lower copy is usually treated in a movement chain in Chuj: the structurally highest copy will be privileged for pronunciation, while the lower copy will be deleted. Following a number of previous works (Chomsky 1993, 1995; Pesetsky 1998; Franks 1999; Bošković 2001; Bošković and Nunes 2007), I therefore assume the following PF principle, which privileges pronunciation of nominals in the head of movement chains, regardless of linear order (see e.g. Chomsky 1993: 35):

(403) **PF principle on pronunciation of copies in movement chains**
In Chuj and Ch'ol, the structurally highest copy in a movement chain is privileged for pronunciation as an R-expression (or classifier pronoun).

Since reflexive sentences do not exhibit high-absolutive syntax (402-c) (see also §6.4), a movement-based account thus correctly predicts that the R-expression will appear in subject position in (402-a). More generally, the movement-based account naturally derives Condition C effects. R-expressions that are covalued with pronouns (formally copies here) that they c-command will, as per (403), be consistently realized at the top of movement chains, which in turn means that they will always be "free".

Comparing the derivation for the reflexive sentence in (402) with that of the Chuj extended reflexive sentence in (398), reflexives differ from extended reflexives in three critical ways: (i) they do not undergo object raising, (ii) the covalued possessor/subject nominal is externally merged only once in the derivation (and not twice), and (iii) the same nominal in the possessor is then internally merged in subject position (represented in a right-side specifier to show VOS order, but see §6.2 above on word order and linear order). This is schematized in the examples (404) and (405), repeated from above. I assume that (404) is also the derivation for Ch'ol reflexives. In (404), the principle on movement chains (403) will force *Xun* to appear in subject position (deriving Condition C). In (405), the anti-cataphora principle (397) will come into effect, forcing *Ana* to appear in the linearly first externally-merged nominal bearing the same index.

(404) Chuj/Ch'ol reflexive → no object raising, 'Xun' externally merged once and re-merged in subject

[saw [$_{\text{OBJ}}$ self [$_{\text{POSS}}$ <Xun>]] [$_{\text{SUBJ}}$ Xun]]

(405) Chuj extended reflexive → object raising, '*Ana*' externally merged twice followed by deletion of linearly second 'Ana'

[sold [$_{\text{OBJ}}$ cow [$_{\text{POSS}}$ Ana$_7$]]$_i$ [$_{\text{SUBJ}}$ ~~Ana$_7$~~] <OBJ>$_i$]

Notice that a movement-based theory can also be extended to derive all of the data seen in section 6.1 for Ch'ol, which shows Condition C effects in extended reflexive sentences. Take, for instance, the Ch'ol sentence in (406-a). Ch'ol is a low-absolutive language, which means that the external argument will c-command the internal argument, and therefore the possessor of the internal argument in extended reflexive constructions like (406-a). This sentence would therefore be derived exactly as a reflexive is derived in (404).

(406) Ch'ol extended reflexive (low-absolutive)

a. *Tyi ichoño iwakax ajAna.*
Tyi i-choñ-o [$_{OBJ}$ i-wakax] [$_{POSS}$ *pro*] [$_{SUBJ}$ aj-Ana].
PFV A3-sell-TV A3-cow *pro* CLF-Ana
'Ana$_1$ sold her$_1$ cow.'

b. Numeration: { Ana, cow, sell, T^0 ... }

c. Possessor A-movement to subject position
[sold [$_{OBJ}$ cow [$_{POSS}$ <Ana>]] [$_{SUBJ}$ Ana]]

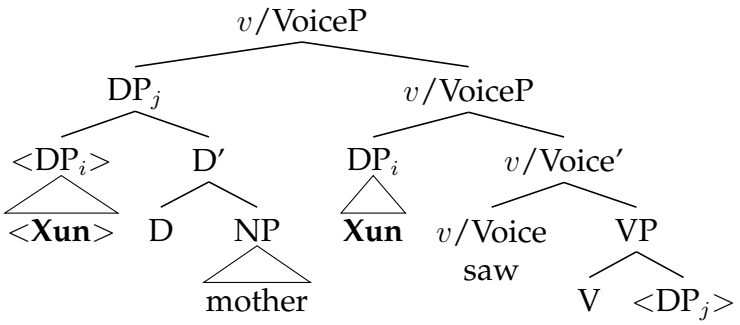
In (406-b), the numeration only contains one instance of the lexical item *Ana*. *Ana* is first externally merged in the possessor position, and then internally merged in the position of the external argument. As is always the case in Ch'ol, PF will delete the lower copy inside the movement chain, leading to the "pronominalization" of the lower copy, and thereby deriving the effects of Condition C.

Finally, it is important to ask why internal merge of the possessor into the subject is only possible with reflexives, and not with Chuj extended reflexives. One possibility is that the latter would lead to an ineffable structure for the semantics, and is thereby ruled out. Following previous work, I assume that the lower copy inside of a movement chain gets interpreted as a bound variable at LF (see e.g., Heim and Kratzer 1998; Sauerland 1998; Fox 2002; and Poole 2017 on trace/copy conversion). Assuming further that variable binding requires c-command (Reinhart 1983, Grodzinsky and Reinhart 1993, Büring 2005), consider what happens if the object raises past the subject:

(407) Internal merge of subject + object raising leads to an illicit structure

a. Numeration: { Xun, mother, see, T^0 ... }

b. *



In (407), the possessor is internally merged in the subject. The object from which the possessor moved subsequently undergoes remnant movement over the subject, which I assume cannot reconstruct for binding (see §6.2.2).[29] The crucial point to notice is that after the object undergoes raising, the higher copy of *Xun* does not c-command its copy/trace. If variable binding requires c-command, as standardly assumed, then this sentence will not be interpretable at LF. I thus propose that, under normal circumstances (though see §6.4.3), the derivation for an extended reflexive in Chuj can only proceed if the possessor and subject are both externally merged, as in (398-c) above. This in turn will always feed the application of the anti-cataphora constraint in (397). In semantic terms, we predict that the two co-indexed expressions in Chuj extended reflexives will only (accidentally) corefer (Reinhart 1983): the sameness of their semantic value arises only because they bear the same index. Therefore, the two nominals in (405) will be semantically independent of each other, insofar as they are not in a binder-bindee relationship in the semantics.

On the other hand, in the case of Chuj and Ch'ol reflexive sentences (404), as well as Ch'ol extended reflexives (406), no object raising occurs. A derivation with internal merge of the possessor into the subject is therefore perfectly licit for the purposes of semantics: the copy in the subject will c-command the copy in the possessor, and the possessor will be interpreted as a bound variable at LF.

In sum, I have proposed that the conditions on free and bound nominals can both be understood as conditions imposed by the interfaces, if one adopts a movement-based theory of syntactic binding. The resulting distinction essentially boils down to differences in the way the interfaces treat internally merged and externally merged nominals, or, in the terminology of Chomsky (2013), *copies* and *repetitions*. Under this theory, syntactically-bound expressions are treated as tails in a movement chain, and are expected to be deleted if PF privileges pronunciation of the structurally highest copy (403). This could well be what underlies the binding conditions in (288). Repetitions of co-indexed lexical items, on

[29] Issues for (407) might also arise due to more general constraints on remnant movement (see e.g. Müller 1996 and Grewendorf 2015).

the other hand, may be subject to independent PF constraints. In some languages, such as English, no specific PF constraints on the realization of repetitions appear to be imposed, as evidenced by the availability of cataphora and backwards ellipsis.[30] In others, such as Chuj and Ch'ol, language-specific PF constraints on repetitions, which essentially ban backwards ellipsis (397), come into effect.

This approach of course raises many questions. For instance, Bruening (2021) recently notes a number of arguments against movement-based approaches to syntactic binding in English. Though these arguments do not necessarily extend to Mayan, the stance I have taken is that the binding conditions are universal, and so data from English remain relevant for a theory that reduce patterns of syntactic binding to movement.[31]

6.6 Summary

This chapter started with a puzzle laid out in section 6.1: while Ch'ol patterns as expected in terms of the classical binding conditions, Chuj seems, at first glance, to largely ignore them. Instead, we saw that Chuj resorts to linear precedence in order to constrain the distribution of covalued nominals. This seemed to cast doubt on the widely-held assumption that the binding conditions reflect a universal property of language (e.g., Grodzinsky and Reinhart 1993, Reuland 2010, 2011).

[30]The PF constraints do not have to be trenchant: it might be the case that certain configurations, but not all, are subject to constraints on free expressions. Bruening (2014) takes the presence of linear precedence effects as a diagnostic for syntactic binding, which leads him to the conclusion that there are binding relations between the two covalued nominals in sentences like *Penelope cursed Peter*$_1$ *and slandered him*$_1$ (Bruening 2014: (37a), cited from Langacker 1969), since it is not possible to get a joint reference reading by flipping the order of the R-expression and the pronoun. Given the Mayan data seen in this chapter, however, we could reinterpret these data as an indication that there are sometimes linear precedence effects between free expressions, even in English.

[31]A lingering question is how to handle syntactic binding within larger, long-distance domains, without running into issues such as island violations. One possibility is that there is no long-distance syntactic binding in Chuj. Based on Aissen 1992, 2000, Royer (2022) argues that complement clauses and clausal adjuncts obligatorily extrapose in Chuj. If this is so, then covaluation between expressions inside a main clause and expressions inside complement clauses or clausal adjuncts would necessarily involve free covaluation, and would therefore not have to be derived through movement.

In section 6.2, I argued that denying the universality of the binding conditions is not necessary, since a convergence of factors is responsible for the apparent inapplicability of the binding conditions in Chuj. The main proposal was that in every configuration in which the binding conditions are inoperative, there are no c-command relations between the relevant covalued nominal expressions, and so the binding conditions are predicted not to apply. Central to this proposal was the claim that, contrary to Ch'ol, Chuj exhibits 'high-absolutive syntax', independently-proposed to underlie a number of syntactic phenomena, including the Ergative Extraction Constraint (EEC) (Coon et al. 2014; Coon et al. 2021). The lack of Condition C effects in Chuj, but not in Ch'ol, were thus taken as another, pervasive symptom of a deep syntactic difference between two types of Mayan languages. The fact that a new correlate of high-absolutive syntax has been identified also offers strong empirical support for high-absolutive approaches to the EEC.

In section 6.3, I then argued that the linear precedence effects seen in Chuj are part of a larger, and perhaps cross-Mayan, constraint on the treatment of free nominals. Contrary to languages like English, Chuj and Ch'ol both ban free pronouns from being covalued with an R-expression which they precede. Since the Chuj sentences which at first glance appeared to violate Condition C actually involve free nominals, the importance of linear precedence in deriving these sentences falls out naturally.

In section 6.4, I then argued that, upon closer inspection, there is evidence that Chuj is sensitive to the binding conditions after all. Most notably, I showed that there are exceptional cases where the internal argument does not raise over the subject, the main one being when the internal argument is a reflexive object. In such cases, linear precedence does not matter for the realization of covalued expressions, and the binding conditions apply as expected. These findings led to the important conclusion that constraints on linear precedence in Chuj apply exclusively in cases of free nominals, and not in cases of syntactic binding.

The result is a unified set of constraints on pronominalization in Chuj and Ch'ol, despite the two languages looking entirely disparate at the beginning of the chapter. That is, both languages are subject to the binding conditions, and both languages show a ban on cataphora for free nominals. The observed differences between Chuj and Ch'ol simply boil down to idiosyncratic syntactic properties of the former language, such as high-absolutive syntax, which obscures the application of the binding conditions under most circumstances.

There are important avenues for future work. For instance, in an effort to make progress on the understanding of patterns of nominal covaluation in Mayan, I have focused on a relatively small sample of syntactic domains. However, it would be interesting to explore how instances of long-distance binding unfold in Mayan languages like Chuj and Ch'ol. As we saw in section 6.1.1, Chuj also features overt classifier pronouns, which can be covalued with antecedents within larger domains (see e.g. Royer to appearb). It would therefore be worthwhile to explore the range of environments in which overt pronouns can be distributed, and whether or not overt pronouns can also be conclusively shown to be subject to the binding conditions. Another interesting area of future work concerns the distinction between syntactic binding and semantic binding. In this chapter, I mostly ignored patterns of variable binding, since a comprehensive description of patterns of quantificational binding and bound variable anaphora in Mayan is still pending. It would be interesting to see to what extent syntactic and semantic binding exhibit overlap, and whether they are amenable to a unified analysis.

In section 6.5, I also provided some speculation about how bound and free nominals may be differentiated in grammar. The main proposal was that the difference between the two could reduce to how PF interprets nominal identity under internal merge and external merge. I hypothesized that while syntactically-bound expressions could be treated as copies in a movement chain, a PF constraint on linear precedence forces R-expressions to be realized as the first of a series of externally merged expressions that bear the same

index. Regardless of whether this theory is on the right track, this discussion will hopefully serve as a basis for future work on nominal covaluation in Mayan and beyond. In a way, Mayan languages offer an ideal vantage point to further explore patterns of nominal covaluation. For one, languages like Chuj and Ch'ol exhibit a number of language-specific properties, such as systematic variation in the position of the internal argument and a constraint on linear precedence, which allow for a clear delimitation of free and bound nominals. Second, the clear morphosyntactic constitution of reflexive anaphors, and their radically-simplified distribution (e.g., there are no cases of 'exempt' anaphora), make reductionist theories of binding, such as movement-based approaches, not only straightforward, but highly appealing.

Finally, I close this chapter with a comment on what I think is an important theoretical implication of this work, having to do with the status of indices in grammar.

The current analysis of the distribution of free nominals in Mayan relies on the ability for the phonological component to be able to identify which free nominal expressions are covalued. As already noted by Aissen (2000) in related work on Popti', this has important ramifications for the status of indices in grammar. The PF generalization in (397), or more generally linear precedence constraints on free nominals, require the phonological component to have access to information about how nominal expressions are to be contextually interpreted, at least indirectly. In other words, PF needs access to "indices".

(408) **PF principle against cataphora with coreferential expressions:**
If two or more free expressions are co-indexed within the same clause, realize the linearly first as an R-expression, and elide the others. (repeated from (397))

But if the phonological component can see indices, then indices must be syntactically-represented. This conclusion runs counter to much recent work on syntactic binding, which follows Chomsky (1995, 2001) in assuming the Inclusiveness Condition:

(409) **Inclusiveness** (Chomsky 2001, 2-3) (cited from Collins and Groat 2018).
[Inclusiveness] bars introduction of new elements (features) in the course of computation: **indices**, traces, syntactic categories or bar levels, and so on.

Inclusiveness has led to a body of fruitful research on syntactic binding, and the absence of indices in syntax has become widely adopted as an underlying assumption (Hornstein 2001, 2007; Kayne 2002; Zwart 2002; Safir 2004; Rooryck and vanden Wyngaerd 2011; Reuland 2011, etc.). For instance, Reuland (2017, 371) recently states: "as is uncontroversial since Chomsky (1995), [...] syntactic indices cannot be part of UG".

But the distribution of free nominals in Mayan languages like Chuj and Ch'ol provides an interesting challenge for Inclusiveness. It is not with bound pronouns, but with free pronouns, that the existence of indices in syntax becomes crucial. That is, though it is possible to derive patterns of syntactic binding without resorting to syntactically-represented indices—this is what the above-cited derivational approaches aim to do—it is difficult to imagine how we could state a phonological constraint such as (408) without appealing to syntactically-represented indices, at least if we want to keep to the T architecture of grammar. After all, the only way PF can see that two free expressions are covalued is for these expressions to be *identified* as covalued, which requires a device such as indices. The data discussed in this chapter therefore support several recent proposals, including some of the proposals made in the previous chapters of this work, which rely on the availability of indexed expressions in syntax (see e.g. Heim and Kratzer 1998; Elbourne 2008; Schwarz 2009; Clem 2019; Arregi and Hanink 2018; Jenks 2018; Hanink 2018, 2021; Jenks and Konate to appear).

Chapter 7

Conclusion

At the beginning of this thesis, we set out to explore three questions about the nature of nominal expressions, with a focus on definite, indefinite, demonstrative, and pronominal elements. These are repeated below:

(410) **Questions**

Q1 What morphosyntactic elements are involved in the composition of nominal expressions used to pick out or quantify over entities?

Q2 What semantic and pragmatic distinctions are expressed by different combinations of functional items within the nominal domain?

Q3 What conditions are imposed on co-referential and bound nominal expressions within sentences?

Chapters 3–5 focused on the first two questions. I showed that the radical decomposition of the extended nominal domain in Chuj, as well as in other languages belonging to the Q'anjob'alan sub-family of Mayan languages, provides key insight into the morphosyntactic composition of nominal expressions. In particular, an in-depth analysis of different combinations of functional items within the nominal domain guided us to identify a wealth of fine-grained semantic and pragmatic distinctions established by different functional items in Chuj. These distinctions involved different flavours of definiteness and indefiniteness, with many of the same building blocks combining in different ways

to deliver subtle meaningful contrasts, both within and across languages. The combinations of functional items proposed to underlie the relevant flavours of definiteness and indefiniteness, both null and overt, are summarized in Table 7.1, repeated from Table 5.3.

Table 7.1: Singular DP configurations in Chuj

	Configuration	Output in Chuj	Rough translation
①	ι + CLF + NP	*nok' tz'i'*	'the dog'
②	ι + CLF + NP + DEIX$_i$	*nok' tz'i' chi'*	'the/that dog' (anaphoric)
③	ι + CLF + *pro*$_i$	*nok'*	'it'
④	ι + CLF + <NP>	*nok'*	'it'
⑤	$\exists$ + NP	*jun tz'i'*	'a dog'
⑥	$\exists$ + CLF + NP	*jun nok' tz'i'*	'a (certain) dog'
⑦	$\exists$ + NP + DEIX$_i$	*jun tz'i' chi'*	'that dog'
⑧	$\exists$ + CLF + NP + DEIX$_i$	*jun nok' tz'i' chi'*	'that dog' (exophoric)

Lines ①–③ involve expressions which I have classified as "definite", whereas lines ④–⑦ involve those which I have classified as "indefinite". As made salient from this table, noun classifiers play a critical role in the composition of both definite and indefinite expressions. One of the major goals of this thesis was therefore to provide a unified analysis of these elements, one which can explain their semantic contribution within both definite and indefinite DPs.

In the definite domain, a total of three flavours of definiteness were identified, which I proposed in Chapter 5 all share the property of instantiating ι in the head of DP. First, building on diagnostics put forth by Schwarz (2009, 2012), I showed that Chuj establishes an overt distinction between *weak* and *strong* definite expressions, thereby echoing previous patterns seen across a number of unrelated languages (see e.g., Schwarz 2019 and Aguilar-Guevara et al. 2019). Roughly, while weak definites, which involve combining a noun classifier with a noun ①, trigger a uniqueness presupposition (for singular expressions), strong definites, which involve adding additional index-introducing morphology to a weak definite base, are used when the satisfier of the NP is either discourse anaphoric or deictically accessible within an immediate situation. Building on Bi and Jenks 2019, I also extended this theory of definiteness to pronouns, which I argued come in two guises

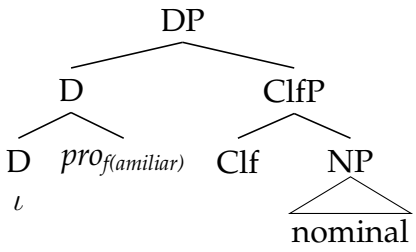
according to the weak/strong definite distinction. On the one hand, based on previous work (Postal 1966, Cooper 1979, Heim 1990), I argued that classifier pronouns can sometimes be E-type pronouns: weak definite determiners which combine with a covert index-introducing predicate. In such cases, they represent a strong definite description. On the other hand, I argued that Chuj classifier pronouns sometimes arise as a result of NP ellipsis, in which case they in fact denote a weak definite description (Elbourne 2001, 2005). I further suggested that more than one flavour of "strong" definiteness may be encoded in Chuj. In particular, I presented evidence that while situational uniqueness must hold for strong definites that combine with a deictic particle (line ②), anaphoric uses of pronouns (line ③) involve an indexical argument that applies in the scope of the uniqueness trigger, yielding a result equivalent to the denotations for strong definites provided in previous work (e.g., Schwarz 2009, Arkoh and Matthewson 2013, Jenks 2018, Hanink 2021). The resulting theoretical picture, which falls out as an immediate consequence of the rich extended nominal domain in Chuj, is that definiteness can in fact be a decompositional notion. This aligns with much recent work, which has proposed a decomposition of definiteness, in some form or the other (e.g., Simonenko 2014, Coppock and Beaver 2015, Hanink 2018, Ahn 2019).

In the indefinite domain, I proposed that different flavours could also be expressed, with, again, much reason to be convinced that these differences are derived compositionally. In particular, I argued that while all indefinite expressions involve an overt indefinite quantifier in the head of the DP, functional projections below the DP in the extended nominal domain serve to impose restrictions, encoded via semantic presuppositions, on the domain of this quantifier. For example, inspired by Schwarzschild (2002), I argued that the same lexical entry for noun classifiers used in definite descriptions could combine with an indefinite determiner to yield a *singleton indefinite*, which I argued explains why indefinite + classifier expressions force specific (or seemingly 'wide scope') interpretations of indefinites. This provided an explanation for the otherwise puzzling fact that

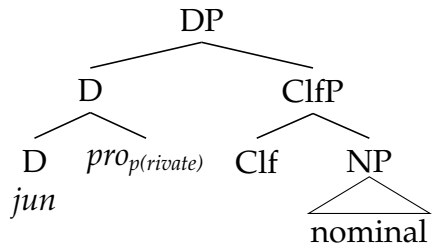
noun classifiers, which effectively exhibit a weak definite semantics, could felicitously combine with indefinite determiners. Likewise, I showed that the same deictic particles used to form strong definites could also combine with indefinite determiners, delivering meanings reminiscent of previously noted indefinite uses of demonstratives in other languages (see e.g., Arsenijević 2018, Doran and Ward 2019). While more work is needed to flesh out the semantic contribution and distribution of indefinite demonstratives in Chuj, the entries and syntax for deictic particles proposed in this thesis ultimately explain why deictic particles are compatible with both definite and indefinite bases.

Since there is overlap in the morphosyntactic composition of definiteness and indefiniteness in Chuj, it is interesting to consider what is the overarching factor that makes definiteness different from indefiniteness. I proposed that this distinction should essentially boil down to what is instantiated in the head of DP, namely ι versus *jun*. More specifically, I argued that these different heads select different kinds of free situation variables, in accordance with the kinds of conditions they impose on the discourse participants' ability to retrieve the intended contextual domain restriction with which the relevant expression is delimited. These two options are schematized again below:

(411) Weak definite description



(412) Indefinite+classifier DP



In the case of definite descriptions, I argued that ι must combine with a *familiar situation pronoun*, one whose value is known to both the speaker and their audience. Consequently, the content of the semantic presupposition of functional items below ι was also interpreted relative to this familiar situation pronoun. In the case of indefinite determiners, on the other hand, I argued that these involve a *private situation pronoun*, one whose value the

speaker assumes is not necessarily known to their audience. In such cases, I argued that the presuppositional content of functional items below the indefinite determiner does not behave as a run-of-the-mill presupposition, insofar as it does not result in a pragmatic presupposition. This allowed us, for example, to provide a unified analysis of noun classifiers, all while keeping to the core proposal that they encode a uniqueness and existence presupposition. That is, when the non-assertive content of noun classifiers is evaluated based on a private situation pronoun, it yields not a canonical weak definite presupposition, but a singleton indefinite (that there is a unique satisfier of the NP in a situation unidentifiable to the hearer). In other words, despite initial reason to believe that a weak definite entry for noun classifiers should be incompatible with indefinite expressions, the view entertained here allows us to unify the contribution of functional items within the extended nominal domain in Chuj, all while accounting for the interpretive distinctions that arise from combining these items with different DP heads.

The resulting analysis of different functional items within the nominal domain in Chuj leads to a number of interesting empirical and theoretical consequences. For one, the elaborate decomposition of the nominal domain in Chuj leads to the possibility that certain functional categories generally regarded as "primitives" may in fact not be. Instead, we saw that functional items generally described as *definite articles*, *indefinite articles*, *demonstratives*, or *pronouns*, form partially overlapping heterogeneous classes in Chuj, with several pieces partaking in the formation of each kind of expression. This leads to the expectation that perhaps other languages work this way (despite not showing a clear decomposition on the surface), an expectation that aligns with a number of recent works that have argued for a decomposition of seemingly simplex morphemes within the extended nominal domain (see e.g., Déchaine and Wiltschko 2002, Leu 2008, 2015, Coppock and Beaver 2015, Arsenijević 2018, Hanink 2018, Ahn 2019, Hsu and Syed 2020).

A second consequence concerns the relation between lexically encoded semantic presuppositions and how these map to pragmatic presuppositions. That is, I argued that

the semantic presupposition associated with a given lexical item in Chuj only results in a pragmatic presupposition in a subset of cases. This amounts to saying that *Stalnaker's Bridge*, which implies that semantic presuppositions should translate into pragmatic presuppositions, is too strong (see e.g., von Fintel 2008 on Stalnaker's Bridge). Instead, under the current system, whether or not a semantic presupposition is mapped to a pragmatic presupposition depends on whether this expression combines with a familiar or private situation pronoun. I also suggested that since *Maximize Presupposition* applies to *pragmatic* presuppositions, only expressions which combine with a familiar situation pronoun should show effects of *Maximize Presupposition*. This can in turn explain why noun classifiers are obligatory in Chuj definite descriptions, but optional with specific indefinites.

Crucially, the proposal that semantic presuppositions do not always map to pragmatic presuppositions is not novel. Matthewson (2006), for instance, argues that not all languages seem to possess pragmatic presuppositions in the sense of Stalnaker 1974, St'át'imcets being the example at hand. Building on Gauker's (1998) approach to presuppositions, Matthewson argues for a macro-linguistic parameter: while some linguistic expressions require their presuppositional content to be in the Common Ground (and therefore involve "Stalnaker Presuppositions"), others do not (and therefore involve "Gauker Presuppositions"). This theory could be recast under the theory proposed in this thesis. Perhaps, free situation variables in St'át'imcets are never specified as *familiar*. If this were the case, then we could potentially explain why the semantic presuppositions associated with determiners and pronouns in St'át'imcets and other Salish languages seem to generally not lead to Stalnaker pragmatic presuppositions (Matthewson 1999, 2008, Davis 2006, Gillon 2006). More generally, it will be interesting to determine to what extent the proposals put forth in this thesis for Chuj could be extended to explain related phenomena about the non-assertive component of lexical items in other languages.

Finally, in Chapter 6, we zoomed out of the internal syntax of nominal expressions to focus on how covalued nominal expressions interact within Chuj and Ch'ol sentences,

thereby providing insight into the third question posed in (410) (what conditions are imposed on co-referential and bound nominal expressions within sentences?). To make progress, I focused on patterns of syntactic binding between non-quantificational expressions. Based on previous work on Popti' (Craig 1977, Hoekstra 1989, Trechsel 1995, Aissen 2000), I showed that while Ch'ol behaves entirely as expected with regards to the classic Binding Conditions, Chuj appears to consistently tolerate violations of Condition C, often privileging linear precedence as the determining factor in the distribution of R-expressions and pronouns. The Chuj data thus initially seemed to cast doubt on a long tradition to treat the Binding Conditions as universal (e.g., Grodzinsky and Reinhart 1993, Reuland 2010, 2011).

I argued that the difference between Chuj and Ch'ol could be largely explained if Chuj, contrary to Ch'ol, exhibits 'high-absolutive' syntax, a proposal that was independently put forth to account for a deep syntactic parameter within the Mayan language family (Coon et al. 2014, Coon et al. 2021). In short, I argued that high-absolutive syntax creates configurations in which the internal argument asymmetrically c-commands the external argument, bleeding otherwise expected binding relations from the external argument into the internal argument. The violations of Condition C in Chuj are thus only apparent. I further argued (i) that linear precedence effects in Chuj are a reflex of a more general anti-cataphora constraint on free nominals, which can also be shown to apply to Ch'ol, and (ii) that there are corners of Chuj where the binding conditions do apply, and that in such cases linear precedence is irrelevant for the distribution of covalued nominals. As a result, we arrived at the following two-fold generalization about the distribution of covalued nominal expressions in Chuj and Ch'ol:

(413) **Generalization about covalued expressions in Chuj and Ch'ol**

a. If a nominal is bound under c-command, it is subject to structurally-determined Binding Conditions (linear precedence is irrelevant).

b. If a free pronoun is covalued with an R-expression, the R-expression must linearly precede the free pronoun (linear precedence is relevant).

Crucially, the Binding Conditions are active in Chuj, even though idiosyncratic syntactic properties of the language sometimes bleed their application in cases where binding condition violations would otherwise be expected. The general lesson is that despite initial evidence to doubt the universality of the Binding Conditions, a universalist approach can not only be maintained, but is supported by the Chuj data.

In conclusion, this thesis reinforces the view that work on understudied languages provides critical clues into topics at the core of debates in linguistic theory. Indeed, we saw that Chuj's generous decomposition of elements inside the extended nominal domain provide a unique vantage point into the primitive building blocks of different kinds of nominal expressions. We also saw that Chuj and Ch'ol exhibit several language-specific properties, such as systematic variation in the position of the internal argument, as well as a constraint on linear precedence, which shed important light into the kinds of conditions, both universal and not, that regulate the distribution of covalued nominal expressions across languages. While much work remains—within Mayan and cross-linguistically—Mayan languages promise to be highly relevant for future theoretical endeavours into the syntax and semantics of the nominal domain and other connected topics.

Bibliography

Abbott, Barbara. 2004. Definiteness and indefiniteness. In *The Handbook of Pragmatics*, eds. Laurence R. Horn and Gregory Ward, 122–150. Oxford, UK: Blackwell.

Abels, Klaus, and Ad Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15:25–74.

Abney, Steven P. 1987. The English noun-phrase in its sentential aspect. Doctoral dissertation, MIT, Cambridge, MA.

Abusch, Dorit. 1994. The scope of indefinites. *Natural Language Semantics* 2:83–135.

Aguilar-Guevara, Ana. 2014. *Weak definites: Semantics, lexicon and pragmatics*. Landelijke Onderzoekschool Taalwtenschap.

Aguilar-Guevara, Ana, Julia Pozas Loyo, and Violeta Vázquez-Rojas Maldonado. 2019. *Definiteness across languages*. Language Science Press.

Ahn, Byron Thomas. 2015. Giving reflexivity a voice: Twin reflexives in English. Ph.D. dissertation, UCLA, Los Angeles, CA.

Ahn, Dorothy. 2019. The determinacy scale: A competition mechanism for anaphoric expressions. Ph.D. dissertation, Harvard University, Cambidge, MA.

Ahn, Dorothy, and Kathryn Davidson. 2018. Where pointing matters: English and Korean demonstratives. In *Proceedings of North East Linguistics Society 48*.

Aikhenvald, Alexandra. 2000. *Classifiers: a typology of noun categorization devices*. Oxford University Press.

Aissen, Judith. 1992. Topic and focus in Mayan. *Language* 68:43–80.

Aissen, Judith. 1996. Pied-piping, abstract agreement, and functional projections in Tzotzil. *Natural Language and Linguistic Theory* 14:447–491.

Aissen, Judith. 1997. On the syntax of obviation. *Language* 73:705–769.

Aissen, Judith. 1999. Agent focus and inverse in Tzotzil. *Language* 75:451–485.

Aissen, Judith. 2000. Prosodic conditions on anaphora and clitics in Jakaltek. In *The Syntax of Verb Initial Languages*, eds. Andrew Carnie and Eithne Guilfoyle. OUP.

Aissen, Judith. 2011. On the syntax of agent focus in K'ichee'. In *Proceedings of FAMLi*, eds. Kirill Shklovsky, Pedro Mateo Pedro, and Jessica Coon. MIT Working Papers in Linguistics.

Aissen, Judith. 2017. Correlates of ergativity in Mayan. In *Oxford Handbook of Ergativity*, eds. Jessica Coon, Diane Massam, and Lisa Travis. New York: OUP.

Aissen, Judith, Nora C. England, and Roberto Zavala Maldonado, eds. 2017. *The Mayan Languages*. Routledge Language Famliy Series. New York: Routledge.

Aldridge, Edith. 2004. Ergativity and word order in Austronesian languages. Doctoral dissertation, Cornell University, Ithaca, NY.

Alexiadou, Artemis. 2003. Some notes on the structure of alienable and inalienable possessors. In *From NP to DP Volume II: The expression of possession in noun phrases*, eds. Martine Coene and Yves D'Hulst, 167–188. John Benjamins Publishing Company.

Alexiadou, Artemis, Liliane Haegeman, and Melita Stavrou. 2007. *Noun phrase in the generative perspective*. de Gruyter Mouton.

Alonso-Ovalle, Luis, and Paula Menéndez-Benito. 2010. Modal indefinites. *Natural Language Semantics* 18:1–31.

Alonso-Ovalle, Luis, and Paula Menéndez-Benito. 2018. Projecting possibilities in the nominal domain: Spanish *uno cualquiera*. *Journal of Semantics* 35:1–41.

Alonso-Ovalle, Luis, and Justin Royer. 2020. Random choice modality: The view from Chuj (Mayan). In *Proceedings of Sinn und Bedeutung 24*, eds. Michael Franke, Nikola Kompa, Mingya Liu, Jutta L. Mueller, and Julianne Schwab, 48–65. Osnabrück: Osnabrück University.

Alonso-Ovalle, Luis, and Justin Royer. 2021. Random choice from likelihood: The case of Chuj. *Journal of Semantics* 38:483–529.

Alonso-Ovalle, Luis, and Justin Royer. 2022. Modality in the nominal domain: The Chuj *Yalnhej* dps. In *Semantics and Linguistic Theory*, eds. Nicole Dreier, Chloe Kwon, Thomas Darnell, and John Starr, volume 31, 61–81. Linguistic Society of America.

AnderBois, Scott, and Robert Henderson. 2015. Linguistically establishing discourse context: two case studies from Mayan languages. In *Methodologies in Semantic Fieldwork*, eds. Ryan Bochnak and Lisa Matthewson, 207–232. Oxford University Press.

Anderson, Stephen. 1976. On the notion of subject in ergative languages. In *Subject and Topic*, ed. C.N. Li. New York: Academic Press.

Anvari, Amir. 2018. A problem for Maximize Presupposition! (Locally). *Snippets* 33:1–2.

Arcos López, Nicolás. 2009. Los clasificadores numerales y las clases nominales en ch'ol. Master's thesis, CIESAS, México.

Arkoh, Ruby, and Lisa Matthewson. 2013. A familiar definite article in akan. *Lingua* 123:1–30.

Arregi, Karlos, and Emily A. Hanink. 2018. Switch reference in Washo as multiple subject agreement. In *NELS 48: Proceedings of the Forty-Eighth Annual Meeting of the North East Linguistic Society*, eds. Sherry Hicklebridge and Max Nelson, volume 1, 39–48. GLSA.

Arsenijević, Boban. 2018. Atypical demonstratives in an articleless language. In *Atypical demonstratives*, eds. Marco Coniglio, Andrew Murphy, Eva Schlachter, and Tonjes Veenstra, 161–198. De Gruyter.

Assmann, Anke, Doreen Georgi, Fabian Heck, Gereon Müller, and Philipp Weisser. 2015. Ergatives move too early: On an instance of opacity in syntax. *Syntax* 18:343–387.

Ayres, Glenn. 1980. A note on the Mayan reflexive. *Journal of Mayan Linguistics* 1–2:53–59.

Ayres, Glenn. 1983. The antipassive "voice" in Ixil. *International Journal of American Linguistics* 49:20–45.

Bale, Alan, and Jessica Coon. 2014. Classifiers are for numerals, not for nouns: Consequences for the mass/count distinction. *Linguistic Inquiry* 45:695–707.

Bale, Alan, Jessica Coon, and Nicolás Arcos-López. 2019. Classifiers, partitions, and measurements: Exploring the syntax and semantics of sortal classifiers. *Glossa: A journal of general linguistics* 4.

Barker, Chris. 2012. Quantificational binding does not require c-command. *Linguistic Inquiry* 43:614–633.

Barwise, Jon, and Robin Cooper. 1981. Generalized quantifiers and natural language. In *Philosophy, language, and artificial intelligence*, eds. J. Kulas, J. H. Fetzer, and T. L. Rankin, volume 2, 241–301. Springer.

Barwise, Jon, and John Perry. 1983. *Situations and Attitudes*. MIT Press.

Beaver, David, and Kai von Fintel. 2013. Variable costs. Slides from CIL 19.

Beaver, David, and Kai von Fintel. 2019. Variable costs. Slides from ESSLLI 2019.

Becker, Laura. 2021. *Articles in the world's languages*. Berlin, Germany: De Gruyter.

Bennett, Ryan, Jessica Coon, and Robert Henderson. 2016. Introduction to Mayan linguistics. *Language and Linguistics Compass* 10:453–468.

Berlin, Brent. 1968. *Tzeltal Numeral Classifiers: A Study in Ethnographic Semantics*. Mouton.

Bervoets, Melanie. 2014. Indefinite determiners in Q'anjob'al. In *Proceedings of FAMLi 2: Formal approaches to Mayan Linguistics*, eds. Lauren Eby Clemens, Robert Henderson, and Pedro Mateo Pedro, 35–48. MIT Working Papers in Linguistics 74.

Bi, Ruyue Agnes, and Peter Jenks. 2019. Pronouns, null arguments, and ellipsis in Mandarin Chinese. In *Proceedings of Sinn und Bedeutung 23*, eds. María Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally, and Cristina Real-Puigdollers, volume 1, 127–142.

Bianchi, Valentina. 2009. A note on backward anaphora. *Rivista di Grammatica Generativa* 34:3–34.

Bielig, Louisa. 2015. Resumptive classifiers in Chuj high topic constructions. BA Thesis, McGill University.

Bittner, Maria, and Ken Hale. 1996. The Structural Determination of Case and Agreement. *Linguistic Inquiry* 1–68.

Blass, Regina. 1990. *Relevance relations in discourse: A study with special reference to Sissala*. Cambridge University Press.

Bobaljik, Jonathan David. 2002. A-chains at the PF interface: Copies and 'covert' movement. *Natural Language and Linguistic Theory* 20:197–267.

Bochnak, Ryan, and Lisa Matthewson, eds. 2015. *Methodologies in Semantic Fieldwork*. Oxford University Press.

Bochnak, Ryan, and Lisa Matthewson. 2020. Techniques in complex semantic fieldwork. *Annual Review of Linguistics* 6:261–283.

Boeckx, Cedric. 2001. Scope reconstruction and A-movement. *Natural Language and Linguistic Theory* 19:503–548.

Boeckx, Cedric, Norbert Hornstein, and Jairo Nunes. 2007. Overt copies in reflexive and control structures: A movement analysis. In *University of Maryland Working Papers in Linguistics*, eds. A. Conroy, C. Jing, C. Nakao, and E. Takahashi, volume 15, 1–46. UMWPiL.

Bohnemeyer, Jürgen. 2015. A practical epistemology for semantic elicitation in the field and elsewhere. In *Methodologies in Semantic Fieldwork*, eds. Ryan Bochnak and Lisa Matthewson, 13–46. Oxford University Press.

Bošković, Željko. 2001. *On the nature of the syntax–phonology interface: Cliticization and related phenomena*, volume 60 of *North-Holland Linguistic Series: Linguistic Variations*. Amsterdam: Elsevier.

Bošković, Željko, and Jairo Nunes. 2007. The Copy Theory of Movement: A view from PF. In *The Copy Theory of Movement*, eds. Jairo Nunes and Norbert Corver, volume 107 of *Linguistics Today*, 13–74. Oxford: Blackwell Publishers.

Bowern, Claire. 2008. *Linguistic Fieldwork: A Practical Guide*. New York, NY: Palgrave MacMillan.

Brasoveanu, Adrian, and Donka F. Farkas. 2011. How indefinites choose their scope. *Linguistics and Philosophy* 34:1–55.

Breheny, Richard. 2003. Exceptional-scope indefinites and domain restriction. In *Proceedings of Sinn und Bedeutung 7*, ed. Matthias Weisgerber, volume 7, 38–52.

Bricker, Victoria R. 1977. Pronominal inflection in the Mayan languages. In *Occasional Paper 1*. New Orleans: Middle American Research Institute at Tulane University and the Zemurray Foundation.

Brodkin, Dan, and Justin Royer. to appear. Ergative anaphors and high absolutive syntax. In *Proceedings of WCCFL 39*, ed. Gabriela de la Cruz Sánchez.

Bruening, Benjamin. 2009. Algonquian languages have A-movement and A-agreement. *Linguistic Inquiry* 40:427–445.

Bruening, Benjamin. 2014. Defects of defective intervention. *Linguistic Inquiry* 45:707–719.

Bruening, Benjamin. 2020. The head of the nominal is N, not D: N-to-D movement, hybrid agreement, and conventionalized expressions. *Glossa: A journal of general linguistics* 5:1–19.

Bruening, Benjamin. 2021. Generalizing the presuppositional approach to the Binding Conditions. *Syntax* 24:417–461.

Brugè, Laura. 1996. Demonstrative movement in Spanish: A comparative approach. *University of Venice Working Papers in Linguistics* 6:1–53.

Buenrostro, Cristina. 1992. Morfología verbal del chuj. Tesis de licenciatura en lingüística, INAH, Mexico.

Buenrostro, Cristina. 1994. El cambio de voz en chuj. In *Memorias del Segundo Encuentro de Lingüística en el Noroeste*, ed. Zarina Estrada-Fernández, 27–40. Universidad de Sonora.

Buenrostro, Cristina. 1995a. La negación en chuj de San Mateo Ixtatán. In *Vitalidad e influencia de las lenguas indígenas en Latinoamérica: Memorias del segundo coloquio M Swadesh*,

eds. Ramón Arzápalo Martín and Yolanda Lastra, 382–391. México, DF: Universidad Autónoma de México.

Buenrostro, Cristina. 1995b. Yuxtaposición de frases nominales en chuj. *Anales de Antropología* 32:269–281.

Buenrostro, Cristina. 1996. Algunos aspectos semánticos de la posesión en chuj. In *Análisis semánticos*, ed. Josefina García Fajardo, 37–51. El Colegio de México.

Buenrostro, Cristina. 1998a. Fenómenos de interferencia en el chuj. In *Estudios de lengua y cultura amerinida II-Lenguas, literaturas y medios*, eds. Julio Calvo-Pérez and Daniel Jorques Jiménez, 145–158. Valencia: Universitat de València.

Buenrostro, Cristina. 1998b. Orden de palabras en chuj. In *Memorias de IV Encuentro de Lingüística en el Noroeste*, eds. AndrAcosta Félix, Zarina Estrada-Fernández, and Max Figueroa Esteva, 217–228. Hermosillo: Universidad de Sonora.

Buenrostro, Cristina. 2000. La voz pasiva en chuj. *Anales de Antropología* 34:145–158.

Buenrostro, Cristina. 2001. Los refugiados guatemaltecos en la frontera sur. *Contacto interlingüístico e intercultural en el mundo hispano* 1:83–89.

Buenrostro, Cristina. 2002. La voz antipasiva y el enfoque de agente en el chuj de San Mateo Ixtatán. *Anales de Antropología* 36:229–247.

Buenrostro, Cristina. 2004. El sufijo -an en el Chuj de San Mateo Ixtatán. *Anales de Antropología* 38:255–267.

Buenrostro, Cristina. 2005. La voz en chuj y tojolab'al. *Anales de Antropología* 39:219–230.

Buenrostro, Cristina. 2007. Oraciones de complemento en Chuj de San Mateo Ixtatán. *Anales de Antropología* 41:239–266.

Buenrostro, Cristina. 2009. *Chuj de San Mateo Ixtatán*. Mexico City: El Colegio de México.

Buenrostro, Cristina. 2010a. Estructura de la palabra verbal en chuj de San Mateo Ixtatán. *Lingüística Mexicana* 5:193–216.

Buenrostro, Cristina. 2010b. Some typological differences between Chuj and Tojolab'al. In *A new look at language contact in Amerindian languages*, eds. Claudine Chamoreau, Zarina Estrada-Fernández, and Yolanda Lastra, 37–48. Munich: Lincom.

Buenrostro, Cristina. 2011. Sintaxis y morfología de las marcas de persona de número en chuj. In *Estudios morfológicos, sintácticos y semánticos de Lingüística Aplicada*, eds. María del Carmen Morúa and Gerardo López Cruz, 125–141. Hermosillo: Universidad de Sonora.

Buenrostro, Cristina. 2013a. El exódo: Una narración chuj. *Tlalocan* XIX:17–58.

Buenrostro, Cristina. 2013b. La voz en Chuj de San Mateo Ixtatán. Ph.D. dissertation, El Colegio de México, Mexico City.

Buenrostro, Cristina. 2014. La voz pasiva en chuj de San Mateo Ixtatán. *Revista UniverSOS* 11:167–190.

Buenrostro, Cristina. 2015a. El morfema *-ok* y la modalidad irrealis en el chuj de San Mateo Ixtatán. In *Estudios de lenguas amerindias 3: Escenarios de diversidad lingüística*, eds. Zarina Estrada-Fernández, Ana Fernández Garay, and Albert Alvarez Gonzalez, 55–75. Universidad de Sonora.

Buenrostro, Cristina. 2015b. La evolución de los verbos de movimiento en chuj. In *Lenguas, lenguaje, y lingüística: Contribuciones desde la lingüística general*, eds. Adriana Gorde-juela Senosiáin, Dámaso Izquierdo Alegría, Felipe Jiménez Berrio, Albert de Lucas Vicente, and Manuel Cadado Velarde, 103–116. Pamplona: Universidad de Navarra.

Buenrostro, Cristina. 2017. Los clasificadores nominales y la definitud en el Chuj de San Mateo Ixtatán. In *Investigaciones actuales en Lingüística, Vol. III: Sintaxis*, eds. Silvia

Gumiel-Molina, Manuel Leonetti, and Isabel Pérez-Jiménez, volume 3. España: Obras Colectivas Humanidades: Universidad de Alcalá.

Buenrostro, Cristina. 2018. Uso social del chuj. In *Estao del desarrollo económico y social de los pueblos indígenas de Chiapas*, eds. José del Val, Carolina Sánchez, and Carlos Zolla, 154–156. Universidad Nacional Autónoma de México, Secreteria para el desarrollo sustentable de los Pueblos Indígenas.

Buenrostro, Cristina. 2022. Artículo indefinido en Chuj. Paper presented at Fom and Analysis in Mayan Linguistics.

Buenrostro, Cristina. to appear. *Estudios de lengua y cultura chuj*. México, DF: Universidad Nacional Autónoma de México.

Buenrostro, Cristina, José Carmen Díaz, and Roberto Zavala. 1989. Sistema de clasificación nominal del Chuj. In *Memorias del Segundo Coloquio Internacional de Mayistas*, volume II. Mexico City: UNAM.

Buenrostro, Cristina, and Justin Royer. 2022. Los demostrativos adnominalea en el Chuj de San Mateo Ixtatán. Ms. UNAM and McGill.

Burge, Tyler. 1973. Reference and proper names. *The journal of philosophy* 70:425–439.

Büring, Daniel. 2003. On d-trees, beans, and b-accents. *Linguistics and Philosophy* 511–546.

Büring, Daniel. 2004. Crossover situations. *Natural Language Semantics* 12:23–62.

Büring, Daniel. 2005. *Binding Theory*. Cambridge: Cambridge University Press.

Burukina, Irina. 2019. Reflexive functional head, verbal and nominal predicates. In *Proceedings of the 36th West Coast Conference of Formal Linguistics*, eds. Richard Stockwell, Maura O'Leary, Zhonhshi Xu, and Z. L. Zhou, 91–98. Cascadilla Proceedings Project.

Campana, Mark. 1992. A movement theory of ergativity. Ph.D. dissertation, McGill University, Montreal.

Campbell, Lyle. 1977. *Quichean Linguistic Prehistory*. Berkeley: University of California Press.

Can Pixabaj, Telma. 2015. Complement and purpose clauses in K'iche'. Ph.D. dissertation, University of Texas Austin, Austin, TX.

Can Pixabaj, Telma. 2017. K'iche'. In *The Mayan Languages*, eds. Judith Aissen, Nora C. England, and Roberto Zavala Maldonado, 461–499. Routledge.

Cardinaletti, Anna, and Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In *Clitics in the Languages of Europe*, ed. Henk van Riemsdijk, 145–234. Berlin: Walter de Gruyter.

Carlson, Greg, and Rachel Sussman. 2005. Seemingly indefinite definites. In *Linguistic evidence: Empirical, theoretical, and computational perspectives*, eds. S. Kesper and M. Reis, 71–85. Mouton De Gruyter.

Carlson, Greg, Rachel Sussman, Natalie Klein, and Michael Tanenhaus. 2006. Weak definite noun phrases. In *Proceedings of NELS 36*, eds. C. Davis, Amy Rose Deal, and Y. Zabbal, 179–196.

Carolan, Elizabeth. 2015. An exploration of tense in Chuj. In *Revue de l'édition 2014 de VocUM*, eds. Marc Pomerleau and Eve-Marie Gendron-Pontbriand. Montreal, QC: Université de Montréal.

Charlow, Simon. 2014. On the semantics of exceptional scope. Ph.D. dissertation, New York University, New York, NY.

Charlow, Simon. 2020. The scope of alternatives: Indefiniteness and islands. *Lingusitics and Philosophy* 43:427–472.

Charnavel, Isabelle, and Dominique Sportiche. 2016. Anaphor binding: What French inanimate anaphors show. *Linguistic Inquiry* 47:35–87.

Charnavel, Isabelle, and Dominique Sportiche. 2021. Taking a strong position on strictly read reflexives. Paper presented at the 39th West Coast Conference in Formal Linguistics (WCCFL).

Chastain, Charles. 1975. *Reference and Context*. Minneapolis: University of Minnesota Press.

Chemla, Emmanuel. 2008. An epistemic step for anti-presupposition. *Journal of Semantics* 25:141–173.

Chierchia, Gennaro. 1995. *Dynamics of meaning*. Chicago: University of Chicago Press.

Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6:339–405.

Chierchia, Gennaro. 2001. A puzzle about Indefinites. In *Semantic interfaces*, eds. Carlo Cecchetto, Gennaro Chierchia, and Maria Teresa Guasti, 51–89. CSLI Publications.

Cho, Jacee. 2016. The acquisition of different types of definite noun phrases in l2-english. *International Journal of Blingualism* 21:367–382.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The View from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, eds. Kenneth Hale and Samuel Jay Keyser, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:33–49.

Christodoulou, Christiana. 2008. Anaphoric relations with Greek pronouns revisited. In *Proceedings of the 2007 annual conference of the Canadian Linguistic Association*.

Christophersen, Paul. 1939. *The articles: A study of their theory and use in English*. Copenhagen: Einar Munksgaard.

Chung, Sandra. 1989. On the notion of "nul anaphor" in Chamorro. In *The Null Subject Parameter*, 143–184. Springer.

Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36:315–332.

Cisnero, Carlos. 2019. Definiteness in Cuevas Mixtec. In *Definiteness across languages*, eds. Ana Aguilar-Guevara, Julia Pozas Loyo, and Violeta Vázquez-Rojas Maldonado, 39–82. Language Science Press.

Clark, Hebert H. 1975. Bridging. In *Thinking: readings in cognitive science*, eds. P. N. Johnson-Laird and P. C. Wason, 411–420. Cambridge University Press.

Clem, Emily. 2017. Two types of binding: Evidence from tswefap pronominals. In *NELS 47: Proceedings of the Forty-Seventh Annual Meeting of the North East Linguistic Society*, eds. A. Lamont and K. Tetzloff, volume 1, 173–184.

Clem, Emily. 2019. Agreement, case, and switch-reference in Amahuaca. Ph.D. dissertation, University of California, Berkeley.

Clemens, Lauren Eby, and Jessica Coon. 2018. Deriving verb initial order in Mayan. *Language* 94:237–280.

Collins, Chris. 2018. Quantifier domain restriction as ellipsis. *Glossa* 1:1–7.

Collins, Chris, and Erich Groat. 2018. Copies and repetitions. Ms. NYU.

Comunidad Lingüística Chuj. 2003. *Spaxti'al slolonelal Chuj: Vocabulario Chuj*. Academia de Lenguas Mayas de Guaatemala.

Coon, Jessica. 2009. Interrogative possessors and the problem with pied-piping in Chol. *Linguistic Inquiry* 40:165–175.

Coon, Jessica. 2010a. Complementation in Chol (Mayan): A Theory of Split Ergativity. Doctoral dissertation, MIT, Cambridge, MA.

Coon, Jessica. 2010b. VOS as Predicate Fronting in Chol. *Lingua* 120:354–378.

Coon, Jessica. 2016a. Mayan morphosyntax. *Language and Linguistics Compass* 10:515–550.

Coon, Jessica. 2016b. Unergatives, antipassives, and the status of roots in Chuj. In *Proceedings of CILLA: The Conference on Indigenous Languages of Latin America-VII*. Austin, TX: Archive of the Indigenous Languages of Latin America.

Coon, Jessica. 2017a. Little-v agreement and templatic morphology in Ch'ol. *Syntax* 20:101–137.

Coon, Jessica. 2017b. Two types of ergative agreement: Implications for case. In *MIT Working Papers in Linguistics*, eds. Claire Halpert, Hadas Kotek, and Coppe van Urk, 361–370. Cambidge, MA: MITWPL.

Coon, Jessica. 2018. Distinguishing adjectives from relative clauses in Chuj (with help from Ch'ol). In *Heading in the right direction: Linguistic treats for Lisa Travias*, eds. Laura Kalin, Jozina van der Klok, and Ileana Paul, 90–99. Montreal, QC: McGill Working Papers in Linguistics.

Coon, Jessica. 2019. Building verbs in Chuj: Consequences for the nature of roots. *Journal of Linguistics* 55:35–81.

Coon, Jessica, Nico Baier, and Theodore Levin. 2021. Mayan Agent Focus and the ergative extraction constraint: Facts and fictions revisited. *Language* 97:269–332.

Coon, Jessica, and Elizabeth Carolan. 2017. Nominalization and the structure of progressives in Chuj Mayan. *Glossa* 2:1–35.

Coon, Jessica, and Robert Henderson. 2011. Two binding puzzles in Mayan. In *Representing Language: Essays in Honor of Judith Aissen*, eds. Rodrigo Gutiérrez Bravo, Line Mikkelsen, and Eric Potsdam, 51–67. University of California, Santa Cruz: Linguistic Research Center.

Coon, Jessica, Pedro Mateo Pedro, and Omer Preminger. 2014. The role of case in A-bar extraction asymmetries: Evidence from Mayan. *Linguistic Variation* 14:179–242.

Coon, Jessica, and Justin Royer. 2020. Nominalization and selection in two Mayan languages. In *Nominalization: 50 years on from Chomsky's Remarks*, eds. Artemis Alexiadou and Hagit Borer, 139–168. Oxford University Press.

Cooper, Robin. 1979. The interpretation of pronouns. In *Syntax and semantics 10*, eds. F. Henry and H. Schnelle. New York: Academic Press.

Cooper, Robin. 1983. *Quantification and syntactic theory*. Reidel: Dordrecht.

Cooper, Robin Hayes. 1975. Montague's semantic theory and transformational syntax. Ph.D. dissertation, University of Massachusetts at Amherst.

Coppock, Elizabeth, and David Beaver. 2015. Definiteness and determinacy. *Linguistics and Philosophy* 38:377–435.

Córdova Hernández, Lorena. 2014. Esfuerzos de revitalización de la lengua chuj en contextos fronterizos multilingües del estado de Chiapas. Ph.D. dissertation, Centro de Investigaciones y Estudios Superiores en Antropología Social, México, DF.

Craig, Colette. 1986a. Jacaltec noun classifiers: A study in language and culture. In *Noun classes and categorization*, ed. Colette Craig, 263–294. Amsterdam, Philadelphia: John Benjamins Publishing Company.

Craig, Colette. 1990. Clasificadores nominales: Una inovación q'anjob'al. In *Lecturas sobre la lingüística maya*, eds. Nora C. England and Stephen R. Elliot, 253–269. Guatemala: CIRMA.

Craig, Colette Grinevald. 1977. *The Structure of Jacaltec*. Austin, TX: University of Texas Press.

Craig, Colette Grinevald. 1986b. Jacaltec noun classifiers. *Lingua* 70:241–284.

Cresti, Diana. 1995. Indefinite topics. Ph.D. dissertation, Massachusetts Institute of Technology, Cambidge, MA.

Curiel Ramírez del Prado, Alejando. 2017. Tojolabal. In *The Mayan Languages*, eds. Judith Aissen, Nora C. England, and Roberto Zavala Maldonado, 570–609. Routledge.

Davis, Henry. 2006. The status of Condition C in St'át'imcets. In *MIT Working Papers in Endangered Languages: Studies in Salishan*, eds. S. Bischoff, L. Butler, P. Norquest, and D. Siddiqi, 49–92. MIT Working Papers in Linguistics.

Davis, Henry. 2009. Cross-linguistic variation in anaphoric dependencies: Evidence from the Pacific Northwest. *Natural Langauge and Linguistic Theory* 27:1–43.

Dawson, Virginia Ellen. 2020. Existential quantification in Tiwa: Disjunction and indefinites. Ph.D. dissertation, University of California, Berkeley, Berkeley, CA.

Day, Christopher. 1973. *The Jacaltec Langauge*, volume 12 of *Language Science Monographs*. Indiana University.

Dayal, Veneeta. 2019. Singleton indefinites and the Privacy Principle: Certain puzzles. In *The semantics of Plurals, Focus, Degrees, and Times: Essays in Honor of Roger Schwarzschild*, eds. Daniel Altschuler and Jessica Rett, 57–80. Springer.

Deal, Amy Rose. 2016. Syntactic ergativity: Analysis and identification. *Annual Review of Linguistics* 2:165–185.

Déchaine, Rose-Marie, and Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33:409–442.

Despić, Miloje. 2013. Binding and the structure of NP in Serbo-Croatian. *Linguistic Inquiry* 44:239–270.

Despić, Miloje. 2015. Phases, reflexives, and definiteness. *Syntax* 18:201–234.

Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam, Philadelphia: John Benjamins Publishing Company.

Domingo Pascual, Pascual Martín. 2007. *Stzolalil Stz'ib'chaj ti' Chuj / Gramática Normativa Chuj*. Guatemala: Academia de Lenguas Mayas de Guatemala.

Donnellan, Keith S. 1966. Reference and definite descriptions. *Philosphical Review* 75:281–304.

Doran, Ryan B., and Gregory Ward. 2019. A taxonomy of uses of demonstratives. In *The Oxford Hanbook of Reference*, eds. Jeanette Gundel and Barbara Abbott. Oxford, UK: Oxford University Press.

Doron, Edit. 1988. The semantics of predicate nominals. *Linguistics* 26:281–301.

Douglas, Jamie, Rodrigo Ranero, and Michelle Sheehan. 2017. Two types of syntactic ergativity in Mayan. In *Proceedings of GLOW in Asia XI: Volume 2*, 41–56. Cambidge, MA: MIT Working Papers in Linguistics.

Drummond, Alex, Dave Kush, and Norbert Hornstein. 2011. Minimalist construal: Two approaches to A and B. In *The Oxford Handbook of Linguistic Minimalism*, ed. Cedric Boeckx. Oxford University Press.

Duncan, Lachlan. 2007. Analytic noun incorporation in Chuj and K'ichee' Mayan. In *Proceedings of the LFG07 Conference*, eds. Miriam Butt and Tracy Holloway King, 163–183. Stanford: CSLI Publications.

Ebert, Karen. 1971. Zwei Formen des bestimmten Artikels. In *Probleme und Fortschritte der Tansformationsgrammatik*, ed. D. Wunderlich, 159–174. München: Hueber.

Elbourne, Paul. 2008. Demonstratives as individual concepts. *Linguistics and Philosophy* 31:409–466.

Elbourne, Paul D. 2001. E-type anaphora and NP-deletion. *Natural Language Semantics* 9:241–288.

Elbourne, Paul D. 2005. On the acquisition of Principle B. *Linguistic Inquiry* 36:333–366.

Elbourne, Paul D. 2013. *Definite descriptions*. Oxford University Press.

Elias, Paulina. 2019. The role of directionals and positional and locative construction in Chuj. In *Proceedings of the Workshop on the Structure and Constituency of Languages of the Americas 23*, eds. D. K. E. Reisinger and Roger Yu-Hsiang Lo, 40–52. Vancouver, BC: University of British Columbia Working Papers in Linguistics.

Enç, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry* 22:1–26.

Endriss, Cornelia. 2009. *Quantificational topics*. Dordrecht: Springer.

England, Nora. 1983. *A grammar of Mam, a Mayan Language*. Austin: University of Texas Press.

England, Nora. 1991. Changes in basic word order in Mayan languages. *International Journal of American Linguistics* 57:446–86.

England, Nora. 2001. *Introducción a la Gramática de los Idiomas Mayas*. Guatemala: Cholsamaj.

England, Nora C. 1988. *Introducción a la lingüística: Idiomas mayas*. Antigua, Guatemala: Proyecto Lingüístico Francisco Marroquín.

England, Nora C. 1990. El Mam: Semajanzas y diferencias regionales. In *Lecturas sobre la lingüística maya*, eds. Nora C. England and Stephen E. Elliott, 221–252. La Antigua, Guatemala: Centro de Investigaciones Regionales de Mesoamérica.

England, Nora C. 1994. *Autonomía de los Idiomas Mayas: Historia e Identidad*. Guatemala: Cholsamaj: Centro Educativo y Cultura Maya.

England, Nora C. 1997. Topicalización, enfoque y énfasis. In *Cultura de Guatemala, año XVIII*, volume 2, 273–288. Universidad Rafael Landívar.

England, Nora C. 2017. Mam. In *The Mayan Languages*, eds. Judith Aissen, Nora C. England, and Roberto Zavala Maldonado, 500–532. Routledge.

Erlewine, Michael Yoshitaka. 2016. Anti-locality and optimality in Kaqchikel Agent Focus. *Natural Language and Linguistic Theory* 34:429–479.

Etxeberria, Urtzi, and Anastasia Giannakidou. 2019. Definiteness, partitivity, and domain restriction: A fresh look at definite descriptions. In *Definiteness across languages*, eds. Ana Aguilar-Guevara, Julia Pozas Loyo, and Violeta Vázquez-Rojas Maldonado, 419–452. Berlin: Language Science Press.

Evans, Gareth. 1977. Pronouns, quantifiers, and relative clauses (i). *Canadian Journal of Philosophy* 8:467–536.

Fara, Delia Graff. 2015. Names are predicates. *Philosophical review* 124:59–117.

Farkas, Donka F. 1981. Quantifier scope and syntactic islands. In *Proceedings of CLS 17*, 59–66. Chicago.

Farkas, Donka F. 2002. Specificity distinctions. *Journal of Semantics* 19:213–243.

Farkas, Donka F., and Adrian Brasoveanu. 2021. Kinds of (non)-specificity. In *The Wiley Blackwell Companion to Semantics*, eds. Daniel Gutzmann, Lisa Matthewson, Cécile Meier, Hotze Rullmann, and Thomas Ede Zimmermann. John Wiley & Sons. Ms. University of California Santa Cruz.

Farkas, Donka F., and Henriëtte de Swart. 2003. *The Semantics of Incorporation: From Argument Structure to Discourse Transparency*. University of Chicago Press.

Felipe Diego, Mateo, and Juan Gaspar Juan. 1998. *Diccionario del idioma Chuj*. La Antigua, Guatemala: Proyecto Lingüístico Francisco Marroquín.

Felipe Gómez, Pedro Federica. 2022. Análisis del proceso de derivación de sustantivos en el idioma Chuj, variante de San Mateo Ixtatán. Ms. Universidad de San Carlos de Guatemala.

Fishman, Joshua A. 1991. *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Multilingual Matters.

Fishman, Joshua A. 2001. From theory to practice (and vice-versa). In *Can threatened languages be saved?*, ed. Joshua A. Fishman, 451–483. Multilingual Matters.

Fodor, Janet Dean, and Ivan A. Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5:355–398.

Fox, Danny. 1999. Reconstruction, binding theory and the interpretation of chains. *Linguistic Inquiry* 30:157–196.

Fox, Danny. 2002. Antecedent-Contained Deletion and the copy theory of movement. *Linguistic Inquiry* 33:63–96.

Franks, Steven. 1999. Optimality theory and clitics at PF. In *Proceedings of Formal Approaches to Slavic Linguistics*.

Freeze, Ray. 1992. Existentials and other locatives. *Language* 68:553–595.

Frege, Gottlob. 1892. On sense and reference. In *Translations from the Philosophical Writings of Gottlob Frege*, eds. P. Geach and M. Black, 56–78. Oxford: Blackwell.

García Pablo, Gaspar, and Pascual Martín Domingo Pascual. 2007. *Stzolalil Sloloni-Spaxtini heb' Chuj: Gramática Descriptiva Chuj*. Academia de Lenguas Mayas de Guatemala.

Gauker, Christopher. 1998. What is a context of utterance? *Philosophical Studies* 91:149–172.

Gillon, Carrie Samantha. 2006. The semantics of determiners: domain restriction in Skwxwú7mesh. Ph.D. dissertation, University of British Columbia, Vancouver, BC.

Giusti, Giuliana. 2015. *Nominal syntax at the interfaces: A comparative analysis of languages with articles*. Cambridge Scholars Publishing.

Givón, Talmy. 1969. Studies in Chibemba and Bantu grammar. Ph.D. dissertation, University of California at Los Angeles, Los Angeles, CA.

Givón, Talmy. 1981. On the development of the numeral 'one' as an indefinite marker. *Folia Linguistica Historica* 2:35–53.

Givón, Talmy. 1985. The pragmatics of referentiality. In *Meaning, form and use in context*, ed. D. Schiffrin. Washington, D.C.: Georgetown University Press.

Gómez Cruz, José. 2017. Estructuras morfosintácticas de tojol-ab'al en perspectiva comparativa: el caso de una lengua maya mixta. Ph.D. dissertation, CIESAS, Ciudad de México.

Greenberg, Joseph. 1978. How does a language acquire gender markers? In *Universals of Human Language*, ed. Joseph Greenberg. Stanford University Press.

Grewendorf, Günther. 2015. Problems of remnant movement. In *Remnant Movement*, ed. Günther Grewendorf, 3–31. Walter de Gruyter.

Grimshaw, Jane, and Sara Thomas Rosen. 1990. Knowledge and obedience: The developmental status of the binding theory. *Linguistic Inquiry* 21:187–222.

Grinevald, Colette. 2000. A morphosyntactic typology of classifiers. In *Systems of nominal classification*, ed. Gunter Senft, 50–88. Cambridge University Press.

Grinevald, Colette. 2016. The Jakaltek Popti' noun classifier system: Changes due to Spanish contact. In *Language contact and change in the Americas: Studies in honor of Marianne Mithun*, eds. Andrea L. Berez-Kroeker, Diane M. Hintz, and Carmen Jany, 273–296. Amsterdam, Philadelphia: John Benjamins Publishing Company.

Grinevald, Colette, and Marc Peake. 2012. Ergativity and voice in Mayan: A functional-typological approach. In *Ergativity, Valency and Voice*, eds. Gilles Authier and Katharina Haude, 15–49. Berlin: Mouton de Gruyter.

Grodzinsky, Yosef, and Tanya Reinhart. 1993. The innateness of binding and coreference. *Linguistic Inquiry* 24:69–101.

Grodzinsky, Yosef, Kenneth Wexler, Yu Chin Chien, Susan Markovitz, and Julie Solomon. 1993. The breakdown of binding relations. *Brain and Language* 45:396–422.

Groenendijk, Jeroen, and Martin Stokhof. 1991. Dynamic predicate logic. *Linguistics and Philosophy* 14:39–100.

Gundel, Jeanette, Nancy Hedberg, and Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69:247–307.

Hanink, Emily. 2021. DP structure and internally headed relatives in Washo. *Natural Language & Linguistic Theory* 39:505–554.

Hanink, Emily A. 2018. Structural sources of anaphora and sameness. Ph.D. dissertation, University of Chicago.

Hankamer, Jorge, and Ivan Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7:391–428.

Harizanov, Boris, and Vera Gribanova. 2019. Whither head movement? *Natural Langauge and Linguistic Theory* 37:461–522.

Haspelmath, Martin. 1999. Explaining article-possessor complementarity: Economic motivation in noun phrase syntax. *Language* 75:227–243.

Hawkins, John A. 1978. *Definiteness and indefiniteness: a study in reference and grammaticality prediction.*. London: Croom Helm.

Hedberg, Nancy, Emrah Görgülü, and Morgan Mameni. 2009. On definiteness and specificity in Turkish and Persian. In *Proceedings of the 2009 Annual Meeting of the Canadian Linguistic Association*.

Heim, Irene. 1982. The semantics of definite and indefinite noun phrases. Doctoral dissertation, University of Massachusetts, Amherst, MA.

Heim, Irene. 1990. E-type pronouns and donkey anaphora. *Linguistics and Philosophy* 13:137–177.

Heim, Irene. 1991. Artikel und Definitheit. In *Semantics: An International Handbook of Contemporary Research*. Berlin: de Gruyter.

Heim, Irene. 1998. Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach. In *The interpretive tract*, eds. Uli Sauerland and Orin Percus, 205–246. MIT Working Papers in Linguistics 25.

Heim, Irene. 2007. Forks on the road to Rule I. In *NELS 38: Proceedings of the 38th Annual Meeting of the North East Linguistics Society*, eds. M. Abdurrahman, A. Schardl, and M. Walkow.

Heim, Irene. 2008. Features on bound pronouns. In *Phi theory: Phi-features across modules and interfaces*, eds. Daniel Harbour, David Adger, and Susana Béjar, 35–56. Oxford University Press.

Heim, Irene. 2011. Definiteness and indefiniteness. In *Semantics*, eds. Klaus von Heusinger, Claudia Maienborn, and Paul Portner, volume 2, 996–1025. de Gruyter.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell Publishers.

Henderson, Robert. 2007. Observations on the syntax of adjunct extraction in Kaqchikel. In *Proceedings of Congreso de Idiomas Indígenas de Latinoamérica (CILLA) III*. AILLA, University of Texas at Austin.

Henderson, Robert. 2012. Morphological alternations at the intonational phrase edge: The case of K'ichee'. *Natural Language and Linguistic Theory* 30:741–787.

Henderson, Robert. 2014. Dependent indefinites and their post-suppositions. *Semantics and Pragmatics* 7:1–58.

Henderson, Robert, and Jessica Coon. 2018. Adverbs and variability in Kaqchikel Agent Focus: A reply to Erlewine (2016). *Natural Langauge and Linguistic Theory* 36:149–173.

Henderson, Robert, Paulina Elias, Justin Royer, and Jessica Coon. 2018. The composition of stativity in Chuj. In *Heading in the right direction: Linguistic treats for Lisa Travias*,

eds. Laura Kalin, Jozina van der Klok, and Ileana Paul, 100–109. Montreal, QC: McGill Working Papers in Linguistics.

Hendrick, Randall. 2005. Tongan determiners and semantic composition. *Language* 907–926.

von Heusinger, Klaus. 2011. Specificity. In *An International Handbook of Natural Language Meaning*, eds. Klaus von Heusinger, Claudia Maienborn, and Paul Portner, volume 2. Mouton De Gruyter.

Hicks, Glyn. 2009. *The derivation of anaphoric relations*. Linguistik Aktuell.

Himmelmann, Nikolaus P. 1997. *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Tübingen: Max Niemeyer Verlag.

Hoekstra, Eric. 1989. A parameter for anaphor binding: The case of Jacaltec. In *Configurationality: The Typology of Asymmetries*, eds. L. Marácz and P. Muysken. Foris.

de Hoop, Helen. 1997. A semantic reanalysis of the partitive constraint. *Lingua* 103:151–174.

Hopkins, Nicholas. 1967. The Chuj language. Ph.D. dissertation, University of Chicago, Chicago, IL.

Hopkins, Nicholas. 2012a. *A Dictionary of the Chuj Mayan Language*. Tallahassee, FL: Jaguar Tours.

Hopkins, Nicholas. 2012b. Noun classifiers of the Chuchumatán Mayan languages: A case of diffusion from Otomanguean. *International Journal of American Linguistics* 78:411–427.

Hopkins, Nicholas A. 1970. Numeral classifiers in tzeltal, jacaltec and chuj (mayan). In *Papers from the sixth regional meeting of the Chicago Linguistic Society*, 23–35.

Hopkins, Nicholas A. 1972. Compound place names in Chuj and other Mayan languages. In *Meaning in Mayan Languages: Ethnolinguistic studies*, ed. Munro S. Edmonson, 165–182. The Hague: Mouton.

Hopkins, Nicholas A. 1980a. Chuj animal names and their classification. *Journal of Mayan Linguistics* 2:13–39.

Hopkins, Nicholas A. 1980b. A San Mateo Chuj text. *International Journal of American Linguistics* 89–106.

Hopkins, Nicholas A. 2006. Review of the linguistics of Maya Writing. *Anthropological Linguistics* 48:404–412.

Hopkins, Nicholas A. 2018. Mayan languages collection of Nicholas Hopkins. Archive of the Indigenous Languages of Latin America, ailla.utexas.org. Accessed 2018.

Hopkins, Nicholas A. 2021. *Chuj (Mayan) Narratives: Folklore, History, and Ethnography from Northwestern Guatemala*. Boulder, CO: University Press of Colorado.

Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 69–96.

Hornstein, Norbert. 2001. *Move! A minimalist theory of construal*. Oxford: Blackwell Publishers.

Hornstein, Norbert. 2007. Pronouns in a minimalist setting. In *The Copy Theory of Movement*, 351–385. John Benjamins.

Hou, Liwen. 2011. Agent focus in Chuj reflexive constructions. BA Honours Thesis, McGill University, Montréal, QC.

Hou, Liwen. 2013. Agent focus in Chuj reflexive constructions. BA Honours Thesis, McGill University, Montréal, QC.

Hsu, Brian, and Saurov Syed. 2020. Cooccurrence of nonarticle determiners as evidence for split DPs and bundled projections. *Language* 96:114–134.

Huang, Cheng-Teh James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT, Cambridge, MA.

Imanishi, Yusuke. 2019. Parametrizing split ergativity in Mayan. *Natural Language & Linguistic Theory* 38:151–200.

Ingason, Anton. 2016. Realizing morphemes in the Icelandic noun phrase. Ph.D. dissertation, University of Pennsylvania, Philadelphia, PA.

Ionin, Tania. 2006. *This* is definitely specific: Specificity and definiteness in article systems. *Natural Language Semantics* 14:175–234.

Irani, Ava. 2019. On (in)definite expressions in American Sign Language. In *Definiteness across languages*, eds. Ana Aguilar-Guevara, Julia Pozas Loyo, and Violeta Vázquez-Rojas Maldonado, 113–152. Language Science Press.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambidge, MA: MIT Press.

Jackendoff, Ray. 1977. *X-bar syntax: A study of phrase structure*. Cambridge, MA: MIT Press.

Jenks, Peter. 2015a. Noun phrase interpretations questionnaire. Handout from talk given at the 4th internation conference on language documentation and conservation.

Jenks, Peter. 2015b. Two kinds of definites in numeral classifier languages. In *Proceedings of Semantics and Linguistic Theory 25*, ed. 103-124.

Jenks, Peter. 2018. Articulated definiteness without articles. *Linguistic Inquiry* 49:501–536.

Jenks, Peter, and Rassidatou Konate. to appear. Indexed definiteness. *Glossa* .

Jukes, Anthony. 2006. Makassarese (basa mangkasara'): A decription of an Austronesian language of South Sulawesi. Ph.D. dissertation, The University of Melbourne, Melbourne.

Kadmon, Nirit, and Fred Landman. 1993. Any. *Linguistics and philosophy* 16:353—-422.

Kamp, Hans. 1981. A theory of truth and semantic representation. In *Formal methods in the study of language*, eds. Jeroen Groenendijk, Theo Janssen, and Martin Stokhof, volume 1, 277–322. Mathematical Centre Tracts.

Kaplan, David. 1989. Afterthoughts. In *Themes from Kaplan*, eds. Joseph Almog, John Perry, and Howard Wettstein, 481–563. Oxford University Press.

Kaplan, Emily. 2021. A guatemalan town remakes itslef in indiana. *National Geographic* .

Karttunen, Lauri. 1976. Discourse referentes. In *Syntax and semantics*, ed. James D. McCawley, volume 7, 363–385. Academic Press.

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and philosophy* 1:3–44.

Katzir, Roni. 2011. Morphosemantic mismatches, structural economy, and licensing. *Linguistic Inquiry* 42:45–82.

Kaufman, Terrence. 1969. Some recent hypotheses on Mayan diversification. *Language Behavior Research Lab, Working Paper* 26.

Kaufman, Terrence. 1974. *Idiomas de Mesoamerica*. Guatemala: Editorial Jose De Pineda Ibarra.

Kaufman, Terrence. 1976. Archaeological and linguistic correlations in Mayaland and associated areas of Meso-America. *World Archaeology* 8:101–118.

Kayne, Richard. 2002. Pronouns and their antecedents. In *Derivation and Explanation*, eds. Samuel David Epstein and T. Daniel Seeley. Oxford: Blackwell Publishers.

Kazanina, Nina. 2005. The acquisition and processing of backwards anaphora. Ph.D. dissertation, University of Maryland.

Kazanina, Nina, and Colin Phillips. 2001. Coreference in child Russian: Distinguishin syntactic and discourse constraints. In *Proceedings of the 25th annual Boston University Conference on Language Development*, 413–424.

Keshet, Ezra. 2008. Good intensions: Paving two roads to a theory of the De re/De dicto distinction. Ph.D. dissertation, Massachusetts Institute of Technology.

King, Jeffrey C. 1988. Are indefinite descriptions ambiguous? *Philosophical studies: An international Journal of Philosophy in the Analytical Tradition* 53:417–440.

King, Jeffrey C. 2001. *Complex demonstratives: A quantificational account*. Cambidge, MA: MIT Press.

King, Jeffrey C. 2018. Strong contextual felicity and felicitous underspecification. *Philosophy and Phenemenological Research* 97:631–657.

Kornfilt, Jaklin. 1997. *Turkish*. London: Routledge.

Kotek, Hadas, and Michael Yoshitaka Erlewine. 2016. Unifying definite and indefinite free relatives: Evidence from mayan. In *Proceedings of NELS 46*, 241–254.

Kotek, Hadas, and Mitcho Erlewine. 2019. *Wh*-indeterminates in Chuj (Mayan). *Canadian Journal of Linguistics* 64:62–101.

Kratzer, Angelika. 1986. Conditionals. *Chicago Linguistics Society 22* 2:1–15.

Kratzer, Angelika. 1989. An investigation of the lumps of thought. *Linguistics and Philosophy* 12:607–653.

Kratzer, Angelika. 1998. More Structural Analogies Between Pronouns and Tense. In *Proceedings of the 8th Semantics And Linguistics Theory conference (SALT 8)*, eds. Devon Strolovitch and Aaron Lawson, 92–110. Ithaca, NY: CLC Publications.

Kratzer, Angelika. 2003. A note on choice functions in context. Ms. University of Massachusetts, Amherst.

Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40:187–237.

Kratzer, Angelika. 2019. Situations in Natural Language Semantics. In *Stanford Encyclopedia of Philosophy*. Stanford, CA: CSLI Publications.

Kratzer, Angelika. 2021. Situations in Natural Language Semantics. In *Stanford Encyclopedia of Philosophy*. Stanford, CA: CSLI Publications.

Kratzer, Angelika, and Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. In *Proceedings of the 3rd Tokyo Conference on Psycholinguistics*, ed. Y. Otsu, 1–25.

Krifka, Manfred. 2011. Varieties of semantic evidence. In *Semantics: An International Handbook of Contemporary Research*, eds. Claudia Maienborn, Klaus von Heusinger, and Paul Portner, 242–267. Mouton De Gruyter.

Labelle, Marie. 2008. The French reflexive and reciprocal *se*. *Natural Language & Linguistic Theory* 26:833–876.

Langacker, Ronald. 1969. On pronominalization and the chain of command. In *Modern studies in English*, eds. David A. Reibel and Sanford A. Schane, 160–186. Englewood Cliffs.

Larsen, Tomas W., and William M. Norman. 1979. Correlates of ergativity in Mayan grammar. In *Ergativity: Towards a theory of grammatical relations*, ed. Frans Plank, 347–370. London/New York: Academic Press.

Lasnik, Howard. 1999. Chains of arguments. *Current Studies in Linguistics Series* 32:189–216.

Law, Danny. 2011. Linguistic inheritance, social difference, and the last two thousand years of contact among lowland mayan languages. Ph.D. dissertation, University of Texas Austin.

Law, Danny. 2014. *Language contact, inherited similarity and social difference: The story of linguistic interaction in the Maya lowlands*. Amsterdam: John Benjamins.

Lebeaux, David. 1984. Locality and anaphoric binding. *The Linguistic Review* 4:343–363.

Lebeaux, David. 2009. *Where does the binding theory apply?*. MIT Press.

Lees, Robert B., and Edward S. Klima. 1963. Rules for English pronominalization. *Languages* 1:17–28.

Legate, Julie Anne. 2014. Split ergativity based on nominal type. *Lingua* 148:183–212.

Lemon, Tyler. 2019. Noun classifiers in Todos Santos Mam. Ms. UC Berkeley.

Lesure, Cora. 2016. La morfofonología del chuj y la representación ortográfica. Paper presented at *Form and Analysis in Mayan Linguistics (FAMLi IV)*.

Leu, Thomas. 2008. The internal syntax of determiners. Ph.D. dissertation, New York University, New York, NY.

Leu, Thomas. 2015. *The Architecture of Determiners*. Oxford University Press.

Lewis, David. 1975. Adverbs of quantification. In *Formal semantics in natural language*, ed. Edward Keenan, 3–15. Cambridge University Press.

Limón, Fernando. 2009. *Historia chuj a contrapelo: Huellas de un pueblo con memoria*. ECOSUR.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, Usem and Iinterpretation of Language*, eds. Rainer Bäuerle, Christop Schwarze, and Arnim von Stechow, 303–323. Berlin: De Gruyter.

Little, Carol-Rose. 2020a. Left branch extraction, object shift, and freezing effects in tumbalá ch'ol. *Glossa: A journal of general linguistics* 5.

Little, Carol-Rose. 2020b. Mutual dependencies of nominal and clausal syntax in Ch'ol. Ph.D. dissertation, Cornell University.

Little, Carol Rose, Mary Moroney, and Justin Royer. 2022a. Classifying classifiers: Evidence from Ch'ol, Shan and Chuj for two types of numeral classifiers. In *Proceedings of SULA 11*, ed. Yixiao Song.

Little, Carol Rose, Mary Moroney, and Justin Royer. 2022b. Two types of numeral classifiers across languages. Ms. University of Oklahoma, University of Rochester, McGill University.

López, María Teresa Rodríguez, and Álvaro Cabelleros Herrera. 2019. Los chuj de méxico y Guatemala: Ejes para una articulación transfronteriza. Región Transfronteriza México-Guatemala.

López García, Diego Neper Josué. 2022. Pérdida de sonidos en la variante Chuj de San Sebastián Coatán, Huehuetenango. Ms. Universidad de San Carlos de Guatemala.

Lust, Barbara. 1986. Introduction. In *Studies in the acquisition of anaphora*, ed. Barbara Lust, volume 1. Dordrecht: D. Reider Publishing Company.

Lust, Barbara, Yu-Chin Chien, Chi-Pang Chiang, and Julie Eisele. 1996. Chinese pronominals in universal grammar: A study of linear precedence and command in Chinese and English children's first language acquisition. *Journal of East Asian Linguistics* 5:1–47.

Lyons, Christopher. 1999. *Definiteness*. Cambridge, UK: Cambridge University Press.

Maclaran, Rose. 1982. The semantics and pragmatics of english demonstratives. Ph.D. dissertation, Cornell University, Ithaca, NY.

Martí, Luisa. 2003. Contextual variables as pronouns. *Semantics and liguistic theory* 13:240–257.

Mateo Pedro, Pedro, and Jessica Coon. 2018. Chuj oral tradition: Collection of Pedro Mateo Pedro and Jessica Coon. Archive of the Indigenous Languages of Latin America, ailla.utexas.org. Accessed 2018-2022.

Mateo Toledo, Eladio. 2017. Q'anjob'al. In *The Mayan Languages*, eds. Judith Aissen, Nora C. England, and Roberto Zavala Maldonado, 533–569. New York: Routledge.

Matthewson, Lisa. 1998. *Determiner systems and quantificational strategies: Evidence from Salish*. The Hague: Holland Academic Graphics.

Matthewson, Lisa. 1999. On the interpretation of wide-scope indefinites. *Natural Language Semantics* 7:79–134.

Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70:369–415.

Matthewson, Lisa. 2006. Temporal semantics in a supposedly tenseless language. *Linguistics and Philosophy* 29:673–713.

Matthewson, Lisa. 2008. Pronouns, presuppositions, and semantic variation. In *Proceedings of Semantics and Linguistic Theory XVIII*, eds. T. Friedman and S. Ito.

Matthewson, Lisa. 2013. Strategies of quantification in St'át'imcets and the Rest of the World. In *Strategies of quantification*, eds. Gil Kook-Hee, Stephen Harlow, and George Tsoulas, 15–38. Oxford, UK: Oxford University Press.

Maxwell, Judith. 1978. Chuj clause collapsing. In *Papers in Mayan linguistics*, ed. Nora C. England, 127–135. University of Missouri: Miscellaneous Publications in Anthropology.

Maxwell, Judith. 1981. How to talk to people who talk *chekel* 'different': The Chuj (Mayan) solution. Ph.D. dissertation, University of Chicago, Chicago, IL.

Maxwell, Judith. 1983. A simplified code: Chuj (Mayan) as written by speakers. *Journal of Mayan Linguistics* 4:3–10.

Maxwell, Judith. 1987. Some aspects of Chuj discourse. *Anthropological Linguistics* 29:489–506.

Maxwell, Judith. 1990. El discurso en chuj. In *Lecturas sobre la lingüística maya*, eds. Nora C. England and Stephen Elliott, 445–459. La Antigua, Guatemala: CIRMA: Centro de investigaciones regionales de mesoamérica.

Maxwell, Judith. 1993. Chuj. In *Encyclopedia of world cultures: Middle America and the Caribbean*, eds. David Levinson, James W. Dow, and Robert Van Kemper, volume VIII, 70–74. New York, NY: G. K. Hall Company.

Maxwell, Judith. 1995. *Textos chujes de San Mateo Ixtatán*. Palos Verdes: Fundación Yax Te'.

Maxwell, Judith M. 1976. Chuj intransitives: Or when can an intransitive verb take an object? *Journal of Mayan Linguistics* 1:128–140.

Mayer, Mercer. 1969. *Frog, where are you?*. New York, NY: Dial Books for Young Readers.

McNally, Louise. 2016. Existential sentences crosslinguistically: Variations in form and meaning. *Annual Review of Linguistics* 2:211–231.

Medina, Alfonso, and Cristina Buenrostro. 2003. Características cuantitativas de la flexión verbal del chuj. *Estudios de Lingüística Aplicada* 38:15–31.

Mendes, Gesoel, and Rodrigo Ranero. 2021. Chain reduction via Substitution: Evidence from Mayan. *Glossa* 6.

Merchant, Jason. 2019. Ellipsis: A survey of analytical approaches. In *The Oxford Handbook of Ellipsis*, eds. Jeroen van Craenenbroeck and Tanja Temmerman. Oxford University Press.

Mikkelsen, Line. 2005. *Copular sentences: Specification, predication, and equation*. John Benjamins Publishing Company.

Milsark, Gary. 1974. Existential Sentences in English. Doctoral dissertation, MIT, Cambridge, MA.

Mohanan, K. P. 1983. Functional and anaphoric control. *Linguistic Inquiry* 14:641–674.

Mohanan, Karuvannur Puthanveettil. 1981. Grammatical relations and anaphora in Malayalam. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA.

Mondloch, James. 1981. Voice in Quiche-Maya. Ph.D. dissertation, SUNY Albany.

Montague, Richard. 1973. The Proper Treatment of Quantification in Ordinary English. In *Approaches to natural language: Proceedings of the 1970 Stanford workshop on grammar and semantics*, eds. Kaarlo Hintikka, Julius Moravcsik, and Patrick Suppes, 221–242. Dordrecht: Reidel.

Montes de Oca Vega, Mercedes. 2004. Los disfrasismos: ¿núcleos conceptuales mesoamericanos? In *La metáfora en Mesoamérica*, ed. Mercedes Montes de Oca Vega, 225–251. México, DF: Universidad Nacional Autónoma de México.

Moroney, Mary. 2021. Updating the typology of definiteness: Evidence from bare nouns in Shan. *Glossa: a journal of general linguistics* 6.

Mosel, Ulrike, and Even Hovdhaugen. 1992. *Samoan Reference Grammar*. Oslo: Scandinivian University Press.

Moseley, Christopher, ed. 2010. *Atlas of the World's Languages in Danger*. Paris: UNESCO Publishing, 3rd edition edition.

Müller, Gereon. 1996. A constraint on remnant movement. *Natural Language and Linguistic Theory* 14:355–407.

Myler, Neil. 2014. Building and interpreting possession sentences. Ph.D. dissertation, New York University, New York, NY.

Neale, Stephen. 1990. *Descriptions*. Cambidge, MA: MIT Press.

Newman, Elise. 2020. On the movement/(anti)agreement correlation in Romance and Mayan. Ms. MIT.

Ordóñez, Francisco. 1995. The antipassive in Jacaltec: A last resort strategy. *CatWPL* 4:329–343.

Partee, Barbara. 1986. Noun phrase interpretation and type-shifting principles. In *Studies in discourse representation theory and the theory of generalized quantifiers*, eds. Jeroen Groenendijk, Dick De Jongh, and Martin Stokhof. Foris.

Partee, Barbara Hall. 1973. Some structural analogies between tenses and pronouns in English. *The Journal of Pilosophy* 70:60–609.

Pascual, Adán F. 2007. Transitividad y dependencia sintáctica y discursiva en Q'anjob'al. Master's thesis, CIESAS, México.

Patel-Grosz, Pritty, and Patrick G. Grosz. 2017. Revisiting pronominal typology. *Linguistic Inquiry* 48:259–297.

Payne, Thomas E. 1997. *Describing Morphosyntax*. Cambridge, UK: Cambridge University Press.

Percus, Orin. 2000. Constraints on some other variables in syntax. *Natural Language Semantics* 8:173–229.

Percus, Orin. 2006. Antipresuppositions. In *Theoretical and empirical studies of reference and anaphora: Toward the establishment of generative grammar as an empirical science*, ed. Ayumi Uyema, 52–73. Japan Society for the Promotion of Science.

Pesetsky, David. 1995. *Zero Syntax: Experiencers and Cascades*. Cambridge, MA: MIT Press.

Pesetsky, David. 1998. Some optimality principles of sentence pronunciation. In *Is the Best Good Enough? Optimality and Competition in Syntax*, eds. Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky, 337–383. Cambridge, MA: MIT Press and MITWPL.

Piedrasanta, Ruth. 2009. *Los Chuj, Unidad y rupturas en su espacio*. Guatemala City, Guatemala: Amrar Editores.

Polian, Gilles. 2007. El sufijo *-uk* y la diversificación de la modalidad *irrealis* en tseltal. In *Conference on Iindigenous Languages of Latin America III*.

Polinsky, Maria. 2017. Syntactic ergativity. In *Blackwell Companion to Syntax*, eds. Martin Everaert and Henk van Riemsdijk. Blackwell Publishers, 2nd edition edition.

Pollard, Carl, and Ivan A. Sag. 1992. Anaphors in English and the scope of binding theory. *Linguistic Inquiry* 23:261–303.

Poole, Ethan. 2017. Movement and the Semantic Type of Traces. Ph.D. dissertation, University of Massachusetts at Amherst.

Portner, Paul. 2002. Topicality and (non-)specificity in Mandarin. *Journal of Semantics* 19:275–287.

Portner, Paul, and Katsuhiko Yabushita. 2001. Specific indefinites and the information structure theory of topics. *Journal of Semantics* 18:271–297.

Postal, Paul. 1966. On so-called pronouns in English. *Monograph series on language and linguistics* 19:177–206.

Preminger, Omer. 2014. *Agreement and Its Failures*. Cambridge, MA: MIT Press.

Prince, Ellen. 1981. Towards a taxonomy of given-new information. In *Radical pragmatics*, ed. P. Cole, 223–256. New York: Academic Press.

Prince, Ellen. 1992. The zpg letter: subjects, definiteness, and information-status. In *Discourse descriptions: diverse linguistic analyses of a fund-raising text*, eds. W. C. Mann and S. A. Thompson, 295–325. Philadelphia: John Benjamins.

Quezada, Sergio Aguayo, and Laura O'Dogherty. 1986. Los refugiados guatemaltecos de Campeche y Quintana Roo. *Foro Internacional* 27:266–295.

Ramsay, V. 1985. Classifiers and referentiality in Jacaltec. In *Proceedings of the First Annual Meeting of the Pacific Linguistic Conference*, eds. Scott DeLancey and R. Tomlin, 289–312. University of Oregon.

Reinhart, Tanya. 1976. The syntactic domain of anaphora. Ph.D. dissertation, MIT, Cambidge, MA.

Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. London: Croom Helm.

Reinhart, Tanya. 1997. Strategies of anaphora resolution. In *OTS Working Papers in Linguistics*. Utrecht Institute of Linguistics.

Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657–720.

Reuland, Eric. 2001. Primitives of binding. *Linguistic Inquiry* 32:439–492.

Reuland, Eric. 2010. The universality of binding principles. In *Structure Preserved: Studies in syntax for Jan Koster*, eds. Jan-Wouter Zwart and Mark de Vries, 277–287. John Benjamins Publishing Company.

Reuland, Eric. 2011. *Anaphora and Language Design (Linguistic Inquiry Monographs)*. MIT Press.

Reuland, Eric. 2016. Universals and variation: Encoding anaphoric dependencies. In *Finiteness matters*, ed. Kristin Melum Eide, 93–120. John Benjamins.

Reuland, Eric. 2017. Grammar of binding in the languages of the world: Unity versus diversity. *Cognition* 168:370–379.

Reuland, Eric, and Sergey Avrutin. 2004. Interpretive dependencies: the case of backward anaphora. *Linguistic Analysis* 34.

Reuland, Eric, and Jan Koster. 1991. *Long distance anaphora*. Cambridge University Press.

Ritter, Elizabeth. 1991. Two functional categories in noun phrases: Evidence from Modern Hebrew. In *Perspectives on Phrase Structure: Heads and Licensing*, ed. Susan D. Rothstein, 37–62. San Diego, CA: Academic Press.

Ritter, Elizabeth. 1995. On the Syntactic Category of Pronouns and Agreement. *Natural Language and Linguistic Theory* 13:405–443.

Roberts, Craige. 1989. Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy* 2:683–721.

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. In *Working papers in Linguistics 49*, eds. Jae-Hak Yoon and Andreas Kathol, 91–136. Columbus, OH: Ohio State University.

Roberts, Craige. 2002. Demonstratives as definites. In *Information sharing: Reference and presupposition in languages generation and interpretation*, eds. K. van Deemter and R. Kibble, 89–106. CSLI Publications.

Roberts, Craige. 2003. Uniqueness and definite noun phrases. *Linguistics and Philosophy* 26:287–350.

Robertson, John. 1977. A proposed revision in Mayan subgrouping. *International Journal of American Linguistics* 43:105–120.

Robertson, John. 1980. *The structure of pronoun incorporation in the Mayan verbal complex*. New York: Garland.

Robertson, John. 1992. *The History of Tense/Aspect/Mood/Voice in the Mayan Verbal Complex*. Austin, TX: University of Texas Press.

Rodrigues, Cilene. 2010. Possessor raising through thematic positions. In *Movement Theory of Control*, eds. Norbert Hornstein and Maria Polinsky, 119–146. Linguistik Aktuell, John Benjamins.

Rooryck, Johan, and Guido vanden Wyngaerd. 2011. *Dissolving Binding Theory*. Oxford University Press.

Ross, John R. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT, Cambridge, MA.

Royer, Justin. 2017. Noun and numeral classifiers in Chuj (Mayan). In *Proceedings of the Berkeley Linguistics Society*, eds. Julia Nee, Margaret Cychosz, Dmetri Hayes, Taylor Lau, and Emily Remirez, volume 2, 29–38.

Royer, Justin. 2019. Domain restriction and noun classifiers in Chuj. In *NELS 49: Proceedings of the forty-ninth annual meeting of the North-East Linguistics Society*, eds. Maggie Baird and Jonathan Pestesky, volume 3, 87–97.

Royer, Justin. 2020. An apparent syntax-prosody mismatch as right extraposition: Evidence from mayan. In *NELS 50: Proceedings of the Fiftieth Annual Meeting of the North East Linguistic Society*, ed. Mariam Asatryan Aryana Whitma, Yixiao Song.

Royer, Justin. 2021a. Binding and coreference in Mayan and the pervasiveness of high-absolutive syntax. Ms. McGill.

Royer, Justin. 2021b. Headless relative clauses in Chuj. In *Headless relative clauses in languages of Mesoamerica*, eds. Ivano Caponigro, Harold Torrence, and Roberto Zavala, 327–361. Oxford, UK: Oxford University Press.

Royer, Justin. 2021c. Object raising, binding, and coreference in chuj. In *NELS 51: Proceedings of the Fifty-first Annual Meeting of the North East Linguistic Society*, eds. Alessa Farinella and Angelica Hill, 191–204.

Royer, Justin. 2022. Prosody as syntactic evidence: The view from mayan. *Natural Language & Linguistic Theory* 40:239–284.

Royer, Justin. to appeara. Decomposing definiteness in Chuj. *Canadian Journal of Linguistics* .

Royer, Justin. to appearb. Decomposing definiteness: The case of noun classifiers in chuj. *Canadian Journal of Linguistics* .

Royer, Justin, and Luis Alonso-Ovalle. 2019. Random choice modality in Chuj: The case of 'komon'. In *Proceedings of Form and Analysis in Mayan Linguistics 5*, eds. Ryan Bennett, Robert Henderson, Pedro Mateo Pedro, and Megan Harvey.

Royer, Justin, Pedro Mateo Pedro, Elizabeth Carolan, Jessica Coon, and Matal Torres. 2022. *A*tz'am k'ik' atz'am: The story of Xuwan and a grammatical sketch of Chuj. *Tlalocan* 27:215–286.

Royer, Justin, Pedro Mateo Pedro, Elizabeth Carolan, Jessica Coon, and Matal Torres. to appear. Atz'am k'ik' atz'am: The story of xuwan and a grammatical sketch of chuj. *Tlalocan* .

Russell, Bertrand. 1905. On denoting. *Mind* 14:479–493.

Safir, Ken. 2004. *The syntax of anaphora*. Oxford: Oxford University Press.

Safir, Ken. 2008. Coconstrual and narrow syntax. *Syntax* 11:330–355.

Safir, Ken. 2014. One true anaphor. *Linguistic Inquiry* 45:91–124.

Salzmann, Martin. 2020. The NP vs. DP debate. why previous arguments are inconclusive and what a good argument could look like. evidence from agreement with hybrid nouns. *Glossa: A journal of general linguistics* 5.

Sauerland, Uli. 1998. The meaning of chains. Doctoral dissertation, MIT, Cambridge, MA.

Sauerland, Uli. 2008. Implicated presuppositions. In *The discourse potential of underspecified structures*, ed. A. Steube, 581–600. Mouton de Gruyter.

Sauerland, Uli. 2013. Presuppositions and the alternative tier. In *Proceedings of Semantics and Linguistic Theory (SALT) 23*, ed. Todd Snider, 156–173. Linguistic Society of America.

Schlenker, Philippe. 2005a. Minimize restrictors! (notes on definite descriptions, Condition C and epithets). In *Proceedings of SuB 9*, 385–416.

Schlenker, Philippe. 2005b. The Lazy Frenchman's Approach to the Subjunctive (Speculations on Reference to Worlds and Semantic Defaults in the Analysis of Mood). In *Romance Languages and Linguistic Theory 2003: Selected Papers from Going Romance 2003*, eds. Twan Geerts, Ivo van Ginneken, and Haike Jacobs, 269–309. Amsterdam: John Benjamins.

Schlenker, Philippe. 2012. Maximize presupposition and Gricean reasoning. *Natural Language Semantics* 20:391–429.

Schoorlemmer, Maaike. 1998. Possessors, articles, and definiteness. In *Possessors, predicates and movement in the determiner phrase*, eds. Artemis Alexiadou and Chris Wilder, 55–86. Amsterdam: John Benjamins.

Schumann, Otto. 1981. La relación lingüística chuj-tojalabal. In *Los legítimos hombres: Aproximación antropológica al grupo tojolabal*, ed. Mario Humberto Ruz, 129–169. México, DF: Universidad Autónoma de México.

Schwarz, Bernhard. 2001. Two kinds of long-distance indefinites. In *Proceedings of the thirteenth Amsterdam Colloquium*, eds. M. J. B. Stokhof and R. A. M. van Rooij, 1–8.

Schwarz, Bernhard. 2004. Indefinites in verb phrase ellipsis. *Linguistic Inquiry* 35:344–353.

Schwarz, Bernhard. 2011. Long distance indefinites and choice functions. *Language and Linguistic Compass* 5:880–897.

Schwarz, Florian. 2009. Two types of definites in natural language. Ph.D. dissertation, University of Massachusetts.

Schwarz, Florian. 2012. Situation pronouns in Determiner Phrases. *Natural Language Semantics* 20:431–475.

Schwarz, Florian. 2013. Two kinds of definites cross-linguistically. *Language and Linguistic Compass* 7:534–559.

Schwarz, Florian. 2014. How weak and how deinite are Weak Definites? In *Weak referentiality*, eds. Ana Aguilar-Guevara, Bert Le Bruyn, and Joost Zwarts, 213–235. John Benjamins Publishing Company.

Schwarz, Florian. 2019. Weak vs. strong definite articles: meaning and form across languages. In *Definiteness across languages*, eds. Ana Aguilar-Guevara, Julia Pozas Loyo, and Violeta Vázquez-Rojas Maldonado. Language Science Press.

Schwarzschild, Roger. 2002. Singleton indefinites. *Journal of Semantics* 19:289–314.

Šereikaitė, Milena. 2019. Strong vs. weak definites: evidence from Lithuanian adjectives. In *Definiteness across languages*, eds. Ana Aguilar-Guevara, Julia Pozas Loyo, and Violeta Vázquez-Rojas Maldonado, 83–112. Language Science Press.

Sharvy, Richard. 1980. A more general theory of definite descriptions. *The Philosophical Review* 89:607–624.

Shima, Etsuro. 2000. A preference for move over merge. *Linguistic Inquiry* 31:375–385.

Silk, Alex. 2016. *Discourse contextualism: A framework for contextualist semantics and pragmatics*. Oxford, UK: Oxford University Press.

Simonenko, Alexandra. 2014. Grammatical ingredients of definiteness. Ph.D. dissertation, McGill University, Montreal, QC.

Simpson, Andrew. 2017. Bare classifier/noun alternations in the Jinyun (Wu) variety of Chinese and the encoding of definiteness. *Linguistics* 55:305–331.

Singh, Raj. 2011. Maximize presupposition! and local contexts. *Natural Language Semantics* 19:149–168.

Speas, Margaret, and Naoki Fukui. 1986. Specifiers and projections. In *MIT Working Papers in Linguistics*, volume 8. Cambridge, MA: MITWPL.

Sportiche, Dominique. 2006. Reconstruction, binding and scope. In *The Blackwell companion to syntax 4*, eds. Martin Everaert and Henk van Riemsdijk, 35–93. Blackwell Publishers.

Stalnaker, Robert. 1974. Pragmatic presuppositions. In *Semantics and Philosophy*, eds. Milton Munitz and Peter Unger, 197–213. New York, NY: New York University Press.

Stiebels, Barbara. 2006. Agent focus in Mayan languages. *Natural Language and Linguistic Theory* 24:501–570.

Strawson, Peter F. 1950. On referring. *Mind* 59:320–344.

Su, Yi-Ching. 2020. Backward/forward anaphora in child and adult Mandarin Chinese. *Language Acquisition* 27:187–216.

Szabolcsi, Anna. 1983. The possessor that ran away from home. *The Linguistic Review* 3:89–102.

Szabolcsi, Anna. 1987. Functional categories in the noun phrase. In *Approaches to Hungarian*, ed. Istvan Kenesei, volume 2, 167–190. JATE Press.

Tada, Hiroaki. 1993. A/A-bar partition in derivation. Ph.D. dissertation, MIT, Cambridge, MA.

Tai, James H. Y. 1973. Chinese as a SOV language. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, volume 1, 659–671. Chicago Linguistic Society.

Takahashi, Shoichi. 2010. Traces or copies, or both: Part I: Characterizing movement properties. *Language and Linguistics Compass* 4:1091–1103.

Tonhauser, Judith, David Beaver, Craige Roberts, and Mandy Simons. 2013. Toward a taxonomy of projective content. *Language* 66–109.

Trechsel, Frank R. 1995. Binding and coreference in Jakaltek. In *Grammatical Relations: Theoretical approaches and empirical questions*, eds. Clifford S. Burgess, Katarzyna Dziwirek, and Donna Gerdts, 449–471. Chicago University Press.

Tyler, Matthew. 2021. Two kinds of external possession in Mississippi Choctaw. *Syntax* 24:77–122.

van Urk, Coppe. 2015. A uniform syntax for phrasal movement: A case study of Dinka Bor. Ph.D. dissertation, Massachusetts Institute of Technology, Cambidge, MA.

Vázquez Álvarez, Juan J. 2011. A grammar of Chol, a Mayan language. Ph.D. dissertation, University of Texas Austin, Austin, TX.

Vázquez Hernández, María Rosalinda. 2022. Las cláusulas relativas en el tojol-ab'al. Master's thesis, Centro de Investigaciones y Estudios Superiores en Antropología Social, México, DF.

Vázquez-Rojas Maldonado, Violeta, B. G. Gonzalez, and A. R. Corte. 2017. Cuestionario para identificar frases nominales de referencia definida "simple". Ms. El Colegio de México.

Velleman, Leah Bridges. 2014. Focus and movement in a variety of K'ichee'. Ph.D. dissertation, University of Texas Austin, Austin, TX.

Vinogradov, Igor. 2012. Modality in the grammar of Modern Mayan Languages. In *Proceedings of the First Central European Conference in Linguistics for postgraduate Students*, 332–345.

von Fintel, Kai. 1994. Restrictions on Quantifier Domains. Ph.D. dissertation, University of Massachusetts, Amherst.

von Fintel, Kai. 1999. Quantifier domain selection and pseudo-scope. Paper presented at the Cornell conference on context-dependence.

von Fintel, Kai. 2004. Would you believe it? the king of france is back! (presuppositions and truth-value intuitions). In *Descriptions and Beyond*, eds. Marga Reimer and Anne Bezuidenhout, 315–341. Oxford University Press.

von Fintel, Kai. 2008. What is presupposition accommodation, again?n accommodation, again? *Philosphical Perspectives* 22:137–170.

von Heusinger, Klaus. 2002. Specificity and definiteness in sentence and discourse structure. *Journal of Semantics* 19:245–274.

von Heusinger, Klaus. 2019. Indefiniteness and specificity. In *The Oxford Hanbook of Reference*, eds. Jeanette Gundel and Barbara Abbott, 146–167. Oxford University Press.

von Heusinger, Klaus, and Jaklin Kornfilt. 2005. The case of the direct object in Turkish: Semantics, syntax and morphology. *Turkic languages* 3–44.

von Heusinger, Klaus, and Jaklin Kornfilt. 2017. Partitivity and case marking in Turkish and related languages. *Glossa: A journal of general linguistics* 2.

von Heusinger, Klaus, and Roya Sadeghpoor. 2020. The specificity marker *-e* with indefinite noun phrases in Modern Colloquial Persian. In *Nominal anchoring: Specificity, definiteness, and article systems across languages*, eds. Kata Balogh, Anja Latrouite, and Robert D. Van Valin, 115–147. Berlin, Germany: Language Science Press.

Westerståhl, Dag. 1984. Determiners and context sets. *Generalized quantifiers in natural language* 4:46–71.

Wilkins, David. 2018. The demonstrative questionnaire: "THIS" and "THAT" in comparative perspective. In *Demonstratives in cross-linguistic perspective*, eds. Stephen C. Levinson, Sarah Cutfield, Michael J. Dunn, N. J. Enfield, and Sérgio Meira, 43–71. Cambridge University Press.

Williams, Barbara, and Kenneth L. Williams. 1966. Chuj. In *Lenguas de Guatemala: Cuadernos del Seminario de Integración Social Guatemalteco*, ed. Marvin K. Mayers, 311–330.

Williams, Barbara, and Kenneth L. Williams. 1971. Chuj pantheon. In *According to our ancestors: Folk texts from Guatemala and Honduras*, ed. Mary Shaw, 97–108/325–328. Summer Institute of Linguistics of the University of Oklahoma.

Williams, Edwin. 1997. Blocking and anaphora. *Linguistic Inquiry* 28:577–628.

Wilson, George. 1978. On definite and indefinite descriptions. *The Philosophical Review* 87:48–76.

Wiltschko, Martina. 2013. Descriptive relative clauses in Austro-Bavarian German. *Canadian Journal of Linguistics* 58:157–189.

Winter, Yoad. 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 399–467.

Wolter, Lynsey Kay. 2006. That's that: The semantics and pragmatics of demonstrative noun phrases. Ph.D. dissertation, University of California Santa Cruz, Santa Cruz, CA.

Woolford, Ellen. 1991. VP-internal subjects in VSO and non-configurational languages. *Linguistic Inquiry* 22:503–540.

Zainuddin Hakim. 1991. *Rupama: certia rakyat Makassar*. Jakarta: Departamen Pendidikan dan Kebudayaan.

Zavala, Roberto. 1992. *El Kanjobal de San Miguel Acatán*. México: UNAM.

Zavala, Roberto. 2000. Multiple classifier sytems in Akatek (Mayan). In *Systems of Nominal Classification*, ed. Gunter Senft, 114–146. Cambridge: Cambridge University Press.

Zavala, Roberto. 2007. Inversion and obviation in Mesoamerica. In *Endangered Languages*, eds. Peter Austin and Andrew Simpson, volume Linguistische Berichte Sonderheft 14, 267–306. Hamburg: Helmut Buske Verlag.

Zavala Maldonado, Roberto. 2017. Alignment patterns. In *The Mayan Languages*, eds. Judith Aissen, Nora C. England, and Roberto Zavala Maldonado, 226–258. New York: Routledge.

Zwart, Jan-Wouter. 2002. Issues related to a Derivational Theory of Binding. In *Derivation and explanation in the Minimalist Program*, eds. Samuel David Epstein and T. Daniel Seeley, 269–304. Oxford: Blackwell Publishers.