UNIVERSITY OF TROMSØ UIT
FACULTY OF HUMANITIES, SOCIAL SCIENCES AND EDUCATION
CENTER FOR ADVANCED STUDY IN THEORETICAL LINGUISTICS

Context and (Co)reference

in the syntax and its interfaces

Sandhya Sundaresan

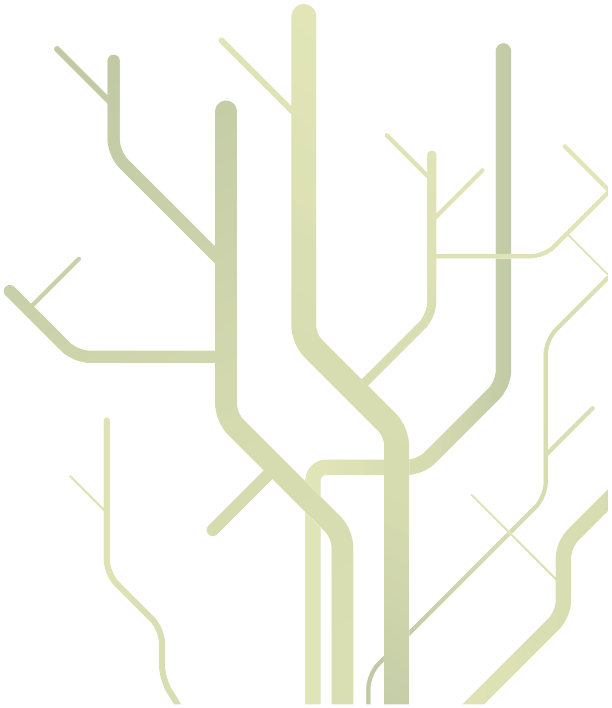
A thesis submitted for the degree of :
Philosophiae Doctor to the Faculty of Humanities, Social Sciences and Education, University of Tromsø
Doctor Philosophiae to the Faculty of the Humanities, University of Stuttgart

Thesis submitted under a co-tutelle agreement between:

Universität Stuttgart

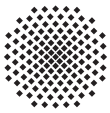
UNIVERSITY OF TROMSØ UiT

November 2012



Context and (Co)reference
in the syntax and its interfaces

Sandhya Sundaresan



Universität Stuttgart

A thesis submitted under a co-tutelle agreement
between the University of Stuttgart and the University of Tromsø
for the degree of:

Doctor Philosophiae to the Faculty of the Humanities
University of Stuttgart,

Philosophiae Doctor to the Faculty of Humanities,
Social Sciences and Education, University of Tromsø

November, 2012

Contents

Acknowledgements

Writing this dissertation has been an exercise in humility. The process of dissertation-writing has been, for me, a gradual unravelling of how much I did *not* know and still needed to learn. For helping me, often protesting, with this realization, I am grateful to Gillian Ramchand. I have to thank her for taking me on as an advisee back when I was still trying to find my academic footing in Tromsø (or indeed anywhere) and for then showing me what advisorship can be like when it is uncompromisingly and unapologetically good. Gillian has a knack for cutting through the fluff and rhetoric and getting right into the unyielding heart of the matter. For her uncompromising stance on quality and her clarity of thought which got me and this dissertation out of many logical dilemmas, for teaching me to be guided by the data even (and especially) when it leads to sticky theoretical situations, and for keeping me frustrated and on my toes through most of the actual writing process, I am very thankful to Gillian.

Artemis Alexiadou was my second advisor. I thank her for taking a leap of faith and agreeing to be my advisor when I returned from academic hibernation and decided to give this linguistics thing another shot. I am also grateful to her for the opportunity to teach linguistics courses in Stuttgart, which I was surprised to discover that I truly enjoyed, and which played a non-trivial role in resparking my interest in the field. Money is always useful, even/especially to a floundering academic: I also thank Artemis for helping me get the LGFG stipend which supported part of my research.

Thomas McFadden found Tamil case and nominalizations so delightfully anomalous that he wanted to co-write a paper about it with me back in 2009. Although the dissertation has since proceeded along its own (and in many ways rather different) trajectory, the initial spark for writing it was ignited by that paper and the many heated discussions (at the whiteboard and chalkboard, on notepaper and bar napkins) on clausal finiteness and subject reference surrounding it. For this alone

I am exceedingly grateful. Tom is such a "regular guy" that it's easy to forget how eccentrically smart he really is, a fact that was brought home to me more than once in the course of writing this dissertation. For reading and rereading the various drafts of various chapters which I shoved under his nose from time to time, for offering insightful input, especially on the more technical morphological and syntactic points of this dissertation, and for gamefully hearing out my half-baked ideas on all things linguistic and being genuinely excited about some of them, I am thankful to him.

Thanks to Peter Svenonius for meeting with me on more than one occasion and for his insightful and meticulous feedback on earlier chapter drafts. The discussion of feature-structure and the clausal functional sequence in this dissertation, in particular, have benefitted a great deal from Peter's input. Michal Starke taught my favorite class at CASTL and perhaps one of the most interesting linguistic seminars I have ever taken (Introduction to Nanosyntax). I am also grateful to Michal for meeting with me to discuss my work. Marit Westergaard and Gillian have both fought for me on various administrative issues more than once – notably two weeks before submitting this dissertation, when certain burocratic forces at the University of Tromsø let it come to pass that I might not be able to submit the dissertation here at all. I also thank Marit for doing her best to make me feel welcome at CASTL. A hat tip also to Tore Bentz for help with various burocratic issues: I have truly enjoyed every one of your emails to me saying that my travel funding has been approved. It seems only right to acknowledge CASTL itself: it would be difficult to find a more intellectually stimulating environment for a linguist, not just high up here in the Arctic, but anywhere.

Others have discussed various aspects of my research with me, in formal meetings as well as over food and drinks, and/or provided written feedback on the earlier avatars of this dissertation. These are, in no particular order: Pavel Caha, Bettina Gruber, Alessandra Giorgi, Pranav Anand, Philippe Schlenker, Rajesh Bhatt, Daniel Hole, Sergey Minor, Tarald Taraldsen, Gennaro Chierchia, Pavel Rudnev, Angelika Kratzer, Hedde Zeijlstra (the one and only! – you win this bet, Hedde), Ken Safir, Denis Delfitto, Mark Baker, Rahul Balusu, Liliane Haegeman, Halldór Sigurðsson, Veneeta Dayal, Luca Ducceschi, Guglielmo Cinque, Anna Szabolcsi, Luigi Rizzi, K.A. Jayaseelan, R. Amritavalli, Ayesha Kidwai, David Adger, Idan Landau, Per Erik Solberg, and Björn Lundquist. If I have left anyone out, and I hope I haven't, please blame my recent sleeplessness, not my intent. I am also grateful to the audiences at the

various conferences and workshops where earlier variants of this dissertation (particularly Part II and Part III) were presented for their helpful feedback. Thanks are also due to my many native speaker informants: to the Tamil speakers around the world (too many to individually list here) who were gracious enough to suffer through long and laborious sessions answering questions about agreement, anaphora, and monsters in my survey (I promise to make them more fun next time!), to Rosmin Mathew for meeting with me about Malayalam on more than one occasion, to Rahul Balusu and Sindhu Herur for help with Telugu and Kannada judgments, respectively, and to Kristine Bentzen, Madeleine Halmøy, and Per Erik Solberg for help with Norwegian. Finally, thanks to Martin Krämer for proofreading the German summary of this dissertation, to Kristine, again, for translating the Norwegian one, and to Pavel Iosad for Linux-related help on various occasions.

A lift of the pint (in random order because it's too late to alphabetize!) to Antonio Fábregas, Martin Krämer, Anna Wolleb, Iulia Zegrean, Violeta Martínez-Paricio, Pavel Caha, Sergey Minor, Björn Lundquist, Peter Jurgec, Pavel Iosad, Marleen van de Vate, Olga Urek, Minjeong Son, Luca Ducceschi, Alexander Pfaff, Inna Tolskaya, Éva Dékany, Minjeong Son, Kristine Bentzen, Kaori Takamine, Gillian Ramchand, Madeleine Halmøy, Irene Franco, Eefje Boef, Juan Romeu Fernández (*slap!*), Kristine Bentzen, Madeleine Halmøy, Sebastian Bican, Marta Velnic, Tammer Castro, Sandra Ronai, Andrea Márkus, Rosmin Mathew, and Michal Starke (ping-pong player extraordinaire) for making Tromsø such a fun place to be, and not just because of the linguistics. Thanks, in particular, to Anna and Odd for lending us their car for almost a year in Tromsø and again now, this last week before dissertation-submission – we've enjoyed it to the fullest! To friends and colleagues elsewhere – Hedde Zeijlstra, Bettina Gruber, Bill Poser, Jeff Parrott, Mike Putnam, Damien Hall, Allison Silver Adelman, Tatjana Scheffler, Sudha Arunachalam, Clare Lindner, Katie Ruhl, Rachael Chipkin, Sara Gibson, Ramshreyas Rao, Sriram Thiagarajan, and my good friends Daniel Ezra Johnson and Joel Wallenberg, two of the smartest and most delightfully eccentric people I know: thanks to all of you for being a part (past, present, and present perfect) of my adventures on three different continents. A significant part of the dissertation-writing took place in Crete, and while having a view of the bright blue Aegean didn't necessarily help me understand monsters or the intricacies of long-distance binding better, it did make thinking about them a hell of a lot of fun.

My sister, Sudha Lakshmi has, despite having been released into the

world five years earlier, always made *me* look like the beta version and my brother-in-law Greg Demirchyan likes pineapple vodkas more than anyone reasonably should: thanks to both of them for hosting us at their apartment in Berkeley for more than a month in Winter 2011-2012, thereby ensuring that our working vacation could be conducted amidst sunlight and greasy Mexican food aplenty. My mother-in-law Irene McFadden has always been very sweet and supportive over the years, and I am grateful. To thank my parents for my existence is trivial, to thank them for an intellectually stimulated, fun-filled and happy one is not: I thank you for both.

To Tom – who deserves his own paragraph, and whom it is necessary but not sufficient to thank – my necessary and sufficient condition.

Abbreviations

I use the following abbreviations. Where possible, I have followed the conventions of the Leipzig Glossing Rules.

1	First person
2	Second person
3	Third person
ABL	Ablative
ACC	Accusative
ADDR	Addressee
ADJ	Adjective/Adjectival
ADV	Adverb/Adverbial
ALL	Allative
ANAPH	Anaphor
ANT	Anterior
ASP	Aspect
AUX	Auxiliary
BEN	Benefactive
CAUS	Causative
CL	Clitic
COMP	Complementizer
COND	Conditional
DAT	Dative
DECL	Declarative
DEF	Definite
EMPH	Emphatic
EPIST	Epistemic
EVAL	Evaluative
EVID	Evidential
EXCL	Exclamative
F	Feminine
FUT	Future
G	Gender

GEN	Genitive
GER	Gerund
HAB	Habitual
HON	Honorific
IMP	Imperative
INF	Infinitive
INS	Instrumental
INTR	Intransitive
LOC	Locative
LOG	Logophoric
M	Masculine
N	Neuter
NACT	Non-active
NEG	Negative
NMLZ	Nominalizer
NOM	Nominative
NPI	Negative Polarity Item
NUM	Number
OBJ	Object
OBL	Oblique
P	Person
PRF	Perfect
PL	Plural
PRS	Present
PROG	Progressive
PTCP	Participle
PST	Past
Q	Question
REFL	Reflexive
REL	Relative
SBJ	Subject
SBJV	Subjunctive
SG	Singular
TR	Transitive

Abstract

It is well known that referentially defective nominals fall into two broad categories: pro-forms whose reference seems structurally constrained (local anaphors, OC PRO) and those which are discourse-pragmatically conditioned (logophors, deictic pronouns, indexicals). Nevertheless, a strict binary distinction cannot be maintained because most actually straddle the syntax-discourse divide: e.g. deictic pronouns can be variable-bound, indexicals may be "shifted" under certain intensional operators, and logophors and long-distance anaphors often look and behave alike.

The central thesis of this dissertation is that a proper subset of pro-forms can receive a unified analysis under an enriched grammatical model that posits the syntactic representation of mental and/or spatio-temporal perspective. To this end, I present novel evidence from verbal agreement triggered under anaphora to show that even so-called "logophoric" reference involves an indelible syntactic core. I propose that perspective is featurally represented on a silent pronominal operator in the specifier of a Perspectival Phrase (PerspP) at the phasal-edge of certain CPs, PPs, DPs, and AspPs and may be exploited to yield a unified account of anaphora and agreement patterns triggered under it. Anaphora involves two distinct dependencies: an Agree relationship between the anaphor and the operator in the [Spec, PerspP] of its minimal phase, which is the equivalent of syntactic binding, and a conceptual relationship between the antecedent and this operator, which is the equivalent of non-obligatory control. Thus, all binding is local and syntactic; all antecedence is non-local and (primarily) non-syntactic. I also illustrate that perspective must be kept conceptually and structurally distinct from the Kaplanian utterance context and the intensional "context" responsible for indexical shift.

The main language of investigation is the Dravidian language Tamil but crosslinguistic comparisons are made with: Abe, Aghem, Amharic, Czech, Donna Sɔ, Dutch, English, French, German, Greek, Icelandic, Italian, Japanese, Kannada, Korean, Malayalam, Mupun, Navajo, North Sami, Norwegian, Romanian, Russian, Slave, Swahili, Telugu, Uyghur, and Zazaki. The Tamil judgments are bolstered by the results of an

online survey conducted among 38 native speakers around the world.

To the memory of my grandfathers, who were the first doctor and first linguist I knew.

Chapter 1

Introduction

The central role of perspective in guiding linguistic dependencies has long been recognized in the literature, particularly in the realm of nominal anaphora (Clements 1975, Kuno 1987, Sells 1987, among others). However, the traditional wisdom has been to treat perspective as a purely discourse-pragmatic concept, one that, furthermore, doesn't interface with the core grammatical modules of syntax and semantics. As such, evidence indicating that binding is sensitive to structural constraints, like locality, minimality, and c-command (Chomsky 1981, Pica 1987, Huang and Tang 1991, Progovac 1993), has automatically been taken as evidence *against* the involvement of perspective. Conversely, evidence supporting binding sensitivity to discourse-pragmatic factors, as in the case of the so-called "logophoric" phenomena, has automatically been taken as evidence against the involvement of structure. This has led to a splitting up of binding phenomena into (at least) two categories: those that are structurally regulated and those that are conceptually driven. The problem is that such a dichotomy doesn't seem to be empirically justified: the putatively conceptual binding phenomena also manifest sensitivity to structural constraints and the putatively structural ones show the effects of thematic and discourse-pragmatic factors.

In this dissertation, I conduct a detailed analysis of binding patterns in the Dravidian language Tamil to show that a more elegant and unified approach to binding, in Tamil and languages like it, might be achieved if we give up the assumption that perspective is not structural. The foundational thesis is that the perspectival relationship that holds between an individual and a predication is structurally represented on a functional projection and is central to determining local and long-distance dependencies in the grammar. I argue that a given phase has a "perspectival center", defined as in (1) below:

(1) **Formal representation of a Perspectival Center:**

i. The *perspectival center* contains the coordinates pertaining to the time, location, world, and/or mental information of a salient perspective holder. These are hosted in a silent pronominal operator in the specifier of a Perspectival Phrase (PerspP).
ii. Certain phases (at least some PPs, DPs, AspPs, CPs) contain a PerspP by virtue of what they inherently "mean".
iii. A phase has at most one PerspP.
iv. The phase containing a successfully bound anaphor must contain a PerspP. The operator in [Spec, PerspP] Agrees with the anaphor in its minimal phase and variable-binds it at LF.
v. A potential antecedent may not be asymmetrically c-commanded by the PerspP it holds a perspective towards. The relationship between the potential antecedent and the operator in [Spec, PerspP] is one of non-obligatory control.

In this connection, I will argue that binding is a two-step, cross-modular process. The first involves a purely syntactic relationship, formalized as Agree, between the anaphor and the operator in [Spec, PerspP] of its local phase. The second is a dependency between this operator and the DP that will end up being construed as the linguistic antecedent of the anaphor, which may be construed as an instantiation of non-obligatory control (Williams 1980). This relationship is predominantly conceptual in nature – involving sensitivity to discourse-pragmatic factors such as speaker-intent, common ground, presupposition, and salience, which in turn restrict the way in which the assignment function applies to the operator in [Spec, PerspP]. There is thus no direct connection between the anaphor and its antecedent. All binding is local and all antecedence is non-local.

Primary empirical evidence for this will come from local and long-distance binding and indexical shift and related agreement patterns in Tamil. Tamil is an ideal language to explore in connection to this thesis because it presents robust evidence for anaphora in a wide array of linguistic environments. The Tamil simplex anaphor *ta(a)n* may be anteceded clause-internally by a co-argument, across clauses by a c-commanding, or sometimes non-c-commanding, DP and also "logophorically" by an entity in the salient discourse which is not represented by a DP within the sentence. This makes it possible to see binding in all its facets and distill that which is common and essential to all. The language also utilizes a strategy for co-argument binding involving a verbal mor-

pheme *ko*⌊ which, from its non-binding related uses, sheds light on the interaction between perspective and thematic relations within the clause. A very important, and typologically rather unique, property of Tamil is that it allows anaphors to occur in agreement-triggering position (Annamalai 1999). The verbal agreement triggered under the anaphor will allow me to present novel evidence showing that long-distance as well as the so-called "logophoric" binding involve a core structural component which will, in turn, provide important support for the hybrid binding model described above. This will also be seen to illuminate a special case of indexical shift (von Stechow 2002, Schlenker 2003b, Anand 2006), which I label "monstrous agreement", in this language, which obtains in conjunction with anaphora and is realized on the agreement itself.

This dissertation is divided into four main parts. Part I looks at long-distance anaphora in Tamil and builds up the case for a perspectival analysis of this phenomenon. Much of the evidence will be geared toward showing that the phenomena labelled variously as "long-distance binding", "backward binding", and "logophoric binding" all involve a core syntactic subcomponent as well as a conceptual one and that perspective is central to both. Based on this evidence, I will develop the fundamental two-step binding model described above, and motivate a specific proposal about where in the structure the Perspectival Phrase is hosted and what the feature-compositions of the anaphor and the operator in [Spec, PerspP] are. The intuition that an anaphor is a "perspective seeker" and that an antecedent is a "perspective-holder" toward the minimal PerspP containing the anaphor, will be central to this discussion. On this basis, I wil propose concrete syntactic derivations for the various descriptive categories of anaphora including the so-called "logophoric" binding.

Part II extends the insights of Part I to the case of co-argument binding. I show that here too perspective plays a central role and indeed the two-step binding model must be applied here as well. Much of the discussion will be centered around the description of *ko*⌊, a verbal suffix that must, in the standard case, mark structures involving co-argument anaphora. Particularly instructive will be the uses of *ko*⌊ outside the domain of anaphora and the contexts where co-argument binding is possible without *ko*⌊, as in the case of psych-predicate structures. The former will help us understand that the use of perspective is not confined to the context of anaphora: we will see that it plays a central role in the instantiation of unaccusativity, for instance. At the same time, we will also see that unaccusatives and reflexives are not the only ones privileged to instantiate perspectival relations, thus calling into question standard analyses in the literature that conflate the two: *ko*⌊ will be seen

to mark many non-reflexive transitives and unergatives. The latter will yield insight into what makes psych-predicates special and how argument structure is relevant for perspectival relationships. Co-argument binding will also bring to light another structural constraint on perspectival relations in general, namely that an intended perspective-holder must not be properly contained inside the PerspP that it holds a perspective towards. Extended to the case of anaphora, this means that an intended antecedent may not be properly embedded inside the minimal PerspP containing the anaphor.

Part III zooms in on a special type of long-distance binding attested in Tamil. This involves the anaphor *ta(a)n* in subject position in the clausal complement of a propositional speech predicate. The agreement under the anaphor is marked 1st-person but tracks the 3rd-person antecedent of the anaphor. I will show that this agreement, just like the "regular" agreement from Part I, is not triggered by the anaphor itself but by the pronominal operator in [Spec, PerspP]. I will argue, moreover, that when the agreement triggered by the operator is 1st-person, as described above, the operator itself is a shifted 1st-person indexical. In addition to showing that indexical shift must be syntactically implemented, this type of structure will also allow me to clarify the distinction between "perspective" and "context", as defined in the sense of Kaplan (1989) – two concepts that are often wrongly conflated in the literature. Part of this clarificatory process will involve illustrating that, contra Schlenker (2003b), the categories of anaphor/logophor and obligatorily shifted indexical must be kept distinct.

Part IV discusses the broader theoretical implications and empirical outlook of this dissertation. I propose that a wide range of languages like Icelandic, Japanese, and Italian whose anaphors display properties very similar to those described for Tamil and spatial anaphora in Norwegian and Dutch may also be accommodated under the hybrid PerspP binding model developed here. The dissertation also has significant implications for the distinctions between anaphora and indexicality in general and for the way in which certain argument-structural phenomena like the GET-passive and the nature of the relationship between unaccusativity and reflexivity should be handled.

My analysis will be couched within the Minimalist framework (Chomsky 2000; 2001), and I assume a Y-modular architecture with late insertion (Halle and Marantz 1993). This means that the narrow syntactic portion of the derivation has access only to abstract syntactic features, not to phonological or (purely) semantic information. I further assume that features can be valued or unvalued, but do not adopt an additional

interpretable-uninterpretable distinction.

A few words are appropriate here about the primary language of investigation. Tamil is a language of the Dravidian family, spoken natively in Tamil Nadu, India, and also in parts of Sri Lanka, Singapore, Malaysia, Mauritius, and Réunion by approximately sixty-five million people (`http://www.ethnologue.com/show_language.asp?code=tam`). It is an SOV language, is head-final, exhibits subject as well as *pro*-drop and rampant scrambling, and shows very rich agreement. Tamil is also severely diglossic (see Schiffman 1995, among others): the spoken language is not typically written, and the written form is often quite different, especially in phonological and morphological ways, from the spoken one. In this dissertation, I have tried to be as faithful to the spoken version as possible. As such, the transliterations of the Tamil examples throughout, reflect as much as possible colloquial pronunciation and are not a faithful rendering of the written form. However, in special cases, I have taken the liberty of retaining the written form so as to be able to delineate morphemic boundaries more clearly. These differences are, in any case, orthogonal to the issues discussed here, as the spoken and written forms of the language are not distinguished in their possibilities for anaphora. But anyone who looks in this work for instances of "pure" Tamil will be likely disappointed.

Regarding the specific transliteration used, I have mostly adopted the standard IPA symbols, e.g. using /ɖ, ʈ, ɭ, ɻ/ for the retroflexes and /ʤ/ for the voiced palatal affricate, reserving /j/ for the palatal glide. For better readability, I have however adopted /š, č/ for the voiceless palatal fricative and affricate, respectively, and /ŭ/ for the back unrounded pronunciation of phonemic short /u/ in non-initial syllables.

I am a native speaker of Tamil and, to a large extent, the Tamil sentences presented here are based on my own judgments. However, all the Tamil data is additionally informed by the results of a grammaticality survey on anaphoric patterns which was carried out online in three parts and taken by Tamil native speakers of different dialects around the world. Parts I and II, which tested the nature of the relationship between anaphora, indexicality, and agreement were taken by 38 and 19 respondents, respectively. Part III, which tested properties of binding alone was taken by 15 respondents. The results of the survey were analyzed to identify points of significant dialectal variation as well as to single out those invariant properties that were shared by all. The results of this microvariation are discussed where relevant but unless otherwise noted, the dialect of investigation, reported through this dissertation, is my own.

Part I

Long-distance binding in Tamil

Chapter 2

Background issues

2.1 Introduction

The focus of this series of chapters in Part I is the long-distance binding of the Tamil anaphor *ta(a)n*, in both subject and object position. Prototypical examples of this phenomenon are given in (2)-(3). In (2), *ta(a)n* is the subject of the embedded clause, and in (3), it is in embedded direct object position:

(2) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ ʤejkka-poo-r-aan-nnŭ]
Raman[NOM] ANAPH[NOM] prize-ACC win-go-PRS-3MSG-COMP
kaɳḍupiḍi-čč-aan.
find.out-PST-3MSG
"Raman$_i$ found out [$_{CP}$ that he$_{\{i,*j\}}$ was going to win the prize.]"

(3) Raman$_i$ [$_{CP}$ Seetha tann-æ$_{\{i,*j\}}$ kaadali-kkir-aaḷ-
Raman[NOM] Seetha[NOM] ANAPH-ACC] love-PRS-3FSG-
ŭnnŭ] kaɳḍupiḍi-tt-aan.
COMP find.out-PST-3MSG
"Raman$_i$ found out [$_{CP}$ that Seetha$_j$ loved him$_{\{i,*j\}}$.]"

In the course of Part I, I will examine the properties of *ta(a)n*-binding in sentences like these as well as ones where the binding relation appears to be less straightforward. These will include structures that instantiate "backward binding" of *ta(a)n* – a phenomenon where the anaphor seems to c-command its antecedent rather than the other way around (Minkoff 2003) – and "logophoric" reference, where *ta(a)n* refers to an entity in the salient discourse (not overtly represented in the sentence) which holds a perspectival relationship toward the sentence in which *ta(a)n* occurs. I develop a precise account of the conditions that a linguistic entity has to fulfill in order to qualify as a potential antecedent of *ta(a)n*. It will

be seen that these conditions are not deterministic but tendential, and influenced by thematic and discourse factors (which are themselves, in many cases, nebulous). Despite this underlying vagueness which, I will show, is inherent to the antecedent-anaphor relationship, the conditions will be flexible enough to cover the full range of apparently distinct empirical patterns yet precise enough to predict possibilities on the potential antecedence of *ta(a)n*.

The nature of these conditions also helps disentangle the individual contributions of thematic and discourse properties of potential antecedence from the syntactic and LF-semantic restrictions on the binding of *ta(a)n*. I will construe binding as a two-step process. The first step will involve a syntactic Agree relationship between the anaphor *ta(a)n*and a silent pronominal operator in the specifier of a "perspectival-center" in the local left periphery of the *ta(a)n*-phase. The perspectival center is a linguistic object containing coordinates pertaining to the identity, time, location, and potentially world of an entity which holds a perspective toward the situational predication in which *ta(a)n* is a thematic participant. The perspectival center is analogous to the enriched intensional "index" of Lewis (1979) and the "internal logophoric center" of Bianchi (2003) but is broader in both conception and empirical applicability because, unlike these, it can be introduced by strategies other than complementation by a propositional predicate. This modification will be crucial because it will allow me to implement long-distance binding into adjunct PPs, DPs and CPs, as well as logophoric and backward-binding phenomena. The second step of the binding process will involve LF wellformedness conditions which regulate the successful association of the silent operator onto the entity in the evaluation context whose linguistic representation will be the actual antecedent of *ta(a)n*. The relationship between the DP denoting this actual antecedent and the silent operator is construed as one of non-obligatory control (Williams 1980).

A major contribution of this model is that it allows the unification of so-called "logophoric" binding – a phenomenon that has posed a nontrivial challenge to analyses of binding in the literature – with standard long-distance binding as well as more problematic binding phenomena like backward binding, among others. We will see that this same model may also be seamlessly extended to structures involving the local binding of *ta(a)n*. This analysis thus represents a synthesis of the purely syntactic and conceptual approaches to long-distance binding which have been proposed in the past, each of which on its own offers important insights but fails to account for the full range of empirical patterns seen to characterize binding relationships crosslinguistically.

2.2 Tamil *ta(a)n*: a primer

Tamil *taan* is a morphologically simplex anaphor. Its nominative form, typically used when it occurs in syntactic subject position, is *taan*. All other cases, such as accusative, dative, genitive, comitative, ablative, instrumental, locative, and genitive, involve suffixes which attach to the oblique stem of the anaphor, which is *tan-*. Henceforth, I will use *ta(a)n* as a cover-term for its various surface forms. In addition to these case-suffixes, *ta(a)n* can also be marked for number: the singular is unmarked, and the plural is marked with the morpheme *-gaḷ*, which occurs sandwiched between the nominal stem and case suffixes, if any. The basic paradigms are in Table 2.1.[1]

	Singular	Plural
Nom	*taan*	*taaŋ-gaḷ*
Acc	*tann-æ*	*taŋ-gaḷ-æ*
Dat	*tan-akkŭ*	*taŋ-gaḷ-ŭkkŭ*
Gen	*tann-ooḍæ*	*taŋ-gaḷ-ooḍæ*
Ins	*tann-aal*	*taŋ-gaḷ-aal*
Com	*tann-ooḍŭ*	*taŋ-gaḷ-ooḍŭ*
Loc/All	*taŋ-giṫṯæ*	*taŋ-gaḷ-giṫṯæ*
Abl	*taŋ-giṫṯæ-rŭndŭ*	*taŋ-gaḷ-giṫṯæ-rŭndŭ*

Table 2.1: Case and number paradigms for Tamil *ta(a)n*

Moving on to its antecedence properties, *ta(a)n* can only take 3rd-person antecedents; 1st and 2nd-person antecedents are strictly ruled out. However, once the antecedent is fixed as 3rd person, its gender is irrelevant: i.e. *ta(a)n* can take 3rd-person antecedents with masculine, feminine, and neuter gender.[2] This constellation of properties has led to

[1] A quick clarification about some of these forms: in colloquial Tamil, the comitative and genitive forms are homophonous, and surface as the suffix: *-ooḍæ*, though they were morphophonologically distinguished in slightly older stages of Tamil. I present the older, and superficially distinct, forms of these cases here for purposes of clarity. The suffix *-kiṫṯæ* marks the locative as well as allative forms of a noun. However, when used as a locative suffix, *kiṫṯæ* only attaches to animate nouns – all inanimate noun stems take the locative suffix *-læ* instead. Since *ta(a)n* only refers to animate entities, only the *kiṫṯæ* locative suffix is relevant. Finally, notice that the ablative case form is built on the locative/allative suffix: however, I have treated it as a primitive case-form and category here, in keeping with the standard case classifications made in Tamil descriptive grammars (Schiffman 1995, Asher and Annamalai 2002).

[2] Tamil has no grammatical gender. As such, its masculine, feminine, and neuter

proposals that *ta(a)n* itself is specified as 3rd-person and is unmarked for gender (Annamalai 1999) – however, since these properties, strictly speaking, characterize the antecedent of *ta(a)n*, rather than *ta(a)n* itself, and in light of recent theoretical proposals questioning the ϕ-featural status of anaphors crosslinguistically (Kratzer 2009, Reuland 2011), we will make no such assumptions at this juncture. Nevertheless, the paradigms are clear and exceptionless, as illustrated by the sentences below. (4a) and (4b) show that 1st and 2nd-person nominals – i.e. indexical pronouns – cannot serve as antecedents for *ta(a)n*, respectively; (5a)-(5c) show that masculine, feminine, as well as neuter nouns may antecede this anaphor:

(4) ***1st-person and *2nd-person antecedents:**

a. * Naan$_{Auth}$ [$_{CP}$ Seetha$_i$ tann-æ$_{\{Auth,i\}}$ paar-tt-aaɭ-ŭnnŭ] so-nn-een.
I[NOM] Seetha[NOM] ANAPH-ACC see-PST-3FSG-COMP say-PST-1SG
"I$_{Auth}$ said [$_{CP}$ that Seetha loved me$_{Auth}$.]" (Intended)

b. * Nii$_{Addr}$ [$_{CP}$ pasaŋ-gaɭ$_i$ tann-æ$_{\{Addr,i\}}$ aɖi-tt-aaŋ-gaɭ-ŭnnŭ] nene-tt-aaj.
You[NOM] boys-PL[NOM] ANAPH-ACC hit-PST-3M-PL-COMP think-PST-2SG
"You$_{Addr}$ thought [$_{CP}$ that the boys hit you$_{Addr}$.]" (Intended)

(5) **Gender of antecedent is irrelevant:**

a. Vivek$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ Seetha-væ$_j$ paar-tt-adaagæ] namb-in-aan.
Vivek[NOM] ANAPH[NOM] Seetha-ACC see-PST-NMLZ believe-PST-3MSG
"Vivek$_i$ believed [$_{CP}$ that he$_{\{i,*j\}}$ saw Seetha$_j$]."

b. [Ellaa poɳ-gaɭ-ŭkk=um]$_i$ [$_{DP/PP}$ [$_{CP}$ Rajinikanth$_j$ [taŋ-gaɭ-ooɖæ]$_{\{i,*j\}}$ pees-in-aar] eŋgiradŭ-læ] rombæ sandoošam
All girls-PL-DAT=CL Rajinikanth[NOM] ANAPH-PL-GEN speak-PST-3MSG that-LOC very happy
"[All girls]$_i$ were very happy about the fact that Rajinikanth$_j$ spoke to them$_{\{i,*j\}}$."

c. Koɻendæ$_i$ [$_{CP}$ tan-akkŭ$_{\{i,*j\}}$ rombæ pasi-nnŭ] aɻŭ-d-adŭ.
child[NOM] ANAPH-DAT very hunger[NOM-COMP] weep-PST-3NSG

forms mark male, female, and neuter entities, respectively.

"The child$_i$ wept [$_{CP}$ that it$_{\{i,*j\}}$ was very hungry]."

As regards its surface distribution, we have seen that *ta(a)n* can occur in either the subject or object position of a clause. This property is illustrated again in the sentences below:

(6) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ ʤejkka-poo-r-aan-nnŭ]
Raman[NOM] ANAPH[NOM] prize-ACC win-go-PRS-3MSG-COMP
kaɳɖupiɖi-tt-aan.
find.out-PST-3MSG
"Raman$_i$ found out [$_{CP}$ that he$_{\{i,*j\}}$ was going to win the prize.]"

(7) Raman$_i$ [$_{CP}$ Seetha tann-æ$_{\{i,*j\}}$ kaadali-kkir-aaɭ-ŭnnŭ]
Raman[NOM] Seetha[NOM] ANAPH-ACC] love-PRS-3FSG-COMP
kaɳɖupiɖi-tt-aan.
find.out-PST-3MSG
"Raman$_i$ found out [$_{CP}$ that Seetha$_j$ loved him$_{\{i,*j\}}$.]"

Regardless of which position *ta(a)n* itself occurs in, it typically takes a c-commanding nominal in syntactic subject position as its antecedent – a general property, observed for anaphors in typologically unrelated languages like Icelandic and Italian (Koster and Reuland 1991a, Giorgi 2006), that has been termed the "subject orientation" of an anaphor. As such, a DP in oblique object position, as in the examples below, or in direct object position (in sentences where *ta(a)n* is in an adjunct phrase) cannot normally serve as an antecedent for *ta(a)n*:

(8) * Raman Krishnan-kiʈʈærundŭ$_i$ [$_{CP}$ taan$_i$ paris-æ ʤejkka-poo-r-aan-nnŭ] kaɳɖupiɖi-tt-aan.
Raman[NOM] Krishnan-ABL ANAPH[NOM] prize-ACC win-go-PRS-3MSG-COMP find.out-PST-3MSG
"Raman found out from Krishnan$_i$ [$_{CP}$ that he$_i$ was going to win the prize.]" (Intended)

(9) * Naan Raman-kiʈʈæ$_i$ [$_{CP}$ Seetha$_j$ tann-æ$_i$ kaadali-kkir-aaɭ-ŭnnŭ] so-nn-een
I[NOM Raman-ALL Seetha[NOM] ANAPH-ACC] love-PRS-3FSG-COMP find.out-PST-3MSG
"I told Raman$_i$ [$_{CP}$ that Seetha$_j$ loved him$_i$.]"(Intended)

In addition, when *ta(a)n* is in object position, it typically resists being bound by its clausemate subject. This again is not an isolated characteristic of Tamil but has been noted for simplex anaphors in languages as wide-ranging as Malayalam, Japanese, Korean, Dutch and Icelandic

(Reinhart and Reuland 1993, Jayaseelan 1997, among others). Such an apparent anti-locality property has been termed a "Condition B" effect in the literature, alluding to the GB-era Condition B rule (Chomsky 1981) which states that a pronoun must be free in its governing category (or local domain). However, both because this term is theory-laden and because it suggests a stronger connection to deictic pronouns than might actually be warranted, I will not use it to describe this property. At the same time, I am also loathe to describe it as a type of anti-locality effect since that presupposes, without adequate evidence at this point, that it is purely structural in nature. As such, I will simply label it the "Ban on Clausemate Subject Antecedence". Here are some representative examples of this ban:[3]

(10) * Raman$_i$ tann-æ$_i$ aḍi-tt-aan.
Raman[NOM] ANAPH-ACC hit-PST-3MSG
"Raman$_i$ hit himself$_i$." (Intended)

(11) * Krishnan [$_{CP}$ Raman$_i$ tann-æ$_i$ aḍi-tt-
Krishnan[NOM] Raman[NOM] ANAPH-ACC hit-PST-
aan-nnŭ] paar-tt-aan.
3MSG-COMP see-PST-3MSG
"Krishnan saw [$_{CP}$ that Raman$_i$ hit himself$_i$]." (Intended)

(12) Krishnan$_j$ [$_{CP}$ Raman tann-æ$_j$ aḍi-tt-
Krishnan[NOM] Raman[NOM] ANAPH-ACC hit-PST-
aan-nnŭ] paar-tt-aan.
3MSG-COMP see-PST-3MSG
"Krishnan$_j$ saw [$_{CP}$ that Raman hit him$_j$]."

(10) shows that in a simple sentence, object *ta(a)n* is typically ruled out altogether – since, in the standard case, the matrix subject is the only available antecedent.[4] (11) and (12) form a minimal pair: the only difference between them lies in the identity of the subject that is intended as the antecedent for *ta(a)n* in embedded object position. In (11), the intended antecedent of *ta(a)n* is its own clausemate subject, *Raman*; in (12), the superordinate subject *Krishnan* serves as antecedent. Just as the ban on clausemate subject antecedent would lead us to expect, (11)

[3]The patterns given here are representative of my own dialect of Tamil and quite robust. In the course of my online survey, I have found that, for some of these dialects, the ban on clausemate subject antecedent for *ta(a)n* doesn't actually hold. For these speakers, therefore, the sentences given in (10) and (11) above would be grammatical. This is discussed in greater detail at the end of Part II.

[4]The non-standard case is a free indirect discourse situation – where a "logophoric" entity that is not overtly represented in the sentence would serve as the antecedent.

is fully ungrammatical whereas (12) is licit.

Another important property of long-distance binding in Tamil is that it really is long-distance – in principle, *ta(a)n* can be bound by any superordinate subject regardless of how many clausal boundaries may intervene.[5] (13) below exemplifies such a structure:

(13) [$_{CP}$ Raman Anand-kiṭṭæ [$_{CP}$ Seetha tann-æ$_i$
Raman[NOM] Anand-ALL Seetha[NOM] ANAPH-ACC
kaappaatt-in-aaḷ-ŭnnŭ] so-nn-aan-nnŭ] Krishnan$_i$
save-PST-3FSG-COMP] say-PST-3MSG-COMP Krishnan[NOM]
paar-tt-aan.
saw-PST-3MSG
"Krishnan$_i$ saw [$_{CP}$ that Raman told Anand [$_{CP}$ that Seetha saved him$_i$.]]"

In (13), the matrix subject *Krishnan* binds *ta(a)n* across other DPs like *Anand* and even across another potential antecedent, the intermediate subject *Raman*. As such, the relationship between the anaphor and its antecedent seems not to be subject to clause-locality or Relativized Minimality.

The possibility of such structures in a number of languages around the world has always posed a bit of a puzzle. However, in some languages, like Chinese, so called "blocking effects" have been observed in structures involving dependencies across multiple clauses, such as that in (13), where the binding dependency seems to be sensitive to the presence of an intermediate nominal (potential antecedent or even otherwise) or functional element. Huang and Tang (1991), for instance, report for Chinese that the intervening potential antecedents must all match in ϕ-feature value with the actual (higher) antecedent in order for the long-distance binding relation to go through;[6] Giorgi (2006) reports parallel intervention effects in the functional domain in long-distance binding structures in Italian, arguing that while the subjunctive mood allows long-distance binding

[5]Restrictions on interpretability are, of course, imposed due to difficulties in processing – but there is no *theoretical* restriction on how far away an antecedent must be from the anaphor.

[6]The details of this condition have since been hotly contested, however. Huang and Liu (2001), for instance, claim that only medial indexicals (i.e. 1st or 2nd-person pronouns) count as interveners for the purposes of binding in Chinese. See also Cole et al. (2001) for discussion of varieties of blocking effect across languages. Regardless, the broader point is that, for at least some speakers, some type of intervention effect has been observed in long-distance binding structures in Chinese, suggesting that intervening nominals are not totally invisible to a long-distance binding relationship in this language.

across it, the indicative blocks it.

In Tamil, however, there is no clear evidence for a nominal or functional blocking effect of this sort. (14) shows that a medial subject with a mismatched 2nd-person feature doesn't block binding of *ta(a)n* by a superordinate 3rd-person subject, like *Raman*. Thus, (14) appears to be just as grammatical as the minimally varying (15) where the medial subject *Anand* matches in ϕ-feature value (= 3MSG) with the matrix subject, and the antecedent, *Raman*. (16) also varies minimally from (14). Here, it is the mood of the medial clause in each that is different: (14) has an indicative medial clause whereas (16) has a subjunctive one. Here also, we see that both sentences are equally grammatical, showing that blocking is not induced by the nature of the clausal mood either.

(14) Raman$_i$ [$_{CP}$ [$_{CP}$ Krishnan tan-akkŭ$_i$ nuurŭ ruubaaj
Raman[NOM] Krishnan[NOM] ANAPH-DAT hundred rupees
kuḍu-tt-aan-nnŭ] nii nambu-gir-aaj-nnŭ]
give-PST-3MSG-COMP you[NOM] believe-PRS-2SG-COMP
so-nn-aan.
say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that you believe [$_{CP}$ that Krishnan gave him$_i$ a hundred rupees]]."

(15) Raman$_i$ [$_{CP}$ [$_{CP}$ Krishnan tan-akkŭ$_i$ nuurŭ ruubaaj
Raman[NOM] Krishnan[NOM] ANAPH-DAT hundred rupees
kuḍu-tt-aan-nnŭ] Anand nambu-gir-aan-ŭnnŭ]
give-PST-3MSG-COMP you[NOM] believe-PRS-2SG-COMP
so-nn-aan.
say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that Anand believes [$_{CP}$ that Krishnan gave him$_i$ a hundred rupees]]."

(16) Raman$_i$ [$_{CP}$ [$_{CP}$ Krishnan tan-akkŭ$_i$ nuurŭ ruubaaj
Raman[NOM] Krishnan[NOM] ANAPH-DAT hundred rupees
kuḍu-tt-aan-nnŭ] nii namba-laam-ŭnnŭ] so-nn-aan.
give-PST-3MSG-COMP you[NOM] believe-SBJV-COMP say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that you might believe [$_{CP}$ that Krishnan gave him$_i$ a hundred rupees]]."

Most of the native speaker informants who took my online survey, with the exception of a systematic dialect-group of speakers who seem to reject finite (i.e. fully agreeing) embedded clauses across the board (see also Annamalai 1999, for some discussion of this pattern), agree on the major empirical patterns outlined here. As such, the empirical properties given

here for *ta(a)n* may be taken to be largely uncontroversial.

The key facets of the data described here raise interesting challenges, many non-trivial, and resist a straightforward syntactic treatment. The non-locality of long-distance binding, which can be treated as its definitional property, raises the first problem: in generative frameworks, such as Minimalism, where locality is assumed to be a central constraint on syntactic operations, it is unclear how such a property is to be syntactically derived. The apparent lack of Relativized Minimality effects is another issue: briefly, why doesn't a potential antecedent in a medial clause act as an intervener for binding by a higher antecedent? Related to this is the optionality issue governing anaphoric antecedence: given that there is more than one potential antecedent in a multiply embedded clause, what is the principle by which one is optionally chosen over another? And how can this be achieved within a purely deterministic syntactic system? A comprehensive theory of binding must be able to answer these questions as well as others, such as the apparent subject-orientation and the ban on clausemate subject antecedence on the part of *ta(a)n*.

2.3 Syntactic vs. conceptual accounts of long-distance binding

As already noted, the cluster of properties described above is not unique to Tamil. Similar characteristics have been observed in binding patterns in other languages, like Icelandic, Dutch, Italian and Japanese (citations as above). The non-locality of long-distance bound anaphors, in particular, has always posed a bit of a problem for the theory. The Binding Conditions developed in the GB binding theory of Chomsky (1981) state that:

> A. An anaphor is bound in its governing category.
> B. A pronominal is free in its governing category.
> C. An R-expression is free.
> – where X is a **governing category** for Y iff X is the minimal category containing Y, a governor of Y, and an accessible SUBJECT for Y.[7]

[7] Chomsky (1981, p. 250) defines "government" as follows:

(1) Consider the structure (i):
  i. $[_{\beta} \ldots \gamma \ldots \alpha \ldots \gamma \ldots]$, where
    a. $\alpha = X^0$ or is coindexed with γ

Under this strictly tripartite conception of DPs, it is immediately clear that long-distance anaphors cannot be straightforwardly accommodated. The reason is that, under the above formulation, such anaphors should count as pronominal - since they are technically free within their governing categories. At the same time, it is clear that they are unlike standard/deictic pronominals in that, except for very specialized "*logophoric*" contexts, they cannot point to entities in the larger discourse.[8] As Safir (2003) points out, the ability to be used for deixis is at the core of what makes something pronominal (vs. anaphoric): thus, one could ostensively point to an individual in a given context with the expression "her!" but never "herself!". With respect to these semantic properties, long-distance bound anaphors seem to pattern together with locally bound ones in the sense that they cannot refer deictically.

The theories that have been proposed over the years to deal with these patterns, and the non-trivial challenges they pose, can be classed into two broad camps. One of these is the syntactic camp, which treats the long-distance binding facts discussed here as being representative of a primarily syntactic phenomenon and proposes various structural rules to derive them. The other side is comprised of those who believe that long-distance binding encodes a relationship that is primarily semantic or discourse-pragmatic in nature; the proponents of this thesis thus attempt to analyze the binding patterns in terms of non-structural conceptual rules.

Regardless of how the differences in these approaches are internally classified, though, none of them, as far as I know, provides a comprehensive and unified analysis of *all* the descriptive properties of long-distance anaphors. The jury is still out, therefore, as to what the correct analysis of long-distance binding patterns should be. I will argue that it should involve a principled synthesis of the two types of approach.

---

b. where ϕ is a maximal projection, if ϕ dominates γ then ϕ dominates α

c. α c-commands γ

In this case, α *governs* γ.

[8]Here, and elsewhere in the dissertation, my use of the term "logophoric" will be purely descriptive. It will be used merely as a descriptive label, in other words, to characterize a phenomenon where an anaphoric pro-form refers to an entity in the salient discourse whose perspective is linguistically represented. I explicitly do not mean to suggest that logophoricity is a primitive phenomenon, one that is underlyingly different from other sorts of anaphoric dependency. In fact, a central goal of this dissertation will be to argue against such a view.

2.3.1 The syntactic camp

Two distinct schools of thought can be distinguished within the syntactic camp. The first was the movement approach which proposed that long-distance binding dependencies be reconstrued as a series of local ones, with locality being defined in the standard sense (a governing category, within the GB framework of Chomsky (1981) and others, and a phase in the current Minimalist system (Chomsky 2001)). The second is based on the notion that locality itself is relativizable, both across languages and according to different types of pro-form within a given language. One instantiation of this idea was the claim that long-distance anaphors had a larger binding domain than did locally bound ones (and deictic pronouns); under such a view, locality was not seen to be violated in structures involving long-distance binding after all, because the anaphor in such structures was held to still be bound within its own governing category/locality domain. In the GB era, both approaches were quite popular due to the existence of different types of empirical support for each, which, in turn, made it difficult to decide between them.

2.3.1.1 The movement approach

This was the idea, versions of which were proposed in Chomsky (1986a), Pica (1987), Huang and Tang (1991) and others, that a long-distance binding dependency was the result of a series of local movements on the part of the anaphor from its base position to a position that was local to its antecedent. The I-to-I movement analysis was a popular instantiation of this hypothesis. Pica (1987), for instance, proposed that a long-distance bound anaphor moves out of its VP base position at LF and ends up on an Infl head where it is in a local configuration with a subject DP which provides it with requisite features that it itself lacks: this subject is construed as the antecedent of the anaphor. Pica likened the head movement of the anaphor to affix-hopping or clitic climbing, a comparison that seemed to be empirically supported by the monomorphemic status of many long-distance anaphors crosslinguistically. Even more than that, the very monomorphemicity of such anaphors was held to be at the root of their ability to undergo such head movement; in contrast, Pica argued, locally bound anaphors tended to be morphophonologically complex, were thus phrasal, and couldn't undergo head movement in this manner.

The idea is an attractive one. First, it offered a potential explanation for what made long-distance bound anaphors be long-distance bound

(and locally bound anaphors be locally bound); in doing so, it provided an answer for what had hitherto been one of the most puzzling aspects of long-distance anaphora, namely how and why long-distance anaphors were able to be bound outside their governing categories in apparent violation of Binding Condition A. In addition to that, it also derived the subject-orientation of so many long-distance anaphors: the only DP that was both local to the anaphor and c-commanded it in its putative landing site at Infl was the specifier of that head – namely, the syntactic subject of that IP. Ergo, only a subject DP could serve as an antecedent for the anaphor. There was a potentially reasonable explanation for what caused the anaphor to move in the first place, as well: such movement was held to be triggered by the (putative) ϕ-feature defectiveness of the anaphor. Such a defectiveness was supposedly problematic for the anaphor because of the idea, going back to Bouchard (1984), that all NPs/DPs needed a full specification of ϕ-features in order to be interpretable. As such, the anaphor needed to get its ϕ-features valued, or it would crash at the interfaces.

This is the fundamental idea behind the movement approach to binding, but variations to this basic thesis have since been proposed. For instance, Huang and Tang (1991) offer a modified version of the movement thesis to account for perceived blocking effects in long-distance binding structures in Mandarin Chinese. This is the observation that, although the anaphor *ziji* itself doesn't seem to place ϕ-featural restrictions on its antecedence, once its antecedent is determined in a certain sentence, all intervening potential antecedents (subjects) must match its ϕ-features. Such a condition, they argue, must be the result of intervention effects which, they propose, can be derived if the anaphor is assumed to move cyclically from one Infl to another, crucially matching ϕ-features along the way, as Pica also proposed. But they reject the idea, crucial to the Pica-style analysis, that the movement of the anaphor is a version of head movement, in favor of an approach involving IP adjunction due to QR of the anaphor. Nevertheless, the core idea is the same, namely that long-distance binding is the result of covert anaphoric movement at LF, triggered by feature-defectiveness on the part of this anaphor, to an Infl position where it is in a local configuration with its subject antecedent.

In their influential paper on reflexivity, Reinhart and Reuland (1993) also assume a version of the movement analysis to capture the complementarity between long-distance bound anaphors and deictic pronouns, such as the fact that the former, but not the latter, may function as ECM/small-clause subjects in Dutch. To explain such facts, the authors propose the following structural wellformedness rule which regulates the

distribution of long-distance bound anaphors and deictic pronouns:

> **Chain Condition** (Reinhart and Reuland 1993):
> A maximal A-chain ($\alpha_1 \ldots \alpha_n$): has exactly one link, α_1, which is both [+R] and marked for structural Case AND exactly one θ-marked link (i.e. the θ-criterion on A-chains).
> An NP is [+R] iff it carries a full specification for ϕ-features and structural Case. Anaphors are specified [-R]; deictic pronouns are [+R].

Under this rule, having a deictic pronoun serve as the tail of an A-Chain (as in ECM subject position) would violate the Chain Condition – since there would then be two links in the Chain (the pronoun itself, and the matrix subject) which are marked [+R]. However, a long-distance anaphor would be licit because it is considered to be marked [-R].

Although the Reinhart and Reuland (1993) binding thesis differs significantly from other theories of binding during that time, including the various flavors of movement analysis, the Chain Condition is reminiscent of aspects of the movement approach. First of all, a long-distance binding dependency is interpreted as a *structural* dependency between the antecedent and the anaphor. Furthermore, while this dependency is not assumed to be derived by movement, it is still formalized in terms of an A-chain and thus ultimately closely resembles the output of A-movement. Finally, this chain dependency is only assumed to hold between long-distance bound anaphors and their antecedents – local binding is implemented differently.[9] Details such as whether the chain is the result of cyclic movement or a fell-swoop movement are not discussed here since the authors are more interested in wellformedness conditions pertaining to the final representation of the chain rather than in the manner of its derivation – but the basic similarities remain, all the same.

The movement approach, in all its guises, has been an influential one and continues to inform current analyses of long-distance binding. Within the Minimalist tradition, some form of (cyclic) Agree between the anaphor and its antecedent has replaced covert movement of the anaphor and the phase sets locality domains in lieu of the GB-era governing category. But despite these changes, the fundamental idea that binding is ultimately the result of a single local dependency or a series of local structural dependencies (in the long-distance case) between antecedent and

[9]The distribution of locally-bound anaphors is held to be a function of wellformedness conditions pertaining to argument-structural factors. The details of this are captured in Reinhart and Reuland (1993)'s revised Binding Conditions A and B and don't concern us here.

anaphor is preserved. In fact, almost all the major approaches to binding within the Minimalist tradition (Heinat 2008, Kratzer 2009, Hicks 2009, Reuland 2011, Rooryck and vanden Wyngaerd 2011) assume some version of this idea (though many of these deal, primarily, with local binding relations) – although there is considerable internal disagreement as to exactly which features are defective, whether the Agree operation is upward or downward, and other technical details.

2.3.1.2 Relativized "SUBJECT" hypothesis

Despite its success in explaining certain hitherto problematic aspects of long-distance binding, the movement analysis was observed to be inadequate in explaining the full range of empirical patterns associated with this phenomenon. In certain other cases, it seemed to even make the wrong empirical predictions. Another influential strand of analysis, also within the syntactic tradition, proposed a non-movement approach to long-distance binding as a way to deal with some of these problems (Manzini and Wexler 1987, Progovac 1993).

The original Binding Condition A, due to Chomsky (1981), concerned exclusively with the distribution of local anaphors, stated that an anaphor must be bound within its governing category. The governing category, we have seen, was defined as the smallest maximal projection containing the anaphor, a governor for the anaphor, and an accessible SUBJECT. Possible candidates for accessible SUBJECT were: [NP, IP] (the clausal subject in [Spec, IP]), [NP, NP] (the specifier of another NP – as in, a possessor NP), or Agr. It was clear that long-distance anaphors, precisely by virtue of being long-distance bound, violated Binding Condition A (as per its standard definition). The core thesis of the relativized SUBJECT approach was that the distribution of long-distance anaphors could be accommodated within Binding Condition A after all, as long as the choice of accessible SUBJECT was relativized. In brief, the idea was that binding domains are relativized, both across languages and across different classes of anaphor within a given language. What was kept constant, across both these domains, was the nature of the binding relation between the anaphor and its antecedent. Koster and Reuland (1991b, 2) present clear and concise descriptions of these two main tenets, which are reproduced below:

> **b** is a governing category for **a** if and only if **b** is the minimal category containing **a**, a governor of **a**, and F (F an opacity factor).
>
> **a** binds **b** iff **a** c-commands **b** and **a** and **b** are coindexed.

> **F** may assume values such as (accessible) SUBJECT, **Tense, Agr,** and **Comp**. These opacity factors are taken from a universal set, with particular anaphors differing in the value selected ... Similarly, languages may differ in the opacity factors they make available. Anaphors with an opacity factor beyond the SUBJECT are classified as long-distance anaphors.

For instance, Progovac (1993), points out that the traditional choices of accessible SUBJECT given above don't form a natural class since they contain a combination of phrasal categories and heads, observing that movement operations in the syntax are relativized according to the categorial status of the moved element: heads move to other heads, phrases to other phrases (Chomsky 1986a); similarly, rules of government (essentially minimality, in the Minimalist sense) also seem to be relativized according to the category of the elements in question, so that heads govern other heads, and phrases other phrases. Extending this notion into the binding domain, Progovac proposes that long-distance and local anaphors should each have different accessible SUBJECTS on the basis of the type of category each anaphoric class belongs to, claiming that:

> A reflexive R must be bound in the domain D containing R, a governor for R, and a SUBJECT.
> If R is an X^0 (monomorphemic) reflexive, then its SUBJECTS are X^0 categories only, that is, Agr (as the only salient (c-commanding) head).
> If R is an X^{max} (morphologically complex) reflexive, its SUBJECTS are X^{max} specifiers, therefore [NP, IP] and [NP, NP] (Progovac 1993, 757).

The significance of this revision lies in the observation that long-distance bound anaphors tend to be monomorphemic, whereas locally bound ones tend to be morphologically complex (Pica 1987). Given the relativized conditions above, this means that the former take Agr as their only possible SUBJECT whereas the latter take possessor NPs and clausal subjects as SUBJECT. As such, only the former type of anaphor may be long-distance bound – across specifiers and other maximal categories – which is the desired result. Interestingly, this restriction also derives the subject-orientation of long-distance anaphors. Since such an anaphor is in a local configuration with its SUBJECT (i.e. Agr), whose own features are "anaphoric" on those of the subject DP (Borer 1989), Progovac proposes that it ends up referring to this subject "by coindexation transitivity" (Progovac 1993, 770).

In a deeper sense, the relativized SUBJECT approach isn't that different from a movement analysis – in both cases, the anaphor ends up being in a syntactic dependency relationship with an Agr/Infl head; subject-orientation and long-distance binding are both derived as a function of this dependency. But the fact that the relativized SUBJECT approach doesn't derive the anaphor-Infl dependency as a result of anaphoric movement has some definite empirical advantages. The first is the fact that the non-movement analysis alone can derive structures involving binding into islands – such as in the Chinese sentence below (originally from Huang and Tang (1991), formatting mine) – similar effects are found in long-distance binding structures in other languages, Malayalam (Jayaseelan 1997) and Tamil among them:

(9) Zhangsan$_i$ bu xihuan [$_{CP}$ neixie piping ziji$_{i,j}$ de ren$_j$].
Zhangsan not like those criticize ANAPH REL person
"Zhangsan$_i$ does not like [those people]$_j$ who criticize self$_{i,j}$."

As Progovac points out, subjacency is respected elsewhere in Chinese – for instance, island violations are not tolerated in sentences involving standard A-movement. This type of data offers perhaps the most serious challenge to a movement approach – it is simply not clear how the grammaticality of (9) is to be accommodated under such an analysis.

Other empirical challenges to the movement approach exist as well, many of them well documented (Koster and Reuland 1991b, Progovac 1993, Huang 2000, Büring 2005, Giorgi 2006). For instance, it is unclear how the anaphor is able to skip intervening heads like v and (species of) C and move directly from one Infl to another. This is especially odd considering, as Progovac points out, that clitic movement is crosslinguistically quite local and limited. Since, in general, phrases as well as heads are capable of movement, it is also unclear what prevents XPs (which locally bound morphologically complex anaphors are taken to be) from also moving up to Infl – not via head-movement, naturally, but via phrasal adjunction – and being similarly long-distance bound.

It seems clear that the movement approach, despite its strengths, does suffer from some non-trivial empirical issues. But the relativized SUBJECT hypothesis has its problems, as well. A fundamental issue has to do with its central assumption that there are two types of governing category within any given language: one for long-distance bound anaphors and a separate one for locally bound ones – in order to be able to subsume both under Binding Condition A. Progovac proposes that this is not a stipulation but an epiphenomenon of the fact that locally-bound reflexives tend to be complex (which Progovac assumes means that these are

elements of category XP) whereas long-distance bound ones tend to be simplex (which is assumed to mean that such anaphors are of type X), but the fact is that morphologically complex anaphors may be bound long-distance (as in English *himself*, for instance); conversely, simplex anaphors may be locally bound, in certain psych-predicate constructions in Tamil, as we will see. The relativized SUBJECT approach to binding seems to predict a strict universal complementarity between the distributions of local and long-distance bound anaphors – one that is simply not supported by the world's languages.

To be fair, non-complementarity effects such as these are problematic for the different flavors of movement approach as well (with the possible exception of the analysis of Reinhart and Reuland (1993) who, by taking argument-structural factors into consideration, are able to explain many of these effects), since I-to-I movement on the part of the long-distance bound anaphor is supposed to be motivated by its supposed clitic-like status. Finally, it is a bit harder to see how the relativized SUBJECT analysis might be transposed to a Minimalist model of the grammar, given that a phase – the Minimalist equivalent of the GB-era governing category – isn't typically thought to be relativizable in this manner.[10]

There is a whole array of data which challenges the fundamental premise, central to both movement- and relativized SUBJECT approaches, that long-distance binding is encoded in the syntax. This in turn triggered a completely different approach to binding – one which claimed that binding dependencies were regulated extra-syntactically and, in many cases, outside the confines of grammar altogether.

2.3.2 The conceptual camp

The empirical issues motivating the conceptual approach to binding may be thematically classified into three broad groups:

(i) Special binding properties in sentences involving psych predicates, showing that thematic considerations play a role in binding relations.

(ii) Binding structures showing sensitivity to the sentience of the antecedent and, in this context, a binding construction involving an

[10]Though see Hicks (2009) for a different type of phase-relativization: Hicks distinguishes between phases that are relevant for LF and those that are relevant for PF and derives distributional distinctions and overlap between deictic pronouns and anaphors as a function of mismatch between the two (putative) phase-types.

anomalous structural relationship between the anaphor and its antecedent, termed "sub-command" by Huang and Tang (1991).

(iii) Structures where the anaphor seems to be bound by an extra-sentential antecedent, a phenomenon termed "logophora", after Clements (1975).

Let us consider these in turn.

2.3.2.1 Psych-predicates and binding: argument-structural considerations

It has been observed that binding behaves quite differently in the context of psych-predicates (Beletti and Rizzi 1988, Reinhart and Reuland 1993). First, a simplex anaphor, though typically long-distance bound, may be bound locally if it is the co-argument of a psych-predicate (Reinhart and Reuland 1993) – a property that has been observed crosslinguistically. Second, in structures involving certain psych-predicates, the phrase containing the anaphor seems to c-command the antecedent rather than the other way around – a phenomenon termed "backward binding" (Minkoff 2003, Giorgi 2006, among others). Instances of both phenomena are given directly below:

(10) LOCAL BINDING OF SIMPLEX ANAPHOR:

a. Jan$_i$ gedragt zich$_{\{i,*j\}}$/*zichzelf.
Jan behaved ANAPH/*ANAPH-SELF
"Jan$_i$ behaved himself$_i$." (**Dutch**, Reinhart and Reuland (1993))

(11) BACKWARD BINDING:

a. [$_{CP}$ Taan$_{\{i,*j\}}$ waliya aaḷ aana enna] Raman-a$_i$
ANAPH[NOM] great man is that Raman-DAT
tooṉṉi.
occurred/seemed
"It seemed to Raman$_i$ [$_{CP}$ that he$_{\{i,*j\}}$ was a great man.]" (**Malayalam**, Jayaseelan (1997))

b. La-propria moglie preoccupa molto Gianni.
self's wife worries a lot Gianni
"Gianna is worried by self's wife." (**Italian**, (Giorgi 2006))

c. [$_{CP}$ Yosiko ga zibun$_i$ o nikundeiru koto] ga Mitiko$_i$
Yosiko SBJ ANAPH$_i$ OBJ be hating COMP] SBJ Mitiko$_i$
o zetuboo e oiyatta.
OBJ desperation to drove

"That Yosiko hated her$_i$ drove Mitiko$_i$ to desperation." (**Japanese**, Sells (1987))

d. That slanderous article about herself$_i$ tipped Sue$_i$ over the edge. (Minkoff 2003)

Neither the movement analysis nor the relativized SUBJECT analysis is equipped to handle these types of structures. The fact that it is possible to locally bind a monomorphemic anaphor, like Dutch *zich* in (10a), is truly puzzling under these approaches. Both types of analysis predict such anaphors to be long-distance bound and *only* long-distance bound due to their structural dependency on a superordinate Infl. The possibility of being bound by a superordinate oblique DP, as in the Malayalam sentence in (11a), is also unexpected under these approaches since they are precisely tailored to rule *out* this possibility. One could, of course, claim that the surface objects of psych-predicates were (c-commanding) subjects at some point in the derivation (in D-structure, if framed in GB terms) and that anaphors may target D-structure or S-structure antecedents – as has been proposed by Beletti and Rizzi (1988). But even this isn't entirely satisfactory – for instance, how is one to reconcile the idea that binding is an LF-phenomenon (as assumed by both movement and relativized SUBJECT approaches) with the idea that the anaphor may target its D-structure antecedent? And under a predominantly derivational rather than representational system like Minimalism, how is a D-structure vs. S-structure difference in subjecthood to be formally encoded? Finally, what is the relevance of all this to the thematic status (psych vs. non-psych) of the predicate?

Needless to say, the backward binding phenomenon illustrated in (11) poses a genuine problem for the syntactic camp. The movement and the relativized SUBJECT approaches both assume that the anaphor must be c-commanded by its antecedent: the former ensures this by means of upward movement on the part of the anaphor to a superordinate Infl directly below its actual antecedent and the latter by assuming, as per the traditional definition of the term (Chomsky 1981), that the accessible SUBJECT for an element X must be the head of the antecedent of X in [Spec, IP]. The structures given above violate three principles that were presumed to characterize long-distance binding within the syntactic camp: (10) violates the ban on clausemate subject antecedence for simplex anaphors, discussed earlier, and (11) the idea that long-distance bound anaphors are always and only subject-oriented (for certain definitions of "subject"), and that anaphors are always c-commanded by their antecedents. Finally, these analyses, by virtue of being impervious to

the role of argument-structure in binding, also cannot straightforwardly explain the relevance of the psych-predicate status of the verbs in these sentences.

2.3.2.2 Sentience restrictions and sub-command

Anaphors in many languages cannot take a non-sentient antecedent. This is true even if the relevant entity is a syntactic subject. The minimal pair below illustrates this phenomenon for Chinese (Huang and Tang 1991, formatting mine):

(12) Wo bu xiaoxin dapo-le ziji de yanjing.
I not careful break-ASP ANAPH POSS glasses
"Not being careful, I broke my own glasses."

(13) * Yanjing$_i$ diao-dao dishang dapo-le ziji$_{\{i,*j\}}$.
glasses drop-to floor break-ASP ANAPH
"[The glasses]$_i$ dropped to the floor and broke themselves$_{\{i,*j\}}$."
(Intended)

Notice that neither of the structural approaches to binding discussed so far, namely the movement and relativized SUBJECT analyses, is equipped, as it stands, to deal with such a restriction. Within the movement hypothesis, covert LF movement of the anaphor is held to be triggered by its need for ϕ-features: thus, as long as a c-commanding DP in a local configuration with the anaphor is able to provide the anaphor with these features, it should be able to be count as an antecedent. The Chinese anaphor *ziji* may take 1st, 2nd, as well as 3rd-person antecedents (Huang and Tang 1991): the movement approach thus predicts sentences like (13) to be grammatical, contrary to fact. The relativized SUBJECT approach suffers the same fate since it also assumes that ϕ-feature defectiveness is at the heart of anaphoricity, and inanimates have ϕ-features, too.

Further complicating matters is the phenomenon termed "sub-command" by Huang and Tang (1991). The authors of this work noticed that, in some sentential contexts, the Chinese anaphor *ziji* may be anteceded by a DP that is contained within the subject DP – i.e. by a non-c-commanding DP – but that this is sensitive to the animacy of the containing DP:

(14) Wo de jiaoao hai-le ziji.
I 's pride hurt-ASP ANAPH
"[My$_i$ pride]$_j$ hurt self$_{i/*j}$."

(15) Wo de meimei hai-le ziji.
I 's sister hurt-ASP ANAPH
"[My$_i$ sister]$_j$ hurt self$_{j/*i}$."

The descriptive generalization based on such patterns appears to be the following: a (possessor) DP contained inside a subject DP may bind an anaphor just in case the subject DP is inanimate (thus itself disqualified from antecedence). Again, this type of effect is not restricted to Chinese. Giorgi (2006) shows that the same effect obtains in Italian with this structure and Jayaseelan reports similar effects for Malayalam *ta(a)n.* Huang and Tang (1991) and others attempt to derive such phenomena by means of a structural restriction which they label "sub-command", defined as below (Huang and Tang 1991, 266):

> "β sub-commands α iff β is contained in an NP that c-commands α or that sub-commands α, and any argument containing β is in subject position.
> The condition under which the c-command requirement may be relaxed is stated [below]: A reflexive α may take an NP β as its binder if
>
> a. β sub-commands α, and
> b. there is no NP γ, γ a potential binder for α, such that γ is closer to α than β is."

Note that the definition above doesn't itself capture the animacy restrictions on sub-command. This is separately hardwired into the definition of "potential binder": specifically, the authors assume that only animate NPs/DPs may qualify as potential binders. But this is, of course, stipulative: i.e. it is unclear why such a restriction should hold in the first place.

2.3.2.3 The logophora problem

The term "logophor" was originally coined by Clements (1975) to denote certain designated pro-forms, originally observed in African languages, which refer to entities "whose speech, thoughts, feelings, or general state of consciousness are reported" (Clements 1975, 141). An example containing such a pro-form is given below from Tuburi, a language spoken in parts of Chad (Sells 1987, 447); the plural logophor *sā:rā*, marked in boldface below, represents the mental perspective of the sayer, the matrix subject "they":

(16) à (ríng) wò gā tí **sā:rā** tʃÍ **sā:rā**
pro (say) PL COMP head LOG hurt LOG
"They$_i$ said [$_{CP}$ that they$_i$ had headaches]."

Interestingly enough, it was observed that long-distance anaphors in many languages could refer in this manner, as well, particularly in sentences modelled in the so-called "free indirect discourse" style, a narrative structure that involves a mixture of direct and indirect speech, made from the perspective of a 3rd-person narrator (Banfield 1982, Schlenker 2004). Here are some illustrative examples of logophoric reference within this narrative style from English (Austen 1816, Chapter XVIII, 321), Icelandic (Sells 1987), and Malayalam (Jayaseelan 1997). In all these examples, the anaphoric antecedent is extra-sentential; in (17), the antecedent (the novel's protagonist, Emma, from whose perspective the narrative is reported) isn't even asserted in the immediately surrounding discourse:[11]

(17) "With Tuesday came the agreeable prospect of seeing him again, and for a longer time than hitherto; of judging of his general manners, and by inference, of the meaning of his manners towards **herself**; of guessing how soon it might be necessary for her to throw coldness into her air . . . " .

(18) Formaðurinn$_1$ varð óskaplega reiður. Tillagan væri
the chairman became furiously angry. the proposal was.SBJV
avívirðileg. Væri henni beint gegn sér$_1$ persónulega?
outrageous. was.SBJV it aimed against ANAPH personally?
"The chairman became furiously angry. The proposal was outrageous. Was it aimed at him personally?"

(19) Aarum sahaayik'k'illa enna Raaman-a manassilaayi taan
No one will.not.help that Raman-DAT realized. ANAPH[NOM]
ini ottak'k'a aana. Tanta bhaarya poolum
from now on alone is. ANAPH-GEN wife even
tan-ne upeeksikkum.
ANAPH-ACC will abandon

[11] Hicks (2009, 136) additionally presents examples from the Southern Hiberno dialect of English, spoken in parts of Ireland, where the anaphor can be the entire subject, as in the sentence below:

i. Did himself go out last night?

Tom McFadden (p.c.) informs me that potentially similar usage is also attested in versions of American English, where the anaphor is used in special circumstances as a pro-form indicating intimacy between two people. For instance, a husband might address a gift to his wife as being "From himself to herself". Given that the anaphoric forms here seem to be used quite deictically, however, it doesn't seem to me that these are instances of true logophora, though Hicks (2009) does treat the sentence in (i) as such. More research must be undertaken to ascertain the proper nature of such pro-forms, but their existence is, nevertheless, interesting.

> "Raman$_i$ realized that no-one would help. Self$_i$ was alone from now on. Even [self's$_i$ wife]$_j$ would abandon self$_{i/*j}$"

This sort of "logophoric" reference on the part of an element which looks, for all intents and purposes, just like an anaphor constitutes another major challenge for a syntactic treatment of long-distance binding. Simply put, it is unclear how a syntactic dependency is to be made sensitive to an entity that is simply not overtly asserted in the syntactic structure at all.

A common analytic strategy for dealing with such sentences has been simply to claim that such elements, despite looking like anaphors on the surface, are underlyingly a different type of animal altogether. They are not anaphors at all, but "logophors" – elements whose reference is derived by means of extra-linguistic factors pertaining to discourse salience, narrative structure, and language use (Hellan 1991, Kuno 1987). Therefore, according to this strategy, they do not actually pose a challenge to syntactic treatments of long-distance anaphora at all.

Such, indeed, is the approach of Reinhart and Reuland (1993) to sentences like (19) above. Indeed, the more radical move has been to claim that even anaphors *with* linguistic antecedents whose distribution doesn't seem to be obviously structurally regulated are instances of logophora rather than anaphora. This, for instance, is Minkoff (2003)'s analysis of backward binding structures in English, like (11d). Logophoricity has thus become a bit of a catch-all label for anaphoric phenomena which cannot be explained by standard rules of binding. As Büring (2005, 72) puts it, "logophoric approaches are only as restrictive as their underlying theory of logophoricity, an area where more work is required." However, some analyses of logophoricity do stand out for their level of detail and rigor, Sells (1987), Kuno (1987), Koopman and Sportiche (1989) notable among them. Below, I briefly review Sells' and Kuno's analyses of logophoric and long-distance binding dependencies – based, in Sells' case, on converging empirical patterns from a wide array of languages and in Kuno's on data taken predominantly from Japanese and English.

Sells (1987), drawing on evidence from a variety of discourse logophoric phenomena in African languages like Mundang, Tiburi and Gokana as well as the more familiar long-distance binding structures in Japanese and Icelandic, proposes that there is no monolithic notion of logophoricity. Rather, he argues, the concepts of logophoricity and long-distance binding can be decomposed into three primitive notions which he terms, "source", "self", and "pivot", and describes as follows (Sells 1987, 457):

> SOURCE: one who is the intentional agent of the communication.
> SELF: one whose mental state or attitude the content of the proposition describes.
> PIVOT: one with respect to whose space-time location the content of the proposition is evaluated.

The core idea of Sells' proposal is that a logophoric pronoun "will link to some NP *in virtue of the fact that it is associated with a particular role*" (Sells 1987, p. 456, italics his). In other words, discourse logophoric as well as long-distance binding antecedents are claimed to instantiate one or other of these three primitive roles. The choice of which role is instantiated by an antecedent DP is, to some extent, parametrized, Sells claims. But it is also a function of the core predicate involved in the binding relationship. Thus, psych-predicates tend to take arguments with a SELF role whereas the arguments of propositional speech-predicates tend to be associated with a SOURCE role. Furthermore, Sells conjectures that there may be an implicational hierarchy between the three roles with SOURCE being the most specific class subsuming both SELF and PIVOT, as follows:

SOURCE » SELF » PIVOT

To demonstrate a concrete implementation of Sells' idea, consider the following psych-predicate structure in Japanese (Sells 1987, p. 454, ex. 29) (formatting mine):

(20) [Yosiko ga zibun$_i$ o nikundeiru koto] ga mitiko$_i$ o
Yosiko SBJ [ANAPH$_i$ OBJ be-hating COMP] SBJ mitiko OBJ
zetuboo e oiyatta.
desperation to drove
"[$_{CP}$ [$_{CP}$ That Yosiko hated her$_i$] drove Mitiko$_i$ to desperation.]"

In this sentence, the antecedent *Mitiko* is associated with the SELF role – since the sentence involves a matrix psych-predicate-phrase (*drive to desperation*) which deals with Mitiko's mental state. By virtue of the implicational hierarchy between the different roles, described above, this automatically entails that *Mitiko* is also associated with the PIVOT role. The anaphoric pronoun *zibun* targets this role and in turn links it to the antecedent that is associated with this role, namely *Mitiko*. In this sense, the true antecedents of anaphors are the roles themselves; the binding relation to a c-commanding DP is mediated by these roles and is, thus,

established only indirectly. This same point has been made, albeit in different terms, by Borer (1989) and, more recently, in Kratzer (2009).

A closely related conceptual account is that of Kuno (1987) who proposes that anaphoric dependences in Japanese (and languages like it) are regulated by their sensitivity to a conceptual property that he terms "empathy" and defines as follows:

(21) **Empathy** (Kuno 1987, 206):
"*Empathy:* Empathy is the speaker's identification, which may vary in degree, with a person/thing that participates in the event or state that he describes in a sentence.
Degree of Empathy: The degree of the speaker's empathy with x, $E(x)$, ranges from 0 to 1, with $E(x) = 1$ signifying his total identification with x, and $E(x) = 0$ a total lack of identification.

In other words, the empathy relationship captures the extent to which an event is described from the point-of-view of an individual denoted by a DP in the sentence rather than from that of the speaker of the utterance. Kuno shows that binding phenomena in Japanese may lexically instantiate this choice; thus, the verbs *yaru* and *kureru* in Japanese both mean GIVE, but the former represents the giving event from the perspective of the AGENT (the giver) and the latter denotes it from the perspective of the GOAL (the receiver). These differences in "camera angle" (to invoke an analogy from Kuno (1987)) are also shown to directly correlate with possibilities of anaphoric binding. Specifically, the restriction appears to be that the anaphor must be bound by the DP that denotes the participant with the highest degree of empathy in a given clause. Thus, the sentence in (22) is illicit because it is anteceded by the GOAL *Taro* whereas the empathy is associated with the non-antecedent AGENT *Hanako*, as signalled by the use of the verb *yatta* (from *yaru*). However, when the verb is changed to *kure(ru)*, indicating that the empathy "locus" is now the GOAL *Taro*, then *Taro* may antecede the anaphor , as in (23):

(22) * Taro$_i$-wa [Hanako-ga **zibun**$_i$-ni **yat-**ta] hon-o yon-da.
Taro-TOP Hanako-NOM ANAPH-DAT give-PST book-ACC read-PST
"Taro$_i$ read the book Hanako gave him$_i$." (Intended)

(23) * Taro$_i$-wa [Hanako-ga **zibun**$_i$-ni **kure-**ta] hon-o yon-da.
Taro-TOP Hanako-NOM ANAPH-DAT give-PST book-ACC read-PST
"Taro$_i$ read the book Hanako gave him$_i$."

This is an abbreviated and informal summary of Sells' and Kuno's analyses. While my own analysis of the binding patterns in Tamil, and languages like it, will turn out to be rather different from those of these

works, it will be seen to reflect, to a large degree, the fundamental intuitions of both. I will, specifically, avail myself of Sells' idea that anaphoric as well as discourse logophoric relations are regulated by their sensitivity to the primitive concepts captured by the labels "SOURCE", "SELF", and "PIVOT". Relatedly, I will propose, in line with Kuno, that anaphors, in Tamil and languages with similar binding patterns, is a "perspective seeker" and that a potential antecedent is a "perspective holder" with respect to the minimal predication containing the anaphor. The key difference between my analysis and these will be that mine argues for a syntactic treatment of empathic/perspectival information (captured also by Sells' "roles") within a cartographic model (Cinque 1999, Speas 2004) set within the larger Minimalist framework.

2.3.3 Structural vs. conceptual accounts: how do we decide?

We have seen two types of data so far. There is the "core" set of facts involving long-distance binding structures which seem to obey structural principles of the grammar, such as subject-orientation, c-command, and the ban on clausemate subject antecedence (which might instantiate a type of anti-locality effect). The structural accounts of Chomsky (1981), Pica (1987), Huang and Tang (1991), Reinhart and Reuland (1993), Progovac (1993, among others) discussed here have been motivated by this type of data. The other set of data looks anomalous in comparison: this is the constellation of long-distance binding facts that suggests that the structural conditions perhaps don't matter and that binding dependencies are regulated by thematic, semantic and discourse-pertinent factors instead. The conceptual accounts of Kuno (1987), Sells (1987), and Hellan (1991) were motivated by such data. In other words, both structural and conceptual analyses seem to be motivated by different sorts of empirical evidence. Are the choices mutually exclusive? If so, which is correct, and on what basis can we adjudicate between the two? If not, why not and how do we prove it? There are three logical options. One is that binding conditions are regulated by structural conditions alone, another is that they are conditioned by conceptual factors alone, and the third option is that they are governed by a combination of structural and conceptual mechanisms.

In order for the all-structural option to be correct, it would have to be provable that *all* long-distance binding dependencies, including those that seem to violate core structural principles actually do obey constraints that are formulable in structural terms. In other words, we

should be able to show that logophoric dependencies, backward binding, and sub-command structures actually *do* obey structural constraints, even though they don't appear to do so on the surface. There is already some precedent for such an analytic approach: recent analyses have suggested enriching the syntactic feature system so that certain types of discourse-pertinent information are present in the syntactic structure, particularly in the left periphery (Cinque 1999, Speas and Tenny 2003, Giorgi 2010), and participate in syntax-internal operations. Others like Hale and Keyser (1993) have proposed structural accounts of thematic role relationships which could offer ways to dealing with the specialness of binding under psych predicates.

The all-conceptual analytic option would, on the other hand, require showing that even those binding patterns that do seem to obey structural constraints do so only epiphenomenally; in other words, it would require proving that even these binding dependencies are really motivated by sensitivity to conceptual factors such as mental or physical point-of-view and discourse salience and that the conformity to structural conditions like c-command, subject-orientation and (anti-)locality is incidental to, or perhaps indirectly derived from, this. There is some motivation for such an approach, as well. Both standard long-distance anaphors and the class of pro-form labelled "logophor" are governed by strikingly similar semantico-pragmatic requirements: briefly, the representation of a mental attitude, point-of-view or experience of a salient entity. We will see evidence for this in the discussion to come.

The third logical class of solution would be to suggest that binding dependencies are governed by both structural and conceptual factors. There are two distinct imaginable ways of pursuing this option. The first would be to claim that there are two underlyingly distinct classes of pro-form – anaphors and logophors, with the former being sensitive to structural, and the latter to conceptual, factors. The second way of doing this would be to say that each individual pro-form is sensitive to a combination of structural and conceptual restrictions.

One type of support for the first of these sub-options would come from empirical evidence showing that the structurally "anomalous" and structurally straightforward binding structures form two strictly non-overlapping sets and, as such, don't ever enter into competition with one another. This, in essence, is the strategy that has been widely adopted within the primarily structural camp with respect to logophoricity: the standard approach has been to claim that this phenomenon is fundamentally different from anaphoricity, and involves inherently different sorts of pro-form ("logophors" vs. "anaphors"). A major problem with

this approach, at least with respect to the logophoricity phenomenon, is that distinguishing two underlyingly distinct (i.e. fully non-overlapping) classes of referentially defective pro-form doesn't seem empirically warranted. First, both anaphors and putative "logophors" frequently have identical morphophonological shapes – treating them as separate underlying elements would effectively entail reducing such identity to instances of accidental homophony – a rather dubious move given the crosslinguistic pervasiveness of this identity relation. Second, they both tend to have very similar interpretations in many languages – both referring to the point-of-view of a salient individual. Support for the second sub-option would involve showing that *every* instance of long-distance binding involves a combination of structural and conceptual factors. These factors might be modularized, within a Y-model of the grammar, with the syntactic module taking care of some aspects of the binding relation, and the LF and PF modules regulating the conceptual and morphophonological properties of binding, respectively. On the other hand, certain conceptual properties might be a part of the lexical information that a linguistic entity is born with – and would thus condition where the entity can be merged in the structure and inform the nature of syntactic dependencies in that position.

The correct analytic choice must ultimately be decided on an empirical basis. Nothing about what we have seen so far gives us enough information to do that just yet. In the next chapter, I provide a detailed exposé of the long-distance binding facts with an aim to doing just this for Tamil and for languages with analogous binding properties. I will conclude that what is needed is an account that combines structural and conceptual factors in a systematic way.

Chapter 3

Tamil long-distance binding under the magnifier

We ended our discussion of the core long-distance binding patterns in Section (2.2) with the following empirical observations:

(i) **Subject-orientation:** Tamil *ta(a)n* is bound by syntactic subjects, not objects.

(ii) **Ban on Clausemate Subject Antecedence:** *ta(a)n* may not be bound by its own clausemate subject.

(iii) **Non-locality:** There is no upward bound on the distance between the anaphor and its antecedent, modulo processing.

(iv) **Non-Minimality:** There are no apparent Relativized Minimality effects in the binding relation. An antecedent may bind *ta(a)n* across other potential antecedents.

(v) **Antecedence Optionality:** While *ta(a)n* may take only one antecedent at a time, in a multiply embedded structure, a number of different DPs may be potential antecedents, and the choice among them is non-deterministic.

The observations pertaining to subject-orientation and ban on clausemate antecedence seem to be a function of syntactic constraints. Others, such as non-locality, non-minimality, and antecedence optionality, on the other hand, seem to violate syntactic principles or, at the very least, resist a straightforward syntactic analysis. Here, I zoom in on these properties and inspect them in much greater detail so as to figure out whether the long-distance binding facts of Tamil conform to structural or conceptual principles or a mixture of the two, with the ultimate view

of developing an optimal analytic solution for these patterns. We will see that the optimal solution involves a principled combination of structural and conceptual approaches.

3.1 Subject orientation of *ta(a)n*

Let us start with the subject-orientation of *ta(a)n*. This refers to the fact that in long-distance binding structures in Tamil, *ta(a)n* can refer to a superordinate subject but, crucially, not to a superordinate object. Sentence (24) illustrates this point again: the superordinate *Anand* may not function as an antecedent for *ta(a)n*; only the superordinate subjects *Raman* and *Seetha* may serve this function:

(24) [$_{CP}$ [$_{CP}$ Raman$_i$ Anand-kiṭṭæ$_j$ [$_{CP}$ Krishnan$_k$
Raman[NOM] Anand-ALL Krishnan[NOM]
tann-æ$_{\{i,l,*j,*k,*Auth\}}$ kaappaatt-in-aan-nnŭ] so-nn-aan-nnŭ]
ANAPH-ACC save-PST-3MSG-COMP say-PST-3MSG-COMP
Seetha$_l$ nene-tt-aaḷ-ŭnnŭ] naan$_{Auth}$ paar-tt-een.
Seetha[NOM] think-PST-3FSG-COMP I[NOM] see-PST-1SG
"I$_{Auth}$ saw [$_{CP}$ that Seetha$_l$ thought [$_{CP}$ that Raman$_i$ told Anand$_j$ [$_{CP}$ that Krishnan$_k$ saved him$_{\{i,l,*j,*k,*Auth\}}$.]]]"

Under the fairly standard assumption that subjects occupy [Spec, TP] and that indirect objects are located somewhere in the *v*P, perhaps in an ApplP (Pylkkänen 2008), this initially looks like a constraint that is easily formulable in purely structural terms. Indeed, such data did motivate the I-to-I movement (Pica 1987, Huang and Tang 1991) and relativized SUBJECT (Manzini and Wexler 1987, Progovac 1993) analyses, both purely structural in spirit.

These subject antecedents predominantly surface with nominative case marking but, in languages like Icelandic and Tamil with rich case systems, it is not just canonical nominative subjects that are capable of functioning as antecedents for binding. "Quirky" dative nominals with EXPERIENCER θ-roles, such as *Ramanŭkkŭ* in Tamil (25), may also bind the simplex anaphors in these languages:

(25) Raman-ŭkkŭ$_i$ [$_{CP}$ Seetha$_j$ tan-næ$_{\{i,*j\}}$ paar-tt-aaḷ-ŭnnŭ]
Raman-DAT Seetha[NOM] ANAPH-ACC see-PST-3FSG-COMP
tooṇ-itt-adŭ.
occur-PST-3NSG
"It occurred to Raman$_i$ [$_{CP}$ that Seetha$_j$ saw him$_{\{i,*j\}}$]"

Under a narrow view of "subject" as an XP with canonical nominative case, the possibility of sentences like (25) might be taken to undermine the (syntactic) subject-orientation hypothesis. However, experiencer datives such as these are standardly treated as occupying syntactic subject position in Icelandic (see Zaenen et al. 1985) and can also be shown to do so in Tamil.[1] As such, binding dependencies such as that in (25) don't challenge the idea that simplex anaphors in such languages target syntactic subjects. In the following sections, however, I will discuss three sets of facts which do constitute real challenges to this notion.

3.1.1 Challenge I: *non-sentient antecedent

We noted in Section (2.2) that, although *ta(a)n* can only take 3rd person antecedents, the gender feature on its antecedent is irrelevant – masculine, feminine, and neuter entities are all licit. There is, however, a sentience restriction to antecedence, just as we observed for Chinese (13), above. Specifically, non-sentient entities cannot serve as antecedents for *ta(a)n*. This restriction holds even when the antecedent is in syntactic subject position. In (26), below, the syntactic subject *gaḍigaaram* (CLOCK) may thus not serve as an antecedent for *ta(a)n* even though it is the syntactic subject:[2]

[1]Zaenen et al. (1985) use V2 word-order and conjunction reduction facts in Icelandic as partial diagnostics in favor of such a position. These tests don't work for Tamil which is not V2 and also has no clear way to express sentential coordination. A more probative test, also used by the authors, is the ability of a dative argument to be replaced by null PRO in an infinitive. Under the common, and fairly uncontroversial, assumption that controlled PRO always and only occupies syntactic subject position, this in turn shows that a dative experiencer may occupy syntactic subject position in this language. Tamil, we see, allows this possibility, just like Icelandic – (i) shows that *piḍi* (LIKE) takes a dative experiencer; (ii) shows that this same experiencer argument can be expressed as controlled PRO:

i. Raman-ŭkkŭ Krishnan-æ piḍi-tt-adŭ.
Raman-DAT Krishnan-ACC like-PST-3NSG
"Raman liked Krishnan."

ii. Raman$_i$ [$_{CP}$ PRO$_i$ Krishnan-æ piḍikk-æ] mujarcci-sej-d-aan.
Raman[NOM] Krishnan-ACC like-INF attempt-do-PST-3MSG
"Raman$_i$ tried [$_{CP}$ PRO$_i$ to like Krishnan$_j$]"

[2]The sentence given in (26) *would* work in a special discourse context, involving a fairy-tale like "Beauty and the Beast", for instance, where the clock is anthropomorphized and endowed with sentience and volition – but this, of course, is not the default reading of this sentence, and that just proves the point.

(26) * Gaɖigaaram$_i$ taraiy-læ viɻ-ŭndŭ tann-æ$_i$ oɖeччu-ko-ɳɖ-adŭ.
clock[NOM] floor-LOC drop-ASP ANAPH-ACC break-*ko*ɭ-
PST-3NSG
"The clock dropped to the floor and broke itself." (Intended)

Sentences like (26) show that syntactic subjecthood is not a sufficient condition for *ta(a)n*-antecedence, thereby weakening the correlation between *ta(a)n*- antecedence and syntactic subjecthood.

3.1.2 Challenge II: logophoricity

Structures involving the so-called "logophoric" reference of *ta(a)n* also challenge the idea that the antecedent of *ta(a)n* must be a syntactic subject, but they do so in the trivial sense that the antecedent is not overtly represented (as a subject or object) in the sentential structure at all.

ta(a)n cannot, in the typical case, occur as the subject of a matrix/root clause – this is illustrated below:

(27) Avan/*Taan$_i$ Krishnan-æ$_j$ aɖi-tt-aan.
he[NOM]/ANAPH[NOM] Krishnan-ACC hit-PST-3MSG
"He$_i$/*Himself$_i$ hit Krishnan$_j$."

However, given the right set of discourse factors, *ta(a)n* can refer "logophorically" – a possibility we already addressed for other languages like Icelandic and Malayalam, in Section (2.3.2.3). In such cases, it *can* appear as a matrix subject, to a lesser degree as a matrix object and, perhaps to an even lesser degree, as an embedded subject/object:[3]

(28) LOGOPHORIC *ta(a)n* AS A MATRIX NOMINATIVE AND "QUIRKY" DATIVE SUBJECT:

[3]Observations pertaining to this tendential ranking are based on the results of my survey. Further research must be undertaken to ascertain both the empirical robustness and theoretical origins of this tendency. It seems plausible, however, that it has its origins in extra-grammatical factors pertaining to the salience of the antecedent (which, in turn, seems to be sensitive to the overt/covert morphophonological status of the entity). When in matrix object position, the DP in matrix subject position, though not itself a potential antecedent is, in the default case, more salient than any silent entity in the discourse. When *ta(a)n* is an embedded subject or object, there are other potential antecedents in higher clauses which, being overtly expressed in the structure, are even more optimal candidates for *ta(a)n*-antecedence. Neither of these factors deterministically rules out the possibility of logophoric reference altogether, in such cases – but, all else being equal, it does appear to play a role.

a. Raman$_i$ Krishnan-kiṭṭæ$_j$ polamb-i-naan. tan-
Raman[NOM] Krishnan-ALL complain-PST-3MSG. ANAPH-
akkŭ$_{\{i,*j\}}$ een inda vidi-joo.
DAT why this fate=CL?
"Raman$_i$ complained to Krishnan$_j$. Curious why he$_{\{i,*j\}}$ suffered this fate." (Rough translation)

b. Raman-ŭkkŭ$_i$ oṉṉum-ee purija-læ.
Raman-DAT nothing[ACC]-EMPH understand-NEG.
taan$_i$ maṭṭum een ippaḍi ellaam kashṭappaḍa-ṇum?
ANAPH-NOM alone why like.this all suffer-must?
"Raman$_i$ didn't understand at all. Why should he$_{\{i,*j\}}$ alone suffer like this?

(29) *ta(a)n* AS MATRIX OBJECT:

a. Raman$_i$ innikki rombæ sandoosha-pa-ṭṭ-aan. ennaal
Raman-DAT today very happy-feel-PST-3MSG. Because
neettikki Krishnan$_j$ tann-æ$_{\{i,*j\}}$ viɻaa-læ nallaa paaraaṭṭ-
yesterday Krishnan ANAPH-ACC gala-LOC well praise-
in-aan.
PST-3MSG
"Raman$_i$ is very happy today. Because yesterday Krishnan$_j$ praised him$_{\{i,*j\}}$ a lot at the gala."

Recall that the logophoricity phenomenon in languages like Icelandic, Italian, English and the like was a strong motivation for conceptual theories of long-distance binding. Even linguists with a primarily structural approach to binding, such as Reinhart and Reuland (1993) and, more recently, Hicks (2009), Reuland (2011) and Rooryck and vanden Wyngaerd (2011), within the Minimalist tradition, are forced to assume that logophoricity is a function of different, perhaps extra-grammatical, processes, since their syntactic theories can't deal with it. The fact that *ta(a)n* can refer logophorically, as in the sentences given above, is thus a potentially non-trivial blow to a unified structural treatment of all the binding patterns involving *ta(a)n*.

3.1.3 Challenge III: backward-binding

Long-distance binding sentences involving "backward binding" (Minkoff 2003), examples of which have already been illustrated for English and Japanese, in (11), where the antecedent doesn't c-command the anaphor – represent another type of challenge. Examples of this phenomenon in Tamil are given below:

(30) [$_{DP}$ Seetha$_i$ tann-æ$_{\{*i,j\}}$ verŭ-tt-aaḷ enbadŭ] Krishnan-æ$_j$ rombæ kašṭappaḍŭtt-ij-adŭ.
Seetha[NOM] ANAPH-ACC hate-PST-3FSG that[NOM] Krishnan-ACC very bother-PST-3MSG
"[$_{CP}$ That Seetha$_i$ hated him$_{\{*i,j\}}$] bothered Krishnan$_j$ very much."

(31) [$_{CP}$ [$_{DP}$ Taan$_{\{i,*j\}}$ avvaḷavŭ eeṭæ-jaagæ irŭnd-adŭ]
ANAPH[NOM] so poor-ADJ be-PST-3NSG.NOM
Raman-æ$_i$ rombæ-vee baadi-jirŭ-kkir-adŭ.]
Raman-ACC very-EMPH affect-be-PRS-3NSG
"[$_{DP}$ His$_{\{i,*j\}}$ having been so poor] has really affected Raman$_i$ very much."

(32) [$_{CP}$ [$_{DP}$ Taan$_{\{i,*j\}}$ avvaḷavŭ eeṭæ-jaagæ irŭnd-adŭ]
ANAPH[NOM] so poor-ADJ be-PST-3NSG.NOM
[$_{DP}$ Raman-ooḍæ$_i$ uḍal-æ] rombæ-vee baadi-jirŭ-kkir-adŭ.]
Raman-GEN body-ACC very-EMPH affect-be-PRS-3NSG
"[$_{DP}$ His$_{\{i,*j\}}$ having been so poor] has really affected [$_{DP}$ [$_{DP}$ Raman's$_i$] health]."

(33) [$_{CP}$ [$_{DP}$ Taan$_{\{i,j\}}$ avvaḷavŭ eeṭæ-jaaga irŭnd-adŭ] [$_{DP}$
ANAPH[NOM] so poor-ADJ be-PST-3NSG.NOM
avan-ooḍæ$_i$ aṇṇaav-æ]$_j$ rombæ-vee baadi-jirŭ-kkir-adŭ.]
Raman-GEN brother-ACC very-EMPH affect-be-PRS-3NSG
"[$_{DP}$ His$_{\{i,j\}}$ having been so poor] has really affected [$_{DP}$ [$_{DP}$ his$_i$] brother]$_j$."

In (30), *ta(a)n* is the object of the sentential subject; it is bound by *Krishnan* which is the direct object of the sentence and thus doesn't seem to c-command *ta(a)n*. In (31), *ta(a)n* is the subject of the sentential subject – it, too, is bound by the non c-commanding direct object of the sentence (here, *Raman*). In both these sentences, it is clear that the subject-orientation condition is violated, since the antecedents are both direct objects (a fact signalled by the accusative-case marking on these DPs).

What is less clear, however, is whether the c-command relation is also violated. The reason for this uncertainty is that Tamil manifests pervasive scrambling. Thus, although Tamil is basically an SOV language (this is the pragmatically unmarked order, among other things), the relative surface positions of sentential arguments cannot be taken as proof of their underlying structure. However, (32) shows conclusively that the antecedent doesn't c-command the anaphor. In this sentence, not only the anaphor, but also the antecedent, are contained inside an-

other XP: *ta(a)n*, in this sentence, is the subject of the clausal subject and its antecedent *Raman* is the possessor DP inside the matrix direct object. *Raman* clearly doesn't c-command the anaphor from this position. (33) shows that (32) isn't just an instantiation of "sub-command", the effect that Huang and Tang (1991) try to capture in Chinese, as discussed earlier. In (33), the possessive pronoun *avanoo*ɖ*æ* (HIS) is contained within a DP which is itself sentient (HIS BROTHER). However, the pronoun can still bind *ta(a)n* from this position, as can the complex DP it is embedded within.[4]

Sentences like (30)-(33) show that neither syntactic subjecthood nor c-command is a necessary condition for antecedence of *ta(a)n*. Given that the relevant antecedents for *ta(a)n* cannot be readily explained either in terms of the grammatical function (subject vs. object) of the antecedent or in terms of hierarchical relations between the anaphor and its antecedent, a straightforward account of *ta(a)n*-antecedence in purely structural terms seems quite difficult.

3.2 Evidence for a conceptual treatment of perspective

The possibilities of backward binding ((30)-(33)) and logophoricity ((28)-(29)) and the sentience restriction on the antecedence of *ta(a)n* (26) all solidly undermine the idea, central to both the movement- and relativized SUBJECT analyses, that the antecedence of *ta(a)n* is always and only a superordinate entity in syntactic subject position. But what, then, is the relevant factor conditioning antecedence in these structures, if not syntactic subjecthood?

Let us take a closer look at the structures above. Starting with the backward binding examples in (30)-(32), we see that argument-structural factors play a role. Notice that each of these sentences involves a psych-predicate and the antecedent of *ta(a)n* is the EXPERIENCER argument of that predicate. Based on this type of evidence, we might propose the following antecedence condition for *ta(a)n*:

(34) **Condition for potential *ta(a)n*-antecedence (Version 1):** The EXPERIENCER argument of a psych-predicate qualifies as a potential antecedent for *ta(a)n*.

[4] It should be noted that, without the proper discourse environment, it is easier for the direct object (HIS BROTHER) to bind *ta(a)n*, than it is for the pronoun, for reasons which will become clear.

As a cautionary note, bear in mind that this description does not involve conditions on the *actual* antecedence of *ta(a)n*, but merely delineates those conditions that have to hold for a DP to qualify as a *contender* for actual antecedence. This will be an important distinction in the model being developed here. Returning to the details of (34), the first observation is that it looks too specific. One might be tempted to argue that the thematic role of the antecedent is irrelevant since, at least in (31) and (32), the antecedent of *ta(a)n* is, in fact, the only other DP in the sentence. In (30), there is another DP in the sentence that *ta(a)n* could refer to, namely its clausmate subject, *Seetha*. However, one could still claim that *Seetha* is ruled out as a potential antecedent here because of the Ban on Clausemate Subject Antecedence.

A more convincing argument in support of the significance of lexical-conceptual factors for antecedence comes from the last sentence in this set – namely (33). There are two DPs that *ta(a)n* could refer to, in this sentence: one is the possessive pronoun *avanoo*ɖ*æ* (HIS) embedded inside the direct object of the clause; the second is the direct object *avanoo*ɖ*æ a*ɳɳ*aa* (HIS BROTHER) itself. Both DPs are possible antecedents for *ta(a)n*, as the marking of referential indices shows. In Footnote (4) above, I observed that, in the pragmatically unmarked case, it is much easier for the direct object DP to antecede *ta(a)n* than for the possessive pronoun inside this DP to do so. The relevance of this observation to this discussion is as follows. Note that the direct object is the EXPERIENCER of the psych-verb *baadijirŭkkiræd*ŭ (HAS.AFFECTED) in this sentence. The possessive pronoun *avanoo*ɖ*æ* (HIS) inside this DP is not an EXPERIENCER, nor is its mental state obviously represented in the sentence – at least, in the pragmatically unmarked case. However, given the right discourse circumstances, this possessive pronoun *could* be associated with a mental perspective; for instance, if (33) were part of an introspective series of thoughts or assertions from the point of view of the pronominal referent. It is precisely in such cases that *avanoo*ɖ*æ* can bind *taan* in this sentence.

The antecedence possibilities in (33) show us that an EXPERIENCER θ-role might not be the relevant factor for antecedence. After all, *avanoo*ɖ*æ* is not assigned an EXPERIENCER role by the matrix verb (or the possessive semantics of the genitive). However, it *can* be associated with the semantics of mental experience by virtue of information present in the salient discourse. This suggests that what is relevant is not thematic roles (which are typically held to be assigned by the predicate of the sentence to its arguments) but the more general conceptual semantics associated with particular θ-roles.

What regulates antecedence in the sentences illustrating logophoric

reference in (28)-(29? In (28a), the antecedent *Raman* is complaining to *Krishnan* – the proposition containing the anaphor represents the content of his complaint; in (28b), the antecedent is in a bewildered state of mind – the proposition containing the anaphor reflects what he is bewildered by; finally, in (29), the antecedent is an individual who is claimed to be very happy and the sentence containing anaphor tells us what he is so happy about. In all these sentences, therefore, the antecedent is again associated with a semantics of mental experience and the anaphor is in a clause that elaborates on this mental experience.

Given this discussion, the ban on non-sentient antecedents, illustrated by sentences like (26), is perhaps not that surprising, after all. This non-sentience restriction can be readily explained under the assumption that *ta(a)n always and only* takes an individual whose mental state is structurally represented, as its antecedent. Non-sentient subjects, such as *gaḍigaaram* (CLOCK) in (26) are not capable of bearing a mental point-of-view in the first place, due to the trivial fact that they are non-sentient/don't have a mind. In such cases, if there is no other other potential antecedent in the structure, ungrammaticality results, just as in (26).

Building on this discussion, let us propose the following updated descriptive condition for the antecedence of *ta(a)n* – at least for the backward binding, non-sentient, and logophoric cases discussed here:

(35) **Condition for potential *ta(a)n*-antecedence (Version 2):**
A potential antecedent of *ta(a)n* is a nominal which has a psychological/mental attitude with respect to a proposition in which the anaphor is a participant (thematic argument). This psychological/mental attitude is linguistically or discourse-saliently available to this proposition.[5]

3.2.1 How pervasive is the antecedence condition?

An important question to ask at this point is whether the condition given in (35) applies only to the "problematic" cases of backward binding, logophoricity, and ban on non-sentience antecedence, discussed above, or whether it is more generally implemented. If we find that the antecedence condition is only relevant in the former case, that is a reason to assume that there are two non-overlapping sets of phenomena: "long-distance anaphora", on the one hand, which is perhaps regulated by structural

[5]The term "availability" is used in a very informal and intuitive sense for now. It will be defined formally in due course.

principles, and "logophora" which is governed by the lexical-conceptual and discourse factors outlined in (35). If, on the other hand, we find that the antecedence condition holds more generally, then we would have reason to pursue a unified analysis of both the "problematic" and straightforwardly structural long-distance patterns in Tamil.

Consider again a standard long-distance binding structure in Tamil, such as (36), repeated from (13) above:

(36) [$_{CP}$ Raman$_j$ Anand-kiṭṭæ$_k$ [$_{CP}$ Seetha$_l$ tann-
Raman[NOM] Anand-ALL Seetha[NOM] ANAPH-
æ$_{\{i,j,*k,*l\}}$ kaappaatt-in-aaḷ-ŭnnŭ] so-nn-aan-nnŭ] Krish-
ACC save-PST-3FSG-COMP] say-PST-3MSG-COMP Krish-
nan$_i$ paar-tt-aan.
nan[NOM] saw-PST-3MSG
"Krishnan$_i$ saw [$_{CP}$ that Raman$_j$ told Anand$_k$ [$_{CP}$ that Seetha$_l$ saved him$_{\{i,j,*k,*l\}}$.]]"

Possible antecedents for *ta(a)n* in this sentence are the matrix subject *Krishnan* and the intermediate subject *Raman*; the clausemate subject *Seetha* is ruled out as an antecedent because of the Ban on Clausemate Subject Antecedence, and the intermediate object *Anand* is also excluded from antecedence because of the subject-orientation of *ta(a)n*. (36) thus represents a "well-behaved" long-distance binding structure.

But let us look closer at the interpretive properties of this sentence, with a focus on the antecedents. The matrix subject *Krishnan* is the subject of a propositional perception predicate: Krishnan is the perceiver. The intermediate subject *Raman*, the other possible antecedent, is the subject of a speech predicate; Raman is the speaker. Do these entities satisfy the antecedence condition given in (35)? Actually, they do. The antecedents are both entities that bear a mental perspective or attitude toward the proposition in their scope. Krishnan's mental state (specifically, his mental perception of a propositional event) involves, and is thus available to, the innermost clause containing the anaphor. Similarly, *Raman's* status as the source of information about an event is accessible to the clause containing the anaphor.

This suggests that the antecedence condition given in (35) is relevant not only to the determination of the principles governing antecedence in sentences involving logophoric reference and backward binding but also for the more run-of-the-mill long-distance binding sentences like (36). How do we deal with this observation? We, of course, still have the analytic option of proclaiming that *ta(a)n*-antecedence in sentences like (36) has to do with structural conditions pertaining to syntactic subject-

hood, whereas antecedence in the more "problematic" cases is governed by the descriptive condition given in (35). But this is a clearly less appealing option both in its empirical scope and theoretical elegance. A subject-orientation account of antecedence would only be able to account for standard cases of long-distance binding such as (36); we would *still* need some version of (35) to explain the logophoric and backward binding phenomena and sentience restriction on subject-antecedence. Also since one unified explanation is better than two (Occam's Razor) – a treatment of anaphoric antecedence in terms of (35) is also theoretically more attractive.

For these reasons, I will adopt (35) as a tentative description of the conditions governing the antecedence of *ta(a)n* while remaining agnostic, for the time being, about exactly *how* it is to be formally implemented and, in this context, also about whether it is to be structurally or conceptually implemented. A more precise description and formal implementation of this condition, involving a hybrid syntactic-conceptual account will follow in due course.

3.2.2 Deriving the ban on object antecedence

Why can the intermediate object *Anand* not be an antecedent for *ta(a)n* in (36)? Shouldn't this be possible, especially given that Anand, being human, is technically capable of sentience? Not necessarily. As Dowty (1991, 573) puts it: "Sentience means more than a presupposition that an argument is a sentient being; it is rather sentience with respect to the event or state denoted by the verb." In this sense, Anand is clearly not sentient with respect to the situational predication involving the anaphor. The entity denoted by the GOAL/RECIPIENT *Anand* in (36) is a passive receiver of information. He may already happen to bear an opinion or attitude towards the proposition that is communicated to him; alternatively, he may come to bear an attitude as a result of hearing this information. However, these are properties of the world, not of the linguistic representation of this sentence. To put it another way, in the unmarked discourse scenario, the embedded proposition in (36) is not presenting the perspective of Anand. As such, *Anand* doesn't qualify as an antecedent for *ta(a)n*, just as the antecedence condition in (35) would lead us to expect. This suggests that when the object of the sentence *is* sentient in Dowty's sense – i.e. bears an attitude towards the proposition in which the anaphor is an argument – it should be able to antecede *ta(a)n*. This prediction has already been seen to be borne out – the backward binding structures in (30)-(33) all instantiate this pattern –

showing that our analysis here is on the right track.

This shows that the ban on object antecedence in structures like (36) is not *directly* due to the objecthood status of such nominals. This is merely an epiphenomenon of the fact that syntactic objects, for *independent* reasons pertaining to the way in which thematic roles are mapped onto grammatical function tend to be arguments which do not bear a mental attitude toward the proposition selected by their predicate. To sum up, the subject-orientation of an anaphor does tend to be true as a *descriptive* generalization which captures the idea that entities that satisfy the conceptual requirements for *ta(a)n*-antecedence (as delineated in (35)) tend to be syntactic subjects – but this is a misleading label for this effect because it suggests a direct correlation between syntactic subjecthood and *ta(a)n*-antecedence which is actually not empirically attested.

3.2.3 Taking stock

We ended the previous chapter with an empirical profile for long-distance binding in Tamil which included a series of properties that seem to resist a structural treatment, namely: non-locality, non-minimality, and optionality of the anaphoric antecedent. In this chapter, we have thus far zoomed in on one of the properties of long-distance binding in Tamil which seemed initially more amenable to a structural treatment: the so-called "subject orientation" of *ta(a)n*. However, upon closer examination, we have been forced to conclude that this is not a syntactic restriction but a conceptually motivated one: evidence in support of this conclusion has come from backward binding, logophoricity and restrictions on animacy on the part of the antecedent. As such, a more inclusive and accurate characterization of antecedence has been seen to be in terms of perspective-holding, and not in terms of grammatical function.

We've also seen that the choice of anaphoric antecedent for *ta(a)n* is not fully deterministic. One source of indeterminacy comes from the fact there is more than one factor that contributes to the nature and degree to which an entity is mentally involved in the proposition in which the anaphor is contained. One of them is the type of thematic relationship the entity bears with its predicate, as we have seen. There is a second source, however, and this is the relationship between the entity and the salient discourse. Thus, in (37) below (repeated from (33)), the possessive pronoun *avanooḍæ* (HIS), though not itself associated with an AGENT or EXPERIENCER role, may nevertheless qualify for antecedence of *ta(a)n* just in case the immediate discourse confers it with the right

level of mental perspective/involvement with the proposition containing the anaphor:

(37) [$_{CP}$ [$_{DP}$ Taan$_{\{i,j\}}$ avva[avŭ ee[æ-yaaga irŭnd-adŭ]
ANAPH[NOM] so poor-ADJ be-PST-3NSG.NOM
[$_{DP}$ avan-ood[æ$_i$ an[aav-æ$_j$] rombæ-vee baadi-jirŭ-kkir-adŭ.
Raman-GEN brother-ACC very-EMPH affect-be-PRS-3NSG
"[$_{DP}$ His$_{\{i,j\}}$ having been so poor] has really affected [$_{DP}$ [$_{DP}$ his$_i$] brother]$_j$."

The choice of antecedent might be indeterminate even within a particular proposition because of the dual influences on potential antecedence from both thematic and discourse factors. Thus, in (38) below, both the matrix subject CAUSER, *Krishnan* and the CAUSEE EXPERIENCER *Raman* qualify as potential antecedents for *ta(a)n*:

(38) Krishnan$_i$ [$_{CP}$ Seetha tann-æ$_{\{i,??j\}}$ kaadali-kkir-aa[-
Krishnan[NOM] Seetha[NOM] ANAPH-ACC love-PRS-3FSG-
ŭnnŭ] Raman-æ$_j$ nenekka-vej-tt-aan.
COMP Raman-ACC think-CAUS-PST-3MSG
"Krishnan$_i$ made Raman$_j$ believe [$_{CP}$ that Seetha loved him$_{\{i,??j\}}$]"

In the pragmatically unmarked case, *Krishnan* would be strongly preferred over *Raman* as an antecedent for *ta(a)n*. This is because, although *Raman* denotes a believer, thus an attitude-holder, with respect to the proposition containing *ta(a)n*, he is made to hold this belief by *Krishnan*. As such, the embedded proposition containing *ta(a)n*, is more likely to be viewed from Krishnan's perspective. However, the discourse circumstances could be altered – e.g. by situating this sentence in a discourse structure that reported primarily on Raman's point-of-view – such that the embedded proposition could be viewed just as easily from Raman's perspective as from Krishnan's. This shows that although the thematic relationship between a DP and its predicate might predispose it to a greater or lesser degree to be a perspective-holder, the discourse-context also makes a crucial contribution.

Based on this type of data, we might propose that the condition for potential antecedence of *ta(a)n*, is actually something like this:

(39) **Condition for potential *ta(a)n*-antecedence (Pre-Final version):**

i. A potential antecedent for *ta(a)n* is a nominal which has a psychological/mental perspective with respect to a proposi-

tion in which the anaphor is a participant (i.e. thematic argument).

ii. The information pertaining to this mental perspective is available (in a manner to be made more precise) to the anaphor.

iii. The degree to which a nominal may qualify for potential antecedence is a function of its relationship both with its clausemate verb (thematic factors) and with the salient discourse. Specifically, it is a function of how likely it is that the minimal proposition containing the anaphor is viewed from the mental perspective of the entity denoted by that nominal.

As such, nothing we have seen so far, with respect to the Tamil long-distance binding patterns, conclusively shows that the relationship between the anaphor and its antecedent is implemented as early as the narrow syntax. The types of data discussed so far suggest rather that long-distance binding is a phenomenon characterized by conceptual restrictions alone and that it would be best treated via a purely conceptual account. Such an account could be made to take into consideration thematic roles as well as discourse-pragmatic ones such as those proposed in Sells (1987) or adapt the empathy analysis in Kuno (1987) to the Tamil data.

However, in the section below, I will show that, while the arguments developed here in favor of a conceptual view of perspective are correct, the conclusion drawn from this, namely that a *purely* conceptual analysis is warranted, is incorrect.

3.3 Evidence for a syntactic treatment of perspective

In this section, I will present two pieces of evidence for a structural subcomponent of perspective. The first comes from the Ban on Clausemate Subject Antecedence. The second comes from verbal agreement triggered when *ta(a)n* is in syntactic subject position. Both pieces of evidence illustrate that the type of perspectival relationship described thus far between an anaphor and its antecedent must be syntactically represented.

3.3.1 The structural nature of the Ban on Clausemate Subject Antecedence

The ban on clausemate subject antecedence is seen in sentences like (40) (repeated from (36)). However, here we see for the first time, real evidence that there is a structural component to perspective:

(40) [$_{CP}$ Raman$_j$ Anand-kiṭṭæ$_k$ [$_{CP}$ Seetha$_l$ tann-
Raman[NOM] Anand-ALL Seetha[NOM] ANAPH-
æ$_{\{i,j,*k,*l\}}$ kaappaatt-in-aa[-ŭnnŭ] so-nn-aan-nnŭ] Krish-
ACC save-PST-3FSG-COMP] say-PST-3MSG-COMP Krish-
nan$_i$ paar-tt-aan.
nan[NOM] saw-PST-3MSG
"Krishnan$_i$ saw [$_{CP}$ that Raman$_j$ told Anand$_k$ [$_{CP}$ that Seetha$_l$ saved him$_{\{i,j,*k,*l\}}$.]]"

Let us start with the case of *Seetha*. *Seetha* is the AGENT of the predicate SAVE in the innermost clause; the anaphor *tannæ* is the PATIENT of this same verb. The agentive θ-role of *Seetha* cannot be responsible for its non-antecedence – the intermediate subject *Raman* is agentive and still allowed to function in this capacity. The problem must rather be that the entity denoted by *Seetha* is, for some reason unable to hold a perspective toward the eventuality containing *ta(a)n*. What could this reason be? Observe that *Seetha* in (40) denotes an individual who is herself a part of the event described by the verb: in other words, *Seetha* is a co-argument of *ta(a)n*. In Part II, I will show that the failure of perspective-holding on the part of *Seetha* has to do with this property, arguing specifically that the DP that denotes a perspective-holder cannot be properly contained inside the situational predication that it holds a perspective towards.

We have been attributing the non-antecedence of subjects like *Seetha* to a Ban on Clausemate Subject Antecedence and had entertained the possibility that this is just a structural anti-locality effect. But we can now argue that this anti-locality is a reflection of a structural aspect of the wellformedness conditions on perspective holding. Note that in Tamil, local binding of *ta(a)n is* possible in psych-predicate structures:

(41) Raman$_i$ tann-æ$_{\{i,*j\}}$ verŭ-tt-aan.
Raman[NOM] ANAPH-ACC hate-PST-3MSG
"Raman$_i$ hated himself$_{\{i,*j\}}$."

(42) Raman-ŭkkŭ$_i$ tann-æ$_{\{i,*j\}}$ piḍi-tt-adŭ.
Raman-[DAT] ANAPH-ACC like-PST-3NSG

"Raman$_i$ liked himself$_{\{i,*j\}}$."

(43) [$_{DP}$ Seetha-vŭkkŭ$_i$ tann-æ$_{\{i,j\}}$ piḍikka-læ-nnŭ] Krishnan$_i$
Seetha[DAT] ANAPH[ACC-SG] like-NEG-COMP Krishnan[NOM]
paar-tt-aan.
see-PST-3MSG
"Krishnan$_j$ saw [$_{CP}$ that Seetha$_i$ didn't like him$_j$/herself$_i$.]"

At first blush, sentences like (41)-(43) seem to constitute an exception to the ban on clausemate subject antecedence and to the structural implementation of it introduced above. However, in Part II, I will argue that psych-predicate sentences instantiate a larger structure than do non-psych ones, so that the antecedent in (41)-(43) is actually outside the minimal predication containing the anaphor in these cases. As such, far from being counter-examples, such structures confirm the structural version of the ban on clausemate subject antecedence and clarify the notion of perspective-holding.

We have noted that several constraints on anaphora that intially looked structural, like the subject-orientation of *ta(a)n*, turned out on closer inspection to be motivated by conceptual properties instead. Based on such data, one might have been tempted to pursue a purely conceptual route to anaphoricity in Tamil. The Ban on Clausemate Subject Antecedence, however, constitutes the first piece of real evidence in favor of a role for structure in binding in Tamil. To be sure, the evidence that I have presented for this so far is only suggestive. In order to make a conclusive argument for it, we need to inspect local binding patterns in Tamil in great detail. This is the concern of Part II. We will thus adopt this position provisionally for now pending more definitive evidence in Part II.

3.3.2 The structural nature of perspectival agreement

The nature of verbal agreement triggered under subject *ta(a)n* presents conclusive evidence in favor of the structural representation of perspective. Tamil uniformly manifests subject agreement on the verb. Verbal agreement triggered under *ta(a)n* in subject position furnishes convincing evidence that perspective involves a syntactic core. In brief, the agreement triggered under *ta(a)n* in subject position always tracks the antecedent of *ta(a)n*.[6] Consider the examples below, all of which involve

[6]When *ta(a)n* is not in subject position, clausal agreement straightforwardly reflects the ϕ-features of its clausemate (non-anaphoric) subject.

ta(a)n in subject position – the verbal agreement triggered under *ta(a)n* is highlighted in boldface:

(44) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ tookkapoo-gir-**aan**-nnŭ]
Raman[NOM] ANAPH[NOM] prize-ACC lose.go-PRS-3MSG-COMP
namb-in-aan.
believe-PST-3MSG
"Raman$_i$ believed [$_{CP}$ that he$_{\{i,*j\}}$ would lose the prize]."

(45) Seetha$_i$ Raman-kiṯṯæ$_j$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ tookkapoo-
Seetha[NOM] Raman-ALL ANAPH[NOM] prize-ACC lose.go-
gir-**aa**ḷ-ŭnnŭ] so-nn-aaḷ.
PRS-3FSG-COMP say-PST-3FSG
"Seetha$_i$ told Raman$_j$ [$_{CP}$ that she$_{\{i,*j\}}$ would lose the prize]."

(46) Maya$_i$ [$_{CP}$ Raman$_j$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ
Maya[NOM] Raman[NOM] ANAPH[NOM] prize-ACC
tookkapoo-gir-**aa**ḷ-nnŭ] namb-in-aan-ŭnnŭ] paar-tt-aaḷ.
lose.go-PRS-3MSG-COMP believe-PST-3MSG-COMP see-PST-3FSG
"Maya$_i$ saw [$_{CP}$ that Raman$_j$ believed [$_{CP}$ that she$_i$/*he$_j$ would lose the prize]]."

(47) Maya$_i$ [$_{CP}$ Raman$_j$ [$_{CP}$ taan$_{\{*i,j\}}$ paris-æ
Maya[NOM] Raman[NOM] ANAPH[NOM] prize-ACC
tookkapoo-gir-**aan**-nnŭ] namb-in-aan-ŭnnŭ] paar-tt-aaḷ.
lose.go-PRS-3MSG-COMP believe-PST-3MSG see-PST-3MSG
"Maya$_i$ saw [$_{CP}$ that Raman$_j$ believed [$_{CP}$ that *she$_i$/he$_j$ would lose the prize]]."

In (44), the agreement under *ta(a)n* is marked 3MSG and matches the ϕ-features on the matrix subject *Raman*, which is the antecedent of *ta(a)n*. The other sentences in this list show that the agreement features on the verb in the *ta(a)n*-clause reflect the ϕ-features of the antecedent alone, not just an arbitrary superordinate DP. For instance, (45) shows that a non-potential antecedent, like the oblique object *Raman*, cannot control this agreement – the agreement reflects the ϕ-features of the antecedent, here the matrix subject *Seetha*. The structures given in (46) and (47) are potentially the most significant in this constellation: they show that even potential antecedents cannot control the agreement-marking in the *ta(a)n*-clause; the agreement tracks the ϕ-features of the actual antecedent. Thus, when *Maya* is the intended antecedent, the embedded verb is marked with the 3FSG suffix *-aa*ḷ, as in (46), and when *Raman* is the intended antecedent, it surfaces instead as *-aan* (3MSG).

The obvious conclusion one might draw from these patterns is that agreement is triggered by *ta(a)n* itself. After all, when the embedded subject is *non*-anaphoric, the agreement on the clausemate verb straightforwardly reflects the ϕ-features of this DP:

(48) Maya$_i$ [$_{CP}$ Raman$_j$ [$_{CP}$ nii$_{Addr}$ paris-æ tookkapoo-
Maya[NOM] Raman[NOM] you[NOM] prize-ACC lose.go-
gir-**aaj**-ŭnnŭ] namb-in-aan-nnŭ] paar-tt-aaḷ.
PRS-2SG-COMP believe-PST-3MSG-COMP see-PST-3FSG
"Maya$_i$ saw [$_{CP}$ that Raman$_j$ believed [$_{CP}$ that you would lose the prize]]."

However, there is reason to think that, when the embedded subject is an anaphoric element, like *ta(a)n*, the agreement on the clausemate verb comes from elsewhere. First, recall that *ta(a)n* itself doesn't "care" about the gender feature on its antecedent, suggesting that it, at the very least, lacks a gender feature altogether. Indeed, there is a rich literature, based on robust crosslinguistic evidence, which argues that anaphors have no ϕ-features at all (Kratzer 2009) or are, at least, ϕ-defective in significant ways (Pica 1987, Reinhart and Reuland 1993, Heinat 2008, Reuland 2011). A related strand of research has shown, even more relevantly to the point, that anaphors are incapable of triggering regular ϕ-agreement altogether (see Rizzi 1990, Woolford 1999, Tucker 2011, on the "Anaphor Agreement Effect"). This would lead us to expect that an anaphor in agreement-triggering position cannot value all the ϕ-features of a clausemate verb by itself. The data in the sentences given in (44)-(47) above support this idea. In these sentences, the verbal agreement under *ta(a)n* involves a full set of ϕ-features including person, number, and gender. This suggests that the agreement ultimately has a different source and is not triggered directly by *ta(a)n*.

Even more convincing evidence for this point comes from the fact that, in certain types of structures, a different sort of agreement may be triggered under *ta(a)n*, altogether. To see this more clearly, consider the sentences in (50) and (49) below:

(49) Raman$_i$ [$_{CP}$ taan$_j$ ʤej-pp-**aan**-nnŭ]
Raman[NOM] ANAPH[NOM]$_i$ win-FUT-3MSG-COMP
nene-tt-aan.
think-PST-3MSG
"Raman$_i$ thought [$_{CP}$ that he$_{\{i,*j\}}$ would win]"

(50) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ ʤej-pp-**een**-nnŭ] so-nn-aan.
Raman[NOM] ANAPH[NOM]$_i$ win-FUT-1SG-COMP say-PST-3MSG

"Raman$_i$ said [$_{CP}$ that he$_{\{i,*j\}}$ would win]"

In (49), the agreement under *ta(a)n* reflects the 3MSG ϕ-features of the antecedent of *ta(a)n*, just as we have seen with the sentences in (44)-(47) above. In (50), however, the agreement under *ta(a)n* manifests different ϕ-features from those of the antecedent of *ta(a)n*: in particular, the agreement is 1SG whereas the antecedent is still 3MSG. Under an account where the agreement on the verb under *ta(a)n* in subject position is always directly triggered by *ta(a)n* itself, we would be unable to explain this difference. We would either have to posit that *ta(a)n* in (50) is somehow different from that in (49) or claim that the ϕ-matching effect on *ta(a)n*-antecedence seen in sentences like (49) and (44)-(47) is purely accidental. Neither of these is a particular elegant option. More worryingly, the 1st-person agreement under *ta(a)n* predominantly obtains under propositional speech verbs like *soll* (SAY) in (50): a correlation that neither of these alternatives will be able to capture.

The nature and derivation of the different agreement patterns under *ta(a)n* is one of the main concerns of Part III of this dissertation. The conclusion that I will argue for there is that the agreement in all these sentences is in fact triggered by the syntactic instantiation of perspective on a pronominal operator in the local phase of *ta(a)n*. This operator will also be argued to "stand in" for the antecedent of *ta(a)n* in local as well as long-distance binding configurations, yielding the effect that the agreement on the verb under *ta(a)n* tracks the antecedent of *ta(a)n*. The special 1st-person agreement on this verb in sentences like (50) will be shown to be the result of this operator's being a shifted 1st-person indexical (Kaplan 1989, von Stechow 2002, Schlenker 2003b, Anand 2006).

For now, let us simply consider the implications of the agreement patterns under *ta(a)n* above, for the nature and representation of perspective. What they show (and show conclusively, I believe) is that the ϕ-features of the antecedent of *ta(a)n* are already "known" at the point in the derivation where the agreement features on the clausemate verb of *ta(a)n* get decided. It is taken as fairly uncontroversial that agreement is a morphosyntactic phenomenon – it is typically taken to be the result of a formal Agree operation in the "narrow" syntax, where a DP with valued ϕ-features checks unvalued (or uninterpretable) ϕ-features on a functional head like T or v (Chomsky 2001). But if ϕ-feature agreement is implemented in the syntactic module and, if as the sentential patterns in (44)-(47) show, this agreement tracks the antecedent of *ta(a)n*, then this must mean that the ϕ-features of the nominal that gets interpreted as the antecedent of *ta(a)n* are represented in the Narrow Syntax. The

logic of this argumentation may be represented as follows:

Observation I: ϕ-feature agreement under subject *ta(a)n* is not directly triggered by *ta(a)n*.

Observation II: ϕ-feature agreement under subject *ta(a)n* tracks the antecedent of *ta(a)n*.

Assumption: ϕ-feature agreement is locally implemented in the Narrow Syntax.

Conclusion I: The ϕ-features of the nominal that gets interpreted as the antecedent of *ta(a)n* are represented on a local entity in the Narrow Syntax.

Conclusion II: The antecedent is itself not a local entity with respect to the anaphor. Thus, the local entity hosting the ϕ-features of the antecedent must be distinct from both the antecedent and the anaphor.

Interestingly enough, this same argument can be used to show that logophoric binding also involves a core syntactic component in Tamil. Consider the following sentences: in all of them, *ta(a)n* is the matrix subject and refers logophorically – i.e. to an extra-sentential antecedent with a mental perspective toward the minimal proposition in which it is contained.

(51) LOGOPHORIC *ta(a)n* AS A MATRIX NOMINATIVE:

a. Seetha$_i$ naḍandadæ-patti nallaa joosi-tt-aaḷ.
Seetha[NOM] happening-ACC-about deeply reflect-PST-3FSG.
Taan$_i$ een ivvaḷavŭ kašṭappaṭṭ-irŭ-kk-**aa**ḷ?
ANAPH[NOM] why this.much suffer-PRF-PRS-3FSG
"Seetha$_i$ reflected deeply about what had happened. Why had she$_i$ suffered this much?"

b. Raman-ŭkkŭ$_i$ oṉṉum-ee puriya-læ. Taan$_i$
Raman-DAT nothing[ACC]-EMPH understand-NEG. ANAPH[NOM]
maṭṭum een ippaḍi ellaattaiyum tolæččŭkkoṇḍ-ee irŭ-nd-**aan**?
only why like.this everything lose-PROG-EMPH be-PST-3MSG?
"Raman$_i$ didn't understand at all. Why did he$_{\{i,*j\}}$ alone keep losing things constantly?"

Just as with the long-distance binding patterns in (44)-(47) above, the agreement under *ta(a)n* reflects the ϕ-features of the antecedent – here, the extra-sentential attitude-holder toward the proposition containing *ta(a)n*. Thus, in (51a), the verbal agreement under subject *ta(a)n* is 3FSG which is the same as that on *Seetha*, a feminine nominal; in (51b), the verbal agreement under *ta(a)n* is 3MSG, matching the ϕ-features of the logophoric antecedent, *Raman*. Any other agreement-values than those given here are impossible and lead to strict ungrammaticality. Notice that these verbs show full person, number, and gender agreement – recall, again, that *ta(a)n* itself doesn't appear to be marked for gender. This suggests, as before, that the features on the verbal agreement are not directly triggered by *ta(a)n* itself.

But this must mean, as it did for the long-distance binding cases above, that even a logophoric antecedent of *ta(a)n* is already determined in the syntax. This is an important discovery: not only does it show that "logophoric" binding has a syntactic component, it also provides empirical support for our analytic position that a unified approach to long-distance binding and logophoricity is warranted. This yields the following conclusion:

(52) **Unified Binding Hypothesis:**

i. Logophoric as well as long-distance anaphoric binding involve a core syntactic component.

ii. A unified approach to logophoric and anaphoric phenomena is empirically warranted.

3.3.3 The hybrid nature of perspective

We have just seen two pieces of evidence showing that all cases of long-distance and so-called logophoric binding involve a core structural component: the Ban on Clausemate Subject Antecedence which, we have proposed, is actually a perspectivally motivated syntactic condition of anti-locality, and the perspectival agreement facts discussed immediately above. The agreement facts, in particular, have demonstrated that the features of the DP that ends up being construed as the antecedent of *ta(a)n* are already syntactically represented and available in the local domain of *ta(a)n*, thus may trigger the agreement on the verb under *ta(a)n* in subject position. In addition, it seems fairly clear that the entity that hosts these features cannot be the antecedent DP itself. After all, the very definitions of both long-distance and logophoric binding involve the idea that the antecedent is explicitly *not* represented in the local domain

of the anaphor. As such, this element must not be the antecedent itself, but an object that stands in for the antecedent.

This is where the nature of perspective enters the picture. All the cases of long-distance and logophoric binding investigated thus far have involved the central role of perspective. In particular, an antecedent has been seen to be a type of perspective-holder toward the minimal predication containing the anaphor. The most intuitive way to combine these conclusions would thus be to say that the element that stands in for the antecedent is a syntactic representation of the perspective holder. In sentences involving an anaphor, such as those we have seen so far, this perspective-holder will also stand in for the anaphoric antecedent.

This paves the way for distinct roles for structural and conceptual information pertaining to anaphora. In the next chapters, I will motivate a cross-modular binding model that involves two distinct relationships. The first, I will propose, is a syntactic dependency that holds between the anaphor and a pronominal operator that stands in for the anaphoric antecedent in the local phase of the anaphor. The second is a more conceptual relationship that holds between the DP that is construed as the actual antecedent of the anaphor and this operator which, I will propose, instantiates a non-obligatory control relationship (Williams 1980). Such a model will allow us to capture the structural conditions on binding seen above as well as its more conceptual aspects involving factors such as non-locality, non-minimality, and antecedence optionality and indeterminacy.

Based on structures such as these I will thus propose the following:

(53) **The syntactic nature of perspective:**

i. The relationship between the anaphor and an entity, containing information pertaining to the antecedent, is syntactic in nature, thus constrained by syntactic principles of locality and minimality.

ii. This shows that a core component of long-distance binding (in Tamil and languages with similar binding patterns) is (narrow-) syntactic.

In the next chapter, I will develop a formal account of the nature and representation of perspective and, relatedly, of the two-step binding model in which it plays the mediating role.

Chapter 4

The nature and representation of "perspective"

How and where is the perspective of an antecedent represented?

In order to properly answer this, we need to tease apart two analytic issues. The first has to do with whether long-distance binding is possible into a particular clause or not. The second question, conditional on an affirmative answer to the first, is what the potential and likely antecedents for an anaphor in that clause are. We have, so far, concentrated on patterns of long-distance binding that *are* possible in Tamil. As such, our focus so far has been on finessing an answer to the second question, namely: nailing down the conditions that influence and determine *ta(a)n*-antecedence. Let us now look at the first issue more closely with a view toward understanding the formal representation of the perspective of an antecedent.

4.1 What conditions whether long-distance binding is possible?

It is fairly straightforward to show that long-distance binding is not possible in all cases. In order for a long-distance binding relationship to be established, a clause must, at the very least, be capable of allowing long-distance dependencies. This means that, in root clauses, only logophoric reference is possible, standard long-distance binding is not. Embedded clauses, on the other hand, by virtue of being subordinate to a higher clause, have superordinate arguments which could, in principle, func-

tion as antecedents for an anaphor. Whether these potential antecedents can, in fact, bind into a subordinate clause is another matter, however. Among embedded clauses, there are adjuncts as well as complements. Long-distance binding of *ta(a)n* is possible into both. However, we will see that the factors affecting the representation of antecedence perspective are different in each case. In the case of binding into complements, the *selectional properties* of the immediately superordinate predicate will be seen to be crucial. In all other cases of binding, however, including binding into adjuncts and logophoric binding, however, there is no selection involved: the representation of antecedence perspective must be due to properties that are strictly internal to the adjunct itself. Let us address these in turn.

4.1.1 Propositional predicates and the representation of a mental perspective

Looking specifically at complement clauses, most theories agree that their relationship to the immediately superordinate clause is regulated by the selectional properties of the superordinate predicate (Grimshaw 1979, Pesetsky 1982, Marantz 1984, among many others) although opinions vary considerably with respect to the origin and formal implementation of this property. Propositional predicates have been observed to be special in this regard because the clausal complements of such predicates are *attitude-reports* whose truth-value may be evaluated, to varying degrees, from the perspective of an *attitude-holder* (rather than relative to the actual context of utterance).[1] Non-propositional predicates, on the other hand, typically do not represent the mental attitude of an individual associated with that predicate. The notable exception to this is the class of psych-predicates involving verbs like FRIGHTEN and AMUSE (Beletti and Rizzi 1988) which represent the mental state of their EXPERIENCER argument – we will deal with this in more detail in Part II.

The significance of this set of facts for binding theory, of course, is that long-distance binding obtains in the complements of propositional- and psych-predicates, as we have seen for Tamil. However, long-distance binding has been shown to be sensitive to even more fine-grained syntactico-semantic distinctions among propositional predicates. Such predicates have been observed to differ in the degree to which their clausal complement is "shielded" from evaluation against the actual world, time,

[1] More formally: "if the complement of an attitude sentence presupposes p, then that sentence as a whole presupposes that the attitude-holder believes p" (Karttunen 1974, via Heim (1982)).

location and other parameters pertaining to the utterance context (see, for instance Stowell 1982, Wurmbrand 2001, Giorgi 2010, for discussions about Double Access vs. Sequence of Tense Readings under different propositional verbs) – leading to the categorization of verbs into different syntactico-semantic classes (Cinque 1999, Cristofaro 2005, Speas 2004). This has been correlated to a distinction with respect to the identity and range of propositional predicates that allow long-distance binding of anaphors in their scope. For instance, Culy (1994) shows that some languages, like the Chadic language Mupun, allow anaphoric elements only to be bound in the complement of verbs of saying; others, like Donna Sɔ of the Niger Congo family, allow long-distance anaphors under verbs of thought as well as under verbs of speech while yet others are even more lax, allowing anaphors to be bound under the scope of all kinds of propositional predicate. Tamil seems to be a very promiscuous language relative to many others with respect to its long-distance binding possibilities: it allows long-distance binding under all classes of propositional predicate, ranging from the propositional complements of speech-, thought-, knowledge-, and direct perception verbs to subjunctive and control complements.

These types of data suggest that there is a tight correlation between the argument-structural properties of a predicate and the possibility of long-distance binding in its scope. Recent proposals within the cartographic tradition (Rizzi 1997, Cinque 1999, Speas and Tenny 2003, Speas 2004, Bianchi 2003, Giorgi 2006; 2010) attempt to capture such correlations by proposing that there is a designated *syntactic* position, in the left-periphery of clauses, which contains information pertaining to the mental perspective of a superordinate attitude-holder. The idea is that the syntactic representation of this mental perspective is at the core of what the attitude-predicate inherently "means".

Given this background, we may now propose that the representation of perspective is as follows:

(54) **Formal representation of mental perspective (Version 1):**

i. The mental perspective of an anaphoric antecedent is syntactically represented in the left periphery of the clause containing the anaphor.

ii. The representation of such a mental perspective is contingent on the selectional properties of a superordinate attitude predicate.

The description in (54) will account for the representation of mental perspective in structures involving cross-clausal binding dependencies,

specifically into a propositional complement.

The problem, however, is that there are other types of long-distance binding relations that it will not be able to account for, such as:

(i) Logophoric binding

(ii) Backward binding

(iii) Binding into spatio-temporal PPs and possessive DPs

(iv) Binding into clausal adjuncts

Such structures don't involve binding into the clausal complements of attitude verbs. As such, the generalization in (54) needs to be extended to accommodate them.

4.1.2 Logophoric and backward binding

In the case of long-distance binding into the propositional complements of attitude verbs, we have proposed that the syntactic representation of mental perspective in the left-periphery of the clausal complement is at the core of what these predicates "mean". We can make a similar case for logophoric and backward binding patterns as well. The data we have discussed so far, for Tamil but also for other languages like Icelandic, Italian, Japanese, and Malayalam – show that *ta(a)n* represents the mental perspective of an antecedent in such structures too. In the case of logophoric reference, the salient discourse typically involves attitude-predicates such as verbs of saying, thinking, and feeling which, in turn, introduce the attitude-holder that serves as the anaphoric antecedent – this can be ascertained from the logophoric patterns given in (17), (18), (28), (29) in Chapter 3; see also Clements (1975), Banfield (1982), Sells (1987), Bianchi (2003), Schlenker (2004) for further data and discussion. Structures involving backward binding also involve attitude-predicates – the backward binding sentences given in (30)-(37) for Tamil and in (11) for Japanese and Italian all involve the binding of *ta(a)n* by the EXPERIENCERS of psych-predicates, for instance. This suggests that the representation of the mental perspective of an antecedent is regulated by factors pertaining to the lexical-conceptual semantics of an attitude predicate in these cases as well.

There are two significant differences, however. First, the antecedent DP in backward binding structures is not (structurally) superordinate to the clause containing the anaphor (see again the sentences given

in (30)-(33), this being precisely why the phenomenon is called "backward" binding. The relevant condition that subsumes backward binding, logophoricity and long-distance anaphoric binding into clausal complements is not structural superordinateness but factors that are more discourse-pragmatic in nature, like discourse salience, common ground, conversational implicature, and the like. The second important difference is that the anaphoric clause is not *selected* by the attitude predicate in question in logophoric and backward binding structures.

Let us propose a revised generalization of (54) based on these observations:

(55) **Formal representation of mental perspective (Version 2):**

i. The mental perspective of an anaphoric antecedent is syntactically represented in the left periphery of the clause containing the anaphor.

ii. The representation of such a mental perspective is contingent on the lexical-conceptual properties of an attitude predicate which is discourse-pragmatically prominent, in ways having to do with discourse-salience, common ground, conversational implicature, and the like.

However, even this is too restrictive, as the following discussion on long-distance binding into adjuncts shows.

4.2 Binding into PP adjuncts and DP complements

As mentioned earlier, Tamil allows binding into phrasal adjuncts: this includes clausal adjuncts as well as adjunct spatio-temporal PPs and possessive DPs. The problem with these structures is that, for many, *ta(a)n*-antecedence doesn't seem to involve the representation of a mental perspective at all, but something more abstract. We thus not only need to revise our conception of how antecedence perspective is represented (55) but also our descriptive condition on antecedence, given in (39).

4.2.1 Binding into spatio-temporal PPs and possessor DPs

Consider the following sentences:

(56) Oblique *ta(a)n* in spatio-temporal PP:

a. Raman$_i$ tan-akkŭ$_{\{i,*j\}}$ meelæ orŭ plane-æ paar-tt-aan.
Raman.NOM ANAPH-DAT above a plane-ACC see-PST-3MSG
"Raman$_i$ saw a plane above himself$_{\{i,*j\}}$."

b. Raman$_i$ tan-akkŭ$_{\{i,*j\}}$ appŭram va-nd-æ
Raman[NOM] ANAPH-DAT after come-PST-REL
paaḍagan-æ aḍi-tt-aan.
singer-ACC hit-PST-3MSG
"Raman$_i$ hit the singer who appeared after him(self)$_{\{i,*j\}}$"

(57) POSSESSOR *ta(a)n*: INSIDE GENITIVE DP:

a. Raman$_i$ [$_{DP}$ tann-ooḍæ$_{\{i,*j\}}$ mugatt-æ] kaṇṇaaḍi-læ
Raman[NOM] ANAPH-GEN face-ACC mirror-LOC
paar-tt-aan.
see-PST-3MSG
"Raman$_i$ saw his$_{\{i,*j\}}$ face in the mirror."

b. Raman$_i$ [$_{PP}$ [$_{DP}$ tann-ooḍæ$_{\{i,*j\}}$ viiṭṭ-ŭkkŭ] uḷḷæ]
Raman[NOM] ANAPH-GEN house-DAT inside
irŭ-kk-aan.
be-PRS-3MSG
"Raman$_i$ is [$_{PP}$ inside [$_{DP}$ his$_{\{i,*j\}}$ house]]."

Such patterns show two things for Tamil. First, they illustrate that long-distance binding is not always cross-clausal. Second, they suggest that an attitude predicate and, by transitivity, a mental perspective associated with the attitude-holder argument of such a predicate are not always required for the establishment of long-distance binding relations.

On the other hand, the anaphoric antecedent might be taken to have a *spatio-temporal* perspective toward the DP/PP containing the anaphor, in such sentences. As we have seen, Sells (1987), discussing logophoric and long-distance binding patterns in a variety of languages, convincingly demonstrates that anaphors track not only the communicative SOURCE or mental SELF of their antecedents, but also their physical perspective (a role he labels PIVOT). Similarly, Kuno (1987) shows that in structures instantiating the so-called "empathy" phenomenon in Japanese, the anaphor may track the physical perspective of its antecedent.

Additional empirical evidence that this line of reasoning is on the right track comes from interpretive differences that obtain as a function of the use of anaphors vs. deictic pronominals in such sentences. A coreferent deictic pronoun may replace *ta(a)n* in all the sentences above, a possible break-down in complementarity that has also been observed in similar structures in other languages (Reinhart and Reuland 1993).

However, in such cases, the relevant XP takes on the perspective of the speaker; in the *ta(a)n*-sentences given above, on the other hand, the perspective is that of the anaphoric antecedent (see Rooryck and vanden Wyngaerd 2011, Kuno 1987, Sells 1987, among others, for more data and discussion). We can show this distinction most clearly using minimal pairs like the following, where the perspective of the speaker and that of the anaphoric antecedent are clearly different; the choice of pro-form (anaphor vs. deictic pronoun) in each sentence tracks this difference:

(58) **Oblique *ta(a)n* vs. deictic pronoun in spatio-temporal PPs:**

a. Tan-akkŭ$_{\{i,*j\}}$ pinnaalæ orŭ perijæ poṯṯi irŭ-kk-æ, Raman-ANAPH-DAT behind one big box be-PRS-REL Raman-aalæ$_i$ vaɳḍi-æ ooṯṯæ-muḍija-læ.
INS car-ACC drive-could-NEG
"With a big box behind him$_{\{i,*j\}}$, Raman$_i$ couldn't drive the car."

b. Avan-ŭkkŭ$_{\{i,j\}}$ pinnaalæ orŭ perijæ poṯṯi irŭ-kk-æ, Raman-he-DAT behind one big box be-PRS-REL Raman-aalæ$_i$ vaɳḍi-æ ooṯṯæ-muḍija-læ.
INS car-ACC drive-could-NEG
"With a big box behind him$_{\{i,j\}}$, Raman$_i$ couldn't drive the car."

(59) ***ta(a)n* vs. deictic pronoun inside possessive DP:**

a. Raman$_i$ tann-ooḍæ$_{\{i,*j\}}$ eḍædŭ-pakkattŭ-læ irŭ-nd-æ
Raman ANAPH-DAT left-side-LOC be-PST-REL
paamb-æ ko-nn-aan.
snake-ACC kill-PST-3MSG
"Raman$_i$ killed the snake that was to his$_{\{i,*j\}}$ left."

b. Raman$_i$ avan-ŭkkŭ$_{\{i,j\}}$ eḍædŭ-pakkattŭ-læ irŭ-nd-æ
Raman he-DAT left-side-LOC be-PST-REL
paamb-æ ko-nn-aan.
snake-ACC kill-PST-3MSG
"Raman$_i$ killed the snake (that was) to his$_{\{i,j\}}$ left."

For the minimal pair in (58), assume the following scenario: I (the speaker) am standing behind Raman's car; Raman is sitting in the driver's seat of the car, facing away from me. The sentence in (58a) has the reading that the big box is by the rear bumper of the car (on the side of the trunk, perhaps next to me). Raman cannot *back the car out* because the big box is in the way. This is because the location of the big box is interpreted from Raman's spatial perspective in the driver's seat of the

car – a reading that crucially favors the use of *ta(a)n*. The sentence in (58b), on the other hand, has the reading that the big box is by the front bumper of the car (on the side of the headlights, on the other side of the car from me). Raman cannot *drive the car forward* because the big box is in the way. This is because the spatial perspective in this sentence is mine, the speaker's, not Raman's, a reading that favors the use of the deictic pronoun. For the minimal pair in (59), assume that Raman and I are standing face-to-face: as such, his right is my left; his left is my right. (59a) has the reading that the snake is to the left of Raman – the "left-ness" of the snake is evaluated from Raman's spatial perspective, an interpretation that favors the use of *ta(a)n*; (59b), on the other hand, has the snake to the right of Raman – the concept of "left-ness" is evaluated from my perspective. This sentence uses the deictic form over the anaphoric one.

The existence of such patterns has motivated proposals that the internal structure of spatial, and temporal PPs contains certain types of contextual information pertaining to the time and location of a speaker (Svenonius (2008) uses the possibility of expressing proximal vs. distal distinctions within spatial PPs as support in favor of a deictic projection internal to the PP which accesses the speaker's location, for instance) or to a perspective-holder within the sentence (see Rooryck and vanden Wyngaerd 2011, for discussion of "observer-centered"/deictic vs. "object-centered"/"intrinsic frames" to capture this distinction). Returning to the larger issue of *ta(a)n*-binding, data such as these show that the idea of a mental perspective for *ta(a)n*-antecedence is too restrictive; *ta(a)n* must be allowed to access the spatial and temporal perspective of its antecedent, as well.

Crucial evidence that a spatio-temporal PP or possessor DP of the kind illustrated above hosts its own perspectival center, comes from the possibility of sentences like (60) below:

(60) Raman$_i$ [$_{CP}$ Seetha$_j$ [$_{DP}$ tann-oo ḍ æ{i,j} paṇatt-æ] [$_{DP}$
Raman[NOM] Seetha ANAPH-GEN money-ACC
tan-akkŭ$_{\{i,j\}}$ pakkatt-ŭlæ] oḷi-tt-aaḷ-ŭnnŭ] paar-tt-aan.
ANAPH-DAT near-LOC hide-PST-3FSG-COMP see-PST-3MSG
"Raman$_i$ saw [$_{CP}$ that Seetha$_j$ hid [$_{DP}$ his$_i$/her$_j$ cash] right near [$_{DP}$ him$_i$/herself$_j$]]."

The sentence above involves both a possessor DP and a spatial PP. The anaphors inside these projections may both refer to *Raman* or both to *Seetha*. This in itself is not surprising, as we have just seen that binding into both possessor DPs and spatio-temporal PPs is possible in Tamil.

What is interesting is the reading under which one of the anaphors refers to *Raman* while the other refers to *Seetha*. In other words, readings like the following:

(61) "Raman$_i$ saw [$_{CP}$ that Seetha$_j$ hid [$_{DP}$ her$_j$ cash] right near [$_{DP}$ him$_i$]]." – and,

(62) "Raman$_i$ saw [$_{CP}$ that Seetha$_j$ hid [$_{DP}$ his$_i$ cash] right near [$_{DP}$ herself$_j$]]."

If the sentence in (60) had only one perspectival center (e.g. in the embedded CP complement), we would expect both anaphors to simultaneously refer to Seetha or both simultaneously to Raman since both anaphors would ultimately get their reference from this same perspectival center. I thus take the availability of readings like those in (61) and (62) as conclusive proof that (certain) possessor DPs and spatio-temporal PPs in Tamil are each capable of hosting their own perspectival center.

4.2.2 Binding into adjunct CPs

Long-distance binding of *ta(a)n* is possible into purposive, temporal, causal, concessive, conditional and manner adjunct clauses (as well as into relative clauses). Here are some illustrative sentences:

(63) CONDITIONAL ADJUNCT:

a. Raman$_i$ [$_{CP}$ Anand$_j$ tan-akkŭ$_{\{i,*j\}}$ paṇam ta-nd-
Raman$_i$ Anand ANAPH-DAT money[NOM] give-PST-
aal-daan] veelæ-jæ sej-v-aan.
COND-ONLY work-ACC do-FUT-3MSG
"Raman$_i$ will do the work [$_{CP}$ only if Anand$_j$ pays him$_{\{i,*j\}}$.]"

(64) TEMPORAL ADJUNCT:

a. Raman$_i$ [$_{CP}$ Seetha$_j$ tann-æ$_{\{i,*j\}}$ kiḷḷ-in-æ poṛŭdŭ]
Raman Seetha ANAPH-ACC pinch-PST-REL time
sattamaagæ ka-tt-in-aan.
loudly yell-PST-3MSG
"Raman$_i$ yelled loudly [$_{CP}$ when Seetha$_j$ pinched him$_{\{i,*j\}}$]."

(65) CAUSAL ADJUNCT:

a. Raman$_i$ [$_{CP}$ Seetha$_j$ tann-æ$_{\{i,*j\}}$ tiṭṭ-in-adu-naalæ]
Raman Seetha ANAPH-ACC scold-PST-3NSG-CAUS
viiṭṭæ-viṭṭŭ ooḍ-in-aan.
house-leaving run-PST-3MSG
"Raman$_i$ ran away from the house [$_{CP}$ because Seetha$_j$ scolded him$_{\{i,*j\}}$]."

What makes long-distance binding possible in such structures? The answer is very similar to what we came up with for the spatio-temporal PP and possessor DP structures above. The use of the anaphor is licensed because it represents the perspective – spatial, temporal, and mental – of its antecedent. Interestingly, we can use very similar types of empirical evidence (as for the PP and DP cases above) to support this hypothesis. Adjunct structures such as these also allow coreferent deictic pronouns in place of *ta(a)n*. In such cases, there is, again, an interpretive difference: the use of the deictic pronoun induces a sentential interpretation from the perspective of the speaker; the use of *ta(a)n*, on the other hand, involves the perspective of the antecedent.

As before, this is more clearly shown in contexts where the perspective of the speaker and anaphoric antecedent markedly differ. Consider the minimal pairs below:

(66) *ta(a)n* **vs. deictic pronoun in causal adjunct:**

a. Ramanæ$_i$ poruttæ varækkum, avan$_i$ [$_{CP}$ Seetha$_j$ tann-æ$_{\{i,*j\}}$
Raman-ACC concerning until, he Seetha ANAPH-ACC
tiṭṭ-in-adŭ-naalæ daan viiṭṭæ-viṭṭŭ ooḍ-in-aan].
scold-PST-3NSG-CAUS only house-leaving run-PST-3MSG-EVID
"As far as Raman$_i$ is concerned, he$_i$ ran away from the house [$_{CP}$ only because Seetha$_j$ scolded him$_{\{i,*j\}}$]."

b. Ennæ poruttæ varækkum, Raman$_i$ [$_{CP}$ Seetha$_j$ avan-æ$_{\{i,j\}}$
Me-ACC concerning until, Raman Seetha he-ACC
tiṭṭ-in-adŭ-naalæ daan viiṭṭæ-viṭṭŭ ooḍ-in-aan].
scold-PST-3NSG-CAUS only house-leaving run-PST-3MSG-EVID
"As far as I am concerned, Raman$_i$ ran away from the house [$_{CP}$ only because Seetha$_j$ scolded him$_{\{i,j\}}$]."

(67) **Restrictive vs. non-restrictive readings in relative clauses:**

a. Raman$_i$ [$_{CP}$ [$_{DP}$ tann-æ$_{\{i,*j\}}$ kaadalitt-æ poṉṉŭ] rombæ
Raman[NOM] ANAPH-ACC love-REL girl[NOM] very
buddhisaali-nnŭ] so-nn-aan.
smart-COMP say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that [$_{DP}$ the girl who loved him$_{\{i,*j\}}$] was very smart.]"

b. Raman$_i$, [$_{CP}$ [$_{DP}$ Seetha$_j$, [$_{CP}$ avan-æ$_{\{i,j\}}$ kaadalitt-
Raman[NOM] Seetha[NOM], he-ACC love-PST-
aaḷ-ee anda poṉṉŭ]], rombæ buddhisaali-nnŭ]
3FSG-EMPH that girl[NOM], very smart-COMP

so-nn-aan.
say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that [$_{DP}$ Seetha, [$_{DP}$ the girl who loved him$_{\{i,*j\}}$]] – was very smart.]"

The judgments are quite clear. In the minimal pair in (66) involving the binding into a causal adjunct, the use of *ta(a)n* is clearly favored in the sentence that is explicitly from the antecedent, Raman's, perspective, namely (66a). In its minimally varying counterpart in (66b), which is from the perspective of me, the speaker, the preferred pro-form is clearly the deictic pronoun – the use of *ta(a)n* is quite marked in this context. The differences in grammaticality judgments in the structures illustrated under (67) are even more striking. (67a) involves a restrictive relative clause – the use of *ta(a)n* clearly favors the reading that the restriction (that the girl being referred to is the one who was in love with him) is made from Raman's perspective. In (67b), on the other hand, this information is part of a non-restrictive appositive, which can only be made from my (the speaker's) perspective; in this sentence, the use of *ta(a)n* is strictly ruled out.

4.3 Formalizing the observations: perspectival center and potential antecedence

Taken together, the patterns involving long-distance binding into CP, PP, and DP adjuncts reinforce our observation that the antecedent of *ta(a)n* is the nominal entity that has a mental, spatial or temporal perspective toward the phrase in which the anaphor is contained. I will now attempt to capture these intuitions in more precise terms. Fillmore (1997) proposes that every sentence has a deictic center which is a reference point with respect to which deictic expressions are to be interpreted. The deictic center includes, among other things, the present time, location, and thematic information pertaining to the speaker; a similar notion is that of Kaplan (1989)'s context which is envisioned as a tuple containing coordinates pertaining to the $Speaker, Addressee, Time$, and $World$ of the actual context of utterance.

Extending these insights, I introduce the notion of a "*perspectival center*" which contains information pertaining to the time, world, location, and mental attitude of the anaphoric antecedent. The perspectival center can also be seen as being on a par with Lewis (1979)'s enriched intensional index which is supposed to contain information pertaining to

the time, world, and location of an attitude-holder. Also clearly related is the concept of the logophoric center developed in Bianchi (2003, 3) which is described as: "a speech or mental event, with its own participants and temporal coordinates, which constitutes the centre of deixis." Bianchi distinguishes between an external and internal logophoric center – the former is envisioned as an object that is anchored to a context of utterance but the latter is seen as "a contextually introduced speech or mental event distinct from the speech event (the utterance)." My conception of perspectival center roughly corresponds to Bianchi's idea of an internal logophoric center. However, it is broader in scope and application than both Lewis' and Bianchi's versions: first, it may be associated with other eventualities besides those of speech and attitude and second, it may be introduced by linguistic strategies other than complementation, a point I return to later.

With these considerations in mind, the perspectival center is defined as follows:

(68) **Formal representation of a Perspectival Center** (Version 1):

i. The *perspectival center* contains the coordinates pertaining to the time, location, world, and/or mental information of a salient perspective holder.
ii. Certain predicational structures (at least some PPs, DPs, CPs) contain a perspectival center by virtue of what they inherently "mean". In a proper subset of these cases, the representation of the perspectival center in a phrase can be traced back to the selectional properties of its immediately superordinate predicate.
iii. A situational predication has at most one perspectival center.
iv. The predication containing a successfully bound anaphor must contain a perspectival center.

With this definition in place, I present the following as the final version of the antecedence-condition for *ta(a)n*:

(69) **Condition for potential *ta(a)n*-antecedence (Final version):**

i. A potential antecedent of *ta(a)n* is a nominal which has a mental, temporal or spatial perspective with respect to a CP, PP, or DP in which the anaphor is a participant (i.e. thematic argument).
ii. This information about the antecedent is represented as part

of the perspectival center in the minimal CP, PP or DP containing the anaphor.

This potential antecedence condition is descriptively adequate in that it accounts for *ta(a)n*-antecedence in all the long-distance binding structures we have seen so far. The description of the perspectival center, given in (68) gives us a concise descriptive account of how the perspective of the antecedent is linguistically represented. Both definitions will be central to the formal implementation of the long-distance binding patterns in the following section.

4.3.1 The relationship between the anaphor and perspectival center

Before we do that, however, let us turn to another question that we started this section with, namely: how is the perspectival center made available to the anaphor? We have already addressed this issue in Section 3.3 of the previous chapter. There we examined two pieces of evidence – namely, the Ban on Clausemate Subject Antecedence and the perspectival agreement on the verb under *ta(a)n* in subject position, which showed that the antecedence of *ta(a)n* is sensitive to structural restrictions. Based on this, we proposed the generalization in (53), repeated as (70) below:

(70) **The syntactic nature of perspective:**

i. The relationship between the anaphor and an entity containing information pertaining to the antecedent is syntactic in nature, thus constrained by syntactic principles of locality and minimality.

ii. This shows that a core component of long-distance binding (in Tamil and languages with similar binding patterns) is (narrow-) syntactic.

The entity that contains information pertaining to the antecedent is, of course, nothing other than the perspectival center, defined as in (68). Thus, as per (70), the relationship between the anaphor and the perspectival center in its local domain is syntactic in nature.

But we are in a position to make even more precise claims than these. As per the condition on potential antecedence given in (69), the perspectival center must contain information pertaining to the mental, temporal, and/or spatial perspective of the actual antecedent with respect to the minimal phase (CP, DP, PP) containing the anaphor. We can now capture this idea in structural terms by claiming that the perspectival center

is located in a syntactic position that is high enough to have scope over the anaphor, its predicate, and potentially other co-arguments – the thematic layer of the phase, in other words. Support in favor of this position comes from work within the cartographic tradition arguing that certain types of discourse-pertinent information are syntactically represented at the edge of a phase (Rizzi 1997, Cinque 1999, Bianchi 2003, among others). I will adopt this approach and formalize these conclusions as follows:

(71) **The relationship between an anaphor and perspectival center:**

(i) A long-distance bound anaphor X and the perspectival center Y are in the same minimal phase. Relevant phase domains are: CP, PP, and DP.[2]

(ii) The perspectival center is syntactically represented in a functional projection in the left periphery of the CP, PP or DP phase containing the anaphor. It c-commands the anaphor, its predicate and co-argument(s) from this position.

(iii) A given phase has at most one perspectival center.

4.3.2 The relationship between a potential antecedent and the perspectival center

We have just seen that the relationship between the perspectival center and an anaphor is entirely syntactic in nature. In contrast, it seems highly unlikely that the relationship between a potential antecedent and the perspectival center is structural.

It does, indeed, seem to be the case that structures involving long-distance binding into the complement of an immediately superordinate predicate have to do with the syntactico-semantic selectional properties of this predicate. But such structures only represent one particular type of long-distance binding dependency. Other kinds of long-distance binding, such as logophoric binding, backward binding, binding into adjuncts, as well as binding across multiple phasal boundaries don't involve such a selectional relationship. It also seems clear that many (perhaps all) of these patterns don't obey core wellformedness principles that are standardly assumed to underlie syntactic operations. In backward binding structures, the antecedent doesn't c-command the anaphor, in binding across multiple clauses the antecedent seems to bind *ta(a)n* across other potential antecedents in apparent violation of Relativized Minimality,

[2]In Part II, we will update this list with AspP which, we will show, may also host a perspectival center.

and in logophoric patterns, the antecedent isn't even syntactically represented in the same sentence as the anaphor. In all these structures, the relationship is also non-local.

All this suggests that the relationship between the linguistic antecedent of the anaphor and the perspectival center in the minimal phase of this anaphor is conceptual in nature. In fact, the nature of this relationship is very reminiscent of the non-obligatory control (or NOC) dependency discussed in Williams (1980). Consider the similarities between Williams' own characterization of this relationship, given below, and the type of relationship we have just described as holding between an anaphor and its antecedent:

(72) **Non-obligatory control** (Williams 1980, 212):

a. No antecedent is necessary.
b. If there is an antecedent, it need not c-command.
c. The antecedent may follow S [the clause containing PRO].
d. The antecedent is not uniquely determined.
e. Lexical NP can appear in the position of PRO.

An example of NOC is given below:[3]

(73) $[_{CP}$ EC$_i$ to leave] would be Max$_i$'s pleasure.
(74) $[_{CP}$ EC$_{arb}$ to leave] would be a pleasure.
(75) She$_i$ is relying on Max$_j$ $[_{CP}$ EC$_{\{i,j\}}$ to get everything done].

(73) above shows that the antecedent may follow the clause containing the controlled element and need not c-command it; (74) shows that there need not be a syntactically represented antecedent at all; (75) shows that this antecedent is not uniquely determined. The description of NOC, given in (72), is strikingly similar to what we have observed for the nature of the relationship between the intended antecedent and the perspectival center.[4]

[3]I have glossed the controlled subject in the examples here as "EC" for "empty category" because there is some debate in the literature as to what sort of element this should be taken to be: Hornstein (1999), for instance, that this element is *pro*, not PRO.

[4]The potential exception is the final property given in (72), i.e. the possibility of having an overt NP/DP in the place of the silent controlled element. This does not apply to the binding cases we have been discussing. But this is not a serious problem since there are NOC environments where no overt subject is possible as well; conversely, there are obligatorily controlled (OC) environments where an overt subject is possible (Szabolcsi 2009, Sundaresan To Appear).

Based on the absence of discernible syntactic effects in such patterns, I will propose that the relationship between a potential antecedent and the perspectival center is as described below:

(76) **The Antecedence-Perspectival Center Relationship:**
   i. This refers to the relationship between a potential antecedent and the perspectival center in the minimal phase of the anaphor.
   ii. The establishment of this relationship qualifies the potential antecedent as the actual antecedent of the anaphor.
   iii. This relationship is predominantly conceptual in nature (potential exception: anaphoric binding into a complement CP/DP/PP, by the perspective-holder argument of an immediately superordinate predicate).
   iv. It instantiates a type of non-obligatory control between the intended antecedent (controller) and the perspectival center (controllee).

The simplest assumption would be to claim that this relationship gets established at LF. Koopman and Sportiche (1989), Speas (2004), Baker (2008) and others have argued that discourse-pertinent information such as mental perspective and evidentiality are encoded on functional heads in the left-periphery of phases; they also show that such heads host a null operator in their specifier which is the real binder of an anaphoric element in its scope. I will propose that the information contained in the perspectival center is hosted inside one such functional projection in the left-periphery of a phase; I will label this projection Perspectival Phrase (PerspP). In line with the proposals in the literature cited above, I will also assume that the element in [Spec, PerspP] hosts a silent operator which binds *ta(a)n* at LF. If we assume that this operator is a pronoun that is born with its own ϕ- (and other) features, the mapping to antecedence can be conducted at LF by the assignment function. Observe that our envisioning the perspectival operator as a silent pronoun fits in rather nicely with our conception of the relationship in (76) as instantiating a type of non-obligatory control.

We might envision the mapping by assignment function to be conducted in the following manner. The range of the assignment function consists of salient individuals in the evaluation context. In order for the function to successfully map the features on the operator in [Spec, PerspP] to an individual in its range, two conditions must simultaneously hold:

(i) The linguistic representation of the individual must match the ϕ-feature values of the operator in [Spec, PerspP] in the same eval-

uation context. In other words, the choice of referent must be *consistent* with all of the information about its possible reference derived from its ϕ-featural specification in the evaluation context.[5]

(ii) The individual must be a potential antecedent. As per the potential antecedence condition in (69), this means that the individual must hold a mental, spatial, and/or temporal perspective toward the minimal XP containing the anaphor.

Heim and Kratzer (1998) propose that ϕ-features introduce presuppositional restrictions on the denotation of a nominal in terms of partial functions that apply to the assignment function. I will follow them in making this same assumption about ϕ-features and will, further, extend this intuition to potential antecedence as well. That is, I will assume that the semantico-pragmatic and thematic conditions on potential antecedence introduce partial functions that restrict the possibilities on reference assignment for the operator in [Spec, PerspP]. This presupposes, of course, that the assignment function can access the relevant components of the discourse context which determine potential antecedence. But this isn't a controversial assumption. After all, the assignment function must be able to access certain types of information like pragmatic salience and common ground in order to be able to determine the membership of the individuals in its range. If there is more than one individual in the range of the assignment function that satisfies both the partial functions carrying presuppositional restrictions on ϕ-features and potential antecedent, the decision of which individual will be chosen for assignment in a given utterance will depend on the intention of the speaker, common ground, and other discourse-pertinent factors.

4.4 A two-stage model of long-distance binding

We have discussed three distinct but related aspects of long-distance binding so far. The first has to do with factors conditioning the antecedence potential of a nominal: we have seen that this is regulated by factors pertaining to the thematic relationship between the nominal and its predicate and is also influenced by its relationship with the salient

[5]The relevance of the choice of evaluation context will become apparent in Part III, which is devoted to the investigation of a particular instantiation of indexical shift (Kaplan 1989) in Tamil, which I call "monstrous agreement".

discourse. Based on our discussion of a variety of long-distance binding facts, we have come up with the description of potential antecedence given in (69). In short, a potential antecedent denotes a perspective-holder with respect to the minimal situational predication containing the anaphor. However, perspective is not merely relevant to the establishment of anaphoric dependencies but has more general uses. In Part II, for instance, we will see that it is central to certain argument-structural relationships involving unaccusatives and psych-predicates, among others. As such, the establishment of potential antecedence is strictly orthogonal to binding: i.e. a DP may be characterized as a perspective-holder with respect to a predication even in the absence of an anaphor in that predication.

A central thesis has been the idea that perspective must be structurally represented. We have, in particular, proposed that perspectival information is represented in a "perspectival center" defined, as in (68), as a tuple containing coordinates pertaining to the mind, location, time, and/or world of a perspective-holder (including a potential antecedent). The perspectival center is hosted in a Perspectival Phrase (PerspP) at the edge of certain phases and mediates the relationship between an anaphor and its antecedent in two separate stages.

The first stage involves a relationship between the intended antecedent and a silent pronominal operator in [Spec, PerspP]. We have argued that this relationship is predominantly conceptual in nature and instantiates a type of non-obligatory control between the intended antecedent (controller) and the silent pronominal operator (controllee). The second stage pertains to the relationship between the operator in [Spec, PerspP] and the anaphor. On the strength of empirical evidence drawn from the Ban on Clausemate Subject Antecedence, which we have argued instantiates a type of anti-locality effect, and agreement patterns on clausemate verbs under subject *ta(a)n*, we have shown that this relationship must be a *local syntactic dependency*, defined as in (71) above.

One of the central goals thus far has been to motivate the conclusion that both types of relationship are necessary for the establishment of long-distance binding. The combination of conceptual and structural relationships in this model also helps explain the unique combination of syntactic and conceptual properties that characterize long-distance anaphora. The cross-modular nature of the analysis allows us to relegate the more tendential and underlyingly vague properties of long-distance binding (such as the conditions on potential antecedence, non-locality, non-minimality, and antecedence-optionality) to LF semantic or pre-theoretical conceptual mechanisms and reserve the role of the syn-

tax for explaining the exceptionless and determinate ones (such as the Ban on Clausemate Subject Antecedence, and agreement under subject *ta(a)n*). The result is a unified analysis of long-distance binding, so-called "logophoric" and backward binding phenomena. In Part II, we will see that this approach also lends itself to an explanation of local binding of *ta(a)n*.

The two-step binding model may be summarized as follows:

(77) **Two step binding model:**
Binding in Tamil and languages like it involves two distinct relationships:

i. The first is the relationship between the intended antecedent DP and a silent pronominal operator in the specifier of the minimal PerspP containing the anaphor. This instantiates a type of non-obligatory control: the antecedent is the controller and the silent pronominal operator the controllee.
ii. The second is the relationship between this operator and the anaphor. This is a syntactic dependency which, therefore, obeys conditions of locality, minimality, and c-command. The operator is the binder and the anaphor the bindee.
iii. There is thus no direct relationship between the anaphor and its antecedent, only an indirect one mediated by the silent pronominal operator in [Spec, PerspP].
iv. Thus, all binding is local. All antecedence is non-local.[6]

[6]We will, however, continue to use the terms "local binding" and "long-distance binding" in their traditional senses for purely descriptive purposes.

Chapter 5

Formalizing the proposal

In this chapter, I formalize the intuitions from the previous ones. I assume that features are specified as valued or unvalued and represented as ordered attribute-value pairs. Many of the specific labels and feature-definitions listed here are inspired by the formal feature classifications proposed in Adger and Ramchand (2005) and more recently adapted for anaphors by Hicks (2009), but my analysis will be seen to differ from these in non-trivial ways.

Recall that long-distance binding is envisioned as a two-stage process in my system: the relationship between the antecedent and the perspectival center is one, the relationship between the anaphor and the perspectival center is the other. Crucially, only the latter is assumed to be syntactic – thus our toolbox of features and structural rules only has to be able to deal with those empirical properties that directly result from the nature of this relationship. Only three properties we have observed thus far fall into this category:

(I) The syntactic factors that trigger a dependency between the anaphor and the perspectival center and the formal nature of this dependency.

(II) The Ban on Clausemate Subject Antecedence in the general case and the absence of this ban in the case of psych-predicate structures.

(III) The 3rd-person antecedence restriction on *ta(a)n*.

All the other patterns we have observed, such as the tendency toward subject orientation on the part of *ta(a)n*, the optionality of antecedence, and apparent relativized Minimality violations are not strictly a part of the relationship between the anaphor and the perspectival center. These

have to do with pragmatic, semantic and thematic factors that condition potential antecedence (see again the condition in (69)) and with the nature of the relationship between an antecedent and the perspectival center, which we are assuming is non-structural.

Here, I will restrict my attention to (I) above. (II) and (III) having to do with the Ban on Clausemate Subject Antecedence and the 3rd-person antecedence restriction on *ta(a)n* will be discussed in detail in Part II and Part III of this dissertation, respectively. The central concern of this chapter will thus be to develop a formal implementation of the relationship between the anaphor and anaphoric antecedent, defined in (71). I will be working within a Minimalist Framework (Chomsky 2001) and, as such, assume a Y-model of the grammatical architecture, with a "narrow" syntactic module that feeds the LF (meaning) and PF (sound) interfaces. I will also assume a Late Insertion model of exponence under which all morphophonological information is introduced post Spell-Out: the syntax only deals with abstract features/feature-bundles and hierarchical structure. In Minimalism, a formal syntactic dependency is typically envisioned as a form of Agree between a Probe and a Goal. I will thus assume that the syntactic dependency between the anaphor and the perspectival center instantiates Agree. It only remains to be seen which is the Probe and which the Goal in this relationship. While downward probing – namely, the idea that Agree between a Probe and Goal proceeds down a phase – is assumed in most versions of Agree (Chomsky 2000; 2001, Bošković 2007, among others), some recent theories have argued in favor of upward probing, where the Goal c-commands the Probe and Agree proceeds up the phase, either for all instances of Agree (Zeijlstra 2010) or for a designated proper subset of them (Baker 2008, Wurmbrand 2011).

As always the decision should be made empirically. In this case, that involves figuring out what features are involved in the dependency relation in the first place. This will help us decide which element is actually deficient in, thus a probe for, that feature.

5.1 Some current hypotheses about anaphoric features

With respect to this question, recent analyses of binding within the Minimalist tradition can be categorized into two broad camps. Binding analyses such as those in Reuland (2001b; 2011), Kratzer (2009) and Rooryck and vanden Wyngaerd (2011) assume, in keeping with GB-era

intuitions on this subject (Pica 1987, Progovac 1993, Reinhart and Reuland 1993) and drawing on Bouchard (1984)'s observation that a nominal needs a full set of ϕ-features in order to be LF-interpretable, that the root of anaphoricity is ϕ-deficiency. The analyses differ amongst themselves with respect to the nature and number of ϕ-features that are lacking and to what extent the deficiency is parametrized across languages: Kratzer (2009), for instance, proposes that all anaphors are "born minimal" whereas Reuland (2001b; 2011) assumes that this is a matter for parametric variation. All analyses within this camp, however, ultimately agree with the core idea that:

(NOMINAL) ANAPHORICITY $\leftrightarrow$ ϕ-DEFICIENCY.

The fact that so many long-distance bound anaphors from such a rich typological range of unrelated languages fail to mark the full range of ϕ-distinctions in the given language lends empirical support to this claim. The approach is also theoretically economical in that it doesn't posit the existence of features on nominals specifically for binding.

The second type of approach, an admittedly less popular one, rejects this bijectional mapping relation between anaphoricity and ϕ-defectiveness and proposes that nominal anaphoricity involves the deficiency of a different type of feature altogether. Hicks (2009) represents such an analytic stance, proposing that the root of anaphoricity on a nominal is the lack of a reference-index. The force of Hicks' argumentation lies in the observation that, while ϕ-features restrict the domain of reference (in the manner described in Heim and Kratzer 1998, for instance), they don't deterministically exhaust it. Thus, two DPs like *Philip* and *Oswald* may refer to two entirely different individuals but still share the same set of ϕ-features: 3rd person, masculine, and singular. A plausible response from the ϕ-deficiency camp might be to contend that this type of difference is encoded, not in the syntax, but at LF where an assignment function maps linguistic entities to salient entities in the evaluation context. ϕ-defectiveness flags a nominal as being anaphoric in the syntax and restricts the domain of the assignment function in the form of partial functions at LF, they may claim. However, Hicks also points out that there are languages with anaphors that don't lack any ϕ-features whatsoever, Modern English being one of them. This, I think, is a more significant problem for the ϕ-deficiency camp. On the other hand, under Hicks' analysis, the strong crosslinguistic tendential relationship between anaphoricity and ϕ-defectiveness must be relegated to accident or be shown to be a function of independent factors.

Thus, there are pros and cons to both approaches. In the section below, I will provide independent evidence from the binding of deictic pronominal forms in Tamil to argue in favor of Hicks' approach.

5.2 Against a ϕ-feature account of anaphoricity

In this section, I will propose that the syntactic representation of an anaphor like *ta(a)n* involves a different sort of feature-deficiency than that resulting from the absence of one or more ϕ-features. My conclusions will be based both on the types of conceptual arguments put forth by Hicks, as elucidated briefly above, and on additional evidence from Tamil which makes a treatment of *ta(a)n*-anaphoricity as ϕ-deficiency rather difficult to maintain.

A major empirical challenge to the NOMINAL ANAPHORICITY $\leftrightarrow$ ϕ-DEFICIENCY idea comes from languages with anaphors that don't seem to lack any ϕ-features whatsoever. As we have seen, Modern English is such a language. However, English is not the only language with this property. Heinat (2008) discusses examples from San Lucas Quiaviní Zapotec and Thai, among others, to show that R-expressions may be long-distance bound. The following examples are from Heinat (2008, p. 151):

(78) **San Lucas Quiaviní Zapotec**:
R-ralloh Gye'eihlly$_i$ [$_{CP}$ r-yu'lààa'z Lia Paamm Gye'eihlly$_i$].
HAB-think Mike HAB-like F Pam Mike
"Mike thinks [$_{CP}$ Pam likes Mike$_{\{i,*j\}}$]." (literal)

(79) **Thai**:
Aa-jaan$_i$ chɔɔp mă̆a tee nák-rian hă̆i aa-jaan$_i$
teacher like dog that student give teacher
"[The teacher]$_i$ likes the dog that the student gave [the teacher]$_i$."

The possibility of such patterns is undeniably problematic for an approach that treats anaphoricity as *always and only* resulting from ϕ-deficiency on the part of a nominal.

There are three plausible defenses the ϕ-deficient camp may make to data such as these. The first would be to say that structures like (78) and (79) don't involve anaphoric binding but accidental coreference. Accidental coreference between deictic pronouns and R-expressions is not unknown – thus, (80) below is perfectly grammatical as an ironic statement in English:

(80) Everyone loves Bill. Bill_i, in particular, really loves Bill_i.

However, it appears that this argument will not hold, at least not for Zapotec; Heinat shows that sloppy readings obtain under VP ellipsis of such structures, a sure sign that variable binding rather than accidental coreference is involved.

The second argument might be to claim that long-distance binding in particular is not a purely syntactic phenomenon – an analytic position that has been discussed in some detail here – thus cannot be taken to argue against hypotheses pertaining to the syntactic features on an anaphor. Other types of evidence from Zapotec, however, militate against the position that these patterns are restricted to long-distance binding, and are thus potentially unproblematic. Heinat (2008, 153) shows that R-expressions in this language may be locally bound by a co-argument antecedent. The sentence below attests to this as well as to the sloppy identity reading under ellipsis:

(81) **Sloppy readings under co-argument binding of an R-expression: Zapotec**
B-gwi'ih Gye'eihlly$_i$ lohoh Gye'eihlly$_{\{i,*j\}}$ zë'cy-cahgza' Li'eb$_j$
PRF-look Mike at Mike likewise Felipe
"Mike$_i$ looked at himself$_{\{i,*j\}}$ and Felipe did too." (i.e. Felipe$_j$ looked at himself$_j$/*Mike)

A third type of argument would be to claim that sentences like these represent a form of syncretism between R-expressions/pronouns, on the one hand, and anaphors, on the other. This is the position that Rooryck and vanden Wyngaerd (2011) adopt to deal with sentences which involve the local binding of apparent pronouns, as in the Brabant Dutch example below (Rooryck and vanden Wyngaerd 2011, 35, Ex. 53):

(82) Jan$_i$ heeft 'm$_{\{i,j\}}$ gewasse.
Jan has him washed.
"Jan$_i$ washed him$_j$", or
"Jan$_i$ washed himself$_i$"

The authors account for the apparent lack of Condition B effects by proposing that, in such structures, the bound element is underlyingly not a pronoun at all, but an anaphor. However, for independent reasons having to do with the availability (or lack thereof) of distinct forms in this language, both anaphors and pronouns are spelled out the same way on the surface. In other words, they treat sentences like (82) as instantiations of a morphological syncretism between pronouns and anaphors,

rather than as a violation of Principle B. Such an analysis is possible within the Late Insertion approach of Distributed Morphology (Halle and Marantz 1993) that the authors adopt.

Let us consider the specifics of their analysis. The authors subscribe to the ϕ-deficiency hypothesis of anaphoricity. Accordingly, in their feature-system, anaphors are born with unvalued ϕ-features and enter into an Agree relationship with a local antecedent which results in their sharing the ϕ-features of said antecedent. Pronouns, on the other hand, are claimed to be born with valued ϕ-features. This leads to the issue of how the system can distinguish between a pronoun and anaphor that have the same ϕ-features (the latter via feature-valuation and the former by virtue of having been born with such features) at Spell-Out. To solve this problem, they propose that there is a formal distinction between features that are part of a feature-sharing relationship and features that are inherent and, furthermore, that the interfaces are sensitive to this difference. The distinction is indicated notationally by marking shared feature values with a "*".[1]

To return to the Brabant Dutch example in (82), the authors assume that the bound pro-form is underlyingly an anaphor, as I have said. As such, it enters the derivation with no valued ϕ-features and participates in (a feature-sharing) Agree relation with the 3MSG antecedent *Jan*. As a result, the anaphor has the following features at Spell-Out: {P:*3**, N:*sg**, G:*m**}. A 3MSG deictic pronoun, on the other hand, is born with its fully specified set of ϕ-features: at Spell-Out, it is featurally-marked as: {P:*3*, N:*sg*, G:*m*}. What's special about Brabant Dutch (in contrast to Standard Dutch), the authors claim, is that it lacks a Spell-Out rule that makes specific reference to a nominal with shared ϕ-features, i.e. there is no Distributed Morphology-style Vocabulary Item that singles out ϕ-features annotated with a "*". Instead, they claim, there is a single Vocabulary Item for 3MSG on a nominal which is underspecified for whether those features are inherent or part of a feature-sharing relation. The specific Spell-Out Rule is shown below (Rooryck and vanden Wyngaerd 2011, 36):

(83) {P:*3(*)*, N:*sg(*)*, G:*m(*)*} ↔ *'m*/__ accusative case, weak.

Given the Subset Principle, it follows that "The Absence of Principle B effects is a direct result of the absence of dedicated reflexive forms in

[1]The authors argue that there is independent reason to think that feature-sharing as a mechanism is distinguished from feature valuation at the interfaces. As such, they propose, the "*" marking on shared features should not be treated as a second-order feature but merely as a notational mnemonic.

the relevant parts of the paradigm, and of the way lexical insertion rules are formulated and ordered" (Rooryck and vanden Wyngaerd 2011, 36). Brabant Dutch, the authors show, does indeed lack a dedicated reflexive form. But this conception of things makes the important prediction that a language that does have a dedicated reflexive form should show Principle B effects in its pronominal forms.

And herein lies the problem. Tamil is a language which, for a non-trivial subset of native speakers, allows local (as well as long-distance) binding of a deictic pronominal form (see also Annamalai 1999, for discussion of this point). Thus, the following sentences are both well-formed with binding of the object by the subject:

(84) **Local binding with deictic pronoun and *ta(a)n*:**

a. Raman-ŭkkŭ$_i$ avan-æ-yee$_{\{i,j\}}$ piḍikka-læ.
Raman[NOM] he-ACC-EMPH like-NEG
"Raman$_i$ didn't like (even) himself$_i$/him$_j$."

b. Raman$_i$ tann-æ-yee$_{\{i,*j\}}$ piḍikka-læ.
Raman[NOM] ANAPH-ACC-EMPH like-NEG
"Raman$_i$ didn't like (even) himself$_{\{i,*j\}}$."

Structures like (84a) show that the deictic 3MSG form *avan-æ* may be locally bound, just like the Brabant Dutch sentence in (82). The problem, however, is that, unlike Brabant Dutch, Tamil *does* have a dedicated reflexive form, namely *ta(a)n*. For many speakers, both *ta(a)n* and a deictic pronominal form like *avan* may be locally and long-distance bound. There is, of course, a difference in interpretation between the *avan-* and *ta(a)n*-sentences above, as we have already discussed in some detail in sentences involving long-distance binding into adjunct PPs, DPs, and CPs: the use of *ta(a)n* favors an interpretation from the point of view of the antecedent, whereas the use of the pronoun favors a reading from the perspective of the utterance-context speaker. Perhaps relatedly, the *avan*-form has a non-coreferent (i.e. deictic) as well as an anaphoric reading in the sentence given in (84a), whereas its *ta(a)n* counterpart (84b) only has the anaphoric reading.

The possibility of locally binding a deictic pronominal form, as in (84a) is entirely unexpected within Rooryck and vanden Wyngaerd (2011)'s system. Under their approach, there is no acceptable way to capture the distinctions and similarities between *avan* and *ta(a)n*. On the one hand, one could assume, along the lines of Brabant Dutch *'m* in (82), that *avan* is an underspecified form compatible with lexical insertion both in environments where 3MSG features are inherent and those where they are

the result of feature-sharing with an antecedent. One could envision a Spell-Out rule like this, for instance:

(85) {P:*3(*)*, N:*sg(*)*, G:*m(*)*} ↔ *avan*

This would allow *avan* to show up in both deictic and anaphoric contexts. However, this would leave no clear way to distinguish *ta(a)n* from the anaphoric use of *avan*. On the other hand, if we try to distinguish *avan* from *ta(a)n* by saying that the former only spells out pronouns with inherent 3MSG features (i.e. a deictic 3MSG pronoun) while the latter spells out pro-forms with feature-shared ϕ-features (e.g. 3*sg*), we would expect (84a) to violate Principle B.

One could, of course, still claim that *avanai* is not locally bound by its antecedent in (84a), but accidentally corefers with it, as in (86) below:

(86) If everyone hates John, then it surely follows that John$_i$ hates him$_i$ too.

But such sentences have been shown to be pragmatically marked and involve "guise"-like readings (Reinhart 1983b, Heim 1998). The Tamil sentence in (84a) lacks both of these interpretative properties, however, suggesting that this doesn't involve mere accidental coreference between a deictic pronoun and another referent. Finally, although I have only discussed the challenges posed by Tamil for a specific analysis within the ϕ-deficiency approach to anaphoricity, I believe that this data is, in fact, problematic for all analyses that conflate the notions of ϕ-defectiveness and referential-defectiveness. This is because a distinction in terms of ϕ-features alone will not yield the full range of observed differences and similarities between *avan* and *ta(a)n*-forms in Tamil. That is, we need some feature in addition to the various possible configurations of ϕ-features in order to indicate anaphoricity, at least of the kind exhibited by Tamil *ta(a)n*, and other languages with similar properties.

Based on this type of evidence from Tamil, the possibility of locally bound R-expressions in other languages like Zapotec, as well as the conceptual arguments of Hicks (2009) showing that ϕ-defectiveness $\neq$ referential defectiveness, I will reject the hypothesis that anaphoricity is defined by ϕ-deficiency. In what follows, I will propose a new feature to do this job.

5.3 Introducing the DEP feature

I propose that an anaphor as well as the perspectival center are characterized by a DEP feature. The DEP feature syntactically flags an element

as being one of two elements in a binding relationship: the binder or the bindee. Values for DEP are arbitrarily assigned integers or letters. Two elements are assumed to converge in reference if they have matching values for DEP; this matching value is manipulated by the assignment function at LF to map both the binder and bindee to the same entity in the evaluation context.

The DEP feature is similar in many ways to Hicks (2009)'s VAR feature and Adger and Ramchand (2005)'s ID feature on nominals. But there are some significant differences both in conception and implementation. The most important of these is that DEP does *not* characterize all nominals. Hicks, for instance, proposes that the VAR feature encodes the information necessary to determine whether a variable co-varies with another or not, upon entering the derivation. Such information is necessary to encode differences between bound variable and deictic interpretations of ϕ-featurally identical pronouns, for instance. In my system, the DEP feature is only assumed to be a part of any two elements A and B that are involved in a syntactic dependency which will result in the interpretation that A binds B (or vice-versa).

The intuition behind this conception of things is that *syntax doesn't care about reference, only about (syntactically-derived) coreference.*[2] If all DPs bore features indicating information about their reference (even if this is only the relative reference conferred on a nominal upon its being merged in the derivation, in order to distinguish it from another nominal in the structure, as Hicks' VAR feature is conceived to be, and not absolute reference) – we would expect to see its syntactic effects in other phenomena besides binding, like agreement or intervention effects. But this is never the case.[3]

[2]"Coreference" is a theoretically loaded term which is often used to distinguish variable binding from other types of referential match. This is explicitly not the sense in which I am using the term: all I mean by "coreference" is that two entities have the same reference, regardless of how they came by it.

[3]Switch-reference and obviation patterns in certain languages, which require explicit referential disjointedness between two designated linguistic elements, might be potential instances of syntactically implemented non-coreference. However, the syntactic derivation of referential disjointedness is a general problem for *all* Agree-based theories of reference. Specifically, positing the existence of special referential features like VAR or DEP on all nominals will not help to *syntactically* derive non-coreference, as I show. Furthermore, switch reference and obviation effects – unlike Condition B effects for locally bound pronouns, for instance, which is robustly attested in the world's languages – are not universally attested linguistic phenomena (Stirling 1993). Extensive research, which is well outside the scope of this dissertation, must be undertaken to see to what extent these phenomena need to be syntactically implemented in the first place and whether and how they correlate with other linguistic properties

One might argue that such features are necessary to encode *non*-coreference, as between a deictic pronoun and a c-commanding DP, on pain of violating Condition B. However, deriving Condition B effects syntactically is a challenge for most Agree-based theories because it involves forcing featural *dis*agreement between a pronoun and a locally c-commanding DP, which is an operation that is not in the standard Minimalist toolbox. Hicks, in particular, has a problem precisely because he assumes that *all* DPs have a VAR feature. The assumption that it's unvalued on anaphors correctly forces them into an Agree relationship with their antecedents, deriving Condition A. The assumption that it's valued on deictic pronouns and R-expressions obviates agreement between two such mutually local DPs. Thus, binding via Agree is blocked, as desired. However, accidental coreference is still not blocked, since these features are given arbitrary values for VAR upon entering the derivation. Indeed, Hicks' system generates, in arbitrary cases, syntactically encoded accidental coreference which thus requires the assumption of additional grammatical mechanisms – a type of (Neo-)Gricean economy principle for instance – to filter them out. Independent evidence that Conditions B and C are sensitive to phonological (Hicks 2009) and semantico-pragmatic constraints (Reinhart 1983a, Reinhart 1983b, Heim 1998) further supports the idea that neither is derived by narrow syntactic operations. Thus, there is no evidence that non-coreference is syntactically encoded.

My conception that the only syntactic dependency in long-distance binding relations is that holding between the anaphor and an anaphoric center, rather than between the anaphor and its antecedent nominal (directly or indirectly, by mediation via functional heads), is crucial in allowing me to maintain the position that only these elements have a DEP feature. Hicks' feature system is used primarily to deal with local binding configurations in languages like English, where he argues that the antecedent participates in an Agree relationship with the anaphor in its phase (the presence of mediating functional heads doesn't deter from this point). This means, for Hicks, that the antecedent nominal must itself have a valued VAR feature in order to value the unvalued VAR feature on the anaphor. Given the long-distance and logophoric binding patterns above, I have, however, explicitly argued that the relationship between the antecedent and the perspectival center (and, by transitivity, the anaphor) cannot be structural. The antecedent nominal thus does not itself enter into a syntactic Agree relationship for binding and there

unique to these languages.

are no implications for its feature specifications.

Based on this discussion, I propose that the DEP feature is defined only on the perspectival center and the anaphor. It starts out valued on the perspectival center and unvalued on the anaphor – the simplest way to capture the intuition that it is the anaphor that is referentially defective.[4] The anaphor thus probes upward to get this feature valued by the perspectival center. Why should the perspectival center alone be privileged to host a valued DEP feature? As we have seen, Koopman and Sportiche (1989) and Baker (2008) have argued that discourse-pertinent information such as mental perspective and evidentiality are encoded on functional heads in the left-periphery of phases; they also show that such heads host a null operator in their specifier which is the real binder of the anaphor. One might surmise, therefore, that the DEP feature is one that such an element hosts by virtue of the syntactico-semantics associated with its position in [Spec, PerspP].

I thus define the DEP feature as follows:

(87) **The DEP feature:**

i. A DEP feature marks two DPs X and Y that are in a syntactic binding dependency with one another.

ii. DEP takes integers or letters as value. The assignment function maps these values to salient entities in the evaluation context. Two elements with matching DEP values will thus denote the same entity in the evaluation context and are construed to be in a binder-bindee relationship with one another.

iii. An anaphor is a nominal with an unvalued DEP feature – this is the syntactic correlate of anaphoricity; the operator in the specifier of the PerspP is a nominal with a valued DEP feature.

iv. The anaphor may have one or more ϕ-features in addition to the DEP feature, some of which may themselves (but need not) be unvalued.

v. ϕ-features, if any, constrain the domain of mapping possibilities for the reference index at LF, but don't (directly) have

[4] I will follow Hicks (2009)'s argumentation with respect to his VAR feature in assuming that a linguistic element is not born with its DEP-feature pre-valued. Rather, it is born with an instruction to the grammar to assign an *arbitrary* value to its DEP-feature as soon as it is merged in the structure. The only condition on value-assignment is that the value may not already have been assigned to another DP in the local phase. Thus it is, in some sense, impossible to know what the value of DEP on a DP can be before it is merged in the structure. See Hicks (2009) for arguments as to why this conception of things doesn't violate the Inclusiveness Condition.

anything to do with flagging a nominal as an anaphor in the syntax.

5.4 Other features on *ta(a)n*

In addition to the DEP feature, I will assume, in line with Kratzer (2009) and others, that *ta(a)n* has unvalued ϕ-features that must be checked in the course of the derivation. As we have seen, there is robust crosslinguistic evidence that anaphors cannot trigger ϕ-agreement, a descriptive generalization that assuming unvalued ϕ-features on *ta(a)n* captures since this will ensure that the anaphor cannot itself value these features on its clausemate T head (see also Kratzer 2009, for discussion of this point). However, in contrast to those who argue that ANAPHORICITY $\leftrightarrow$ ϕ-DEFICIENCY, I do not buy into the view that ϕ-unvaluedness is the property that defines anaphoricity on *ta(a)n*: rather, it is the unvalued DEP-feature that defines an anaphor. In other words, we could assume instead that *ta(a)n* has no ϕ-feature *attributes* at all, and my account would still carry over essentially unchanged. The advantage of assuming that *ta(a)n* has unvalued ϕ-features is that it makes it possible to claim that all agreement goes through the subject position in Tamil, regardless of whether the subject is born with its ϕ-features or itself gets them valued in the course of the derivation. This is a point I discuss in greater detail in Part III. To sum up, therefore, *ta(a)n* has the following feature-specification: [DEP: __, P: ___, NUM: ___, G: ___].[5]

In the next section, I will walk through some sample derivations for well-formed structures manifesting the core properties of long-distance binding in Tamil. Recall, again, that many of these properties don't have anything to do with the syntactic dependency between the anaphor and perspectival center in the model developed here. I.e. the dependency between the anaphor and perspectival center is syntactically local and minimal and proceeds via Agree between the anaphor (Probe) and operator in [Spec, PerspP] (Goal) for the DEP feature. This much is constant regardless of whether the antecedent is several clauses away or just a single clause away, whether the minimal phase containing the anaphor has been selected by a superordinate attitude predicate or is an adjunct, or whether the sentence involves backward binding, logophoricity, or long-distance binding with a structurally represented antecedent. This is the attractiveness of the model: it reduces the burden on the syn-

[5]Presumably it has a case-feature as well. I do not represent case-features here for purposes of brevity.

tactic computation by relegating the more tendential and non-structural aspects of long-distance binding to other modules of the grammar.

I will now present detailed syntactic derivations only for a subset of the sentences discussed in this series of chapters, specifically for those where the derivation might differ slightly due to the category of the minimal phase containing the anaphor or to the structural position of the anaphor itself. In particular, I will work through:

- Binding of object *ta(a)n* across single and multiple clausal boundaries.
- Logophoric and long-distance binding of subject *ta(a)n*.
- Binding into adjunct PPs.

5.5 Long-distance binding of object *ta(a)n*

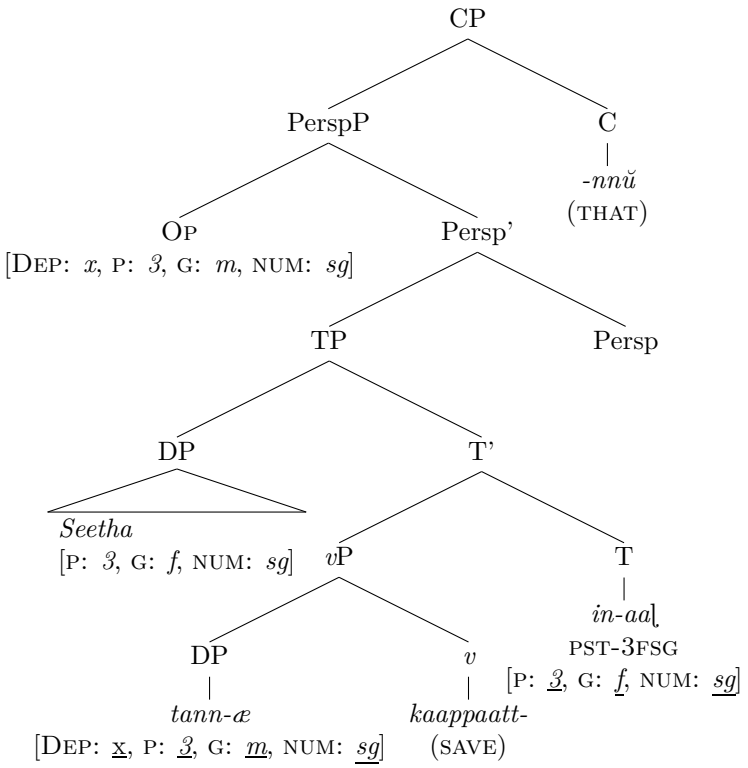
Let us start with long-distance binding of object *ta(a)n* across single and multiple clausal boundaries. Consider the sentence below, (repeated from (13):

(88) [$_{CP}$ Raman$_j$ Anand-kiʈʈæ$_k$ [$_{CP}$ Seetha$_l$ tann-æ$_{\{i,j,*k,*l\}}$
Raman[NOM] Anand-ALL Seetha[NOM] ANAPH-ACC
kaappaatt-in-aaɭ-ŭnnŭ] so-nn-aan-nnŭ] Krishnan$_i$
save-PST-3FSG-COMP] say-PST-3MSG-COMP Krishnan[NOM]
paar-tt-aan.
saw-PST-3MSG
"Krishnan$_i$ saw [$_{CP}$ that Raman$_j$ told Anand$_k$ [$_{CP}$ that Seetha$_l$ saved him$_{\{i,j,*k,*l\}}$.]]"

Here is the tree-structure for the CP containing the anaphor – the only part of the sentence that is involved in the syntactic aspect of binding, regardless of the actual antecedent of *ta(a)n*, crucially before the application of the relevant Agree operation between the anaphor and the pronominal operator in [Spec, PerspP]:[6]

[6]Some clarificatory points about the trees are in order. The tree structure here and those following are essentially abbreviated. For instance, the *v*P internal structure is not really depicted. Here and elsewhere in the trees, valued features, as on *ta(a)n* and T, are underlined to differentiate them from inherent features. However, this is a purely visual mnemonic for convenience and should not be treated as a featural diacritic of any sort. Finally, note that the trees represent the state of the syntactic derivation after Agree has already taken place.

(89)



ta(a)n enters the derivation with unvalued DEP and ϕ-features. As such, it is a Probe within the minimal CP phase to get these features checked. The closest element that c-commands it and has a valued DEP feature as well as valued ϕ-features is the operator in [Spec, PerspP] in the CP. *ta(a)n* thus enters into an Agree relationship with this operator, which results in the operator valuing its DEP feature as [DEP: x] and its ϕ-features as 3MSG, as indicated.[7] At LF, the matching [DEP: x] feature on both the operator and the anaphor results in them getting identified as a binder-bindee pair. Following Heim and Kratzer (1998), I assume that the choice of binder vs. bindee is made on the basis of asymmetric c-command. Here, the operator asymmetrically c-commands the anaphor; as such, the operator is the binder, and the anaphor, the bindee.

We already know from our discussion of factors conditioning potential antecedence that the matrix subject *Krishnan* and intermediate superordinate subject *Raman* are the only two DPs that satisfy the potential antecedence condition; the entities, Anand and Seetha, denoted by the

[7] We may need to assume, here and elsewhere, that *ta(a)n* first moves to [Spec, *v*P], so that it is at the edge of the lower phase and can be visible to the operator in [Spec, PerspP]. In Part II, I will argue that, in certain cases, there is a PerspP projected just above *v*P/VoiceP. However, the basic components of the derivation will not be affected by this.

other DPs *Anand* and *Seetha*, respectively,[8] will be represented in the range of the assignment function but will not satisfy the potential antecedence condition introduced as a partial function, since they don't have the relevant perspectival semantics with respect to the minimal CP containing the anaphor. Thus, x will not be mapped to Anand and Seetha. The linguistic representations of Raman and Krishnan in the structure also happen to satisfy the ϕ-feature requirement of being 3MSG, thus g may map x to either of these entities. The choice between them ultimately depends on speaker-intent. If $x \rightarrow$ Raman, we get the effect of binding across a single clausal boundary. If we get $x \rightarrow$ Krishnan, we derive the effect of binding across multiple clausal boundaries.

5.6 Logophoric and long-distance binding of subject *ta(a)n*

Consider the sentences below. (90) involves logophoric reference of *ta(a)n*. (91) reframes the pair of sentences in this example as a single complex clause involving long-distance binding of subject *ta(a)n*:

(90) Raman$_i$ Krishnan-kiṭṭæ$_j$ polamb-in-aan. Taan$_i$
Raman[NOM] Krishnan-ALL complain-PST-3MSG. ANAPH[NOM]
vaaɻkkæ-læ rombæ kašṭappaṭṭ-irŭ-kk-aan.
life[ACC] very difficulty-felt-be-PRS-3MSG.
"Raman$_i$ complained to Krishnan$_j$. He$_{\{i,*j\}}$ had suffered very much in life."

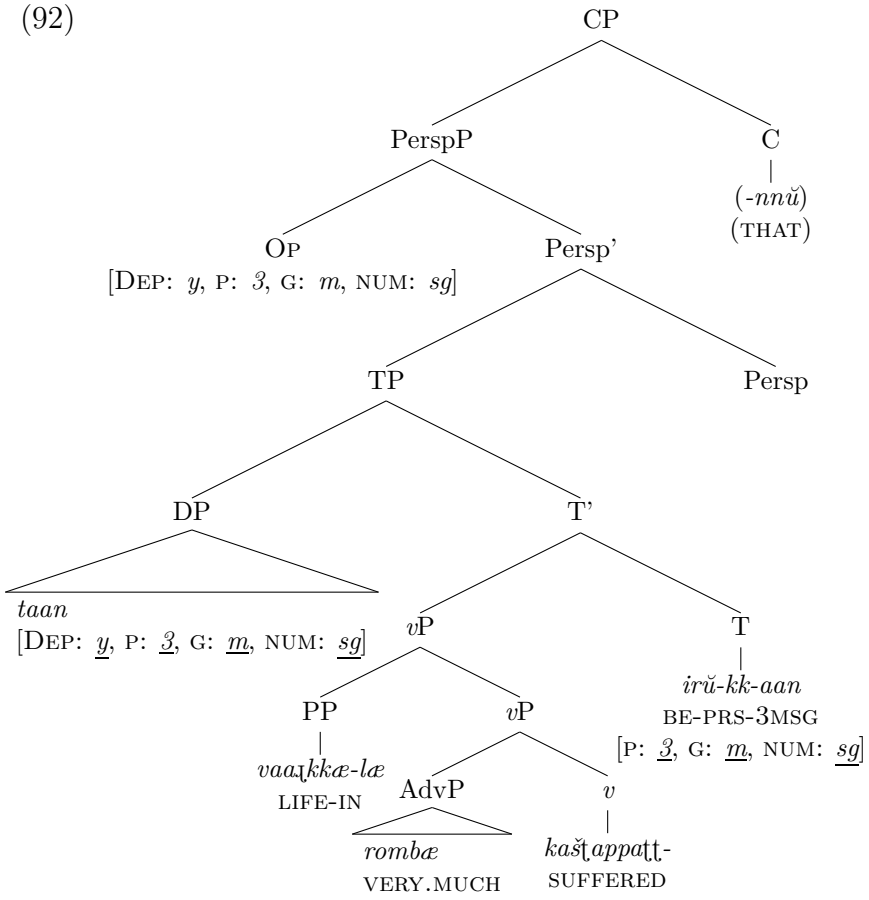
(91) [$_{CP}$ Raman$_i$ Krishnan-kiṭṭæ$_j$ [$_{CP}$ taan$_{\{i,*j,*k\}}$ vaaɻkkæ-læ
Raman[NOM] Krishnan-ALL ANAPH[NOM] life[ACC]
rombæ kašṭappaṭṭ-irŭ-kk-aan-nnŭ polamb-in-ad-æ]
very difficulty-felt-be-PRS-3MSG complain-PST-3NSG-ACC
Maya paar-tt-aaḷ.
Maya[NOM] see-PST-3FSG.
"Maya$_k$ saw/observed [$_{CP}$ [$_{DP}$ Raman's$_i$ complaining to Krishnan$_j$
[$_{CP}$ that he$_{\{i,*j,*k\}}$ had suffered very much in life.]]]"

As before, the only relevant structure for the syntactic component of binding is the minimal CP containing the anaphor. This is the same for

[8] As is probably already clear, I am using the italicized formatting of these names to indicate the linguistic representations of the actual persons in the evaluation context. Their non-italicized analogs represent the individuals in the evaluation context.

both (90) and (91) (modulo the presence of the overt complementizer in the latter). Here is its tree-structure:

(92)



The anaphor *ta(a)n* probes upward to get its DEP feature valued. The closest c-commanding Goal is the operator in [Spec, Persp]. Thus, *ta(a)n* enters into an Agree relationship with this Goal with the result that it ends up with the following feature-specification at the end of the syntactic derivation: [DEP: *y*, P: *3*, G: *m*, NUM: *sg*]. At LF, as usual, the matching *y* feature on the operator and anaphor results in them being construed as a binder-bindee pair under semantic variable binding. The asymmetric c-command relationship between them results in the c-commanding operator variable-binding the anaphor. At LF, the assignment function *g* will try to map *y* to one of the individuals in its range. Crucially these individuals will be selected not just from the sentential structure but also from the salient discourse. This is an assumption that standard views on reference assignment have to make anyway, in order to be able to deal with deictic reference for pronouns and R-expressions at LF.

In the sentence in (90), the range of the assignment function g will consist of the set: {Raman, Krishnan}. The mapping of the DEP-feature value to one of these individuals will, however, be restricted such that the individual chosen must satisfy the presuppositions placed by the ϕ-features on the operator. The second requirement is that the chosen individual fulfill the thematic, semantic and discourse requirements for potential antecedence, also implemented as partial functions on the denotation of the operator. Both *Raman* and *Krishnan* fulfill the ϕ-feature requirement, since both are specified 3MSG. However, only *Raman* denotes a mental perspsective holder with respect to the situational predication containing the anaphor – thus only it satisfies the potential antecedence presupposition on its denotation. This yields: $y \rightarrow$ Raman by g with the result that *ta(a)n* refers to Raman "logophorically", as desired.

In the long-distance embedded sentence in (91), the range of the assignment function g is: {Maya, Raman, Krishnan} (among many others). The DPs denoting Raman and Krishnan fulfill the 3MSG ϕ-feature requirements on proper assignment; The DP that denotes Seetha, being 3FSG, doesn't fulfill the ϕ-feature requirements, however, and is disqualified on these grounds, despite fulfilling the potential antecedence requirement. The DP *Raman* also fulfills the requirements on potential antecedence since it denotes an individual that holds mental perspective toward the minimal event predication containing the anaphor; the DP *Krishnan*, as the GOAL argument, doesn't (in the pragmatically unmarked case) qualify as a potential antecedent. The end result is that *Raman* is the only possible antecedent for subject *ta(a)n* in this structure, in the pragmatically unmarked case.

One final point is due. The embedded agreement on the clausemate verb in (92) is due to an Agree relationship between T and the operator in [Spec, Persp] in the narrow syntax. T starts out with unvalued PERSON, NUMBER and GENDER features. It tries to get these valued by the clausemate subject in [Spec, TP], as usual but in this case, the subject is *ta(a)n* which also has unvalued ϕ-features. I propose, in line with Pesetsky and Torrego (2007), that T and *ta(a)n* enter into a feature-sharing relationship for ϕ-features. These get simultaneously checked on both by the operator in [Spec, PerspP] in the local CP phase. At PF, the features on T get spelled out as *-aan*. We also have the desired result that the agreement under *ta(a)n* tracks the features of the antecedent of *ta(a)n*. I will discuss this agreement process in greater detail in Part III, but this is the basic idea.

5.7 Binding into adjunct PPs and DPs

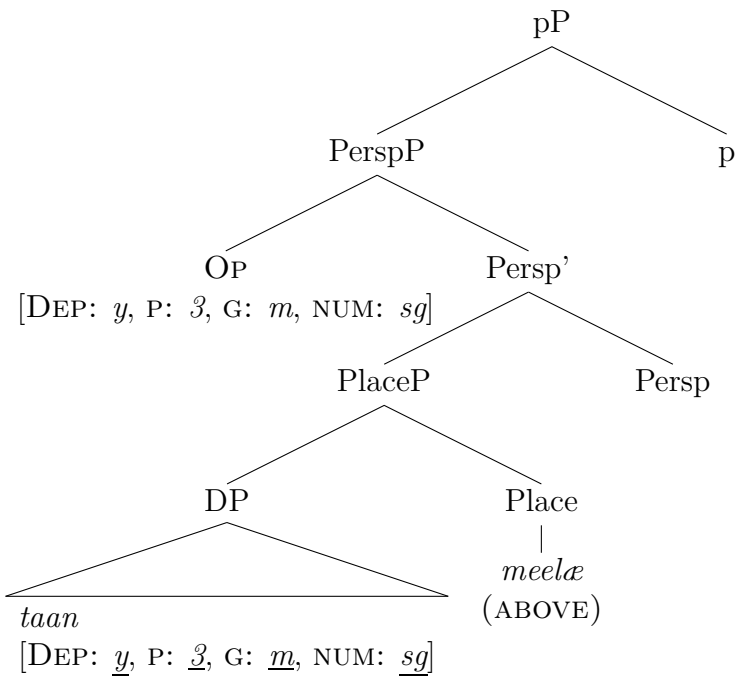
Consider the following binding structures involving long-distance binding of *ta(a)n* into a spatio-temporal PP and a possessor DP, respectively (the sentences are repeated from (56a) and (57a):

(93) Raman$_i$ [$_{PP}$ tan-akkŭ$_{\{i,*j\}}$ meelæ] orŭ plane-æ paar-tt-aan.
Raman.NOM ANAPH-DAT above a plane-ACC see-PST-3MSG
"Raman$_i$ saw a plane above himself$_{\{i,*j\}}$."

(94) Raman$_i$ [$_{DP}$ tann-oo$\underline{d}$æ$_{\{i,*j\}}$ mugatt-æ] ka$\eta\underline{\eta}$aa$\underline{d}$i-læ
Raman[NOM] ANAPH-GEN face-ACC mirror-LOC
paar-tt-aan.
see-PST-3MSG
"Raman$_i$ saw his$_{\{i,*j\}}$ face in the mirror."

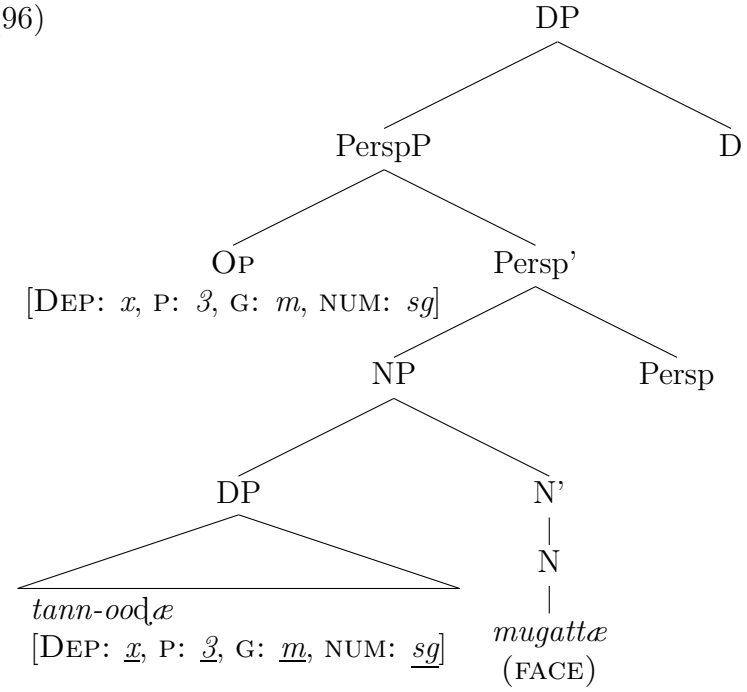
Once again, only the minimal PP, DP or CP phase containing the anaphor is relevant for the syntactic component of long-distance binding – in other words, the PP and DP containing *ta(a)n* in (93) and (94), respectively. Here are the tree structures for each:[9]

(95)



[9]The internal structure of the PP is based on that given in Svenonius (2008).

(96)



The anaphor *ta(a)n* in both structures starts out with unvalued DEP and ϕ-features. In (93)/(95), it ends up with the features: [DEP: *y*, P: *3*, G: *m*, NUM: *sg*] upon Agree with the operator in [Spec, PerspP] at the left-periphery of the PP phase. In (94)/(96), it ends up being featurally specified as: [DEP: *x*, P: *3*, G: *m*, NUM: *sg*] upon Agree with the operator in [Spec, Persp] at the left-periphery of the DP phase.

At LF, the oblique anaphor in (93) and the possessive anaphor in (94) are variable-bound by their respective operators in [Spec, PerspP], since this operator has the same value for DEP and also asymmetrically c-commands the anaphor in each. The assignment function g in (93) has a range consisting of Raman and potentially other salient individuals from the discourse, yielding e.g. {Raman, Bill, Anand, Maya, Seetha}. However, the DP denoting Raman is the only one that simultaneously satisfies the ϕ-feature restriction and potential antecedence restriction on assignment, by virtue of being specified 3MSG (like the operator) and bearing a spatial perspective towards the minimal PP containing the anaphor. Thus, the DEP value y is mapped to Raman with the result that its linguistic correlate *Raman* is construed as the antecedent of *ta(a)n* in (93). In (94), the assignment function might have the range: {Raman, Krishnan, Vishnu, Sudha, Champa}, among others. Here again, the DP that denotes Raman is the only one that simultaneously satisfies the presuppositional requirements on ϕ-matching and potential antecedence. Thus,

$x \rightarrow$ Raman, with the effect that *Raman* is construed as the antecedent of *ta(a)n* in (94).

There are two issues we haven't yet discussed. This is the Ban on Clausemate Subject Antecedence which, we have proposed, is really an anti-locality condition triggered by conditions of wellformedness on the representation of perspective. In this connection, we must also discuss the issue of binding in psych-predicate structures where this anti-locality effect seems to be somehow circumvented. The other issue is the 3rd-person antecedence of *ta(a)n* which clearly requires stating something special, since we are explicitly stating that *ta(a)n* has no valued ϕ-features (thus, no person feature). We will defer discussion of the Ban on Clausemate Subject Antecedence and the apparent exception to this rule in psych-predicate structure to Part II which deals with the issue of local binding in Tamil. The 3rd-person antecedence restriction on *ta(a)n* will be discussed in detail in Part III where the nature of the ϕ-feature specification of *ta(a)n* will be revisited in the context of indexical shift under speech predicates.

Part II

Tamil local binding and the "*kol*" morpheme

Chapter 6

Issues in local binding: introducing "*ko*⌊"

6.1 Introduction

This chapter extends the insights on long-distance binding in Tamil from Part I to the case of the local binding of *ta(a)n*. We have already addressed one aspect of this phenomenon, namely the fact that, in the typical case, a clausemate subject of *ta(a)n* may not serve as its antecedent. Thus, in the sentence in (97) below, the object anaphor *tannæ* may take the matrix subject *Raman*, but not its own clausemate subject *Krishnan*, as its antecedent:

(97) Raman$_i$ [$_{CP}$ Krishnan$_j$ tann-æ$_{\{i,*j\}}$ kaɳɳaaḍi-læ paar-tt-aan-nnŭ]
Raman[NOM] Krishnan[NOM] ANAPH[NOM] mirror-LOC see-PST-3MSG-COMP
nene-tt-aan.
think-PST-3MSG
"Raman$_i$ thought [$_{CP}$ that Krishnan$_j$ saw himself$_{\{i,*j\}}$ in the mirror.]"

At the same time, we have seen that *ta(a)n may* be bound by a clausemate subject in certain special cases, like in psych-predicate structures. This is illustrated in (98) below:

(98) Raman-ŭkkŭ$_i$ tann-æ$_{\{i,*j\}}$ piḍikka-læ.
Raman-[DAT] ANAPH-ACC like.INF-NEG
"Raman$_i$ disliked himself$_{\{i,*j\}}$."

What is it about the psychiness of an EXPERIENCER co-argument, like that in (98), that allows it to be a perspective-holder of the *ta(a)n*-eventuality (which is trivially also its own)? Conversely, what is it about

a non-psychy co-argument, like that in (97) that prevents it from being a perspective-holder towards the *ta(a)n*-eventuality (which is trivially also its own eventuality)?

These are intriguing questions and they are made more intriguing by the fact that local binding of *ta(a)n* *is* possible even in non-psych-predicate structures under certain special conditions. In sentences that don't involve a psych-predicate, local binding of *ta(a)n* is effected by the presence of a morpheme *ko*ɭ which is suffixed onto the verbal stem.[1] Thus, in (99) below, the presence of *ko*ɭ on the embedded verb appears to facilitate local binding of the object anaphor *tannæ* by its clausemate embedded subject *Krishnan* [2]

(99) Raman [$_{CP}$ Krishnan$_j$ tann-æ$_j$ kaɳɳaaɖi-læ
Raman[NOM] Krishnan[NOM] ANAPH-ACC mirror-LOC
paarttŭ-ko-ɳɖ-aan-nnŭ] nene-tt-aan.
see-*ko*ɭ-PST-3MSG-COMP think-PST-3MSG
"Raman thought [$_{CP}$ that Krishnan$_j$ saw himself$_j$ in the mirror.]"

In the face of data such as that in (99), it is tempting to analyze *ko*ɭ as a type of "SELF-"morpheme, much like Norwegian *selv* or Dutch *zelf* which have been argued to to "reflexive-mark" – essentially detransitivizing – a predicate, thus facilitating local binding. At the same time, such an analysis seems rather implausible. After all, there are two separate pieces that seem to contribute to the local antecedence of *ta(a)n*: *ta(a)n* itself and *ko*ɭ. In contrast, standard detransitivizing analyses of reflexivity involve a single element that marks both types of phenomena. Furthermore, such an analysis seems inconsistent with the facts pertaining to the local binding of *ta(a)n* in psych-predicate structures. Empirical evidence argues against such a position as well: as I will show, it is difficult to maintain the position that *ko*ɭ is a reflexive-marker because its presence is neither universally necessary nor sufficient for the encoding of local binding relations in Tamil. As a way to understand the larger puzzles surrounding the local binding phenomenon in Tamil, we will devote our attention to the investigation of *ko*ɭ – a morpheme that, incidentally, is characterized as being notoriously hard to describe in descriptive grammars and typological studies on Tamil (see e.g. Steever 2005).

[1] This is true for my own dialect of Tamil. There is some microvariation in this area, as I will discuss in Section 9.4.

[2] I have merely glossed *ko*ɭ as *ko*ɭ – since its meaning is not yet clear and its determination is our major focus.

The fact that local binding is possible in psych-structures in the absence of *ko*ɭ (as in (98)) shows that the presence of this morpheme is not necessary for local binding. However, there are also uses of *ko*ɭ that have nothing to do with local binding whatsoever, which shows that *ko*ɭ is not sufficient for local binding either. For instance, *ko*ɭ may be *optionally* suffixed to the unaccusative predicate *taræ* (OPEN) as the the unaccusative minimal pairs in (100) and (101) show:

(100) Kadavŭ tara-nd-adŭ.
door[NOM] open-PST-3NSG
"The door opened."

(101) Kadavŭ tarandŭ-ko-ɳɖ-adŭ.
door[NOM] open-*ko*ɭ-PST-3NSG
"The door opened-*ko*ɭ."

More strikingly, *ko*ɭ can mark regular transitive verbs in structures where *ta(a)n* is entirely absent, as in (102) below:

(102) Raman Krishnan-æ aɖittŭ-ko-ɳɖ-aan.
Raman[NOM] Krishnan-ACC hit-*ko*ɭ-PST-3MSG
"Raman hit-*ko*ɭ Krishnan."

As such, a more intuitive and consistent analysis would be to say, informally speaking, that the presence of *ko*ɭ on a verb essentially turns it into a more "psych-y" version of itself, essentially equivalent to a psych predicate. The external argument of the predicate is still considered an argument of this psych-y version and, as such, qualifies as a potential antecedent for *ta(a)n* in much the same way as *Raman* in (98). This, indeed, is the position I will end up arguing for in this series of chapters.

In doing so, I will also clarify the notion of "perspective", used in the linguistic sense. Specifically, I will show that in order for an individual to hold a perspective toward a predicational structure, that individual may not be properly contained (i.e. embedded) inside that predicate. This is a structural restriction that must be added as a wellformedness condition on perspective-holding – thus, on potential antecedence of *ta(a)n*. This restriction, indeed, will be seen to be behind what we have been calling the Ban on Clausemate Subject antecedence (in non-pysch predicate structures). We will see that the addition of *ko*ɭ allows a co-argument of *ta(a)n* to "step outside" its eventuality so as to be able to hold a perspective toward it, thus by extension qualifying as a potential antecedent for *ta(a)n*, both in the case of psych-predicate structures like (98) and those like (99). In the process, I will also show that the notion of perspective or point-of-view marks not only binding relations but also unaccusativity –

interesting given that unaccusativity been observed to go hand-in-hand with reflexive marking in many languages (Reinhart and Reuland 1993, Embick 2004b, Chierchia 2004, Levin and Rappaport-Hovav 2005, among others). Most importantly, it will also be seen that the model of long-distance binding developed in Part I can be extended, with very minimal variations, to the local-binding of *ta(a)n*.

Much of the set of chapters here in Part II will be devoted to nailing down a precise syntactico-semantics for *ko*⌊ which properly captures both its licit and illicit use in reflexives, unaccusatives and non-reflexive transitives. To this end, I investigate the compatibility and effects of this morpheme with the different lexical-conceptual verb-classes delineated in Levin (1993) supplemented by judgments from native speaker respondents of my survey. On the basis of this, I develop a precise denotation and argument structure for *ko*⌊. In particular, I will show that *ko*⌊ is a semi-functional restructuring predicate in the sense of Wurmbrand (2001), one that introduces a Perspectival Phrase (PerspP) in its complement, much like a propositional predicate has been argued to do in Part I. However, I will show that there are some additional syntactico-semantic restrictions on *ko*⌊. Foremost among these is the idea that *ko*⌊, unlike an applicative, doesn't increase the valency of the predicate in its complement. Rather, it allows the highest argument of the main eventive predication in its scope to appear above the minimal PerspP as its specifier – in other words, it behaves like a raising or control verb.

6.2 Setting the parameters: what *ko*⌊ is and isn't

Here, I present an overview of the distributional and interpretational properties of *ko*⌊ in Tamil:

(i) In the pragmatically unmarked case, *ko*⌊ tends to be absent in structures involving the long-distance binding of *ta(a)n*.

(ii) In the pragmatically unmarked case, *ko*⌊ tends to be present in structures involving the local binding of *ta(a)n*.

(iii) *ko*⌊ also frequently occurs in unaccusative structures.

Properties (i) and (ii) thus reflect a dichotomy between local and long-distance binding in Tamil and suggest that *ko*⌊ is nothing other than a local reflexivity marker, much like the SELF-morpheme in languages like

those of the Germanic language family. This impression is only strengthened by (iii), since unaccusatives and reflexives have been shown to be identically marked in a wide range of typologically unrelated languages – an observation that has been explained by proposing that reflexives themselves involve a sort of unaccusative argument structure.

In the sections below, I will first briefly review the basic patterns of *ko*ɭ in local binding and unaccusative structures. I will then move on to discussing whether there is any credence to the idea that *ko*ɭ is just a marker that shows up in unaccusative structures and, by extension, whether it makes sense to view local binding sentences in Tamil as involving some sort of unaccusative structure underlyingly. In this context, I will discuss an analysis by Lidz (2004) which proposes an analysis along these lines for *ko*ɭ in the related Dravidian language Kannada. However, based on independent empirical evidence from non-reflexive transitives and *ko*ɭ-less unaccusatives in Tamil, I will conclude that such an approach is not viable here. The facts from (i)-(iii) will thus have to be explained in some other way.

6.2.1 *ko*ɭ in reflexives and unaccusatives: basic paradigms

*ko*ɭ doesn't seem to affect long-distance binding in Tamil. Rather, in the pragmatically unmarked case, long-distance binding is established in the absence of *ko*ɭ on the relevant predicate, as illustrated in (103) below:

(103) Krishnan$_i$ [$_{CP}$ Raman$_j$ tann-æ$_{\{i,*j\}}$ aḍi-tt-aan-nnŭ]
Krishnan[NOM] Raman[NOM] ANAPH-ACC hit-PST-3MSG-COMP
tavaraagæ nene-tt-aan.
wrongly think-PST-3MSG
"Krishnan$_i$ wrongly thought [$_{CP}$ that Raman$_j$ hit himself$_{\{i,*j\}}$.]"

When it comes to the local binding of *ta(a)n*, however, the presence of *ko*ɭ seems to be the norm rather than the exception. Thus, the *ko*ɭ-less sentence in (104) is ungrammatical;[3] the standard way to express local binding is with the addition of *ko*ɭ to the main predicate, as shown in (105):

(104) * Raman$_i$ tann-æ$_i$ aḍi-tt-aan.
Raman[NOM] ANAPH-ACC hit-PST-3MSG

[3]There is a group of speakers identified in my survey who allow local binding in the general case without *ko*ɭ, i.e. who find sentences like 104 grammatical. I will discuss this dialect in Section 9.4.

"Raman$_i$ hit himself$_i$." (Intended)

(105) Raman$_i$ tann-æ$_{\{i,*j\}}$ aɖittŭ-ko-ɳɖ-aan.
Raman[NOM] ANAPH-ACC hit-*ko*ɭ-PST-3MSG
"Raman$_i$ hit himself$_{\{i,*j\}}$."

This suggests that *ko*ɭ facilitates the *local* binding of the anaphor *ta(a)n*, rather like the "SELF" morpheme – as in the minimal pairs from Dutch below (Reinhart and Reuland 1993):

(106) * Jan$_i$ haat zich$_{\{i,*j\}}$.
Jan hates ANAPH[ACC]
"Jan$_i$ hates himself$_{\{i,*j\}}$." (Intended)

(107) Jan$_i$ haat zich-zelf$_{\{i,*j\}}$.
Jan hates ANAPH[ACC]-SELF
"Jan$_i$ hates himself$_{\{i,*j\}}$."

This impression is, if anything, strengthened by the observation, noted above, that *ko*ɭ frequently marks unaccusative predicates:

(108) Kadavŭ muuɖi-ko-ɳɖ-adŭ.
door[NOM] close-*ko*ɭ-PST-3NSG
"The door shut-*ko*ɭ."

(109) Paanæ oɖænʤŭ-ko-ɳɖ-adŭ.
pot[NOM] break-*ko*ɭ-PST-3NSG
"The pot broke-*ko*ɭ."

As mentioned above, this is because unaccusative and reflexive paradigms in languages of the Slavic (Medová 2009), Germanic (Schäfer 2008), and Romance (Reinhart and Siloni 2004) language families, among others, are often marked with the same types of morpheme, motivating valency reduction analyses which claim that reflexives and unaccusatives share an identical structural subcomponent. There are two main schools of thought for this idea within the Lexicalist and Non-lexicalist traditions. The first proposes that reflexivization is essentially a process of detransitivization – via argument-suppression or θ-absorption – in the lexicon (Reinhart 2000) or in the syntax (Reinhart and Siloni 2004). The second proposes that reflexive structures have a type of unaccusative structure to begin with which in turn feeds morphophonological similarities between the two: Embick (2004b)'s analysis of passives, reflexives, and unaccusatives in Greek within the Distributed Morphology framework (Halle and Marantz 1993) is an example of such an approach.

However, both approaches share the intuition that reflexives share important structural similarities with unaccusatives accounting for the

many surface similarities between them. Given that *kol* marks both unaccusatives and reflexives in Tamil, it is thus important to ascertain whether this means that local reflexives in Tamil also involve an unaccusative structure – a similarity which is, furthermore, morphologically instantiated as "*kol*".

6.2.2 The unaccusativity approach to local reflexivity

Consider the following sets of unaccusative and local reflexive structures (formatting mine):

(110) GERMAN *sich*:

a. Hans$_i$ hat sich$_{\{i,*j\}}$ ins Gesicht geschlagen.
Hans[NOM] has ANAPH[ACC] in.the face hit.PTCP
"Hans$_i$ hit himself$_{\{i,*j\}}$ in the face."

b. Die Tür hat sich geschlossen.
The door has ANAPH[ACC] closed.
Lit: "The door closed."

(111) CZECH *se* (Medová 2009):

a. Madlenka$_i$ se$_{\{i,*j\}}$ učesala.
Madlenka[NOM] ANAPH[ACC] comb.FSG
"Madlenka$_i$ combed herself$_{\{i,*j\}}$."

b. Království se rozpadlo.
kingdom[NOM] ANAPH[ACC] disintegrated.NSG
"The kingdom disintegrated."

(112) GREEK non-active morphology (Alexiadou and Anagnostopoulou 2004):

a. I Maria htenizete.
the Maria[NOM] combs.NACT
"Maria$_i$ combs herself$_{\{i,*j\}}$."

b. I supa kegete.
the soup[NOM] burns.NACT
"The soup is burning."

(113) FRENCH clitic *se-* (Reinhart and Siloni 2004):

a. Jean$_i$ s$_{\{i,*j\}}$'est introduit á Paul.
Jean SE-is introduced to Paul.
"Jean$_i$ introduced himself$_{\{i,*j\}}$ to Paul."

b. Jean s'est évanoui.
Jean SE-is fainted.

"Jean fainted."

The data above consists of sentence-pairings showing standard ways to represent unaccusatives and reflexives in different languages – and shows that the morphemes that mark these phenomena are often one and the same. The fact that this commonality cuts across a whole range of typologically unrelated languages suggests that this is the result of principled syntactico-semantic similarities between unaccusativity and reflexivity that are encoded at the level of UG.

Analyses vary considerably with respect to how these similarities are to be formally derived and the extent to which the phenomena of unaccusativity and reflexivity can be conflated. With respect to the first point, proposals like those of Grimshaw (1982), Chierchia (2004) argue that reflexive and unaccusative structures have the same argument structure. They treat reflexivization as an arity-reducing operation which detransitivizes the transitive predicate to yield a Burzio (1986)-style unaccusative whose surface subject is actually the underlying object. However, approaches like those of Reinhart and Siloni (2004) argue that reflexive structures seem more unergative than unaccusative in nature (a point that Medová (2009) incidentally also makes) and propose instead that "the morphological similarity often attested between reflexives and unaccusatives is not due to their common argument structure, but to the basic operation at the heart of their derivation" (Reinhart and Siloni 2004, 159). The authors retain the idea that reflexivization is the result of a valency reduction operation which yields an unaccusative structure but propose that other empirical differences between unaccusatives and reflexives should be derived as a function of where in the derivation (lexicon vs. syntax) this operation occurs.

Within the non-Lexicalist tradition (Marantz 1984, Embick 2004b, Alexiadou and Anagnostopoulou 2004, Schäfer 2008), such an approach is fundamentally not viable, given that a generative lexicon is assumed not to exist. The relationship between unaccusatives and reflexives must thus be got at differently. Embick (2004b), examining "non-active" morphology that marks reflexives, passives, unaccusatives and middles in Greek, argues that this common morphological marking reflects a common syntactic subcomponent that is shared across these different structures. In particular, he argues that they all lack a position for the external argument in [Spec, vP]. Under Embick's analysis, local reflexives in Greek have the structure of passives, with the antecedent binding the anaphor from internal, rather than external, argument position. The anaphor itself is assumed to be more verb-y than nominal, either incorporated or

compounded with the main verb, with an interpretation of *self*-V. The reflexive structure thus lacks a [Spec, *v*P] and "the sole DP argument comes to be interpreted as agentive in a derivative fashion" (Embick 2004b, 146), just like the agent of a passive structure.

Working within the Late Insertion model of Distributed Morphology (Halle and Marantz 1993), Embick further proposes that morphological spell-out rules are sensitive to the *absence* of an external argument in the syntax; as such, the absence of this structural position in unaccusatives, reflexives, and passives is marked by (the same) special morphology, as formalized below:

(114) $v \leftrightarrow v\text{-}X/__$No external argument (where "-X" is the feature/signal associated with non-active morphology.)

Given the Y-modular architecture of the grammar, the PF module isn't sensitive to LF-semantic distinctions between unaccusatives, reflexives, and passives – such as those discussed in Reinhart and Siloni (2004). Thus, Embick's model is able to simultaneously account for the morphological similarities and syntactico-semantic differences between these different types of structures without having to resort to positing the same valency reduction mechanism in two different generative modules, as Reinhart and Siloni do.

6.2.3 An Embick-style analysis for Kannada *kol*: Lidz (2004)

Lidz (2001; 2004), Lidz and Williams (2005) offer an adaptation of Embick's model to capture the distribution of *kol* in Kannada, a Dravidian language closely related to Tamil. The *kol*-patterns in Kannada are largely similar to those in Tamil, so it is worthwhile to look at Lidz' analysis in some detail.

In unaccusative structures in Kannada, just like in Tamil, *kol* appears optionally suffixed to the unaccusative predicate, as shown in the minimal pair below:

(115) Huuvu udur-i-tu.
flower[NOM] wilt-PST-3NSG
"The flower wilted."

(116) Huuvu udur-i-ko-ŋḍ-itu.
flower[NOM] wilt-PST-KOL-3NSG
"The flower wilted."

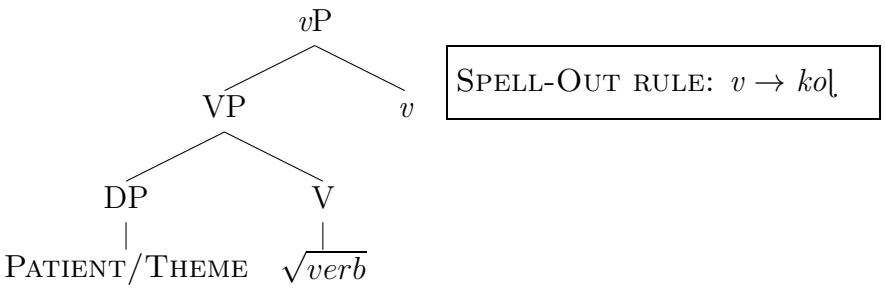
Based on this type of data, Lidz proposes, in line with Embick (2004b) that unaccusatives (with or without *koḷ*) involve a structure which lacks an external argument position in the syntax. Unlike Embick, however, Lidz argues that *koḷ*-unaccusatives alone involve an agent/causer semantics on (the specifierless) *v*, much like the passive in Embick's system, and labels *koḷ* a "monadic causative"; *koḷ*-less unaccusatives on the other hand, are like unaccusatives in Embick's model – the *v* head in such structures is not associated with an agentive semantics. Evidence for this contrast comes from Lidz' observation that *koḷ*-unaccusatives may co-occur with an adjunct external causer, but that their *koḷ*-less variants may not:

(117) * Gaal-ige huuv-u udur-i-tu.
wind-DAT flower-ACC wilt-PST-3NSG
"The flower wilted (due) to the wind." (Intended)

(118) gaal-ige huuv-u udur-i-koṇḍ-itu.
wind-DAT flower-ACC wilt-PST-KOL.3NSG
"The flower wilted (due) to the wind."

Lidz' argument-structure for *koḷ* is illustrated below – the agentive/causer semantics that is putatively associated with *koḷ* is claimed to be the result of the combined semantics of the V-*v* complex:[4]

(119)



The analysis, as well as that of Embick (2004b), does face some challenges: for instance, it is not clear how a spell-out rule is to be made sensitive to the *absence* of syntactic structure, especially if the missing

[4]To quote Lidz (2004, 12-13): "It is simply the relation between verbs [*v*-V] that creates the causative role. In this sense, the 'agent' role is not actually assigned by any syntactic mechanism to the [spec, *v*P]. Rather, causation implicates a causer and an NP found in this position is free to be interpreted as such. If there is no NP in that position, as in (23), then the event is still construed as causative."

element is a specifier – a point that both Lidz and Embick concede. It is also unclear to what extent the presence of an adjunct, such as that of the adjunct causer PP in (118), can be taken as evidence for thematic structure. Nevertheless, the analysis successfully derives the unaccusative structures in Kannada, given above.

However, the monadic causative account of *koḷ* in Kannada cannot be straightforwardly extended to local reflexives which, just like their Tamil counterparts, must involve *koḷ* in the default case:

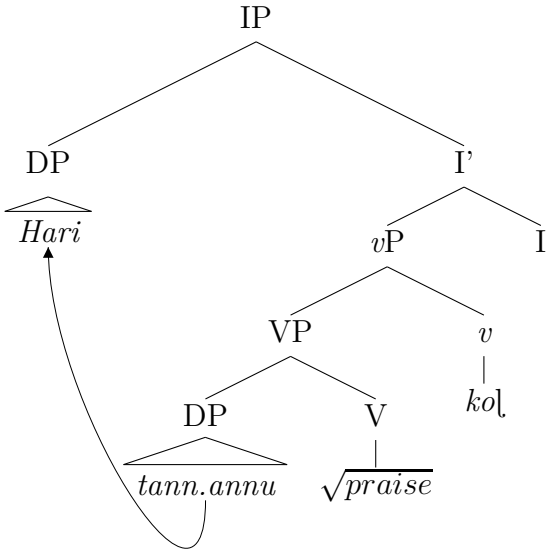
(120) * Hari tann-annu hogaḷ-id-a.
Hari[NOM] ANAPH-ACC praise-PST-3MSG
"Hari praised himself." (Intended)

(121) Hari tann-annu hogaḷi-ko-ṇḍ-a
Hari[NOM] ANAPH-ACC praise-*koḷ*-PST-3MSG
"Hari praised himself."

As we saw above, Embick was able to accommodate local reflexives in Greek under his specifierless *v*P account because (a subset of) these reflexives were seen to be adverbial rather than nominal – thus, local reflexive structures in Greek seemed to have only one DP argument to begin with, corresponding to the anaphoric antecedent. Other valency reduction analyses, such as that of Marantz (1984) for Romance reflexives involving *se-* were able to exploit the clitic-nature of this element to get it to end up in a local relationship with *v*. Neither option is possible for Kannada, however, where the reflexive is neither a clitic nor an adverb, but a full-on nominal argument of the verb. In other words, structures like (121) seem to be fully transitive – thus should not be able to feed *koḷ*-insertion, under the current analysis.

Lidz' solution is to claim that, although the structure in (121) doesn't start out with the right structural environment for *koḷ*-instertion (i.e. a specifierless *v*P), it *becomes* such an environment in the course of the derivation due to independent factors. He proposes that there is a binding chain-relation between the object anaphor *tann-annu* and the antecedent *Hari* in [Spec, *v*P]. In addition, *Hari* is posited to move from [Spec, *v*P] to [Spec, TP] for case reasons, creating an A-movement chain between these two positions. The net result is the following three-membered chain: [Spec, TP]$_{Hari}$–[Spec, *v*P]$_{t_{Hari}}$–DP$_{tannannu}$. Such a chain is, however, argued to be ill-formed because the medial link in [Spec, *v*P] has an unvalued case feature that cannot be checked. As such, Lidz proposes, this position is deleted entirely, yielding a specifier-less *v*P, which in turn feeds *koḷ*-insertion in *v*. The final structure looks like that in (122) below:

(122)



Although Lidz' analysis does differ from that of Embick (2004b) for local reflexives in Greek and Romance, it is clearly related. In Embick's analysis, a local reflexive structure doesn't have a [Spec, *v*P] at any point in the derivation – an analytic stance that's possible given that the anaphor itself, in languages like Greek, is not a DP that competes for an argument slot, but is a more "verb-y" element. Under Lidz' treatment, the structure starts out as fully transitive but *becomes* intransitive in the course of the syntactic derivation. Crucially, the structure that is ultimately input to PF for Spell-Out is equivalent for both – specifically, lacking an external argument position which feeds insertion of "reflexive" morphology.

I have devoted some space and time to discussing Lidz' analysis for *koL* because of its potential relevance for Tamil. But it should be borne in mind that Lidz' analysis as well as that of Embick (2004b) are similar in spirit to other unaccusative treatments of reflexivity, including those within the lexicalist tradition. Regardless of the differences, which are largely due to variations in the grammatical framework (lexicalist vs. non-lexicalist) assumed, these approaches all share the intuition that the similarities between unaccusatives, reflexives, and passives can be tied to the reduced valency of the predicate-classes involved in each. In other words, reflexivity and unaccusativity are both treated as species of voice phenomena.

Given this rich tradition and Lidz' own suggestions with respect to reflexivity and unaccusativity in Kannada, can this idea be extended to Tamil *koL*? Although theoretically appealing, it appears that the answer is negative.

6.3 Against Tamil *koɭ* as a marker of valency reduction

There is one clear reason why treating *koɭ* as a voice-marker in Dravidian – thus treating local reflexivity in this language as a form of unaccusativity – is misguided. In languages like Greek, and the languages of the Romance, Slavic, and Germanic families, there is a single element that signifies both reflexivity and unaccusativity. This makes it easy to treat local reflexivity as being the result of a detransitivizing operation where the reflexive ends up adjoined to the verb via clitic climbing or noun-incorporation or some other mechanism in these other languages. In contrast, in local reflexive structures in Tamil and Kannada, there are *two* distinct elements that contribute to local reflexivity: *koɭ* and the anaphor *ta(a)n*. This proved to be a problem for Lidz (2004), as we just saw, who was forced to delete the external argument from [Spec, *v*P] to allow for *koɭ*-insertion in a local binding structure.

In addition, there are two clear empirical arguments that mitigate against a reduced valency treatment of *koɭ*-structures in Tamil, both of which essentially show that the presence of *koɭ* is entirely orthogonal to the valency of the predicate it occurs suffixed onto:

(i) *koɭ* attaches to verb-stems that are *already* marked as intransitive. Such verbs remain intransitive after *koɭ*-suffixation.

(ii) *koɭ* appears in non-reflexive transitive structures. For a subset of such sentences, *koɭ* is suffixed onto the transitively marked stem of the main predicate. Such verbs remain transitive after *koɭ*-suffixation.

I address these in turn below.

6.3.1 Argument 1: *koɭ* attaches to intransitive verbs

Consider the sentences below, both of which involve change-of-state unaccusatives with *koɭ*:

(123) Paanæ oɖænʤŭ-ko-ɳɖ-adŭ.
pot[NOM] broke-*koɭ*-PST-3NSG
"The pot broke-*koɭ*."

(124) Saʈʈæ (vejjal-læ) suruŋgi-ko-ɳɖ-adŭ.
shirt[NOM] (heat-LOC) shrink-*koɭ*-PST-3NSG
"The shirt shrank-*koɭ* (in the heat)."

What is interesting about these sentences (as well as about many other change-of-state unaccusatives too numerous to individually list here) is that the change-of-state verb in each is overtly distinguished from its transitive counterpart primarily by means of a systematic voicing difference on the coda of the stem-final syllable. This is illustrated in Table (6.1).

VERBAL ROOT	INTRANSITIVE STEM	TRANSITIVE STEM
oɖæ (BREAK)	*oɖæ-nʛ-*	*oɖæ-čč-*
surugŭ (SHRINK)	*suru-ŋg-*	*suru-kk-*
vaɭar (GROW)	*vaɭar-nd*	*vaɭar-tt-*
veɖi (BURST)	*veɖi-nʛ-*	*veɖi-čč-*
kiri (TEAR)	*kiri-nʛ-*	*kiri-čč-*
maɖi (FOLD)	*maɖi-nʛ-*	*maɖi-čč-*

Table 6.1: Transitive and intransitive verbal stems

In sentences like (123)-(124) above, the *koɭ* morpheme is affixed to verbs that are *already* intransitive. In fact, these sentences can just as easily occur without *koɭ*, as the *koɭ*-less counterparts of (123) and (124) demonstrate:

(125) Paanæ oɖæ-nʛ-adŭ.
pot[NOM] broke-PST-3NSG
"The pot broke."

(126) Saʈʈæ (vejjal-læ) suruŋg-ij-adŭ.
shirt[NOM] (heat-LOC) shrink-PST-3NSG
"The shirt shrank (in the heat)."

Crucially, the DPs representing the pot and shirt in (125) and (126), respectively, are affected arguments (or PATIENTS) that undergo a change. In other words, the sentences in (125)-(126) have an unaccusative syntactico-semantics even before the addition of *koɭ*. The addition of *koɭ* must, therefore, be contributing a meaning that is orthogonal to unaccusativity per se, a semantics that is added *on top of* the voice/eventive layer of the sentence in question.

I investigate the precise meaning contribution of *koɭ* in Chapter 7, and will discuss these transitivity alternations again in Section 7.1 of that chapter in the context of what they can tell us about the precise position of *koɭ* in the structure. For now, it suffices to note that the kind of data presented here is fatal for treatments of reflexivity such as that of Chierchia (2004) under which *koɭ* would itself be treated as a type of

detransitivizer. It should be noted that such data isn't problematic for Lidz who notes similar alternations for Kannada. However, the pattern we will discuss directly below is fatal for a Lidz- and Embick-style analysis as well as for lexicalist treatments.

6.3.2 Argument 2: *ko*ɭ attaches to transitive verbs

Even more damning evidence against a reduced-valency treatment of *ko*ɭ comes from the fact that this morpheme may be attached to fully transitive and non-reflexive predicates which, furthermore, retain both properties even after *ko*ɭ-suffixation.[5]

However, *ko*ɭ can occur without *ta(a)n*, not only in anticausative and body-part constructions, but also in transitive structures involving two fully non-coreferent nominal arguments. Thus,*ko*ɭ may occur in a transitive sentence like (127); that its presence is optional (with this predicate) is shown by the grammaticality of the minimally varying sentence in (128):

(127) Raman taɳɲi-jæ koʈʈi-kko-ɳɖ-aan.
Raman[NOM] water-ACC pour-*ko*ɭ-PST-3MSG
"Raman poured-*ko*ɭ the water."

(128) Raman taɳɲi-jæ koʈʈ-in-aan.
Raman[NOM] water-ACC pour-PST-3MSG
"Raman poured the water."

This pattern is illuminated by new data from the results of my Tamil survey. To test the meaning contribution of *ko*ɭ in such sentences and also

[5]Some of this data is discussed by Lidz himself for Kannada and involves the optional appearance of *ko*ɭ in sentences involving coreference between an agentive external argument and the possessor of a body-part object (Lidz 2001; 2004, for Kannada), as in (i) and (ii) from Tamil below:

i. Raman$_i$ tann-oodæ$_{\{i,*j\}}$ kaɳɳ-æ tara-nd-aan.
Raman[NOM] ANAPH-GEN eye-ACC open-PST-3MSG
"Raman$_i$ opened his$_{\{i,*j\}}$ eyes."

ii. Raman$_i$ tann-oodæ$_{\{i,*j\}}$ kaɳɳ-æ tarandŭ-ko-ɳɖ-aan.
Raman[NOM] ANAPH-GEN eye-ACC open-*ko*ɭ-PST-3MSG
"Raman$_i$ opened his$_{\{i,*j\}}$ eyes."

However, Lidz argues that the availability of such sentences isn't necessarily problematic for a reduced valency approach to reflexivity because they are similar enough to reflexives to warrant an identical analytic treatment. Lidz, for instance, proposes that such sentences involve "near reflexivity" with a binding chain relation that exists between two near-identical entities, x and $f(x)$, corresponding to Raman and Raman's eyes, above.

to see to what extent structures like (127) are grammatical in the first place, the respondents of my survey were asked to rate the grammaticality of minimal pairs like (129) and (130) below:

(129) Raman Krishnan-æ paarttŭ-ko-ɳɖ-aan.
Raman[NOM] Krishnan-ACC see-*koɭ*-PST-3MSG
"Raman saw-*koɭ* Krishnan."

(130) Raman Krishnan-æ paar-tt-aan.
Raman[NOM] Krishnan-ACC see-PST-3MSG
"Raman saw Krishnan."

In order to disambiguate the meaning of the *koɭ*- and *koɭ*-less variants and, by extension, to nail down the contribution of *koɭ*, the survey-takers were asked to judge these sentences against discourse scenarios such as the ones below:

SCENARIO I:

Krishnan is a petty thief who, as Raman has just realized, has stolen Raman's cash and escaped in the nick of time. Before Krishnan gets away, however, Raman makes sure to get a good look at Krishnan's face – just in case he ever sees him again.

The majority (12 out of 16) of the respondents found the *koɭ*-variant in (129) to be grammatical in this scenario.[6] Thus, whereas (130) seems to merely assert that Raman saw Krishnan, its *koɭ*-counterpart in (129) seems to have the additional reading that this seeing event was beneficial to or otherwise affected Raman in some way.

As a control case, the survey-takers were also asked to rate the grammaticality of these same sentences under a different scenario where a self-benefactive or self-affective reading is pragmatically ruled out:

SCENARIO II:

Raman and Krishnan are playing hide-and-seek. Today, it's Krishnan's turn to hide but Raman can't find him for the longest time. At last, just as Raman is about to give up, he happens to catch sight of Krishnan crouching under the table in the corner of the room.

[6] My Telugu and Kannada native-speaker informants confirm that this is true for the corresponding sentences in Telugu and Kannada, as well.

Under this scenario, the results of the survey-takers were clearly different: most of them (13 out of 16) judged (130) to be grammatical; in contrast, very few (2 out of 16) considered (129) acceptable with this reading. Furthermore, the addition of an adverb expressing non-volitionality, such as *edeechiyaa* (ACCIDENTALLY), was considered to be marked in (129) but acceptable in (130). The interpretive difference between structures with (129) and without (130) *ko*ḷ follows in the manner described for Scenario 1. I.e. while the structure in (130) is the pragmatically unmarked way to capture the proposition that Raman saw Krishnan, the sentence in (129) has the additional meaning that this seeing event was beneficial to or otherwise affected Raman. The suffixation of *ko*ḷ appears to be sensitive to this interpretive difference.

The precise conditions under which *ko*ḷ is licensed and, by extension, what *ko*ḷ (or, more precisely, the element that *ko*ḷ spells out) means are the concern of much of the rest of this series of chapters. For now, it suffices to note that the possibility of *ko*ḷ in a clearly non-reflexive structure such as that in (130) is fatal for unaccusative analyses of reflexivity both within the Lexicalist and Non-Lexicalist frameworks since, in sentences such as these, the predicate remains transitive and non-reflexive even after *ko*ḷ-suffixation.[7]

Now that we have ascertained what *ko*ḷ is not, we can move on to an investigation of what it *is*. To this end, we will discuss the morphological, syntactic, and semantic properties of *ko*ḷ in the next chapter.

[7]Interestingly, my Kannada native speaker informants report that *ko*ḷ may also occur in non-reflexive structures in Kannada. Thus, the following Kannada sentences are fully grammatical – indeed, the *ko*ḷ-less version is dispreferred for these verbs:

i. Ramaa Krishnan-annu tabbi-ko-ɳḍ-aa/*tabb-id-aa
rama[NOM] Krishnan-ACC hug-*ko*ḷ-PST-3MSG/*hug-PST-3MSG
"Rama hugged Krishnan."

ii. Ramaa magu-vannu tegedu-koɳḍu-hoodaa/*tege-d-aa
Rama[NOM] child-ACC take-*ko*ḷ-GO/take-PST-3MSG
"Rama took the child."

A thorough and large-scale investigation of Kannada *ko*ḷ with different verb-classes, akin to what is being done in this portion of the dissertation, must be carried out to get a better sense of what the morpheme means in this language. However, the very possibility of sentences such as (i) and (ii) suggests that *ko*ḷ is no more a monadic causative in Kannada than it is in Tamil.

Chapter 7

The quest for a meaning: where, what, how and why is *ko*⌊?

In the Dravidian literature, the *ko*⌊ morpheme is often singled out as being a notoriously difficult morpheme to describe. For instance, Schiffman (1995) describes *ko*⌊ as "one of the most complex of the Tamil aspectual verbs", with a range of apparently disparate or internally inconsistent meanings, among them: self-benefaction or self-affectedness, volitionality, accident, inchoation from a state, the simultaneity or completion of an action. Annamalai (1999) states that the meaning of *ko*⌊ in Tamil is rather vague and cannot be easily captured, in part also because it is so sensitive to discourse-pragmatic variables. Steever (2005), in his book about auxiliary verb constructions in Tamil, writes that "Due to the broad range of circumstantial meanings it conveys, *ko*⌊⌊*æ* 'hold, contain' is perhaps the most difficult of indicative auxiliaries to analyze" (Steever 2005, 207). However, our goal in this chapter will be to do precisely this: we will investigate the morphological, categorial, distributional, and semantic properties of this elusive morpheme in order to pin down exactly what it is, and what it isn't.

We will start from first principles, looking at where in the linear sequence of the predicate cluster this morpheme appears. Based on this investigation, we will see that *ko*⌊ is base-merged directly above an aspectual morpheme, built on top of VoiceP, which creates a result state out of an event. However, while this shows that *ko*⌊ is a "verb-y" element that combines with a telic verbal predicate, it still doesn't tell us what kind of verb-y entity it is. Is it, for instance, a lexical element, like a verb, or a functional element, like an auxiliary, or Pylkkänen (2008)-style ap-

plicative head? Or is it something in-between, like Wurmbrand (2001)'s "semi-functional" restructuring verb – an animal with hybrid lexical and functional properties? We will investigate this issue in some detail, on the basis of which we will argue that *ko*ɭ̣ is indeed a semi-functional predicate, much like direct perception verbs and *lassen*-type causatives in German.

Building on this result, we will move on to a thorough investigation of what *ko*ɭ̣ means, examining the compatibility of this morpheme with a wide range predicates culled from the Levin (1993) verb-classes. *ko*ɭ̣ will be seen to contribute the meaning that the highest argument of the event in its scope *comes to hold*, in a physical or mental sense, the result state of this event – much like *get* in English (McIntyre 2010). I will show that the supposedly paradoxical readings of accident and volitionality are not mutually inconsistent at all, but are a function of whether the "coming to hold" semantics of *ko*ɭ̣ is understood as a physical or mental coming-to-hold, respectively.

We will then turn our attention to the topic that ties this series of chapters with the others in the dissertation, namely: what is the relationship between *ko*ɭ̣ and local binding in Tamil? I will propose that PerspP – the phrasal projection which was argued, in Part I, to contain the operator which Agrees with *ta(a)n* in the syntax and binds it at LF – appears not only on CPs, DPs, and PPs, as proposed, but also on the AspP complement of *ko*ɭ̣. In other words, we will see that PerspP is a property of *phases*. I will show, furthermore, that local antecedence of *ta(a)n* is precluded in the absence of *ko*ɭ̣ because the intended antecedent, in such cases, doesn't asymmetrically c-command the minimal PerspP containing *ta(a)n* – a structural condition that, I will argue, is independently necessary to qualify the entity denoted by a DP as a perspective-holder (and, by extension, to qualify the DP as a potential anaphoric antecedent). As a result of our investigations, we will be forced to clarify and update the definition of "perspective" developed in Part I.

7.1 Morphology and linear sequence of verbal elements

In this section, we will use the relative ordering of morphemes on the verbal root as a metric for understanding where in the verbal structure *ko*ɭ̣ is syntactically merged. This is, of course, based on the assumption that the surface ordering faithfully reflects the underlying syntactic structure, but there is no reason to assume otherwise, since these morphemes cannot

be re-ordered relative to one another.

Consider the unaccusative and transitive variants of a fairly standard SOV sentence in Tamil like those below, organized as minimal pairs with and without *ko*ɭ:

(131) *o*ɖ*æn*ʤ-, BREAK (Unaccusative):

a. Paanæ oɖæ-nʤ-adŭ.
pot[NOM] break-INTR.PST-3NSG
"The pot broke."

b. Paanæ oɖænʤŭ-ko-ɳɖ-adŭ.
pot[NOM] break.INTR-*ko*ɭ-PST-3NSG
"The pot broke-*ko*ɭ."

(132) *o*ɖ*æčč*-, BREAK (Transitive):

a. Maya paanæ-jæ oɖæ-čč-aaɭ.
Maya[NOM] pot-ACC break-TR.PST-3FSG
"Maya broke the pot."

b. Maya paanæ-jæ oɖæččŭ-kko-ɳɖ-aaɭ.
Maya[NOM] pot-ACC break.TR-*ko*ɭ-PST.3FSG
"Maya broke-*ko*ɭ the pot."

The verbal complexes of the *ko*ɭ-less and *ko*ɭ forms consist of the following linear sequences of morphemes:

(133) **Linear sequence of verb-forms without** *ko*ɭ**:**

131a. *o*ɖ*æ-nʤ-adŭ* = VERB-ROOT–INTR.PST–3NSG

132a. *o*ɖ*æ-čč-aa*ɭ = VERB-ROOT–TR.PST–3FSG

(134) **Linear sequence of verb-forms with** *ko*ɭ**:**

131b. *o*ɖ*æ-nʤŭ-ko-*ɳɖ*-adŭ* = VERB-ROOT–INTR.PST–KOɭ–PST–3NSG

132b. *o*ɖ*æ-ččŭ-ko-*ɳɖ*-aa*ɭ = VERB-ROOT–TR.PST–KOɭ–PST–3FSG

Both the *ko*ɭ-less sequences in (133) and their *ko*ɭ-variants in (134) involve a transitivity marker above the verb-root. Transitivity is marked by means of a morphophonological distinction on the coda of the stem-final syllable, as has already been illustrated in Table 6.1. Specifically, in the intransitive variant (*o*ɖ*æn*ʤ-), the coda is a sequence of nasal followed by a homorganic voiced affricate while in the transitive variant (*o*ɖ*æčč*-), it is a geminate voiceless affricate. Note that the cluster which shows this voicing distinction is in fact also the realization of the morpheme that looks like a past-tense marker. Assuming that this intransitive-intransitive distinction is, in fact, a realization of a syntactic Voice head

(Kratzer 1996), we have evidence that *koɭ* attaches above Voice. Independent support for this conclusion comes from the fact that, in transitive *koɭ*-structures, *koɭ* modifies the external (AGENTIVE or EXPERIENCER) argument, not the internal, one in its scope. In unaccusatives, however, the meaning-contribution of *koɭ* applies to the internal (THEME/PATIENT) argument:

(135) **Transitives: *koɭ* applies to external AGENT argument:**

i. Raman paanæ-jæ oɖæ-čč-aan.
Raman[NOM] pot-ACC break-TR.PST-3MSG
"Raman$_i$ broke-*koɭ* the pot."

(136) **Unaccusatives: *koɭ* applies to internal THEME/PATIENT argument**

i. Paanæ oɖæ-nʂ-adŭ.
Pot-ACC break-INTR.PST-3MSG
"The pot broke-*koɭ*."

What is the nature of the past-tense like morpheme that also realizes the transitivity distinction? In the *koɭ*-less unaccusative and transitive sequences given in (133), this morpheme does indeed seem to denote real tense. We can see this because the past morpheme may be replaced with the present to yield a present reading for the breaking event. Observe, incidentally, that the voicing distinctions marking transitivity continue into the present as well:

(137) ***koɭ*-less unaccusative (131a) - past and present:**

broke.INTR: *oɖæ-nʂ-adŭ* = VERB-ROOT–INTR.PST–3NSG

breaks.INTR: *oɖæ-gir-adŭ* = VERB-ROOT–INTR.PRS–3NSG

(138) ***koɭ*-less transitive (132a) - past and present:**

broke.TR: *oɖæ-čč-aaɭ* = VERB-ROOT–TR.PST–3FSG

breaks.TR: *oɖæ-kkir-aaɭ* = VERB-ROOT–TR.PRS–3FSG

In contrast to (137) and (138), the past-tense-like morpheme that appears directly before *koɭ* in the *koɭ*-variants doesn't seem to represent "real" tense. First, it seems to be "frozen" and cannot be replaced by a different tense form, like the present:

(139) ***koɭ* unaccusative (131b) - past and present:**

broke.INTR: *oɖæ-nʂŭ-koɭ* = VERB-ROOT–INTR.PST–*koɭ*

breaks.INTR: **oɖæ-gir-koɭ* = VERB-ROOT–INTR.PRS–*koɭ*

(140) *koɭ* **transitive (132b) - past and present:**

broke-*koɭ*.TR: *oɖæ-ččŭ-koɭ* = VERB-ROOT– TRANS.PST–*koɭ*

breaks-*koɭ*.TR: **oɖæ-kkirŭ-koɭ* = VERB-ROOT–TRANS.PRS–*koɭ*

Second, there *is* another tense marker in the unaccusative and transitive *koɭ* sequences – the morpheme that occurs *after koɭ*: this morpheme seems to indicate *real* tense, because it can be replaced with a present tense marker, just like we saw for the *koɭ*-less variants in (137) and (138), to yield a present-tense reading. The "past"-tense marker to the left of *koɭ* stays frozen, as we can see:

(141) oɖæ-ččŭ-ko-ɳɖ-aaɭ
break-TR.PST-*koɭ*-PST-3FSG
"She broke-kol"

(142) oɖæ-ččŭ-koɭ-gir-aaɭ
break-TR.PST-*koɭ*-PRS-3FSG
"She breaks-kol"

These examples show two things: first, that the real tense marking comes directly before agreement in both the *koɭ*- and *koɭ*-less variants; and second, that the morpheme directly to the left of *koɭ* in the unaccusative and transitive *koɭ*-sequences above doesn't signifiy real past tense but something else. The minimal pairs given below confirm these intuitions:

(143) *aɖi-* (HIT):

a. aɖi- tt- aan ∼ aɖittŭ- kko- ɳɖ- aan
hit- PST- 3MSG ∼ hit- *koɭ*- PST- 3MSG

b. aɖi- kkir- aan ∼ aɖittŭ- kkoɭ- gir- aan
hit- PRS- 3MSG ∼ hit- *koɭ*- PRS- 3MSG

c. aɖi! ∼ aɖittŭ- kkoɭ!
hit.IMP ∼ hit- *koɭ*.IMP

(144) *vaaŋg-* (BUY):

a. vaaŋg- in- aan ∼ vaaŋgi- kko- ɳɖ- aan
buy- PST- 3MSG ∼ buy- *koɭ*- PST- 3MSG

b. vaaŋgŭ- gir- aan ∼ vaaŋgi- kkoɭ- gir- aan
buy- PRS- 3MSG ∼ buy- *koɭ*- PRS- 3MSG

c. vaaŋgŭ! ∼ vaaŋgi- kkoɭ!
buy.IMP ∼ buy- *koɭ*.IMP

(145) *soll-* (SAY):

a. so- nn- aan ∼ solli- kko- ɳɖ- aan
say- PST- 3MSG ∼ say- *koɭ*- PST- 3MSG

b. soll- gir- aan ∼ solli- kkoḷ- gir- aan
say- PRS- 3MSG ∼ say- *koḷ*- PRS- 3MSG
c. sollŭ! ∼ solli- kkoḷ!
say.IMP ∼ say- *koḷ*.IMP

The examples above confirm that:

(i) The verb-stem that *koḷ* attaches to consists of: verb-root + something that looks like a past-tense marker. In the forms given under (143) and (144), this is fairly easy to see. In (145), this is obscured by the irregular form of the simple past *so-nn-aan* but the *-i-* at the end of the stem *solli is* a past-marker, as seen also in *vaaŋgi* in (144).

(ii) The "past-tense" marker that precedes *koḷ* is invariant. It appears regardless of the actual tense of the verb-form which occurs after *koḷ*. This is shown in the present-tense *koḷ* forms given in: (143b), (144b) and (145b). Indeed, this "past" morpheme even shows up in non-tensed forms like imperatives, as in (143c), (144c), and (145c).

In descriptive grammars of Tamil, this marker is simply glossed as PAST (see e.g. Annamalai 1997). However, Amritavalli and Jayaseelan (2005) argue, on the basis of evidence from negation, serial verb constructions in Malayalam and other Dravidian languages, and appearance in a series of clearly nonfinite forms like gerundivals, that these markers represent aspect, not tense. Thus, it is not so surprising that these forms should appear inside *koḷ* and, furthermore, be orthogonal in value to the outer tense marker that does appear. Let us assume that this is indeed the case. We now have an answer to where in the syntactic structure the head that spells out *koḷ* is merged. I will thus propose that the linear order of relevant morphemes is as given below:[1]

(146) **Syntactic position of** *koḷ*:
VerbRoot - (Voice) - Asp - *koḷ* - Tense - Agr

7.2 The nature of Asp

What is the nature of the Asp head that *koḷ* attaches to? Consider the *koḷ*-sentences below:

[1]The Voice head only occurs in transitive structures. Since *koḷ* may occur in transitives or intransitives, the Voice head is marked as being optional, in the generalized structure given here.

(147) *muuɖ̆ŭ* (CLOSE):

a. Kadavŭ muuɖ-i-jadŭ.
door[NOM] close-PST-3NSG
"The door closed."

b. Kadavŭ muuɖ-i-kko-ɳɖ-adŭ.
door[NOM] close-*koɭ*-PST-3NSG
"The door closed-*koɭ*."

(148) *minŭminŭ* (SPARKLE):

a. Vairam minŭminŭ-kkir-adŭ.
diamond[NOM] sparkle-PRS-3NSG
"The diamond sparkles."

b. * Vairam minŭminŭ-ttŭ-koɭ-gir-adŭ.
diamond[NOM] sparkle-ASP-*koɭ*-PRS-3NSG
"The diamond sparkles." (Intended)

The consequences of adding *koɭ* to *muuɖ̆ŭ* (CLOSE) and *minŭminŭ* (SPARKLE) are rather different, as the sentences above show: in brief, *koɭ* may be readily added to the former (147b), but is ungrammatical (148b) with the latter. Why should this be?

The answer has to do with the aspectual properties of these predicates. An event of door-closing is inherently *telic* because it leads to a result state of the door being closed; on the other hand, an event of diamond-sparkling isn't telic – a diamond could keep sparkling forever and, worded rather informally for the time being, the world would be no different after the sparkling than it was before. A standard diagnostic for distinguishing between telic and atelic predicates is the felicity of their combination with temporal adverbials or PPs (Dowty 1979). The former are expected to be compatible with "in an hour" PPs, however the latter are not. A second diagnostic, discussed in Borik and Reinhart (2004, who credit Verkuyl (1972) for its origin) has to do with interpretive differences under modification by two conjoined temporal PPs. We can test both diagnostics with sentences constructed around closing and sparkling predicates – this is shown below for English:

(149) **Adverbial modification:**

a. The door closed in an hour.

b. * The diamond sparkled in an hour.

(150) **Conjunction of temporal PPs:**

a. The door closed on Tuesday and on Wednesday. (UNAMBIGUOUS)

b. The diamond sparkled on Tuesday and on Wednesday. (AMBIGUOUS)

The adverbial modification test in (149) shows that telic predicates alone may be modified by a PP *in an hour*: thus (149a) is fully grammatical while (149b) is not. The conjunction test in (150) shows that telic predicates like *close* yield an unambiguous reading under modification by two conjoined temporal PPs; thus, (150a) can only mean that there were two separate door-closing eventualities – one on Tuesday and the next on Wednesday. In contrast, atelic predicates are ambiguous in this structure. Thus, (150b) can mean that there were two separate sparkling eventualities, one on Tuesday and another on Wednesday (analogous to the interpretation of the door-closing eventuality); alternatively, it could mean that there was a single eventuality of diamond-sparkling which carried on over (at least) two days, from Tuesday through Wednesday. This second reading is simply impossible with a telic predicate like CLOSE in (150a).

How do the corresponding Tamil verbs *muuɖ̆ŭ* (CLOSE) and *minŭminŭ* (SPARKLE) fare with regard to these diagnostics? With respect to the adverbial modification test, (151b) shows that a PP meaning "in an hour" cannot modify a predicate that doesn't involve a telos;[2] (151a), in contrast, shows that such modification is possible with a telic predicate which does involve such a result state. With respect to the conjunction diagnostic as well, these verbs behave just like their English counterparts. In (152a), *muuɖ̆ŭ* (CLOSE) yields the unambiguous reading that there were two separate door-closing eventualities, just like its English equivalent in (150a); in contrast, (152b) is ambiguous in the same way as (150b) is: it could either mean that there are two separate sparkling eventualities, one on Tuesday and another on Wednesday, or it could mean that there is a single eventuality of diamond sparkling that continued over these two days:

(151) **Adverbial modification in Tamil:**

a. Kadavŭ orŭ maŋŋæratt-ŭlæ muuɖ̆-ij-adŭ.
door[NOM] one hour-LOC close-PST-3NSG
"The door closed in an hour."

b. * Vairam orŭ maŋŋæratt-ŭlæ minŭminŭ-tt-adŭ.
diamond[NOM] one hour-LOC sparkle-PST-3NSG
"The diamond sparkled in an hour."

[2] As in English, the sentence *is* licit under a reading where "in an hour" indicates the time taken until the sparkling comes about – e.g. if I start with a dull diamond and clean it until it sparkles.

(152) **Conjunction of temporal PPs in Tamil:**

a. Kadavŭ tiŋgækkiɻæmæ=jum budaŋgiɻæmæ=jum muuɖ-
door[NOM] tuesday=CL wednesday=CL close-
ij-adŭ.
PST-3NSG
"The door closed on Tuesday and on Wednesday" (**Unambiguous**)

b. Vairam tiŋgækkiɻæmæ=jum budaŋgiɻæmæ=jum
diamond[NOM] tuesday=CL wednesday=CL
minŭminŭ-tt-adŭ.
sparkle-PST-3NSG
"The diamond sparkled on Tuesday and on Wednesday" (**Ambiguous**)

The grammaticality judgments above show that the Tamil equivalents of BREAK and CLOSE are telic and atelic, just like their English counterparts.[3]

Telic predicates may also be compatible with modification by "for an hour" type PPs, as long as the part of the event leading up to the result state has some duration. So, in English, (153) is possible as long as the door closes very slowly and we are focussing on the process rather than the end result of closing:

(153) The door closed for an hour.

Once again, this same flexibility holds true for inherently telic predicates in Tamil, as shown in (154) below:

[3]Pederson (2008) claims that predicates like BREAK in Tamil systematically differ from their counterparts in English in that they lack a target state. Thus, Pederson reports that a sentence like (i) is perfectly licit in Tamil, although its English translation is internally inconsistent, thus nonsensical (Pederson 2008, 331, formatting mine):

i. Aijar teeŋgaaj-æ oɖæ-čč-aar. Aanaal teeŋgaaj
brahmin[NOM] coconut-ACC break-TR.PST-3MSG.HON. But coconut[NOM]
oɖæja-villæ.
break.INTR.INF-NEG
"The brahmin broke the coconut. But the coconut didn't break."

I do not share these judgments, however. Thus, the sentence in (i) seems just as internally contradictory to me as its English equivalent – in other words, *oɖæ* (BREAK) in Tamil is inherently telic in my grammar, just like *break* in English. I will thus ignore this reported use of BREAK and BREAK-like verbs in Tamil for the time being and focus on those judgments that I do know to be true based on my own native speaker intuitions.

(154) Kadavŭ orŭ maŋŋæramaa muuɖ-ij-adu.
door[NOM] one hour.ADV close-PST-3NSG
"The door closed for an hour."

This is where things get interesting. *ko*ɭ-suffixation is impossible in sentences like (154) which involve modification by a PP like "for an hour". However, modification by "in an hour" or the absence of any temporally modifying PP renders *ko*ɭ suffixation on such predicates fully grammatical:

(155) * Kadavŭ orŭ maŋŋæramaa muuɖi-ko-ɳɖ-adŭ.
door[NOM] one hour.ADV close-*ko*ɭ-PST-3NSG
"The door got closed for an hour." (Intended)

(156) Kadavŭ orŭ maŋŋæratt-ŭlæ muuɖi-ko-ɳɖ-adŭ.
door[NOM] one hour-LOC close-*ko*ɭ-PST-3NSG
"The door closed-*ko*ɭ in an hour."

(157) Kadavŭ muuɖi-ko-ɳɖ-adŭ.
door[NOM] close-*ko*ɭ-PST-3NSG
"The door closed-*ko*ɭ."

These grammaticality patterns can be explained if we assume that the addition of *ko*ɭ to a verb forces it to be interpreted as telic. The sentence in (155) is ungrammatical with *ko*ɭ because the addition of the AdvP "for an hour" makes it resist a telic interpretation. If this idea is correct, we expect *ko*ɭ-suffixation to be degraded on other predicates that similarly resist a telic interpretation.

This expectation is, indeed, confirmed – *ko*ɭ is degraded with inherent statives and bodily process, light, and sound emissives. These predicates can all be independently shown to be atelic because they fail the "in an hour" test given above and yield ambigious readings under PP conjunction. Examples of the former are reproduced below:[4]

(158) * Seetha orŭ maŋŋærattŭ-læ tumm-in-aaɭ.
Seetha[NOM] one hour-LOC sneeze-PST-3FSG
"*Seetha sneezed in an hour."

(159) * Vivek orŭ maŋŋærattŭ-læ Madras-læ irŭ-kkir-aan.
Vivek[NOM] one hour-LOC Madras-LOC be-PRS-3MSG

[4]In Section 7.2.1, we will see that activity predicates behave differently. They are not inherently telic, and thus do not normally combine felicitously with "in an hour", but neither are they *incompatible* with telicity, and can yield telic predications when combined with the right modifying elements – just as in English *run* is atelic, but *run to the store* is telic. Activities are thus compatible with the addition of *ko*ɭ, which thus forces a telic interpretation as expected.

"*Vivek is in Madras in one hour."

The sentences below show that *ko*ɭ is degraded with such predicates:

(160) BODILY EMISSIVES - (*tummŭ*, SNEEZE):

a. Seetha sattamaagæ tumm-in-aaɭ.
Seetha[NOM] loudly sneeze-PST-3FSG
"Seetha sneezed loudly."

b. * Seetha sattamaagæ tummi-ko-ɳɖ-aaɭ.
Seetha[NOM] loudly sneeze-*ko*ɭ-PST-3FSG
"Seetha sneezed loudly." (Intended)

(161) LIGHT EMISSIVES - (*minŭminŭ*, SPARKLE):

a. Vairam minŭminŭ-kkir-adŭ.
diamond[NOM] sparkle-PRS-3NSG
"The diamond sparkles."

b. * Vairam minŭminŭttu-koɭ-kir-adŭ.
diamond[NOM] sparkle-*ko*ɭ-PRS-3NSG
"The diamond sparkles." (Inten ded)

(162) SOUND EMISSIVES:

a. Seetha Krishnan-ooɖæ keʈʈæ guɳatt-æ paʈri
Seetha[NOM] Krishnan-GEN terrible nature-ACC about
kiiччĭ-iʈʈ-aaɭ.
screech-PST-3FSG
"Seetha screeched about Krishnan's terrible nature."

b. ?? Seetha Krishnan-ooɖæ keʈʈæ guɳatt-æ paʈri
Seetha[NOM] Krishnan-GEN terrible nature-ACC about
kiiччiʈʈŭ-ko-ɳɖ-aaɭ.
screech-KOL-PST-3FSG
"Seetha screeched about Krishnan's good nature." (Intended)

(163) BODILY PROCESS VERBS - (*naɖŭŋgŭ*, TREMBLE):

a. Anand tiɖiirunŭ naɖŭŋg-in-aan.
anand[NOM] suddenly shiver-PST-3MSG
"Anand shivered suddenly."

b. * Anand tiɖiirunŭ naɖŭŋgi-ko-ɳɖ-aan.
anand[NOM] suddenly shiver-*ko*ɭ-PST-3MSG
"Anand shivered suddenly." (Intended)

(164) INHERENT STATIVES - (*irŭ*, BE):

a. Vivek Madras-læ irŭ-kkir-aan.
Vivek[NOM] Madras-LOC be-PRS-3MSG

"Vivek is in Madras."

b. * Vivek Madras-læ irŭndŭ-*ko*ḷ-gir-aan.
Vivek[NOM] Madras-LOC be-*ko*ḷ-PRS-3MSG
"Vivek is in Madras." (Intended)

As mentioned earlier, the various classes of emissives such as those in (160)-(162), bodily process verbs like those given under (163) and, finally, inherent statives like (164) are all atelic, thus do not inherently contain a telos state; furthermore, they all resist the addition of a result state to create a telos. *ko*ḷ-suffixation, however, must itself be contingent on the structure in its scope's being interpreted as telic. This would lead explain why such sentences are all ungrammatical or sharply degraded with *ko*ḷ.

Based on this discussion, we may propose the following:

(165) A *ko*ḷ-structure is always interpreted as being telic.

But we are now faced with a choice. In Section (7.1), we saw that *ko*ḷ doesn't attach to the bare verbal root. Rather, it attaches to what we have identified to be an aspectual marker. Thus, there are three logical sources of telicity: the verbal root, the aspectual head above the root, or *ko*ḷ itself. Which of these is the correct choice?

7.2.1 Where is telicity encoded?

Where is telicity structurally encoded in *ko*ḷ-structures? On the one hand, the fact that the telicity of verbs like *muuḍŭ* (BREAK) is an inherent part of their lexical meanings suggests that it is simply encoded at the level of the verbal root. On the other hand, it is fairly easy to show that *ko*ḷ doesn't just attach to predicates whose roots are inherently telic – it attaches to any verb that is, in theory, *compatible* with a telic interpretation even it is itself not necessarily telic.

Support for the latter comes from the fact that *ko*ḷ may, in fact, be productively added to most predicate-classes. Given the appropriate discourse-context, it can be added to activities (e.g. *ooḍŭ-* (RUN), *teeḍŭ-* (SEARCH)), achievements like (*kaṇḍŭpiḍi-* (DISCOVER), *aarambi-* (START)), as well as accomplishments:

(166) **Activity verb** + *ko*ḷ:

Raman veegamaagæ pooṯṯi-læ ooḍi-kko-ṇḍ-aan.
Raman[NOM] fast race-LOC run-*ko*ḷ-PST-3MSG

"Raman ran quickly in the race (for his own sake)." (Rough translation)

(167) **Achievement verb** + *koɭ*:

Maya višijatt-æ kaɳḍŭpiḍittŭ-kko-ɳḍ-aaɭ.
Maya[NOM] news-ACC discover-*koɭ*-PST-3FSG
"Maya went out and discovered the news for herself." (Rough translation)

(168) **Accomplishment verb** + *koɭ*:

Seetha pustagatt-æ paḍittŭ-ko-ɳḍ-aaɭ.
Seetha[NOM] book-ACC read-*koɭ*-PST-3FSG
"Seetha read the book to/for herself." (Rough translation)

Activities like *ooḍŭ* (RUN) in (166) lack a telos, while achievements like *kaɳḍŭpiḍi* (DISCOVER) in (167) and accomplishments like *paḍi* (READ) in (168) have an inherent telos. The *koɭ*-predicates in the sentences above, however, are *all* interpreted as being telic.

This shows that the semantics of telicity in *koɭ*-structures must have its origin, not in the verbal root itself, but in some element outside it. Note that such an approach also accommodates verbs which are inherently telic, like CLOSE, as there is nothing to prevent the addition of an external result state on top of a predicate with an internally represented 'target state' (see Kratzer 2000, for related discussion on the formation of resultant state participles from stems with a target state argument). The fact that *koɭ* is compatible with all three aspectual classes of predicate thus shows that telicity must be due to the addition of a result state added external to the verbal root. This narrows down our possibilities but it doesn't exhaust them: we must still determine whether this result state is supplied by the aspectual morpheme directly above the verbal root or by *koɭ* itself. Independent empirical tests show that a telic interpretation is obtained even when the aspectual morpheme appears suffixed on a verb in the absence of *koɭ*, suggesting that the result state is introduced by the Asp head and not by *koɭ*. I discuss two such pieces of evidence below.

First, this aspectual marking also appears obligatorily below other morphemes like *koɭ* – e.g. under *kuḍŭ-* (Lit: GIVE) or *vej-* (Lit: KEEP) both of which can be used as full verbs but can also have a less "meaty" interpretation in which case they appear in the same position as *koɭ* in the linear predicate-sequence. *kuḍŭ-* and *vej*-structures are also always interpreted as being telic, even if the verbal roots involved themselves

are not:

(169) Raman (maɳikkaɳakkaa) paɖi-čč-aan.
Raman[NOM] for.hours read-PST-3MSG
"Raman read (for hours)."

(170) * Raman (*maɳikkaɳakkaa) pustagatt-æ paɖi-ččŭ-kuɖŭ-
Raman[NOM] for.hours book-ACC read-ASP-GIVE
tt-aan/ paɖi-ččŭ-vej-čč-aan.
-PST-3MSG/ read-ASP-KEEP-PST-3MSG
"Raman finished reading (*for hours)." (rough translation for both)

The main verb *paɖi* (READ) in the sentences above denotes an open-ended activity and is thus atelic – a reading that thus allows modification by the adverbial *maɳikkaɳakkaa* (FOR HOURS), as (169) shows. However, once either *kuɖŭ* or *vej* are added, the resultant structures must be interpreted telically – this is why modification by *maɳikkaɳakkaa* is no longer grammatical in these cases.

Second, the aspectual morpheme also marks every verb, except the final one, in a serial verb sequence, as shown in (171). While it may occur on most verbs, it is degraded when suffixed onto those that lack an inherent telos and resist the addition of a telos, like inherent states. For this reason, it is in fact difficult to form a serial verb construction out of such verbs. This is shown in (172):

(171) Maya paal-æ aɭæ-ndŭ kaačč-i kuɖi-kkir-aaɭ.
Maya[NOM] milk-ACC measure-ASP boil-ASP drink-PRS-3FSG
"Maya measures (out), boils, and drinks the milk."

(172) * Maya irŭ-ndŭ verŭttŭvaaɻ-gir-aaɭ.
Maya[NOM] be-ASP live-PRS-3FSG
"Maya is, hates, and lives." (Intended)

Claiming that, in predicates involving *koɭ*, the derived result state is added not by *koɭ* itself but by the aspectual morpheme that is immediately below *koɭ* and above the verbal root, would explain all these properties. I will assume this, therefore, and propose the following:

(173) The aspectual morpheme between the verbal root and *koɭ* encodes a derived result state. This ensures that the predicate under *koɭ* is interpreted as being telic – i.e. as an inchoation that culminates at the result state introduced by Asp_{res}.

We can now give the aspectual head between the verbal root and *koɭ* a name: let us call it Asp_{res}.

7.2.2 The denotation of Asp_{res}

What is the denotation of Asp_{res}? Kratzer (2000) proposes the following denotation for a target state "stativizer" in German and English(see also the Asp_R head in Embick 2004a):

(174) $\lambda R_{<s,s,t>} \lambda s_s \exists e_s . R(s)(e)$

Kratzer's stativizer head does not itself introduce a target state – this is rather a part of the denotation of the predicate that it attaches to; what it does is to require, by virtue of its denotation, that the predicate in its complement have unsaturated state and event arguments, i.e. that this predicate relate a state to an event. The stativizer head binds off the event variable and passes the state further along in the structure.

Our denotation for Asp_{res} must, however, be different, in keeping with our results pertaining to the compatibility of *koḷ* with different aspectual classes of predicate, above. I.e. we have seen that *koḷ* doesn't require that its predicate have unsaturated state and event variables. In fact, we saw that it may readily co-occur with verbs that explicitly don't have a target state/inherent telos, – e.g. activity verbs, like *oo*ḍ*ŭ* (RUN). Thus, the predicate that Asp_{res} combines with must itself be eventive; the result state variable must then be *introduced* by Asp_{res}, not merely passed along, as Kratzer's stativizer in (174) does; but we will maintain that the unsaturated event variable of the main predicate is existentially bound, just as in Kratzer's version. This yields the following denotation for Asp_{res}:

(175) $[\![Asp_{res}]\!] = \lambda R_{<s,t>} \lambda s_s \exists e . R(e) \wedge Result(e, s)$

7.3 What syntactic category does *koḷ* instantiate?

Now that we have ascertained the nature of the aspectual head below *koḷ* and, more generally, the aspectual property of the predicate that *koḷ* takes as complement, we can move on to examining what syntactic category *koḷ* itself instantiates. The fact that *koḷ* occurs as part of a larger predicate-cluster suggests that it is some sort of "verb-y" element, but within this domain there are, of course, many distinct options to choose from.

In Tamil descriptive grammars and typological studies of the language (Schiffman 1995, Steever 2005, among others), *koḷ* is described as an auxiliary. However, the term "auxiliary" is rather ill-defined in the

theoretical literature and has been used to refer to a range of functional elements with mutually inconsistent properties. For many, an auxiliary is synonymous with an element that appears in Aux or T. However, Steever (2005) himself explicitly states that his use of the term "auxiliary" should not be understood to represent an element that corresponds precisely to a syntactic category like Aux or T. In fact, his use of the label "auxiliary verb construction" referring to the type of structure *ko*ɭ yields when it attaches to a predicate, seems to me to be more in line with that of a complex predication in the Butt (1995)-sense where, in fact, it would be categorized as a light verb and explictly *not* as an auxiliary. To the extent that "auxiliary" has a consistent use, it is defined in morphological terms to refer, broadly speaking, to a functional element with verbal properties which occurs above the main lexical verb as a separate morphophonological word, i.e. as an ingredient of periphrasis. But this says nothing about the syntactic category it represents – clear English auxiliaries are thus standardly seen to be instantiations of a series of distinct functional heads such as T, Mod, Perf, *v*, and so on.

I will thus stay away from what might prove to be nothing more than a superficial exercise in labelling and set aside the issue of whether *ko*ɭ should be classified as an auxiliary or not. I will focus instead on analytic options that make substantive claims about what kind of syntactic element *ko*ɭ actually is. Specifically, I will consider the following possibilities for *ko*ɭ:

(i) that it is a serial verb.

(ii) that it is a Pylkkänen (2008)-style applicative. head.

(iii) that it is a light verb in a Butt (1995)-style complex predication.

Each of these analytic alternatives has its own empirical implications for the syntactic and semantic properties of *ko*ɭ. I will discuss them in turn and show that none of these is ideal, concluding that a fourth option – namely one that proposes that *ko*ɭ is a semi-functional restructuring verb in the sense of Wurmbrand (2001) is called for.

7.3.1 Option 1: *ko*ɭ is a serial verb

Aikhenvald and Dixon (2006) define a serial verb construction as "a series of verbs which acts together as one. They are monoclausal; their intonational properties are those of a monoverbal clause; they generally have just one tense, aspect, mood, and polarity value." Although object DPs and verbal modifiers may intervene between the verbs in a subset

of languages that exhibit this phenomenon, the clausal subject appears to be "shared" across all the verbs. Tamil is a language that manifests serial verb formation (see Jayaseelan 2004, for an exposition of the serial verb formation in Malayalam and some of the other major Dravidian languages). *koḷ*, we will recall, occurs suffixed to a verbal stem in all the examples we have seen so far. Thus, it is not immediately apparent whether *koḷ* is itself just another serialized verb or the spell-out of a functional head on top of the verbal root.

To determine this, we need to compare the properties of uncontroversial serial verb structures and those involving *koḷ* and see whether they are identical or not. The standard serial-verb structures in (176)-(178) all vary slightly but significantly from one another, in a way which casts light on their fundamental properties:

(176) Champa maaŋgaa-væ pari-čč̆ŭ uri-čč̆ŭ narŭkk-i
Champa[NOM] mango-ACC pluck-ASP peel-ASP chop-ASP
uppŭpoo-ṯṯŭ uurave-čč̆ŭ uurugaase-ɳʂŭ saappi-ṯṯ-aaḷ.
brine-ASP marinate-ASP pickle-ASP eat-PST-3FSG
"Champa plucked, peeled, chopped, brined, marinated, pickled, and ate the mango."

(177) Champa maaŋgaa-væ pari-čč̆ŭ uri-čč̆ŭ narŭkk-i
Champa[NOM] mango-ACC pluck-ASP peel-ASP chop-ASP
uppŭpoo-ṯṯŭ uuravæ-čč̆ŭ uurugaase-ɳʂŭ saappiḍ-uv-aaḷ.
brine-ASP marinate-ASP pickle-ASP eat-FUT-3FSG
"Champa will pluck, peel, chop, brine, marinate, pickle, and eat the mango."

(178) Champa maaŋgaa-væ maratt-ŭlirŭndŭ pari-čč̆ŭ uri-čč̆ŭ
Champa[NOM] mango-ACC tree-ABL pluck-ASP peel-ASP
tuɳḍam tunɳḍamaagæ narŭkk-i uppŭpoo-ṯṯŭ eɳɳæ-læ nallaa
small small chop-ASP brine-ASP oil-LOC well
uurave-čč-ŭ kaaramaagæ uurugaase-njŭ saappiḍ-uv-aaḷ.
marinate-ASP spicy pickle-make-PST eat-FUT-3FSG
"Champa will pluck the mango from the tree, peel it, chop it into small pieces, brine it, marinate it in oil well, pickle it (nice and) spicy, and eat it."

(179) Champa maaŋgaa-væ pari-čč̆ŭ uri-čč̆ŭ narŭkk-i
Champa[NOM] mango-ACC pluck-ASP peel-ASP chop-ASP
uurave-čč̆ŭ uppŭpoo-ṯṯŭ saappi-ṯṯŭ uurugaa-sej-v-aaḷ.
marinate-ASP brine-ASP eat-ASP pickle-make-FUT-3FSG
"Champa will pluck, peel, chop, marinate, brine, eat, and pickle the mango."

The sentence in (176) involves a series of lexical verbs, each of which can also occur alone in a clause with its meaning unchanged: this is the standard way to express the intended meaning of this sentence. (177) shows that the event as a whole may still be set in the future. This tense is marked on the ultimate verb in the string which is also marked for agreement; all preceding verbs still bear the frozen "past-tense" suffix, glossed here as ASP. The structure in (178) brings to our attention another important property of serial verbs in Tamil, namely that each verb in a sequence may be individually modified by an adjectival, adverbial, or prepositional phrase which precedes the verb in question. Finally, (179) shows us that the lexical verbs in a serial-verb sequence may be reordered relative to each other: all 7! permutations of the seven verbs in (179) are theoretically possible (modulo pragmatic markedness, but the resultant sentence would nevertheless be grammatically licit). A theoretical analysis of the internal structure of these constructions would take us too far afield. For the purposes of the current discussion, it is enough to keep in mind that lexical verbs in a serial verb formation in Tamil each bear a frozen past-tense suffix and may be reordered and further modified at will.

Now consider the sentences below, the members of which differ from one another only with respect to the presence vs. absence of *kol* on the verb:

(180) Kadavŭ tarakk-ir-adŭ.
door[SG-NOM] open-PRS-3NSG
"The door is opening."

(181) Kadavŭ tara-ndŭ-kol-gir-adŭ.
door[NOM] open-ASP-*kol*-PRS-3NSG
"The door is opening-*kol*."

At first glance, it looks like *kol* is a serial verb just like the others we've seen in (176)-(179) above. It occurs adjacent to another verb whose stem is marked with a frozen aspectual suffix just like in the other cases. *kol* itself, being the final verb in the two-membered sequence in (181) has the (real) tense and agreement information suffixed to it. However, there is already one noticeable difference between (181) and the serial verb constructions: the meaning contributed by *kol* is significantly less "meaty" than that contributed by the fully lexical verbs in (176)-(179). I.e. although *kol* introduces a perspectival semantics, it crucially doesn't denote a separate event the way these other verbs do the main event is still the door-opening event; *kol* merely adds some *extra* information about the manner in which the door opened.

There is indeed a slightly older use of *koɭ* as a verb meaning HAVE or TAKE or HOLD: Krishnamurti (2003, 463) claims, for instance, that modern *koɭ* comes from Proto-Dravidian **koɭ/konɳ* meaning TAKE/SEIZE/BUY. The sentence in (182) illustrates this use. Such a sentence is, however, not entirely grammatical in Modern colloquial Tamil (see Steever 2005, and Annamalai (1999) for discussion, and Jayaseelan (2004) for related discussion about Malayalam *koɭ*).[5]

(182) * Sudha uɳarcci ko-ɳɖ-aaɭ.
Sudha[NOM] emotion-ACC get/have-PST-3FSG
"Sudha got emotional." (Intended meaning).

That said, we will see that the kind of meaning contribution that *koɭ* makes is, perhaps unsurprisingly, very closely linked to its original lexical meanings.

For the purposes of the current discussion, the fact that *koɭ* cannot occur as the main verb of a clause alone suggests that it doesn't instantiate a full-on lexical verb. Additionally, *koɭ*, like other functional information, follows all the lexical verbal material in a serial verb sequence (183). That is, although *koɭ* may be succeeded by other functional morphemes, a lexical verb-stem may never occur after it (184) with the rough meaning given below:

(183) Sudha dosai-jæ araeccŭ-ŭ vaatt-ŭ saappiʈʈ-ŭ-kko-ɳɖ-aaɭ.
Sudha[NOM] dosa-ACC grind-ASP bake-PST eat-ASP-*koɭ*-PST-3FSG
"Sudha ground, baked, and ate the dosas (for herself)"

(184) * Sudha dosai-jæ araeccŭ-ŭ vaatt-ŭ koɳɖ-ŭ saappi-ʈʈ-aaɭ.
Sudha[NOM] dosa-ACC grind-ASP bake-ASP *koɭ*-ASP eat-PST-3FSG
"Sudha ground, baked, and ate the dosas (for herself)" (Intended)

Finally, *koɭ* itself may not be further modified by an adjunct phrase. Thus, the adverb *veegamaagæ* (QUICKLY) may modify the verb OPEN

[5]Purists may protest that the sentence in (182) is indeed grammatical in "sen Tamil" (literally: "pure Tamil", the term used for the formal and classical varieties of Tamil) which is still the modern written form of this diglossic language. It is important to remember, however, that the sentence in (182) will typically never be uttered in the standard spoken varieties of Tamil. As such, it is degraded in its colloquial use.

(186) but when it is placed in a position where it could only modify *koḷ*, the resultant sentence is sharply ungrammatical (187):

(185) Kadavŭ tarand-ŭ-kko-ɳḍ-adŭ.
door[NOM] open-ASP-*koḷ*-PST-3NSG
"The door opened (itself)"

(186) Kadavŭ veegamaagæ tarand-ŭ-kko-ɳḍ-adŭ.
door[NOM] quickly open-ASP-*koḷ*-PST-3NSG
"The door opened (itself) quickly."

(187) * Kadavŭ tarand-ŭ veegamaagæ ko-ɳḍ-adŭ.
door[NOM] open-ASP quickly *koḷ*-3NSG
"The door opened (itself) quickly." (Intended)

Based on this constellation of facts from serial verb constructions, I conclude that *koḷ* in modern Spoken Tamil is not a full-fledged lexical verb but spells out a head that is more functional in nature.

7.3.2 Option 2: *koḷ* spells out an applicative head

Pylkkänen (2008) argues that a range of argument-types crosslinguistically, like beneficiary, goal and source DPs, are not "core" arguments of the lexical predicate but are introduced by special (and frequently silent) functional heads which she terms applicatives, in analogy to the term used in the Bantu literature (Marantz 1993) for overt morphemes which play this role. She also develops an articulated mapping between the typological inventory of applicative types across languages and their syntactico-semantic contributions in designated functional projections in the argument structure. Applicatives in Pylkkänen's system are thus formally classified in terms of whether they are merged high or low and also with respect to what types of complements (verb or root or phase or something else) they combine with compositionally.

Given the functional nature of *koḷ* just discussed, we might propose that *koḷ* is nothing but a Pylkkänen-style applicative – i.e. a functional head that introduces a particular type of thematic argument and relates it to the rest of the structure in its scope. Such a proposal is attractive because it would capture what, we will end up proposing, *koḷ* does, in some sense, namely that it creates a perspectival relationship between an argument and the rest of the predicational structure in its scope. There is a crucial distinction, however: this is the fact that *koḷ* does not introduce a *new* thematic argument but adds additional thematic information to an argument that has already been merged below it. As far as I know, none of the applicatives of the kind discussed by Pylkkänen are imbued with

this property – rather, they all represent functional heads that create a new thematic argument and relate it to the predicate in their scope (Butt 2005, comparing applicatives with verbal elements similar to *koɭ* makes the same point).

I will thus propose that, although *koɭ* does assign a thematic role to a DP in its specifier, as a Pylkkänen (2008)-style applicative is argued to do, it is fundamentally different because this DP is also assigned a thematic role by the main event predication in its scope.

7.3.3 Option 3: *koɭ* is a light verb in a complex predication

A third option is that *koɭ* spells out part of a complex predicate structure. Butt (2005) defines a complex predicate as follows:

(188) **Definition of a complex predicate:** Butt (2005, 1)
"A complex (polyclausal) argument structure that corresponds to a monoclausal functional structure (a single subject; a single primary event predication)."

What type of predicate might *koɭ* represent in the complex-predicate structure? Turning again to Butt (2005, 1), we see that "a complex predicate consists of a main predicational element (noun, verb or adjective) and a light verb that is usually the syntactic head of the construction". In our discussion of serial verbs in Section 7.3.1 we have seen syntactic evidence showing that *koɭ* is not a main (verbal) predicate; we must thus consider whether *koɭ* is a light verb.

Given the templatic definition of a complex predicate in (188), this initially seems plausible. *koɭ*-structures *are* monoclausal – this can be shown by the fact that there is a single overt subject and that the clausal agreement reflects the ϕ-features of this subject, just like in simple *koɭ*-less structures:

(189) **Intransitive clause:**

a. Paanæ odæ-nʤ-**adŭ**.
pot[NOM] break-PST-3NSG
"The pot broke."

b. Paanæ odæ-nʤŭ-kko-ɳɖ-**adŭ**.
pot[NOM] break-ASP-*koɭ*-PST-3NSG
"The pot broke-*koɭ*."

(190) **Transitive clause:**

a. Paijan pustagatt-æ paɖi-tt-**aan**.
boy[NOM] book-ACC read-PST-3MSG
"The boy read the book."
b. Paijan pustagatt-æ paɖi-ttŭ-kko-ɳɖ-**aan**.
boy[NOM] book-ACC read-ASP-*ko*ɭ-PST-3MSG
"The boy read-*ko*ɭ the book."

But despite this simplicity of functional structure (to borrow Butt's terminology for a moment), a *ko*ɭ-structure is complex argument-structurally, as we have already seen. The thematic selectiveness of *ko*ɭ with respect to the verb in its complement is also characteristic of many light verbs (Butt 2005). All this lends credence to the idea that *ko*ɭ is a light verb in a complex predication.

At the same time, there *are* important differences. The main difference is that elements that are categorized as light verbs in the Butt (1995) sense determine the valency of the predicates they combine with – i.e. the light verb determines whether the complex predicate construction is transitive or intransitive. However, as we have seen in some detail in Section 6.2, *ko*ɭ doesn't affect the valency of the predicate it combines with: it may attach to transitives, unergatives, and unaccusatives alike without altering their valencies. Additional support for the idea that the meaning contribution of *ko*ɭ is divorced from voice comes from our finding, in Section 7.1, that *ko*ɭ is merged above the Kratzerian Voice head.

This is an important distinction which makes a light-verb analysis of *ko*ɭ, at least in the sense of Butt (1995), unviable.

7.3.4 Proposal: *ko*ɭ is a semi-functional restructuring verb

In the sections above, we have looked fairly closely at exactly what sort of syntactic object *ko*ɭ spells out. On the one hand, the meaning of *ko*ɭ is too insubstantial to warrant its treatment as a full-on lexical verb that occurs "serialized" onto other predicates – a conclusion that receives further support from the fact that *ko*ɭ-structures behave quite differently from standard serial-verb constructions in Tamil. At the same time, *ko*ɭ doesn't seem to represent a more functional element like a Pylkkänen (2008)-style applicative head. Finally, we have just seen that an analysis of *ko*ɭ as a light verb in a Butt (1995)-style complex predication is also not empirically supported.

Here, I will show that *ko*ɭ corresponds most closely to a semi-functional

restructuring verb in the sense of Wurmbrand (2001). A restructuring predicate in the sense of Wurmbrand (2001) refers to a predicate that is part of a monoclausal structure, in the sense that there is a single T head and a single subject. It doesn't mean, however, that the predicate comprises a single event. To this extent, Wurmbrand's definition of restructuring verb corresponds rather closely to Butt's description of a complex predicate in (188). The relevant difference for us is the fact that Wurmbrand's restructuring predicates includes a wider membership of verbs, many of which do not have valency-altering properties.

Wurmbrand distinguishes three classes of restructuring predicate: lexical, pure functional, and semi-functional. Restructuring verbs of the lexical kind, such as German *versuchen* (TRY), are considered to be full verbs, thus instantiate the syntactic category V. As such, they also exhibit some optionality in the size and nature of their complements – i.e. they don't only appear in restructuring contexts – and, furthermore, are capable of assigning θ-roles.[6] It is clear that *koḷ* is not such a predicate: our comparison of this morpheme with prototypical full verbs in serial verb constructions has already shown this. Furthermore, as we've also seen, it cannot show up as the sole verb of a clause in colloquial registers of Tamil. Pure functional restructuring verbs, on the other hand, are explicitly not full verbs: Wurmbrand proposes that that they may occupy nodes such as Mod. Such verbs, in direct contrast to their lexical counterparts, bear a tight relationship with their complements: there is no optionality in the size or nature of their complements. Wurmbrand proposes that raising verbs (e.g. German *scheinen* (SEEM)), and most modals instantiate this type of predicate. However, *koḷ* doesn't seem to be a purely functional predicate in this sense either: a crucial difference has to do with the fact that pure functional verbs don't assign θ-roles. In contrast, we will see that the meaning-contribution of *koḷ* does place thematic restrictions on the type of DP that may merge in its specifier.

This distinction leads us directly to Wurmbrand's third type of restructuring verb: the semi-functional kind (e.g. motion verbs like *come* and *go* and direct perception verbs like *see*). These are characterized as being like pure functional verbs in realizing functional heads as opposed

[6]The question of whether a restructuring predicate may assign a thematic role at all is a matter of some contention in the literature. For instance, Cinque (2004) argues that a restructuring verb is always functional, even when it doesn't seem to be, and thus may never assign a thematic role. The level of detail and intricacy that adjudicating the pros and cons of this argument requires is beyond the scope of the current discussion and also, at some level, orthogonal to its concerns. I thus do not take a stance in this debate here but merely acknowledge its existence.

to lexical ones and in placing a tighter restriction in the structural properties of their complements. We have already seen evidence of this in our investigation of the distributional and modificational restrictions on *ko⌊* in contrast to those on full verbs in a serial verb construction. However, semi-functional verbs are like restructuring predicates of the lexical kind in the sense that they *do* assign θ-roles to their subjects. Wurmbrand argues that such predicates occupy the *v* area of the clause, a point that jibes quite well with the position of *ko⌊* in the linear sequence, given in (146) above. Without getting too bogged down in the *nomenclatural* details of what exactly a semi-functional restructuring predicate is, we can nevertheless see that the kind of element that *ko⌊* represents bears a close correspondence with such an entity. I will thus propose that *ko⌊* is a semi-functional restructuring verb of this type but be explicit about precisely what is entailed by my use of this term.

In classifying *ko⌊* as a restructuring *verb*, we place it on a par with other selectional predicates (like THINK) which introduce a *new* functional sequence in their scope with a potentially complex event structure. There are morphological, syntactic, and semantic advantages to such a claim. On the morphological end, we have observed that the aspectual morpheme that *ko⌊* occurs directly above looks the same as the past-tense morpheme (that occurs above *ko⌊*). On the strength of this type of evidence, Amritavalli and Jayaseelan (2005) have argued that there is no tense marking in Dravidian, only aspectual marking. Under an approach that treats *ko⌊* as restarting the functional sequence afresh in its complement, this kind of recursion is, however, precisely what we expect. I.e. under this analysis, we would expect the aspectual marking to also occur above *ko⌊* with a covert tense marker above it yielding "real" tense effects. On the syntactic side, putting *ko⌊* on a par with full-fledged verbs (like propositional THINK) will prove to be significant because I will end up arguing, based on the investigation of locally-anteceded *ta(a)n*-structures, that *ko⌊* selects a predicate-denoting complement of a particular size. I.e. what we have with *ko⌊* is a case of real embedding, where the phrase that *ko⌊* selects constitutes an independent functional sequence, rather than a continuation of the sequence including *ko⌊* itself. I will claim that it is autonomous functional sequences like this that are the potential locus for Perspectival anchoring. On the semantic side, this type of analysis helps to underline another principled distinction between the type of head *ko⌊* spells out and a Pylkkänen (2008)-style applicative head: whereas Pylkkänen argues that an applicative head is event-modificational (combining with its complement via Event Identification (Kratzer 1996)), we are proposing that it is predicational and combines with its complement

via standard function application.

In calling *koլ* a *restructuring* verb, we are explicitly distinguishing it from full-fledged verbs by claiming that the functional structure it selects is smaller than the ones that these do, with the result that it is part of a monoclausal structure, with a single T head and subject, and not a biclausal one. Finally, in classifying *koլ* as a *semi-functional* restructuring verb, I mean the following:

(i) It spells out a functional head and not a lexical one. Henceforth I will simply call this head F.

(ii) It bears a tight structural connection with its complement. Specifically, it always selects a resultative AspP whose Asp_{res} head has the following denotation:
$[\![Asp_{res}]\!] = \lambda R_{<s,t>} \lambda s_s \exists e.R(e) \wedge Result(e, s)$

(iii) It assigns a θ-role to a DP in its specifier, one that, incidentally, already has another θ-role assigned to it by the main event predication in its scope.[7]

I will elaborate on these points in the course of the remainder of the chapter, but this sets the broad parameters of its nature.

7.4 The interpretation and distribution of *koլ*

Now that we have established what type of syntactic entity *koլ* is and where in the argument structure it is merged, we can move on to the question of what it actually means. To this end, we will investigate the compatibility of this morpheme with the various classes of predicate described in Levin (1993) and see, furthermore, what kind of interpretive difference the addition of *koլ* makes in such cases. Our investigation will reveal that, for the majority of the fifty seven verb-classes in Levin's sample, the addition of *koլ* is truly optional, with its suffixation to such verbs making a subtle but productive change in meaning. However, not all verbs will turn out to be quite so neutral in this sense: some will be seen to prefer or even require *koլ* whereas others will be seen to disprefer or reject it altogether. Nevertheless, we will see that, contrary to the claims made in the descriptive literature, the meaning-contribution of *koլ* *is* consistent for all verb-classes and, furthermore, that the various

[7] This is, of course, a violation of the θ-criterion, a point I return to later.

degrees of its compatibility across different verb-classes *are* predictable from its meaning.

The addition of *koɭ* will be seen to contribute the interpretation that the highest argument of the main event predication in its scope *comes to hold*, in a mental or physical sense, the result state of this event. In other words, the highest argument will be interpreted as a mental or physical *locus* from which the result state of the main event predication is viewed. We will propose a precise denotation for *koɭ* that captures these properties and integrates them with what we already know about the syntactic status and structural position of this morpheme.

7.4.1 Verbs that optionally take *koɭ*: a first definition

As mentioned above, the majority of verbs discussed in Levin (1993) optionally occur with *koɭ* – this can thus be treated as the default pattern. Examples of verbs that belong in this category are: PUT-verbs like *uuttŭ* (POUR), PUSH/PULL VERBS like *taɭɭŭ* (PULL), CONCEALMENT-class predicates like *oɭi* (HIDE), verbs of the POKE-class like *kuttŭ* (PIERCE), THROWING-verbs like *odæ* (KICK), CREATION-class predicates such as *vaarŭ* (BAKE), CONTACT-verbs like *taʈʈŭ* (TAP), HOLD/KEEP verbs like *puɖi* (HOLD), DESTROY-verbs like *naasam-sej* (DESTROY), and the transitive variants of change-of-state verbs like *taræ* (OPEN), *arŭ* (SEVER) and *oɖæ* (BREAK).

What is the meaning that *koɭ* contributes when it is added to such verbs? We will try to determine this by comparing the *koɭ*-sentences with their *koɭ*-less counterparts below:

(191) *oɭi* (HIDE):

a. Raman paɳatt-æ oɭi-čč-aan.
Raman[NOM] money-ACC hide-PST-3MSG
"Raman hid the money."

b. Raman paɳatt-æ oɭi-ččŭ-ko-ɳɖ-aan.
Raman[NOM] money-ACC hide-ASP-*koɭ*-PST-3MSG
"Raman hid-*koɭ* the money."

(192) *oɖæ* (BREAK):

a. Raman kaal-æ oɖæ-čč-aan.
Raman[NOM] leg-ACC break-PST-3MSG
"Raman broke (his) leg."

b. Raman kaal-æ oɖæ-ččŭ-ko-ɳɖ-aan.
Raman[NOM] leg-ACC break-ASP-*koɭ*-PST-3MSG

"Raman broke-*koɭ* (his) leg."

(193) *uuttŭ* (POUR):

a. Mansi paal-æ uutt-in-aaɭ.
Mansi[NOM] milk-ACC pour-PST-3FSG
"Mansi poured the milk."

b. Mansi paal-æ uutt-i-ko-ɳɖ-aaɭ.
Mansi[NOM] milk-ACC pour-ASP-*koɭ*-PST-3FSG
"Mansi poured-*koɭ* the milk."

(194) *poottŭ* (COVER):

a. Raman poorvæ-jæ poott-in-aan.
Raman[NOM] blanket-ACC cover-PST-3MSG
"Raman put on/covered the blanket."

b. Raman poorvæ-jæ poott-i-ko-ɳɖ-aan.
Raman[NOM] blanket-ACC cover-ASP-*koɭ*-PST-3MSG
"Raman cover-*koɭ* the blanket."

(195) *muuɖŭ* (CLOSE):

a. Geetha kadav-æ muuɖ-in-aaɭ.
Geetha door-ACC shut-PST-3FSG
"Geetha shut the door."

b. Geetha kadav-æ muuɖ-i-ko-ɳɖ-aaɭ.
Geetha door-ACC shut-ASP-*koɭ*-PST-3FSG
"Geetha shut-*koɭ* the door."

Let us analyze the contribution of *koɭ* for each of these sentences. The sentence in (191a) has the simple meaning that Raman hid the money. With the addition of *koɭ* in (191b), the sentence takes on the extra reading that Raman hid the money and, furthermore, that he hid the money on his own person (say in the pockets of his pants). Given the right discourse context, a different reading is also possible, namely that Raman hid the money for his own benefit; in this case, it is not necessary that the money be hidden on Raman's person. In the next pairing, (192a) has the interpretation that Raman broke his own or someone else's leg; with the addition of *koɭ*, in (192b), the sentence takes on the reading that the breaking event happened to Raman's own leg or the pragmatically odder one that Raman deliberately broke his (or someone else's) leg with a view to ultimately benefitting himself.

Interestingly, these two readings seem to have rather different effects on the interpretation of the resultant sentence. Under the former type of reading, the AGENT of the main event acquires a more "PATIENT"-like interpretation by virtue of being the physical locus of the outcome

of the event. In the latter, however, the interpretation seems rather different – the entity denoted by the external argument, by going out of his/her way to get the outcome (in other words, the result state) of the event to benefit or otherwise affect himself/herself, seems to be more AGENT- than PATIENT-like. As noted earlier, the existence of this kind of apparent tension has been used as empirical support, in the literature (see e.g. Schiffman 1995), for the idea that *ko̱l* resists a uniform semantic characterization.[8] With the next two pairs of sentences, given in (193) and (194), our intuitions about the meaning contribution of *ko̱l* seem to carry over pretty exactly. The *ko̱l*-variant in (193b) either has the reading that Mansi poured the milk *on* herself or *for* herself. Similarly, (194b) has the reading that Raman ended up with the blanket covering himself or that he covered someone else with a blanket for his own benefit.

Our final minimal pair in (147) has a slightly different reading. The *ko̱l*-less sentence in (195a) asserts simply that Geetha shut the door. With the addition of *ko̱l* in (195b), we again get one of two possible readings. One of the readings is entirely analogous to what we have seen so far – this is the interpretation that Geetha shut the door in order to benefit herself (e.g. because she wanted some quiet from loud noises outside). The other reading, however, cannot *literally* be that Geetha shut the door *on* herself – it is unclear what this would mean in any case. In other words, *Geetha* isn't interpreted as being "PATIENT"-like in the same way as in the other cases seen thus far. The resultant interpretation is the more figurative one that Geetha shut herself in – i.e. that Geetha shut the door, thereby enclosing herself in the resulting physical space. In other words, the resultant "shutness" of the door is interpreted from Geetha's physical point-of-view (see also Selvanathan 2009, for a similar proposal).

Thus, the addition of *ko̱l* seems to yield readings along two different dimensions: a mental one and a physical one. The mental reading is highly reminiscent of what Sells (1987) characterizes as "SELF" and which he defines as representing "one whose mind is being reported" (Sells 1987, 455). Analogously, the physical reading under *ko̱l* seems to correspond to Sells' "PIVOT"-role; Sells states: "I understand PIVOT in a very physical sense . . . ; if someone makes a report with Mary as the PIVOT, that person is understood as literally standing in Mary's shoes" (Sells 1987, 455-456). Given these initial intuitions, the two sets of readings we have discerned

[8] A similar type of apparent tension has been observed for the case of *get*-passives in English, seen in sentences like "Mary got her teeth pulled out." This sentence could either mean that Mary was the unfortunate victim of a teeth-pulling event or that Mary went out of her way to get them pulled out (McIntyre 2011).

for each of the *ko*⌊-sentences above may be summarized as follows:

(196) **Informal description of the interpretation of a *ko*⌊-sentence:**

Reading I: For an event P and an AGENT X of P, the outcome (or result state) of P comes to be located in X's mental space. I.e. X comes to hold the outcome of P in his/her mind.

Reading II: For an event P and an AGENT X of P, the outcome (or result state) of P comes to be located in X's physical space. I.e. X comes to hold the outcome of P on his/her physical person.

We observed above that the two types of readings have different effects – yielding a more "PATIENT"-like interpretation of the external argument in Reading II and a more "AGENT"-like one in Reading I. But we now see that these effects must be *implicational* and not a part of the event-semantics because the two readings clearly share a common core. We can now describe the contribution of *ko*⌊ as follows:

(197) **Contribution of *ko*⌊ – Version 1:**
ko⌊ attaches to the derived result state of a main event predication such that the AGENT of this event *comes to hold* the result state of this event in their mental or physical space.

What determines whether the AGENT represents a mental or physical locus of the result state of the main event? The answer, based again on our investigation of the minimal pairs listed under (191)-(147), seems to be that this is due to a combination of what the verb in question itself "means" and extra retrictions imposed by discourse-pragmatic information. Certain verbs, like SHRINK and POUR, are clearly predisposed towards an interpretation along the spatial plane whereas others lend themselves more readily towards a reading along the mental one. At the same time, there *is* a certain degree of flexibility: the final choice of one or other can only be made after due consideration of relevant discourse factors pertaining to such variables as speaker intent, common ground, presupposition, conversational implicature, salience, and the like. In other words, there is an underspecification in the type of semantics that *ko*⌊ introduces: *ko*⌊ simply contributes a *coming-to-hold* semantics without specifying whether the locus of this holding is the mental or physical space of the AGENT.

7.4.2 *ko*ɭ in unaccusatives – updating the definition

The definition in (197) above states that the perspectival semantics of *ko*ɭ applies to the AGENT of the main event. But this cannot be quite right. Recall our observation from the previous chapter, that *ko*ɭ may be suffixed onto the *in*transitive variants of change-of-state verbs:

(198) Paanæ oɖæ-nʛŭ-ko-ɳɖ-adŭ.
pot[NOM] break-INTR.ASP-*ko*ɭ-PST-3NSG
"The pot broke-*ko*ɭ."

(199) Saʈʈæ (vejjal-læ) suru-ŋgi-ko-ɳɖ-adŭ.
shirt[NOM] (heat-LOC) shrink-INTR.ASP-*ko*ɭ-PST-3NSG
"The shirt shrank-*ko*ɭ (in the heat)."

The pot and shirt in (198) and (199), respectively, are clearly not the AGENTS of the breaking and shrinking events described in these sentences – they are the PATIENTS corresponding to entities undergoing the change-of-state process denoted by the verbs. (198) and (199) are, in other words, unaccusative structures.

Furthermore, as we've already seen, the addition of *ko*ɭ is strictly optional in such sentences. Thus the *ko*ɭ-less counterparts of (198) and (199) are fully grammatical:

(200) Paanæ oɖæ-nʛ-adŭ.
pot[NOM] break-INTR.PST-3NSG
"The pot broke."

(201) Saʈʈæ (vejjal-læ) suru-ŋg-ij-adŭ.
shirt[NOM] (heat-LOC) shrink-INTR-PST-3NSG
"The shirt shrank (in the heat)."

Most importantly, these sentences are still interpreted as unaccusatives – recall that *ko*ɭ doesn't affect the valency of the predicate it attaches to – showing that what *ko*ɭ contributes is not unaccusativity but something else. The meaning differences between the *ko*ɭ- and *ko*ɭ-less minimal pairs above may be described as follows. Whereas the sentences in (200) and (201) merely state that the pot broke and that the shirt shrank in the heat, respectively, their *ko*ɭ-variants in (198) and (199) may be roughly translated as: "The pot got/became broken" and "The shirt got/became shrunk in the heat", respectively. In other words, with the addition of *ko*ɭ, the breaking and shrinking events take on an additional PIVOT-like reading wherein the outcome of these events is evaluated from the physical space of the pot and shirt, respectively.

koɭ frequently marks change-of-state unaccusatives in Tamil, as we've already noted. The contribution of *koɭ* in these structures may be uniformly characterized in the manner given above for SHRINK and BREAK. This is illustrated below:

(202) *karæ-* (DISSOLVE):

a. Maattŭræ taɳɳi-læ karæ-nʤ-adŭ.
pill[NOM] water-LOC dissolve-PST-3NSG
"The pill dissolved in the water."

b. Maattŭræ taɳɳi-læ karæ- nʤŭ-ko-ɳɖ-adŭ.
pill[NOM] water-LOC dissolve-ASP-*koɭ*-PST-3NSG
Lit: "The pill got dissolved in the water."
The result state of the pill's dissolving in the water comes to be located in the physical space of the pill.

(203) *vaɭæ-* (BEND/CURVE):

a. Aarŭ vaɭæ-nʤ-adŭ.
river[NOM] bend-PST-3NSG
"The river bent/curved."

b. Aarŭ vaɭæ-nʤŭ-ko-ɳɖ-adŭ.
river[NOM] bend-ASP-*koɭ*-PST-3NSG
"The river got bent/curved."
The result state of the river's bending comes to be located in the physical space of the river.

(204) *kiri* (TEAR):

a. Pustagam kiri-nd-adŭ.
book[NOM] tear-PST-3NSG
"The book tore."

b. Pustagam kiri-ndŭ-ko-ɳɖ-adŭ.
book[NOM] tear-ASP-*koɭ*-PST-3NSG
"The book got torn."
The result state of the book's tearing comes to be located in the physical space of the book.

The interpretive parallel to the meaning contribution of *koɭ*-transitives, such as those discussed in the previous section, is clear. In both types of structures, the addition of *koɭ* yields a structure wherein the outcome of the main event comes to be mentally or spatially located on one of the arguments of that event.

But the choice of argument isn't random. Rather, THEMES/PATIENTS are associated with a *koɭ*-semantics only when they are the sole arguments of the clause, as in unaccusative structures like those above. But in

transitives, where both AGENTS and PATIENTS/THEMES are present, it is invariably the former that get associated with the semantics of *ko*l̩. The sentences in (205) (repeated from (198)) and (206) serve to drive this point home:

(205) Paanæ oɖæ-nʤŭ-ko-ŋɖ-adŭ.
Pot[NOM] break-INTR.ASP-*ko*l̩-PST-3MSG
"The pot got broken."

(206) Raman paanæ-jæ oɖæ-ččŭ-ko-ŋɖ-aan.
Raman[NOM] pot-ACC break-TR.ASP-*ko*l̩-PST-3MSG
"Raman got the pot broken." (lit)

In the unaccusative structure in (205), the entity that undergoes the change of state, represented by the affected or PATIENT-like argument, is the pot. The semantics of *ko*l̩ applies to this argument – which is, in fact, the sole argument of the sentence. In the transitive variant given in (206), the affected entity is still the internal PATIENT argument corresponding to the pot. However, in this case, the semantic-contribution of *ko*l̩ applies, not to this affected internal argument, but to the external AGENT argument *Raman*: the sentence in (206) states that Raman broke the pot and that Raman comes to then evaluate the outcome of this event from his mental or physical center. I.e. the result state of this pot-breaking event is rebounded back to Raman who is thus the entity that comes to represent the mental or physical locus of this result state.

There are no exceptions to this rule, suggesting that the designated argument is not chosen on the strength of its thematic relationship with its predicate but on a "blind" structural basis. Based on this discussion, I propose the following updated description of the semantics of *ko*l̩:

(207) **Contribution of *ko*l̩ – Version 2:**
*ko*l̩ attaches to the derived result state of a main event predication such that the highest argument of this event *comes to hold* the result state of this event in their mental or physical space.

In the following sections, I will look at verbs that are degraded with *ko*l̩. We will see that these verbs fall into two groups:

(i) Verbs whose denotations already encode the semantics of *ko*l̩.

(ii) Verbs whose denotations are incompatible with the semantics of *ko*l̩.

7.4.3 Verbs which already possess a semantics similar to *koḷ*

The following types of predicate will be seen to be degraded with *koḷ* because their semantics already encodes something akin to that of *koḷ*:

(i) Psych-predicates

(ii) Inherently-directed motion verbs

(iii) Alter-benefactive verbs

7.4.3.1 Psych-predicates and *koḷ*

The class of psych-verbs has received a great deal of attention in the Dravidian literature precisely because such verbs are typically incompatible with *koḷ* (Lidz 2004, for Kannada and Annamalai (1999) for Tamil) – a property we can now explain. A psych predicate, by its very nature, represents the inner mind – capturing the thoughts, feelings, opinions, and so on – of a sentient entity. In other words, the EXPERIENCER of a psych-predicate already represents a mental locus, not just with respect to the result state of the psych-event but also with respect to its initiation and process components. As such, its denotation already contains the meaning of *koḷ* (described as in (207) above); assuming that the EXPERIENCER of this verb is also its highest argument, *koḷ*-suffixation is ungrammatical or degraded.

This is illustrated with the eventive psych-predicate structures below which may have a dative experiencer subject (as in (208b)) or a nominative one (as in (209b))

(208) *aattiram-vaa* (ANGER-COME) – Dative EXPERIENCER:

a. Sudha-vŭkkŭ rombæ aattiram-va-nd-adŭ.
Sudha[NOM] very anger-come-PST-3NSG
"Sudha got very angry."
Lit.: "(Much) Anger came to Sudha"

b. * Sudha-vŭkkŭ rombæ aattiram-va-ndŭ-ko-ɳḍ-adŭ.
Sudha[NOM] very anger-come-ASP-*koḷ*-PST-3NSG
"Sudha got very angry." (Intended)

(209) *bayappaɖŭ* (FEAR) – Nominative EXPERIENCER:

a. Veena rombæ bayappa-ʈʈ-aaḷ.
Veena[NOM] very fear-PST-3FSG
"Veena felt very scared."

b. ? Veena rombæ bayappa-ʈʈŭ-kko-ɳɖ-aaɭ.
Veena[NOM] very fear-ASP-*koɭ*-PST-3FSG
"Veena got very scared." (Intended)

Stative psych-verbs may also occur with dative or nominative EXPERIENCERS and are also degraded with *koɭ*:

(210) *piɖi* (LIKE) – Dative EXPERIENCER:
a. Radha-vŭkkŭ Krishnan-æ piɖi-tt-adŭ.
Radha-[DAT] Krishnan-ACC like-PST-3NSG
"Radha liked Krishnan."
b. ??/* Radha-vŭkkŭ Krishnan-æ piɖi-ttŭ-ko-ɳɖ-adŭ.
Radha-DAT Krishnan-ACC like-ASP-*koɭ*-PST-3NSG
"Radha liked Krishnan (for her own benefit)." (Intended)

(211) *verŭ* (HATE) – Nominative EXPERIENCER:
a. Maya kaɳakk-æ verŭ-tt-aaɭ.
Maya[NOM] math-ACC hate-PST-3FSG
"Maya hated math."
b. ?? Maya kaɳakk-æ verŭ-ttŭ-ko-ɳɖ-aaɭ.
Maya[NOM] math-ACC hate-ASP-*koɭ*-PST-3FSG
"Maya hated math (for her own benefit)." (Intended)

There are two types of distinctions between the psych predicates illustrated here, and they both have an effect on *koɭ*-compatibility. One is the eventive vs. stative distinction which is evidenced by the sentences in (208)-(209) on the one hand, and those in (210)-(211) on the other, respectively. As we saw, both types of psych-predicates are degraded with *koɭ*, but stative psych predicates are a bit more degraded because they resist *koɭ*-suffixation for an additional reason, namely that, being inherently stative, they are not compatible with the addition of the result-state aspectual marker whose presence *koɭ* requires for its own suffixation. The second distinction has to do with the case-marking on the experiencer subject. Dative experiencers tend to be significantly more degraded with *koɭ* than do nominative ones. In the final chapter of this series on *koɭ*, I will model this distinction by proposing that *koɭ* requires the argument in its specifier to be marked nominative.

Moving on, recall that the semantics of *koɭ* applies to the highest argument of the event in its scope. This means that, if the EXPERIENCER argument is not the subject of a psych-predicate structure, we should expect *koɭ*-suffixation to be possible, in theory. This is indeed the case, as the sentences below illustrate:

(212) Raman Maya-væ bayamurutt-in-aan.
Raman[NOM] Maya-ACC frighten-PST-3MSG
"Raman frightened Maya."

(213) Raman Maya-væ bayamurutt-i-ko-ɳɖ-aan.
Raman[NOM] Maya-ACC frighten-ASP-*koɭ*-PST-3MSG
"Raman$_i$ frightened Maya (for his$_i$ own benefit)."

In (212)-(213), the EXPERIENCER *Maya* is the internal argument of the psych-verb, thus is not the highest argument – as such, the semantics of *koɭ* is not applied to it. Instead, in (213), *koɭ* modifies the agentive external argument *Raman* and yields the reading that the result state of the frightening event is represented in Raman's mind.

So far, these facts are just what we would expect given the description of *koɭ*'s meaning contribution in (207). But one point remains to be clarified. We have observed with other predicates (like POUR or COVER) that the semantics of *koɭ* may be interpreted as being along the mental or physical dimensions. We have just seen that, when the EXPERIENCER is the highest argument of the event, as in (208)-(209a) above, *koɭ*-suffixation is ruled out because the EXPERIENCER already serves as a mental locus in these cases. But why can't the EXPERIENCER then serve as a spatial locus with the addition of *koɭ*? The simple answer seems to be that psych verbs like those in (208)-(209a) lack a physical component to their denotations altogether. I.e. their meanings are encoded entirely on the mental plane – thus, such sentences are fully ungrammatical with *koɭ*.

7.4.3.2 Inherently directed motion verbs and *koɭ*

The semantics of an inherently directed motion verb includes a specification of the path and goal of motion. Thus, a verb of this type already includes the meaning that the motion ends up in a predetermined (physical) location or locative entity. A verb of this type is also degraded with *koɭ*, as might be expected, since its highest argument cannot be associated with a *koɭ*-semantics – i.e. it cannot be interpreted as the entity that becomes the physical locus of the result state of the motion. This is illustrated below:

(214) *viɻŭ* (FALL-DOWN):

a. Raman kiiɻæ viɻŭ-nd-aan.
Raman[NOM] down fall-PST-3MSG
"Raman fell down."

b. ?? Raman kiiɻæ viɻŭ-ndŭ-ko-ɳɖ-aan.
Raman[NOM] down fall-ASP-*ko*ɭ-PST-3MSG
"Raman fell down." (Intended)

(215) *vaa* (COME):

a. Raman viiʈʈ-ŭkk-kiʈʈæ va-nd-aan.
Raman[NOM] house-DAT-ALL come-PST-3MSG
"Raman came near the house."

b. ?? Raman viiʈʈ-ŭkk-kiʈʈæ va-ndŭ-ko-ɳɖ-aan.
Raman[NOM] house-DAT-ALL come-ASP-*ko*ɭ-PST-3MSG
"Raman came near the house." (Intended)

Interestingly, however, such sentences may be rendered grammatical if the highest argument is interpreted as being an AGENT rather than a THEME – in other words, if it may be interpreted as representing the *mental* locus of the outcome of the event. In a discourse scenario where the entity denoted by the highest argument of a directed-motion verb is construed as having deliberately initiated the event with a particular result state in mind, this argument will automatically count as coming to hold the mental locus of this result state. As such, *ko*ɭ-suffixation should be possible on such verbs in such cases. This reasoning seems to be correct. The sentence in (216) below is grammatical under the reading, enhanced by the addition of the agentive adverb *veeɳummŭnnŭ* (DELIBERATELY), that the entity denoted by *Raman* deliberately fell down in order to benefit himself in some way. Similarly, (217) is licit under the reading that Raman deliberately approached the house with a particular end in mind:

(216) Raman veeɳummŭnnŭ kiiɻæ viɻŭ-ndŭ-ko-ɳɖ-aan.
Raman[NOM] deliberately down fall-ASP-*ko*ɭ-PST-3MSG
"Raman fell down to benefit himself" (Rough translation)

(217) Raman veeɳummŭnnŭ viiʈʈŭkk-kiʈʈæ va-ndŭ-ko-ɳɖ-aan.
Raman[NOM] veeɳummŭnnŭ house-DAT-ALL come-ASP-*ko*ɭ-PST-3MSG
"Raman came near the house to benefit himself." (Rough translation)

The grammaticality of these sentences shows that, while psych-predicates like those discussed in Section 7.4.3.1 above may not be construed as involving a spatial component, spatial motion verbs like these *can* be construed as deliberate actions, involving (mental) volition and consciousness, given the right discourse context.

7.4.3.3 Alter-benefactive verbs and *koɭ*

Verbs like TEACH and GIVE, which belong to the class of alter-benefactives, involve the notion that the outcome of the denoted event benefits someone other than the AGENT – specifically, that it benefits a GOAL or BENEFICIARY. *koɭ* we have seen, on the other hand, seems to impose the reading that the AGENT receives (or comes to hold) the result state of the event in his/her mental space – an interpretation that a semantics of self-benefaction subsumes. Unsurprisingly, therefore, *koɭ*-suffixation is degraded, in the default pragmatic context, with such verbs:

(218) Maya Guruv-ŭkkŭ kaŋakk-æ sollikkuɖŭ-tt-aaɭ.
Maya[NOM] Guru-DAT math-ACC teach-PST-3FSG
"Maya taught Guru math."

(219) ?? Maya Guruv-ŭkkŭ kaŋakk-æ sollikkuɖŭ-ttŭ-ko-ɳɖ- aaɭ.
Maya[NOM] Guru-DAT math-ACC teach-ASP-*koɭ*-PST- 3FSG
"Maya taught Guru math." (Intended)

At the same time, there is no reason why the outcome of an event shouldn't, given the right pragmatic situation, be interpreted as benefitting someone else, as long as the AGENT continues to represent the mental locus of the event's result state. Here again, an adverb like *veeŋummŭnnŭ* (INTENTIONALLY) or ʤ*aakkaradæjaa* (CAREFULLY), which enforces the reading that the AGENT has a vested interest in the outcome of the main event improves the resultant *koɭ*-sentence considerably. (219) below is fully grammatical under a reading where Maya carefully teaches Guru mathematics in order, say, to ensure that she herself doesn't get fired from the job:

(220) Maya Guruv-ŭkkŭ kaŋakk-æ ʤaakkaradæjaa
Maya[NOM] Guru-DAT math-ACC carefully
sollikkuɖŭ-ttŭ-ko-ɳɖ-aaɭ.
teach-ASP-*koɭ*-PST-3FSG
"Maya$_i$ taught Guru math carefully (in order to benefit herself$_i$)."

Both the degradedness of *koɭ* with alter-benefactives in the pragmatically unmarked case and its improvement to near or full grammaticality under the right discourse context – are precisely what we expect given our description of the meaning contribution of *koɭ* in (207) above.

7.4.4 Verbs which are incompatible with a *koɭ*-semantics

The second set of predicates that are degraded with *koɭ* constitutes verbs which, by virtue of what they mean (and, by extension, of the types of thematic arguments they select), are incompatible with the mental or physical "coming to hold" semantics of *koɭ*. Here, I examine three such classes of predicate:

(i) Weather-verbs.

(ii) Raising predicates.

(iii) Predicates of (dis)appearance and occurrence.

The verbs belonging to the first two classes are considered not to have thematic participants at all (Chomsky 1981). As such, it should be trivially impossible for such verbs to co-occur with *koɭ*. The third class involves verbs that denote the creation or disappearance of their highest argument – in other words, the argument is not present through all the relevant stages of the event, thus doesn't count as a participant of the event in its entirety. Unsurprisingly, *koɭ*-suffixation is degraded in these cases as well:

(221) Weather verbs:

a. Maɻæ pen-ʤ-adŭ/*pen-ʤŭ-ko-ɳɖ-adŭ
rain[NOM] pour-PST-3NSG/*pour-ASP-*koɭ*-PST-3NSG
"It rained." (Lit: "The rain poured.")

b. Minnal minn-ij-adŭ/*minn-i-ko-ɳɖ-adŭ.
lightning[NOM] strike-PST-3NSG/*strike-ASP-*koɭ*-PST-3NSG
"Lightning struck."

(222) Raising verbs:

a. Vaɳɖi$_i$ [$_{TP}$ t$_i$ nagar-æ] aarambi-tt-adŭ.
truck[NOM] move-INF begin-PST-3NSG
"[The truck]$_i$ began [$_{TP}$ t$_i$ to move]."

b. * Vaɳɖi$_i$ [$_{TP}$ t$_i$ naga-æ] aarambi-ttŭ-ko-ɳɖ-adŭ.
truck[NOM] move-INF begin-ASP-*koɭ*-PST-3NSG
"[The truck]$_i$ began [$_{TP}$ t$_i$ to move]." (Intended)

(223) (Dis)appearance and occurrence verbs:

a. Raman se-tt-aan/*se-ttŭ-ko-ɳɖ-aan.
Raman[NOM] die-PST-3MSG/*die-ASP-*koɭ*-PST-3MSG
"Raman died."

b. Neettikkŭ oru vibaadattŭ naḍa-nd-adŭ/*naḍa-ndŭ-ko-ṇḍ-adŭ.
yesterday one accident[NOM] happen-PST-3NSG/*happen-ASP-*ko*ɭ-PST-3NSG
"Yesterday, an accident occurred."

Interestingly, even here, *ko*ɭ may be optionally suffixed on the predicate if the sole argument may be somehow construed as being agentive and as having participated in the event with a particular end in mind. For the types of sentences considered above, this would only be possible in a fantasty or dream scenario, where accidents could happen on purpose or dead people could come back to life and observe their own dying events in flashback. Far from invalidating our intuitions about the semantic contribution of *ko*ɭ, the possibility of such repairability actually strengthens them because it confirms our intuition that the addition of *ko*ɭ introduces a semantics which must be associated with a DP that has participated in all the relevant stages – initiation, process, and result – of the main event. If the designated argument has participated in some but not all stages of the event (as is the case in disappearance and occurrence verbs for instance) – the semantics introduced by *ko*ɭ may not be associated with it, yielding ungrammaticality.

7.4.5 Verb-classes which require or prefer *ko*ɭ

We have just looked at a series of predicates for which *ko*ɭ-suffixation is degraded. At the other end of the spectrum are verbs which prefer or require *ko*ɭ-suffixation. The verbs belonging to this set are either themselves telic or compatible with the addition of a derived result state. Furthermore, they all encode a middle-*like* interpretation (Kemmer 2003) corresponding, in this case, to the idea that the outcome of the main event is somehow rebounded back to one of the arguments of that event. While *ko*ɭ-structures are not middle constructions in the traditional sense because they do not represent a type of voice phenomenon, or interact with the valency of the predicate – a point we have already discussed in some detail – the descriptive effect of a middle is created with the addition of *ko*ɭ.[9]

[9]In Part IV, I will discuss two types of constructions, namely *get*-passives in English and *kriegen*-passives in German (see McIntyre 2005, for a discussion of the "middle" properties of English *get*), which encode a very similar semantics and propose that the analysis of *ko*ɭ being developed here be extended to them.

Table 7.1: Verbs that require *koɭ*-suffixation for a given meaning

Bare Stem	**Meaning**	**Stem + *koɭ***	**Meaning**
vej	PUT	*vejttŭ-kkoɭ*	KEEP
puɖi	CATCH	*puɖittŭ-kkoɭ*	HOLD
vaaŋgŭ	BUY	*vaaŋgi-kkoɭ*	ACCEPT
kaʈʈŭ	TIE	*kaʈʈi-kkoɭ*	HUG
paar	SEE	*paarttŭ-kkoɭ*	BE-CAREFUL.IMP

7.4.5.1 Verbs which require *koɭ*

There are many predicates on whose stem the suffixation of *koɭ* appears to be obligatory, at least for a given meaning of that verb. These verbs can be descriptively further classified into two sub-types. For the first sub-class of verbs, *koɭ* appears to be a part of the verbal stem itself: that is, there is no freely occurring *koɭ*-less form of the predicate. The second type of verb does have a morphological stem that occurs without *koɭ*, but the meanings of the *koɭ*-less stem are very different from those formed with the *koɭ*-variant. Non-compositional (or idiomatized) uses of *koɭ* can be observed only with a few predicates: *ottŭkkoɭ* (ADMIT, ACCEPT) is one of them.

(224) Raman tapp-æ ottŭkko-ɳɖ-aan/*ott-aan.
Raman[NOM] mistake-ACC admit-PST-3MSG
"Raman admitted the mistake."

(225) Krishnan paris-æ ottŭkko-ɳɖ-aan/*ott-aan.
Krishnan[NOM] prize-ACC accept--PST-3MSG
"Krishnan accepted the prize."

purindŭkkoɭ is another verb that appears to obligatorily require the presence of *koɭ*, as shown below:

(226) Seetha kaɳakk-æ purindŭkko-ɳɖ-aaɭ/*puri-nd-aaɭ.
Seetha[NOM] math-ACC understand-PST-3FSG
"Seetha understood math."

The second sub-class of predicates, namely the ones for which a *koɭ*-less stem may be used, albeit with a very different and not obviously compositional meaning, has a somewhat larger membership. Table (7.1) presents a comprehensive list of the verbs in this group. All the verbs in (7.1) involve, as part of their inherent meaning, a target result state – i.e.

they are inherently telic. Furthermore, the meaning of each of verb necessarily involves the notion that its highest argument (AGENT or THEME) is the mental or physical locus of the outcome of the eventive predication. In our discussion of psych-verbs and inherently directed motions predicates in Sections 7.4.3.1 and 7.4.3.2, respectively, we saw that the verbs were able to encode this meaning as part of their own lexical-conceptual semantics. The verbs discussed here employ the alternative strategy of doing so via the obligatory suffixation of *koḷ*.

7.4.5.2 Verbs which prefer *koḷ* I: Grooming verbs

Grooming verbs often tend to occur with *koḷ*:

(227) *vaarŭ*, COMB:

a. Krishnan talai-jæ vaar-in-aan.
Krishnan[NOM] hair-ACC comb-PST-3MSG
"Krishnan combed his hair" (Preference: disjoint reading)

b. Krishnan talai-jæ vaar-i-ko-ɳḍ-aan.
Krishnan[NOM] hair-ACC comb-ASP-*koḷ*-PST-3MSG
"Krishnan combed his (own) hair" or
"Krishnan$_i$ combed (someone else's) hair for his$_i$ own benefit"

(228) *maattŭ*, CHANGE:

a. Leela poḍavai-jæ maatt-in-aaḷ.
Leela[NOM] sari-ACC change-PST-3FSG
"Leela changed her sari." (Preference: disjoint reading)

b. Leela poḍavai-jæ maatt-i-ko-ɳḍ-aaḷ.
Leela[NOM] sari-ACC change-ASP-*koḷ*-PST-3FSG
"Leela changed the sari for herself.", or
"Leela changed the sari on herself" i.e. "Leela changed her (own) sari"

The events described by these verbs are all compatible with a result state. Furthermore, these verbs also involve, as part of their inherent meaning, the idea that the initiator or theme argument of the event is also physically or mentally affected by its outcome in some way. This frequently leads to the effect of coreference in the meanings of the resultant sentences – an "inherently reflexive" interpretation that has also been observed in grooming structures in other languages, like Dutch (Reinhart and Reuland 1993). However, it is important to bear in mind that this coreference effect is just that – an effect. In both (227b) and (228b), the coreferent reading is not the only one available; both sentences are also

compatible with a disjoint reading that is nevertheless beneficial to (or affects in some other way) the AGENT/INITIATOR of the combing or dress-changing event. In both cases, there is a requirement that the highest argument (typically an AGENT or THEME) of an event represent a mental or physical locus toward the result state of the event – a requirement that is fulfilled by the suffixation of *koḷ*.

7.4.5.3 Verbs which prefer *koḷ* II: postural verbs

This class of predicate involves verbs like *okkaarŭ*, SIT, *nillŭ* STAND, and *paḍŭ* (LIE). In their basic use, where these verbs simply describe the posture of their (sole) argument, they are atelic. But they naturally lend themselves to a telic interpretation of the argument's *assuming* the relevant posture, with the potential addition of further specifying material: e.g. in English, each of the postural verbs has a particle that yields the corresponding telic counterpart (as in *sit* vs. *sit down*). Under their telic reading such predicates in Tamil readily attach to *koḷ* because the sole argument functions as a spatial locus for the result state of the event:

(229) *nillŭ* (STAND):

a. Lata ni-nn-aaḷ.
Lata[NOM] stand-PST-3FSG
"Lata stood for hours."

b. Lata ni-nnŭ-ko-ṇḍ-aaḷ.
Lata[NOM] stand-ASP-*koḷ*-PST-3FSG
"Lata stood up." (PHYSICAL LOCUS READING)
"Lata stood (for her own benefit)." (MENTAL LOCUS READING)

(230) *okkaarŭ* (SIT):

a. Lata okkaar-nd-aaḷ.
Lata[NOM] sat-PST-3FSG
"Lata sat."

b. Lata okkaar-ndŭ-ko-ṇḍ-aaḷ.
Lata[NOM] sit-ASP-*koḷ*-PST-3FSG
"Lata sat down." (PHYSICAL LOCUS READING)
"Lata sat (for her own benefit)." (MENTAL LOCUS READING)

Of course, the argument may also be construed as having participated in the event agentively, in which case it may represent the mental locus of the event's result state. This requires a more specialized discourse context, however, since the verbs themselves represent an action in the spatial domain.

7.4.5.4 Verbs which prefer *koɭ* III: Self-benefactive verbs

The verbs in this list are not syntactically ditransitive – but built into their conceptual interpretation is the idea that the verbal event could be beneficial to the AGENT. Such verbs stand in direct contrast to the types of alter-benefactive predicates discussed in (218) and (219) which were seen to be degraded with *koɭ*. Self-benefactive verbs, on the other hand, are eminently compatible with *koɭ*, because inherent in their meaning is the idea that the AGENT "receives", in a physical or mental sense, the outcome of the event. Aspectually, as well, such verbs are eminently compatible with *koɭ*, being inherently telic as accomplishments and achievements:

(231) *paɖi* LEARN:

a. Gayatri pariccæ-kkŭ nallaa paɖi-cc-aaɭ.
Gayatri[NOM] exam-ACC well study-PST-3FSG
"Gayatri studied well for the exam."

b. Gayatri pariccæ-kkŭ nallaa paɖi-ccŭ-ko-ɳɖ-aaɭ.
Gayatri[NOM] exam-ACC well study-ASP-*koɭ*-PST-3FSG
"Gayatri studied well for the exam (for herself)"

(232) *keeɭ* HEAR:

a. Chandra višiyatt-æ keeʈ-ʈ-aaɭ.
Chandra[NOM] news-ACC hear-PST-3FSG.
"Chandra heard the news."

b. Chandra višiyatt-æ keeʈ-ʈŭ-ko-ɳɖ-aaɭ.
Chandra[NOM] news-ACC hear-ASP-*koɭ*-PST-3FSG
"Chandra listened to the news.", or
"Chandra heard the news (for her own benefit)"

The *koɭ*-variant given in (231b) conveys the idea that the AGENT Gayatri studied for the exam and has the additional reading that this studying event was either conducted for self-benefit or that she actually learned something (acquired knowledge) from her studying event. In either case, the resultant interpretation is that the Gayatri represents a mental viewpoint toward the target state of the event. Similarly, the addition of *koɭ* to the verb *keeɭ* in (232b) changes the meaning of this verb from HEAR to an event that was initiated more deliberately, with a projected end in mind – like LISTEN.

7.4.6 Summary of results

The main conclusions we have arrived at in this section may be summarized as follows:

Verbs that "like" *koɭ***:** i. Telic predicates that involve a change in physical location or person of a THEME such that the result state of the event is evaluated from its physical viewpoint. E.g. grooming verbs, postural verbs, change-of-state unaccusatives, "reciprocal" verbs like HUG, verbs that take *ta(a)n* as an argument.

ii. Telic predicates that may be construed as having been conducted by an AGENT with a result state in mind or may undergo a mental change as a result of their actions; E.g. self-benefactive verbs, verbs like UNDERSTAND, "reciprocal" verbs like MARRY, verbs that take *ta(a)n* as an argument.

Verbs that are truly optional with *koɭ***:** Verbs that are not telic but are compatible with a telos. Verbs which are compatible, given the right discourse context, with a reading where their highest argument is the mental or spatial locus with respect to the outcome of the event. E.g. most predicate classes: PUT-, PUSH/PULL, CONCEALMENT-, POKE-, THROWING-, CONTACT-, HOLD/KEEP, DESTROY-, and the transitive variants of change-of-state verbs.

Verbs that are degraded with *koɭ***:** i. Predicates that are incompatible with an applied result state. E.g. inherently stative verbs like *irŭ* (BE), and involuntary emissives, like *tummŭ* (SNEEZE).

ii. Predicates that are incompatible with the holding of a mental or physical viewpoint toward their result state by one of their arguments, like involuntary directed motion verbs, and alter-benefactives. E.g. *kiiɻæ viɻŭ* (FALL DOWN) and *kuɖŭ* (GIVE).

Our conclusions with respect to *koɭ* and the type of structure it occurs in may be summarized as follows:

(233) **Properties of a** *koɭ* structure:

i. *koɭ* is a semi-functional restructuring verb: it spells out a functional head, selects a resultant AspP, and assigns a θ-role to the DP in its specifier.

ii. The DP that gets assigned a θ-role by *koɭ* is the highest argument of the event predication under *koɭ*; when it is merged in the specifier of *koɭ*, it denotes an entity that comes to be

the mental or physical locus of the result state of the main event predication in the scope of *ko*⌊.

iii. The AspP that is selected by *ko*⌊ denotes the result state of an event predication.

There is one other type of structure with which *ko*⌊ occurs very frequently which we have deliberately postponed discussion of until all the relevant facts about *ko*⌊ were amassed. This is the local reflexive structure involving a locally anteceded *ta(a)n* in directed object position. Since understanding the nature of local reflexivity in Tamil is a major concern of our investigations – one which, furthermore, will connect this series of chapters with the others in the dissertation – we will devote the following chapter to investigating its properties relative to *ko*⌊.

Chapter 8

Back to binding: local reflexivity and *koɭ*

We saw that, in the default case, the presence of *koɭ* is required for the local binding of *ta(a)n* to go through, yielding minimal pairs like the following:

(234) * Maya$_i$ tann-æ$_{\{i,*j\}}$ aɖi-tt-aaɭ.
Maya[NOM] ANAPH-ACC hit-PST-3FSG
"Maya$_i$ hit herself$_{\{i,*j\}}$." (Intended)

(235) Maya$_i$ tann-æ$_{\{i,*j\}}$ aɖi-ttŭ-kko-ɳɖ-aaɭ.
Maya[NOM] ANAPH-ACC hit-ASP-*koɭ*-PST-3FSG
"Maya$_i$ hit herself$_{\{i,*j\}}$."

In the minimal pair above, the addition of *koɭ* to the predicate essentially makes it possible for the matrix subject, *Maya* to bind *ta(a)n*. Why should the clausemate subject not be able to serve as an antecedent for *ta(a)n* in the absence of *koɭ*? I.e. what is the extra property that *koɭ* brings to the table that somehow lifts this restriction and allows it to antecede?

Our discussion of the well-formedness conditions on potential anaphoric antecedence in Part I of this dissertation hints at the answer. There, we saw that long-distance binding of *ta(a)n* is possible across complement CPs, as well as into spatio-temporal and causal PPs, DPs, and CPs and in so-called "logophoric" structures – as illustrated below:

(236) Maya$_i$ [$_{CP}$ Raman$_j$ tann-æ$_{\{i,*j\}}$ terŭvŭ-læ paar-tt-aan-nnŭ] nene-
Maya Raman ANAPH-ACC street-LOC see-PST-3MSG think-
tt-aaɭ.
PST-3MSG
"Maya$_i$ thought [$_{CP}$ that Raman$_j$ saw her$_{\{i,*j\}}$ on the street.]"

(237) Ravi$_i$ Krishnan-kiṭṭæ polamb-in-aan. Taan$_i$
Ravi[NOM] Krishnan-ALL complain-PST-3MSG. ANAPH[NOM]
maṭṭum een ippaḍi ellaam kashṭappaḍa-ṇum?
alone why like.this all suffer-must?
"Ravi$_i$ complained to Krishnan$_j$. Why should he$_{\{i,*j\}}$ alone suffer like this?"

(238) Vivek$_i$ [$_{DP}$ tann-ooḍæ$_{\{i,*j\}}$ mugatt-æ] aḍi-tt-aan.
Vivek[NOM] ANAPH-GEN face-ACC hit-PST-3MSG
"Vivek$_i$ hit his$_{\{i,*j\}}$ face."

(239) Koḻændæ$_i$ tan-akku$_{\{i,*j\}}$ meele orŭ plane-æ paar-tt-adŭ.
child[NOM] ANAPH-DAT above a plane-ACC see-PST-3NSG
"[The child]$_i$ saw a plane above itself$_{\{i,*j\}}$."

(240) Seetha$_i$ [$_{CP}$ Anand$_j$ tan-akkŭ$_{\{i,*j\}}$ paṇam tandæ-
Seetha$_i$ Anand ANAPH-DAT money[NOM] give-PST-GER-
(a)ppŭram] veelæ-jæ sej-v-aaḷ.
AFTER work-ACC do-FUT-3FSG
"Seetha$_i$ will do the work [$_{CP}$ after Anand$_j$ pays her$_{\{i,*j\}}$.]"

In all these cases, the antecedent of *ta(a)n* was observed to be chosen from among a set of individuals who all hold a mental (e.g. (236), (237), (240)), spatial ((238) and (239)) or temporal (240) perspective toward the minimal predicational structure containing *ta(a)n*, as described more precisely below:

(241) **Condition for potential *ta(a)n*-antecedence (Final version):**

i. A potential antecedent of *ta(a)n* is a nominal which has a mental, temporal or spatial perspective with respect to the minimal CP, PP, or DP in which the anaphor is a participant (i.e. thematic argument).
ii. This information is represented as part of the perspectival center in the minimal CP, PP or DP containing the anaphor.

There is an obvious connection between the conditions on potential antecedence and the kind of meaning that *koḷ* has been observed to contribute. We have *independently* argued here that the syntactico-semantic contribution of *koḷ* is to modify an event such that the highest argument of that event comes to be the mental or physical locus for the result state of this event. But this can be said to also be true for all potential antecedents of *ta(a)n* – such as those in the long-distance binding

sentences in (236)-(240), for instance.[1] In light of this discussion, it is hardly surprising that *ta(a)n* may be anteceded by a co-argument in the concomitant presence of *ko*ɭ. Consider the sentence in (235) again:

(242) Maya$_i$ tann-æ$_{\{i,*j\}}$ aɖi-ttŭ-kko-ŋɖ-aaɭ.
Maya[NOM] ANAPH-ACC hit-ASP-*ko*ɭ-PST-3FSG
"Maya$_i$ hit herself$_{\{i,*j\}}$."

As we've seen, the semantics of *ko*ɭ is associated with the highest argument of the event which, in this case, is the referentially independent subject *Maya*. As such, with the addition of *ko*ɭ, *Maya* will come to represent a mental or physical locus for the result state of the hitting event. In other words, *Maya* will count as a potential perspective-holder toward the *ta(a)n*-eventuality. This constellation of properties is enough to ensure that *Maya* qualifies as a potential antecedent of *ta(a)n*, as per (241) above.

8.1 Clarifying the notion of "perspective" in light of *ko*ɭ

What is perhaps more surprising is why this shouldn't be possible in the absence of *ko*ɭ. Consider again the ungrammatical sentence involving local binding of *ta(a)n* in the absence of *ko*ɭ, below:

(243) * Maya$_i$ tann-æ$_{\{i,*j\}}$ aɖi-tt-aaɭ.
Maya[NOM] ANAPH-ACC hit-PST-3FSG
"Maya$_i$ hit herself$_{\{i,*j\}}$." (Intended)

The fact that *Maya* is not able to antecede *ta(a)n*, given (241), must be because this DP is, for some reason, unable to qualify as a perspective-holder towards the *ta(a)n*-eventuality (which is trivially also its own eventuality). But why should this be the case?

It couldn't have anything to do with the AGENTIVE θ-role on *Maya* or other properties of this DP such as its ϕ-feature content, animacy,

[1] The predicates that introduce these DPs may imbue them with additional properties beyond this: for example, the attitude-verb *nene* (THINK) contributes the additional information that the viewpoint of the attitude-holder towards the embedded predicate is one of thought rather than, say, one of speech or perception. But the basic requirement that the embedded predicate is viewed from the mental locus of the attitude-holder is met. The fact that *ko*ɭ, unlike *nene* (THINK), introduces *just* this bare minimum suggests that it is less "meaty" than such a verb – a property that fits in nicely with our treatment of it as a semi-functional restructuring predicate, rather than a fully-fledged lexical one.

or grammatical function in the clause. After all, if the anaphor is in a different clause, this same DP *Maya* may antecede it, as illustrated in (244) below:

(244) [$_{CP}$ Krishnan$_k$ tann-æ$_{\{i,*j,*k\}}$ aḍikkum-poḻuḏu] Maya$_i$
Krishnan[NOM] ANAPH-ACC hitting-while Maya[NOM]
Raman-æ$_j$ aḍi-tt-aaḷ.
Raman-ACC hit-PST-3FSG
"[$_{CP}$ While Krishnan$_k$ hit her$_{\{i,*j,*k\}}$], Maya$_i$ hit Raman$_j$."

This might suggest that the problem is simply that a DP may not serve as an antecedent to an anaphor in its own clause – a descriptive restriction that I have labelled the Ban on Clausemate Subject Antecedence. But this cannot be quite correct, either. After all, as we already saw in Part I, a co-argument of *ta(a)n* may antecede it in psych-predicate structures, as in (245). Conversely, a non-co-argument may be disbarred from anteceding it, as in (246):

(245) Maya tann-æ$_{\{i,*j\}}$ veru-tt-aaḷ.
Maya[NOM] ANAPH-ACC hate-PST-3MSG
"Maya$_i$ hated herself$_{\{i,*j\}}$."

(246) * Maya$_i$ Raman-æ$_j$ aḍi-tt-aaḷ. Appŭram
Maya[NOM] Raman-ACC hit-PST-3FSG. And then
Krishnan tann-æ$_i$ aḍi-tt-aan.
Krishnan[NOM] ANAPH-ACC hit-PST-3MSG
"Maya$_i$ hit Raman$_j$. And then Krishnan$_k$ hit her$_i$." (Intended)

The condition(s) that disallow *Maya* as a potential antecedent in (244) must thus be something else. To see what it is, we need to be more precise about the nature of "perspective" and its linguistic representation.

Perspective at its most intuitive represents an asymmetric relationship between two objects – one is the object that holds the perspective (call it the "anchor") and the other is the object that is being viewed (call it the "object of scrutiny"). Perspective in the linguistic sense retains this basic intuition. It is an asymmetric relationship (which can be modelled as a two-place predicate) between an individual (the anchor) and a predicational structure (the object of scrutiny). When we assert that an individual holds a perspective towards a predicate, we are asserting that the space or time or world or mind components of the predicate are evaluated against the corresponding coordinates of the individual. But this means that the principles governing whether a particular individual (like *Maya* in (244) and (243)) can count as the perspective-holder towards

a particular predicate don't have to do with properties of the individual alone. They *also* have to do with the properties of the predicational structure and the nature of the relationship between the individual and the predicate. In other words, the conditions on potential antecedence delineated in (280) are incomplete because they talk about properties pertaining to the individual (the anchor) in isolation.

Returning to the example in (244), we can now see that it is not just the properties of the agentive subject *Maya* – 3rd-person, animate, and so on – which qualify it as a perspective-holder toward the *ta(a)n*-eventuality. The nature of the relationship between the two clauses – specifically the fact that the embedded one is temporally linked to the matrix via the temporal subordinator *poɻŭdŭ* (WHILE) – also plays a key role in facilitating this. In contrast, in (246), *Maya* is not able to antecede *ta(a)n* because the conjunction *appŭram* (AND THEN) doesn't serve as the right kind of *linker* or *relator* to enable the establishment of a perspectival relationship between *Maya* and the *ta(a)n*-predicate. Extending this logic to the ungrammatical case in (243), we must thus conclude that it is the absence of such a relator that prevents *Maya* from locally anteceding *ta(a)n*.

8.2 A structural restriction on linguistic perspective

In Part I, we proposed that certain CPs, DPs, and PPs have a *perspectival center* which we defined as follows:

(247) **Formal representation of a Perspectival Center:**

i. The *perspectival center* contains the coordinates pertaining to the time, location, world, and/or mental information of a salient perspective holder.

ii. Certain predicational structures (at least some PPs, DPs, CPs) contain a perspectival center by virtue of what they inherently "mean". In a proper subset of these cases, the representation of the perspectival center in a phrase can be traced back to the selectional properties of its immediately superordinate predicate.

iii. A predicational structure has at most one perspectival center.

iv. The predication containing a successfully bound anaphor must contain a perspectival center.

This shows that a perspectival center is nothing but a linguistic *relator* or *linker*, in the sense discussed above. Assuming this to be true, it is tempting to suggest that *Maya* is not able to antecede *ta(a)n* in (243) simply because this CP lacks a perspectival center altogether. But this cannot be quite right – after all, if this clause were to be embedded under another predicate, e.g. an attitude verb, long-distance binding of *ta(a)n* would be possible into it:

(248) Raman$_i$ [$_{CP}$ Maya$_j$ tann-æ$_{\{i,*j\}}$ aḍḭ-tt-aaḷ-ŭnnŭ]
Raman[NOM] Maya[NOM] ANAPH-ACC hit-PST-3FSG-COMP
nene-tt-aan.
think-PST-3MSG
"Raman$_i$ thought [$_{CP}$ that Maya$_j$ hit him$_{\{i,*j\}}$.]"

The ungrammaticality of (243) seems, rather, to be because the perspectival-center in the C-layer asymmetrically c-commands the DP *Maya* which is the intended perspective-holder of that CP. This is a structural configuration that seems to be disallowed. In other words, the following restriction seems to hold:

(249) **Restriction on the relationship between the predicational structure (object of scrutiny) and individual (anchor):**

Intuition:

In order for a perspectival relationship to be established between an anchor and an object of scrutiny, the anchor may not be properly contained inside the object of scrutiny.

Linguistic instantiation:

The perspectival center (in [Spec, PerspP]) immediately containing the predicational structure may not asymmetrically c-command the DP that is the linguistic representation of that individual.

The condition in (249) is admittedly a stipulation but it is an inherently reasonable one to make. We have been treating linguistic perspective as a two-place predicate that encodes an asymmetric relationship between a perspective-holder and a predicational structure. Assuming, as is standard in compositional approaches to semantics (Heim and Kratzer 1998), that semantic predicate relationships are structurally realized, this in turn means that one of the arguments of the relation cannot be wholly contained inside the other. They must be distinct objects which can occupy two distinct structural positions, just as (249) states.[2]

[2]Incidentally, the condition in (249) is reminiscent of the restriction, proposed in Chomsky (1981) to correctly rule out structural conditions where one coreferent DP

With these intuitions in place, we can now define linguistic perspective as follows:

(250) **Definition of linguistic perspective:**

i. Linguistic perspective denotes a two-place predicate that relates an individual that exists in a specific space, time, world, and, if sentient, has a mind, with a predicational structure.

ii. When we assert that an individual holds a perspective towards a predication, we are asserting that the predication is evaluated against the space, time, world, or mind of this individual.

iii. The individual may not hold a perspective toward a predication that it is wholly embedded within.

8.3 Why *ko⌊* is typically necessary for co-argument binding

In long-distance binding structures involving the legitimate binding of *ta(a)n*, this condition is trivially satisfied, since the DP that ends up being construed as the antecedent is already outside the eventuality containing *ta(a)n* – in a higher phase (in the typical case), in a lower phase (in the case of backward binding) or in the larger discourse (in cases of logophoric reference) – to begin with. But in structures like (243), the DP that is the intended antecedent *is* embedded within the predicational structure that we would have it hold a perspective towards. Thus, (249) is violated. Returning now to the grammatical case of *ta(a)n*-binding instantiated in (242), the addition of *ko⌊* must somehow make it possible to avoid a violation of the c-command condition in (249) – i.e. by placing *Maya* in a position where it is no longer asymmetrically c-commanded by the immediately containing perspectival center in [Spec, PerspP].

Recall that *ko⌊* makes the minimum contribution necessary for a DP to function as a perspective-holder – and that it is, indeed, much like an attitude-verb (e.g. *nene* (THINK)) in this sense. The natural way to implement this would thus be to propose that *ko⌊* introduces a structure

is contained inside another. This restriction, called the "i-within-i condition" which is defined as follows:

i. *$[_A \ldots \text{B} \ldots]$ where A and B bear the same index.

The two conditions are not identical by any means, but there is a basic intuitive similarity between the two, I believe.

that contains a PerspP. We have seen that *koḷ* selects an AspP – what we are proposing here then is that an AspP contains a projection for PerspP. Independent empirical support for the idea that AspP contains a PerspP comes from sentences like (251) below:

(251) Seetha$_i$ rombæ dukkappa-ṭṭ-aaḷ. Raman$_j$ [$_{CP}$ taan$_j$
Seetha[NOM] very.much sad-PST-3FSG. Raman ANAPH[NOM]
tannæ$_i$ kaadalikka-læ-nnŭ] solli-jirukka veeṇḍaam.
ANAPH-ACC love-NEG-COMP say-PTCP need.not
"Seetha$_i$ was very sad. Raman$_j$ needn't have said [$_{CP}$ that he$_j$ didn't love her$_i$]."

There are two instances of *ta(a)n* in the second sentence in (251): one is the embedded subject in [Spec, TP] and the other is the embedded direct object in the complement of VP. Crucially, each of these instances of *ta(a)n* is anteceded by a different DP – subject *ta(a)n* is long-distance bound by its matrix subject *Raman*, but object *ta(a)n* is "logophorically" bound by *Seetha* in the previous sentence. Under the model of binding being developed here, where *ta(a)n* is actually bound by an operator in [Spec, PerspP], the presence of two instances of *ta(a)n* with two referentially distinct binders automatically entails the presence of two distinct PerspPs. The accustomed one in the embedded CP layer will bind *ta(a)n* in [Spec, TP] of the embedded clause. This means that the additional one binding the object must be in a lower position, presumably in the *v*P layer.

Claiming that AspPs contain a PerspP projection is consonant with what we have observed about the distribution of PerspP. The empirical evidence in Part I has independently shown that a PerspP is present on CPs, DPs, and PPs. These are all maximal projections with a special syntactic status – they have all been argued to be phases. It is entirely reasonable to think that the AspP that *koḷ* takes as complement is also a phase. First of all, it corresponds essentially with the *v**P level standardly taken to be a phase. Second of all, we are claiming that *koḷ* selects an independent functional sequence as its complement, a property that plausibly corresponds with phasehood. This would allow us to say that having a PerspP is a property of phases.

In light of the discussion thus far, we may now update the definition in (247) as follows:

(252) **Formal representation of a Perspectival Center** (Final version):

i. The *perspectival center* contains the coordinates pertaining to the time, location, world, and/or mental information of

a salient perspective holder. These are hosted in a silent pronominal operator in the specifier of a Perspectival Phrase (PerspP).

ii. Certain phases (at least some PPs, DPs, AspPs, CPs) contain a PerspP by virtue of what they inherently "mean".

iii. A phase has at most one PerspP.

iv. The phase containing a successfully bound anaphor must contain a PerspP. The operator in [Spec, PerspP] Agrees with the anaphor in its minimal phase and variable-binds it at LF.

v. A potential antecedent may not be asymmetrically c-commanded by the PerspP it holds a perspective towards. The relationship between the potential antecedent and the operator in [Spec, PerspP] is one of non-obligatory control.

Consider again the sentence in (242). Assuming that the agentive subject *Maya* is externally merged in [Spec, VoiceP], a PerspP at the left edge of the c-commanding *Asp*P phase would still asymmetrically c-command it. Thus, as discussed above, *ko*ɭ must be making an additional contribution, namely allowing *Maya* to escape this PerspP, either by raising out of its thematic base position or via some sort of control configuration. In other words, I am proposing that *ko*ɭ selects a structure with a PerspP and allows a sentient DP, like *Maya*, in its scope to be represented in a position where it asymmetrically c-commands the minimal PerspP containing *ta(a)n*, rather than the other way around. It further yields the interpretation that *Maya* comes to represent the mental or physical locus of the result state of the *ta(a)n*-predicate. As a result of both properties simultaneously coming to hold, *Maya* is able to serve as a potential antecedent of *ta(a)n* in (242).

Finally, note that *Maya* still only serves as a *potential* antecedent for *ta(a)n*. The introduction of *ko*ɭ has made it *possible* for this DP to antecede *ta(a)n*, but it doesn't force this to be the case because the operator in [Spec, PerspP], as we have already seen, is not syntactically dependent on another element in the structure. It is a null deictic pronoun (a little *pro*) which only needs to be assigned to a referent (corresponding to the actual antecedent of *ta(a)n*) at LF. As such, if (242) were to be embedded under an intensional predicate, as in (253), we would have two potential antecedents – *Maya* and the attitude-holder of the c-commanding intensional verb. Indeed, such examples are ambiguous with respect to the identity of the antecedent of *ta(a)n*, as the sentence below shows:

(253) Raman$_i$ [$_{CP}$ Maya$_j$ tann-æ$_{\{i,j\}}$ aḍi-ttŭ-kko-ŋḍ-aaɭ-ŭnnŭ]
Raman[NOM] Maya[NOM] ANAPH-ACC hit-ASP-*ko*ɭ-PST-3FSG-COMP

nene-tt-aan.
think-PST-3MSG
"Raman$_i$ thought [$_{CP}$ that Maya$_j$ hit him$_i$/her$_j$.]"

8.4 But what about psych-predicates?

In Section 7.4.3.1, we saw that psych predicates are degraded, to varying degrees depending on whether they are eventive vs. stative and whether they involve dative vs. nominative EXPERIENCER subjects, with *ko*̲. However, *ta(a)n* may still be locally anteceded in such structures, as we've already seen. This is strictly independent of whether the psych-predicate is stative or eventive and whether the EXPERIENCER subject is marked dative or nominative – those factors matter for *ko*̲-suffixation but not for the local antecedence of *ta(a)n*. Below is a representative example of locally-anteceded *ta(a)n* in a psych-predicate structure:

(254) Raman-ŭkkŭ$_i$ tann-æ-jee$_{\{i,*j\}}$ pid̲ikk-æ-læ.
Raman-DAT ANAPH-ACC-EMPH like-INF-NEG
"Raman$_i$ didn't like (even) himself$_{\{i,*j\}}$."

(255) * Raman-ŭkkŭ$_i$ tann-æ-jee$_{\{i,*j\}}$ pid̲ittŭ-ko̲-æ-læ.
Raman-DAT ANAPH-ACC-EMPH like-*ko*̲-INF-NEG
"Raman$_i$ didn't like (even) himself$_{\{i,*j\}}$." (Intended)

In (254), the dative EXPERIENCER subject *Raman* antecedes *ta(a)n* despite the (obligatory) absence of *ko*̲. Given (249) and the surrounding discussion above, this must mean that *Raman* is not asymmetrically c-commanded by the minimal PerspP containing *ta(a)n* in this sentence. How do we model this?

First of all, the minimal PerspP for *ta(a)n* in psych-predicate structures cannot be the resultant AspP (that occurs in the complement of *ko*̲ in *ko*̲-structures) but the PerspP in the clausal left-periphery. That the resultant AspP (AspP$_{res}$) is not involved in local anaphoric antecedence in such structures is easily shown – after all, it is entirely absent in a sentence like (254). We saw that *ko*̲-marked predicates have two sets of aspectual markers – a "frozen" aspectual form below *ko*̲ and the aspectual form that combines with tense which occurs above it. In contrast, psych-predicate structures evidence only one set – the second kind that yields real tense effects:

(256) Raman$_i$ tann-æ$_{\{i,*j\}}$ verŭ-kkir-aan.
Raman[NOM] ANAPH-ACC hate-PRS-3MSG
"Raman$_i$ hates himself$_{\{i,*j\}}$."

(257) Raman tann-æ verŭ-pp-aan.
Raman[NOM] anaph-ACC hate-FUT-3MSG
"Raman will hate Seetha."

We have argued that the AspP-res selected by *ko*ļ marks a phase-boundary and contains a PerspP which then serves as the minimally c-commanding PerspP for *ta(a)n* in *ko*ļ-structures involving this anaphor. However, if psych-predicates lack this phasal projection altogether, then the minimal phase containing *ta(a)n* must be the CP. Assuming, further, as we have done, that there is a unique PerspP per phase, this must mean that the minimal PerspP for *ta(a)n* in psych-predicate structures like (254), (256), and (257) must be the PerspP in the C layer.

Combining this with our previous argument that the local antecedent of *ta(a)n* must be outside the PerspP that minimally c-commands *ta(a)n*, we arrive at the conclusion that EXPERIENCER subjects that locally antecede *ta(a)n* must be merged above the PerspP in the C layer. I will thus propose that an EXPERIENCER θ-role is assigned, not in [Spec, VoiceP], but above PerspP in the C layer. This entails that the argument-structure of a psych-predicate is structurally larger than those of other types of verbs, a point I will discuss in more detail in Section 9.3.2.[3] At the same time, I am claiming that the functional structure of the minimal CP containing *ta(a)n* in a psych-predicate structure is simpler than that of a CP containing *ko*ļ: the former involves a single extended functional sequence, whereas the latter involves two functional sequences (though it does still count as mono-clausal in the Wurmbrand-sense by virtue of having a single T head and a single subject).

Both the incompatibility of psych-verbs with *ko*ļ and the possibility of local antecedence may be captured under this proposal – a point that I will formally model in the next chapter.

[3]Observe that such a proposal could also be used to account for the possibility of backward binding in local cases. Adger and Ramchand (2006) argue that psych predication in Scottish Gaelic involves experiencers that are base-generated higher than other stative subjects

Chapter 9

Formally modelling *ko*⌊ and local binding

9.1 Introduction

This chapter develops a formal syntax and semantics for *ko*⌊-structures, including those involving the local binding of *ta(a)n*. I will assume that thematic relations are syntactically represented, along the lines of Hale and Keyser (1993), Ramchand (2008) and others. I will also assume, in line with Kratzer (1996) that an external argument is not the argument of a predicate but is introduced in a Neo-Davidsonian manner, by a separate functional head like Voice. Finally, I will assume, as I have been doing all along, that spell-out rules apply to the syntactic output and that morphophonological exponents are thus inserted "late" in the grammatical derivation.

Let us recapitulate what we already know with respect to *ko*⌊ and the type of structure it occurs in.

(258) **Properties of a** *ko*⌊ structure (updated from (233):

i. *ko*⌊ is a semi-functional restructuring verb: it spells out a functional head, selects a resultant AspP, and assigns a θ-role to the DP in its specifier.

ii. The AspP that is selected by *ko*⌊ denotes the result state of an event predication. It also constitutes a phase, thus contains a PerspP, which is a projection that uniquely marks a phase.

iii. The DP that gets assigned a θ-role by *ko*⌊ is the highest argument of the event predication under *ko*⌊; when it is merged in the specifier of *ko*⌊, it denotes an entity that comes to be the mental or physical locus of the result state of the main event predication in the scope of *ko*⌊.

In the following sections, we will present a formal syntactic and semantic model for *koɭ*-structures including those involving the local antecedence of *ta(a)n*. I will first propose a denotation for *koɭ* in line with the "coming to hold" semantics that it introduces to the DP in its specifier. We will then see how to model the idea that this DP is also the highest thematic participant of the main event in its scope, to which end I will consider two analytic options. One possibility is that *koɭ* is a control verb – in other words, that the argument in its specifier obligatorily controls a null operator in the event predication under *koɭ*, along the lines of a recent analysis for adjectival passives proposed by Bruening (To Appear). The alternative would be to claim that *koɭ* is a thematic raising verb (Ramchand 2008): i.e. that the highest argument of the event predication under *koɭ* raises up to the specifier of the *koɭ*-phrase, as for instance suggested for certain types of complementizers in Irish with similar raising effects by McCloskey (2002). Each option comes with its own pros and cons – I will discuss these briefly but remain agnostic for now about which one is correct for *koɭ*. We will conclude with step-by-step derivations of the various types of *koɭ*-structures discussed in the previous chapters.

9.1.1 Developing a denotation for *koɭ*

In Section 7.2.2 of Chapter 7, we proposed that the Asp_{res} head that occurs above the verb has the following denotation:

(259) $[\![Asp_{res}]\!] = \lambda R_{<s,t>} \lambda s_s \exists e.R(e) \wedge Result(e,s)$

(259) asserts that the Asp_{res} head takes an eventive proposition (of type $< s,t >$), binds off the event and introduces a state which is the derived result of that event. We have said that the head that *koɭ* spells out is built on top of Asp_{res} – a fact that we have modelled by claiming that *koɭ* selects the resultant AspP of which Asp_{res} is the head.

However, the resultant *koɭ*-structure that is thus developed itself seems to be eventive, not stative. This is shown by the "immer noch" test from Kratzer (2000), below:

(260) The glasses were (still) broken yesterday.
(261) The glasses (*still) broke yesterday.
(262) Kaŋŋaaɖi neettikki (*innum) oɖen-ʠŭ-ko-ŋɖ-adŭ.
Glasses yesterday (*still) break-ASP-*koɭ*-PST-3NSG
"The glasses (*still) got broken yesterday."

Under the relevant stative reading, the adverb *yesterday* in (260) only modifies the time of the target state (of brokenness) of the glasses.[1] Under this reading, the event of the glasses breaking might have happened the day before yesterday – all that is being asserted is that the glasses were in a broken state yesterday; thus, an adverb like *still* is licit. This shows that the target state is available for further modification at the propositional level while the event has been bound off. Compare this with a normal eventive past-tense sentence like (261) where *yesterday* clearly only modifies the event of glass-breaking – *still* is clearly ungrammatical in this sentence. We can see how *koḷ*-structures in Tamil behave in comparison. (262) has only the reading that the event of glass-breaking happened yesterday; as such, *innum* (STILL) is ungrammatical in this sentence. This shows that the target state has been bound off at the propositional level at which modification occurs.[2,3]

I will take this type of data to show that, although the event variable pertaining to the main event predication is bound off below *koḷ*, *koḷ* itself introduces a new sub-event variable. In addition, I propose that *koḷ* binds off the result state variable introduced by Asp_{res}. This, combined with the conclusions about *koḷ* given in (258), yields the following denotation for this morpheme:

(263) $[\![ko\text{ḷ}]\!] = \lambda Q_{<s_s,t>}\lambda x\lambda e'\exists s.Q(s)\wedge Get(e')\wedge Locus(e',x)\wedge Theme(e',s)$

This denotation states that *koḷ* takes a stative proposition as its argument and binds off the state. It further takes an individual and a(n) (sub-)event and relates them to the result state. Specifically, x KOL [AspP] means "x comes to hold the derived result state denoted by Asp_{res}". This is a nice result because it is very close to the fully lexical meaning of *koḷ* in older stages of the language, where it means HOLD.

9.1.2 The denotation of Persp

Consider our linear ordering for the *koḷ*-structure from (146), now with the Persp head under AspP included:

(264) **Structural position of *koḷ*:**
[VerbRoot - (Voice) - Persp - Asp_{res}] - $\text{F}_{ko\text{ḷ}}$ - Tense - Agr

[1] There is also an eventive passive reading for this sentence (but "*still*" has to first be removed) but Kratzer argues that this has a different underlying structure.

[2] Thanks to Tom McFadden (p.c.) for alerting me to the significance of this point.

[3] Note the similarity to the "get broken" reading in English, which also has only an eventive reading under temporal adverbial modification. I will suggest in Part IV that this similarity is not accidental and argue that the semantics of *koḷ* is in some ways quite similar to that of GET.

We have derived a denotation for *ko*⌊ and a denotation for Asp_{res}. What is the denotation of the Persp head (that introduces a perspectival center operator in its specifier) that occurs sandwiched between the two?

The perspectival head, I propose, is much like Voice in the Kratzer (1996) sense. It introduces an argument in its specifier and relates it to the proposition in its scope. It does not combine with the proposition by function application but by Event Identification, an operation that Kratzer defines as follows:

(265) **Event Identification:**

i. "Event Identification makes it possible to chain together various conditions for the event described by a sentence. It takes a function f and a function g (order irrelevant) as input and yields a function h as output. Input functions f and output functions h are of type $< e, < s, t >>$. Input functions g are of type $< s, t >$" (Kratzer 1996, 122).
Formally: $\lambda x_e \lambda e_s [f(x)(e) \& g(e)]$

ii. I.e. $< e, < s, t >> \quad < s, t > = < e, < s, t >>$

The denotation of Persp is thus very simple – as illustrated below:

(266) $[\![Persp]\!] = \lambda x \lambda s.PerspHolder(e, x)$

The Persp head itself represents a function of type $< e, < s, t >>$. I.e. it takes a proposition (type $< s, t >$) as its complement and introduces an individual (type $< e >$) which is saturated by the argument that is merged in its specifier. The Persp head and the proposition in its scope combine by the Event Identification operation defined in (265) above – the output is thus also a function of type $< e, < s, t >>$. In general, a Persp head may combine with a stative or eventive proposition. In *ko*⌊-structures, I propose that the Persp head combines with an eventive one – this is because, under the linear sequence given in (264), Persp attaches directly below Asp_{res} which, being a stativizer, takes an eventive proposition and yields a stative one.

9.1.3 Implementing the "raising" effect of *ko*⌊

There is one final point with respect to *ko*⌊ that still needs to be ironed out. We have observed that *ko*⌊ relates the highest argument of the main event predication in its scope to the result state of that event. Essentially then, the addition of *ko*⌊ allows this highest argument to occupy two distinct syntactic positions at the same time: the first being its thematic position and the second its derived position.

There are two ways to implement this idea. One way would be to treat a *ko*⌊-structure as being analogous to a control configuration; the other would be to treat it as a type of raising structure. Both strategies are attested in the literature. For instance, Bruening (To Appear) proposes that adjectival passives in English have a control-type structure. In a sentence like "I bought a broken table", for instance, a single referent has to be associated with the main predicate "bought" but also have a thematic relationship with "broken" – in other words, this argument must be simultaneously associated with two predicates, just like the perspective-holder DP in a *ko*⌊-structure. Bruening's strategy is to claim that the argument position in the verbal structure underlying "broken" is filled by an empty operator which is eventually controlled by the relevant referent. Such an implementation is attactive for *ko*⌊ given that aspectual phrases are similar to adjectival phrases in many ways – in these particular cases, the connection is even tighter because we are essentially dealing with participial morphology in both.

The second strategy essentially involves raising. A single DP must move from one thematic position to another, thereby acquiring more than one θ-role. Such an implementation will necessarily violate the θ-criterion. But this isn't necessarily such a terrible thing, as argued in Hornstein (1999), Ramchand (2008), among others. The θ-criterion was not adopted because of clear evidence that a single argument cannot be thematically related to two distinct predicates but because it could be used to block certain unwanted derivations that the GB-theory otherwise wrongly predicted to be possible. Hornstein argues that many hitherto problematic syntactic issues can be more elegantly derived if we relax the restriction on the assignment of θ-roles. Furthermore, as we have seen, the other ways to analyze *ko*⌊, discussed above, would preserve the sanctity of the θ-criterion. But they would do so at the expense of ignoring independent empirical evidence or by presupposing a meaning for *ko*⌊ that is quite unlike what we have observed for it so far.

In the syntax, raising must be implemented by means of a formal mechanism which prevents external Merge of a DP in the specifier of *ko*⌊ and ensures that this position is filled by remerge (i.e. internal Merge) of a lower argument in the local domain. There are a number of technical ways that this could be implemented and at this stage I do not have enough evidence to distinguish among them. One possibility might be to propose that there is a featural diacritic on the head that *ko*⌊ spells out which explicitly stipulates that its specifier be filled by remerge of a DP in its (local) c-command domain. This would be analogous to the OP feature proposed by McCloskey (2002) to distinguish the version of C in

Irish that forces *wh*-movement from the one that accompanies resumptive pronouns.

Either the control or raising implementation is, in theory, compatible with the analysis of *ko*ɭ-structures in Tamil being developed here. Both control and raising verbs are treated as instances of types of restructuring predicate by Wurmbrand (2001), thus either analysis would be compatible with our treatment of *ko*ɭ as a predicate of this kind. Each approach also brings its own pros and cons to the table. A raising analysis is appealing because we could model the fact that *ko*ɭ attributes thematic properties to the highest argument of the event predication in its scope as a function of Relativized Minimality effects on A-movement: i.e. we could claim that the highest argument is chosen simply because it is the minimally closest argument to the specifier of a *ko*ɭ-phrase. This minimality effect must be separately stipulated for a control-type analysis, however. At the same time, assuming as we have done that the resultant AspP selected by *ko*ɭ constitutes a phase, it is unclear how this argument is able to escape out of this phasal domain to remerge in the specifer of the *ko*ɭ-phrase.[4] In this respect, a(n) (LF-)semantic control analysis such as that of Bruening (To Appear) clearly fares better.

I will remain agnostic for now about which of these is the correct strategy. The choice of one over the other will, in any case, be entirely orthogonal to our intuitions about the meaning-contribution of *ko*ɭ, and the nature of its denotation, given above. For concreteness, in the tree-structures illustrating the step-by-step derivation of various *ko*ɭ-sentences below, I model the raising analysis because it allows for significantly simpler structures.

9.1.4 Final definition for *ko*ɭ

We now have all pieces of the puzzle in place. Here is the final definition of *ko*ɭ in Tamil based on our findings thus far:

(267) **Properties of a *ko*ɭ structure** (Final version):

i. *ko*ɭ is a semi-functional restructuring verb: it spells out a functional head, selects a resultant AspP, and assigns a θ-role to the DP in its specifier.

ii. The AspP that is selected by *ko*ɭ denotes the result state of an event predication. It also constitutes a phase, thus contains

[4]Standard raising analyses explicitly assume, for this reason, that raising does not happen across a phase boundary, and that raising infinitives e.g. are TPs and thus not phasal.

a PerspP, which is a projection that uniquely marks a phase.
$[\![Asp_{res}]\!] = \lambda R_{<s,t>}\lambda s_s \exists e.R(e) \wedge Result(e,s)$
$[\![Persp]\!] = \lambda x \lambda s.PerspHolder(e,x)$

iii. *koɬ* allows the highest argument of the main event predication in its scope to be in two distinct syntactic positions at the same time: it is thus either a control or (thematic-)raising predicate.

iv. The DP that occurs in the specifier of *koɬ* denotes an entity that comes to be the mental or physical locus for the result state of the main event predication in the scope of *koɬ*.

v. $[\![koɬ]\!] = \lambda Q_{<s_s,t>}\lambda x \lambda e' \exists s.Q(s) \wedge Get(e') \wedge Locus(e',x) \wedge Theme(e',s)$

9.2 Where *koɬ* is licit: step-by-step derivations

Here, I proceed through each type of structure where we have observed the presence of *koɬ* to be licit:

(i) Change-of-state unaccusatives with *koɬ*.
(ii) Non-reflexive transitives with *koɬ*.
(iii) Transitive reflexives with *koɬ*.

9.2.1 Change-of-state unaccusatives with *koɬ*

Consider the minimal pair below:

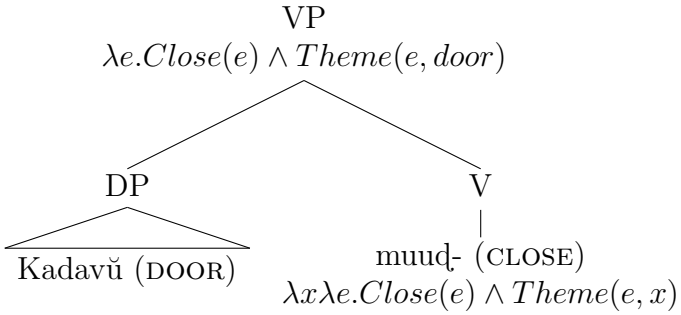
(268) Kadavŭ muuɖ-i-ttŭ.
door[NOM] close-PST-3NSG
"The door closed."

(269) Kadavŭ muuɖ-i-ko-ɳɖ-adŭ.
door[NOM] close-ASP-*koɬ*-PST-3NSG
"The door got (itself) closed."

Following Burzio (1986) and many others, I assume that the intransitive unaccusative sentence in (268) has the following argument-structure. For the sake of clarity and simplicity, I depict only the thematically relevant information – which, in this case, is the structure up to VP.[5]

[5]As mentioned earlier, I do not represent the inherent target-state of *muuɖŭ* (CLOSE) in the structures below since it will presumably be bound off by an even-

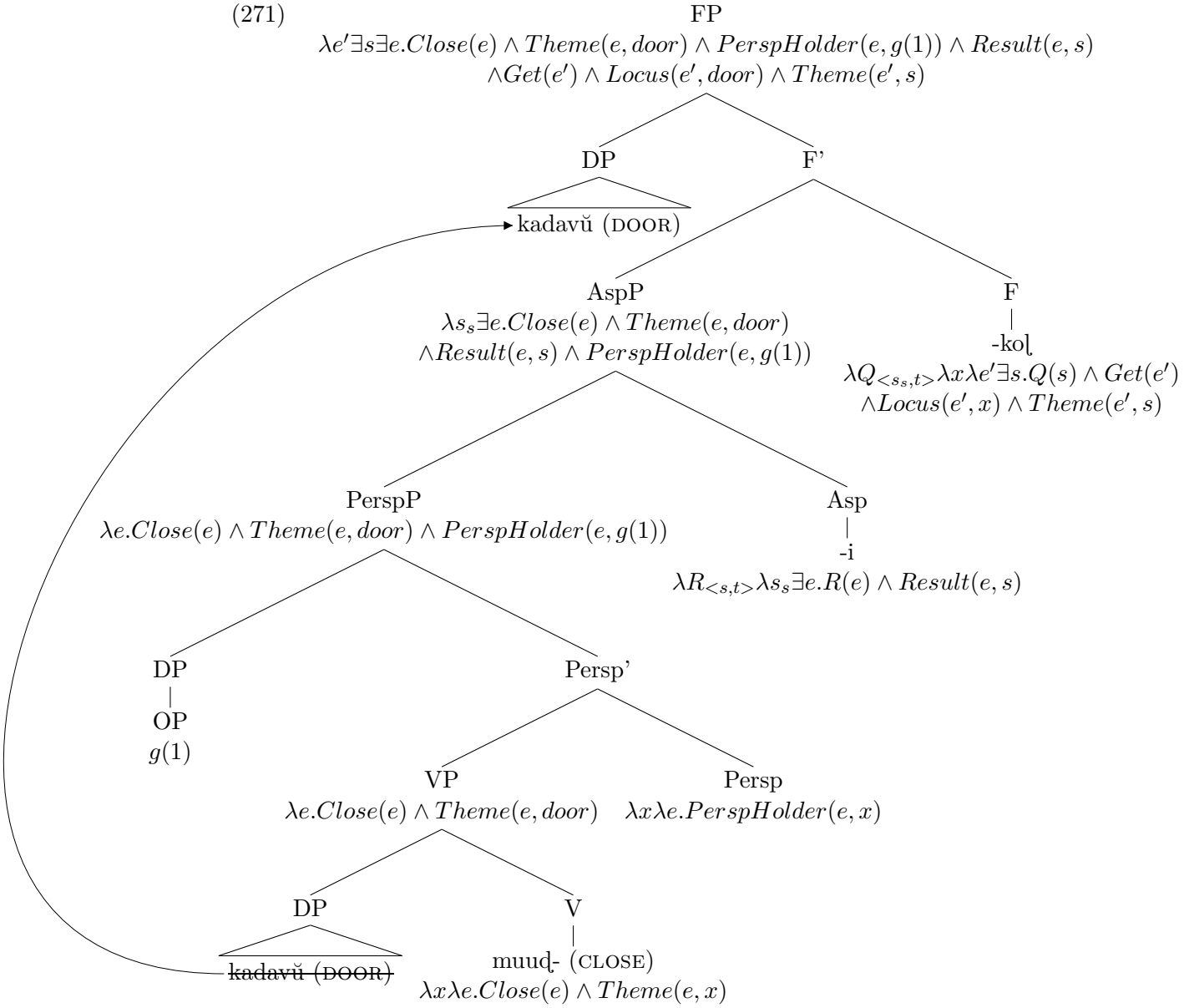
(270)



The structure for the *ko̥*-variant in (269) is as depicted below:

tivizer below Kratzerian (agentive) Voice anyway. However, I do represent the state introduced by Asp_{res} since this is necessary for *ko̥*-suffixation, as we have seen.

(271)



The argument-structure illustrated in (271) shows that the THEME argument, *kadavŭ* (DOOR) thematically raises from its VP-internal base position to its derived position in [Spec, FP], as discussed above. As a result, it gets associated with the semantics of the F head as a result of which it denotes an entity that comes to be the physical locus of the result state of the door-closing event. The operator in PerspP is a deictic pronoun containing the coordinates of the perspectival center. If it ends up getting mapped onto the referent of the DP *the door* in [Spec, ViewP], then this will be the perspective that the sentence expresses. However,

as a deictic pronoun, it could also be mapped onto the referent of a potential perspective-holder DP in a higher clause or in the discourse, in which case it will denote the perspective of this individual in the relevant evaluation context.

9.2.2 Non-reflexive transitive structure

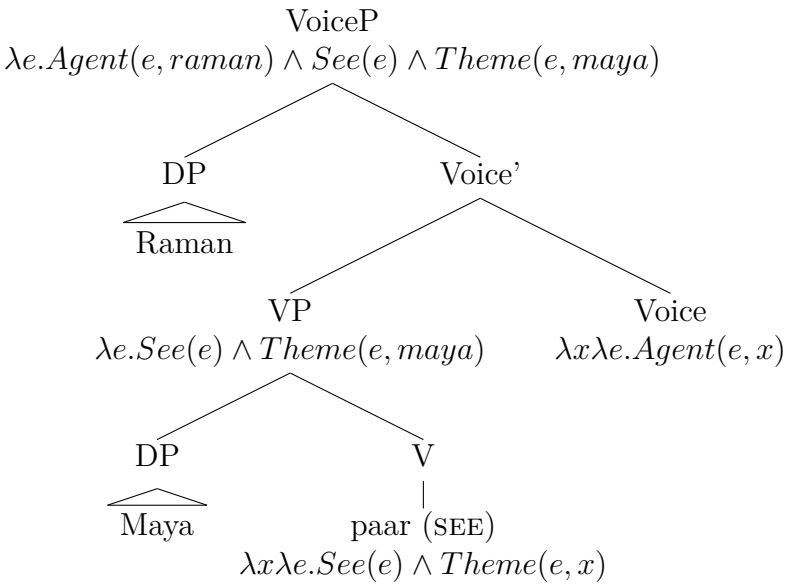
We saw that a majority of transitive verbs in Tamil, with the exception of certain verb-classes such as the so-called "alter-benefactive" predicates, optionally (sometimes preferably) allow the suffixation of *koḷ*. Here, I present argument-structures for the following *koḷ*- and *koḷ*-less structures:

(272) Raman Maya-væ paar-tt-aan.
Raman[NOM] Maya-ACC see-PST-3MSG
"Raman saw Maya."

(273) Raman Maya-væ paar-ttŭ-ko-ɳḍ-aan.
Raman[NOM] Maya-ACC see-ASP-*koḷ*-PST-3MSG
"Raman saw Maya (for his own benefit/for his own sake)."

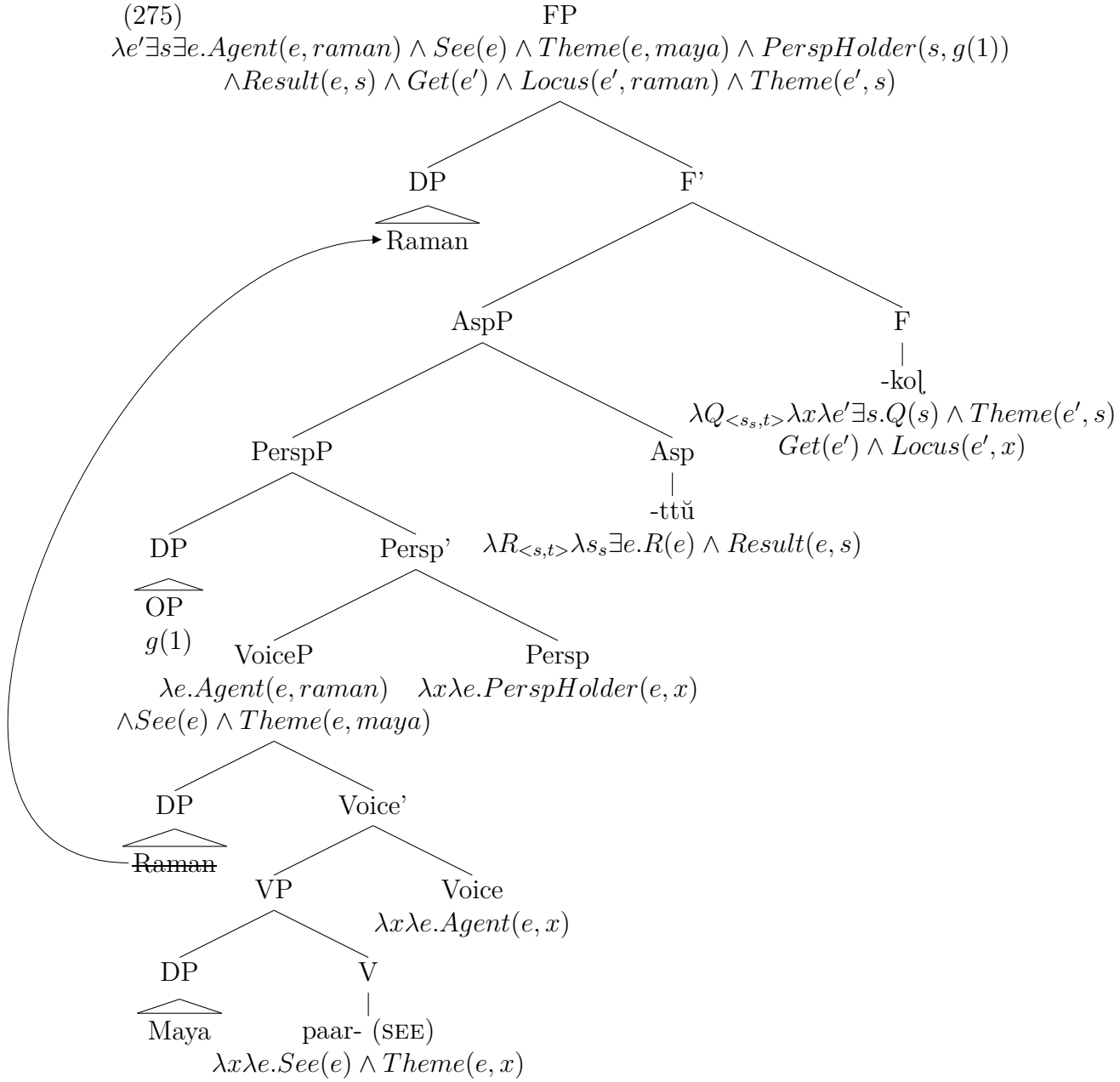
Under a Kratzer (1996) model with the external argument introduced in a Neo-Davidsonian manner by a functional Voice head, the sentence in (272) would have the following thematic structure:

(274)



The V head that spells out *koḷ*, we have assumed, is merged above the core proposition (represented by VoiceP in (274) above) and takes this proposition as its argument. The structure for the *koḷ*-transitive in (273) is thus as follows:

(275)



The denotation of F merely requires that the individual that occupies [Spec, FP] be a thematic participant of the result state description. It doesn't specify which argument this should be. And yet, we know that the DP that occupies this position must be the highest argument of the event predication under F. How do we formally model this restriction? I.e. specifically with respect to this sentence, how do we ensure that the AGENT *Raman* is the one that raises up to this position and not

Maya? I propose that this is mechanistically decided by the nature of the structure. Essentially, due to reasons of relativized minimality, the argument that is internally merged into [Spec, FP] will always be the one that is minimally close to it – thus, in a transitive structure this will always be the argument that is merged in [Spec, VoiceP] and never an argument that is merged lower down, like the Theme argument *Maya*.[6]

In this case, the F head assigns a mental Sells-like SELF role to the AGENT argument, here *Raman*. The root node of this structure thus states that there is a seeing event whose AGENT is *Raman* and whose theme is *Maya* and claims, crucially furthermore, that Raman comes to be the mental locus for the result state of this same event. Recall that this is indeed, entirely compatible with the results that were obtained when the sentences in (272) and (273) (in Section 6.3 of Chapter 6) were tested on 16 native speakers of Tamil. The sentence in (273) was deemed perfect just in those situations where the AGENT stood to benefit from or otherwise be affected by the seeing event; in scenarios where such a reading was explicitly ruled out – as, for instance, in a situation where Raman accidentally caught sight of Maya, the use of *koḷ* was considered to be significantly degraded in the sentence. The sentence additionally asserts that there is a Perspective-Holder whose mental or spatial perspective the entire clause represents. This perspective-holder may be the AGENT *Raman*, but it may just as well be some other salient perspective-holder in the immediate discourse or higher clause. This mapping is decided based on discourse-pragmatic factors pertaining to speaker intent, common ground, salience, and the like.

9.3 Syntax and semantics for the local binding of object *ta(a)n*

Consider the structure in (276) below:

(276) Raman$_i$ tann-æ$_{\{i,*j\}}$ aḍi-ttŭ-ko-ṇḍ-aan.
Raman[NOM] ANAPH-ACC hit-ASP-*koḷ*-PST-3MSG
"Raman$_i$ hit himself$_{\{i,*j\}}$."

The same structure without *koḷ* was found to be quite odd:

[6]Of course, if we ended up pursuing a Bruening (To Appear)-style control, rather than raising, analysis this would have to be separately stipulated as a condition on control, as discussed earlier.

(277) ?? Raman$_i$ tann-æ$_{\{i,*j\}}$ aḍi-tt-aan.
Raman[NOM] ANAPH-ACC hit-PST-3MSG
"Raman$_i$ hit himself$_{\{i,*j\}}$."

As per our discussion in the previous chapter, local binding (in the typical case) is possible just in case *koḷ* is suffixed onto the verb because it is the presence of *koḷ* that allows the co-argument DP of *ta(a)n* (i.e. *Raman* in (276)) to appear in a position above the minimal PerspP containing *ta(a)n*. Since *koḷ* additionally ascribes the semantics of a mental or physical viewpoint to this DP, it thus also qualifies as a potential antecedent – thus is now eligible to antecede *ta(a)n* by controlling the operator in [Spec, PerspP].

In our discussion of long-distance binding facts in Part I, we also showed that it is this operator that actually Agrees with *ta(a)n* in the syntax, a relationship that is construed at LF as semantic binding. We also proposed that *ta(a)n* has an unvalued DEP-feature as well as unvalued ϕ-features. The pronominal operator in [Spec, PerspP] on the other hand, was assumed to have valued ϕ-features and a valued DEP-feature, allowing the operator and *ta(a)n* to enter into an Agree relationship in the syntax for DEP and presumably also ϕ. The definition of DEP is reproduced below for convenience:

(278) **The DEP feature:**

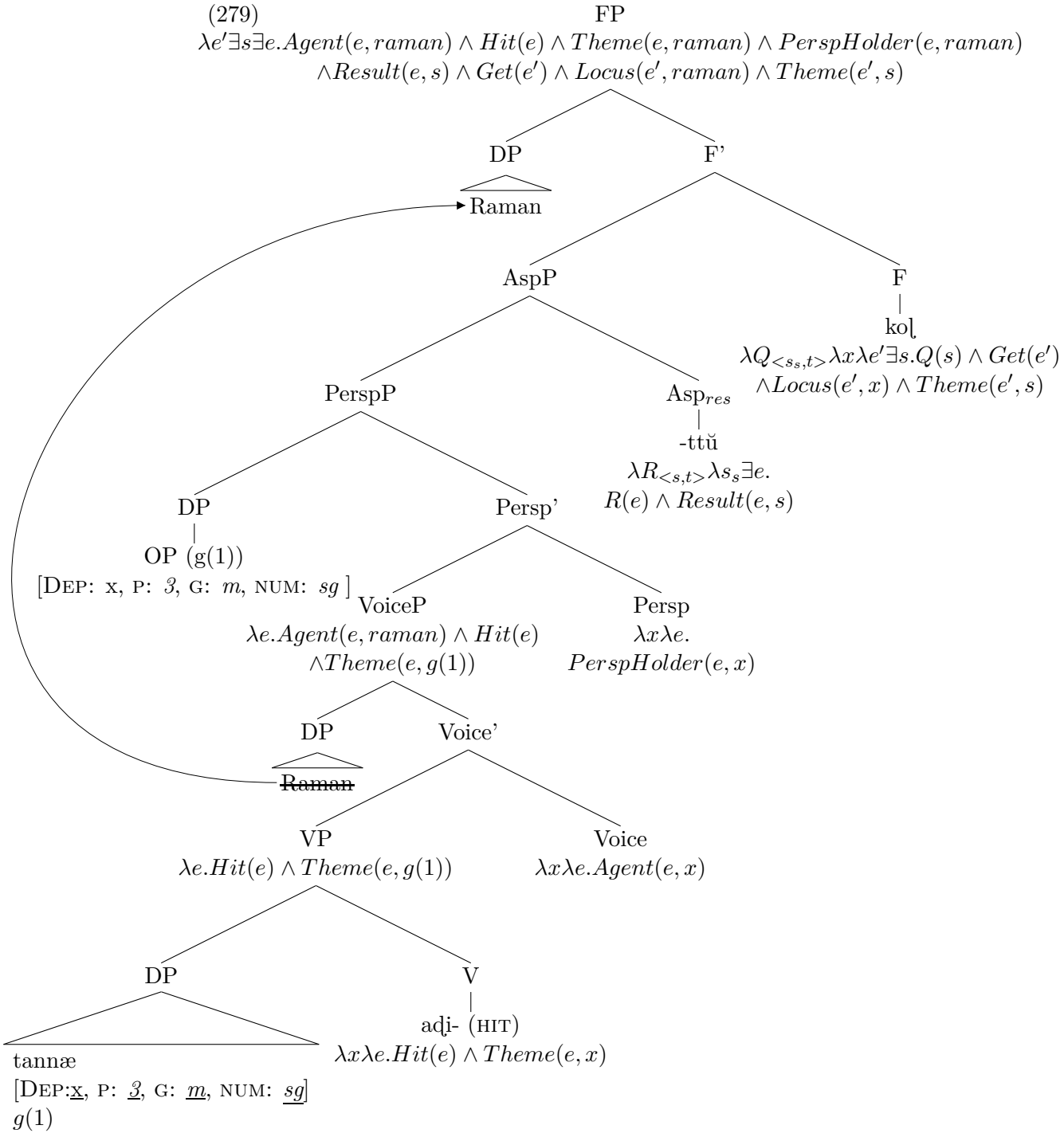
i. A DEP feature marks two DPs X and Y that are in a syntactic binding dependency with one another.

ii. DEP takes integers or letters as value. The assignment function maps these values to salient entities in the evaluation context. Two elements with matching DEP values will thus denote the same entity in the evaluation context and are construed to be in a binder-bindee relationship with one another.

iii. An anaphor is a nominal with an unvalued DEP feature – this is the syntactic correlate of anaphoricity; the operator in the specifier of the Perspectival Center is a nominal with a valued DEP feature.

In both local and long-distance binding configurations, therefore, *ta(a)n* is assumed to enter into an Agree relationship with the operator in [Spec, PerspP] for DEP; at LF, this relationship combined with the fact that the latter asymmetrically c-commands the former, will ensure that the latter binds the former rather than the other way around. Presuppositional constraints on reference assignment at LF will then regulate the mapping between DEP and the antecedent of *ta(a)n*. This is the

general idea, but we will walk through a derivation of the local binding structure below, to be clear:



There are several points worth discussing in the tree structure above. First, as mentioned just now, the local binding relation is instantiated as Agree between *ta(a)n* and the operator in [Spec, Persp-CenterP], since the operator is the minimally closest element with a valued DEP-feature in the *ta(a)n*-clause; as a result, *ta(a)n* will get its DEP-feature valued as *x*. At LF, this will result in the operator variable-binding *ta(a)n*, as indicated in the tree. Since one-many mapping is disallowed, the assignment function *g* will map the DEP-value *x* on the operator and *ta(a)n* to the same individual, in the context of evaluation, in its range. LF well-formedness conditions will ensure that the mapping is to one of the potential antecedents in its range, in the manner described in Part I. In the meantime, *koɭ* confers a viewpoint (mental or spatial, in this case) to the AGENT of VoiceP, namely *Raman*. This will cause *Raman* to move to [Spec, FP], from which position it will represent the mental/spatial locus with respect to the target result state of the event proposition in VoiceP. In other words, it will hold a perspective toward the event-description in which *ta(a)n* is a thematic argument, thus satisfying the condition on potential antecedence of *ta(a)n*, repeated below:

(280) **Condition for potential *ta(a)n*-antecedence (Final version):**

i. A potential antecedent of *ta(a)n* is a nominal which has a mental, temporal or spatial perspective with respect to the minimal CP, PP, or DP in which the anaphor is a participant (i.e. thematic argument).
ii. This information is represented as part of the perspectival center in the minimal CP, PP or DP containing the anaphor.

If the intent of the speaker is that *Raman* be construed as the actual antecedent of *ta(a)n* (in the sentence given in (276), *Raman* is in fact the only available antecedent, so this is trivial), the assignment function will map the DEP-value *x* to Raman, thus yielding the effect of local binding of *ta(a)n*. This is in fact the reading that is depicted in the tree – *Raman* is not just the potential antecedent, but the actual antecedent of *ta(a)n*.

Now consider a sentence like (281): it involves the sentence in (276) embedded under an attitude verb:

(281) Krishnan$_j$ [$_{CP}$ Raman$_i$ tann-æ$_{\{i,j\}}$ aḍi-ttŭ-ko-ŋḍ-
Krishnan[NOM] Raman[NOM] ANAPH-ACC hit-ASP-*koɭ*-PST-
aan-nnŭ] nene-tt-aan.
3MSG-COMP think-PST-3MSG
"Krishnan$_j$ thought [$_{CP}$ that Raman$_i$ hit himself$_{\{i,j\}}$]."

The anaphor *tannæ* in embedded direct object position now has two potential antecedents – the clausemate subject *Raman* and the matrix subject *Krishnan*. Recall that, even in cases of so-called "long-distance" binding, the syntactic and LF-semantic relationships between the operator in [Spec, PerspP] and *ta(a)n* are strictly phase-local. Thus, in both cases, the only relevant part of the structure for these modules of grammar is the minimal PerspP containing *ta(a)n*. Since the embedded clause contains *koɭ*, this is the AspP below *koɭ*. In other words, we end up with the same tree structure as that illustrated in (279) above. The only difference is that, until we know which of the two potential antecedents is intended to be the actual antecedent of *ta(a)n* – the operator in [Spec, PerspP] below the resultant AspP will remain specified as: $g(1)$ – i.e. as a type of deictic pronoun. If the speaker decides to make *Raman* the antecedent of *ta(a)n*, $g(1)$ will be mapped to Raman, in the evaluation context for this proposition; alternatively, if the speaker decides to make *Krishnan* the antecedent of *ta(a)n*, $g(1)$ will be mapped to Krishnan, in the evaluation context for the proposition.

9.3.1 Long-distance binding and *koɭ*

The discussion above serves as a nice segue into the question of why *koɭ* is not involved in the long-distance binding of *ta(a)n*. We have, in fact, already hinted at the answer to this question, but it is worth discussing more explicitly. Consider the long-distance binding patterns in (282) and (283) below:

(282) Raman$_i$ [$_{CP}$ Krishnan$_j$ tann-æ$_{\{i,*j\}}$ kaɳɳaaɖi-læ paar-
Raman[NOM] Krishnan[NOM] ANAPH mirror-LOC see-
tt-aan-nnŭ] nene-tt-aan.
PST-3MSG-COMP think-PST-3MSG
"Raman$_i$ thought [$_{CP}$ that Krishnan$_j$ saw himself$_{\{i,*j\}}$ in the mirror.]"

(283) Raman$_i$ [$_{CP}$ Krishnan$_j$ tann-æ$_{\{i,j\}}$ kaɳɳaaɖi-læ paar-ttŭ-
Raman[NOM] Krishnan[NOM] ANAPH mirror-LOC see-ASP-
ko-ɳɖ-aan-nnŭ] nene-tt-aan.
koɭ-PST-3MSG-COMP think-PST-3MSG
"Raman$_i$ thought [$_{CP}$ that Krishnan$_j$ saw himself$_{\{i,j\}}$ in the mirror.]"

The *koɭ*-variant in (283) induces local binding whereas its *koɭ*-less counterpart in (282) induces long-distance binding by the matrix subject *Raman*. We have already seen why the presence of *koɭ* is required for local

binding, thus the pattern in (283) is just as expected. But why is the absence of *ko*ɭ apparently required in the long-distance binding case in (282)?

The answer is simple: it is not that *ko*ɭ has to be absent. The condition is actually a weaker one: *ko*ɭ is simply not required to be present for the implementation of a long-distance binding relationship. This is because, when a DP is not a co-argument of *ta(a)n*, it is already outside the minimal PerspP containing *ta(a)n*; thus, there is no problematic containment relationship that would prevent it from holding a viewpoint towards the event-description as a whole. Since one of the contributions of *ko*ɭ is precisely to provide a way for a DP to get outside the event description containing *ta(a)n*, there is simply no need for it here.

At the same time, its presence does no harm. I.e. it doesn't block long-distance binding across it any more than an intervening attitude verb would do so – thus, the matrix subject *Raman* still qualifies as a potential antecedent for *ta(a)n* in (283). When *ko*ɭ occurs in a proposition that doesn't contain *ta(a)n*, its function is still the same. It takes the highest argument of the main event predication in its scope and allows it to appear in its specifier (either by control or raising, as we have seen). As such, it ascribes to this argument a semantics of viewpoint that says, simply, that this argument now holds a mental or physical viewpoint toward the result state of the main event that it has been a participant of. Whether the argument also denotes an entity that holds a perspective toward the *ta(a)n*-eventuality is thus an independent question. If the argument is animate, marked 3rd-person, and is also related to the *ta(a)n*-eventuality by an appropriate temporal, spatial, or mental subordinator (or other relator) – then it may qualify to do so. If not, it won't.

To sum up, the presence of *ko*ɭ in a proposition matters for the binding of *ta(a)n* only if *ta(a)n* is also a thematic participant in that same proposition. Otherwise, the two properties are unrelated to one another. Thus, in the sentences below, *ko*ɭ doesn't block *ta(a)n*-binding across it; but neither does it facilitate new binding possibilities for it:

(284) Raman$_i$ [$_{CP}$ Krishnan$_j$ tann-æ$_{\{i,*j\}}$ kaɳɳaaḍi-læ paar-
Raman[NOM] Krishnan[NOM] ANAPH mirror-LOC see-
tt-aan-nnŭ] nene-ttŭ-kko-ɳḍ-aan.
PST-3MSG-COMP think-ASP-*ko*ɭ-PST-3MSG
"Raman$_i$ thought [$_{CP}$ that Krishnan$_j$ saw him$_{\{i,*j\}}$ in the mirror.]"

(285) Raman$_i$ [$_{CP}$ Krishnan$_j$ tann-æ$_{\{i,j\}}$ kaɳɳaaḍi-læ
Raman[NOM] Krishnan[NOM] ANAPH mirror-LOC

paar-ttŭ-ko-ɳɖ-aan-nnŭ] nene-tt-aan.
see-ASP-*koɭ*-PST-3MSG-COMP think-PST-3MSG
"Raman$_i$ thought [$_{CP}$ that Krishnan$_j$ saw him$_{\{i,j\}}$ In mirror.]"

(286) Raman$_i$ [$_{CP}$ Krishnan$_j$ tann-æ$_{\{i,j\}}$ kaɳɳaaɖi-læ
Raman[NOM] Krishnan[NOM] ANAPH mirror-LOC
paar-ttŭ-ko-ɳɖ-aan-nnŭ] nenettŭ-kko-ɳɖ-aan.
see-ASP-*koɭ*-PST-3MSG-COMP think-PST-3MSG
"Raman$_i$ thought [$_{CP}$ that Krishnan$_j$ saw him$_{\{i,j\}}$ in the mirror.]"

The interpretation of (284) is something like: "Raman$_i$ thought to himself$_{\{i,*j\}}$ [$_{CP}$ that Krishnan saw him$_{\{i,*j\}}$ in the mirror]." Similarly, the only difference that adding *koɭ* to the embedded *ta(a)n*-clause makes, in (285), is that it allows *ta(a)n*'s clausemate subject *Krishnan* to qualify as a potential antecedent for it. The matrix subject *Raman* being in a higher attitude clause, is a potential antecedent as well (but this has nothing to do with *koɭ*). The sentence in (286) merely combines both these effects.

9.3.2 Psych predicates, local binding, and *koɭ*

Here, we will model the following distinctive properties of psych predicates:

(i) That they are degraded with *koɭ*.

1. That stative psych-predicates are a bit more degraded than eventive ones with *koɭ*.
2. That psych-predicate structures with dative EXPERIENCER subjects are more degraded with *koɭ* than those with nominative EXPERIENCER subjects.

(ii) That their EXPERIENCER subjects are capable of locally anteceding *ta(a)n*.

Consider the psych-sentences below:

(287) Raman-ŭkkŭ$_i$ tann-æ$_{\{i,*j\}}$ piɖi-tt-adŭ/*piɖi-ttŭ-ko-ɳɖ-adŭ.
Raman-DAT ANAPH-ACC like-PST-3NSG/*like-ASP-*koɭ*-PST-3NSG
"Raman$_i$ liked himself$_{\{i,*j\}}$"

(288) Raman$_i$ tann-æ-jee$_{\{i,*j\}}$ mara-nd-aan/??mara-ndŭ-ko-ɳɖ-aan.
Raman[NOM] ANAPH-ACC.EMPH forget-PST-3MSG/??forget-ASP-*koɭ*-PST-3MSG
"Raman$_i$ forgot (even) himself$_{\{i,*j\}}$"

Let us review why *ko*ɭ is prohibited in such structures – a point that was discussed in some detail in Section 7.4.3.1 in Chapter 7. There we proposed that stative psych-predicates (like LIKE, HATE, and KNOW) disallow *ko*ɭ because, being inherently stative, they are incompatible with the derived result state encoded by Asp_{res}. Syntactically, therefore, such verbs lack an Asp_{res} head. Since the presence of *ko*ɭ is dependent on that of the aspectual morpheme encoded by Asp_{res}, such verbs are infelicitous with *ko*ɭ. We also saw that both eventive and stative psych predicates disallow *ko*ɭ because the "mental locus" semantics of *ko*ɭ is, in some sense, already a part of their denotation. Where is this semantics structurally introduced? The answer is contingent to some degree on whether such predicates are compatible with Asp_{res}.

With stative predicates, we have just seen that they are not, thus it is clear that, for these predicates at least, the semantics of mental "experience" must be encoded elsewhere. In Section 8.4, we saw that eventive psych verbs need not have an Asp_{res} projection – in other words, we argued that the argument structure of such verbs instantiates a single extended functional sequence. At the same time, it should in theory be possible to build a derived result state on top of an eventive psych verb via the addition of Asp_{res}. But in such cases, we would also expect *ko*ɭ to be able to attach on top of that. This expectation is borne out:

(289) Raman baya-nd-aan.
Raman[NOM] fear-PST-3MSG
"Raman got scared." (Rough translation)

(290) Raman baya-ndŭ-kko-ɳɖ-aan.
Raman[NOM] fear-ASP-*ko*ɭ-PST-3MSG
"Raman got scared."

The two sentences seem to have near identical meanings – nevertheless, both are in theory possible. These sentences show, first, that the semantics of what we have informally paraphrased as mental "experience" is different from that of $\text{F}_{koí}$ – in particular, we cannot say that in (289), it is realized by a silent spell-out of $\text{F}_{koí}$ because then we would expect *ko*ɭ suffixation to be impossible. In contrast, in (290), the DP is merged in the specifier of $\text{F}_{koí}$ which yields a very similar semantics.

However, this is not the whole story. When the subject EXPERIENCER of an eventive psych-predicate is dative-marked, it is strictly impossible to add *ko*ɭ:

(291) Raman-ŭkkŭ koovam-va-nd-adŭ.
Raman-DAT anger-come-PST-3NSG

"Raman got angry." (Rough translation)

(292) * Raman-ŭkkŭ koovam-va-ndŭ-kko-ɳɖ-adŭ.
Raman-DAT anger-come-ASP-*koɭ*-PST-3NSG
"Raman got angry." (Intended)

I propose that this is because, for some reason, dative-marked DPs are not allowed in the specifier of $F_{koí}$. In other words, I propose that dative marked experiencers may only be merged in a different position.

With respect to the question of how *ta(a)n* may be anteceded by a co-argument in a psych-predicate structure, despite the absence of *koɭ*, we proposed in Section 8.4, that the EXPERIENCER subject is *externally* merged above the minimal PerspP containing *ta(a)n* – which, in psych-predicate structures that do not contain a resultant AspP, either because they are explicitly incompatible with it (like stative psych-verbs) or because they don't require it – is the PerspP in the minimally c-commanding C-layer. This might be interpreted as the designated structural position for the assignment of an EXPERIENCER θ-role – in other words, we might claim that EXPERIENCERS are merged relatively high in the structure.

Returning now to the question of why eventive psych-verbs with dative marked subjects, as in (291), are incompatible with *koɭ*, I propose that this is because dative-marked subjects may *only* be externally merged in the specifier of the functional head in the C layer where high experiencers are introduced. Another way of stating this restriction would be to say that dative-marked DPs in subject position ("quirky" dative subjects) are always and only interpreted as EXPERIENCERS. This is, of course, a stipulation but it is a reasonable one to make. Quirky dative case is generally associated with the first Merge position of a DP (Woolford 2006). Thus, if this case is assigned in a high EXPERIENCER position, as we are claiming, then we would expect that the relevant DPs must be externally merged in this position, not moved there from a lower position having passed through the specifier of $F_{koí}$. Furthermore, there is independent evidence suggesting that EXPERIENCERS are merged structurally rather high – see, for instance Adger and Ramchand (2006) for arguments to this effect.

This state of affairs may be summarized as follows:

(293) **Modelling *koɭ*-incompatibility for psych-predicates:**

i. **Stative psych-verbs:** They are inherent statives. Thus, they lack a (resultant) Asp_{res}; thus they cannot co-occur with *koɭ*.

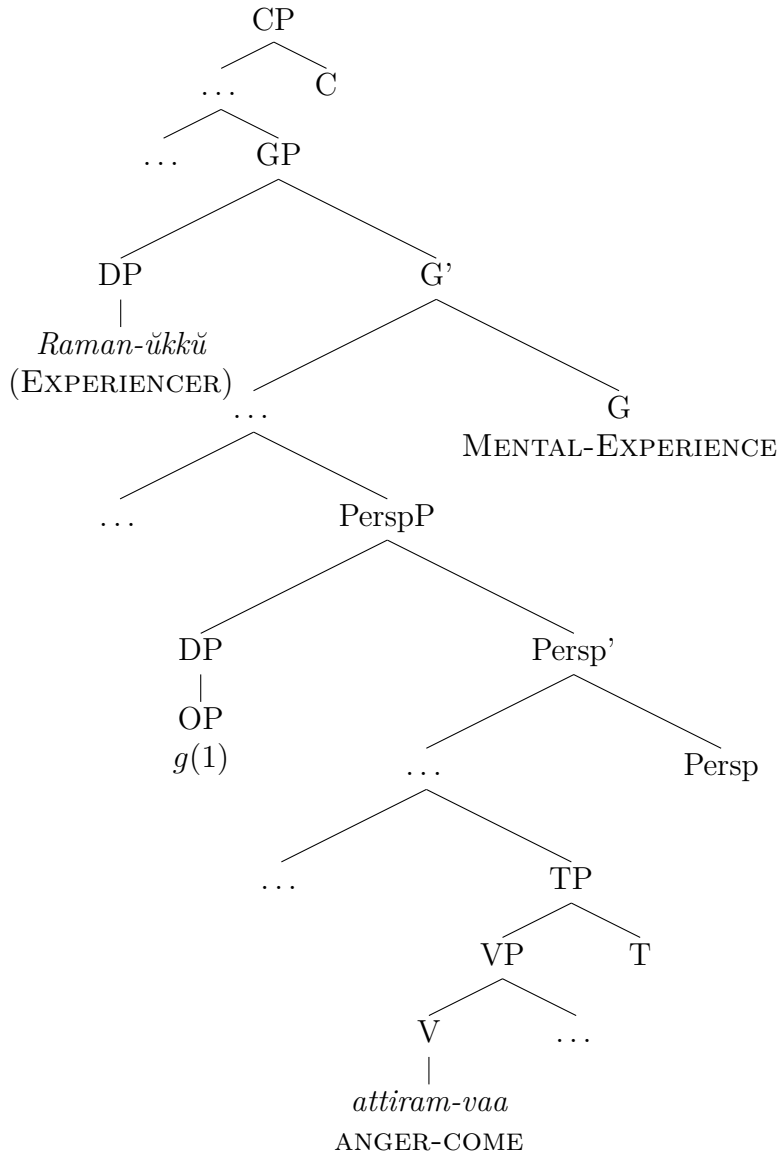
ii. **All psych-verbs:** Possess the *mental locus* semantics of *ko⌊*, or something like it, as part of their denotation (let us nominally label this semantics as MENTAL EXPERIENCE for now).
iii. **Eventive psych-predicates:** These are, in theory, compatible with Asp_{res}, thus by extension with *ko⌊*. Indeed, eventive psych-verbs with nominative-marked structures may optionally instantiate a structure with *ko⌊* (see (290) above).
iv. **Eventive psych verbs with "quirky" dative subjects:** These are independently incompatible with *ko⌊* because of a restriction that dative-marked EXPERIENCERS must be externally merged in the specifier of a functional head in the C layer.

(294) **Modelling binding by a co-argument** EXPERIENCER:

i. A MENTAL EXPERIENCE semantics is assigned by a head in the C-layer to a DP in its specifier. This DP comes to be interpreted as an EXPERIENCER.
ii. This DP is, crucially, above the PerspP in the C layer, which is also the minimal PerspP containing *ta(a)n*, given the absence of a resultant AspP. By virtue of its own thematic properties and the fact that it asymmetrically c-commands this PerspP, it may serve as a potential antecedent for *ta(a)n*.

Based on these conclusions, I now propose the following structure for an eventive psych-predicate with a dative EXPERIENCER subject, as in (291) above:

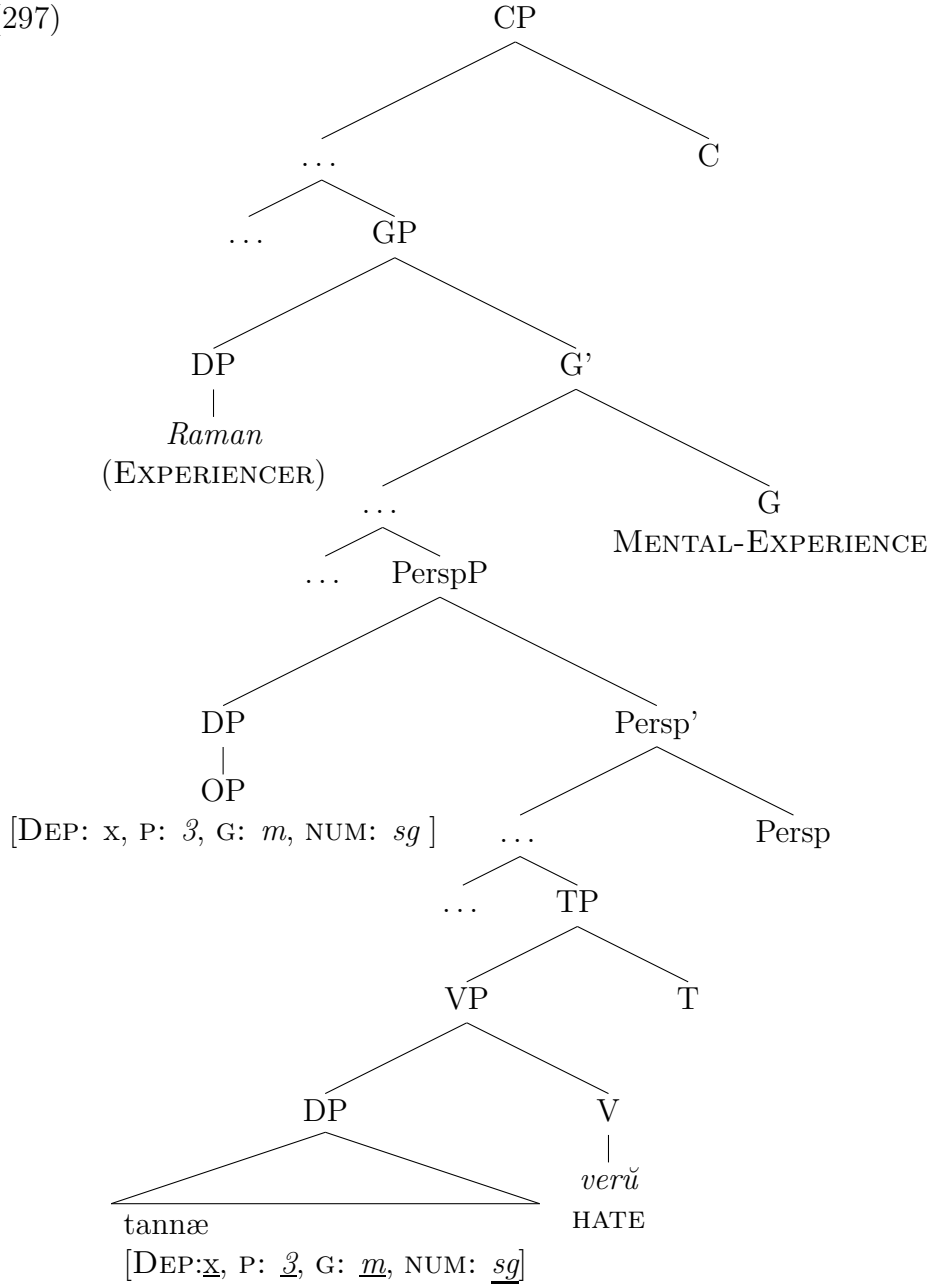
(295) **Eventive Psych Verbs with "quirky" dative** EXPERIENCER **subjects:** *aattiramvaa* **(**GET ANGRY**):**



The relevant parts of the structure of a stative psych-verb – regardless of whether it takes a nominative or dative EXPERIENCER subject – are identical to that in (295). The main difference is that the primary eventuality is stative rather than eventive. A tree-structure for a sentence like (296), which takes a nominative EXPERIENCER subject, is given below. In addition, I have made the direct object be locally-anteceded *ta(a)n*:

(296) Raman$_i$ tann-æ$_{\{i,*j\}}$ veru-tt-aan.
Raman[NOM] ANAPH-ACC hate-PST-3MSG
"Raman$_i$ hated himself$_{\{i,*j\}}$."

(297)



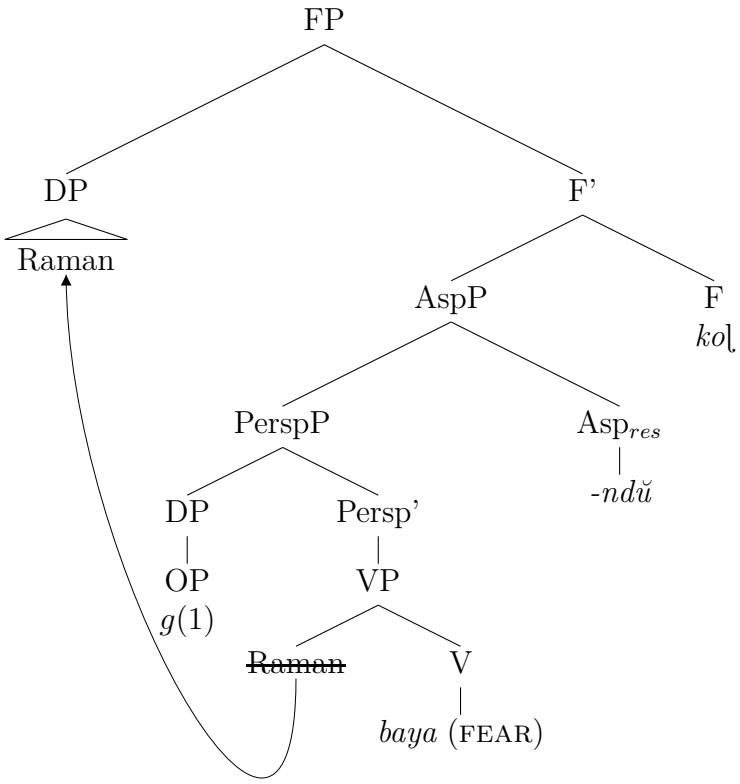
As mentioned above, stative psych-verbs such as these are simply incompatible with an Asp_{res}, thus by extension are incompatible with *ko*ɭ. We have modelled this in (297) by making the projections for Asp_{res} and $\text{F}_{koí}$ be absent in the structure. The minimally closest PerspP to *ta(a)n*, whose operator Agrees with it for the Dep-feature in the syntax and binds it at LF is thus the one in the C-layer. Since the experiencer *Raman* is externally merged above this PerspP and is additionally the-

matically and ϕ-featurally qualified to be a perspective-holder for the *ta(a)n*-eventuality, it is a potential antecedent of *ta(a)n*. If the DEP-value of x gets mapped onto Raman in the evaluation context by the LF assignment function, then *Raman* will be the actual antecedent of *ta(a)n* in this sentence.

To finish, let us see what the structure of an eventive psych verb with a nominative-marked subject looks like when it combines with *koɭ*. The sentence below is repeated from (290):

(298) Raman baya-ndŭ-kko-ɳɖ-aan.
Raman[NOM] fear-ASP-*koɭ*-PST-3MSG
"Raman got scared."

(299)



$$[\![FP]\!] = \lambda e' \exists s \exists e. Locus(e', Raman) \wedge Get(e') \wedge Theme(e', s) \wedge Result(e, s) \wedge PerspHolder(e, g(1)) \wedge Theme(e, Raman) \wedge Fear(e)$$

The tree in (299) shows that *Raman* is interpreted as the entity that comes to represent the mental locus of the result state of the fearing event – just like other arguments that appear in the specifier of *koɭ*. In other words, it is not really an EXPERIENCER in the way in which we have described it. In contrast, *Raman* in the *koɭ*-less variant of (298) –

as in (289) above – is an EXPERIENCER and is merged in the specifier of a functional head in the C layer, just like the EXPERIENCER subjects in (295) and (297) above.

9.4 Capturing microvariation

The data on local reflexives presented here has conformed primarily to my own dialect of Tamil. However, the results of my survey shows that there is systematic microvariation with respect to the interpretation of such data. Here, I will discuss another dialect which manifests the following behavior with respect to local *ta(a)n*-antecedence:

- The addition of *ko*ɭ is strictly optional for local reflexivity. *ko*ɭ seems to mean the same thing for them as it does for me (and other speakers of my dialect of Tamil).
- I.e. *ta(a)n* may be locally anteceded under the absence of *ko*ɭ even in a non-psych-predicate structure.

In other words, for speakers of this dialect, the Ban on Clausemate Subject Antecedence doesn't hold. Thus, sentences like (300) and (301) are both grammatical for them:

(300) Raman$_i$ tann-æ$_{\{i,*j\}}$ aɖi-tt-aan.
Raman[NOM] ANAPH-ACC hit-PST-3MSG
"Raman$_i$ hit himself$_{\{i,*j\}}$."

(301) Raman$_i$ tann-æ$_{\{i,*j\}}$ aɖi-ttŭ-kko-ɳɖ-aan.
Raman[NOM] ANAPH-ACC hit-ASP-*ko*ɭ-PST-3MSG
"Raman$_i$ hit himself$_{\{i,*j\}}$ (for his own benefit)."

This dialect seems to correspond to that reported for Tamil by Annamalai (1999). Annamalai proposes for instance that *ta(a)n* has the same antecedence possibilities in the minimal pairs below:

(302) Kumar$_i$ [$_{CP}$ Raman$_j$ tann-æ$_{\{i,j\}}$ tiʈʈ-i-kko-ɳɖ-aan-nnŭ]
Kumar[NOM] Raman[NOM] ANAPH-ACC scold-PST-3MSG-COMP
so-nn-aan.
say-PST-3MSG
"Kumar$_i$ said [$_{CP}$ Raman$_j$ scolded himself$_j$/him$_i$]"

(303) Kumar$_i$ [$_{CP}$ Raman$_j$ tann-æ$_{\{i,j\}}$ tiʈʈ-in-aan-nnŭ]
Kumar[NOM] Raman[NOM] ANAPH-ACC scold-PST-3MSG-COMP
so-nn-aan.
say-PST-3MSG

"Kumar$_i$ said [$_{CP}$ Raman$_j$ scolded himself$_j$/him$_i$]"

Annamalai proposes that *ko*l contributes a semantics of self-benefaction or self-affectation onto the subject "of some action predicated of it" (Annamalai 1999, 179). Although Annamalai then goes onto treat *ko*l as a voice-marker – an analysis that, we have seen, isn't really viable – his description of *ko*l is reminiscent of what we ourselves have proposed for it. Indeed, the responses of the survey-takers who also exhibited this dialect confirm this proposal: they tended to prefer the use of *ko*l in pragmatic scenarios where the highest argument of an event could be construed as the mental/physical locus of the result state of the event, just as we have described so far. The antecedence possibilities for *ta(a)n* in the *ko*l-sentence in (302) corresponds to that of the grammar we have described so far: both the clausemate subject *Raman* and the matrix subject *Kumar* are possible antecedents for *ta(a)n*– however, Annamalai states that *Raman* is favored as the antecedent in the default pragmatic scenario because of *ko*l's semantics of "self-benefaction".

The difference between the two dialects becomes apparent in the antecedence-possibilities for *ta(a)n* in the *ko*l-less variant in (303). In the grammar we have seen thus far, the only possible antecedent for *ta(a)n* would be the matrix subject *Kumar*; the clausemate subject *Raman* would not be a possible antecedent – a property we have explained by claiming that the minimal PerspP containing *ta(a)n* in this sentence also contains (specifically, asymmetrically c-commands) the intended perspective-holder *Raman*. In the dialect that Annamalai describes and which a statistically significant subset of my survey-takers also seem to speak, the clausemate subject *Raman* may antecede *ta(a)n* in this structure, however, as illustrated in (303).

How do we formally model the properties of this dialect within the broader parameters of the binding analysis being developed here? There are several logical possibilities, as I discuss below. But we will see that not all possibilities are created equal: some are clearly less desirable than others.

9.4.1 A perspective holder may be contained inside the PerspP

Recall that we explained the Ban on Clausemate Subject Antecedence as resulting from a structural restriction that a Perspective Holder (anchor) may never be embedded inside the predicational structure (object of scrutiny) that it holds a perspective towards. We showed that non-

psych predicate structures where *ta(a)n* is anteceded by a co-argument DP (like (300) above) are ruled out because they violate this condition. A simple way to capture the fact that speakers for whom sentences like (300) *are* grammatical might be claim that this structural restriction simply doesn't hold. I.e. we might propose that, for these speakers, a DP may serve as a perspective holder toward the *ta(a)n*-eventuality even if it is asymmetrically c-commanded by the minimal PerspP containing *ta(a)n* – here, the PerspP in the C-layer of the minimal *ta(a)n*-clause.

Under this approach, our analysis of the syntax, semantics, and morphological status of *ko*ɭ could remain untouched. Such an analysis would, thus, technically work. However, it is an unattractive one simply because the condition that a perspective-holder may not be contained inside the predicational structure it holds a perspective towards goes to the heart of what we mean by "perspective". Getting rid of the structural condition would force us to revise our notion of "perspective" from first-principles.

9.4.2 There is a PerspP below VoiceP/*v*P

A different possibility would be to say that, in sentences like (300), where *ta(a)n* is locally anteceded in the absence of *ko*ɭ, there is a PerspP below the clausemate subject – introduced perhaps below *v*P/VoiceP. Such a proposal wouldn't be that controversial given standard Minimalist analyses (Chomsky 2001) which assume that *v*P is a phase. Thus, if we are claiming that there is a unique PerspP for every phase, it would follow that there is a PerspP below VoiceP, crucially below where the co-argument DP is merged.

Such a proposal is more attractive because it allows us to retain the idea that a perspective-holder may not be embedded inside the predicational structure that it holds a perspective towards. However, capturing the fact that speakers of this dialect may optionally add *ko*ɭ to such structures might be more challenging under our current analysis for *ko*ɭ. I.e. we could no longer maintain the idea that the resultant AspP selected by *ko*ɭ is a phase with a unique PerspP if VoiceP/*v*P were also to be treated as a phase.

9.4.3 The clausemate subject has a DEP feature

An alternative would be to model microvariation as a function of differences in the identity of the binder. In the current analysis, the actual binder is the null pronominal operator in the specifier of the minimal PerspP containg *ta(a)n*. We have syntactically modelled this by propos-

ing that this operator bears a valued DEP-feature which then checks the unvalued DEP-feature on *ta(a)n* – an Agree dependency that triggers binding of *ta(a)n* by this operator at LF. In the current model, the relationship between *ta(a)n* and the DP that serves as its actual antecedent is thus indirect, mediated by the operator in [Spec, PerspP]. But we could imagine that, in this new dialect, normal DPs may also have valued DEP-features, allowing them to directly bind *ta(a)n*. In other words, we might imagine that the relationship between *ta(a)n* and the DP that ends up anteceding it is unmediated by the operator in [Spec, PerspP] and obtains directly between anaphor and antecedent DP.

The problem with this analysis is that it is unclear how to derive optionality of antecedence. I.e. if the relationship between the antecedent DP and *ta(a)n* were a direct syntactic one, then we would expect the minimally closest DP to always antecede *ta(a)n*. Furthermore, long-distance binding in general would be problematic both in terms of locality and Relativized Minimality. The only way to make it work, it seems, would be to claim that DPs *optionally* bear a DEP-feature, a stipulation that would explicitly undermine the deterministic manner in which syntactic operations are thought to proceed. A further complication is that, for speakers of this dialect, *ta(a)n* does seem to be subject-oriented: thus, *ta(a)n* in embedded subject position would be anteceded by a superordinate subject even if there is a minimally closer DP in c-commanding object position. One could, of course, try to get around this by claiming that DEP is a feature which targets DPs in syntactic subject position alone: but this, in addition to being an extra stipulation, also predicts a bijectionary mapping between *ta(a)n* and antecedence by a syntactic subject which is not empirically supported.

None of these options are particularly appealing, on theoretical as well as empirical grounds. Below, I propose a final alternative which seems more plausible.

9.4.4 *ta(a)n* contains a PerspP

Finally, we might propose that *ta(a)n* spells out an internally complex DP in this new dialect which includes its own PerspP. This could be construed to be parallel to the contribution of SELF-morphemes in languages like English and Dutch (Reuland 2011), with the difference that the "SELF" morpheme is simply not pronounced here. One way or another, the addition of these morphemes to simplex anaphors seems to render them "locally" bindable. If we translate this into the current system, it could mean that the SELF-morpheme in fact signals the presence of a Persp

head meaning that, in fact, what looks like local binding is actually a type of long-distance binding. We have already seen that PerspP may mark certain DPs, PPs, CPs, and AspPs – thus, it should in theory be possible for one of these DPs to be the anaphor itself.

Incidentally, this analysis is very similar to one proposed by Jayaseelan (1997) for "local" binding phenomena in a range of languages: although Jayaseelan's analysis doesn't include the syntacticization of pragmatic and lexical-conceptual properties like "perspective", it proposes that the SELF-morpheme delineates a locality domain for reflexivity. Such an analysis is potentially also compatible with that in Reuland (2011). Reuland proposes that SELF-marking on a simplex anaphor "protects" that anaphor from forming a direct syntactic chain with its antecedent – the idea that an anaphor should not do so is somewhat reminiscent of my proposal that binding does not involve a direct relationship between the anaphor and the antecedent DP.

We can also easily accommodate the optionality of *ko*ɭ for speakers of this dialect, evidenced in sentences like (301). Since they already have an independent mechanism involving the silent "SELF"-like morpheme to facilitate local binding, *ko*ɭ is not necessary for this purpose. However, we can assume that *ko*ɭ still makes the same syntactic and thematic contributions and, therefore, will be used when the speaker wishes to convey a "coming to hold" semantics for the result state of the main event in its scope, onto the highest argument of this event.

A potential drawback at this juncture is that, given the simplex morphological form of *ta(a)n* which is, furthermore, identical to that found in my dialect, there is no independent evidence for the additional internal structure. Works such as Déchaine and Wiltschko (2002) have argued that pronouns are not primitives, and that different degrees of structural complexity may underlie distinct pronominal forms in natural language. Along these lines, further research would thus have to look for evidence of some additional complexity in the *ta(a)n* found in this new dialect.

I have noted the main analytic possibilities for dealing with the lack of a Ban on Clausemate Subject Antecedence for speakers of this new dialect, while retaining the notion that *ko*ɭ has the same interpretation that we have given it thus far. I have also pointed out the various advantages and drawbacks to each approach. As always, the decision must be made on an empirical basis, but the final alternative proposed here seems to me to be the most attractive of all. At this juncture, however, the evidence on this dialect isn't sufficient to make a clear decision. This is a matter for future research.

Part III

Monsters, anaphora and agreement

Chapter 10

Monstrous agreement: what it is, what it isn't

10.1 The puzzle

The focus of this series of chapters is a special sort of long-distance binding structure that exhibits what I call "monstrous agreement". Consider the minimal pair below:

(304) Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ poo⟦ʈ⟧i-læ ʤejkkapoo-r-**aa**⟦ɭ⟧-ŭnnŭ]
Maya[NOM] ANAPH[NOM] contest-LOC win-PRS-3FSG-COMP
namb-in-aa⟦ɭ⟧.
believe-PST-3FSG
"Maya$_i$ believed [$_{CP}$ that she$_{\{i,*j\}}$ would win the contest.]"

(305) Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ poo⟦ʈ⟧i-læ ʤejkkapoo-r-**een**-nnŭ]
Maya[NOM] ANAPH[NOM] contest-LOC win-PRS-1SG-COMP
so-nn-aa⟦ɭ⟧.
say-PST-3FSG
"Maya$_i$ said [$_{CP}$ that she$_{\{i,*j\}}$ would win the contest.]"

The sentence in (304) is just a standard instance of long-distance binding: the subject anaphor *ta(a)n* is long-distance bound by the matrix subject *Maya*, which is also the AGENT of the intensional predicate that directly selects the *ta(a)n*-clause. In accordance with the antecedence condition developed in Part I, *Maya* holds a mental perspective toward the *ta(a)n*-proposition. In the minimally varying sentence in (305), *ta(a)n* is still long-distance bound by the matrix subject, but the matrix predicate is now the speech-predicate *soll* (SAY), rather than the thought-predicate *nambŭ* (BELIEVE). In addition, the two sentences differ with respect to the agreement marking on the clausemate verb under *ta(a)n*:

in (304), this marking straightforwardly reflects the ϕ-features of *ta(a)n*'s antecedent *Maya*, but in (305), the agreement is 1SG and doesn't obviously agree with anything in the sentence.

In the course of these chapters, I will show that the 1st-person marking on the embedded verb in (305) instantiates indexical shift (Schlenker 2003b, Anand 2006, von Stechow 2002) in the embedded clause – and label it "monstrous agreement". I.e., I will argue that this agreement is triggered by a 1st-person pronoun in the embedded clause whose person-feature is evaluated against the linguistic context introduced by the matrix speech predicate *soll*, rather than against the utterance context, as is standard. I will also show that monstrous agreement only obtains under the scope of speech predicates.[1]

This combination of properties has the following important consequences. First, the minimal pair in (304)-(305) shows conclusively that *ta(a)n* is not itself an obligatorily shifted indexical: after all, the monstrous agreement only obtains in one of these sentences (namely (305)) whereas *ta(a)n* is present in both. In doing so, it debunks a popular hypothesis in the literature (following Schlenker 2003b) – motivated by the observation that logophors have an obligatory 1st-personal *de se* semantics – that logophors/anaphors are nothing but obligatorily shifted indexicals. Second, it refines our understanding of the relationship between anaphora and agreement from Part I of this dissertation. There, we argued that the agreement under *ta(a)n*, in standard long-distance structures involving this anaphor in subject position, could not have been triggered by *ta(a)n* itself but must have been inherited from another element in the local domain. Minimal pairs like (304)-(305) above reinforce this conclusion. After all, assuming that the same underlying element *ta(a)n* occurs in both sentences, we would hardly expect a difference in agreement if it is indeed *ta(a)n* which is triggering it. Third, it shows that, in addition to perspectival information in PerspP, there is a further kind of information which is syntactically represented. PerspP, which we have argued is responsible for the binding of *ta(a)n*, must be present in (304) and (305), since this anaphor is legitimately bound in both cases. Hence, it must be something else – which, I will argue, is information pertaining to the intensional speech context – which is responsible for the indexical shift observed in (305). Furthermore, since this affects morphological agreement, it must be represented as early as the Narrow Syntax.

[1] At the same time, for the dialect that will be primarily discussed here (which is also my dialect), agreement is not obligatorily monstrous under speech verbs, a fact that we will formally model down the line.

10.2 Eliminating the usual suspects

One might suspect that the 1st-person agreement marking on the embedded verb in (305) does not represent indexical shift after all, but something much more pedestrian. In this section, I will consider three such alternatives and rule them out on independent empirical grounds.

10.2.1 1st-person agreement: not default agreement or frozen

One plausible response to the agreement facts in (305) might be to claim that the 1SG agreement on the embedded verb actually exemplifies default agreement. Default agreement in Tamil typically obtains in constructions involving experiencer subjects with dative case (as in 306):

(306) Raman-ŭkkŭ koovam va-nd-adŭ.
Raman-DAT anger[NOM] come-PST-3NSG
"Raman became angry." (Lit. "Anger came to Raman")

However, default agreement in Tamil is always marked 3NSG – thus cannot be responsible for the 1SG agreement on the embedded verb in (305).

The agreement marking on the embedded verb cannot be treated as a lexicalized or frozen form. This is because it co-varies with the number feature on the antecedent of *ta(a)n*. For instance, when the matrix subject is plural, as in (307), the embedded agreement has to be marked plural as well:

(307) Pasaŋ-gaɭ$_i$ [$_{CP}$ taaŋ-gaɭ$_{\{i,*j\}}$ ʤej-pp-oom/*aangaɭ-ŭnnŭ]
boy-PL.NOM [ANAPH-PL.NOM$_i$ win-FUT-1PL/*3MPL-COMP]
so-nn-aaŋ-gaɭ.
say-PST-3M-PL
"The boys said [$_{CP}$ that they$_{\{i,*j\}}$ would win]"

To sum up, the 1st-person agreement under *ta(a)n* in cannot be treated as a type of morphophonological default, or as a frozen or idiomatized form.

10.2.2 1st-person agreeement: not quoted

Quotations may emulate shifted-indexicality effects because they "form a closed domain with respect to syntactic and semantic operators" (p. 81 Anand 2006). Given this, we must consider whether the embedded CP in (305) is just a direct quote.

We can rule out a partial quotative analysis for (305) which proposes that the agreement morpheme alone is quoted. A partially quoted string – a nonsense string like "*Pfffft!*" for instance – should be possible under any type of predicate. Thus, if the embedded CP in (305) did involve a partial quotative, we wouldn't expect the kind of systematic syntactico-semantic predictability with respect to when it obtains, that we get. The other option is that the entire *ta(a)n*-clause is quoted, in such sentences. One way to test whether a clause is quoted might be to see whether it is headed by a complementizer. Such a diagnostic works quite well in English, a language in which quoted elements may never be embedded under a complementizer. But Tamil embeds quotatives and non-quotatives alike under the same complementizer – so this wouldn't work as a test.

There is already one very good argument against a quotative analysis, however. This is the fact that the subject is the anaphor *ta(a)n* – which, precisely by virtue of being anaphoric – would not be expected to be the matrix subject of a quotational string. The grammatical opacity of quotatives also yields some useful diagnostics for testing their existence. For instance, an NPI inside a quotative may not be licensed by an operator outside the quote; similarly, movement is not possible across a quotational boundary. Consider the sentences below, both of which involve special 1st-person agreement under *ta(a)n*:

(308) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ orŭ tappu-m se-ndʒ-een-nnŭ] ottukka-læ.
Raman[NOM] [ANAPH[NOM] one mistake=NPI make-PST-1SG-COMP] admit-NEG
"Raman$_i$ didn't admit [$_{CP}$ that he$_{\{i,*j\}}$ made any mistake.]"

(309) Raman$_i$ jaaræ$_x$ [$_{CP}$ taan$_{\{i,*j\}}$ t$_x$ aḍi-tt-een-nnŭ] so-nn-aan?
Raman[NOM] whom [ANAPH[NOM] t hit-PST-1SG-THAT] say-PST-3MSG
"Who(m)$_x$ did Raman$_i$ say [$_{CP}$ that he$_{\{i,*j\}}$ hit t$_x$]?"

In (308), an NPI within the CP is associated with a NEG operator outside the CP; in (309), the *wh*-element moves from within the CP to a landing-site in the matrix clause. If the resulting sentences are grammatical, we have conclusive evidence that the embedded CP is *not* a quote. As we can see, both (308) and (309) are fully grammatical, showing that the embedded clause is not quoted.

As a control case, we can test what would happen if the embedded CPs *were* quoted:

(310) * Raman$_i$ [$_{CP}$ naan$_i$ orŭ tappu=m se-nʊ̀ʒ-een-
Raman[NOM] [I[NOM] one mistake=even make-PST-NEG-
nnŭ] ottukka-læ.
COMP] admit-NEG
"*Raman didn't say ['I made any mistake.']"

(311) * Raman$_i$ jaar-æ$_x$ [$_{CP}$ naan$_i$ t_x aɖi-tt-een-nnŭ] so-nn-aan?
Raman[NOM] who-ACC [I[NOM] t hit-PST-1SG-COMP] say-PST-3MSG
"*Who(m)$_x$ did Raman$_i$ say [$_{CP}$ 'I hit t$_x$]'?"

The sentences above vary minimally from (308) and (309): the anaphoric subject *ta(a)n* in each has been replaced by the deictic 1SG pronoun *naan* ("I"). These sentences are fully ungrammatical under the quoted reading where *naan* refers to *Raman*, just as expected. Furthermore, when grammatical operations across the CP boundary are not performed, the sentences become licit once again. In (312), the NPI in the embedded CP is licensed by a NEG operator within that CP; in (313), the *wh*-object remains *in situ* within the embedded clause – both are fully grammatical:

(312) Raman$_i$ [$_{CP}$ naan$_{\{i,*j\}}$ orŭ tappu=m sejja-læ-
Raman[NOM] [I one mistake=even make-PST-3MSG-
nnŭ] ottukko-ɳɖ-aan.
COMP] admit.PST-3MSG
"*Raman$_i$ admitted ["I$_{\{i,*j\}}$ didn't make any mistake."]"

(313) Raman$_i$ [$_{CP}$ naan$_{\{i,*j\}}$ jaar-æ aɖi-tt-een-nnŭŭ] kee-
Raman[NOM] [I[NOM] who-ACC hit-PST-1SG-COMP] ask-
ʈʈ-aan.
PST-3MSG
"Raman$_i$ asked [$_{CP}$ 'Whom$_i$ did I$_{\{i,*j\}}$ hit t$_i$?'"]

To sum up, therefore, the 1st-person agreement that obtains under *ta(a)n* is not quoted, nor does it occur in a quoted clause.

10.2.3 1st-person agreement: not a historical accident

One might claim that the 1SG agreement on the embedded verb in structures like (305) is a quirk of Tamil – an accident, perhaps, due to other factors internal to the language and its historical development.[2] However,

[2]This would be similar to the kinds of proposals that have been adduced to "explain" the relative rareness of nominative-marked anaphors crosslinguistically, which has been claimed to represent a typological "gap". In addition to Tamil and Donna Sɔ

this doesn't seem to be a very viable course of action either. Analogous structures have been reported in a small sub-class of African languages as well. Curnow (2002), citing Culy (1994), presents data from Donna Sɔ, a language of the Dogon sub-family of the Niger-Congo family, which has a structure that resembles that in Tamil (305) to a startling degree. I reproduce Curnow's example in (314) below (formatting mine):

(314) Oumar [$_{CP}$ inyemɛ jɛmbɔ paza bolum] miñ tagi.
Oumar [ANAPH[SBJ] sack.DEF drop left.1SG] 1SG.OBJ informed
"Oumar$_i$ told me [$_{CP}$ that he$_{\{i,*j\}}$ had left without the sack.]"

Curnow further argues that the embedded complement in such sentences does not constitute a direct speech report and states that such 1SG agreement, while quite rare, is evidenced in half a dozen other African languages. Further research must be undertaken to ascertain the full range of empirical properties evidenced by such structures – but I take the existence of sentences like (314) as evidence that such agreement is not unique to Tamil or even to the Dravidian language family as a whole. It is indeed possible that further research on this topic will uncover other languages with similar agreement paradigms.

The discussion above has shown conclusively what the 1st-person agreement under *ta(a)n* in structures like (305) is *not*. In the following sections, we will see what it actually is – namely, an instantiation of indexical shift, the phenomenon where a person feature like 1st or 2nd is evaluated, not against the utterance context, as is standard, but against a linguistic context that is introduced by a speech predicate.

10.3 A brief history of indexical shift

Consider the following sentence from the Indo-Iranian language Zazaki (Anand and Nevins 2004):

(315) Hɛsen-i$_j$ (mɨ$_k$-ra) va [$_{CP}$ kɛ ɛz$_{j/k}$ dɛwletia].
Hesen-OBL I-OBL.TO said that I rich.be.PRS
"Hesen said (to me) [$_{CP}$ that {I am, Hesen is} rich.]"

In (315), the indexical pronoun ɛ*z* (I) can refer to the speaker of the utterance context, just like in English; but it can crucially also refer to the speaker of the speech event introduced by the matrix verb – namely,

discussed immediately below, nominative anaphors are however found in e.g. North Sami (Nickel 1994) and Japanese (Kuno 1987).

Hɛsen. ɛ*z* is labelled a "shifted indexical" because of this property. The property itself is termed "indexical shift".

The discussion thus far has conclusively shown that the unexpected 1st-person agreement under *ta(a)n*, in structures like (305), cannot be dismissed as a type of default agreement, a frozen or idiomatized morpheme, as a quoted string, or as a Dravidian-specific historical accident. However, it does bear a strong resemblance to the shifted indexical in Zazaki (315) above. The agreement in (305), just like the indexical in (315), is 1st-person, but it refers to the speaker of the matrix speech context rather than to the speaker of the utterance context. Second, indexical shift crosslinguistically has been reported to obtain predominantly under the scope of speech predicates (Anand 2006, Schlenker 2003b, von Stechow 2002, Speas 1999, Shklovsky and Sudo To Appear) – the minimal pair in (304)-(305) suggests that 1st-person agreement under *ta(a)n* similarly obtains only under the scope of a speech predicate.

In the sections below, I will present an overview of the literature on indexical shift and show that the 1st-person agreement in Tamil (305) is an instantiation of this phenomenon.

10.3.1 What are indexicals anyway?

Consider the following scenario:

(316) **Scenario 1:**

a. Max: "I really hate pineapples on my pizza."
   $[\![I]\!]$ = Max.
b. Moritz: "So do I!"
   $[\![I]\!]$ = Moritz.

Max' proclamation in (316a) and Moritz' reply in (316b) are uttered in specific *contexts* where certain key information about the utterance, like who is speaking, and when and where (s)he is speaking are encoded. It is thus clear that Moritz' response is uttered in a different context from that of Max: the speaker of (316b) is Moritz, whereas that of (316a) is Max, so the two contexts are already distinguished in terms of *who* is speaking. Also, almost certainly (modulo simultaneity of utterance, a very unlikely scenario), (316b) is probably uttered a short time after Max's, thus the *time* of the utterance in (316b) is different from that of (316a) as well. Other parameters like *where* the utterances were made (probably different if, for instance, Max and Moritz are talking on the phone) might differ as well, distinguishing the two contexts even further.

Now compare a slightly different scenario:

(317) **Scenario 2:**

a. Jack: "Max$_i$ really hates pineapples on his$_i$ pizza."
b. Jill: "Yes, he$_i$ really does!"
$[\![he]\!]$ = Max.

Just like in (316), each of the utterances in (317) is associated with a distinct context. But there's a difference in the nature of pro-form used in each. The pronoun *I* in (316) necessarily changes its reference from one utterance to the next, as we have seen. In contrast, the deictic pronoun *he* in (317) doesn't have to denote a distinct entity in every new context: in fact, in this scenario, it uniformly denotes Max in both contexts of utterance. Pronouns like *I*, *you*, *here*, and *now*, called "indexicals", form a natural class because of their context-sensitivity; they stand in contrast to other expressions like *he, the next day*, and *there* which, not being obligated to change their reference from one context to the next, aren't context-sensitive in the same way.

In his seminal paper Kaplan (1989), the philosopher David Kaplan proposed that indexicals are also special because they cannot be manipulated by intensional operators. This was motivated by differences such as the following:

(318) I am here now.
(319) Sandhya is in Berkeley at 7.22pm.

(318) and (319) are both true. However, (319) will be falsified a minute from now, or if I happen to travel outside of Berkeley at 7.22pm tomorrow or if, indeed, I am magically transformed into someone else, as in a dream world. In contrast, (318) cannot be falsified: it is logically true across all possible worlds, locations, and times. In other words, the truth-value of (319) can be influenced by intensional operators that manipulate "circumstances of evaluation", to borrow Kaplan's term, but that of (318) cannot. Furthermore, (318) alone contains indexicals referring to the speaker, time, and location of the utterance context, suggesting that it is this property that allows it to be universally true.

Kaplan's thesis about indexicals, motivated by these observations, is reproduced below:[3]

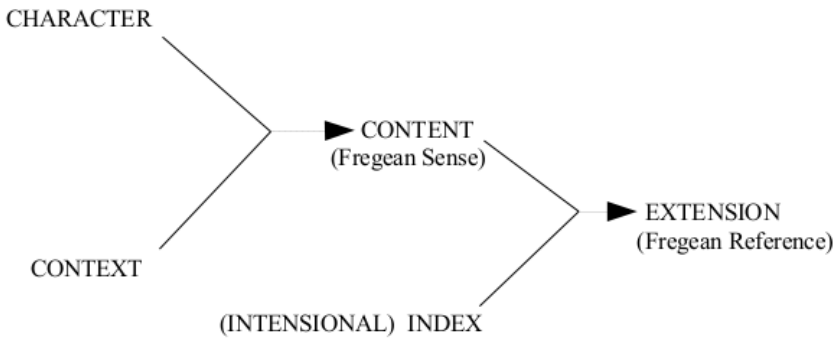
[3]Kaplan treated demonstratives as a type of indexical as well – a classification that has since been questioned in the literature. The difference between demonstratives and the types of indexical expresssions we have been considering here, like *I* and *here*, is that the former but not the latter require a symbolic or actual demonstration to complete their denotation. As such, these latter types of indexicals were labelled as "pure indexicals" by Kaplan.

(320) Quoted from Kaplan (1989, p. 492):

> Principle I: The referent of a pure indexical depends on the context, and the referent of a demonstrative depends on the associated demonstration.
>
> Principle II: Indexicals pure and demonstrative alike, are directly referential.[4]

But how can an element have its reference mediated by contextual information and *simultaneously* be directly referential? The only way out of the quandary, Kaplan argues, is to claim that contexts and circumstances of evaluation (or intensions) represent two ontologically distinct components of meaning. The Kaplanian model illustrated in (321) below (modified from Anand 2006) thus involves two sorts of operators (in contrast to the Fregean one which involves just a single type of intensional operator): those which operate on the "character" of an expression to yield its "content" (corresponding to Fregean "sense" or *Sinn*) and those that manipulate the content of an expression to yield its extension (corresponding to Fregean "meaning" or *Bedeutung*). Denotations of linguistic objects in Kaplan's model are thus "doubly-indexed":

(321) Kaplanian model of meaning, adapted from Anand (2006):



> A *character* is a function from contexts to contents.
> A *content* is a function from circumstances of evaluation (or intensional indices) to extensions.

[4]Kaplan defined the concept of direct referentiality as follows (Kaplan 1989, p. 492): "I intend to use the term 'directly referential' for an expression whose referent, once determined, is taken as fixed for all possible circumstances [evaluations according to intensional parameters], i.e. it is taken as *being* the propositional content."

(322) Denotations of "pure" indexicals in the doubly-indexed Kaplanian model, for c = context, and i = (intensional) index.[5]

$[\![I]\!]^{c,i} = \lambda c \lambda i.Author(c)$
$[\![you]\!]^{c,i} = \lambda c \lambda i.Addressee(c)$
$[\![now]\!]^{c,i} = \lambda c \lambda i.Time(c)$
$[\![here]\!]^{c,i} = \lambda c \lambda i.Location(c)$

10.3.2 Enter Kaplanian "monsters"

Further empirical support for the doubly-indexed model came from the observation that contextual operators behave differently from intensional ones. Kaplan observed that, while intensional operators can freely manipulate intensional expressions or "contents" to yield their extensional equivalents, contextual operators do not seem analogously capable of controlling indexicals (or "characters") in their scope to yield their corresponding intensions (or "contents"). Thus, the sentence in (323) does *not* mean that for some contexts, the individual authors of these contexts like pizza; it can only mean that the author of the *current* context of utterance (namely me, Sandhya) doesn't like pizza.

(323) [[I don't like pizza]]
$\neq \lambda c.\neg Like(Author(c), pizza)$, for c = any context.

By the same reasoning, a sentence like (324) is utterly nonsensical:

(324) In some (utterance) contexts, it is true that I like pizza.

Based on such evidence Kaplan (1989, 510-11) concludes that:

> "no operator can control the character of the indexicals within their scope, because they will simply leap out of its scope to the front of the operator ... Operators like 'In some contexts it is true that' which attempt to meddle with characters, I call *monsters*. I claim that none can exist in English (without sneaking in a quotation device)."

[5] For reasons of consistency, I have represented indexicals as *constant functions* from intensions (content) to extensions. But I could just as well have left the intensional operator out of these denotations altogether and represented the denotations of indexicals as functions from contexts directly to extensions, in accordance with Kaplan's direct referentiality thesis for these expressions. It is unclear at this juncture what substantive differences, if any, there are between the two formulations: the reader is referred to Kaplan (1989) for some speculation on this point.

Kaplan's ban against these "monsters" seems valid enough for the English examples he discusses. But other languages seem to be harder to accommodate.

Schlenker (2003b) discusses two types of counter-evidence against monsters. The first is the class of logophors – 3rd-person pro-forms with an obligatorily 1st-person (or *de se*) semantics, such as those found in many African languages (Clements 1975) and in long-distance binding paradigms in many languages (Sells 1987, Reinhart and Reuland 1993). Schlenker analyzes these forms as the indirect discourse analogs of "I" and treats them as obligatorily shifted indexicals.

The second kind of counter-evidence involves indexical pronouns that are shifted to refer to a context other than the utterance context. We have already seen an example of this in the Zazaki sentence in (315) – Schlenker presents other examples of such "monsters" from Amharic (Schlenker 2003b, p. 68, ex. 53) (formatting mine):

(325) ǰon [$_{CP}$ jəgna nə-ññ] yɨl.all.
John hero be.PF-1SO 3M.say-AUX.3M
"John$_i$ says [$_{CP}$ that he$_{\{i,*j\}}$ is a hero]"
Lit: "John$_i$ says that I$_i$ am a hero", $[\![I]\!] \rightarrow$ John.

Unlike English *I*, the 1st-person pronoun in the embedded CP in (325) can refer to the speaker of the intensional speech predicate in the matrix clause, rather than to the speaker of the utterance context. Schlenker demonstrates convincingly that the embedded CP in (325) represents a form of indirect speech (i.e. is not quoted). He also shows that the indexical obtains this "shifted" denotation only under the scope of intensional predicates like *nə-ññ* (SAY) in (325); in all other contexts, it denotes the speaker of the utterance context, just like in English. Schlenker thus proposes that, at least in some languages, certain types of intensional predicate may introduce contextual operators which manipulate indexicals in their scope, in apparent contradiction to Kaplan's thesis. In other words, he claims that Kaplanian monsters are attested in natural language.

Since the pioneering work of Schlenker (1999) and Schlenker (2003b), other instances indexical shift have come to light. Some of these have been novel data showing indexical shift from hitherto un(der)-studied languages; others have been phenomena that have been known for some time but were re-analyzed as indexical shift as the literature on this subject became better known. The empirical evidence pertaining to this phenomenon is still relatively sparse – nevertheless, some pervasive and systematic patterns emerge even from the relatively small sample of lan-

guages on hand which have been observed to exhibit indexical shift. Indexical shift for person has been reported for Navajo (see Speas 1999, for a description of this phenomenon, though it is not formally analyzed as indexical shift), the Turkic language Uyghur (Shklovsky and Sudo To Appear), and the Athapaskan language Slave (Rice 1986), among others.

However, the person dimension is not the only one that is capable of being shifted: evidence of indexical shift along the modal and temporal domains have been attested as well. Speas (1999) notes for Navajo that indexical temporal adverbials like *tomorrow* may be evaluated with respect to the context of the speech report, which is nothing other than indexical shift with respect to the time parameter of the context. Similarly, Giorgi (2010) claims that tenses in languages like Romanian, Russian (for further discussion of such phenomena in Russian, see also Schlenker 2003a;b; To appear), and Japanese, may be interpreted relative to the time of a matrix speech report rather than against the utterance time, which may also be interpreted as temporal indexical shift. The world parameter is capable of being shifted as well: Schlenker (2003b) claims that the subjunctive mood that characterizes verbs that are embedded under certain speech and thought predicates in German, termed *Konjunktiv I*, is nothing but a modal indexical that has been shifted. Contextual shifting of mood indexicals has also been observed in certain sign languages: see Quer (2005) for discussion of this phenomenon in Catalan Sign Language.

A different sort of parametric variation has to do with whether shifting is obligatory, optional or impossible. English appears to be a language that doesn't allow indexical shift, at least along the person domain. Zazaki, on the other hand, allows optional indexical shift along *all* parameters of evaluation:

(326) 2ND-PERSON INDEXICAL SHIFT:
Hɛseni$_j$ (Ali$_k$-ra) va [$_{CP}$ kɛ **t**i$_{j,k}$ dɛwletia].
Hesen.OBL (Ali-OBL.TO) said that you rich.be-PRS
"Hesen said [$_{CP}$ that {Ali is, you are} rich.]"

(327) TEMPORAL INDEXICAL SHIFT:
Waxto kɛ ma Diyarbekir-de bime, Hɛsen mi-ra va kɛ
when that we Diyarbekir-at were, Hesen.OBL me-at said that
o **ita** ame dina.
he here came world
"When we were in Diyarbekir, Hesen told me he was born here, in Diyarbekir".

(328) LOCATIVE INDEXICAL SHIFT:

Hefte nayeraraver, Hɛsen mi-ra va kɛ o **vizeri** Rojda
week ago, Hɛsen.OBL me-at said that he yesterday Rojda
paci kɛrd.
kiss did
"A week ago, Hɛsen told me that he kissed Rojda 8 days ago, #yesterday."

At the other end of the typological spectrum, Slave has been reported to be a language that obligatorily shifts some person-indexicals under a subset of intensional predicates.

10.3.3 Some prior analyses of indexical shift

Based on such evidence, Schlenker argues that indexical pronouns may be grouped into three typological sub-classes: "rigid" indexicals like *I* in English which are specified to never shift with respect to the utterance context, those like Amharic *jɔgna* which are specified to optionally shift with respect to the utterance context, and logophors/anaphors which shift obligatorily relative to the context of utterance. These three sub-classes of indexical have the following denotations:[6]

Rigid indexical:
$[\![I_{eng}]\!]^{c,g} = [\![x_i + Author * (x_i)]\!]^{c,g} = \mathrm{s(x}_i) \wedge \delta(\mathrm{g(x}_i) \text{ is Author(c*)}).$

Optionally shifting indexical:
$[\![I_{amh}]\!]^{c,g} = [\![x_i + Author(x_i, c_j)]\!]^{c,g} = \mathrm{s(x}_i) \wedge \delta(\mathrm{g(x}_i) \text{ is Author(c}_j)).$

Obligatorily shifting indexical:
$[\![I_{log}]\!]^{c,g} = [\![x_i + Author(x_i, c_j)]\!]^{c,g} = \mathrm{s(x}_i) \wedge \delta(\forall\ \mathrm{c}_j.(\mathrm{c}_j \neq \mathrm{c*}) \rightarrow \mathrm{g(x}_i)$ is Author(c_j)).

Since Schlenker (1999)'s dissertation which first brought this issue to light, other instances of monsters and theoretical treatments thereof have emerged in the literature. Anand (2006) is one such analysis which, while sharing some of the core intuitions of Schlenker, differs in the details of its formal implementation of indexical shift. Shifted indexicality, for Anand, is the result of *context shifting* due to context-overwriting by

[6]Some terminological clarifications: $c*$ is a free variable that denotes the context of utterance; analogously, *Author*$*$ denotes the unique speaker/author of the utterance context. A pro-form is represented as a free variable whose reference is constrained by semantic presuppositions (Heim and Kratzer 1998); an expression followed by δ is construed to be presuppositional.

an operator introduced by an intensional predicate. Schlenker's pronoun-centric model derives the possibility of indexical shift as a function of the denotation of the indexicals themselves, as we have just seen. However, Anand proposes that the difference between the shifted and unshifted readings is not the denotation of the indexical – which, he proposes, remains constant at $\lambda c.Author(c)$ – but has to do with the value of the context that the AUTHOR function ranges over. As discussed above, languages have been observed to differ, not just with respect to whether indexical-shift is optional, obligatory or never attested, but also relative to which type of indexical may shift. For some languages, only 1st or 2nd person indexicals may shift, for others only temporal or modal ones may do so, and for yet others, indexicals along all intensional domains may be shifted. Envisioning, as Schlenker and Kaplan himself do, that a context is an ontological tuple consisting of variables pertaining to different domains of evaluation (e.g. $< Speaker, Addressee, Time, World >$), Anand proposes that this variation be derived in terms of parametric choices pertaining to which contextual coordinates may be overwritten.

To see a concrete implementation of Anand's proposal, consider again the case from Zazaki (315), repeated as (329) below:

(329) Hɛseni$_j$ (mi$_k$-ra) va kɛ ɛz$_{j/k}$ dɛwletia.
Hesen.OBL I.OBL-TO said that I rich.be-PRS
"Hesen said that {I am, Hesen is} rich."

Zazaki is a language that, in principle, allows every contextual coordinate to be shifted (recall the sentential paradigms given in (326)-(328)). In Anand's system, this means that the contextual operator in Zazaki is associated with a universal quantifier which overwrites every coordinate of the utterance context with a corresponding one in the intensional context, and binds it. Formally:

(330) $[\![OP_{\forall}\alpha]\!]^{c,i} = [\![\alpha]\!]^{j,i}$
where j = $<Auth(i), Addr(i), Time(i), World(i)>$.

The unshifted "English-like" reading of the indexical in (329) is derived as follows:

> The matrix speech predicate does not introduce a contextual operator. Thus, there is no context overwriting; the indexical *I* is interpreted against the utterance context (perhaps introduced at root C) and is thus mapped onto the utterance speaker = me, Sandhya (in this utterance context). Formally:

[Hesen said $[_{CP}$ I am rich]]^{c,i,g} = 1 iff
[Hesen said $[_{CP}$ Author(c) am rich]]^{c,i,g}
= [Hesen said $[_{CP}$ $I_{Sandhya}$ am rich]]^{c,i,g}

The shifted reading in (329), on the other hand, has the following derivation:

> The matrix speech predicate has a different denotation in this case; specifically, it introduces the contextual operator defined in (330). This has the effect of overwriting every coordinate of the utterance context in its scope with a corresponding coordinate of the intensional context. Formally:

[Hesen said $[_{CP}$ $OP_{\forall}$ I am rich]]^{c,i,g}
= [Hesen said $[_{CP}$ I am rich]]$^{i',i',g}$ = 1 iff
[Hesen said $[_{CP}$ Author(i') am rich]]$^{i',i',g}$
= [Hesen said $[_{CP}$ I_{Hesen} am rich]]$^{i',i',g}$

Despite their differences, Anand's and Schlenker's analyses both share the common view that indexicals may be manipulated by contextual operators, contrary to what Kaplan's prohibition against monsters would lead us to expect. Other analyses, like that of von Stechow (2002), are more sympathetic to Kaplan's claim, however. The debate, of course, is not about whether certain indexicals may be bound by linguistic operators along the personal, temporal, modal and spatial domains – that this is possible is conclusively proven by sentences like those in Amharic (325) and Zazaki (329) above. Rather, it centers around whether these operators count as *contextual* in the Kaplanian sense, or whether they are merely rich intensional operators that quantify over these same dimensions.

For instance, von Stechow reconstructs a context in the Kaplanian sense as follows: "A [Kaplanian] context is a triple $< x, t, w >$ consisting of a person x, a world w and a time t. For any context c, x_c is the first component of c, t_c is the second component and w_c is the third component of c." He proposes that a contextual variable of this kind (i.e. an ontologically complex *unit*) is never attested in natural language, even in sentences that show effects of indexical shift. Rather, he claims, what we get in such structures are three separate variables x, t and w which may co-vary independently of each other – a Lewis (1979)-style series of enriched intensional indices, in other words. Indeed, given the definition of Kaplanian context above, it is clear that this entity (being essentially a complex primitive) may never be *partially* shifted for some contextual

coordinates but not others as, for instance, Anand suggests; nor can it be claimed, as Schlenker does, that indexicals may be inherently specified to shift along some contextual dimensions but not others. Von Stechow thus concludes that, while "[w]e may say that attitudes quantify over contexts ... they never shift the context of utterance" – which thus still special and different, just as Kaplan envisioned it.

von Stechow (2002) derives the effects of indexical-shift in a different way. He proposes that attitude predicates are "universal quantifiers which take an egocentric proposition [a sentence quantified by a series of intensional indices, as discussed above] as their complement". Indexical shift, in this model, obtains when an indexical pronoun is bound by such an attitude-verbal quantifier.[7] An important component of von Stechow's analysis is the idea that the features of a variable are deleted at LF when they are bound: such feature-deletion is, in fact, a precondition on variable binding in von Stechow's system. The difference between languages that manifest (person) indexical shift, like Amharic or Tamil, and those that don't, like English, is a function of whether the features on the indexical may be deleted by the c-commanding verbal quantifier or not (respectively).

For languages like English, von Stechow proposes that feature-deletion obtains just in case the features of the verbal quantifier match those of the indexical in value – as captured by the following rule (von Stechow 2002, 23):

> **Feature Deletion under Attitudes (English):**
> A verb of attitude deletes the features of the variables it binds under agreement with its checkees. (A checkee a agrees with a checker b if a = *b.)

This ensures that an attitude predicate that is marked 3rd-person will never delete the features of an (1st/2nd) indexical pronoun in its scope and will, by extension, never bind it. As such, an indexical pronoun in English will never "shift" its reference to a 3rd-person attitude-holder that c-commands it, just as desired.

In languages that do evidence indexical shift, such as Amharic, on the other hand, feature-deletion is governed by different rules. In such languages, the "verbal quantifiers [may] delete ([at] LF) the feature **1st** of the person variable they bind, regardless of what their person checkee is" (von Stechow 2002, 25). Thus, in a sentence like: "John said$_{3sg}$ [$_{CP}$ that I

[7] This is an extension of his analysis for bound-variable indexicals, – like "my" in the sentence: "Only I finished drinking **my** beer" – which, perhaps more patently, involve the variable-binding of an indexical by a c-commanding quantifier.

am hungry]" in Amharic, the features of the embedded indexical pronoun "I" could be deleted and then bound by the attitude predicate checkee even though this has mismatched 3SG features. Crucially, this feature deletion will happen at LF, so for the purposes of morphophonology, the pronoun will still surface as "I". The formal binding operations are reproduced below (von Stechow 2002):[8]

(331) Target sentence: "John says I am a hero." (Amharic)
  a. John$^{\text{3rd}}$ says$^{\text{3rd*}}$ λ<x$^{\text{1st}}$ . . . >. . . x$^{\text{1st}}$ am$^{\text{*1st}}$ a hero.
  "John says 'I am a hero.'" (**Shifted reading**)
  b. John$^{\text{3rd}}$ says$^{\text{3rd*}}$ λ<x$^{\text{1st}}$ . . . >. . . y^{1} am$^{\text{*1st}}$ a hero.
  "John says that I am a hero." (**Unshifted reading)**

The basic debate in the literature that is centered around the question of whether monsters exist in natural language or not, seems to me to be based (at least in part) on a fundamental terminological ambiguity about how a Kaplanian context is to be understood. If it is to be understood as referring to the utterance situation, it seems that Kaplan would be right in claiming that monsters don't exist, as von Stechow also points out. If, on the other hand, a context is interpreted as an enriched sequence of personal, temporal, and modal operators introduced by a proper subset of intensional predicates, then evidence from languages like Amharic, Navajo, and Zazaki seem to clearly prove him wrong. I will thus not take a stance with respect to this question here. My use of the terms "shifted indexical" and "monster" should thus be taken to be purely descriptive, referring to an indexical whose reference has been shifted by a minimally c-commanding intensional operator. Similarly, I will use the term "intensional context" to refer to the types of intensional variables along personal, temporal, modal, and spatial coordinates that are introduced by a proper subset of intensional predicates – and which may induce indexical shift in their scope. The intensional context (essentially a tuple of mutually independent intensional variables) is intended to be ontologically distinct from the utterance context (which is a complex primitive in the Kaplanian sense) in the manner discussed by von Stechow (2002) above.

In the course of the next few chapters, I will present evidence, based on the types of Tamil data discussed earlier, to show that the complements of certain speech predicates *syntactically* encode a projection which hosts the author of the embedded utterance. At the same time, this evidence will also drive home the point that the utterance context, under-

[8] A quick note on the notation adopted here. For a privative feature F, [F] = the interpretable counterpart of F; [*F] = the uninterpretable/inherited feature F.

stood in the Kaplanian sense, *is* special and has a different relationship with propositional content than does the kind of intensional context introduced by a speech predicate.

Chapter 11

Tamil "monstrous" agreement under the magnifier

Consider again the Tamil sentence in (305), repeated as (332) below:

(332) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ ʤej-pp-een-nnŭ] so-nn-aan.
Raman[NOM] ANAPH[NOM]$_i$ win-FUT-1SG-COMP say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that he$_{\{i,*j\}}$ would win]"

Given the discussion of indexical shift, the similarities between the Tamil paradigm in (332) and the illustrations of indexical shift given above are quite clear. Both occur predominantly under the scope of speech predicates and involve indexical elements that seem to be evaluated against an intensional context rather than relative to the actual context of utterance. Based on the larger discussion of indexical shift above and the observed parallels between the anomalous 1SG agreement in (332) and other bona fide cases of indexical shift, I will now assume without further argument that the 1SG agreement marking on the embedded verb in structures like (332) is an instantiation of indexical shift for 1st-person and refer to it, henceforth, as "monstrous agreement".

At the same time, the Tamil monstrous agreement facts pose some non-trivial empirical challenges which do not characterize the more "classic" cases of indexical shift discussed in the previous chapter. A major source of the problems is that indexical shift in Tamil is marked on person agreement, rather than on a full-on pronoun. Since agreement features are not inherent to the categorial head on which they are realized, these must ultimately have another source. Tamil is a subject agreement language, thus it might seem reasonable to assume that monstrous agreement on the embedded verb in (332) is triggered by its clausemate anaphoric subject *ta(a)n*. However, in our discussion of long-distance binding in Part I of this dissertation, we independently argued that

ta(a)n is fully featurally defective (i.e. that it is a "minimal" pronoun in the sense of Kratzer (2009)), proposing that it has a set of unvalued ϕ-features and an unvalued DEP-feature. But if this is the case, where does the 1st-person feature on the embedded verb in (332) come from? In Part I, we proposed that the agreement-marking under subject *ta(a)n* is triggered by the silent operator in [Spec, PerspP], which bears the ϕ-features of the DP that will end up serving as the antecedent of *ta(a)n*. The question is whether it is possible to extend this analysis to the monstrous agreement cases, where the features on embedded T differ from those on the anaphoric antecedent.

Potentially complicating matters even further is the fact that, in the dialect being discussed here, monstrous agreement under propositional *soll* (SAY) is optional. I.e. the sentence in (332) above has a minimally varying counterpart in (333), with the agreement under *ta(a)n* faithfully reflecting the ϕ-features of the anaphoric antecedent:

(333) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ ɖʒej-pp-**aan**-nnŭ]
Raman[NOM] ANAPH[NOM]$_i$ win-FUT-3MSG-COMP
so-nn-aan.
say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that he$_{\{i,*j\}}$ would win]"

How do we account for this optionality, assuming that the "same" verb *soll* is being used in both cases?

In this chapter, we will take a closer look at the Tamil patterns in order to investigate this and other questions. We will tease apart the binding relationship between *ta(a)n* and its antecedent from that of indexical shift in the embedded clause – as a result, we will show that the two phenomena are not *directly* related to one another, contra the proposal in Schlenker (2003b, and subsequent). At the same time, we will see that the 1st-person agreement always tracks the antecedent of *ta(a)n*, showing that that there is a one-way dependency between the presence of *ta(a)n* and monstrous agreement. Based on a close examination of the Tamil data, I will propose that the shifted indexical is actually the null pronominal operator in [Spec, PersP] – the DP that enters into an Agree relationship feature with *ta(a)n* in embedded subject position and with the T head under *ta(a)n*.

11.1 Long-distance anaphora vs. indexical shift

Let us look again at the minimal pair that we started this discussion with:

(334) Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ poo̱ṯ̱i-læ ʤejkkapoo-r-**aa**ḷ̱-ŭnnŭ]
Maya[NOM] ANAPH[NOM] contest-LOC win-PRS-3FSG-COMP
namb-in-aaḷ̱.
believe-PST-3FSG
"Maya$_i$ believed [$_{CP}$ that she$_{\{i,*j\}}$ would win the contest.]"

(335) Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ poo̱ṯ̱i-læ ʤejkkapoo-r-**een**-nnŭ]
Maya[NOM] ANAPH[NOM] contest-LOC win-PRS-1SG-COMP
so-nn-aaḷ̱.
say-PST-3FSG
"Maya$_i$ believed [$_{CP}$ that she$_{\{i,*j\}}$ would win the contest.]"

The sentences in (334) and (335) both involve long-distance binding of *ta(a)n* in the subject position of the embedded clause. However, the 1st-person agreement under *ta(a)n*, which I have labelled "monstrous agreement", only obtains under the scope of a matrix speech predicate, as in (335). This set of facts, taken together, leads us to a significant conclusion, namely that, contra the claims of Schlenker (2003b, and subsequent), indexical shift and long-distance binding are not one and the same phenomenon. By extension, anaphors (even those with an obligatory *de se* semantics, like *ta(a)n*) are not the same as obligatorily shifted indexicals. After all, in the Tamil sentences above, *ta(a)n* itself maintains the same overt form; what changes is the form of the agreement under *ta(a)n*, in each case.

This much seems fairly straightforward. But to better understand the nature of the difference, let us review some of the other properties that distinguish long-distance anaphora and indexical shift in Tamil.

11.1.1 Differences in licensing environment

As discussed in detail in Part I, the distribution of *ta(a)n* is quite varied. It may be bound in the propositional complement of all classes of matrix predicate, and may also legitimately occur in spatio-temporal adjunct PPs, DPs, and CPs; in its so-called "logophoric" use, it may even occur in a root clause:

(336) PROPOSITIONAL COMPLEMENT:

Raman$_i$ [$_{CP}$ Seetha tann-æ$_{\{i,*j\}}$ kaadali-kkir-aaḷ-ŭnnŭ]
Raman[NOM] Seetha[NOM] ANAPH-ACC] love-PRS-3FSG-COMP
kaɳḍupiḍi-tt-aan.
find.out-PST-3MSG
"Raman$_i$ found out [$_{CP}$ that Seetha$_j$ loved him$_{\{i,*j\}}$.]"

(337) SPATIAL PP ADJUNCT:

Raman$_i$ tan-akkŭ$_{\{i,*j\}}$ meelæ orŭ plane-æ paar-tt-aan.
Raman.NOM ANAPH-DAT above a plane-ACC see-PST-3MSG
"Raman$_i$ saw a plane above himself$_{\{i,*j\}}$."

(338) TEMPORAL CP ADJUNCT:

Raman$_i$ [$_{CP}$ Seetha$_j$ tann-æ$_{\{i,*j\}}$ kiḷḷ-in-æ poɻŭdŭ]
Raman Seetha ANAPH-ACC pinch-PST-REL time
sattamaagæ ka-tt-in-aan.
loudly yell-PST-3MSG
"Raman$_i$ yelled loudly [$_{CP}$ when Seetha$_j$ pinched him$_{\{i,*j\}}$]."

(339) "LOGOPHORIC" USE:

Raman-ŭkkŭ$_i$ oɳɳum-ee purija-læ. taan$_{\{i,*j\}}$
Raman-DAT nothing[ACC]-EMPH understand-NEG. ANAPH-NOM
maṯṯum een ippaḍi ellaam kašṭappaḍa-ɳum?
alone why like.this all suffer-must?
"Raman$_i$ didn't understand at all. Why should he$_{\{i,*j\}}$ alone suffer like this?

This distribution matches that of long-distance bound anaphors in other languages. Crosslinguistically, anaphors have been reported to occur under the scope of propositional attitude predicates like BELIEVE and THINK, emotive predicates like FEAR and LOATHE, factive predicates like KNOW and REGRET, perception predicates like SEE and HEAR, locational motion verbs like COME and GO, propositional speech predicates like SAY, as well as in DP, PP, and clausal adjuncts and "logophoric" environments, just like in Tamil.

In contrast, monstrous agreement under *ta(a)n*, of the kind observed in (335), has a much more restricted distribution. For many native-speakers, it occurs only in the propositional complement of a speech predicate. (335) illustrates monstrous agreement under *soll* (SAY); manner variants of this predicate may also induce such agreement in their immediate scope:

(340) Seetha$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ Krishnan-æ kaadali-kkir-
Seetha[NOM] ANAPH[NOM] Krishnan-ACC love-PRS-
een-nnŭ] katt-in-aa[.
1SG-COMP shout-PST-3FSG
"Seetha$_i$ shouted [$_{CP}$ that she$_{\{i,*j\}}$ loved Krishnan]"

(341) Krishnan$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ rombæ-vee kašṭappaḍæ-r-
Krishnan[NOM] ANAPH[NOM] very-EMPH suffer-PRS-
een-nnŭ] polamb-in-aan.
1SG-COMP whine-PST-3MSG
"Krishnan$_i$ whined [$_{CP}$ that he$_{\{i,*j\}}$ was suffering a great deal]."

In other words, monstrous agreement doesn't obtain in the scope of other types of propositional predicate – rather, in these cases, the agreement under subject *ta(a)n* matches the ϕ-values of the anaphoric antecedent, as in (334) and (336)-(339) above, and as discussed in detail in the long-distance binding chapters in Part I. Incidentally, this too is characteristic of indexical shift paradigms crosslinguistically, as the examples we have seen in the previous chapter show: in languages where this phenomenon is attested, it has been observed to obtain predominantly in the scope of speech predicates.

To sum up, the distribution of *ta(a)n* is much more general than that of monstrous agreement in Tamil. Furthermore, it patterns itself much more closely with that of anaphors crosslinguistically than with shifted indexicals. This poses a real challenge to the idea that a logophoric/obligatorily *de se* anaphor, and by extension *ta(a)n*, is an obligatorily shifted indexical.

11.1.2 Differences in minimality and optionality restrictions on antecedence

Baker (2008) observes that indexical shift only obtains in the propositional *complement* of a speech predicate which, in turn, ensures that the referent of a shifted indexical is a c-commanding DP in the minimally c-commanding clause. Monstrous agreement in Tamil is also subject to this restriction. Earlier we saw that non-monstrous agreement is also possible under speech predicates in Tamil. When we compare two minimally varying clauses under propositional *soll* (SAY), one without monstrous agreement (342) and the other with (343), we can see clearly that there are differences in antecedence-possibilities between the two:

(342) Seetha$_i$ [$_{CP}$ Maya$_j$ [$_{CP}$ taan$_{\{i,j\}}$ kiḻamb-ir-**aa**ḻ-
Seetha[NOM] Maya[NOM] ANAPH[NOM]$_i$ leave-PRS-3FSG-
ŭnnŭ] so-nn-aaḻ-ŭnnŭ] nene-tt-aaḻ.
COMP say-PST-3FSG-COMP think-PST-3FSG
"Seetha$_i$ thought [$_{CP}$ that Maya$_j$ said [$_{CP}$ that she$_{\{i,j\}}$ was leaving]]"

(343) Seetha$_i$ [$_{CP}$ Maya$_j$ [$_{CP}$ taan$_{\{*i,j\}}$ kiḻamb-ir-**een**-
Seetha[NOM] Maya[NOM] ANAPH[NOM]$_i$ leave-PRS-1SG-
nnŭ] so-nn-aaḻ-ŭnnŭ] nene-tt-aaḻ.
COMP say-PST-3FSG-COMP think-PST-3FSG
"Seetha$_i$ thought [$_{CP}$ that Maya$_j$ said [$_{CP}$ that she$_{\{*i,j\}}$ was leaving]]"

The sentence in (342) is, in fact, just a classic long-distance binding structure, such as we have seen many instances of in Part I. As we have seen, for a DP to qualify as a potential antecedent for *ta(a)n* two independent conditions must simultaneously hold: the DP in question must hold a spatio-temporal and/or mental perspective toward the minimal proposition containing *ta(a)n*, and it must not be asymmetrically c-commanded by the minimal PerspP containing *ta(a)n*. Other factors, such as whether the DP c-commands *ta(a)n*, whether it is one clause or several clauses away, or even whether it is asserted in the sentence at all, are irrelevant. This leads to the situation that in a given sentence involving *ta(a)n*, more than one DP may qualify to antecede this anaphor. In (342), both the matrix subject *Seetha* and the intermediate one *Maya* satisfy the conditions on potential antecedence, thus either of them may antecede *ta(a)n*.

In the monstrous agreement structure in (343), the DP that antecedes *ta(a)n* is *Maya*, showing that these two conditions on potential antecedence still prevail. The difference is that the matrix subject *Seetha*, despite also satisfying the potential antecedence conditions no longer qualifies to antecede *ta(a)n*. In other words, the conditions on potential antecedence are necessary but not sufficient in the monstrous agreement case. In binding with monstrous agreement, there is a Minimality condition that applies on top of the potential antecedence condition for standard long-distance binding structures like (342): in short, the antecedent of *ta(a)n* must be the AGENT (denoting the speaker) of the speech-verb that selects the *ta(a)n*-proposition as its complement.

Based on such differences in Relativized Minimality effects in indexical shift vs. long-distance binding phenomena, Baker (2008) concludes that:

> "logophoric pronouns are very similar to shifted first and second person pronouns, but they are not identical to them. All three must be bound by a designated operator. The difference is that first and second person pronouns must be bound by the closest operator of the relevant kind whereas third person pronouns (including logophorc ones) need not be."

Thus, here again we see that long-distance binding behaves differently from indexical shift – that it is, specifically, subject to less restricted rules.

A related difference between indexical shift and long-distance binding is that the latter, but not the former, manifest optionality with respect to the antecedent of *ta(a)n*. As long as the conditions on potential antecedence, recapitulated above, are satisfied, any DP in the sentence or in the salient surrounding discourse may antecede *ta(a)n*. In contrast, in long-distance binding structures that also involve monstrous agreement, the choice of antecedent for *ta(a)n* is deterministic: it always refers to the AGENT of the immediately superordinate selecting speech predicate.

11.1.3 Summary

We have seen three systematic differences between anaphors and shifted indexicals in this section:

(i) Differences in licensing environment: long-distance anaphors may occur in a range of structural environments that lend themselves to a perspectival semantics on the part of the antecedent nominal, such as adjunct PPs, DPs, and CPs, propositional complements, and root clauses. In contrast, shifted indexicals occur in a very narrow range of environments, specifically in the scope of a speech predicate.

(ii) Relatedly, indexical shift only obtains in the propositional *complement* of a speech predicate: i.e. the clause in which this phenomenon is attested must be directly selected by the speech verb. Long-distance binding, in contrast, may occur in a much wider range of environments: its antecedent may be several clauses away, may not c-command it and, in structures involving "logophoric" binding, may not be clause-internally asserted at all.

(iii) The choice of antecedent for an anaphor is a function of a complex interplay of factors pertaining to its perspective-holder status

> with respect to the minimal situational predication involving the anaphor. The choice of referent for a shifted indexical is rather more deterministic: it is the speaker of the speech event denoted by the immediately superordinate speech verb.

These factors systematically distinguish anaphors from shifted indexicals crosslinguistically. However, it is the presence of minimal pairs like (334)-(335) above, which explicitly show the interaction between indexical shift and long-distance binding, that conclusively prove it.

The discussion above thus shows that an analysis like that of Schlenker (2003b), which essentially conflates the categories of logophor/*de se* anaphor and obligatorily shifted indexical, is incorrect. The analysis of Anand (2006) maintains that there is an inherent difference between obligatorily shifted indexicals and anaphors that are interpreted obligatorily *de se*. Interestingly enough, though, Anand (2006) argues on independent empirical grounds that the anaphor *ta(a)n* in the closely related Dravidian language Malayalam is an obligatorily shifted indexical, not an anaphor. Anand's conclusion is based on the observation that Malayalam *ta(a)n* is insensitive to the "De Re Blocking effect" – a wellformedness condition that states that an obligatorily *de se* anaphor may not be c-commanded by a *de re* element.[1] Anand takes this insensitivity to automatically mean that Malayalam *ta(a)n* is an obligatorily shifted indexical, not an anaphor. Tamil *ta(a)n* may also be legitimately c-commanded by a *de re* element, like a deictic pronoun, as in (344) below, thus should also count as an obligatorily shifted indexical under Anand's reasoning:

(344) [$_{CP}$ [$_{CP}$ Mary tann-æ$_{\{i,j\}}$ virumbŭ-kir-aaḷ-ŭnnŭ] Bill$_j$ avan-kiṭṭæ
Mary ANAPH-ACC love-PRS-3FSG-COMP Bill him-ALL
so-nn-aan-nnŭ] John$_i$ nene-tt-aan.
say-PST-3MSG-COMP John$_i$ think-PST-3MSG
"John$_i$ thought [that Bill$_j$ told him$_i$ [that Mary loves him$_{i,j}$]]"

But, as discussed above, the existence of minimal pairs like that in (334)-(335) show that this conclusion must be false: *ta(a)n* itself may occur in both the unshifted structure in (334) and the shifted one in (335) but the 1st-person agreement on its clausemate verb, which we have shown is an instantiation of indexical shift, only obtains in the latter. Interestingly, minimal pairs like these are lacking in Malayalam. This is because that

[1] A detailed exposition of this condition and the motivation behind its formulation would take us too far afield. The interested reader is referred to Chapter 1, Section 1.4 of Anand (2006) for further details.

language entirely lacks agreement marking on the verb and therefore has neither monstrous agreement nor regular agreement. In other words, Tamil provides crucial novel evidence that the class of shifted indexicals is distinct from that of (obligatorily *de se*) anaphors.

11.1.4 Severing the direct connection between *ta(a)n* and agreement

The conclusion we have just reached, namely that *ta(a)n* itself has nothing to do with indexical shift, has a significant entailment. It shows that monstrous agreement under *ta(a)n* cannot have been *directly* triggered by *ta(a)n* itself but must have a different source.

Interestingly, we independently reached the same conclusion in Part I with respect to *non*-monstrous agreement under *ta(a)n* in standard long-distance binding constructions involving this anaphor in embedded subject position. Our conclusion was based on the fact that the agreement under *ta(a)n* is fully specified for person, number, and gender features; however, *ta(a)n* itself doesn't "care" about the gender feature on its antecedent suggesting that it is unspecified for this feature. Furthermore, the value of the gender feature on this agreement matches that of *ta(a)n*'s antecedent, suggesting that this is the ultimate source of the gender feature on T:

(345) Vivek$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ Seetha-væ$_j$ paar-tt-adaagæ] namb-
Vivek[NOM] ANAPH[NOM] Seetha-ACC see-PST-NMLZ believe-
in-aan.
PST-3MSG
"Vivek$_i$ believed [$_{CP}$ that he$_{\{i,*j\}}$ saw Seetha$_j$]."

(346) [Ellaa poɳ-gaɭ-ŭkk=um]$_i$ [$_{DP/PP}$ [$_{CP}$ Rajinikanth$_j$ [taŋ-gaɭ-
All girls-PL-DAT=CL Rajinikanth[NOM] ANAPH-PL-
ooɖæ]$_{\{i,*j\}}$ pees-in-aar] eŋgiradŭ-læ] rombæ sandoošam.
GEN speak-PST-3MSG that-LOC very happy
"[All girls]$_i$ were very happy about the fact that Rajinikanth$_j$ spoke to them$_{\{i,*j\}}$."

(347) Koɻendæ$_i$ [$_{CP}$ tan-akkŭ$_{\{i,*j\}}$ rombæ pasi-nnŭ] aɻŭ-
child[NOM] ANAPH-DAT very hunger[NOM-COMP] weep-
d-adŭ.
PST-3NSG
"The child$_i$ wept [$_{CP}$ that it$_{\{i,*j\}}$ was very hungry]."

At the same time, there is no conclusive way to show that the non-

monstrous agreement under *ta(a)n* in sentences like (345)-(347) is not triggered by *ta(a)n* itself. After all, one cannot "see" the ϕ-features on *ta(a)n* – thus, in a purely descriptive sense, the agreement facts in these sentences are compatible with the idea that *ta(a)n* is the source of agreement on its clausemate verb. The fact that Tamil uniformly manifests subject-verb agreement elsewhere might even be taken to additionally support such a conclusion.

Real support for the idea that the agremeent under *ta(a)n* in these cases doesn't come from *ta(a)n* itself comes from also looking at structures with monstrous agreement. We have already observed that, for the dialect being discussed here, both monstrous and non-monstrous agreement may be triggered under *ta(a)n* in subject position. This yields minimal pairs like the following:

(348) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ ʤej-pp-**een**-nnŭ]
Raman[NOM] ANAPH[NOM]$_i$ win-FUT-1SG-COMP
so-nn-aan.
say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that he$_{\{i,*j\}}$ would win]"

(349) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ ʤej-pp-**aan**-nnŭ]
Raman[NOM] ANAPH[NOM]$_i$ win-FUT-3MSG-COMP
so-nn-aan.
say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that he$_{\{i,*j\}}$ would win]"

It is much harder to maintain the position that the agreement under *ta(a)n* is directly triggered by *ta(a)n*, in the face of minimal pairs like that above. This is because, claiming that *ta(a)n* is the source of agreement in both would force us to say that there are two different underlying elements *ta(a)n* with different inherent ϕ-features in each. However, there is no independent evidence that the instances of *ta(a)n* in the two sentences differ in this manner. Minimal pairs like (348)-(349) thus show more conclusively than either of these sentences taken by themselves that the agreement under *ta(a)n* is not directly triggered by this element.

Based on this type of data, we proposed that, while *ta(a)n* may be indirectly involved in the Agree relationship that results in the T head acquiring the antecedent's ϕ-features, it is not the *source* of this agreement. Alternatively:

(350) The agreement under subject *ta(a)n* is not *directly* triggered by *ta(a)n*.

Interestingly, this is far from being an anomalous conclusion. Rather,

evaluated against the larger crosslinguistic tapestry, this is precisely the result that we are led to expect.

11.1.5 The Anaphor Agreement Effect (AAE)

Rizzi (1990) observed that an anaphor cannot occur in positions construed with agreement – a generalization that he termed the "Anaphor Agreement Effect" (AAE). This was motivated by minimal pairs like (351) and (352) in Italian, and analogous sentences in Icelandic (not listed here):

(351) A loro import-a solo di se-stessi.
to them matters-3SG only of them-selves
"They$_i$ only matter to themselves$_i$."

(352) * A loro interess-ano solo se-stessi.
to them interest-3PL only them-selves
"They$_i$ only interest themselves$_i$."

Rizzi independently shows that the THEME argument of *interessano* is in the nominative case and triggers agreement, whereas that of *importa* doesn't. As we can see, this distinction seems to directly regulate the grammaticality of these sentences: the anaphoric THEME in (352) in ungrammatical whereas that in (351) is licit. Woolford (1999) presents typological evidence from a much wider spectrum of languages to present a more nuanced version of Rizzi's AAE. Woolford illustrates that, in some languages, the presence of an anaphor in agreement-position doesn't lead to ungrammaticality: rather, the anaphor triggers a special form of agreement which is not attested elsewhere in the agreement paradigm of that language. This is demonstrated for Swahili below:

(353) Ahmed a-na-*ji*/*m-penda mwenyewe.
Ahmed 3SBJ-PRS-REFL/*3OBJ-love himself
"Ahmed$_i$ loves himself$_i$." (emphatic)

(354) Ahmed a-na-m-penda Halima
Ahmed 3SBJ-PRS-3OBJ-love Halima.
"Ahmed loves Halima."

Based on examples like (353)-(354), Woolford modifies Rizzi's original AAE, proposing that (Woolford 1999, p. 264, ex.22)):

(355) "Anaphors do not occur in syntactic positions construed with agreement, unless the agreement is anaphoric."

In a more recent work, Tucker (2011) observes that languages may employ the following linguistic strategies to avoid a violation of the AAE:

(i) The reflexive in agreeing position is deleted altogether yielding an intransitive with inherently reflexive interpretation. Inuit is a language that supposedly exhibits this strategy.

(ii) The verb surfaces with default agreement: the Icelandic and Italian examples originally discussed in Rizzi (1990) fall into this category.

(iii) The verb is marked with a special, "anaphoric" agreement, as in Swahili (353), above.

(iv) The anaphor appears inside a PP or possessor DP and is thus unable to trigger agreement: Greek, West Flemish, and the Malayo-Polynesian language Selayarese supposedly employ this "protected anaphora" strategy (see also Haegeman 2004).

(v) What looks like an anaphor is in fact a detransitivizing morpheme (yielding a similar inherently reflexive interpretation as in (i)) – i.e. not an actual argument and thus trivially not in agreement-triggering position. Tucker proposes that French is such a language.

Based on the empirical generalizations above, Tucker proposes the following re-modified version of the AAE (Tucker 2011, p. 30, ex. 40)):

(356) "Anaphors do not occur in syntactic positions construed with covarying ϕ-morphology."

If we consider the Tamil facts in light of this, it seems at first like Tamil trivially violates the generalization in (356). After all, given that Tamil uniformly manifests subject-verb agreement, *ta(a)n* appears in the subject position that is normally construed with co-varying ϕ-morphology and, in a sentence like (357) (repeated from (349)), the ϕ-features on the embedded verb under *ta(a)n* do match with those that we would expect *ta(a)n* to have, given its antecedent:

(357) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ ʤej-pp-**aan**-nnŭ]
Raman[NOM] ANAPH[NOM]$_i$ win-FUT-3MSG-COMP
so-nn-aan.
say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that he$_{\{i,*j\}}$ would win]"

The point is that, as discussed in Section 11.1.4, the agreement under *ta(a)n* doesn't *covary* with it. This becomes clear when we add structures with monstrous agreement under *ta(a)n* into the mix, as shown in (358), repeated from (348):

(358) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ ʤej-pp-**een**-nnŭ] so-nn-aan.
Raman[NOM] ANAPH[NOM]$_i$ win-FUT-1SG-COMP say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that he$_{\{i,*j\}}$ would win]"

(357) and (358), taken together, show that agreement under *ta(a)n is* special and that it is special in a way that is consistent with the generalization in (356).

At the same time, we see that the *way* in which Tamil avoids a violation of the AAE (as stated in (356)) is different from any of the five crosslinguistic strategies listed above. Neither the antecedent-matching nor monstrous agreement under *ta(a)n*, given in (334) and (335) respectively, is a type of default agreement which, we have seen, occurs under quirky dative subjects and always surfaces as 3NSG. It is also not "special" in the sense of Woolford (1999) – as, incidentally, suggested for Tamil structures like (335) by Woolford herself. For it to count as special in Woolford's sense, the agreement features under *ta(a)n* must not be attested elsewhere in that language – but as we have seen, the agreement in these structures simply corresponds to the 1SG agreement and 3MSG agreement that is triggered by DPs with these ϕ-values. Finally, the verbs under *ta(a)n* are also clearly not detransitivized, nor does *ta(a)n* itself seem to be "protected" from triggering agreement by being enclosed within a DP or PP.

11.1.6 Agreement strategy under *ta(a)n*

What, then, is the strategy that Tamil *does* employ? We have already answered this question in Part I for the standard cases of long-distance binding under subject *ta(a)n* and alluded to it again above. The agreement that obtains under *ta(a)n* in these cases reflects the ϕ-features on the antecedent of *ta(a)n*. This is most easily observed in a sentence where there is more than one potential antecedent to choose from, and where each potential antecedent bears a different specification for ϕ-features:

(359) Raman$_i$ [$_{CP}$ Seetha$_j$ [$_{CP}$ taan$_{\{i,*j\}}$ paɳatt-æ
Raman[NOM] Seetha[NOM] ANAPH[NOM] money-ACC
tirŭɖ-in-**aan**-nnŭ] namb-in-aaɭ-ŭnnŭ] paar-tt-aan.
steal-PST-3MSG-COMP believe-PST-3FSG-COMP see-PST-3MSG
"Raman$_i$ saw [$_{CP}$ that Seetha$_i$ believed [$_{CP}$ that he$_{\{i,*j\}}$ stole the money]]."

(360) Raman$_i$ [$_{CP}$ Seetha$_j$ [$_{CP}$ taan$_{\{*i,j\}}$ paɳatt-æ
Raman[NOM] Seetha[NOM] ANAPH[NOM] money-ACC

tirŭḍ-in-**aa**l̤-ŭnnŭ] namb-in-aal̤-ŭnnŭ] paar-tt-aan.
steal-PST-3FSG-COMP believe-PST-3FSG-COMP see-PST-3MSG
"Raman$_i$ saw [$_{CP}$ that Seetha$_i$ believed [$_{CP}$ that she$_{\{*i,j\}}$ stole the money]]."

In (359), the antecedent of *ta(a)n* is the matrix subject *Raman* which denotes a male individual and thus triggers 3MSG agreement on the matrix verb. The agreement under *ta(a)n* matches these ϕ-features and is also marked 3MSG. In (360), the antecedent of *ta(a)n* is *Seetha* (denoting a female individual) the subject of the intermediate clause: in this case, the agreement under *ta(a)n* is marked 3FSG. These sentences thus show that non-monstrous agreement under *ta(a)n* – which, indeed, represents the default scenario – tracks the antecedent of *ta(a)n*.

Interestingly, monstrous agreement under *ta(a)n* tracks the antecedent of this anaphor as well. Consider the examples below – the embedded verbal agreement form is formatted in boldface for convenience:

(361) Raman$_i$ [$_{CP}$ Krishnan$_j$ [$_{CP}$ taan$_{\{j,*i\}}$ Seetha-væ
Raman[NOM] Krishnan[NOM] ANAPH[NOM] Seetha-ACC
kaadali-kkir-**een**-nnŭ] so-nn-aan-nnŭ] kee l̤vipa-tt̤-aan.
love-PRS-1SG-COMP say-PST-3MSG-COMP overhear-PST-3MSG
"Raman$_i$ overheard [$_{CP}$ that Krishnan$_j$ said [$_{CP}$ that he$_{\{j,*i\}}$ loves Seetha.]]"

(362) Raman$_i$ [$_{CP}$ Krishnan$_j$ [$_{CP}$ taan$_{\{i,*j\}}$ Seetha-væ
Raman[NOM] Krishnan[NOM] ANAPH[NOM] Seetha-ACC
kaadali-kkir-**aan**-nnŭ] so-nn-aan-nnŭ] kee l̤vipa-tt̤-aan.
love-PRS-3MSG-COMP say-PST-3MSG-COMP overhear-PST-3MSG
"Raman$_i$ overheard [$_{CP}$ that Krishnan$_j$ said [$_{CP}$ that he$_{\{i,*j\}}$ loves Seetha.]]"

Monstrous agreement only obtains when the antecedent of *ta(a)n* is the speaker of the immediately superordinate speech predicate, as in (361), i.e. we have 1st-person agreement *because* the antecedent is 1st-person with respect to the speech event denoted by the immediately superordinate speech predicate. In the minimally varying sentence in (362), the antecedent of *ta(a)n* is the matrix subject *Raman* which is the AGENT of the non-speech predicate OVERHEAR – in this case, the agreement under *ta(a)n* still tracks the antecedent but may not be monstrous. This is because the antecedent is not 1st or 2nd-person with respect to the speech event denoted by the immediately superordinate speech predicate – i.e. it is neither the speaker nor the hearer of the intensional speech context.

However, the monstrous and non-monstrous agreement paradigms, taken together, show that the agreement under *ta(a)n* never enters into a direct relationship with the antecedent of *ta(a)n*. If we were to claim that such a direct relationship did exist, we wouldn't be able to explain why the agreement matches the ϕ-features of the antecedent in the non-monstrous case and differs from them in the monstrous one. Rather, what we would like to claim is that the apparent relationship between the agreement and the antecedent of *ta(a)n* comes out indirectly and is mediated by the operator in [Spec, PerspP], as follows. This operator corefers with the antecedent of *ta(a)n* in monstrous and non-monstrous cases alike. If both the antecedent and the operator in [Spec, PerspP] are evaluated against the same context, they will have always the same phi-features, as in (359)-(360). If they are evaluated against different contexts, they will have different person features. Specifically, if the operator is evaluated against the intensional context associated with the speech predicate whose AGENT the antecedent is, as in (361), the person feature on the operator will be 1. The features on the agreement are triggered by the operator in [Spec, PerspP] (just in case the clausemate subject is anaphoric, i.e. *ta(a)n*, and cannot itself value the ϕ-features on T). This yields the *effect* of a relationship between the agreement and the antecedent, in monstrous and non-monstrous cases alike.

This leads us to the following generalization:

(363) The agreement under subject *ta(a)n* indirectly tracks the antecedent of *ta(a)n* in monstrous and non-monstrous cases alike.

The answer to the question of what strategy Tamil employs to avoid a violation of the AAE is thus as follows:

(364) The agreement-marking under an anaphor in agreement-triggering position is originally triggered by some other element in its local domain.

It is interesting to ask, at this juncture, how Tamil seems to be able to "get away" with such a unique strategy, one that is moreover putatively so rare typologically. The answer is that it may not be such a rare strategy as it seems. We have already seen that the Niger Congo language Donna Sɔ seems to manifest a phenomenon that looks a lot like monstrous agreement (as in (365)).

(365) Oumar [$_{CP}$ inyemɛ jɛmbɔ paza bolum] miñ tagi.
Oumar [ANAPH[SBJ] sack.DEF drop left.1SG] 1SG.OBJ informed
"Oumar$_i$ told me [$_{CP}$ that he$_{\{i,*j\}}$ had left without the sack.]"

Other languages, like Amharic, have been reported to manifest indexical shift in clauses containing a *pro*-dropped subject (Delfitto and Fiorin 2011, ultimately from Malamud (2006), formatting mine):

(366) Profäsəru$_i$ [$_{CP}$ *pro*$_i$ bät'am bəzu səra ə-sär-allähu]
professor *pro* very much work 1SG-work.IMP-AUX.1SG
alä.
say.PRF.3SG.MASC
"The professor$_i$ said [$_{CP}$ that he$_i$ works very hard]."

In such cases, the *pro*-dropped subject is tacitly treated as an obligatorily shifted 1st-person indexical that happens to be silent. However, as Delfitto and Fiorin (2011, 219) correctly point out, there is no way to be sure that these silent elements are in fact 1st-person pronouns; they could just as well be anaphors (like *ta(a)n*) in a monstrous agreement configuration. As Delfitto and Fiorin further note, independent evidence that such sentences may be interpreted *de re* is suggestive of the latter option. It is thus possible that anaphors may appear in positions construed with covarying ϕ-morphology but that, for some independent reason, many of these are typically *pro*-dropped.

11.1.7 Where does monstrous agreement come from?

In our discussion of long-distance binding in Part I, we proposed that the agreement under *ta(a)n* is triggered by the silent pronominal operator hosted in the perspectival center of its phase – specifically, in [Spec, Persp-CenterP]. This same operator, we argued, is also responsible for binding *ta(a)n* at LF; this ensures by transitivity that the agreement tracks the antecedent of *ta(a)n*, just as desired.

It's an advantage of this approach that it can be naturally extended to cases of monstrous agreement, as well. We have just seen that monstrous agreement also tracks the antecedent of *ta(a)n* – in other words, that it is 1st-person precisely because the antecedent can be interpreted as 1st-person with respect to a linguistic context associated with the immediately superordinate speech predicate. So it's clear that, in these types of structures as well, the element that is responsible for binding *ta(a)n* – namely the operator in [Spec, Persp-CenterP] – must also be responsible for triggering agreement on its clausemate verb.

The core proposal that will allow us to tie these facts together is as follows:

(367) In clauses with monstrous agreement under *ta(a)n*, the pronominal operator in [Spec, Persp-CenterP] which is responsible for

binding *ta(a)n*, is a shifted 1st-person indexical.

11.1.8 Summary of results

Our conclusions from this section may be summarized as follows:

(368) The relationship between anaphora and agreement:

i. The agreement under *ta(a)n* is not directly triggered by *ta(a)n*. This is an instantiation of the AAE.
ii. All agreement under *ta(a)n* is originally triggered by the pronominal operator in the (phase-)local [Spec, Persp-CenterP] which binds *ta(a)n*. This accounts for the antecedent-tracking property of monstrous and non-monstrous agreement alike.
iii. In cases of non-monstrous agreement, the pronominal operator is unshifted with respect to the utterance context.
iv. In cases of monstrous agreement, the pronominal operator is a shifted 1st-person indexical with respect to the utterance context; i.e. it is interpreted 1st-person with respect to the intensional context introduced by the immediately superordinate speech predicate.

11.2 The uniqueness of speech predicates

Now that we have established these points, we can turn to the next piece of the puzzle, namely: *why* is indexical shift in the *ta(a)n*-clause induced just in case the immediately superordinate verb is a type of speech-predicate? I.e. what is the connection between the "speechiness" of a verb and the availability of indexical shift in its scope? In this section, I will present evidence from different types of crosslinguistic paradigms to show that speech predicates *are* syntactico-semantically special, thus different from all other classes of intensional predicate. Based on the nature of this evidence, it will be seen that the uniqueness of speech verbs has to do with the fact that their complements are structurally larger than those of other predicate-classes – a property that can be easily modelled within a cartographic framework of predicate selection (Cinque 1999). This property will be seen to be at the root of the fact that indexical shift obtains predominantly in the scope of such predicates.

11.2.1 Motivating the uniqueness of speech predicates

The literature on indexical shift doesn't make a principled distinction between speech-predicates and other types of intensional predicate. The logic behind this (often tacit) conflation is a semantic one that presupposes a binary divide between intensional and extensional predicate-types. However, in all the languages for which it has been reported, indexical shift has been shown to privilege speech predicates over other types of intensional verb.

Consider the table below, adapted from Anand (2006) (p. 110), which illusrates the types of verbs induce indexical shift in a small sample of the world's languages:

Language	**Verb**	**Class description**
English	all verbs	no indexical shift.
Amharic, Aghem	SAY	optionally shifts 1st/2nd person indexicals.
Navajo	SAY	optionally shifts 1st/2nd person indexicals.
Zazaki	SAY	optionally shifts all 1st person indexicals.
Tamil, Telugu	SAY	optionally shifts 1st-person verbal agreement.
Donna Sɔ (?)	SAY	obligatorily(?) shifts 1st-person verbal agreement.
Slave	TELL	optionally shifts 1st/2nd person indexicals.
	SAY	obligatorily shifts 1st person indexicals.
	WANT	optionally shifts 1st person indexicals.
Uyghur	all propositional verbs	indexical shift.

Table 11.1: Mini-typology of indexical shift across languages

Table (11.1) shows that *all* these languages allow indexical shift under propositional speech predicates. The fact that, in languages like Slave and Uyghur, indexical shift is induced under the scope of non-speech

predicates like WANT isn't a counter-argument to this claim since, in these languages, indexical shift *also* obtains under the scope of speech predicates. In other words, there is no language that allows indexical shift under non-speech predicates but not under speech-predicates.

That this is not accidental is confirmed by independent evidence from morphosyntactic phenomena crosslinguistically, which also privilege speech predicates.

11.2.2 Typological and syntactic evidence for the uniqueness of speech predicates

Various types of typological and syntactic evidence suggest that, while there is undeniably a relationship between speech- and attitude-predicates, they is nevertheless a deep divide between the two. For instance, based on a detailed survey of a wide typological array of languages, Cristofaro (2005) points out that there is a strong **implicational hierarchy** in the size of complement clauses that are selected by propositional predicates, which in turn corresponds to the lexico-conceptual semantics of these verbs. The implicational hierarchy she proposes is as follows:

(369) COMPLEMENT SIZE AND ANAPHORIC PREDICATE HIERARCHY:
Speech » Thought » Knowledge » Direct perception
Implication for complement size: Speech complements are structurally the largest; Direct perception complements are structurally the smallest.

Interestingly, there is independent crosslinguistically robust empirical support for this same hierarchy from a wide range of seemingly unrelated linguistic phenomena. As already briefly discussed in Part I, Culy (1994) reports, based on an investigation of anaphoric patterns in thirty-two languages, that the choice of structural environment in which an anaphor may be bound is directly regulated by the hierarchy above. In a nutshell, if an anaphor is licit in the complement of a certain class of predicate, it is also licit in the scope of all predicate-classes to the left of it in (369). For instance, if an anaphoric element is licensed in the complement of a direct perception predicate, it is also licensed in the scope of knowledge, thought, and speech predicates; if it is possible in the scope of a knowledge predicate, it is also licensed under thought and speech predicates, but is not necessarily licensed under direct perception verbs; and so on.

Speas (2004) extends Culy's insights to the case of evidentials and shows that the latter are also sensitive to the relative ordering above, though the implicational direction is reversed. To quote Speas (Speas

2004, pp. 263-4): "The more likely a predicate is to induce an anaphoric context, the less likely it is to be a category in the evidential paradigm: *say* is the predicate that is most likely to be anaphoric; hearsay is the category that is least likely to be a part of an evidential paradigm." Hooper and Thompson (1973), classifying predicates according to somewhat different criteria, demonstrate that certain root-transformations such as VP-preposing, topicalization, tag question formation, and left-dislocation can apply to embedded clauses with a degree of ease that can be predicted from the hierarchy above. Thus, root-transformations apply most easily to speech predicates, and least easily to direct perception predicates:

(370) **Root transformation – VP preposing:**

a. Manu intends to marry her, and marry her he will!
b. Manu intends to marry her, and he says [$_{CP}$ that marry her he will!]
c. * Manu intends to marry her, but he doubts [$_{CP}$ that marry her he will!]

(371) **Root transformation – Negative constituent preposing:**

a. Never before had Sally seen such a crowd of daffodils.[2]
b. Sally exclaimed [$_{CP}$ that never before had she seen such a crowd of daffodils.]
c. ?? Sally discovered [$_{CP}$ that never before had she seen such a crowd of daffodils.]
d. * Sally wondered [$_{CP}$ whether ever before had she seen such a crowd of daffodils.]

Thus, in the sentences above, VP-preposing – a root phenomenon, as (370a) shows, is grammatical under a speech predicate, as in (370b) but ungrammatical under a counterfactual one like that in (370c). Similarly, negative constituent preposing, also a root phenomenon (371a), is most easily available under the speech-verb *exclaimed* in (371b) and is either marked or ungrammatical under other types of verbs ((371c)-(371d)).[3] Wiklund et al. (2009) extend Hooper and Thompson's insight to Scandinavian languages and use the possibility of embedded V2 in these languages to confirm the ordering above.

[2]With apologies to William Wordsworth.

[3]Hooper and Thompson (1973) actually claim that verbs like DISCOVER (a member of their "Class E" predicate-type) allow root transformations just as easily as speech predicates do. I do think there is a real difference in grammaticality between the sentences in (371b) and (371c), however, an intuition that native English speakers I have consulted also share.

Finally, in his pioneering work in Cinque (1999), Cinque presents data to show that the relative ordering and interpretive scope of adverbial classes also respects the implicational predicate-hierarchy given in (369). I.e. adverbs that modify speech-verbs, like *frankly*, lie outermost in a sentence with a sequence of adverbs succeeded by evaluative ones like *fortunately*, evidential ones like *allegedly* and *reportedly*, and then by epistemic adverbs like *probably*. These are in turn followed by adverb classes that modify other predicate-types including non-propositional ones. A full range of ordered adverbs for English is reproduced below (Cinque 1999, 47):

(372) *frankly > fortunately > allegedly > probably > once/then > perhaps > wisely > usually > already > no longer > always > completely > well.*

This is a rigid ordering, as Cinque shows: an adverb may not precede another that is higher than itself in the implicational hierarchy. Even more strikingly, Cinque shows that this same hierarchy is reproduced to a greater or lesser degree in a wide array of languages such as those of the Romance family, Norwegian, Bosnian/Serbo-Croatian, Hebrew, Chinese, Albanian, and Malagasy – to name only a few. A parallel hierarchy emerges in the order of functional affixes modifying the various predicate-classes in head-final languages. Cinque shows that in Korean, this same ordering is reflected in the linear sequence of clause-final agglutinative morphemes. Consider the following Korean sentence from Cinque (1999, p. 53) – the idea is that any other relative ordering of suffixes would lead to ungrammaticality:

(373) Ku say-ka cwuk-ess-keyss-kwun-a!
that bird-NOM die-ANT-EPIST-EVAL-DECL
"That bird must have died!"

An analysis that views the fact that indexical shift is *always* induced under the scope of propositional speech predicates as accidental would be at a complete loss to deal with this range of data. Based on this type of evidence, I will now take it as uncontroversial that speech predicates are different from other types of predicate.

11.2.3 Modelling the uniqueness of speech predicates

Based on the observation that the relative ordering of adverbs and head-final affixes crosslinguistically is rigid, in the manner described above,

Cinque (1999) argues that there is a designated position in a clause for the relevant adverbs and functional heads which corresponds to the syntactico-semantic types of the predicates that they modify. He proposes a *universal functional sequence* of heads which is partially reproduced below:

(374) Speech Act » Evaluative » Evidential » Epistemological.

Speas (2004) proposes that different classes of propositional predicates select complements of differing sizes on this hierarchy, and that this can be used to explain the availability of things like evidential marking and logophora in those embedded clauses.

Based on the discussion and empirical arguments above, I will now assume without further argument that the major claims of this approach are, indeed, correct and propose the following:

(375) **Uniqueness of speech predicates:**

i. Speech predicates are underlyingly different from other types of predicate.

ii. The uniqueness of speech predicates is due to the fact that the complements of such predicates are structurally larger than those selected by other predicate-classes.

iii. Specifically, I propose that a propositional speech-predicate selects a Speech-Act Phrase which monotonically contains within it an Evaluative Phrase, an Evidential Phrase, and an Epistemological Phrase, in the relative order given in (374). These phrases constitute part of the C-layer in the left-periphery of a clause.

How do we connect this to the fact that indexical shift obtains predominantly under the scope of speech-predicates crosslinguistically? As we saw, despite their many differences, the standard analyses in the literature all share the intuition that indexical shift obtains when an indexical pronoun is evaluated against an intensional context[4] – specifically one introduced by an immediately superordinate speech predicate – rather than against the Kaplanian utterance context. The fact that indexical shift manifests itself on verbal agreement, a purely morphosyntactic phenomenon, in Tamil additionally shows that these intensional variables

[4]Recall that my use of the term "intensional context" is not intended to suggest that the type of information that is contributed by a speech predicate is ontologically the same as a Kaplanian utterance context. All I mean by "intensional context" is a sequence of mutually independent "rich" intensional indices that operate over modes of person, time, world, and location.

must be represented as early as the Narrow Syntax. In other words, these variables must be represented in the form of features that are legible to the syntactic module.

This isn't a wholly novel idea: recent proposals along these lines have been made by Speas (2004), Sigurðsson (2004), Baker (2008), Giorgi (2010), among others. For instance, describing indexical shift in Slave, a language of the Athapaskan family, Rice (1989, but see also the discussion of these facts in Baker (2008)) reports that there is a regular correspondence between complementizer-deletion patterns and indexical shift phenomena in this language. Specifically, she shows that the complements of verbs that do not induce indexical shift can host overt complementizers whereas those that do manifest indexical shift cannot. This leads to differences in grammaticality like the following (Rice 1989, p. 1273, formatting mine):

(376) No indexical shift → Overt complementizer

John [$_{CP}$ ʔeráke ʔée wihsı̨ gú] kodı̨hshǫ.
John parka 1sg.made comp 3.know.area
"John knows [$_{CP}$ that I made a parka]"

(377) Indexical shift → *Overt complementizer

John [$_{CP}$ ʔeráke ʔée wihsı̨ (*gú)] sedeyı̨dí.
John parka 1sg.made (*comp) 3.told.1sg
"John told me [$_{CP}$ he made a parka]"

As Baker (2008, p. 131, fn.11) points out, such data provide strong evidence that the contextual information responsible for inducing indexical shift is present in the syntax, and additionally suggest that it is represented in the complementizer layer of the clause.

Based on such evidence, Baker (2008, pp. 125-6, ex. 29f) proposes that:

(378) a. All matrix clauses and certain embedded clauses have two special null arguments generated within the CP projection, one designated S (for speaker) and the other A (for addressee).
b. In the absence of an overriding control relationship, S designates the person who produced the CP and A designates the person who the CP was addressed to.

We can implement this idea within the cartographic model of clausal selection described in (375) by proposing the following:

(379) Contextual Feature Generalization (CFG):

a. For any language, only those clausal complements that contain a Speech-Act Phrase are large enough to host intensional contextual features in the syntax.
b. These contextual features are hosted in the specifier of the Speech-Act Phrase.

11.2.4 Modelling the minimality restriction on monstrous agreement

One final point needs to be ironed out. We earlier observed a systematic difference in the minimality restrictions on antecedence for *ta(a)n* in standard long-distance binding structures and those additionally involving monstrous agreement in the *ta(a)n*-clause. We saw that the antecedent of *ta(a)n* in the standard case may be several phasal domains away and, in structures involving the so-called "logophoric" use of *ta(a)n*, may not be even sententially represented. In contrast, it was observed that long-distance *ta(a)n*-binding with monstrous agreement only obtains in the propositional *complement* of a speech predicate – as illustrated again below:

(380) * Maya$_i$ [$_{CP}$ Deepa$_j$ [$_{CP}$ taan$_i$ school-ŭkkŭ
Maya[NOM] Deepa[NOM] ANAPH[NOM] school[DAT]
poo-r-een-nnŭ] nene-tt-aa-ŭnnŭ] so-nn-aa.
go-PRS-1SG-COMP think-PST-3FSG-COMP say-PST-3FSG
"Maya$_i$ said [$_{CP}$ that Deepa$_j$ thought [$_{CP}$ that she$_{\{i,*j\}}$ was going to school.]]"

(381) Deepa$_j$ [$_{CP}$ Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ school-ŭkkŭ
Deepa[NOM] Maya[NOM] ANAPH[NOM] school[DAT]
poo-r-een-nnŭ] so-nn-aa-ŭnnŭ] nene-tt-aa.
go-PRS-1SG-COMP say-PST-3FSG-COMP think-PST-3FSG
"Deepa$_j$ thought [$_{CP}$ that Maya$_i$ said [$_{CP}$ that she$_i$ was going to school.]]"

The data above shows, in other words, that the referent of a shifted indexical in Tamil must be in the immediately superordinate clause. Baker (2008) presents evidence from shifted indexical phenomena in other languages to show that this is actually a more general restriction and models it as a Relativized Minimality condition on indexical binding.

However, it doesn't seem like Relativized Minimality is the driving principle behind the antecedence restriction in monstrous agreement structures like (381). Rather, this seems derivative of a more basic restriction that the antecedent of *ta(a)n* in these cases must be the SPEAKER of

the intensional context that induces indexical shift. But the intensional context itself is introduced in a clause by the speech predicate that selects that clause (as per (379) above). Thus, it must ultimately be due to a *thematic* restriction on the syntactico-semantic selectional properties of SAY. In other words, it must be a component of the meaning of SAY that the AGENT of this predicate is set as the speaker of the intensional context in its complement (again, as per (379)). It is well known that thematic relationships are extremely local – thus, it is entirely expected that indexical shift must involve the direct clausal complement of a speech predicate. Relativized Minimality effects on *ta(a)n*-antecedence in monstrous agreement structures follow from this.

This conception of things is reminiscent of standard descriptions of obligatory control, as in Chierchia (1989), Wurmbrand (2002), Landau (2012). For instance, Wurmbrand states that: "In obligatory control constructions, the antecedent is thematically or grammatically uniquely determined, the antecedent is obligatory, there are strict locality constraints on the relation between the antecedent and the embedded subject, only a sloppy interpretation is available under ellipsis, and only a *de se* interpretation is possible for the embedded subject" (Wurmbrand 2002, 99). Some of these properties – such as the obligatory *de se* reading of the embedded subject – hold true for both monstrous and non-monstrous structures with a *ta(a)n*-subject, alike. Thus, they reflect the properties of the clausal subject anaphor and not those of the speaker of the intensional context (which, I have proposed (379) above, is represented in [Spec, Speech-ActP] in the complement of the speech predicate). What *is* probative is the fact that, it is only in long-distance binding structures that also involve monstrous agreement that the antecedent of *ta(a)n* is obligatorily minimal and uniquely determined on thematic grounds.

(382) I thus propose that the relationship between the speaker of the intensional context represented in the complement of a speech verb and the AGENT of this verb is one of obligatory control.
I.e. it is a component of the meaning of this verb that its AGENT is set as the speaker of the intensional context in its complement.

As both Chierchia (1989) and Wurmbrand (2002) discuss in detail, obligatory control is a semantic relationship. There are thus different possibilities for how it could be represented syntactically. Speaking in terms of the case at hand, we could propose one of three possibilities:

(i) Semantically, the speech verb selects a property, not a proposition. The unsaturated argument, corresponding to the controlled argument, is not syntactically represented at all. Rather, it is supplied

later by a meaning postulate which is part of the denotation of the (speech) predicate, and which asserts that this argument is obligatorily controlled by the AGENT of this predicate.

(ii) Semantically, the speech predicate still selects a property. But the unsaturated argument *is* represented in the syntax, by some sort of silent DP in [Spec, Speech-ActP]. This argument is then abstracted over by a "propositional abstractor" (Chierchia 1989), creating a property out of the proposition – this is the property that the speech-verb selects. Obligatory control is thus derived via variable binding by the AGENT of the speech predicate.

(iii) Semantically, the speech predicate selects for a proposition. Again, there is a (silent) DP in [Spec, Speech-ActP] which represents the speaker of the intensional context. Obligatory control by the AGENT of the selecting speech predicate is thematically specified (again via meaning postulate). This is a possibility explored in Wurmbrand (2002) for certain types of infinitival constructions in German which are propositional but have obligatory control readings.

Any of these analytic options could be workable in the model being developed here. At this time, I do not have any empirical grounds for deciding among them. For the sake of concreteness, I will assume that the speaker of the intensional context in the complement of the speech predicate *is* syntactically represented, as described under (ii) and (iii) above – specifically, in [Spec, Speech-ActP], as proposed in (379). I will also assume that the complement of the speech predicate represents a property, rather than a proposition, as described under (ii). Nothing of substance rests on this choice for now, however, and all the results derived below will transfer seamlessly if we adopt one of the other options. In Section 12.3.1 of the following chapter, I will provide an explicit feature-structure and semantics for the DP in [Spec, Speech-ActP] which denotes the speaker of the intensional context in that clause.

Chapter 12

Analysis: when anaphora meets agreement meets indexicality

In this chapter, we will formalize our conclusions from the previous one to derive a precise syntax and semantics for monstrous agreement structures in Tamil. We will first flesh out the templatic hierarchy in (374) to yield its specific instantiation in Tamil. With all the pieces of the puzzle finally in place, we will then update our toolbox of features, syntactic operations, and LF semantic rules from that developed in Parts I and II for standard binding structures in Tamil. In particular, we will propose a formal denotation for the Speech-Act head and investigate the nature of the feature in [Spec, Speech-ActP] that will ultimately denote the speaker of the intensional context introduced by a speech-predicate. With these formalisms in hand, we will then walk through step-by-step derivations of monstrous and non-monstrous agreement structures like those in (334) and (335).

12.1 The clausal functional sequence in Tamil

Although some version of the functional sequence presented in (374) is assumed to hold universally across languages, it is only a template that helps make broad predictions about hierarchical relationships between different predicate-classes. Subtle differences in the specific instantiations of the categories represented in the Cinque hierarchy do obtain from one language to the next. For instance, as Cinque (1999) and Speas (2004) both note, languages may differ in the type of evidential marker they express. I will argue below that languages may similarly vary in their

choice of perspectival marker. Before proceeding to a formal analysis of the monstrous and "regular" agreement paradigms in Tamil, it is thus important to be clear about the specific nature and expression of the relevant parts of the clausal functional sequence in this language. This is the focus of this section.

12.1.1 Differences between speech and thought complements

The first thing we must do is motivate the distinction between speech and thought complements in Tamil. We have claimed that monstrous agreement obtains predominantly in the scope of speech predicates, drawing a distinction between these and all other intensional predicates including predicates of thought and belief, like *nene* (THINK) and *nambŭ* (BELIEVE). We have modelled the uniqueness of speech predicates in this sense by claiming that such predicates introduce a structurally larger complement than the other classes of intensional predicate – specifically one that contains a projection for Speech-ActP.

A concern is that the class of thought-predicate is believed to also introduce a fairly large complement. Both tend to allow root transformations to more or less equivalent degrees, as discussed in Hooper and Thompson (1973). But under the account being proposed here, we should expect there to be some detectable syntactic differences between the complements of the two classes of predicate in addition to the licensing of monstrous agreement. Such differences are indeed hard to find: the two classes of predicate behave quite similarly in Tamil, just as they have been reported to do in other languages. Nevertheless, speech and thought verbs do differ systematically with respect to some other structural and interpretive properties, showing that the grammar does treat them as belonging to distinct syntactico-semantic classes:

Speaker-oriented adverbs: A primarily speaker-oriented adverb like *satjamaa* (TRUTHFULLY/HONESTLY) may more easily modify the complement of a speech predicate than that of a thought predicate:

(383) Maya [$_{CP}$ Raman satjamaa paɳatt-æ
Maya[NOM] Raman[NOM] honestly money-ACC
toɖæ-læ-nnŭ] saadi-tt-aaḷ.
touch.INF-NEG-COMP swear-PST-3FSG
"Maya swore [$_{CP}$ that Raman honestly didn't touch the money]"

(384) ? Maya [$_{CP}$ Raman satjamaa paɳatt-æ
Maya[NOM] Raman[NOM] honestly money-ACC
toɖæ-læ-nnŭ] nene-tt-aaḷ.
touch.INF-NEG-COMP think-PST-3FSG
"Maya thought [$_{CP}$ that Raman honestly didn't touch the money]" (Intended)

Neg-raising differences: Hooper and Thompson (1973) show that the complements of thought predicates manifest Neg-raising whereas those of speech predicates do not. Thus, in English, a sentence like: "She didn't think John was sick" may mean: "She thought John wasn't sick". However, the same cannot be said of "She didn't say John was sick"; i.e. this doesn't mean "She said John wasn't sick". In other words, the negation in the sentence with the speech verb simply negates the saying, not the proposition under it. The same difference is found in Tamil as well:

(385) Raman [$_{CP}$ Krishnan varŭ-v-aan-nnŭ]
Raman[NOM] Krishnan[NOM] come-FUT-3MSG-COMP
nenekkæ-læ.
think.INF-NEG
"Raman didn't think that Krishnan would come."
= "Raman thought that Krishnan wouldn't come." (Possible reading)

(386) Raman [$_{CP}$ Krishnan varŭ-v-aan-nnŭ]
Raman[NOM] Krishnan[NOM] come-FUT-3MSG-COMP
sollæ-læ.
say.INF-NEG
"Raman didn't say that Krishnan would come."
≠ "Raman said that Krishnan wouldn't come."

Verb-deletion: The propositional speech predicate following a complementizer may be deleted altogether. In this case, the complementizer does "double-duty" and functions both as a subordinator and as the speech-predicate. This is only possible if the deleted verb is one of speech:

(387) Raman [$_{CP}$ ţiičær tirŭmbi va-r-
Raman[NOM] teacher[NOM] return come-PRS-
aar-ŭnnŭ] so-nn-aan.
3MSG.HON-COMP say-PST-3MSG

"Raman said [$_{CP}$ that the teacher was coming back]."

(388) Raman [$_{CP}$ ʈiičær tirŭmbi va-r-
Raman[NOM] teacher[NOM] return come-PRS-
aar-ŭnn]-aan.
3MSG.HON-COMP-3MSG
"Raman said [$_{CP}$ that the teacher was coming back]."
Not: "Raman thought [$_{CP}$ that the teacher was coming back]."

Steever (2002) reports that the finite complementizer *-(ŭ)nnŭ* (Literary *enrŭ*) in these sentences is actually the "conjunctive form of the verb *ena* 'SAY"' (Steever 2002, 97) – thus its ability to function as a speech-predicate on its own is not particularly surprising. However, this seems to be a more general property across typologically unrelated languages. For instance, Sells (1987) describes a very similar property for the Niger Congo language Tuburi, which has a complementizer *gā*; Sells reports that this morpheme "carries the force of speech itself" and that the speech predicate may be left out altogether.

Gerundive formation: Gerundives are most easily formed under speech predicates. Gerundive structures under speech verbs like SAY were consistently rated as some of the most grammatical in my survey, with high mean-values for grammaticality and low standard deviations. Gerundives are less easily formed under other types of predicate, including the class of thought-predicate:

(389) Raman [$_{CP}$ Krishnan kaɖæ-kkŭ poo-r-
Raman[NOM] Krishnan[NOM] store-DAT go-PRS-
ad-aagæ] so-nn-aan.
3NSG-NMLZ say-PST-3MSG
"Raman spoke [$_{CP}$ of Krishnan's going to the store]."

(390) ? Raman [$_{CP}$ Krishnan kaɖæ-kkŭ poo-r-
Raman[NOM] Krishnan[NOM] store-DAT go-PRS-
ad-aagæ] nene-tt-aan.
3NSG-NMLZ think-PST-3MSG
"Raman thought [$_{CP}$ of Krishnan's going to the store]."

While it's not entirely obvious how to explain these distinctions in structural terms, their existence does show that the grammar systematically

distinguishes the complements of speech and thought verbs. It is reasonable, then, to adopt the proposal that they differ in the presence of a Speech-ActP in the former but not the latter.

12.1.2 Where does the PerspP fit in?

Next we must determine where in the clausal structure PerspP is encoded in Tamil. To do this, we must first examine whether it makes sense to propose that it is a part of the clausal functional sequence in the first place. The formal representation of the PerspP, repeated from (252) in Part II, is as follows:

(391) **Formal representation of a Perspectival Center:**

i. The *perspectival center* contains the coordinates pertaining to the time, location, world, and/or mental information of a salient perspective holder. These are hosted in a silent pronominal operator * in the specifier of a Perspectival Phrase (PerspP).

ii. Certain phases (at least some PPs, DPs, AspPs, CPs) contain a PerspP by virtue of what they inherently "mean".

iii. A phase has at most one PerspP.

iv. The phase containing a successfully bound anaphor must contain a PerspP. The operator in [Spec, PerspP] Agrees with the anaphor in its minimal phase and variable-binds it at LF.

v. A potential antecedent may not be asymmetrically c-commanded by the PerspP it holds a perspective towards. The relationship between the potential antecedent and this operator is one of non-obligatory control.

The definition above suggests that the PerspP might not be a part of the clausal functional sequence at all, but orthogonal to it, much like Topic projections in the Rizzi (1997) sense.[1] First, the fact that PerspP may be hosted in non-clausal constituents like PP, AspP, and DP, as we have shown, argues against the idea that it is functional information specific to a CP. Second, as Speas (2004) and Culy (1994) show, evidence from binding patterns across languages explicitly shows that this information may be hosted at different heights for different clauses. For instance, whereas some languages, like Donna Sɔ only seem to allow long-distance binding under speech predicates, suggesting that the PerspP in this language is located fairly high (e.g. at the level of the Speech-ActP)

[1] Thanks to Peter Svenonius (p.c.) for bringing this important point to my attention.

other languages, like Tamil, allow binding under all types of propositional predicate, suggesting that it is positioned rather low.

I thus assume that the PerspP is not part of the rigid functional sequence of a particular structural region like the clause.[2] Rather, I propose, in line with what we have already discussed in Part II of this dissertation, that PerspP is a property of phases in general. It may be interpreted as a kind of anchor which explicitly situates a phase in connection with its sentential and salient discourse environments and relates it in a particular way (along the mental, spatial, or temporal dimensions) to an individual in this environment. In this sense, it is more along the lines of a projection like Topic, although a detailed investigation of the nature and extent of a possible connection between the two is beyond the scope of the current discussion.

This said, we will still assume that PerspP is encoded in the same structural position for all CPs in Tamil, so we should still be able to discuss where in the CP this projection is represented relative to the Speech-ActP. A central part of the current analysis has to do with the idea that the Speech-ActP, which has been argued to be selected predominantly by speech predicates, is the highest projection of a CP. As such, it is important to show that PerspP is indeed lower than it in the structural hierarchy.

Evidence that this is so comes from the observation that *any* propositional predicate may induce long-distance binding of subject or object *ta(a)n* in its scope in Tamil. Relevant illustrations are given below:

(392) SPEECH VERBS:
Krishnan$_i$ Anand-kiṭṭe$_j$ [$_{CP}$ taan{i,*j} rombæ
Krishnan[NOM] Anand-ALL ANAPH[NOM] very
buddhisaali-nnŭ] soll-i-ko-ɳḍ-aan.
intelligent.man[NOM]-COMP say-ASP-*ko*ɭ-PST-3MSG
"Krishnan$_i$ told Anand$_j$ that he$_{\{i,*j\}}$ was very intelligent."

(393) THOUGHT VERBS:
Seetha$_i$ [$_{CP}$ Raman$_j$ tann-æ$_{\{i,*j\}}$ kaadali-kkir-aan-nnŭ]
Seetha[NOM] Raman[NOM] ANAPH-ACC love-PRS-3MSG-COMP

[2] In contrast, monstrous agreement which is dependent on the presence of a Speech-ActP in the *ta(a)n*-clause *is* contingent on the selectional properties of the immediately superordinate predicate, as we have already seen. This, indeed, is why monstrous agreement only obtains under predicates which select a Speech-ActP – like speech predicates – and why the antecedent-anaphor relationship in such structures is strictly minimal. Speech-ActP is also not found in non-clausal constituents. Thus Speech-ActP *is* part of the rigid clausal functional sequence.

tappaa nambŭ-gir-aaḷ.
wrongly believe-PRS-3FSG
"Seetha$_i$ wrongly believes [$_{CP}$ that Raman$_j$ loves her$_{\{i,*j\}}$.]"

(394) KNOWLEDGE VERBS:
Raman-ŭkkŭ$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ pooṭṭi-læ tookka-ppoo-gir-aan-eŋgiradŭ] terijum.
Raman-DAT ANAPH.NOM contest-LOC lose-LEAVE-PRS-3MSG-REL know-PST-3NSG
"Raman$_i$ knew [$_{CP}$ that ANAPH$_{\{i,*j\}}$ was going to lose in the contest.]"
Lit: "Raman$_i$ knew of his$_{\{i,*j\}}$ going-to-lose in the race."

Perhaps even more tellingly, *ta(a)n* can be long-distance bound even under the scope of predicates that are traditionally classed as "non-finite" or "less finite", as in the object control structure in (395a) and gerundival complement in (395b):

(395) OBJECT *ta(a)n* IN INFINITIVAL AND GERUNDIVAL COMPLEMENTS:
a. Rukmini$_i$ Krishnan-æ$_j$ [$_{CP}$ PRO$_{\{j,*i\}}$ tann-æ$_{\{i,*j\}}$ kaappaatt-æ] veeṇḍ-in-aaḷ.
Rukmini[NOM] Krishnan-NOM PRO ANAPH-ACC rescue-INF pray-PST-3FSG
"Rukmini$_i$ prayed to Krishnan$_j$ [$_{CP}$ PRO$_{\{j,*i\}}$ to rescue her$_{\{i,*j\}}$.]"
b. Paijan-ŭkkŭ$_i$ [$_{DP}$ akka$_j$ tann-æ$_{\{i,*j\}}$ school-ŭkkŭ tajjaar-sej-dadŭ] rombæ piḍi-tt-adŭ.
Boy-DAT sister[NOM] ANAPH-ACC school-DAT ready-make-GER much like-PST-3NSG
"[The boy]$_i$ liked (his$_{\{i,*j\}}$) sister's getting him$_{\{i,*j\}}$ ready for school very much."

Additional indirect evidence for the relative ordering of Speech-ActP and PersP comes from interactions with the evidential marker *-aam* which attaches to the end of a clause:

(396) Raman neettikkŭ veelæ-kkŭ rombæ veegam-aa naḍandŭ-poo-n-aan-aam.
Raman[NOM] yesterday work-DAT very quick-ADV walk.ASP-go-PST-3MSG-EVID
"Yesterday, Raman apparently walked to work very quickly."

As we have already seen, languages vary in the type of evidential marker they choose to instantiate which in turn corresponds to the kind of evidence that the morpheme denotes. The evidence could be of a direct

nature as, for instance, if the speaker were to directly experience it; alternatively, the evidence could be less direct, resulting, for instance, from personal inference. Least indirect of all is hearsay – evidence based on a verbal report made to the speaker of the utterance context. By carefully investigating evidential morphemes in a range of languages, Speas argues that different types of evidential markers target different designated projections along the clausal functional sequence. In other words, she proposes that there is no monolithic "Evidential Phrase" category. Based on Speas' categorization, Tamil *-aam* seems to be a (structurally) high evidential corresponding to evidence that is reported to the speaker of the utterance context. But it is unclear at this juncture exactly how high it is in the clausal functional sequence.

Regardless, sentences like (397) below show that the PerspP in a CP may be present in an embedded EvidP in Tamil:

(397) Raman$_i$ Krishnan-kiṯṯærŭndŭ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ
Raman[NOM] Krishnan-ABL ANAPH[NOM] prize-ACC
ʤej-čč-aan-aam-ŭnnŭ] keeḷvipa-ṯṯ-aan.
win-PST-3MSG-EVID-COMP find.out-PST-3MSG
"Raman$_i$ found out from Krishnan [$_{CP}$ that he$_{\{i,*j\}}$ apparently won the prize]."

One piece of potential evidence for the relative hierarchy of PerspP and EvidP comes from interactions with the question and exclamative particles in the following examples:

(398) Raman neettikkŭ veelæ-kkŭ rombæ veegam-aa naḍandŭ-
Raman[NOM] yesterday work-DAT very quick-ADV walk.ASP-
poo-n-aan-aa?
go-PST-3MSG-Q
"Did Raman walk very quickly to work yesterday?"

(399) Raman neettikkŭ veelæ-kkŭ rombæ veegam-aa naḍandŭ-
Raman[NOM] yesterday work-DAT very quick-ADV walk.ASP-
poo-n-aan-ee!
go-PST-3MSG-EXCL
"Raman walked very quickly to work yesterday!"

The question particle *-aa* and exclamative particle *-ee* seem to be Speech-Act Mood markers under the Cinque system delineated above since they mark "the basic illocutionary force of a sentence (Cinque 1999, 84)".

Now, when the evidential marker co-occurs with one of these Speech-Act markers, only one relative ordering is licit – the Speech-Act marker must come outermost. This, indeed, is what we expect, given the Mirror

Principle, if the phrase containing the evidential marker is lower on the clausal spine than the Speech-Act Phrase:

(400) EVIDENTIAL + *Yes/No* QUESTION MARKER:

a. Raman neettikkŭ veelæ-kkŭ rombæ veegam-aa
Raman[NOM] yesterday work-DAT very quick-ADV
naɖandŭ- poo-n-aan-aam-aa?
walk.ASP- go-PST-3MSG-EVID-Q
"Did Raman apparently walk to work very quickly yesterday?"

b. * Raman neettikkŭ veelæ-kkŭ rombæ veegam-aa
Raman[NOM] yesterday work-DAT very quick-ADV
naɖandŭ- poo-n-aan-aa-aam?
walk.ASP- go-PST-3MSG-Q-EVID
"Did Raman apparently walk to work very quickly yesterday?" (Intended)

(401) EVIDENTIAL + EXCLAMATIVE MARKER:

a. Raman neettikkŭ veelæ-kkŭ rombæ veegam-aa
Raman[NOM] yesterday work-DAT very quick-ADV
naɖandŭ- poo-n-aan-aam-ee!
walk.ASP- go-PST-3MSG-EVID-EXCL
"Raman apparently walked to work very quickly yesterday!"

b. * Raman neettikkŭ veelæ-kkŭ rombæ veegam-aa
Raman[NOM] yesterday work-DAT very quick-ADV
naɖandŭ- poo-n-aan-aa-ee-aam!
walk.ASP- go-PST-3MSG-EXCL-EVID
"Raman apparently walked to work very quickly yesterday!" (Intended)

This suggests that the evidential marking is in a position lower than the Speech-Act Phrase which houses the question and exclamative suffixes *-aa* and *-ee*, respectively. Since a detailed investigation of the functional sequence in the clausal left-periphery in Tamil is outside the scope of the current discussion, I simply label the projection for the evidential marker "High Evidential Phrase" (High-EvidP).

Returning then to the position of PerspP, recall that it is able to occur in non-finite clause types, as shown by the possibility of long-distance binding into such clauses in the examples in 395. We might expect these clauses to be structurally reduced (Wurmbrand 2001), and indeed, the

addition of a speech-act interrogative/exclamative marker or an evidential marker to the embedded CPs in these sentences is impossible:

(402) * Rukmini$_i$ Krishnan-æ$_j$ [$_{CP}$ PRO$_{\{j,*i\}}$ tann-æ$_{\{i,*j\}}$
Rukmini[NOM] Krishnan-NOM PRO ANAPH-ACC
kaappaatt-a-aam] veeɳɖ-in-aaḷ
rescue-INF-EVID pray-PST-3FSG
"Rukmini$_i$ prayed to Krishnan$_j$ [$_{CP}$ PRO$_{\{j,*i\}}$ to apparently rescue her$_{\{i,*j\}}$]." (Intended)

(403) * Rukmini$_i$ Krishnan-æ$_j$ [$_{CP}$ PRO$_{\{j,*i\}}$ tann-æ$_{\{i,*j\}}$
Rukmini[NOM] Krishnan-NOM PRO ANAPH-ACC
kaappaatt-a-ee] veeɳɖ-in-aaḷ
rescue-INF-EXCL pray-PST-3FSG
"Rukmini$_i$ prayed to Krishnan$_j$ [$_{CP}$ PRO$_{\{j,*i\}}$ to rescue her$_{\{i,*j\}}$!]." (Intended)

This suggests that, while gerundival and infinitival embedded clauses such as those in (395a) and (395b) are too small to host Speech-Act and evidential markers, they are still large enough to host a perspectival phrase which can bind *ta(a)n*. Based on this type of evidence, I will thus conclude that PerspP is structurally below both Speech-ActP and High-EvidP, leaving us with the following partial hierarchy for CPs in Tamil:

(404) Speech-ActP» High-EvidP» PerspP.

12.2 Summary of results

We now have all the pieces of the puzzle on hand and are ready to analyze the Tamil minimal pair in (305) vs. (304) which triggered the discussion in this series of chapters. Let us lay the pieces on the table:

I. The embedded agreement in (305) reflects indexical shift in the embedded CP of this sentence. It is a type of "monstrous agreement".

II. Contra the claims of Schlenker (2003b, and subsequent), anaphors and obligatorily shifted indexicals are underlyingly different elements. Contra Anand (2006), Dravidian *ta(a)n* is not an obligatorily shifted indexical but an anaphor.

III. The verbal agreement, monstrous or otherwise, that surfaces under *ta(a)n* in subject position is not triggered directly by *ta(a)n* but has

a different ultimate source. This is an instantiation of the Anaphor Agreement Effect/AAE.

IV. Propositional predicates crosslinguistically are hierarchically ordered in a functional sequence with respect to the size of their clausal complements. Tamil obeys the templatic hierarchy provided in (374) and has the following partial hierarchy in the clausal left-periphery:
Speech-ActP (largest) » High-EvidP » PerspP.

V. Information pertaining to the intensional context is structurally represented as early as the Narrow Syntax. Intensional-context features are present in [Spec, Speech-ActP] and are responsible for the syntactic derivation of indexical shift phenomena.

VI. The intensional-context variable introduced by the Speech-Act head is obligatorily controlled. It gets variable-bound by the AGENT of the selecting speech predicate, in the sense of Chierchia (1989). This yields the result that indexical shift always obtains in the propositional *complement* of a speech-predicate and that the antecedent of *ta(a)n* in such structures is always the AGENT of this speech verb

VII. The fact that indexical shift *always* obtains in the immediate scope of speech predicates crosslinguistically is thus not an accident but results directly from the syntactico-semantics of speech-predicates modelled in the manner described above.

These are the core theoretical insights we have developed. Below, I develop a formalization of these ideas.

12.3 Toolbox of features, structures, and rules

Let us consider again the relevant monstrous agreement and non-monstrous agreement paradigms in Tamil:

(405) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ ʤej-pp-een-nnŭ] so-nn-aan.
Raman[NOM]$_i$ ANAPH[NOM]$_i$ win-FUT-1SG-COMP say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that he$_{\{i,*j\}}$ would win.]"

(406) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ ʤej-pp-aan-nnŭ] paar-tt-aan.
Raman[NOM]$_i$ ANAPH[NOM]$_i$ win.FUT-3MSG-comp see.PST-3MSG
"Raman$_i$ saw [$_{CP}$ that he$_{\{i,*j\}}$ would win.]"

In order to formally derive these structures in the syntax, we need to know the syntactic representation of their key players. This involves knowing, at the very minimum:

(i) The features of the intensional context in [Spec, Speech-P].

(ii) The features on the local syntactic binder of *ta(a)n*: the operator in [Spec, Persp-P].

(iii) The properties of ϕ-features, in particular, the nature of 3rd-person.

(iv) The features on *ta(a)n* and those on its antecedent DP.

We already have much of this information. We will thus focus our attention on (i), the relationship between (i) and (ii), and the representation of the 3rd-person feature (iii).

12.3.1 Features on the intensional context in [Spec, Speech-ActP]

We have seen that monstrous agreement in Tamil instantiates indexical shift for 1st-person. But there isn't clear evidence suggesting that other types of indexicals may also shift in this language. Thus, in the monstrous agreement example in (407), Raman is going to hit you, not Krishnan; and he is going to do so today, and did not already do so yesterday:[3]

(407) Raman$_i$ neettikkŭ$_{t-1_{utt}}$ Krishnan-kiṭṭæ$_j$ [$_{CP}$ taan$_{\{i,*j\}}$ onn-æ$_{*j,Addr_{utt}}$ innikkŭ$_{t_{utt},*t-1_{utt}}$ aḍikka-ppoo-r-een-nnŭ] so-nn-aan-aam.
Raman yesterday Krishnan-ALL ANAPH[NOM] you-ACC today hit-go-PRS-1SG-COMP say-PST-3MSG-EVID
"Raman$_i$ apparently told Krishnan$_j$ [$_{CP}$ that he$_{\{i,*j\}}$ was going to hit you$_{Addr_{utt},*j}$ today$_{t_{utt},*t-1_{utt}}$.]]]"

[3]In fact, even the 1st-person pronoun *naan* in Tamil behaves like a rigid Kaplanian indexical, as we already saw. I will propose that this is because *naan* is explicitly specified to *only* refer to the speaker of the utterance context, just as discussed in Schlenker (2003b) for rigid "English-like" indexicals.

I will thus make the more minimal assumption, for now, that only the contextual coordinate for speaker is syntactically represented, thus shiftable, in [Spec, Speech-ActP].

The simplest and most intuitive way to represent the speaker of the intensional context syntactically is as a pronoun. In Part I, we represented the operator in [Spec, PerspP] as a pronoun with a valued DEP-feature. It is tempting to employ the same strategy with the pronoun in [Spec, Speech-ActP], as well. However, the DEP-feature plays a specific role in the derivation of syntactic binding relations, as captured by its definition, repeated below:

(408) **The DEP feature:**

i. A DEP feature marks two DPs X and Y that are in a syntactic binding dependency with one another.
ii. DEP takes integers or letters as value. The assignment function maps these values to salient entities in the evaluation context.
iii. Two elements with matching DEP values will thus denote the same entity in the evaluation world and are construed to be in a binder-bindee relationship with one another.

However, we will see that the DP in [Spec, Speech-ActP] doesn't itself enter into a formal syntactic dependency with another element in its local domain. Rather, we will see that it represents an anchoring point for the interpretation of person features in the local syntactic structure, which restricts reference assignment at LF in systematic ways.

Thus, in the interests of simplicity, I will assume that the DP in [Spec, Speech-ActP] just has a privative D feature, and denotes a variable. I will assume, again for simplicity, that it doesn't have any ϕ-features, since they will play no crucial role in the implementation. It is not an anaphor in the syntactic sense because it has no unvalued features. However, it is an anaphor in the LF-semantic sense because the variable gets λ-abstracted over and ultimately saturated by the AGENT of the selecting speech predicate. I assume, following Chierchia (1989), that the choice of AGENT vs. a DP that is potentially minimally closer (like a superordinate GOAL object) as the controller is explicitly asserted via meaning postulate as part of the denotation of the speech predicate.

This is all we need to say about the contextual feature at present. Let us now review our assumptions about the other main features in the monstrous and non-monstrous agreement paradigms in Tamil.

12.3.2 The relationship between [Spec, PerspP] and [Spec, Speech-ActP]

In the syntax, the DP in [Spec, PerspP] has a valued DEP-feature, valued ϕ-features and a D feature, whereas the one in [Spec, Speech-ActP] has just a D feature. The lack of a DEP-feature on the DP in [Spec, Speech-ActP] means that we don't expect any syntactically driven referential dependency between the two DPs. This in turn means that we neither expect them to be obligatorily coreferent nor obligatorily non-coreferent with one another. In other words, we expect them to be able to corefer on occasion and refer independently otherwise.

This is empirically borne out. Consider the sentences in (410) below – the agreement marking on the embedded verb in the *ta(a)n*-clause is marked in boldface in each:

(409) Krishnan$_i$ [$_{CP}$ Raman$_j$ [$_{CP}$ taan$_j$ paris-æ
Krishnan[NOM] Raman[NOM] ANAPH[NOM] prize-ACC
ʤej-tt-**een**-nnŭ] so-nn-aan-nnŭ] kaɳḍŭpiḍi-tt-aan.
win-PST-1SG-COMP say-PST-3MSG-COMP discover-PST-3MSG
"Krishnan$_i$ discovered [$_{CP}$ that Raman$_j$ said [$_{CP}$ that he$_j$ won the prize.]]"

(410) Krishnan$_i$ [$_{CP}$ Raman$_j$ [$_{CP}$ taan$_i$ paris-æ
Krishnan[NOM] Raman[NOM] ANAPH[NOM] prize-ACC
ʤej-tt-**aan**-nnŭ] so-nn-aan-nnŭ] kaɳḍŭpiḍi-tt-aan.
win-PST-3MSG-COMP say-PST-3MSG-COMP discover-PST-3MSG
"Krishnan$_i$ discovered [$_{CP}$ that Raman$_j$ said [$_{CP}$ that he$_i$ won the prize."]]

The only difference between the sentences in (409) and (410) is the choice of antecedent for *ta(a)n*, which in turn drives the difference in agreement on the clausemate verb under *ta(a)n*. In (409), the antecedent of *ta(a)n* is the AGENT of the immediately superordinate speech predicate *Raman*; as we have come to expect by now, the clausemate verb of *ta(a)n* in embedded subject position shows monstrous (1SG) agreement. In (410), the antecedent of *ta(a)n* is the matrix subject *Krishnan* which is not the AGENT of a speech predicate (and even if it were, it is not in the clause that directly selects the *ta(a)n*-clause). As such, the agreement under *ta(a)n*, which still tracks the antecedent, shows non-monstrous agreement whose ϕ-features match those of *Krishnan*, as expected.

But if we take a closer look at (410), we see that the DPs in [Spec, Speech-ActP] and [Spec, PerspP] denote different entities in the evaluation context. In (410), the DP in [Spec, Speech-ActP] is obligatorily

controlled by the AGENT of the selecting speech predicate which denotes Raman. However, the antecedent of *ta(a)n* in (410) is not *Raman*, but *Krishnan*; the DEP-value of the operator in [Spec, Persp-P] will thus end up denoting Krishnan.

12.3.3 The ϕ-featural relationship between PerspP, *ta(a)n* and T

In the previous chapter, we proposed that the relationship between anaphora and agreement be modelled as follows:

(411) THE RELATIONSHIP BETWEEN ANAPHORA AND AGREEMENT:
- i. The agreement under *ta(a)n* is not directly triggered by *ta(a)n*. This is an instantiation of the AAE.
- ii. All agreement under *ta(a)n* is originally triggered by the pronominal operator in the local [Spec, PerspP] which binds *ta(a)n*. This accounts for the antecedent-tracking property of monstrous and non-monstrous agreement alike.
- iii. In cases of non-monstrous agreement, the pronominal operator in [Spec, PerspP] is unshifted with respect to the utterance context.
- iv. In cases of monstrous agreement, the pronominal operator in [Spec, PerspP] is a shifted 1st-person indexical with respect to the utterance context; i.e. it is interpreted as 1st-person with respect to the intensional context introduced by the immediately superordinate speech predicate.

The pronominal operator in [Spec, PerspP], *ta(a)n*, and the T head under *ta(a)n* have the following set of features:

Operator in [Spec, PerspP]: This is a deictic pronoun. It has a valued DEP-feature, and a full set of valued ϕ-features.

***ta(a)n*:** *ta(a)n* is syntactically and semantically an anaphor. It has an unvalued set of ϕ-features, and an unvalued DEP-feature. It is thus featurally "minimal" in the sense of Kratzer (2009).

T: The T head has an unvalued set of ϕ-features.

The operator in [Spec, PerspP] enters into a syntactic Agree relationship with *ta(a)n* for the DEP-feature. As discussed in detail in Parts I and II, this will be interpreted as a binding relationship between the two elements at LF: the operator will bind *ta(a)n* since it asymmetrically

c-commands the latter. The clausemate T head of *ta(a)n* will probe to get its ϕ-features valued. Tamil manifests subject-verb agreement thus T will try to get its features checked against *ta(a)n*, in sentences where the latter is the syntactic subject. However, *ta(a)n* also has an unvalued set of ϕ-features thus will not be able to value these features on T. The two will thus enter into a feature-sharing relationship (Pesetsky and Torrego 2007) for the relevant features which will continue to probe and ultimately get valued by the operator in [Spec, PerspP]. This is the basic scenario for deriving monstrous as well as non-monstrous agreement under subject *ta(a)n*.

Let us now look at the specific details involved in deriving monstrous agreement in this model, as in a sentence like (412):

(412) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ ʤej-pp-een-nnŭ] so-nn-aan.
Raman[NOM] ANAPH[NOM]$_i$ win-FUT-1SG-COMP say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that he$_{\{i,*j\}}$ would win]"

For monstrous agreement to obtain, two conditions must simultaneously hold: first, the operator in [Spec, PerspP] must be born with a 1st-person feature denoting the speaker of the context of evaluation; second, the context of evaluation must be the intensional context introduced by a superordinate speech predicate, and not the utterance context. The operator will value the DEP-feature on *ta(a)n* and bind it at LF. The value of this DEP-feature will be mapped onto the speaker of the intensional context, due to the presuppositional restriction induced by the 1st-person feature on the operator. The DP that linguistically represents this speaker, namely the AGENT of the superordinate speech predicate, will thus be construed as the antecedent of *ta(a)n*. In the syntax, the operator in [Spec, PerspP] will also have straightforwardly valued ϕ-feature agreement on T and *ta(a)n* as 1st person, but the agreement on T will look "monstrous" because it will track the antecedent of *ta(a)n*, namely the DP denoting the speaker of the intensional context.

When there is no Speech-ActP selected in the *ta(a)n*-clause and thus no representation of an intensional-context, the person feature on the operator in [Spec, PerspP] must be interpreted relative to the utterance context (as this is the only context available). If this operator is born with a 1st-person feature, it will thus be interpreted at LF as the speaker of the utterance context. It will then value the ϕ-features on both *ta(a)n* and T as 1st-person and the DEP-feature on *ta(a)n*, just as before. These Agree operations will go through mechanistically and unproblematically. However, the derivation will crash at LF, specifically at the point where the DEP-values on both *ta(a)n* and the operator in [Spec, PerspP] are

mapped onto the speaker of the utterance context. This, I will propose, is because *ta(a)n* is inherently specified not to refer to a participant of the utterance context: this is assumed to be a part of its denotation. I will return to a fuller discussion of the nature of this restriction in the next section.

What happens in cases where we have 3rd-person agreement in the embedded *ta(a)n*-clause, as in (413) below?

(413) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ ʤej-pp-aan-nnŭ] so-nn-aan.
Raman[NOM] ANAPH[NOM]$_i$ win-FUT-3MSG-COMP say-PST-3MSG
"Raman$_i$ said [$_{CP}$ that he$_{\{i,*j\}}$ would win.]"

This, of course, has to do with what 3rd-person means. Here it is important to distinguish the featural representation of 3rd-person in the syntax and the potential use of underspecification in the spell-out of 3rd-person *forms*, from the semantic interpretation of 3rd-person. While there is a lively debate with respect to the former (Harley and Ritter 2002, Bobaljik 2008), there is general consensus about the latter. The standard assumption is that 3rd-person denotes an entity that is neither the speaker nor the hearer of the context of evaluation.[4]

Consider again the sentence in (413). The agreement under *ta(a)n* tracks the antecedent of *ta(a)n* – namely the AGENT of the matrix speech predicate, *Raman*. Since the *ta(a)n*-clause is selected by this speech predicate, it contains a Speech-ActP which hosts the intensional context introduced by this verb, including the information that *Raman* denotes its speaker. The 3rd-person agreement under *ta(a)n* is triggered by the operator in [Spec, PerspP] – thus, this operator must itself be born with a 3rd-person feature. Under the intended reading, *Raman* is the antecedent of *ta(a)n*, which means that the DEP-feature on the operator and *ta(a)n* must have been mapped onto Raman in the context of evaluation. But this leads to a peculiar state of affairs. Specifically: how could the operator refer to the speaker of the intensional context if it is, as we have argued, born with a 3rd-person feature and 3rd-person explicitly denotes an entity that is *not* the speaker (or hearer) of the evaluation context?

The only reasonable explanation, it seems, must be that the context of evaluation in this case is not the intensional context, but the context of utterance. The operator *is* in fact 3rd-person with respect to the utterance-context. I.e. *Raman* is, indeed, neither the speaker nor the

[4]Thanks to Thomas McFadden and Peter Svenonius for clarifying this distinction with me.

hearer of the utterance context. But this conception of things entails that the operator in [Spec, PerspP] itself doesn't "care" what its context of evaluation is. This might seem to be a controversial assumption under a context overwriting analysis like that of Anand (2006) where it is assumed that there is only one linguistically represented contender for context of evaluation. Under the current approach, however, we have been consistent in claiming that the context of utterance is really a different sort of entity from the intensional context.

Empirical evidence in favor of this position comes from the fact that, even in domains where there is a locally available intensional context, certain elements like *naan* (I) or *nii* (YOU) seem to be able to "pop out" and be interpreted relative to the context of utterance:

(414) Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ enn-æ$_{Auth_{utt}}$ paar-tt-een-nnŭ] so-nn-aa̱.
Maya ANAPH[NOM] me-ACC see-PST-1SG-COMP say-PST-3FSG
"Maya$_i$ said [$_{CP}$ that she$_{\{i,*j\}}$ had seen me$_{Auth_{utt}}$.]."

In (414), the matrix speech predicate induces monstrous agreement under *ta(a)n* in its clausal complement. Thus, the intensional context pertaining to this verb must be represented in the embedded clause. However, the direct object *ennæ* (ME) in this clause must still be interpreted as 1st-person relative to the context of *utterance* – i.e. it refers to me, Sandhya. This type of evidence shows two things. It shows, first, that certain indexicals are prespecified to be interpreted relative to the context of utterance. Second, and relatedly, it illustrates that the utterance context may never be fully overwritten – it's always linguistically present at some level, although it may not be syntactically represented like the intensional context is.

I will assume therefore that the operator in [Spec, PerspP] is truly underspecified with respect to its context of evaluation and that this is thus determined only at LF. For the sentence in (413) to go through, we need the operator (and by extension *ta(a)n*) to denote Raman while still maintaining the presupposition, contributed by its 3rd-person feature, that it is *not* the speaker of the context of evaluation. Thus, (413) can only be consistently interpreted if the context of evaluation is the context of utterance.

But we must now iron out another potential wrinkle. We have already seen how monstrous agreement is derived in sentences like (412). The details of this derivation remain unchanged. However, given our new assumption that the operator in [Spec, PerspP] is underspecified with respect to its context of evaluation (at LF), we must ensure that it *isn't* interpreted relative to the context of utterance when it is born with a

1st-person feature. If it were interpreted against the utterance context, it would denote the speaker of this context resulting, in turn, in *ta(a)n* also denoting this entity. However, *ta(a)n* may independently not refer to the speaker (or hearer) of the utterance context, as we have seen. This, I propose below, is encoded as a semantic restriction in the denotation of *ta(a)n*. Therefore, when the operator in [Spec, PerspP] is born with a 1st-person feature, it will be forced to be evaluated against the intensional context – yielding the monstrous agreement scenario in (412).

In the section below, we will discuss more explicitly how to model this restriction on *ta(a)n*.

12.3.4 Deriving the 3rd-person antecedence restriction on *ta(a)n*

The indexical pronouns *naan* and *nii* are always interpreted deictically, as the speaker and hearer, respectively, of the utterance context. Thus, even in clauses that are embedded under *soll* (SAY), these proforms may not refer to the speaker or hearer coordinates (respectively) of the minimal intensional context introduced by this speech predicate (unless they are interpreted as quoted strings):

(415) Raman [$_{CP}$ naan paris-æ ɖʒej-čč-een-nnŭ]
Raman[NOM] I[NOM] prize-ACC win-PST-1SG-COMP
so-nn-aan.
say-PST-3MSG
"Raman said [$_{CP}$ that I had won the prize.]"
$[\![I]\!]^{c,g}$ = Speaker of the utterance context

(416) Raman Krishnan-kiʈʈæ [$_{CP}$ nii paris-æ
Raman[NOM] Krishnan-ALL you[NOM] prize-ACC
ɖʒej-čč-aaj-ŭnnŭ] so-nn-aan.
win-PST-2SG-COMP say-PST-3MSG
"Raman told Krishnan [$_{CP}$ that you had won the prize.]"
$[\![you]\!]^{c,g}$ = Hearer of the utterance context.

We have observed right from the beginning of the dissertation that *ta(a)n* may not be anteceded the speaker or hearer of the utterance context – given (415) and (416) above, this means that it cannot be anteceded by *naan* or *nii*:

(417) * Naan$_i$ [$_{CP}$ taan$_i$ school-ǔkkǔ poo-r-een-nnǔ]
I[NOM] ANAPH[NOM] school-DAT go-PRS-1SG-COMP
so-nn-een.
say-PST-1SG
"I$_i$ said [$_{CP}$ that I$_i$ am going to school.]" (Intended)

(418) * Nii$_i$ [$_{CP}$ taan$_i$ school-ǔkkǔ poo-r-een/aaj-nnǔ]
you[NOM] ANAPH[NOM] school-DAT go-PRS-1SG/2SG-COMP
so-nn-aaj.
say-PST-2SG
"You$_i$ said [$_{CP}$ that you$_i$ are going to school.]" (Intended)

We have just proposed that this restriction is because *ta(a)n* is inherently specified to *not* denote a participant of the utterance context but haven't really explained it in detail. Now that we have our feature-specifications on all the key players fully determined, we are in a position to do so.

Given our analysis of monstrous agreement structures as involving a 1st-person shifted indexical binding *ta(a)n*, we cannot simply say that *ta(a)n* is itself specified 3rd-person or is incompatible with 1st-person in general, claiming in other words, that 1st-person as evaluated against an intensional context is licit, but that 1st-person as evaluated against the utterance context is not. The antecedence restriction on *ta(a)n* thus cannot be modelled as a restriction against a particular value for person but must take into account the context of evaluation for that person feature as well. The same point can be made with respect to 2nd person as well.

Note that it is not especially stipulative or otherwise odd to state that *ta(a)n* is negatively specified with respect to the utterance context.[5] After all, we know that there are elements in natural language –

[5]At first glance, it seems like this is an unnecessarily strong claim. Peter Svenonius (p.c.) suggests that one might try to get the same result from spell-out effects. For instance, we might propose that an anaphor bound by a 1st-person element will always be spelled out as *naan*, rather than *ta(a)n*, because the former is more highly specified featurally. However, in precisely the monstrous agreement cases, we have a 1st-person element in [Spec, PerspP] binding the anaphor and yet, this anaphor is spelled out as *ta(a)n*, not *naan*. One might try to get out of this quandary by proposing that *naan* can only spell out a 1st-person feature that is evaluated against the utterance context (regardless of how this might be implemented). However, this would have to be further supplemented by a condition stating that other 1st-person features can only be evaluated against non-utterance contexts. In other words, a 1st-person feature that is underspecified with respect to its context of evaluation cannot exist. Furthermore, both kinds of 1st-person feature would have to trigger the same agreement, thus cannot be made to be fully disjoint. Such a solution is thus technically possible but, by this point, all its initial promises of elegance and simplicity are long

the so-called "well behaved" Kaplanian indexicals like *naan* and *nii* in Tamil and their counterparts in English – that are *positively* specified with respect to the utterance context. This shows that the utterance context is in some sense always "present" – thus can be made reference to – though it may well not be syntactically represented the same way as the intensional context is. Indeed, robust evidence from anaphoric patterns crosslinguistically suggests that many (perhaps most) anaphors are negatively specified in the same way as *ta(a)n*.

Note, incidentally, that while my analysis of *ta(a)n* may be initially reminiscent of a Schlenkerian one (Schlenker 2003b) in its use of negative specification for the utterance context, there is a crucial difference. Under Schlenker's analysis *ta(a)n* would be specified to denote a participant (specifically the speaker) of a non-utterance context. I am, however, claiming that *ta(a)n* is specified not to refer to a participant of the utterance context. There is an important difference in the scope of the negation in each, which may be modelled as below:

(419) **Schlenker's version:**
$[\![taan]\!]^{c,g} = Participant(\neg UtteranceContext)$

(420) **My version:**
$[\![taan]\!]^{c,g} = \neg Participant(UtteranceContext)$

In other words, under Schlenker's system, *ta(a)n* would be treated as an obligatorily shifted *indexical* (since an indexical, defined in the Kaplanian sense, is a participant of a Kaplanian context). I.e. it must be a participant but not of the utterance context. For me, however, there is no requirement that it be a participant at all, even of the intensional context. And in fact, it may be explicitly specified not to be a participant of the utterance context.

The logical space where the two approaches make distinct predictions is in structures where the antecedent of *ta(a)n* is a non-participant of the intensional context or if, indeed, there is no intensional context represented in the local domain of *ta(a)n*. Schlenker's system predicts that *ta(a)n* should be infelicitous in such cases whereas my analysis predicts that it should be licit (as long as independent conditions on its reference are met, of course). The sentences below show, however, that *ta(a)n* is perfectly licit in such environments:

---

gone.

(421) Seetha$_i$ [$_{CP}$ Raman$_j$ [$_{CP}$ taan$_i$ kaḷḷ-æ kuḍi-
Seetha[NOM] Raman[NOM] ANAPH[NOM] toddy-ACC drink-
čč-aaḷ-ŭnnŭ] Krishnan-kiṭṭæ so-nn-aapoṛŭdŭ] tuuŋg-in-aaḷ.
PST-3FSG-COMP Krishnan-ALL told-WHEN sleep-PST-3FSG
"Seetha$_i$ slept [$_{CP}$ while Raman$_j$ said [$_{CP}$ that she$_i$ drank the toddy.]]"

(422) Seetha$_i$ [$_{CP}$ taan$_i$ kaḷḷ-æ kuḍi-čč-adŭ-
Seetha[NOM] ANAPH[NOM] toddy-ACC drink-PST-3NSG-
naalæ] tuuŋg-in-aaḷ.
BECAUSE sleep-PST-3FSG
"Seetha$_i$ slept [$_{CP}$ because she had drunk the toddy.]"

In (421), we see long-distance binding of subject *ta(a)n* across a temporal CP adjunct. However, this adjunct contains a speech predicate which then selects the innermost CP containing *ta(a)n*. In other words, the CP containing *ta(a)n* must have a Speech-ActP which holds information pertaining to the intensional context associated with *soll*. Both the speaker and hearer of this intensional context are represented in the intermediary clause – by *Raman* and *Krishnan*, respectively. However, *ta(a)n* doesn't refer to either in this sentence – it refers instead to *Seetha* who is not a participant of any context (utterance or intensional). In (422), there is no intensional context represented in the *ta(a)n* clause at all. The embedded CP containing the anaphor is a causal adjunct and there is no intensional predicate in the entire sentence. Here again, *ta(a)n* is perfectly licit – anteceded by *Seetha* which doesn't denote a participant of any context.

I will assume therefore that the analysis that *ta(a)n* is negatively specified with respect to the utterance context is the correct one. As for the implementation, one could, for instance, attempt to do it in the syntax, via the presence of a (semantically) negative feature on *ta(a)n* – see for instance Harbour (2011) for independent arguments in favor of negatively specified syntactic features. But this seems too strong, especially since there is no reason to assume that the utterance context is itself syntactically represented – its relevance doesn't seem to be subject to Minimality restrictions, for instance. I will therefore assume, simply, that the negative specification for the utterance context is encoded in the denotation of *ta(a)n* and filters out unwanted mappings by the assignment function, at LF (as described in the previous section).

This completes our discussion of the formal featural properties of the key players in Tamil monstrous and non-monstrous agreement paradigms under subject *ta(a)n*. Table (12.1) provides a sample of possible feature-

Table 12.1: SAMPLE FEATURE REPRESENTATIONS OF KEY ELEMENTS

ELEMENT	FEATURE-BUNDLE
Spec, Speech-ActP	D
Spec, PerspP	[DEP: *x*, P: 3, NUM: *sg*, G: F, D]
Anaphor (*ta(a)n*)	[DEP: ___, P: ___, NUM: ___, G: ___, D]
T	[P: ___, NUM: ___, G: ___]
Antecedent (*Raman*)	[P: *3*, NUM: *sg*, G: *m*, D]

specifications for the main linguistic elements in our analysis. We are now ready to proceed to a step-by-step derivation of the monstrous and non-monstrous agreement structures in Tamil.

12.3.5 Step-by-step derivation of monstrous agreement

Consider the sentence below (repeated from (409)):

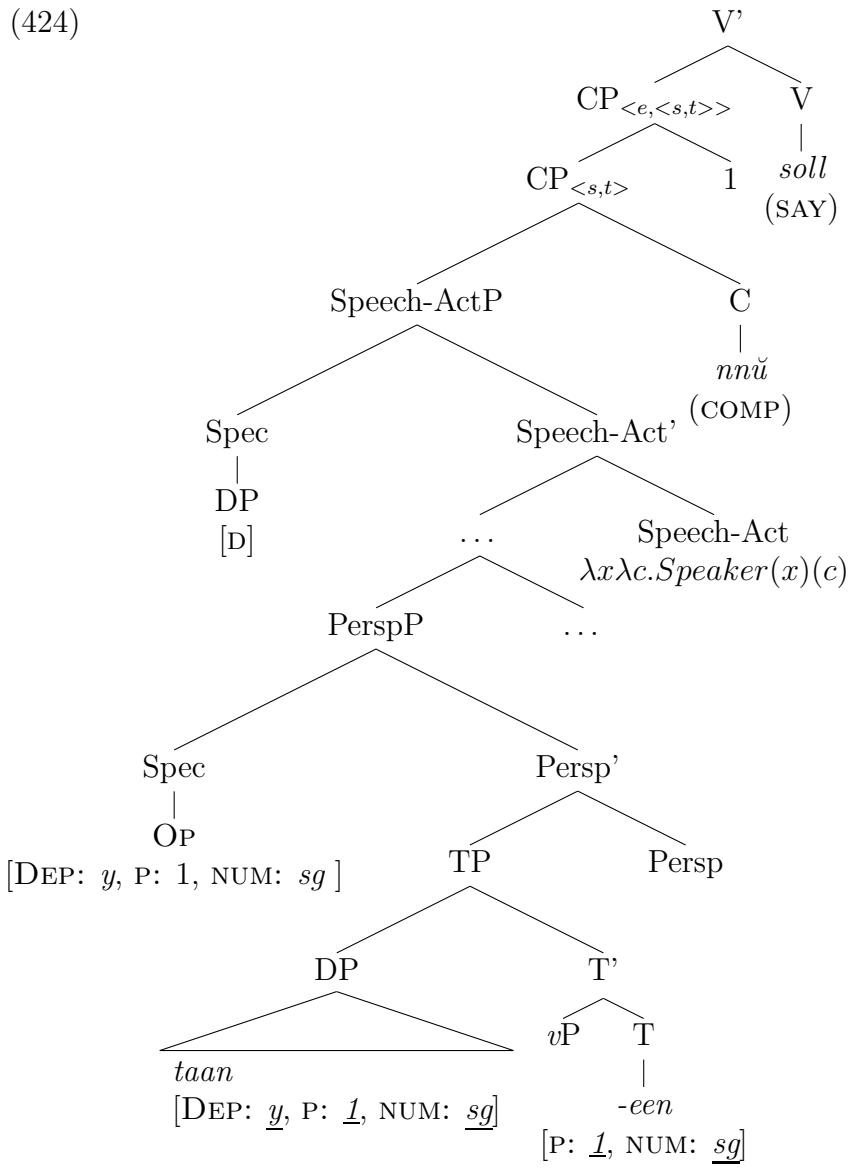
(423) Krishnan$_i$ [$_{CP}$ Raman$_j$ [$_{CP}$ taan$_i$ paris-æ
Krishnan[NOM] Raman[NOM] ANAPH[NOM] prize-ACC
ʤej-čč-**een**-nnŭ] so-nn-aan-nnŭ] kaɳḍŭpiḍi-tt-aan.
win-PST-1SG-COMP say-PST-3MSG-COMP saw-PST-3MSG
"Krishnan$_i$ saw [$_{CP}$ that Raman$_j$ said [$_{CP}$ that he$_j$ won the prize.]]"

The antecedent of *ta(a)n* is *Raman* which is the AGENT of the speech predicate that selects the *ta(a)n*-clause. The agreement tracks this antecedent and is marked 1SG – in other words, the agreement is "monstrous".

Let us now derive this paradigm. As usual, the only clause where any syntactic "action" happens is the clause containing *ta(a)n*. This clause has the following tree structure after Agree and before being shipped off to LF and PF:[6]

[6] As before, I underline features that have been valued as a result of Agree – this is, of course, purely a visual mnemonic for ease of understanding and should not be interpreted as a featural diacritic in the syntax.

(424)



The Agree relation between *ta(a)n* and the operator in [Spec, PerspP] is exactly as described for various long-distance binding structures in Parts I and II. *ta(a)n* probes up its phase to get its DEP feature valued. The closest Goal with a valued DEP feature is the operator in [Spec, PerspP] which values the DEP-feature on *ta(a)n* as [DEP: *y*]. In the meantime, the embedded T head starts out with unvalued ϕ-features and cannot get these valued by its clausemate subject because the ϕ-features on *ta(a)n* are not valued either. *ta(a)n* and T enter into a feature-sharing relationship, and keep probing to get their ϕ-features valued. They get these features valued by the pronominal operator in [Spec, PersP] which

is the minimally closest element with valued ϕ-features. This operator *happens* to be born with 1SG ϕ-features in this case, so it values the ϕ-features on T as such. At PF, these agreement features on T get spelled out as *-een*. At LF, the matching DEP-values on *ta(a)n* and the operator in [Spec, PerspP] will result in the latter binding the former (by virtue of asymmetrically c-commanding it). The assignment function g will ultimately treat these two linguistic elements as a unit for the purposes of reference assignment because it will simply map their common DEP-value onto a single entity in the evaluation context.

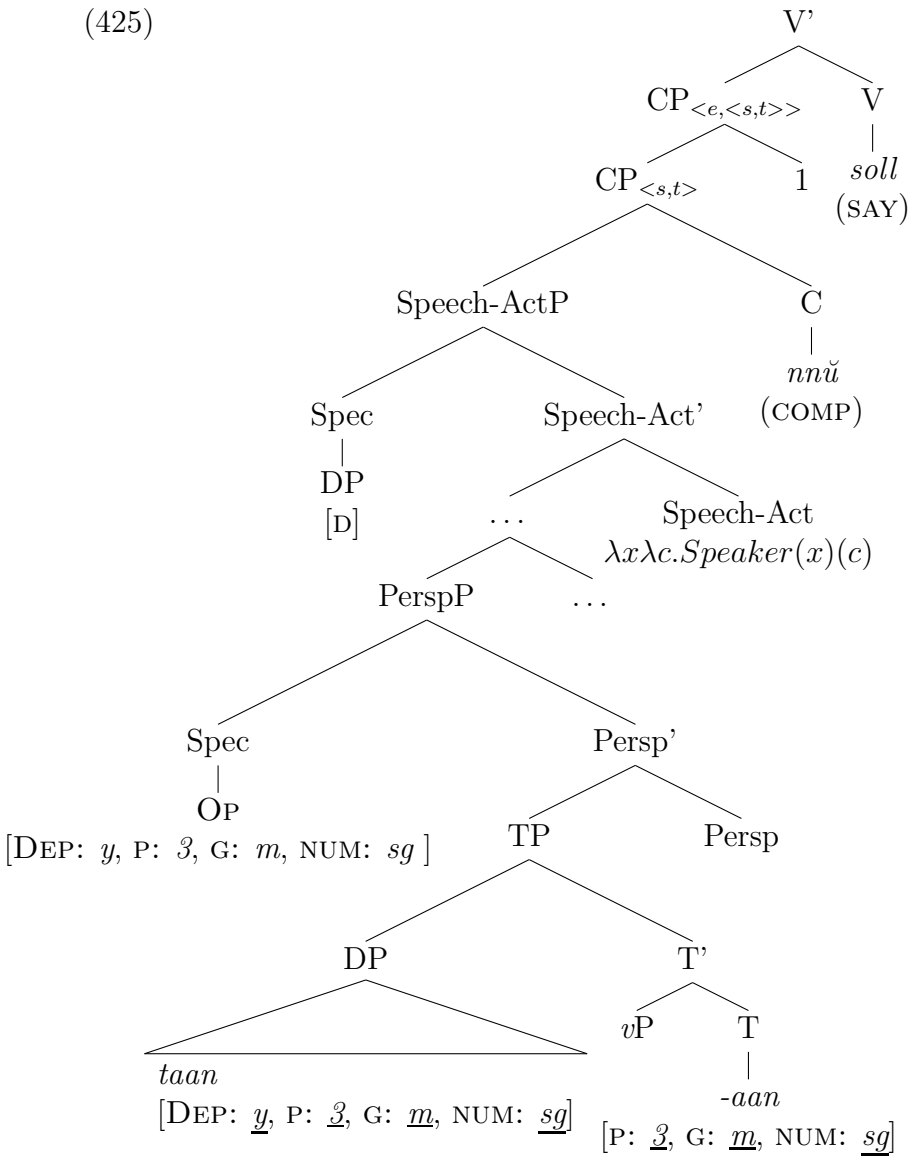
How do the two linguistic elements end up getting mapped onto Raman, rather than Krishnan in (423)? This is where the DP in [Spec, Speech-ActP] enters the picture. Recall that this DP is obligatorily controlled by the AGENT of the selecting speech predicate, thus must denote Raman. Furthermore, Raman will be identified as the speaker of the intensional context introduced by this verb and represented in Speech-ActP. The 1st-person feature on the operator in [Spec, PerspP] will introduce the presupposition that it must denote the speaker of its context of evaluation. As we have seen, the operator itself is underspecified with respect to which context it is interpreted against – thus, it may in theory be evaluated against the utterance context or the intensional one. If the former, it will denote the speaker of the utterance context. Since *ta(a)n* has the same DEP-value this means that *ta(a)n* will also denote this entity. However, at this point the derivation will crash because *ta(a)n* is explicitly specified not to refer to the speaker (or hearer) of the utterance context. The only choice is that the operator and *ta(a)n* denote the speaker of the intensional context. This indeed is what happens – the 1st-person agreement under *ta(a)n* which tracks this antecedent thus looks "monstrous".[7]

12.3.6 Step-by-step derivation of non-monstrous agreement

What if the operator in [Spec, Persp-P] didn't happen to have 1SG ϕ-features, but something else, like 3MSG, as in (425) below?

[7] In other words, agreement, being a purely morphosyntactic phenomenon, is itself blind to the interpretive properties of 1st-person, such as which context it is evaluated against. Thus, there is nothing monstrous about the agreement per se – this is just a surface instantiation of the indexical shift that happens elsewhere in the clause.

(425)



The tree structure above represents the embedded CP of the following sentence:

(426) Krishnan$_i$ [$_{CP}$ Raman$_j$ [$_{CP}$ taan$_{\{i,j\}}$ paris-æ
Krishnan[NOM] Raman[NOM] ANAPH[NOM] prize-ACC
ʤej-čč-**aan**-nnŭ] so-nn-aan-nnŭ] kaɳḍŭpiḍi-tt-aan.
win-PST-3MSG-COMP say-PST-3MSG-COMP saw-PST-3MSG
"Krishnan$_i$ saw [$_{CP}$ that Raman$_j$ said [$_{CP}$ that he$_{\{i,j\}}$ won the prize.]]"

As (426) shows, the antecedent of *ta(a)n* is actually ambiguous in this case, since both *Krishnan* and *Raman* are potential antecedents. As far as

the syntax is concerned, there is no difference in the derivation, regardless of which DP the intended antecedent is. This is because both *Raman* and *Krishnan* have the same ϕ-features – thus the agreement under *ta(a)n* will be 3MSG regardless of which one ultimately gets chosen as the actual antecedent.

The syntactic derivation proceeds in the following manner. The operator in [Spec, PerspP] is born with the features 3MSG. The embedded T head and *ta(a)n* have entered into a feature-sharing relationship for their unvalued ϕ-features, just as described for the monstrous agreement structure above. The operator simultaneously values these ϕ-features on both *ta(a)n* and embedded T. At PF, this 3MSG agreement on T is spelled out as: *-aan*. *ta(a)n* is also a probe for its unvalued DEP-feature: this also gets valued by the DEP-feature on the operator in [Spec, PerspP], as [DEP: $\underline{y}$]. At LF, the operator binds *ta(a)n*, thus ensuring that the two elements denote the same entity in the evaluation context. As far as the operator is concerned, the evaluation context may be either the context of utterance or the intensional context associated with the immediately superordinate speech predicate.

If the intensional context is picked, the 3rd-person feature on the operator will introduce the presupposition that it does not denote the speaker or hearer of this context. This means that the DEP-value on the operator and *ta(a)n* may not denote Raman who is the speaker of the intensional context. Thus, the only alternative is that [DEP: y] is mapped to Krishnan. In this case, we get the effect of linguistic antecedence by the matrix subject *Krishnan*. If the utterance context is chosen as the context of evaluation, however, the 3rd-person feature on the operator and *ta(a)n* will introduce the presuppositional constraint that they may not denote the speaker or hearer of the utterance context. In this case, both Raman and Krishnan qualify, because neither is the speaker (or hearer) of the utterance context. In this case, speaker-intent and other discourse-pragmatic factors must decide which individual gets picked. If Krishnan gets chosen, we have the effect of antecedence by *Krishnan*; if Raman gets chosen, we have the effect of linguistic antecedence by the intermediate subject *Raman*.

12.3.7 Step-by-step derivation of deictic reference under SAY

We saw earlier that certain elements remain stubbornly indexical in the Kaplanian sense even when there are quantified over by an intensional context. We have seen that *naan* (*I*) and *nii* (*you*) are two such elements

in Tamil, as illustrated by the sentences below (repeated from (415) and (416)):

(427) Raman [$_{CP}$ naan paris-æ ʤej-čč-een-nnŭ] so-nn-aan.
Raman[NOM] I[NOM] prize-ACC win-PST-1SG-COMP say-PST-3MSG
"Raman said [$_{CP}$ that I had won the prize.]"
$[\![I]\!]^{c,g}$ = Speaker of the utterance context

(428) Raman Krishnan-kiʈʈæ [$_{CP}$ nii paris-æ ʤej-čč-aaj-ŭnnŭ] so-nn-aan.
Raman[NOM] Krishnan-ALL you[NOM] prize-ACC win-PST-2SG-COMP say-PST-3MSG
"Raman told Krishnan [$_{CP}$ that you had won the prize.]"
$[\![you]\!]^{c,g}$ = Hearer of the utterance context.

To derive this, I propose that *naan* and *nii* are prespecified to be evaluated only against the context of utterance – just as in Schlenker (2003b). I.e. I posit that these elements have the following Schlenkerian denotations for rigid Kaplanian indexicals:[8]

(429) **Denotations for *naan* and *nii*:**

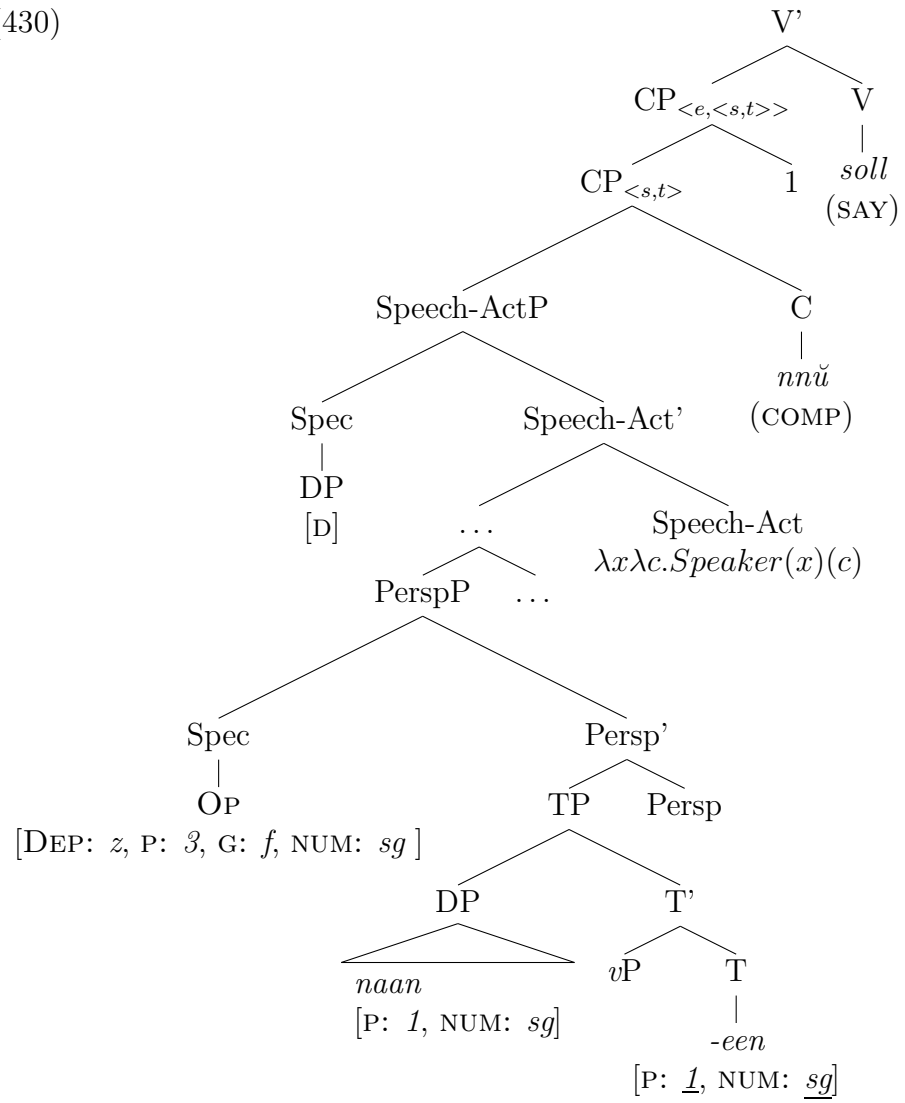
Rigid 1st-person indexical:
$[\![naan]\!]^{c,g} = [\![x_i + Speaker * (x_i)]\!]^{c,g} = \mathrm{s}(\mathrm{x}_i) \wedge \delta(\mathrm{g}(\mathrm{x}_i)$ is Speaker(c*)).
for c* = utterance context;
Speaker* = Speaker(UtteranceContext)

Rigid 2nd-person indexical:
$[\![nii]\!]^{c,g} = [\![x_i + Hearer * (x_i)]\!]^{c,g} = \mathrm{s}(\mathrm{x}_i) \wedge \delta(\mathrm{g}(\mathrm{x}_i)$ is Hearer(c*)).
for c* = utterance context;
Hearer* = Hearer(UtteranceContext)

The embedded CP in (427) has the following tree structure:

[8] Schlenker uses the terms *Author* and *Addressee*. I have changed these to *Speaker* and *Hearer* for reasons of consistency with the previous discussion. Nothing crucial rests on this change, however.

(430)



In the syntax, the T head starts out with unvalued ϕ-features as usual and probes towards its clausemate subject to get these features valued. However, in this case, this subject is not the anaphor *ta(a)n* which also has unvalued ϕ-features but the deictic 1st-person indexical *naan* which *does* have inherently specified ϕ-features, as indicated. This subject will straightforwardly value the ϕ-features on this T head as 1SG which gets spelled out as *-een*, at PF. The T head doesn't "see" the ϕ-features on the operator in [Spec, PerspP] at all – so, the ϕ-features on DEP may be different from those on T, just as indicated. The CP in (430) has a Speech-ActP due to the selectional properties of the immediately superordinate speech predicate *soll* (SAY); thus, the intensional context is syntactically represented in this CP. The operator in [Spec, PerspP]

may thus be evaluated against the utterance context or the intensional one, just as described before. Its DEP-value will be mapped onto one of the individuals of this evaluation context listed in the range of the assignment function at LF; this is the individual whose perspective the embedded CP will report. The DP that linguistically represents this individual must have 3FSG features as well. However, none of this is overtly represented in the sentence above because there is no anaphor that needs to be bound.

At LF, the denotation of *naan* will be as given under (429) above. Thus, although the intensional context is available as a potential context of evaluation, *naan* may not be interpreted against it. *naan* is evaluated against the utterance context instead (which, as we have consistently claimed, is special in that it is available at LF) and thus deictically denotes the speaker of the utterance context, i.e. me, Sandhya. The moral of this story is that the intensional context is only available as a potential context of evaluation in specialized syntactico-semantic environments, but the utterance context is always available – although, unlike the intensional context, it is probably not syntactically represented, thus is presumably accessible only at LF.

12.4 Default agreement under *ta(a)n* in intensional contexts

There is one potential loose end that must be tied up. Consider the sentence below:

(431) Maya [$_{CP}$ tana-kkŭ Krishnan-æ piḍi-tt-adŭ-nnŭ]
Maya[NOM] ANAPH-DAT Krishnan-ACC like-PST-3NSG-COMP
so-nn-aa[.
say-PST-3FSG
"Maya$_i$ said [$_{CP}$ that she$_{\{i,*j\}}$ liked Krishnan]."

The embedded CP containing *ta(a)n* in subject position is evaluated against the intensional context introduced by the matrix speech predicate *soll* (SAY). *ta(a)n* itself is anteceded by the AGENT of this immediately superordinate speech predicate. And yet, the agreement under *ta(a)n* is marked 3NSG. This is default agreement – as shown by its invariance under plurally marked *ta(a)n* (432); furthermore, the agreement may not be monstrous, as shown in (433):[9]

[9]The agreement may not be 3FSG either, to match those on *Maya*.

(432) Pasaŋ-gaɭ [$_{CP}$ taŋ-gaɭ-ŭkkŭ Krishnan-æ piɖi-tt-adŭ-nnŭ] so-nn-aaŋ-gaɭ.
boys-PL.NOM ANAPH-PL-DAT Krishnan-ACC like-PST-3NSG-COMP say-PST-3PL
"The boys$_i$ said [$_{CP}$ that they$_{\{i,*j\}}$ liked Krishnan]."

(433) * Maya [$_{CP}$ tana-kkŭ Krishnan-æ piɖi-tt-een-nnŭ] so-nn-aaɭ.
Maya[NOM] ANAPH-DAT Krishnan-ACC like-PST-1SG-COMP say-PST-3FSG
"Maya$_i$ said [$_{CP}$ that she$_{\{i,*j\}}$ liked Krishnan]." (Intended)

How do we account for this? First of all, notice that *piɖi* (LIKE) in Tamil always takes a "quirky" dative subject (431)– a point that was discussed independently in the context of the *koɭ*-discussion in Part II. Furthermore, verbs with quirky dative subjects in Tamil always trigger default 3NSG agreement on the verb, independent of whether they are embedded under a speech predicate. It is also independent of whether their subject is an anaphor or not, thus is not an instantiation of the AAE:

(434) Seetha-vŭkkŭ Krishnan-æ piɖi-tt-adŭ.
Seetha-DAT Krishnan-ACC like-PST-3NSG
"Seetha liked Krishnan."

These facts show that the default agreement on the verb is really a different type of agreement. It seems to be a direct and exhaustive function of whether the subject is marked dative or not. A potential reason for why dative-marked subjects don't trigger normal agreement might have to do with them being externally merged higher in the clausal structure (Adger and Ramchand 2005). Recall that we suggested precisely this in our discussion of EXPERIENCER dative subjects in psych-predicate constructions, in Part II. How default agreement is precisely derived is a matter of lively debate in the literature: with respect to the current discussion, what matters is that having default agreement under a speech-predicate in Tamil is not problematic for the analysis being developed here.

In fact, this type of data supports a particular point of the analysis, namely the idea that the T head under subject *ta(a)n* doesn't directly Agree for ϕ-features with the operator in [Spec, PerspP] – rather, it is mediated by *ta(a)n* which must thus itself have ϕ-featural attributes. If there were a direct relationship between embedded T and the operator in [Spec, PerspP] – there would be no reason to expect monstrous agreement

to fail in a sentence like (431). The fact that monstrous agreement does fail, however, shows that the T head "sees" the dative marking on *ta(a)n*.

12.5 Deriving microvariation

The focus of this third series of chapters has been a dialect of Tamil that is spoken by a predominant subset of native speakers, myself included. But this is just one of (at least) four major dialects in Tamil. My survey of indexical shift and long-distance binding paradigms amongst native speakers of Tamil yielded the interesting result that there is systematic microvariation with respect to the availability and distribution of monstrous agreement structures. Based on standard deviation- and weighted average values for grammaticality judgments of the relevant sentences, four major dialects could be discerned. These may be organized in terms of increasing restrictiveness with respect to monstrous agreement, as follows:

Dialect 1: It is possible to get monstrous agreement under not just speech predicates like *soll* (SAY) but under certain other classes of propositional predicate, like *nene* (THINK) and *kaɳḍŭpiḍi* (DISCOVER).

Dialect 2: Monstrous agreement occurs only under speech predicates, but it is optional. When it does not obtain, the agreement on the clausemate verb of subject *ta(a)n* matches the ϕ-features on the antecedent of *ta(a)n*. This is my dialect, which has been the focus of this series of chapters.

Dialect 3: Monstrous agreement always and only obtains under a propositional speech predicate like *soll* (SAY). There is no optionality.

Dialect 4: No monstrous agreement is possible at all (for a proper subset of these speakers, no agreement obtains in embedded clauses at all).

Below, I present the broad strokes of these various micro-grammars and show that these differences may be easily accommodated within the broader analysis being developed here with just a few tweaks to the analytic details. More detailed fieldwork must be undertaken to uncover the evidence required to properly study and derive these dialectal differences. As such, the information below is to be understood in the nature of preliminary evidence on the subject.

12.5.1 Dialect 1: optional monstrous agreement under speech and (certain) non-speech verbs

For the speakers of this dialect, monstrous agreement is optionally possible, not just under propositional speech-predicates like *soll* (SAY), but also under other types of propositional predicate, like the evaluative verb *nene* (THINK), and the perception predicate *kaɳɖŭpiɖi* (FIND.OUT). Thus, for these speakers, the following monstrous agreement sentences are all grammatical:

(435) Raman$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ ʤej-pp-**een**-nnŭ]
Raman[NOM] ANAPH[NOM]$_i$ win-FUT-1SG-COMP
nene-tt-aan.
think-PST-3MSG
"Raman$_i$ thought [$_{CP}$ that he$_{\{i,*j\}}$ would win.]"

(436) Seetha$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ ʤej-pp-**een**-nnŭ] kaɳɖŭpiɖi-
Seetha[NOM] ANAPH[NOM]$_i$ win-FUT-1SG-COMP find.out-
tt-aaɭ.
PST-3FSG
"Seetha$_i$ found out [$_{CP}$ that she$_{\{i,*j\}}$ would win.]"

This type of variation can be easily encoded as a function of systematic differences in the selectional properties of these verbs across different speakers. For speakers like me who cannot (very easily) get monstrous agreement under non-speech propositional verbs (like *nene* (THINK)), we have proposed that such a verb selects a smaller complement that, in particular, does not include a projection for a Speech-ActP. Since the Speech-ActP has been argued to be the only phrase that is capable of locally hosting information pertaining to the intensional context, this in turn means that an intensional context will never be represented in the complement of such a verb.

For speakers of Dialect 1 who *do* allow monstrous agreement under non-speech verbs like *nene* (THINK) and *kaɳɖŭpiɖi* (DISCOVER), we must therefore propose that the complements of these verbs are large enough to host a Speech-ActP. As such, the operator in the specifier of Speech-ActP hosts contextual features that are associated with the immediately superordinate verbs, THINK and FIND OUT in (435) and (436), respectively. The 3MSG feature on the operator in [Spec, Persp-P] in (435 and the 3FSG feature on this operator in (436) will thus be evaluated, not against the utterance context, but relative to the contexts introduced by *nene* (THINK) and *kaɳɖŭpiɖi* (FIND.OUT), respectively. Monstrous agreement is derived in the manner described for sentences like (423)

above.

What we are claiming, in a more intuitive sense, is that THINK and DISCOVER mean something slightly different for these people – a difference that has the structural repercussions described above.

12.5.2 Dialect 2: Optional monstrous agreement only under speech predicates

This is the dialect that has been the focus of this series of chapters. In the grammar we have been describing, both monstrous and non-monstrous agreement obtain under the scope of a speech predicate like *soll* (SAY). However, only non-monstrous agreement obtains under other types of predicate:

(437) **Under speech predicates:**

a. Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ ʤej-čč-een-ŭnnŭ
Maya[NOM] ANAPH[NOM] prize-ACC win-PST-1SG-COMP
so-nn-aal]
say-PST-3FSG
"Maya$_i$ said [$_{CP}$ that she$_{\{i,*j\}}$ had won the prize]."

b. Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ ʤej-čč-aal-ŭnnŭ
Maya[NOM] ANAPH[NOM] prize-ACC win-PST-3FSG-COMP
so-nn-aal]
say-PST-3FSG
"Maya$_i$ said [$_{CP}$ that she$_{\{i,*j\}}$ had won the prize]."

(438) **Under non-speech predicates:**

a. * Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ ʤej-čč-een-ŭnnŭ
Maya[NOM] ANAPH[NOM] prize-ACC win-PST-1SG-COMP
nene-čč-aal]
think-PST-3FSG
"Maya$_i$ thought [$_{CP}$ that she$_{\{i,*j\}}$ had won the prize]."
(Intended)

b. Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ ʤej-čč-aal-
Maya[NOM] ANAPH[NOM] prize-ACC win-PST-3FSG-
ŭnnŭ nene-čč-aal]
COMP think-PST-3FSG
"Maya$_i$ thought [$_{CP}$ that she$_{\{i,*j\}}$ had won the prize]."

We have modelled this by proposing that the Speech-ActP that may host information, pertaining to the intensional context introduced by SAY, within the *ta(a)n* clause is only introduced by speech predicates.

The optionality between monstrous and non-monstrous agreement under a speech predicate obtains from the underspecification on the context of evalution for the operator in [Spec, PerspP] in the *ta(a)n*-clause which ensures that it may be interpreted relative to the context of utterance or the intensional context introduced by a selecting speech predicate.

12.5.3 Dialect 3: obligatory monstrous agreement under SAY

For speakers of this dialect, the following bi-conditional relationship between speech predicates and monstrous agreement under subject *ta(a)n*, holds:

SPEECH PREDICATE ↔ MONSTROUS AGREEMENT

In other words, for speakers of this dialect, monstrous agreement only obtains under speech predicates, just as described for Dialect 2 above. However, unlike the speakers of Dialect 2, speakers of this dialect always get monstrous agreement under speech predicates. This is illustrated below:

(439) **Under speech predicates:**

a. Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ ɬej-čč-een-ŭnnŭ so-nn-aaɭ].
Maya[NOM] ANAPH[NOM] prize-ACC win-PST-1SG-
COMP say-PST-3FSG
"Maya$_i$ said [$_{CP}$ that she$_{\{i,*j\}}$ had won the prize]."

b. * Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ ɬej-čč-aaɭ-ŭnnŭ so-nn-aaɭ]
Maya[NOM] ANAPH[NOM] prize-ACC win-PST-3FSG-
COMP say-PST-3FSG
"Maya$_i$ said [$_{CP}$ that she$_{\{i,*j\}}$ had won the prize]."
(Intended)

(440) **Under non-speech predicates:**

a. * Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ ɬej-čč-een-ŭnnŭ nene-čč-aaɭ].
Maya[NOM] ANAPH[NOM] prize-ACC win-PST-1SG-
COMP think-PST-3FSG
"Maya$_i$ thought [$_{CP}$ that she$_{\{i,*j\}}$ had won the prize]."
(Intended)

b. Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ ɬej-čč-aa]-
Maya[NOM] ANAPH[NOM] prize-ACC win-PST-3FSG-
ŭnnŭ nene-čč-aa].
COMP think-PST-3FSG
"Maya$_i$ thought [$_{CP}$ that she$_{\{i,*j\}}$ had won the prize]."

Limiting monstrous agreement to the scope of speech predicates isn't an issue of course – we would simply say, like we have for speakers of Dialect 2, that speech predicates are the only ones in this grammar that are capable of selecting a complement that is large enough to host a Speech-ActP, and, by extension, an intensional context. Monstrous agreement in this dialect would obtain when the operator in [Spec, PerspP] in the propositional complement of a speech predicate is born with a 1st-person feature and thereby denotes the speaker of the context of evaluation (which must be the intensional context in this case since *ta(a)n* may not refer to the speaker of the utterance context). Being 1st-person, it straightforwardly triggers 1st-person agreement on the embedded T head under *ta(a)n* – this agreement tracks the antecedent of *ta(a)n*, and is thus interpreted as being monstrous.

The difficulty has to do with capturing the idea that monstrous agreement is the only one that obtains in this situation. In Dialect 2, we argued that non-monstrous agreement obtains under *ta(a)n* in the scope of a speech predicate (as in (437b)) when the operator in [Spec, PerspP] is evaluated against the utterance context, instead of the intensional one. This option must be ruled out for speakers of Dialect 3. A possible way to do this would be to claim that the operator in [Spec, PerspP] must be evaluated against the minimally closest context. This will ensure that, whenever the intensional context is represented in its local Speech-ActP, it will end up being evaluated against this, and not against the utterance context. The 3rd-person agreement would then introduce the presupposition that the operator is not the speaker or the hearer of the intensional context; this would clash with the intended antecedence by the speaker of the intensional context, as in (437b), essentially ruling out such a configuration.

The difference between Dialects 2 and 3 is that, all else being equal, the former places the intensional and utterance context as equally qualified contenders for the context of evaluation. The latter, however, is sensitive to a principled distinction between the two contexts, namely that the intensional one is structurally represented in a designated position in the hierarchy whereas the latter is not. Notions of minimality may thus pertain to the former, but never the latter.

12.5.4 Dialect 4: Monstrous agreement is impossible under all verbs

For speakers with this dialect, monstrous agreement is impossible under *ta(a)n* in all structural environments:

(441) **Under speech predicates:**

a. * Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ ʤej-čč-een-ŭnnŭ
Maya[NOM] ANAPH[NOM] prize-ACC win-PST-1SG-COMP
so-nn-aa]
say-PST-3FSG
"Maya$_i$ said [$_{CP}$ that she$_{\{i,*j\}}$ had won the prize]."

(442) **Under non-speech predicates:**

a. * Maya$_i$ [$_{CP}$ taan$_{\{i,*j\}}$ paris-æ ʤej-čč-een-ŭnnŭ
Maya[NOM] ANAPH[NOM] prize-ACC win-PST-1SG-COMP
nene-čč-aa]
think-PST-3FSG
"Maya$_i$ thought [$_{CP}$ that she$_{\{i,*j\}}$ had won the prize]." (Intended)

This is similar to the case with English. Given that we have been deriving microvaration as a function of the selectional properties of propositional predicates, this in turn means that, for such speakers, the complements of such predicates *always* lack a Speech-ActP projection. In other words, an intensional-context is never represented in such clauses and all indexicals are evaluated against the utterance context – i.e. indexical shift does not obtain.

An alternative would be to claim, adopting a proposal by Ritter and Wiltschko (2009), that languages differ in terms of whether the anchoring to the utterance context is done via $Person, Time, World$ or $Location$. The dialects that allow monstrous agreement would anchor in terms of $Person$, yielding the possibility of indexical shift for person, while dialect 4, which does not allow monstrous agreement, would anchor in terms of one of the other coordinates. This would make the prediction that dialect 4 should show signs of shift along this other dimension. I discuss this option more generally as a source of crosslinguistic variation in indexical shift patterns, in Part IV.

12.6 Conclusion and outlook

We have studied an intriguing phenomenon that involves the convergence of shifted indexicality, anaphora, and verbal agreement in a single sentence. By disentangling the contributions of the various components of this sentence, we have shown that there is a crucial syntactic aspect to indexical shift which has hitherto received primarily semantic analyses. We have illustrated that speech predicates really *are* special and differ from all other types of predicate, including other types of intensional verb. By modelling this uniqueness in terms of a cartographic hierarchy of propositional complement size, we have been able to explain and predict where indexical shift obtains in Tamil. We have also had the opportunity to investigate the nature of the relationship between anaphora and agreement – a phenomenon that has not been properly understood in the literature. We have been able to show that even when *ta(a)n* is itself in subject position, it does not itself trigger agreement, even when it might seem to. This is because the agreement is in fact tracking the antecedent of *ta(a)n*, or, more precisely, the operator in the perspectival center in its phase. This has provided an independent source of evidence for the two-step binding model that has been developed in this dissertation and in particular for the syntactic step involving the operator in [Spec, PerspP].

We have also seen that a given language may have both optionally shifting and rigidly unshifting indexicals, even within the same modality (person, time, world, or location) of contextual evaluation. For instance, under the analysis proposed here, when the operator in [Spec, PerspP] is born with a 1st-person feature, it represents an indexical that is underspecified with respect to its context of evaluation. I.e., it is an optionally shifting indexical, descriptively speaking. In contrast, the 1SG indexical *naan* is a rigid indexical in the Kaplanian sense: it is specified to be evaluated relative to the utterance context. Note that this is not a strange claim to make. Shifting and unshifting indexicals are simultaneously attested in many languages. This shows that Anand (2006)'s claim that the choice of whether a particular type of indexical (person, temporal, modal, etc) is shifting or not is parametrically "set" for a given language is too restrictive: even within a particular dimension, more than one type of indexical may be empirically attested.

Some outstanding questions remain. It is, for instance, unclear how to model the difference between the utterance context and the type of rich intensional index (which we have been labelling the intensional context) that speech predicates have been seen to introduce. That there is an

ontological difference between the two is clear, however. For instance, we have seen that even in monstrous agreement structures which, we have proposed, involve the representation of an intensional context, the utterance context must, at some level, still be available. This is because both *ta(a)n* and *naan* are able to make reference to it – the former, to ensure that it does *not* refer to a Participant of the utterance context and the latter to ensure that it *does* refer to the Speaker of the utterance context. It doesn't make sense to say that the utterance context is syntactically represented since we would then expect it to interact in familiar structural ways with the intentional context (in monstrous agreement structures), leading to Minimality effects. That is, we might predict that one of the contexts will always take precedence over the other for the interpretation of all elements within a particular clause. However, such structural interference effects don't obtain – rather, monstrous agreement constructions in Tamil show that both the utterance and intensional contexts must be made available to different elements in the clause. In other words, the utterance context is grammatically available, but at a different level of the grammar (presumably at LF), whereas the intensional context is featurally represented in the Narrow Syntax – confirming the Kaplanian intuition that the type of meaning introduced by intensional operators is intrinsically different from that introduced by contextual ones.

macs

Part IV

The broader perspective

13.1 Introduction

This dissertation has explored binding from three distinct perspectives. Part I has involved deconstructing binding into its very essential parts and coming up with the building blocks for a hybrid system that would allow us to incorporate the discourse-pragmatic and less deterministic aspects of this phenomenon within a precise syntactic model. To this end, we have looked at long-distance binding where these discourse-properties, particularly perspective, are most apparent, but shown that a core syntactic component must be present as well. Part II has focussed on clause-internal aspects of binding which are traditionally treated separately because of the impression that this binding, by virtue of being local, must be dealt with entirely in the syntax. The challenge in this series of chapters has thus lain in motivating, not the deterministic aspects of binding, but its more conceptual and discourse-pragmatic characteristics. In the process we have seen that the hybrid perspectival approach developed for long-distance binding does indeed carry over to cases of local binding as well. Part III has had to do with the relationship between anaphora, indexicality, and agreement. We have demonstrated that the classes of anaphor and indexical may not be conflated, contra analyses in the literature that assume otherwise, and that a principled distinction between perspective, which regulates anaphoric dependencies, and the Kaplanian and intensional contexts, which regulate indexical ones, must be made.

The broad topics addressed in this dissertation are thus: long-distance binding and perspective, local binding and argument structure, indexical shift, agreement, and intensionality. These phenomena have been discussed intensively in the literature but seldom if ever all together. The primary language of investigation, Tamil, has led us to consider these empirical phenomena in conjunction because the anaphoric element *ta(a)n* is centrally involved in all of them. As such, we have not only been able to understand their nature in isolation but also in relation to one another.

The central concept that cuts through these domains is that of perspective. The importance of perspective in grammar has long been acknowledged in the literature, particularly in the context of discussions about logophors and long-distance binding. However, the standard assumption has been that perspective itself is a strictly non-structural notion. As such, evidence for the relevance of structure for grammatical dependencies as, for instance, in cases of local binding, has automatically been construed as evidence *against* an approach involving perspective. Conversely, evidence for the relevance of perspective, as in sentences involving logophoric reference, has been automatically taken as evidence

against a syntactic approach. This has led to disparate accounts of phenomena that might profitably be treated together. One of the fundamental theses of this dissertation has been that perspective and structure are not mutually incompatible. Rather, I have shown that perspectival relations are structurally instantiated and drive syntactic, semantic as well as discourse-pragmatic dependencies.

There is a growing body of work that argues in favor of syntacticizing discourse-contextual information in functional structure and, to a certain extent, my work may be seen as a further contribution to this program. At the same time, much of this work isn't very precise about what sorts of information should count as being contextual and what shouldn't and often also makes the wrong cuts between them. Thus, another important contribution of the dissertation has been to bring some clarity to this topic. In particular, I have shown that there is a principled distinction between perspective and Kaplanian context, on the one hand, and between Kaplanian context and the kind of rich intensional operator introduced by a speech predicate (which is often inaccurately treated as a "shifted" Kaplanian context), on the other.

In the following sections, I will explore the broader implications and outlook of the main conclusions of this dissertation, going through the three empirical domains in turn.

13.2 Long-distance binding and perspective

As mentioned above, the standard wisdom in the literature has been that structural and conceptual binding dependencies are mutually incompatible. As such, prior work on long-distance binding has been either predominantly conceptual (Sells 1987, Kuno 1987) or predominantly structural (Pica 1987, Huang and Tang 1991, Progovac 1993, Reinhart and Reuland 1993) but seldom, if ever, both. This has caused a dichotomy between treatments of local binding which has been interpreted as being more structural and treatments of long-distance binding which has been conceived as being more conceptual and non-deterministic. An orthogonal split characterizes treatments of so-called "logophoric" binding, which has been seen to be extra-grammatical, and all other types of binding. Most of these analyses, with perhaps the notable exception of Pollard and Sag (1994), also involve the assumption that binding instantiates a direct dependency between the anaphor and its antecedent. Admittedly, many (perhaps even most) of these analyses do assume that this dependency might be implemented cyclically, in several smaller steps me-

diated by intervening functional heads (Pica 1987, Huang and Tang 1991, Kratzer 2009, Reuland 2011, among others). Nevertheless, the different steps of this relationship are all syntactic and construed as instantiating a single successive-cyclic dependency. In the two-step binding model that has been developed here, the two stages are completely orthogonal to one another and separately implemented in different modules of the grammar.

In the case of long-distance binding, the direct syntactic connection that is traditionally assumed to hold between the anaphor and its antecedent has led, perhaps inevitably, to problems having to do with antecedence, such as its optionality, non-locality, and non-minimality, which are very difficult to derive in purely syntactic terms.

13.2.1 Central contributions of Part I

In Part I, we saw that many properties that look structural on the surface turn out to be more conceptual upon closer examination: for instance, subject orientation was seen to have its "exceptions" in object experiencer antecedents. This might have suggested that binding is purely conceptual in nature. However, we have reviewed two pieces of evidence suggesting that binding involves an indelible syntactic core. The first is the Ban on Clausemate Subject Antecedence, which has been argued to be an anti-locality restriction imposed by a general wellformedness condition on perspective-holding, namely that the DP denoting the perspective-holder toward a PerspP not be asymmetrically c-commanded by the operator in the specifier of that PerspP. The second piece of evidence has had to do with the fact that, in Tamil, agreement under subject *ta(a)n* always tracks the antecedent of *ta(a)n*, but is not directly triggered by *ta(a)n* itself. We have seen, furthermore, that this property characterizes the agreement under *ta(a)n* regardless of whether the binding is long-distance, backward, or logophoric. We have reasoned from this data as follows:

Observation I: ϕ-feature agreement under subject *ta(a)n* is not directly triggered by *ta(a)n*.

Observation II: ϕ-feature agreement under subject *ta(a)n* tracks the antecedent of *ta(a)n*.

Assumption: ϕ-feature agreement is locally implemented in the Narrow Syntax.

Conclusion I: The ϕ-features of the nominal that gets interpreted as the antecedent of *ta(a)n* are represented on a local entity in the Narrow Syntax.

Conclusion II: The antecedent is itself not a local entity with respect to the anaphor. Thus, the local entity hosting the ϕ-features of the antecedent must be distinct from both the antecedent and the anaphor.

The Ban on Clausemate Subject Antecedence shows that there is a relationship between the antecedent and an element in the PerspP towards which its referent holds a perspective, and that this relationship involves a structural component (the only part of the relationship that is structural, as we have shown). The argumentation from the agreement facts provides evidence for the other stage of the binding process, namely the relationship between the anaphor and an element that "stands in" for the antecedent in the local domain of this anaphor. The elegant way to combine both results is to assume that the anaphor and the antecedent both have a relationship with the same element. This, we have proposed, is the operator in [Spec, PerspP]. Thus, the operator in [Spec, PerspP] mediates the relationship between the anaphor and its antecedent, just as desired.

One of the central contributions of Part I has thus been to motivate the idea that perspective and structure are compatible. In fact, I have argued that binding dependencies may be elegantly derived within an enriched grammatical model that involves the structural representation of perspective on a functional projection at the edge of certain phases. Crucially, there is a Perspectival Phrase (PerspP), unique to a phase, which plays a central role in the derivation of binding relationships: an anaphor is a "perspective-seeker", and an intended antecedent is a perspective-holder toward the minimal phase containing the anaphor.

A second major proposal is that the binding operation involves two distinct steps, as observed above. One involves the syntactic dependency between the anaphor and the silent pronominal operator in the specifier of its local PerspP. The second is a predominantly conceptual one (with the exception of the c-command wellformedness restriction described above) between this operator and the referent of the intended antecedent. This second step instantiates a type of non-obligatory control relationship (Williams 1980) between the intended antecedent DP and the silent operator in the specifier of the minimal PerspP containing the anaphor. All binding in this model is local; all antecedence is non-local. A direct connection between the anaphor and its antecedent is thus severed, making it much easier to pursue a cross-modular two-stage approach to anaphora.

13.2.2 Loci for parametric variation

The central attractiveness of this model is that it reduces the burden on the syntactic computation by relegating the more tendential aspects of long-distance binding to semantic and discourse-pragmatic factors at LF. The relationship between the anaphor and the locally c-commanding operator in [Spec, PerspP] in the syntax is thus constant regardless of whether the antecedent is several clauses away, of whether it c-commands the PerspP containing the anaphor and of whether it is even overtly represented within the sentence. Thus, my analysis can simultaneously deal with the non-deterministic aspects of (so-called) long-distance binding – like antecedent optionality, non-locality, and non-minimality – and account for the deterministic properties like antecedent-tracking agreement under *ta(a)n*.

The primary language of investigation in this dissertation is Tamil. But the nature of the analysis proposed is not specific to it. Many languages, like Italian, Icelandic, Japanese, Norwegian, and even English, manifest phenomena that are traditionally thought to be perspectivally regulated, such as logophoricity and backward binding – aspects of which I discuss separately in Section 13.3 below. The central proposal in Part I, namely the thesis that perspective is structurally represented, constitutes a claim about the availability of a certain type of discourse-pragmatic information in the linguistic structure. This is thus something that we expect to be inherent to language as a whole, and not to specific languages alone. More concretely, we expect PerspP to be available to and play a role in the syntax and semantics of all languages.

At the same time, we might expect crosslinguistic variation in the

types of linguistic phenomena that are sensitive to, thus driven by, the structural representation of perspective. The facts discussed throughout the dissertation make it clear that perspective cannot be equated with binding but is potentially relevant to other linguistic phenomena as well, and thus accommodate this logical possibility. The crosslinguistic variation in the different types of attested binding patterns seems to support this point. We have seen that binding patterns in (at least some dialects of) Tamil may be entirely captured in terms of their sensitivity to perspective. In Section 13.3 below, I will show that the so-called "empathy" phenomena discussed for Japanese (Kuno 1987) also encode a type of perspectival relationship and that binding in Japanese appears, to a large degree, to be regulated by its sensitivity to this. Icelandic and Italian also exhibit perspectival binding in certain types of structures, specifically across subjunctive clauses, as I will discuss below. Dutch and Norwegian, on the other hand, both seem less perspectival than these, at least with respect to binding. At the other end of the spectrum, we have languages like English which manifest types of binding relation that are not perspectival at all. Thus, in its standard locally bound use, an anaphor like *himself* is not subject-oriented and may thus be anteceded by a DP that does not denote a perspective-holder toward the minimal predication containing the anaphor. It seems possible, therefore, that languages do fall into a spectrum with respect to how perspectival they are with respect to binding.

That said, it is important to exercise caution in order to correctly distinguish the lack of relevance of PerspP for binding from cases where its relevance is obscured by independent grammatical factors in that language. The issue of logophoricity in Norwegian is a potentially good example of the latter. Norwegian lacks logophoric binding but it is not clear that this should automatically be taken to mean that PerspP doesn't play a role in regulating binding dependencies in this language. This is because Norwegian also lacks long-distance binding across finite CPs, a property that could very easily be related to its lack of logophoricity; after all, it is very difficult to imagine a free indirect discourse narrative consisting solely of non-finite clauses. Other sources of parametric variation might have to do with the nature of the operator in [Spec, PerspP] – specifically, whether it is obligatorily controlled by the intended antecedent or non-obligatorily controlled. This might also yield effects that look non-perspectival – or, perhaps, more accurately – structural, to the point that the relevance of perspective is obscured. This must be determined on a case-by-case basis for every language.

13.3 Crosslinguistic profile of perspectival binding

Here, I survey a range of languages that involve the use of perspective to varying degrees in the instantiation of their binding patterns. In particular, I will look at the influence of "empathy" (Kuno 1987) on the binding of Japanese *zibun*, the role of perspective in long-distance, logophoric, and backward binding into subjunctive clauses in Icelandic and Italian, and the representation of spatial perspective in binding into spatial PPs in Norwegian and Dutch. This brief survey will reveal the ways in which the two-step PerspP model may be adapted to different types of data and also illustrate in a concrete way the sorts of parametric variation that might be accommodated within the perspectival binding model developed in this dissertation.

13.3.1 Binding and "empathy" in Japanese

Kuno (1987) shows that certain lexical items in Japanese overtly mark the perspectival information associated with an eventuality. In Part I, we briefly looked at the examples of *yaru* and *kureru*. These are two verbs that both mean GIVE; they differ solely with respect to the type of perspectival relationship they encode. *Yaru* is used when the giving event is reported from the perspective of the AGENT (the giver), whereas *kureru* is used to signal that the utterance is reported from the perspective of the GOAL/RECIPIENT of giving (the examples below are from Oshima 2007, 2, formatting mine):

(443) *yaru:* AGENT-**oriented:**

Taro-ga Hanako-ni hon-o **yat-**ta.
Taro-NOM Hanako-DAT book-ACC give-PST
"Taro gave Hanako a book."

(444) *kureru:* GOAL-**oriented:**

Taro-ga Hanako-ni hon-o **kure-**ta.
Taro-NOM Hanako-DAT book-ACC give-PST
"Taro gave Hanako a book." (lit)
Reading: "Hanako received a book from Taro."

Crucially, Kuno shows that the binding of Japanese anaphor *zibun* appears to be sensitive to whose perspective is linguistically encoded

in this manner. Specifically, the antecedent of *zibun* (at least in long-distance binding structures) must be the DP whose perspective is being reported. This illustrated by the minimal pair in (445)-(446); in both sentences, the AGENT of GIVE is *Hanako* and the GOAL is the topic-marked *Taro*. (445) is illicit because it is anteceded by the GOAL *Taro* whereas the reported perspective of the embedded clause is that of the AGENT *Hanako*, as signalled by the use of the verb *yatta* (from *yaru*). However, in (446), the verb is changed to *kure(ru)*, indicating that the reported perspective is now that of the referent of the GOAL *Taro*; in this case, *Taro* may antecede the anaphor, as in (446):

(445) * Taro$_i$-wa [Hanako-ga **zibun**$_i$-ni **yat-**ta] hon-o yon-da.
Taro-TOP Hanako-NOM ANAPH-DAT give-PST book-ACC read-PST
"Taro$_i$ read the book Hanako gave him$_i$." (Intended)

(446) Taro$_i$-wa [Hanako-ga **zibun**$_{\{i,*j\}}$-ni **kure-**ta] hon-o yon-da.
Taro-TOP Hanako-NOM ANAPH-DAT give-PST book-ACC read-PST
"Taro$_i$ read the book Hanako gave him$_{\{i,*j\}}$."

As already briefly discussed in Part I, Kuno (1987) proposes that anaphora in Japanese is regulated by its sensitivity to "empathy", defined as follows:

(447) "*Empathy:* Empathy is the speaker's identification, which may vary in degree, with a person/thing that participates in the event or state that he describes in a sentence.
Degree of Empathy: The degree of the speaker's empathy with x, $E(x)$, ranges from 0 to 1, with $E(x) = 1$ signifying his total identification with x, and $E(x) = 0$ a total lack of identification." (Kuno 1987, 206).

Empathy Locus: The individual, denoted by a DP in the clause, that receives the highest degree of empathy in that clause.

There is a clear relationship between perspective, as we have been using it, and empathy. It seems, in particular, that empathy is closely tied, perhaps even identical, to the notion of *potential* perspective-holding. In our discussion of potential antecedence in Part I, we saw that there are various thematic and discourse factors that play a role in deciding how qualified an individual, denoted by a DP in the clause or immediate discourse, is as a perspective-holder with respect to an eventuality. Unsurprisingly then, the differences in binding possibility for *zibun* illustrated by the minimal pair in (445)-(446) can be captured very easily within our perspectival binding model.

In our system, this grammaticality difference would be a direct function of which entity controls the operator in the minimal PerspP containing *zibun*. The use of the verb *yaru* signals that this operator must be controlled by the AGENT of giving, namely *Hanako*, whereas the use of *kureru* forces control of this operator by the GOAL *Taro*. (446) is grammatical because the restriction on the identity of the perspective-holder placed by the choice of verb is consistent with that placed by the intended antecedence of *zibun*. In contrast, (445) is ungrammatical because these restrictions are incompatible. Specifically, the intended antecedent is *Taro* which, in this model, means that *Taro* must control the operator in [Spec, PerspP]; however, the verb *yaru* places the independent restriction that the controller must be the AGENT *Hanako*. Since there is a unique [Spec, PerspP] which can only host one operator, this is impossible.

To the extent that we are correct about empathy and perspective being intuitively similar, it should be possible to extend the perspectival model to all Japanese empathy-based binding phenomena, in the way sketched above. The question is whether all Japanese binding phenomena may be analyzed in terms of empathy in the first place. In a recent paper, Oshima (2007) argues that they may not. He proposes instead that Japanese *zibun* instantiates three distinct binding strategies: local, empathic, and logophoric, and that these must not be conflated. Oshima reports that local *zibun* must be co-argument bound and also seems to be insensitive to empathy restrictions. The antecedence condition for local binding seems to be merely that the DP construed as the antecedent be a co-argument of *zibun*; crucially, this is the DP that is chosen even if there is a different DP in the local domain with a higher empathy value (as indicated by the use of *yaru* vs. *kureru*. This is illustrated in (448) below (Oshima 2004):

(448) Max$_i$-wa Alice-ni zibun$_{\{i,*j\}}$-o e-no moderu-tosite
Max-TOP Alice-DAT ANAPH-ACC picture-GEN model-as
wariatete-**yat/kure**-ta.
assign-BEN-PST
"Max$_i$ assigned himself$_{\{i,*j\}}$ to Alice as a portrait model."

Given the absence of evidence, at this juncture, suggesting that empathy does play a role, we will set aside the issue of local binding here.

Oshima's evidence for suggesting a principled distinction between the (putative) categories of logophoric *zibun* and empathic *zibun* includes the following:

(i) Logophoric *zibun* displays no blocking effects with a 1st-person pronoun, but empathic *zibun* does (Oshima 2007):

(449) Taro$_i$-wa [$_{CP}$ boku-ga *zibun*$_{\{i,*j\}}$-o but-ta] koto-o mada
Taro-TOP I-NOM ANAPH-ACC hit-PST fact-ACC still
urande-i-ru.
resent-ASP-PRS
"Taro$_i$ still resents [$_{CP}$ that I hit him$_{\{i,*j\}}$]."

(450) * Taro$_i$-wa [$_{CP}$ boku-ga *zibun*$_{\{i,*j\}}$-ni kasi-ta okane-o
Taro-TOP I-NOM ANAPH-DAT lend-PST money-ACC
nakusite-simat-ta rasii.
lose-end.up-PST it.seems
"It seems that Taro$_i$ lost the money [$_{CP}$ I lent to him$_{\{i,*j\}}$]"

(ii) Logophoric *zibun* can be bound by a non-subject or extra-sentential antecedent. Empathic *zibun* must be bound by a sentential subject.

Based on these types of differences, Oshima concludes that logophoric *zibun* is an obligatorily shifted indexical along the lines of Schlenker (2003b), whereas empathic *zibun* is a "pov-o-phor" which targets an "empathic center" rather than the Kaplanian context (which, he correctly argues, must be kept distinct).

Observe, however, that it might be possible to develop a unified approach to the empathic and logophoric uses of *zibun* under the perspectival model developed here. First, the distinction between whether the antecedent is extra-sentential or not and whether it is a syntactic subject or not are not significant in the perspectival system: this is because the relationship between the intended antecedent and the operator in [Spec, PerspP] of the minimal phase containing the anaphor is envisioned as a type of non-obligatory control. This means that both subjects and non-subjects should be able to control this operator and the intended antecedent may also be extra-sentential. Tamil *ta(a)n* may also be anteceded by subject and non-subject antecedents; however, this by itself is not enough reason to posit two different categories of *ta(a)n*.

The differences in blocking effect due to 1st-person, seen in (449) and (450), are also not really problematic for us. Oshima's explanation of these differences is to claim that this is because the former involves logophoric *zibun* whereas the latter involves empathic *zibun*. Empathic *zibun* alone is expected to show blocking for 1st-person since the empathy locus of its minimal clause will always be the speaker, not its intended 3rd-person antecedent, causing a conflict with *zibun*'s requirement that it always be anteceded by the empathy locus (i.e. the DP with the highest empathy value). But we might be able to explain these differences within the perspectival model without having to posit two different types of

zibun. What is necessary in order to bind an anaphor in the perspectival approach is that the intended antecedent qualify as a perspective-holder with respect to the minimal PerspP containing the anaphor.

We have seen, however, that there are various thematic and discourse factors which go into determining which linguistic entities may denote perspective-holders. Observe that, in (449), the topicalized subject *Taro* thematically denotes an EXPERIENCER with respect to the embedded clause which is also the minimal PerspP containing *zibun*. In other words, *Taro* is a perspective-holder with respect to this PerspP, and qualifies as a potential antecedent for *zibun* by virtue of this. In (450), on the other hand, *Taro* does not qualify as a perspective-holder on thematic grounds, thus cannot antecede it in the pragmatically unmarked case. The only way for it to qualify as a potential antecedent would be if it were highly discourse salient and the speaker were narrating from the viewpoint of Taro. This is, however, blocked by the presence of a syntactic representation of the speaker in the sentence.

This account is, of course, preliminary and must be researched further before any definitive conclusions are reached. However, these initial considerations are suggestive that a unified approach could, in theory, be achieved. This is not only desirable from a theoretical elegance point-of-view, but also from an empirical one. Oshima proposes that his taxonomy of anaphors into: "pov-o-phor", logophor, and locally bound reflexive be extended to other languages, like English, Icelandic, Ewe, and Mundang. But a problem that Oshima himself notes for distinguishing "pov-o-phors" from logophors, on the one hand, and locally bound anaphors, on the other, is that, in none of these languages does the pov-o-phor have a distinct surface form: it is either homophonous with the local form, with the logophoric one, or with both. This is suprising if they are underlyingly different but expected if they are not.

A second problem for Oshima's approach has to do with his proposal that logophors are obligatorily shifted indexicals – a conflation that I have argued against in great detail in Part III. Under the perspectival model, such a conflation is entirely unnecessary. All the reported properties of logophoric *zibun* reported here may be just as easily accommodated by treating it as a "perspective-seeker" and not as a shifted 1st-person indexical.

13.3.2 Perspectival binding: Icelandic and Italian

Icelandic and Italian both exhibit clear evidence for the role of perspective in binding.[10] In both cases, the generalization seems to be that such binding obtains across subjunctive, but not across indicative, clauses.

We will start with Icelandic. As Hellan (1988) reports, a characteristic property of long-distance binding in Scandinavian is that it obtains primarily across nonfinite clauses. Icelandic is an exception in this regard: while indicative clauses block binding across them, subjunctives allow it (Hellan 1988, Reuland 2001a, Hicks 2009). The examples below are from Hicks (2009, 270):

(451) Jón$_i$ heyr-ð-i [$_{CP}$ að ég **hef-ð-i**
Jon hear.IND.PST-3SG that I have.SBJV-PST-3SG
svikið sig$_{\{i,*j\}}$].
betrayed.PTCP ANAPH
"Jon$_i$ heard [$_{CP}$ that I had betrayed him$_{\{i,*j\}}$]."

(452) * Jón$_i$ heyr-ð-i [$_{CP}$ að ég **haf-ð-i**
Jon hear.IND.PST-3SG that I have.IND-PST-3SG
svikið sig$_i$].
betrayed.PTCP ANAPH
"Jon$_i$ heard [$_{CP}$ that I had betrayed him$_{\{i,*j\}}$]."

Binding across subjunctives may be truly long-distance (Sells 1987, 473, Ex. 77):

(453) Jón$_i$ segir [$_{CP}$ að Haraldur$_j$ sé hér enn [CP þo að
Jon says that Harold is.SBJV here still though that
María kyssi sig$_{\{i,*j\}}$]]
Maria kisses.SBJV ANAPH
"Jon$_i$ says [$_{CP}$ that Harold$_j$ is still here [$_{CP}$ even though Maria is kissing him$_{\{i,*j\}}$]]."

At the same time, the presence of subjunctive marking is not sufficient for long-distance binding to obtain; what matters, it appears, is that the antecedent be construed as a perspective-holder toward the (subjunctively marked) proposition containing *sig*. This yields contrasts like the following Sells (1987, 451):

[10] Strahan (2010) reports similar facts for Faroese which suggest that perspective plays a central role in regulating binding dependencies in this language, as well.

(454) Barnið$_i$ lét ekki í ljós [$_{CP}$ að það hef-ði verið
child.DEF put not in light that there had-SBJV been
hugsað vel um sig$_{\{i,*j\}}$].
thought well about ANAPH
"[The child]$_i$ didn't reveal [$_{CP}$ that she$_{\{i,*j\}}$ had been taken good care of]."

(455) * Barnið$_i$ bar þess ekki merki [$_{CP}$ að það hef-ði verið
child.DEF bore of it not signs that there had-SBJV been
hugsað vel um sig$_i$].
thought well about ANAPH
"[The child]$_i$ didn't look [$_{CP}$ as if she$_i$ had been taken good care of]."

Although the same subjunctive-marked verb occurs in the embedded CP in both sentences above, it is only in (454) that the matrix subject is a perspective-holder with respect to the embedded CP and thus qualifies as an anaphoric antecedent. As Reuland (2001b, 345), describing these sentences, reports: "The difference in acceptability between [(454)] and [(455)] can be attributed to the fact that in [(454)] the report is made from the child's point of view, i.e., it is the child, and not the speaker, who didn't reveal that he/she had been taken good care of, whereas in [(455)], it is the speaker who reports that the child didn't look as if he/she had been taken good care of." Further evidence that Icelandic adopts a perspectival route to binding (at least in the subjunctive cases) comes from the fact that the antecedent may be non-c-commanding (Reuland 2001a, Ex. (456), 343) or even completely absent within the clause (logophoric Ex. (457) from Sigurðsson 1990, via Reuland 2001a), as long as it denotes a perspective-holder toward the minimal subjunctive CP containing the anaphor:

(456) [$_{DP}$ Skoðun Jóns$_i$] er [$_{CP}$ að sig$_{\{i,*j\}}$ vanti hæfileika].
opinion Jon.GEN is that ANAPH.ACC lacks.SBJV talents
"[$_{DP}$ Jon's$_i$ opinion] is [$_{CP}$ that he$_{\{i,*j\}}$ lacks talents]."

(457) María var alltaf svo andstyggileg. þegar Ólafur$_j$ kæmi
Maria was always so nasty. When Olaf come.PST.SBJV
segði hún sér$_{\{i,*j\}}$ áreiðanlega að fara ...
say.PST.SBJV she ANAPH.DAT certainly to leave ...
"Maria was always so nasty. When Olaf would come, she would certainly tell him(self) [the person whose thoughts are being presented – not Olaf] to leave."

The translation of the sentence in (457) clarifies that the antecedent

of *sig* is indeed a perspective holder with respect to the minimal proposition containing this anaphor. We can also clearly show the relevance of perspective-holding for binding in (456) by contrasting it with (458) below. The sentence in (458) contains the same embedded CP, marked with a subjunctive verb, as that in (456). The crucial difference between the two sentences has to do with the nature of the relationship between the intended antecedent *Jón* and the embedded CP in each. In (456), the embedded clause is reported from the perspective of *Jón* which denotes an attitude-holder with respect to the embedded proposition; in (458), however, *Jón* doesn't denote an attitude-holder at all; rather the embedded proposition is related from the speaker's perspective (denoted by *meg*). The possibilities for *sig*-antecedence track this difference: only in (456) is *Jón* a possible antecedent for *sig*:

(458) * $[_{DP}$ Skoðun Jóns$_i$] fær mig til að halda $[_{CP}$ að
opinion Jon.GEN leads me to to.INF believe that
sig$_{\{i,*j\}}$ vanti hæfileika].
ANAPH.ACC lacks.SBJV talents
"$[_{DP}$ Jon's$_i$ opinion] leads me to believe $[_{CP}$ that he$_{\{i,*j\}}$ lacks talents]."

The subjunctive binding facts above show that some instances of binding in Icelandic follow a perspectival route, very similar to that adopted for Tamil. The grammatical sentences above could all be straightforwardly derived within the PerspP-model developed here, under the assumption that there is a unique PerspP projected in the left edge of the subjunctive clauses that is controlled by a perspective holder with respect to that PerspP, and that *sig* in these cases, just like Tamil *ta(a)n*, is a "perspective seeker".

There are still some important open questions. It is not entirely clear, for instance, why a perspectival route to binding is apparently blocked across indicative clauses (compare again (451) and (452)). We might speculate that this is, at least in part, due to the semantics of the indicative/subjunctive distinction in Icelandic. The function of the subjunctive in Icelandic is explicitly to mark obviation from the speaker of the utterance context, "its role being mainly to signal that the perspective-holder of a given construction is distinct from the speaker" (Hellan 1988, 89). As might be expected from this, subjunctive in Icelandic only obtains under certain types of predicate. Sigurðsson (2010) reports that these are predicates that are classified as "assertive" in Hooper and Thompson (1973) and involve, in their standard use, non-factive verbs like BELIEVE, HOPE, SAY, THINK. Sigurðsson also reports that there is a special

subjunctive marking under the so-called "true factives" which includes verbs like: DEPLORE, REGRET and REJOICE claiming that "In modern Icelandic, the most important factor that triggers subjunctive marking in these complements is that the speaker does not take *responsibility* for their truthfulness" (Sigurðsson 2010, 50). In contrast, the indicative does not carry with it this "obviative" function, suggesting that the perspective (represented by the operator in [Spec, PerspP]) is simply preset to be that of the speaker in such cases.

Relatedly, it is also unclear at this juncture whether perspective regulates all types of binding in Icelandic. Binding across infinitives in Icelandic is less obviously perspectival: for instance, both logophoric and non-c-commanding antecedence are ruled out across such clauses. But this could either be taken to mean that binding is non-perspectival in these cases (implemented, for instance, via a direct structural connection between the anaphor and its antecedent), or to mean that it *is* perspectival but subject to more stringent structural constraints (e.g. that the intended antecedent be sententially represented and c-command the operator in [Spec, PerspP]). A proper understanding of these issues will require a detailed investigation like the one conducted here for Tamil.

In Italian, just like in Icelandic, long-distance binding is blocked across indicative clauses (459) but is permitted across subjunctive ones (460) (Giorgi 2006):

(459) * [Quel dittatore]$_i$ ha detto [$_{CP}$ che i notiziari televisivi
that dictator has said that the.PL news television
hanno parlato a lungo delle proprie gesta].
have.IND talked for long about ANAPH.GEN$_i$ deeds
"[That dictator]$_i$ said [$_{CP}$ that the TV news programs talked for a long time about his$_i$ deeds]."

(460) [Quel dittatore]$_i$ spera [$_{CP}$ che i notiziari televisivi
that dictator hopes that the.PL news television
parlino a lungo delle proprie gesta].
talk.SBJV for long about ANAPH.GEN deeds
"[That dictator]$_i$ hopes [$_{CP}$ that the TV news programs will talk for a long time about his$_{\{i,*j\}}$deeds]."

Also like in Icelandic, the binding across subjunctive clauses appears to be regulated by sensitivity to perspective. I.e. the requirement is that the antecedent denote a perspective-holder toward the minimal situational predication containing the anaphor. Thus, in (461), the medial DP *il dittatore* may not be an antecedent, even though the embedded clauses are subjunctive, because it does not denote a perspective holder.

In contrast, the matrix subject *il primo ministro* does qualify as a potential antecedent because it denotes an attitude-holder with respect to the embedded CP containing the anaphor *proprio*:[11]

(461) [Il primo ministro]$_i$ sperava [$_{CP}$ che il dittatore partisse
the prime minister hoped that the dictator left.SBJV
prima [$_{CP}$ che i rivoluzionare sequestrassero il proprio$_{\{i,*j\}}$
first before the revolutionaries sequestered.SBJV the ANAPH.GEN
patrimonio]].
patrimony]]
"[The prime-minister]$_i$ hoped [$_{CP}$ that the dictator$_j$ left first [$_{CP}$ before the revolutionaries$_k$ sequestered his$_{i,*j,k}$ patrimony]]."

The situation with the Italian example above is parallel to that seen with Icelandic (454) and (455) above. As in Icelandic, Italian also allows a non-c-commanding DP to antecede the anaphor as long as this DP denotes a perspective-holder with respect to the minimal predication containing the anaphor. This is shown in (462) and (463) and, even more clearly, in (464) below (Giorgi 2006):

(462) [$_{DP}$ La-propria$_{\{i,*j\}}$ moglie] preoccupa molto Gianni$_i$.
ANAPH$_{\{i,*j\}}$ wife worries a-lot Gianni.
"[$_{DP}$ His$_{\{i,*j\}}$ wife] worries Gianni$_i$ a lot."

(463) [$_{DP}$ Coloro che ambiscono al proprio$_{\{i,*j\}}$ incarico] preoccupano
those who wish.for to.the ANAPH.GEN job worry
molto [il primo ministro]$_i$.
a lot the prime minister
"[$_{DP}$ Those who wish for his$_{\{i,*j\}}$ official position] worry [the prime-minister]$_i$ a lot."

(464) [$_{CP}$ Che la propria$_{\{i,*j\}}$ figlia sia andanta in campeggio
that the ANAPH.GEN daughter has gone to camp
da sola] turba i sogni di Gianni$_i$.
alone disturbs the dreams of Gianni
"That's his$_{\{i,*j\}}$ daughter went to camp alone disturbs Gianni's$_i$ dreams a lot."

Observe that, in all these cases, the antecedent of *il proprio* is an EXPERIENCER and that the minimal predication containing the anaphor

[11] The possibility of clausemate subject antecedence by the innermost subject *i rivoluzionare* should not worry us here. Note that the anaphor is itself embedded inside a possessive DP – thus the Ban on Clausemate Subject Antecedence is not violated in this case.

reports on the experience of this antecedent. Such examples show that what is relevant is not c-command or syntactic subjecthood, on the part of the DP that is construed as the antecedent, but the holding of a perspective toward the minimal situational predication containing the anaphor. Needless to say, such data is also very easily accommodated in the PerspP-based model proposed in this dissertation. As with the Icelandic sentences, we might propose that, in sentences like these, there is a PerspP projected in the left edge of the minimal prediction – here, as in Icelandic, the subjunctive clause – containing the anaphor. The actual antecedent of the anaphor is the DP that non-obligatorily controls the operator in the specifier of this PerspP, and the anaphor is a "perspective seeker".

13.3.3 Binding into spatial PPs: Norwegian and Dutch

On the face of it, binding phenomena in Norwegian and Dutch don't seem nearly as perspectival as those in the other languages discussed here. However, binding into spatial PPs seems to constitute a legitimate exception in both. In both languages, an anaphor contained inside a spatial PP takes the DP whose spatial location is being denoted by that PP as its antecedent.

Let us start with Norwegian. In standard cases of local and long-distance binding, the simplex anaphor *seg* in Norwegian displays subject-orientation – a property that Hellan (1988) captures by proposing that the antecedent must be the subject of the predication containing the anaphor:

(465) **Subject orientation of *seg*:**

a. Leif$_i$ fortalte oss om et forsøk på å hjelpe
Leif[NOM] tell.PST us about an attempt to to.INF help.INF
seg$_{\{i,*j\}}$.
ANAPH
"Leif$_i$ told us about an attempt to help him$_{\{i,*j\}}$."

b. * Vi fortalte Leif$_i$ om et forsøk på å hjelpe
we[NOM] tell.PST John about an attempt to to.INF help.INF
seg$_i$.
ANAPH
"We told Leif$_i$ about an attempt to help him$_i$."

However, as both Hellan (1988) and Hestvik and Philip (2001) note, an object DP may antecede the possessive anaphoric form *sin* if *sin* is

part of a locative PP that denotes the location of this object. This is shown below (Hellan 1988, 74):

(466) Vi fant Jon$_i$ i sengen sin$_{\{i,*j\}}$.
We[NOM] find.PST Jon in bed ANAPH.POSS
"We found Jon$_i$ in his$_{\{i,*j\}}$ bed."

Hellan points out that the subject-of-predication account could be extended to explain the antecedence of *Jon* in sentences like (466), since *Jon* is the subject of the PP predication here. However, as Lødrup (2007) points out, this explanation doesn't carry over to cases of binding involving *non*-locational PPs. In such structures, the complex anaphor *segselv* is used over the simplex form, yielding contrasts like the following:

(467) *mot* (TOWARD, AGAINST):
a. Han$_i$ drar den mot seg$_{\{i,*j\}}$.
He[NOM] pull.PRS it towards ANAPH
"He$_i$ pulls it towards himself$_{\{i,*j\}}$."
b. Forbrukerråd-et argumenterer mot [seg selv]$_{\{i,*j\}}$
consumer.council-DEF argue.PRS against ANAPH self.
"[The consumer council]$_i$ argues against itself$_{\{i,*j\}}$."

(468) *om* (AROUND, ABOUT):
a. De$_i$ spredte en karakteristisk odør om seg$_{\{i,*j\}}$.
they[NOM] spread.PST a characteristic odor around ANAPH
"They$_i$ spread a characteristic odor around themselves$_{\{i,*j\}}$."
b. De$_i$ vil fortelle om [seg selv]$_{\{i,*j\}}$.
they[NOM] will tell about ANAPH self
"They$_i$ will tell about themselves$_{\{i,*j\}}$."

Based on such data, Lødrup (2007) argues that a predicational account is not sufficient and that the semantics of the preposition involved must be taken into consideration as well. His own solution to the problem is to propose that "the simple reflexive is used when the physical aspect of the referent of the binder is in focus" (Lødrup 2007, 183). In all other cases, *seg selv* is used, making this form the elsewhere case. This conception of things correctly explains the use of the simplex forms in sentences like (467a), (468a), all of which involve the physical use of the relevant preposition; it also explains why the complex form occurs in those sentences with binding into non-physical PPs. But it does so at the expense of positing two different types of *seg* – one that is "physically oriented" and one that is not. It is also unclear why *seg* should be physically oriented in this manner in local cases alone.

Consider instead an analysis that starts from the observation that the use of *seg* in the locational PP structures in (466), (467a) and (468a) is necessarily accompanied by the reading that the locative PP in each denotes the location of the antecedent of *seg*. Put another way, the antecedent of *seg* in each of these cases denotes the entity that occupies the locational space indicated by the spatial PP containing *seg*. This suggests that such structures would be better analyzed as instances of perspectival binding of the sort developed here. Recasting the binding relationship as a type of perspectival relationship could also be used to explain the different behavior of the anaphors in locational PPs vs. non-locational ones. We would expect a locational PP to instantiate a spatial perspective in its left periphery which would, however, not be present in a non-locational one. We can attempt, then, to explain distributional differences between simplex and complex anaphoric forms in such PPs as a function of this difference in the availability of a perspective. Such a treatment has the advantage that the differences aren't attributed to properties internal to the anaphors themselves, which can thus retain a uniform feature-structure and denotation in both local and long-distance environments.

Very similar effects obtain in Dutch. Binding in Dutch, for instance, doesn't seem to be as obviously "perspectival" or even predicational like that in the Scandinavian languages, as Hellan (1988) reports. Thus, the simplex anaphor *zich* in Dutch is not "subject"-oriented like Norwegian *seg* or Icelandic *sig*. Furthermore, it appears that long-distance binding is not truly long-distance in this language: binding is not possible across more than one infinitival clause boundary or across any finite clauses. Interestingly enough, Dutch does manifest some logophoric binding but the *zich(zelf)* form is sharply ungrammatical in this case, and a different form *hemzelf* is used instead (Rooryck and vanden Wyngaerd 2011). These facts suggest that the connection between the antecedent and *zich* in Dutch might be more direct than that in the Scandinavian languages and, definitely, than in Tamil.

At the same time, there *is* potential evidence for the representation of spatial perspective in Dutch in the so-called "snake" sentences, namely structures that involve binding into spatial PPs. In Dutch, the use of the anaphoric *zich* vs. deictic *hem* is a function of whose spatial perspective is being reported: the former is used to denote the spatial perspective of the anaphoric antecedent whereas the latter is used to indicate that of the utterance-context speaker or that of the antecedent (Rooryck and

vanden Wyngaerd 2011, 266-7, formatting mine):[12]

(469) [De volwassenen]$_i$ op het schilderij kijken van ons weg, met
The adults on the painting look from us away with
de kinderen [$_{PP}$ achter zich$_{\{i,*j\}}$].
the children behind ANAPH
"[The adults]$_i$ in the picture are facing away from us, with the children placed [$_{PP}$ behind themselves$_{\{i,*j\}}$]."

(470) [De volwassenen]$_i$ op het schilderij kijken van ons weg, met
The adults on the painting look from us away with
de kinderen [$_{PP}$ achter hen$_i$].
the children behind ANAPH
"[The adults]$_i$ in the picture are facing away from us, with the children placed [$_{PP}$ behind them$_i$]."

Rooryck and vanden Wyngaerd note that the interpretation of BEHIND in (469) is from the spatial perspective of the adults whereas that in (470) is from the spatial perspective of the antecedent or of the observer/speaker. Thus, although the Dutch facts are not exactly parallel to those in Norwegian, the tendencies go in the same direction.

This is of course only a tentative proposal. Further research must be undertaken of the binding patterns in both these languages to ascertain to what extent the simplex and complex forms are amenable to the perspectival binding analysis developed here. But the spatial PP binding facts discussed here are suggestive. If this type of analysis is on the right track, the Norwegian and Dutch data might be taken to suggest that another locus for parametric variation might be the type of perspective – mental, spatio-temporal, modal, or underspecified – that is linguistically instantiated and relevant for binding in a given language. In languages like Tamil and Japanese, spatial as well as mental perspective may be relevant for binding, whereas in less obviously perspective-sensitive languages like Norwegian and Dutch, only one type of perspective – specifically the former – influences binding possibilities.

[12]The authors mark the sentence in (470) with a "#" to indicate that not all dialects of Dutch allow a deictic pronoun to be used in place of the simplex anaphor here. In the main text, I am simply describing the interpretive differences between these sentences for those speakers for whom both are already grammatical; I have thus left out this marking here.

13.4 Local binding and argument-structure

In the domain of local binding, the tendency seems to be to assume that the dependencies are structural rather than conceptual. This is partly because local binding, by virtue of its locality, is easier to implement purely syntactically. But it is also because it does seem to involve fewer indeterminate factors than long-distance binding does, at least in languages like English whose binding properties shaped much of the earlier theoretical discussion (Chomsky 1981). These properties have continued to characterize more recent approaches to local binding as well, such as those in Heinat (2008), Hicks (2009), Reuland (2011), Rooryck and vanden Wyngaerd (2011), among others. The central challenge here has, in fact, come from thematic and argument-structural factors which, in many languages, have been shown to influence the distribution of simplex vs. complex reflexives as well as yield systematic differences in the interpretation of the anaphor. An important concern related to this is the connection between reflexivity and unaccusativity. It has been observed in a number of languages that these two seemingly distinct argument-structural categories often receive identical morphosyntactic marking (see Chierchia 2004, Embick 2004b, Lidz 2004, Medová 2009, Schäfer 2008, among others). As such local reflexivity has often been treated as a voice phenomenon, although the issue of how this should be implemented is by no means resolved.

13.4.1 Central contributions of Part II

One of the central contributions of my discussion of local binding in Part II is that perspective may be represented not just in CPs, PPs, and DPs, as discussed in Part I, but also in the thematic-aspectual layer of the extended verb-phrase, specifically in AspP. This places sensitivity to argument-structural considerations, on the part of binding, on a par with discourse-sensitivity at the propositional level. Some verbs, by virtue of what they "mean" lend themselves more easily to a perspectival interpretation, others are neutral with respect to it, and yet others actively resist it. The potential for adding *ko*⌊ to a verb has been shown to track these distinctions. One of the major conclusions in Part II is the idea that the apparent connection between reflexivity and voice phenomena, at least in Tamil, is just another expression of the sensitivity of binding to perspective. Evidence for this from Tamil comes from the fact that, although the same morpheme *ko*⌊ marks both unaccusatives and reflexives, it also marks fully non-reflexive transitives and unergatives. Thus, it

does not directly derive or mark reflexivity or unaccusativity but rather a particular kind of perspective which is compatible with these categories. Another fallout of the idea that perspective may be represented in the thematic layer is that we can maintain that perspective is a property of phases (given that types of *v*P are argued to be phasal (Chomsky 2001)) – in other words, that there is at most one PerspP per phase.

A larger issue that this has brought to light is the Ban on Clausemate Subject Antecedence, banning co-argument antecedence unless there is a special marker, like *ko*l̥ (or potentially Germanic SELF) explicitly allowing it. In Part I, we briefly considered the possibility that this condition might be conceptual, rather than the instantiation of a structural anti-locality effect, due to the fact that it does not seem to hold in psych-predicate structures, at least in Tamil. But we pre-emptively rejected this claim pending further investigation in Part II. In Part II, we argued that local, or more precisely, co-argument binding is not possible without some additional "help" because of what perspective itself means, which requires a perspective-holder to be external to the object of scrutiny. We formalized this by proposing that a DP may not be asymmetrically c-commanded by the operator in the specifier of the PerspP towards which it denotes the intended perspective-holder. Elements like *ko*l̥ in Tamil were seen to offer a way out of this problem by raising relevant DPs to a position outside this PerspP.

The main pay-off of this approach is that it allows a unified account of local and long-distance binding. It also yields an explanation for the facts of the distribution of the element *ko*l̥ in Tamil showing, in particular, that while local binding, in the standard case, needs *ko*l̥, *ko*l̥ doesn't need local binding. This explains why *ko*l̥ may occur in contexts that have nothing to do with binding, such as change-of-state unaccusatives, and non-reflexive transitives and unergatives. In the case of psych-predicates, we argued that the EXPERIENCER subjects are externally merged in a relatively high position that is already above the minimal PerspP containing the anaphor. This idea, if true, would also explain possible psych-predicate structures involving backward co-argument binding, a potentially significant conclusion.

13.4.2 Loci for parametric variation

To the extent that a language makes use of perspective in the derivation of its binding dependencies, we expect that the structural condition that a DP denoting a perspective-holder with respect to a PerspP may not be properly embedded within that PerspP, will hold in this weakest form.

However, it is possible that certain languages require a more stringent version of this – e.g. that the intended antecedent asymmetrically c-command the operator in [Spec, PerspP]. Such a condition would predict that an antecedent always be overtly represented within the sentential structure containing the anaphor; i.e. logophoric and backward binding should be impossible. We again expect that, in such languages, local binding should require the use of something "extra" either to allow the DP to raise out of the minimal PerspP containing both it and the anaphor or to introduce an additional PerspP between the two.

Relatedly, we may expect languages to vary in the precise kind of "extra" element they use for introducing perspective. Faltz (1977) conducts a comprehensive typological inventory of local-reflexive marking and reports that a great many languages do mark local binding relations with something additional. As Reuland (2011) describes it, this tends to be an element like SELF in Germanic or an inalienable body-part (like "head") or something else (presumably *ko*⌊ would fall into this miscellaneous category). Possible sources of variation are the argument-structural position of the PerspP introduced by this extra element. Tamil *ko*⌊, we have shown, introduces a PerspP in the thematic-aspectual layer of the clause. There may also be variation in the precise nature of the PerspP that is thus introduced – specifically with respect to whether the perspective denoted by this phrase is fundamentally spatial, mental or underspecified for either. A more fundamental question, of course, is to what extent the presence of something "extra", like this, in local binding structures in a given language can be taken as evidence for a perspectival route to binding in that language.

Another issue which provides fertile ground for crosslinguistic investigation is the relationship between reflexivity and unaccusativity, observed in many languages. The crosslinguistic pervasiveness of this parallel is suggestive of a deeper connection between the two. This is the type of connection that analyses of local reflexivity as a voice phenomenon seem to target. At the same time, this is called into question by the Tamil facts which show conclusively that the connection between local binding and unaccusativity (to the extent that one even exists) has nothing to do with predicate-valency. One way to resolve this tension might be to start from the observation that, in our analysis of the argument structure of *ko*⌊, Voice and PerspP are in adjacent structural positions. As such, we can imagine that the Voice and Persp heads are, in many languages, realized by a single exponent. This would be akin to the "Voice-bundling" strategy discussed in Pylkkänen (2008) and proposed more generally as a spell-out mechanism in the spanning account of Ramchand (2008).

13.4.3 Potential insights into the GET-passive

A potentially useful connection is that between *koɭ* and GET-passives in languages like English, and German. The argument-structure of *koɭ* is very similar to that for the middle-like readings reported with *get* and *kriegen* in English and German, respectively (McIntyre 2010). All of these seem to involve a BECOME-type semantics with a designated argument of the event predication coming to hold the derived result state of that event. Both types of structures induce meanings like affectedness, volitionality, and self-benefaction. Finally, GET-passives have been analyzed as types of raising or control verbs (Butler and Tsoulas 2006), just like we have proposed for *koɭ*.

An interesting observation with respect to the GET-passive is that it seems to be ambiguous between an AGENT-like and PATIENT-like reading, as in the following sentence:

(471) Susan got her hair cut.

AGENT-**like reading:** Susan went out of her way to get a haircut (e.g. went to a beauty salon).

PATIENT-**like reading:** Someone cut Susan's hair (e.g. while she was asleep).

This sort of tension is typically taken to be the result of a "middle voice" semantics in GET-passives. For instance, McIntyre (2011) proposes that GET-passives involve a Kratzerian Voice head with a middle-semantics and argues that the AGENT-like vs. PATIENT-like readings above are a function of whether the relevant argument is associated with a Causer or Causee semantics, respectively.

The tension between an AGENT-like and PATIENT-like interpretation has also been seen to characterize *koɭ*-structures in Tamil:[13]

(472) *oɖæ* (BREAK):

Raman kaal-æ oɖæčč̆u-ko-ɳɖ-aan.
Raman[NOM] leg-ACC break-*koɭ*-PST-3MSG
"Raman got his leg broken." (AGENT-like and PATIENT-like)

However, to the extent that Tamil *koɭ* is an instantiation of GET, it suggests that GET should not be conflated with Voice, as is standardly done in the literature. As such, we have proposed that this tension in Tamil results from the fact that the perspective denoted by the PerspP

[13]Lidz (2001) reports a similar tension in structures involving Kannada *koɭ*.

under *ko*ɭ is underspecified. I.e. it is consistent with either a mental or spatial interpretation. The mental yields an AGENT-like reading while the spatial one yields a PATIENT-like one. We might thus consider adopting a parallel strategy for the similar dichotomy with GET-passives. This too is a matter for future research.

13.5 Indexical shift, agreement, and intensionality

Part III of this dissertation deals with the phenomenon that I've termed "monstrous agreement" which brings together seemingly disparate issues of shifted indexicality, long-distance binding, clausal agreement, cartography, and intensional predicates. For the most part, these phenomena are discussed separately, not all together. The indexical shift literature aligns itself into two distinct camps: Kaplanian (Kaplan 1989, von Stechow 2002) and non-Kaplanian (Schlenker 1999; 2003b, Anand 2006, Shklovsky and Sudo To Appear), based on analytic positions with respect to whether a Kaplanian context may be manipulated in the scope of intensional predicates or not. This body of literature tends to be primarily semantic in spirit and implementation, thus doesn't concern itself too much with issues pertaining to agreement, clausal complementation, feature-structure, and the like (Baker 2008, being a notable recent exception). However, the issues of anaphora and indexicality are often discussed in the same breath with the Schlenkerian-type analyses conflating the two by analyzing anaphors/logophors as obligatorily shifted indexicals.

A more syntactic viewpoint is taken by practitioners of the cartographic tradition (Cinque 1999, Speas and Tenny 2003, Bianchi 2003, Sigurðsson 2004, Speas 2004, Giorgi 2010), who argue that certain types of contextual information are syntactically represented in designated functional heads in the clausal spine. However, work in this tradition tends to be predominantly syntactic in nature and thus sometimes conflates the notion of what we have called "perspective" with that of "context", in the Kaplanian sense. A third stream of research is concerned with the relationship between anaphora and agreement and with issues pertaining to the directionality of agreement in the syntax (such as, for instance Rizzi 1990, Woolford 1999, Haegeman 2004, Zeijlstra 2010, Tucker 2011, Wurmbrand 2011). These authors are focussed more on details of agreement, the feature-structure of anaphors, and the nature of the relationship between the two and thus don't spend too much time dealing with

issues pertaining to the semantics of these elements.

As such, one of the main contributions of this part of the dissertation is that it brings these various strands together and integrates their results. As a result of this, we are able to view these phenomena in a new light, one that, furthermore, illuminates connections and disconnects that haven't been exposed in more traditional approaches.

13.5.1 Central contributions of Part III

The major conclusions of Part III are, in no particular order, the following:

(i) Anaphors are distinct from obligatorily shifted indexicals, contra Schlenker (2003b, and subsequent).

(ii) Intensional context ≠ perspectival center:

   a. The intensional context is relevant for indexical shift. The perspectival center is relevant for anaphora.
   b. The intensional context has no discourse-pragmatic component whereas the perspectival center does.
   c. The intensional context is represented structurally higher than perspective. Information pertaining to the former is hosted in SpeechActP, which is the highest functional projection in the clausal functional sequence. Information pertaining to the latter is hosted in PerspP; however, this is not strictly a part of the clausal functional sequence but is a projection that is superimposed onto a predetermined functional sequence, like Topic or Focus or Neg (Rizzi 1997). As such, it may occur in different hierarchical positions in different languages and may also occur in different categories of phase in a given language, like (certain) CPs, PPs, DPs, and AspPs.
   d. The pronominal element in [Spec, SpeechActP] is obligatorily controlled by the minimally c-commanding antecedent corresponding to the AGENT of the speech predicate. The pronominal operator in [Spec, PerspP] is non-obligatorily controlled by the intended antecedent of *ta(a)n*.

(iii) The intensional context ≠ Kaplanian utterance context:

   a. The intensional "context" is not really a context in the Kaplanian sense at all. It is, rather, an enriched series of intensional

operators in the sense of Lewis (1979), each of which may covary independently of the other. The Kaplanian context is a complex *primitive* that indicates a deterministic connection between its various coordinates.

b. The intensional context is syntactically represented. The utterance context doesn't seem to be, because it yields no syntactic intervention effects. However, the utterance context is available to restrict reference assignment at LF.

(iv) Speech predicates are special. While long-distance binding, in languages like Tamil, is commonly found in the complement of a wide range of predicates, indexical shift obtains predominantly in the scope of speech predicates. This specialness is modelled within a cartographic framework of clausal selection by proposing that the propositional complements of speech predicates include a larger functional structure which subsumes that of complements of other classes of intensional predicate.

(v) Tamil instantiates the Anaphor Agreement Effect. The agreement under *ta(a)n* in subject position is not triggered directly by *ta(a)n*.

(vi) The agreement under *ta(a)n* is triggered by the operator in the specifier of the minimal PerspP containing *ta(a)n*. This operator in turn is non-obligatorily controlled by the antecedent of *ta(a)n*. This yields the *effect* that agreement under *ta(a)n* seems to track the antecedent of *ta(a)n*.

(vii) A single language may have more than one type of indexical, even within a particular dimension of meaning (person vs. time vs. space vs. world). E.g. the silent pronominal operator in [Spec, PerspP] in monstrous agreement structures in Tamil is an optionally shifting 1st-indexical whereas the overt 1SG *naan* is obligatorily non-shifting.

13.5.2 Loci for parametric variation

The conclusions made above with respect to the distinctions between Kaplanian context, intensional context, and perspective should hold universally. By extension, we expect that the classes of indexical (shifted or otherwise) and anaphor are underlyingly distinct in every language. However, this effect may be obscured on the surface, perhaps entirely, in languages that only allow binding under speech predicates – e.g. the

Chadic language, Mupun (Culy 1994). In such a language, the distribution of a shifted indexical and a long-distance bound anaphor will look very similar, to the point of identity. However, we do expect potential differences between the two: e.g. minimality effects on antecedence is only predicted with shifted indexicality, not with long-distance binding. The idea that SpeechActP is the highest projection in the clausal functional sequence is not expected to vary across languages.

We know for a fact that not all languages allow indexical shift for a particular dimension, like person. This variation might be captured by claiming that, in some languages (like English), a clausal complement that is big enough to project a SpeechActP is never selected. This would be a cartographic implementation of the kind of analysis proposed in Anand (2006) where languages are claimed to differ with respect to the availability of context-overwriting operators in the scope of speech predicates. However, it might be argued that this goes against the spirit of a universal functional sequence and its correlation with predicate-selection.

A different logical source for variation on this point is that languages differ with respect to the kind of indexical they make available. Thus, we might say, the 1SG indexical *I* in English doesn't shift because it is inherently specified to always and only be evaluated against the utterance context. A language like Amharic, on the other hand, would differ in having a 1SG indexical that is underspecified with respect to its context of evaluation. This, for instance, is the Schlenkerian view to parametric variation in indexical shift. Under such a view, we might in theory also expect indexicals that always shift with respect to the utterance context to be attested. Indeed, as we have seen, this is how Schlenker (2003b) would treat an element like *ta(a)n*. However, we have shown clearly that *ta(a)n* is not an *indexical* to begin with. As such, it is simply restricted from being a participant of the utterance context and is not specified to be a participant of a non-utterance context. In other words, *ta(a)n* unlike an obligatorily shifted indexical, has the additional option of not denoting a participant at all. If this analysis can be maintained more generally for elements like *ta(a)n*, we might be able to eliminate reference a context other than the utterance context entirely in the denotations of indexical elements.

A third way to capture this crosslinguistic difference would be to propose, in line with Ritter and Wiltschko (2009), that what varies across languages is the anchoring mechanism, in particular the choice of whether anchoring to the utterance-context is done via the $Person, Time, World$ or *Location*. Under this classification, languages with indexical shift for person would be ones that anchor to the utterance context via $Person$,

while those with indexical shift for tense (i.e. sequence-of-tense effects) or mood (i.e. Konjunktiv I in German) are anchored via the $Time$ and $World$ coordinates, respectively. An interesting question is how to model languages that show indexical shift across different domains, like the Iranian language Zazaki which has been argued (Anand and Nevins 2004) to optionally shift all the contextual coordinates. A reasonable option might be to say that, in such cases, the anchoring content is underspecified. In the model we have developed in this dissertation, anchoring to an intensional context is formally implemented in the syntax via the element in the specifier of SpeechActP. Variation in anchoring mechanism would thus be a direct function of what type of pronoun (personal, temporal, modal, etc) occupies [Spec, SpeechActP]. An advantage of this option is that we can retain the idea that the inventory of functional heads selected by a given predicate-class can be the same across languages.

Aesthetic preferences aside, the choice between these three different options must, of course, be made on empirical grounds. The first option – namely the idea that variation in possibilities for indexical shift is a function of variation in the size of clausal complement selected by a predicate-class across languages – predicts accompanying variation in other manifestations of structural size. E.g. we might expect concomitant differences in the presence or absence of morphology marking additional categories, the availability or lack thereof of certain kinds of adverbial, differences in clausal finiteness, and so on. The second option, encapsulating the idea that variation in indexical shift has to do with the denotations of the indexicals themselves might predict that, within a single language, we might find more than one type of person indexical, more than one type of temporal indexical, and so on. Observe that such variation does seem to be attested in Tamil: we have argued that the operator in [Spec, PerspP] is a shifting indexical whereas the deictic 1SG pronoun *naan* is rigidly unshifting with respect to the utterance context. This shows that this option is linguistically available, at least for some languages. Evidence for the third option, namely that variation is a function of the mode of anchoring to the intensional context, also makes definite empirical predictions. Under this alternative, we would be claiming that a language like English, which doesn't shift for person, *does* shift for some other contextual coordinate. In other words, for every language, we should be able to find at least *one* designated mode of anchoring to the intensional context, causing the effect of shift for that dimension.

While capturing crosslinguistic variation with respect to indexical shift is not completely straightforward, as we have seen, capturing para-

metric variation with respect to the types of predicate that allow long-distance binding in their scope *is* straightforward, in the system being developed here (see Culy 1994, for details on this type of variation). The crucial difference is that we are explicitly not proposing that the PerspP which is responsible for mediating binding dependencies between an intended antecedent and the anaphor occupies a designated position in the functional sequence. Rather, we are claiming that this is a functional projection that is superimposed onto a functional sequence, like Neg. This means the difference between a more restrictive language like Donna Sɔ, which allows long-distance binding only under speech and thought verbs, and a more promiscuous one like Tamil could simply be made to fall out from where in the clausal hierarchy PerspP can be instantiated in each. This is analogous to crosslinguistic variation in for e.g. the position of Topic, Focus or Neg projections (Rizzi 1997).

One final area of relevant crosslinguistic variation, which must be discussed, is the Anaphor Agreement Effect. As we have seen, the variation here is not so much in whether languages observe the effect as in what types of strategies they use to avoid a violation of it. Tucker (2011) delineates a series of language-specific strategies undertaken in this regard. We have proposed that Tamil instantiates a different strategy from these, in which a functional head having failed to get its ϕ-features valued by the anaphor, enters into a feature-sharing relationship with the anaphor for these ϕ-features (Pesetsky and Torrego 2007) and keeps probing in the phase until it finds a Goal that can do so. It is unclear at this juncture why this strategy is not adopted by more languages, but we can start with the following considerations. First, the language will have to allow anaphors in agreement-triggering position in the first place. For some reason, this is very rare. Second, the ϕ-featural specification on the anaphor has to be such that, while it doesn't itself trigger agreement, it also doesn't block further agreement from happening in the phase. Something in this combination of factors must be extremely uncommon, and it is as yet unclear what this is. This is a matter for future research.

13.6 Summing up: Central theses and contributions of the dissertation

The dissertation makes the following theoretical claims. A central thesis is the idea that perspective, defined as in (473) is an important concept in guiding linguistic dependencies:

(473) **Definition of linguistic perspective:**

i. Linguistic perspective denotes a two-place predicate that relates an individual that exists in a specific space, time, world, and, if sentient, has a mind, with a predicational structure.
ii. When we assert that an individual holds a perspective towards a predication, we are asserting that the predication is evaluated against the space, time, world, or mind of this individual.
iii. The individual may not hold a perspective toward a predication that it is wholly embedded within.

We argued, furthermore, that the perspectival relationship is featurally represented at the edge of a phase in the form of a *perspectival center* (474) and plays an important role in regulating referential dependencies, such as that between a nominal anaphor and its antecedent or between clausal agreement and an anaphor, both in the syntax and at LF:

(474) **Formal representation of a Perspectival Center:**

i. The *perspectival center* contains the coordinates pertaining to the time, location, world, and/or mental information of a salient perspective holder. These are hosted in a silent pronominal operator in the specifier of a Perspectival Phrase (PerspP).
ii. Certain phases (at least some PPs, DPs, AspPs, CPs) contain a PerspP by virtue of what they inherently "mean".
iii. A phase has at most one PerspP.
iv. The phase containing a successfully bound anaphor must contain a PerspP. The operator in [Spec, PerspP] Agrees with the anaphor in its minimal phase and variable-binds it at LF.
v. A potential antecedent may not be asymmetrically c-commanded by the PerspP it holds a perspective towards. The relationship between the potential antecedent and this operator is one of non-obligatory control.

(i) Binding is a two-step hybrid process: it has a core syntactic component and the rest of it is semantic/pragmatic:

a. An anaphor is a "perspective seeker". The intended antecedent is a "perspective-holder".
b. The anaphor enters into a local syntactic relationship, not with the DP that ends up being construed as its antecedent, but with the operator in [Spec, PerspP] in its minimal phase.

c. The intended antecedent enters into a non-obligatory control relationship with the operator in [Spec, PerspP].
d. All binding is local. All antecedence is non-local.

(ii) Syntax cares about co-reference, not about reference. DPs that are co-referent in the syntax are marked with a DEP-feature. No other DPs are marked with this feature.

(iii) Local and long-distance binding, and so-called "logophoric" binding may be unified under such a view. Empirical evidence for this comes from clausal agreement under *ta(a)n* in Tamil which, while not triggered by *ta(a)n*, tracks the antecedent of *ta(a)n* in all these cases. This additionally shows that all binding in Tamil (and languages like it) involves an indelible syntactic core.

(iv) The fact that the individual denoted by a DP may not hold a perspective toward the PerspP that the DP is itself embedded within means that co-argument antecedence requires something "extra". This "extra" piece is often morphologically marked yielding the result that so-called local binding looks different from so-called long-distance binding. We can adduce (at least) two strategies for adding this extra piece:

a. **PerspP is introduced high (Tamil):** The addition of an element (like *ko*l̩) in the thematic-aspectual layer which allows the intended antecedent DP to "raise" out of the minimal PerspP containing it and the anaphor, to a position where its referent may hold a perspective toward this PerspP.
b. **PerspP is introduced low:** The addition of an element in the extended projection of the DP containing the anaphor which introduces a PerspP in its scope ensuring that the minimal PerspP for the anaphor is no longer the one that also contains the intended antecedent. As suggested in Part II, it is possible that the dialect of Tamil that allows local *ta(a)n*-binding even in the absence of *ko*l̩ (in psych- as well as non-psych predicate structures) employs this strategy.

(v) The identical morphosyntactic marking on local reflexivity and unaccusativity in a number of typologically unrelated languages should not be automatically treated as evidence that so-called local binding is a species of voice phenomenon. Rather, this effect could be made to derive from the fact that PerspP and Kratzerian

Voice are adjacent functional heads and thus interact and may be easily realized by a single exponent in a given language. Evidence for this comes from Tamil, which shows that *ko*⌊ marking on unaccusatives and local reflexives does not affect predicate valency, thus is not represented in Voice.

(vi) Certain languages may introduce a rich intensional operator (or intensional "context") which anchors a clause to its $Person, Time, World$, and *Location* coordinates. The structural representation of this context drives indexical shift effects.

(vii) The intensional "context" ≠ Kaplanian utterance context:

a. **Conceptual differences:** The latter is a complex primitive whose coordinates are fixed with respect to each other. The latter comprises a series of operators that quantify over the same modalities as the utterance-context: but these operators may covary independently of each other, thus don't constitute a complex *unit* in the same way.
b. **Structural differences:** The intensional context is syntactically represented in the highest projection of the clausal functional sequence, i.e. in SpeechActP. The Kaplanian utterance context is probably not syntactically represented at all, as it doesn't obviously participate in syntactic operations or make its syntactic presence felt via minimality effects. However, the Kaplanian utterance context constrains reference assignment at LF in all languages.
c. Thus the intensional context can simulate the effect of a Kaplanian utterance context but crucially not overwrite it (contra Anand (2006)).

(viii) Intensional context ≠ Perspectival Center:

a. **Conceptual differences:** The intensional context anchors a clause to an immediately superordinate speech event, thereby ensuring that the $Person, Time, World$, or *Location* parameters of that clause are evaluated with respect to the higher one. The values for these parameters are deterministically set based on the properties of the selecting speech predicate. This is why the referent of a shifted indexical must be denoted by a DP in the minimally c-commanding clause (Baker 2008). The perspectival center is less restricted by nature. The same predication may be viewed from distinct perspectives – i.e. the values

of the $Person, Time, World$, or *Location* parameters of the Perspectival Center may vary as a function of differences in speaker-intent, common ground, presupposition, and the like.

b. **Structural differences:** The intensional context is represented on the highest projection of the clausal functional sequence, specifically in SpeechActP. The perspectival center is represented in PerspP which, however, does not have a designated position in the functional sequence. It may rather occur in different positions in different languages – much like a Neg or Focus operator (Rizzi 1997) – and also across different categories of phrase. PerspP is a property of phases more generally: a given phase may have at most one PerspP.

c. The elements in [Spec, SpeechActP] and [Spec, PerspP] are both silent pro-forms. The former is obligatorily controlled by a designated argument of the immediately superordinate speech predicate. The latter is non-obligatorily controlled by a DP denoting a perspective-holder toward the PerspP.

d. The intensional context is responsible for indexical shift. The perspectival center is responsible for anaphora.

(ix) Anaphors cannot directly trigger agreement ("Anaphor Agreement Effect (AAE)"). However, languages may differ in the ways in which they avoid a violation of the AAE. Tamil avoids the AAE by allowing the T head to get its ϕ-features valued by another element in its local domain, specifically the operator in [Spec, PerspP].

Although the central theses of the dissertation have been based on a detailed investigation of binding patterns in Tamil, it appears that many of them may be easily extended to other languages with perspectival phenomena, as we have seen above with empathy phenomena in Japanese, binding across subjunctives in Icelandic and Italian, and across spatial PPs in Norwegian and Dutch. It is to be hoped that the two-step binding model can yield insight into referential dependencies in other languages as well. A further extension would be to see how this model could be used as a template for understanding perspectival dependencies along other dimensions, such as $Time, World$, and *Location* to help further our understanding of other phenomena besides binding, such as sequence-of-tense and -mood, double-access readings, and the properties of clausal finiteness and anchoring more generally.

Bibliography

Adger, David and Gillian Ramchand. 2005. Merge and Move: *Wh*-dependencies revisited. *Linguistic Inquiry* 36 2: 161–193.

Adger, David and Gillian Ramchand. 2006. Psych nouns and predication. In *Proceedings of NELS*, vol. 36.

Aikhenvald, Alexandra and R.M.W. Dixon. 2006. *Serial verb constructions: a cross-linguistic typology*. OUP, Oxford.

Alexiadou, Artemis and Elena Anagnostopoulou. 2004. Voice morphology in the causative-inchoative alternation: evidence for a non-unified structural analysis of unaccusatives. In *The unaccusativity puzzle*, edited by Artemis Alexiadou, Elena Anagnostopoulou, and Martin Everaert, OSTL, pp. 114–136. Oxford University Press, Oxford.

Amritavalli, R. and K.A. Jayaseelan. 2005. Finiteness and negation in Dravidian. In *The Oxford Handbook of Comparative Syntax*, edited by Guglielmo Cinque and Richard S. Kayne, pp. 178–220. OUP, Oxford.

Anand, Pranav. 2006. *De de se*. Ph.D. thesis, Massachusetts Institute of Technology.

Anand, Pranav and Andrew Nevins. 2004. Shifty operators in changing contexts. In *Proceedings of SALT*.

Annamalai, E. 1997. *Adjectival clauses in Tamil*. Institute for the study of language and cultures of Asia and Africa, Tokyo.

Annamalai, E. 1999. Lexical anaphors and pronouns in Tamil. In *Lexical and anaphors and pronouns in selected South Asian languages: a principled typology*, edited by Barbara C. Lust, Kashi Wali, James W. Gair, and K.V. Subbarao, pp. 169–216. Mouton de Gruyter.

Asher, R.E. and E. Annamalai. 2002. *Colloquial Tamil: the complete course for beginners*. Routledge, London.

Austen, Jane. 1816. *Emma*. .

Baker, Mark C. 2008. *The syntax of agreement and concord*. Cambridge University Press, United Kingdom.

Banfield, Ann. 1982. *Unspeakable sentences: Narration and representation in the language of fiction*. Routledge and Kegan Paul, Boston.

Beletti, Adriana and Luigi Rizzi. 1988. Psych verbs and θ-theory. *Natural Language and Linguistic Theory* 6 3: 291–352.

Bianchi, Valentina. 2003. On finiteness as logophoric anchoring. In *Temps et point de vue/Tense and Point of View*, edited by Jacqueline Guéron and L. Tasmovski, pp. 213–246. Université Paris X, Nanterre.

Bobaljik, Jonathan David. 2008. Missing persons: a case study in morphological universals. *The Linguistic Review* 25: 203–230.

Borer, Hagit. 1989. Anaphoric AGR. In *The null subject parameter*, edited by Osvaldo Jaeggli and Ken Safir, pp. 69–109. Kluwer Academic Publishers, Dordrecht.

Borik, Olga and Tanya Reinhart. 2004. Telicity and perfectivity: two independent systems. In *Proceedings of LOLA 8 (Symposium on logic and language)*, edited by L. Hunyadi, G. Rakosi, and E. Toth, pp. 13–34. Debrecen, Hungary.

Bošković, Željko. 2007. On the locality and motivation of Move and Agree: an even more minimal theory. *Linguistic Inquiry* 38: 589–644.

Bouchard, Denis. 1984. *On the content of empty categories*. Dordrecht: Foris.

Bruening, Benjamin. To Appear. Word formation is syntactic: adjectival passives in English. *Natural Language and Linguistic Theory* .

Büring, Daniel. 2005. *Binding Theory*. Cambridge Textbooks in Linguistics. Cambridge University Press, Cambridge.

Burzio, Luigi. 1986. *Italian Syntax*. Reidel, Boston.

Butler, Jonny and George Tsoulas. 2006. *Get*-passives, raising, and control. Ms. University of York.

Butt, Miriam. 1995. *The structure of complex predicates in Urdu*. Ph.D. thesis, Stanford University, Stanford, CA.

Butt, Miriam. 2005. Complex predicate compendium. Handout of talk given in Tromsø.

Chierchia, Gennaro. 1989. Structured meanings, thematic roles, and control. In *Properties, types, and meaning*, edited by Gennaro Chierchia, Barbara Partee, and Raymond Turner, vol. II: Semantic Issues of *Studies in Linguistics and Philosophy*, pp. 131–166. Kluwer Academic Publishers, Dordrecht.

Chierchia, Gennaro. 2004. A semantics for unaccusatives and its syntactic consequences. In *The unaccusativity puzzle: Explorations of the syntax-lexicon interface*, edited by E. Anagnostopoulou A. Alexiadou and M. Everaert, Oxford Studies in Theoretical Linguistics, pp. 22–59. OUP, Oxford.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, Noam. 1986a. *Knowledge of Language: its nature, origin, and use*. Praeger, New York.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In *Step by Step: Essays on minimalism in honor of Howard Lasnik*, edited by Roger Martin, David Michaels, and Juan Uriagereka. MIT Press, Cambridge, Mass.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A Life in Language*, edited by Michael Kenstowicz. MIT Press, Cambridge, Mass.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective*. Oxford Studies in Comparative Syntax. Oxford University Press.

Cinque, Guglielmo. 2004. *Restructuring and functional heads*, vol. 4 of *The cartography of syntactic structures*. OUP, Oxford.

Clements, G.N. 1975. The logophoric pronoun in Ewe: its role in discourse. *Journal of West African Languages* 2: 141–177.

Cole, Peter, Gabriela Hermon, and Cher Leng Lee. 2001. Grammatical and discourse conditions on long distance reflexives in two chinese dialects. In *Long-distance reflexives*, edited by Gabriela Hermon Peter Cole and C.T.J. Huang, pp. 1–46. Academic Press, New York.

Cristofaro, Sonia. 2005. *Subordination*. OUP.

Culy, Christopher. 1994. Aspects of logophoric marking. *Linguistics* 32 6: 1055–1094.

Curnow, Timothy Jowan. 2002. Verbal logophoricity in African languages. In *Proceedings of the 2002 Conference of the Australian Linguistic Society*.

Déchaine, Rose-Marie and Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33 3: 409–442.

Delfitto, Denis and Gaetano Fiorin. 2011. Person features and pronominal anaphora. *Linguistic Inquiry* 42 2: 193–224.

Dowty, David. 1979. *Word-meaning and Montague grammar*. Reidel, Dordrecht.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67 3: 547–619.

Embick, David. 2004a. On the structure of resultative participles in English. *Linguistic Inquiry* 35: 355–392.

Embick, David. 2004b. Unaccusative syntax and verbal alternations. In *The unaccusativity puzzle: Explorations of the syntax-lexicon interface*, edited by E. Anagnostopoulou A. Alexiadou and M. Everaert, Oxford Studies in Theoretical Linguistics, pp. 137–158. OUP, Oxford.

Faltz, Leonard. 1977. *Reflexivization: a study in universal syntax*. Ph.D. thesis, UC Berkeley.

Fillmore, Charles. 1997. *Lectures on deixis*. CSLI, Stanford.

Giorgi, Alessandra. 2006. From temporal anchoring to long distance anaphors. *Natural Language and Linguistic Theory* .

Giorgi, Alessandra. 2010. *Towards a syntax of indexicality*. Oxford Studies in Theoretical Linguistics. Oxford University Press, Oxford.

Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10: 279–326.

Grimshaw, Jane. 1982. On the lexical representation of Romance reflexive clitics. In *The mental representation of grammatical relations*, edited by Joan Bresnan, pp. 87–148. MIT Press, Cambridge, MA.

Haegeman, Liliane. 2004. A dp-internal Anaphor Agreement Effect. *Linguistic Inquiry* 35 4: 704–712.

Hale, Ken and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, edited by Ken Hale and Samuel Jay Keyser. MIT Press, Cambridge, Mass.

Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, edited by Ken Hale and Samuel Jay Keyser. MIT Press, Cambridge, Mass.

Harbour, Daniel. 2011. Valence and atomic number. *Linguistic Inquiry* 42 4: 561–594.

Harley, Heidi and Elizabeth Ritter. 2002. Person and number in pronouns: a feature-geometric analysis. *Language* 78 3: 482–526.

Heim, Irene. 1982. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9: 183–221.

Heim, Irene. 1998. Anaphora and semantic interpretation: a reinterpretation of Reinhart's approach. In *The interpretive tract*, edited by Uli Sauerland and Olin Percus, vol. MIT Working Papers in Linguistics 25, pp. 205–246. Cambridge, MA: MITWPL.

Heim, Irene and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell Publishers.

Heinat, Fredrik. 2008. *Probes, pronouns and binding in the Minimalist Program*. VDM Verlag.

Hellan, Lars. 1988. *Anaphora in Norwegian and the theory of grammar*, vol. 32 of *Studies in Generative Grammar*. Foris Publications, Dordrecht.

Hellan, Lars. 1991. Containment and connectedness anaphors. In *Long-distance anaphora*, edited by Jan Koster and Eric Reuland, pp. 27–48. Cambridge University Press, Cambridge.

Hestvik, Arild and William Philip. 2001. Syntactic vs. logophoric binding: evidence from Norwegian child language. In *Long-distance reflexives*, edited by Peter Cole, Gabriella Hermon, and C.T. James Huang,

vol. 33 of *Syntax and semantics*, pp. 119–139. Academic Press, New York.

Hicks, Glyn. 2009. *The derivation of anaphoric relations*. Linguistik Aktuell. John Benjamins.

Hooper, Joan and Sandra Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4: 465–497.

Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30: 69–96.

Huang, C.C. James and C.T. Jane Tang. 1991. The local nature of the long-distance reflexive in Chinese. In *Long-distance anaphora*, edited by Eric Reuland Jan Koster, pp. 263–282. Cambridge University Press.

Huang, James and Luther Liu. 2001. Logophoricity, attitudes, and *ziji* at the interface. In *Long-distance reflexives*, edited by Peter Cole, Gabriella Hermon, and James Huang, pp. 141–195. Academic Press, New York.

Huang, Yan. 2000. *Anaphora*. OUP, Oxford.

Jayaseelan, K.A. 1997. Anaphors as pronouns. *Studia Linguistica* 51 2: 186–234.

Jayaseelan, K.A. 2004. The serial verb construction in Malayalam. In *Clause structure in South Asian languages*, edited by Veneeta Dayal and Anoop Mahajan, Studies in Natural Language and Linguistic Theory, pp. 67–92. Kluwer, Netherlands.

Kaplan, David. 1989. Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In *Themes from Kaplan*, edited by John Perry, Joseph Almog, and Howard Wettstein, pp. 481–563. Oxford University Press, Oxford.

Karttunen, Lauri. 1974. Presupposition and linguistic context. *Theoretical Linguistics* 1: 181–94.

Kemmer, Suzanne. 2003. *The middle voice*. Typological studies in language. John Benjamins, Philadelphia.

Koopman, Hilda and Dominique Sportiche. 1989. Pronouns, logical variables, and logophoricity in Abe. *Linguistic Inquiry* 20 4: 555–588.

Koster, Jan and Eric Reuland (eds.) . 1991a. *Long-distance anaphora*. CUP, Cambridge.

Koster, Jan and Eric Reuland. 1991b. Long-distance anaphora: an overview. In *Long-distance anaphora*, edited by Jan Koster and Eric Reuland, pp. 1–27. Cambridge University Press, Cambridge.

Kratzer, Angelika. 1996. Severing the external argument from its verb. In *Phrase structure and the lexicon*, edited by Johan Rooryck and Laurie Zaring, pp. 109–137. Kluwer, Dordrecht.

Kratzer, Angelika. 2000. Building statives. In *Proceedings of the 26th annual meeting of the Berkeley Linguistics Society*, pp. 385–399.

Kratzer, Angelika. 2009. Making a pronoun: fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40 2: 187–237.

Krishnamurti, Bhadriraju. 2003. *The Dravidian Languages*. CUP, Cambridge.

Kuno, Susumo. 1987. *Functional syntax – anaphora, discourse and empathy*. Chicago University Press, Chicago.

Landau, Idan. 2012. *Control in Generative Grammar: A Research Companion*. Cambridge University Press.

Levin, Beth. 1993. *English verb classes and alternations: a preliminary investigation*. University of Chicago Press, Chicago.

Levin, Beth and Malka Rappaport-Hovav. 2005. *Argument realization*. Cambridge University Press, Cambridge.

Lewis, David. 1979. Attitudes de dicto and de se. *Philosophical Review* 88: 513–43.

Lidz, Jeffrey. 2001. The argument structure of verbal reflexives. *Natural Language and Linguistic Theory* 19 2: 311–353.

Lidz, Jeffrey. 2004. Causation and refexivity in Kannada. In *Clause Structure in South Asian Languages*. Springer.

Lidz, Jeffrey and Alexander Williams. 2005. On the absence of reflexive benefactives in Kannada. In *The Yearbook of South Asian Languages and Linguistics*, edited by Tanmoy Bhattacharya Rajendra Singh, pp. 237–254. Mouton de Gruyter, Berlin.

Lødrup, Helge. 2007. A new account of simple and complex reflexives. *Journal of comparative Germanic Linguistics* 10: 183–201.

Malamud, Sophia. 2006. *Semantics and pragmatics of arbitrariness*. Ph.D. thesis, University of Pennsylvania, Philadelphia.

Manzini, Rita and Ken Wexler. 1987. Parameters, binding theory, and learnability. *Linguistic Theory* 18: 413–444.

Marantz, Alec. 1984. *On the nature of grammatical relations*. MIT Press, Cambridge, MA.

Marantz, Alec. 1993. Implications of asymmetries in double object constructions. In *Theoretical aspects of Bantu grammar*, edited by Sam Mchombo. CSLI Publications, Stanford.

McCloskey, James. 2002. Resumption, successive cyclicity, and locality of operations. In *Derivation and explanation*, edited by Samuel Epstein and Daniel Seely, pp. 184–226. Blackwell Publishers, Oxford.

McIntyre, Andrew. 2005. The semantic and syntactic decomposition of *get*: an interaction between verb meaning and particle placement. *Journal of semantics* 22: 401–438.

McIntyre, Andrew. 2010. The BECOME = CAUSE hypothesis and the polysemy of *get*. Ms. University of Neuchâtel.

McIntyre, Andrew. 2011. English *get*-passives, middle voice and causative-passive ambiguities. Ms. University of Neuchâtel.

Medová, Lucie. 2009. *Reflexive clitics in the Slavic and Romance languages*. Ph.D. thesis, Princeton University.

Minkoff, Seth. 2003. Consciousness, backward coreference, and logophoricity. *Linguistic Inquiry* 35 3: 485–494.

Nickel, Klaus Peter. 1994. *Samisk grammatikk*. Davvi Girji, Karasjok.

Oshima, David. 2007. On empathic and logophoric binding. *Research on language and computation* 5: 19–35.

Oshima, David Y. 2004. *Zibun* revisited: empathy, logophoricity, and binding. *University of Washington Working Papers in Linguistics* 23: 175–190.

Pederson, Eric. 2008. Event realization in Tamil. In *Crosslinguistic perspectives on argument structure: implications for learnability*, edited by Melissa Bowerman and Penelope Brown, pp. 331–355. Lawrence Erlbaum Associates, New York.

Pesetsky, David. 1982. *Paths and categories*. Ph.D. thesis, MIT.

Pesetsky, David and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In *Phrasal and clausal architecture: syntactic derivation an dinterpretation*, edited by Wendy K. Wilkins Simin Karimi, Vida Samiian, pp. 262–294. Amsterdam: John Benjamins.

Pica, Pierre. 1987. On the nature of the reflexivization cycle. In *Proceedings of NELS 17*, edited by J. McDonough and B. Plunkett. University of Massachusetts, Amherst, GLSA.

Pollard, Carl and Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.

Progovac, Liljana. 1993. Long-distance reflexives: movement to Infl versus relativized SUBJECT. *Linguistic Inquiry* 24: 755–772.

Pylkkänen, Liina. 2008. *Introducing arguments*. MIT Press, Cambridge, Mass.

Quer, Josep. 2005. Context shift and indexical variables in sign languages. In *Proceedings of SALT 15*.

Ramchand, Gillian. 2008. *Verb meaning and the lexicon: a first phase syntax*. Cambridge University Press, Cambridge.

Reinhart, Tanya. 1983a. *Anaphora and semantic interpretation*. Croon Helm linguistics series: Anaphora. Croon Helm.

Reinhart, Tanya. 1983b. Coreference and bound anaphora: a restatement of the anaphora questions. *Linguistics and Philosophy* 6: 47–88.

Reinhart, Tanya. 2000. The theta system: syntactic realization of verbal concepts. *UiL OTS Working Papers* .

Reinhart, Tanya and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24: 657–720.

Reinhart, Tanya and Tal Siloni. 2004. Against an unaccusative analysis of reflexives. In *The unaccusativity puzzle*, edited by Artemis Alexiadou, Elena Anagnostopoulou, and Martin Everaert, OSTL, chap. Oxford, pp. 159–180. Oxford University Press.

Reuland, Eric. 2001a. Anaphors, logophors, and binding. In *Long-distance reflexives*, edited by Peter Cole, Gabriella Hermon, and C.-T. James Huang, pp. 343–370. Academic Press, London.

Reuland, Eric. 2001b. Primitives of binding. *Linguistic Inquiry* 32 3: 439–492.

Reuland, Eric. 2011. *Anaphora and language design*. MIT Press, Cambridge, Mass.

Rice, Keren. 1986. Some remarks on direct and indirect speech in Slave (Northern Athapaskan). In *Direct and Indirect Speech*. Mouton de Gruyter.

Rice, Keren. 1989. *A grammar of Slave*. Mouton de Gruyter, Berlin.

Ritter, Elizabeth and Martina Wiltschko. 2009. Varieties of INFL: TENSE, LOCATION, and PERSON. In *Alternatives to cartography*, edited by Jeroen van Craenenbroeck and, pp. 153–202. De Gruyter, Berlin.

Rizzi, Luigi. 1990. On the anaphor-agreement effect. *Rivista di Linguistica* 2: 27–42.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar*, edited by Liliane Haegeman, pp. 281–337. Kluwer Academic Publishers, Dordrecht.

Rooryck, Johan and Guido vanden Wyngaerd. 2011. *Dissolving Binding Theory*. Oxford Studies in Theoretical Linguistics. Oxford University Press, Oxford.

Safir, Ken. 2003. Anaphors, movement and coconstrual. In *Grammatik i fokus: Grammar in focus: A Festschrift for Christer Platzack*, edited by Lars-Olof Delsing, Cecilia Falk, Gunlög Josefsson, and Halldór Sigurðsson, pp. 283–303. Lund university, Lund, Sweden.

Schäfer, Florian. 2008. *The syntax of (anti-)causatives. External arguments in change-of-state contexts*, vol. 126 of *Linguistik Aktuell*. John Benjamins, Philadelphia.

Schiffman, Harold. 1995. Online grammar of spoken Tamil.

Schlenker, Philippe. 1999. *Propositional attitudes and indexicality: a cross-categorial approach*. Ph.D. thesis, MIT.

Schlenker, Philippe. 2003a. Indexicality, logophoricity, and plural pronouns. In *Research in Afroasiatic grammar II (selected papers from the fifth conference on Afroasiatic languages, Paris, 2000)*, edited by J. Lecarme, pp. 409–428. Benjamins, Philadelphia.

Schlenker, Philippe. 2003b. A plea for monsters. *Linguistics and Philosophy* 26: 29–120. Kluwer Academic Publishers.

Schlenker, Philippe. 2004. Context of thought and context of utterance: a note on free indirect discourse and the historical present. *Mind and language* 19 3: 279–304.

Schlenker, Philippe. To appear. Indexicality and *de se* reports. In *Handbook of Semantics*, edited by von Heusinger, Maienborn, and Portner. Mouton de Gruyter.

Sells, Peter. 1987. Aspects of logophoricity. *Linguistic Inquiry* 18 3: 445–479.

Selvanathan, Nagarajan. 2009. *Taan, ko(n), and anaphoric dependencies in Tamil*. Master's thesis, National University of Singapore.

Shklovsky, Kirill and Yasutada Sudo. To Appear. Indexical shifting in Uyghur and the syntax of monsters. In *Proceedings of NELS 40*.

Sigurðsson, Halldór Ármann. 1990. Long distance reflexives and moods in Icelandic. In *Modern Icelandic syntax*, edited by Joan Maling and Annie Zaenen, pp. 309–346. Academic Press, New York.

Sigurðsson, Halldór Ármann. 2004. The syntax of person, tense, and speech features. *Italian Journal of Linguistics* 16: 219–251. Edited by Valentina Bianchi and Ken Safir.

Sigurðsson, Halldór Ármann. 2010. Mood in Icelandic. In *Mood in the languages of Europe*, edited by Björn Rothstein and Rolf Thieroff, pp. 33–55. John Benjamins, Amsterdam.

Speas, Margaret. 1999. Person and point of view in Navajo. In *Proceedings of the West Coast Conference on Formal Linguistics 18 (WCCFL 18)*.

Speas, Margaret. 2004. Evidentiality, logophoricity and the syntactic representation of pragmatic features. *Lingua* 114 3: 255–276.

Speas, Peggy and Carol Tenny. 2003. Configurational properties and point-of-view roles. In *Asymmetry in grammar*, edited by Anna Maria Di Sciullo, vol. I: Syntax and semantics, pp. 315–344. John Benjamins, Amsterdam.

von Stechow, Arnim. 2002. Feature deletion under semantic binding: tense, person, and mood under verbal quantifiers. In *Proceedings of NELS 33*.

Steever, Sanford B. 2002. Direct and indirect discourse in Tamil. In *Reported discourse: a meeting ground for different linguistic domains*, edited by Tom Güldemann and Manfred von Roncador, pp. 91–108. John Benjamins, Amsterdam.

Steever, Sanford B. 2005. *The Tamil auxiliary verb system*. Routledge, London.

Stirling, Lesley. 1993. *Switch-reference and discourse representation*. Cambridge University Press, Cambridge.

Stowell, Tim. 1982. The tense of infinitives. *Linguistic Inquiry* 13: 561–570.

Strahan, Tania. 2010. PERSPECTIVE study, Icelandic and Faroese. Ms. University of Iceland.

Sundaresan, Sandhya. To Appear. Making sense of silence: distinguishing between "PRO" and "*pro*": a commentary on the paper by Kissock. *Natural Language and Linguistic Theory* .

Svenonius, Peter. 2008. Spatial P in English. Lingbuzz.

Szabolcsi, Anna. 2009. Overt nominative subjects in infinitival complements in Hungarian. In *Approaches to Hungarian*, edited by Marcel den Dikken and Robert M. Vago, vol. 11 of *Papers from the 2007 NYU Conference*. John Benjamins.

Tucker, Matthew A. 2011. Even more on the anaphor agreement effect: when binding does not agree. Manuscript written at UCSC.

Verkuyl, H.J. 1972. *On the compositional nature of the aspects*. Reidel, Dordrecht.

Wiklund, Anna-Lena, Kristine Bentzen, Gunnar Hrafn Hrafnbjargarson, and Þorbjörg Hróarsdóttir. 2009. On the distribution and illocution of V2 in Scandinavian *that*-clauses. *Lingua* 119: 1914–1938.

Williams, Edwin. 1980. Predication. *Linguistic Inquiry* 11: 203–238.

Woolford, Ellen. 1999. More on the anaphor agreement effect. *Linguistic Inquiry* 30 2: 257–287.

Woolford, Ellen. 2006. Lexical case, inherent case, and argument structure. *Linguistic Inquiry* 37 1: 111–130.

Wurmbrand, Susanne. 2001. *Infinitives: restructuring and clause structure*. Mouton de Gruyter, Berlin.

Wurmbrand, Susi. 2002. Syntactic vs. semantic control. In *Studies in comparative Germanic syntax*, edited by C. Jan-Wouter Zwart and Werner Abraham, pp. 95–129. John Benjamins, Amsterdam.

Wurmbrand, Susi. 2011. Tense and aspect in English infinitives. University of Connecticut.

Zaenen, Annie, Joan Maling, and Hóskuldur Thrainsson. 1985. Case and grammatical functions: the Icelandic passive. *Natural Language Linguistic Theory* 3: 441–483.

Zeijlstra, Hedde. 2010. There is only one way to agree. Talk given at GLOW 33, Wroclaw, Poland.

ISBN xxx-xx-xxxx-xxx-x