TAAN, KO(N) and ANAPHORIC DEPENDENCIES IN TAMIL

NAGARAJAN SELVANATHAN

(B.Arts.(Hons.)), NUS

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF ARTS

DEPARTMENT OF ENGLISH LANGUAGE & LITERATURE

NATIONAL UNIVERSITY OF SINGAPORE

2009

Acknowledgements

Many people have had a positive impact on this thesis and this can be found throughout this work. Dr Kim Chonghyuck, for being a caring and demanding supervisor as well as a friend, deserves the most gratitude. As a supervisor, he made sure that I asked the relevant questions over and over again to ensure that I had asked myself every necessary question. Wherever the thesis makes a positive contribution to the literature is due to Dr Kim. Overall, his words of encouragement and his constant reminder to "never give up" played a large part in me completing this thesis.

A very big thanks to Rong Chen and Qizhong is also required. Their sense of humour as well as their companionship was a great pillar of support. The people at LogicMills, especially Mark & Eunice (Thanks for the office use and all the support), Jeremy and Colman also deserve special mention for always challenging me intellectually as well as being great friends throughout this period.

I also have to thank Naadi, Vigneswari, Puvane, Ganash, Balan, Jagatheswari and Tilshana for providing me with the required data. Being non-linguists, I am grateful to them for their extreme patience whenever I asked them for judgements on the same set of data again and again.

Most importantly, I have to thank my parents, Selvanathan and Kamaladevi. Without their unfailing support, doing my Masters degree would only have been a dream. I dedicate this thesis to them.

Table of Contents

Summary of Thesis

In this thesis, we aim to show the distribution of Tamil *taan* and *ko(n)* and explain what leads to such a distribution. We show that both the Standard Binding Theory (SBT) (Chomsky 1981) and the Reflexivity Framework (Reinhart & Reuland 1993), two influential theories of anaphora, are unable to account satisfactorily for all of the Tamil data. SBT runs into a serious contradiction. We will show that out of three Tamil data sets, only one will follow SBT principles as well as the predictions made by SBT. The other two do not. The Reflexivity Framework, on the other hand, fails to account for the Tamil data for different reasons. While it is able to account for two sets of data, it cannot account for the set of data with psych verbs without including a lot of stipulations which render the theory unfalsifiable.

Motivated by the failings of the two theories to account for Tamil *taan*, we move on to explain the two ways in which anaphoric elements are assigned a value. We adopt Hein and Kratzer (1998)'s theory to do this. We show that binding and the assignment function are the two ways and that VP ellipsis data is best explained using this mechanism. However in Tamil, all the VP ellipsis data cannot be explained by simply assuming *taan* to be a particular type of variable.

We finally move on to describe Sells (1987) theory of logophoricity. We adapt his theory for our own application of the theory to Tamil and show that we can now account for all of the data if we consider *taan* to be a logophoric pronoun. This is a claim about *taan* that no one has made before as far as we are aware. We show that not only can all the data be explained, it also provides us with an explanation for the distribution of *ko(n)*. We reanalyze *ko(n)* as a marker which requires a subject-as-pivot reading and show that

all the predictions made by such an analysis pan out. We also provide an account for the complete set of data seen in this thesis using our analysis. We conclude the thesis with an account of the distribution of pronouns and word order issues by incorporating the Chain Condition which was first formulated by Chomsky.

List of Abbreviations

NOM	Nominative Case
ACC	Accusative Case
DAT	Dative Case
LOC	Locative Case
GEN	Genitive Case
Past	Past Tense
Fut	Future Tense
Pres	Present Tense
3sgm	3rd Person Singular Masculine
Comp	Complementizer
COOR	Coordinating Suffix

Table of the Tamil Pronoun System

	1st Person		2nd Person			3rd Person			
	Singular	Plural	Singular	Plural	Honorific	Singular			Plural
	--	--	--	--	--	*Masculine*	*Feminine*	*Honorific*	--
Nominative Case	*naan*	*naa-ngal*	*nii*	*nii-ngal*	*nii-ngal*	*avan*	*aval*	*avar*	*avargal*
Accusative Case	*enn-ei*	*en-gal-ei*	*unn-ei*	*u-ngal-ei*	*u-ngal-ei*	*avan-ei*	*aval-ei*	*avar-ei*	*avar-gal-ei*
Dative Case	*enn-aki*	*en-gal-iki*	*un-aki*	*u-ngal-iki*	*u-ngal-iki*	*avan-iki*	*aval-iki*	*avar-iki*	*avar-gal-iki*

	3rd Person	
	Neuter	
	Singular	*Plural*
Nominative Case	*taan*	*taa-ngal*
Accusative Case	*tann-ei*	*ta-ngal-ei*
Dative Case	*tan-iki*	*ta-ngal-iki*

Chapter 1 Introduction

1.1 The Problem(s)

The objective of this thesis is twofold; One, to provide a characterization of Tamil *taan* and two, to apply theoretical insights that have arisen in the recent studies of anaphors. In the last fifty years, there has not been a consensus on the status of *taan,* which occurs in many other Dravidian languages other than Tamil. While Caldwell (1956) and Mohanan (1981) consider it an anaphor, Amritavalli (1984) considers it to be a pronoun. Lehman (1989) deems it to be a 4th person pronoun while in Annamalai (1999), *taan* seems to have come full circle and has been characterized as an anaphor again. While these authors have concentrated on different Dravidian languages, the aim here is to shed new insight on Tamil *taan* and provide a satisfactory characterization. Furthermore although there has been much work within the linguistic field in the area of anaphora, the application of new theoretical insights to Tamil has not been undertaken. As far as we are aware, the last explicit treatment of Tamil *taan* was Yadurajan (1987), a work which provides many insights but is ultimately an inadequate treatment of the facts.

Emeneau (1967) and Masica (1976) have, on the basis of many formal features, categorized South India as a distinct linguistic area. As a *major member of the Dravidian family of languages* (Asher 1985: ix) which is part of the South Asian linguistic area, we believe that Tamil has an important role to play in the quest towards a theory of anaphora and, eventually, of language. Building upon ideas formulated in, but not limited to, Chomsky (1981) and Reinhart & Reuland (1993), this thesis hopes to further fill the gap in the understanding of *taan* and by extension, anaphora in general.

An anaphor is understood to be an element lacking in complete information which prevents it from being interpreted fully (Lust 1986) and has to refer to an antecedent in order to be interpreted. Given this definition, an anaphor in a given language can be a reflexive, reciprocal, pronoun or even an empty category (Huang 1994). In this thesis, we will focus on overtly realized forms of anaphora. Plus we will also adopt the following terms used by Büring (2005: 3); 'reflexive/ reciprocal' and 'non-reflexive pronoun' in a bid to be theory-neutral. Their usage will be made apparent in due course. Furthermore we refer to anaphors as a collective term denoting reflexives/ reciprocals as well as non-reflexive pronouns.

Chomsky (1981), being one of the first systematic treatments of anaphora in natural languages, outlined the distribution of reflexives/ reciprocals (what he calls anaphors), non-reflexive pronouns (pronominals) and full NPs (r-expressions) with three principles (Principle A, Principle B and Principle C) which have formed the foundations of our understanding of the distribution of these elements. These principles stipulate features inherent within these elements and demarcate domains within which these elements can occur. Even though the Standard Binding Theory (SBT) can explain a fair bit of cross-linguistic data, empirical problems inevitably arise. Principle A and Principle B together determine that the distribution of reflexives/reciprocals and non-reflexive pronouns should be complementary within the same domain. This does fall out in many examples.

1) a. John_i likes $\text{himself}_{i/ *j}$
   b. John_i likes $\text{him}_{*i/ j}$

In (1a) the reflexive in the object position has to refer to *John* while within the same clause, the non-reflexive pronoun *him* cannot. However such complementary breaks

down in sentences like those illustrated in (2) where the reflexive and non-reflexive pronoun can occur in the same position.

2) Lucie saw a picture of herself/ her. [Reinhart & Reuland (1993:661)]

Based on examples like the above Reinhart & Reuland (1993) (R&R, henceforth), generalize that the common thread between examples (1) and (2) has got to do with argumenthood of the anaphoric element in question. Building on this insight, perhaps first suggested by Partee & Bach (1981), R&R claim that the distribution of anaphors lies not in their inherent properties but the property of the predicate of which they are arguments. And in cases where the anaphor is not an argument of the predicate, R&R claim that their formal theory will have nothing to say about it. By situating the mechanism responsible for the distribution of anaphors within the predicate, R&R's theory is able to explain (1) directly and (2) by virtue of not predicting complementarity.

Thus the literature on the syntactic distribution of anaphors can be broadly categorized as following two schools of thought; 1) as a property of the anaphors themselves, as characterized by SBT, and, 2) as a property of the predicate in which anaphors occur, as characterized by R&R. However we find that the Tamil anaphoric system can be explained straightforwardly by neither SBT nor R&R's theory. Consider the Tamil sentences in (3) and (4).

3) a. **Maaren*$_i$ *tann-ei*$_i$*/ aven-ei*$_i$ *adi-t-aan*
Maran.NOM taan-ACC/ 3sgm-ACC beat-past-3sgm
For: 'Maran beat himself'.

b. *Maaren*$_i$ *tann-ei*$_i$*/ aven-ei*$_i$ *aditi-**ko**-nd-aan*
Maran.NOM taan-ACC/ 3sgm-ACC beat-**kon**-past-3sgm
'Maran beat himself'.

(3) comprises of simple clauses with the Tamil anaphor[1], *taan,* and the non-reflexive pronoun *aven* as the object. In (3a) neither *taan* nor *aven* is allowed as the object. In (3b) when there is an additional suffix *ko(n)* on the verb, both *taan* and *aven* are allowed to occur and take the subject as their antecedent[2]. (3b) shows that the complementarity between reflexives and non-reflexive pronouns which is predicted by SBT does not eixst. These sentences also show that *ko(n)* has a crucial part to play in the interpretation of these sentences. Lidz (1995) argues that in Kannada, a language closely related to Tamil, *ko(n)* plays the part of reflexivizing the predicate along the lines proposed by R&R. Perhaps the same applies to Tamil and consequently this would mean that R&R have a straightforward explanation for (3a) and (3b). However, the picture gets complicated. Consider (4) where the verb is a so-called psych verb[3].

4) *Maaren*$_i$ *tann-ei*$_{i/ *j}$*/ aven-ei*$_{*i/ j}$ *veru-tt-aan*
Maran.NOM taan-ACC/ 3sgm-ACC hate-past-3sgm
Maran hates himself/ him.

In (4) *taan* and *aven* appear to have a complementary distribution meaning that the distribution of *taan* and *aven* in these sentences is predicted by SBT. On one hand, in sentences like (3), the property of the verb dictated by the suffix *ko(n)* does seem to determine the occurence of *taan* and the non-reflexive pronoun. On the other hand in (4), it seems to be the inherent property of the anaphors themselves that dictates their

---

[1] We are intentionally vague in addressing *taan* as an anaphor. We will see in this thesis that the characterization of *taan* is problematic even with our theory neutral labels. We will continue referring to *taan* as an anaphor until we reach a suitable point in the thesis where we will be able to label it more accurately.

2 Once again we have to be noncommittal about the terms that we use in this introductory chapter. What we mean when we say that the pronoun and taan 'pick out' or 'take' their antecedent will have to be left to when we get into the thesis proper.

[3] We will refer to these verbs as psych verbs as defined by Sells (1987). A psych verb is a verb which reports the mental or physiological state of an individual.

distribution. As far as we are aware, this rather puzzling distribution of *taan* and the pronoun in Tamil has not been dealt with in the literature and requires an explanation.

Note that SBT and R&R are mainly syntactic accounts for the distribution of anaphors and we show that a satisfactory answer to the problem outlined above can only be reached if we were to take into consideration the semantic and discourse features of the anaphoric system in Tamil as well.

By providing a characterization of *taan*, we hope to show why *taan* seems to fit neither the principles laid out by SBT nor R&R completely as shown by (3) and (4). We will illustrate and explain the distribution of *taan* as well as *ko(n)* and by the end of this thesis, we hope to have contributed significantly to the understanding of the Tamil *taan*.

1.2 What Tamil?

The variety of Tamil that will be dealt with in this paper is the formal, 'higher' variety as found in Singapore. Tamil is described as being in a situation of diglossia (Ferguson 1959). This means that there are two distinct varieties of Tamil found within the same speech community, each serving its own functional load. The higher variety is used in writing and formal situations. The lower variety can be found in informal conversations and, increasingly in Singapore, Tamil language based television series where a higher degree of realism is desired. Even though the lower variety fulfills the daily conversational needs of Tamil speakers and might even be more ubiquitous, there is one main reason why the higher variety is chosen for this thesis.

This thesis hopes to provide an account of Tamil anaphors which will be applicable to most, if not all, varieties of Tamil spoken all over the world and the high variety lends itself to this purpose better. The main difference between the high variety and the low

variety lies in lexical content as well as phonetic realization of word endings (We refer the interested reader to the appendix of Asher (1985) where he provides a detailed commentary on the differences between the high and low variety of Tamil.). As far as the anaphoric system goes, no major differences have been reported in judgments among the different varieties in the literature or encountered in this author's experience with the language. However Britto (1986) shows that even within the high and low varieties of Tamil, there exist many sub-varieties depending on the level of socio-economic stratification. But since the high variety has its own writing system and prescriptive rules, it is more resistant to such division and regardless of whether it is spoken in South India or Singapore it is more stable. The lower variety, with a multitude of social classes is much less homogeneous within even South India, let alone across countries.

The methodology perused in this thesis is simple. Data from Tamil is presented which will then form the foundation for the various arguments made. The approach here is geared more towards a qualitative rather than a quantitative perspective. The judgments presented in this thesis belong to the author and where judgments have been unclear they have been checked with other native speakers of Tamil. The author, himself, is a native speaker of Tamil having acquired it in his childhood and he further learnt it in school where the 'mother tongue' learning policies of Singapore mean that he was exposed to the high variety of Tamil for at least twelve years in a formal setting. The author is still fluent in reading and writing both the high and low varieties of Tamil and is more than suitable to provide Tamil judgments. Data taken from elsewhere are duly acknowledged.

1.3 Thesis Content

The entire thesis is divided into three chapters. In Chapter 2 the Tamil data illustrating the distribution of *taan* will be provided. Essentially the distribution of *taan*, the verbal suffix *ko(n)* as well as their behavior in simple and embedded clauses will be illustrated. We then review two influential syntactic accounts of anaphors; SBT and Reinhart & Reuland (1993)'s Reflexivity Framework. We show that neither are able to account for all of the Tamil data. Even in those sentences where these theories do provide a *prima facie* description, when we delve deeper into predictions that these theories make, we do not find what we would expect. This means that neither major theory is able to account for all the Tamil data that we have here. With this in mind, we turn to other properties of *taan* in the next chapter.

In Chapter 3, we look at the VP ellipsis data in Tamil. We illustrate Heim & Kratzer (1998)'s explanation of VP ellipsis which utilizes the concept of free and bound variables. We will show that while this provides a neat characterization of VP ellipsis cross-linguistically, the classification of *taan* strictly as either a free or bound variable does not work. We then move on to illustrate our own account of *taan*. We will be arguing that *taan* is actually a logophoric pronoun and showing that such a characterization does have a lot of explanatory power. We use Sells' (1987) take on logophoricity and show that it does go some way towards explaining the Tamil data. We will show that by using his primitives of source, self and pivot, we can account for all the Tamil data. Although not the main objective of this thesis, we also provide an explanation of *ko(n)* and show that it should be considered a pivot marker. We also incorporate the Chain Condition which will enable us to derive the differences in the distribution of the

pronoun *aven* and *taan*. Using our account of *taan* as a logophoric pronoun and *ko(n)* as a pivot marker, we then show the coreference possibilities of *taan* in embedded clauses. We also explain why *ko(n)* is incompatible with psych verbs and finally provide an account for the VP ellipsis data. We conclude the thesis with a small section summarizing the thesis with a laundry list of the questions we have answered and questions we have not attempted. The latter will have to guide future research on *taan*.

Chapter 2 *Taan* as a Syntactic Element

The main objective of this chapter is to introduce the basic data set as well as show that existing syntactic accounts of the behavior and distribution of *taan* are not satisfactory. Toward this aim, we will illustrate two influential syntactic accounts of anaphors; SBT and R&R's Reflexivity framework. Both accounts will be shown to be inadequate in explaining *taan*. First, an introduction to the basics of the distribution and interpretation of *taan* is provided.

2.1 Basics of taan

An overview of the data to be presented suggests that *taan* is sensitive to its syntactic position as a subject or non-subject as well as the case it is in. Therefore, we will categorise the data sets into the following; i) *Taan* as a non-subject with a nominative subject, ii) *Taan* as a non-subject with a dative subject, iii) *taan* as a subject with nominative case and iv) *taan* as a subject with dative case. We will first look at clauses with *taan* as a non-subject and a nominative subject.

2.1.1 Nominative Subjects and non-Subject *taan*

The data which follows show *taan* as an object with a nominative subject in a simple sentence.

5) a. **Maaren* *tann-ei*$_j$ *adi-t-aan*
Maran.NOM taan-ACC beat-past-3sgm
For: 'Maran beat himself'.

b. *Maaren*$_i$ *tann-ei*$_{i/ *j}$ *aditi-**ko**-nd-aan*
Maran.NOM taan-ACC beat-**kon**-past-3sgm
'Maran beat himself'.

(5) shows the transitive verbs 'beat'. The (a) and (b) sentences show that without the *ko(n)* suffix, *taan* cannot pick out the subject as the antecedent. The following shows the simple sentence as an embedded clause.

6) a. [*Maaren*$_i$ *tann-ei*$_{*i/\ j/\ *k/*m}$ *adi-t-aan*] *enru* *Somu*$_j$
Maran.NOM taan-ACC beat-past-3sgm comp Somu.NOM
Raman-idam$_k$ *co-nn-aan*
Raman-LOC say-past-3sgm
'Somu told Raman that Maran beat him'.

b. [*Maaren*$_i$ *tann-ei*$_{i/\ *j/\ *k/*m}$ *aditi-**ko**-nd-aan*] *enru* *Somu*$_j$
Maran.NOM taan-ACC beat-**kon**-past-3sgm comp Somu
Raman-idam$_k$ *co-nn-aan*
Raman-LOC say-past-3sgm
'Somu told Raman that Maran beat himself'.

(6a) shows that when the clause without the *ko(n)* suffix is embedded, *taan* can only pick out the matrix subject *Somu* as its antecedent. When the clause with *ko(n)* is embedded, *taan* can now only pick out the embedded subject *Maran* as its antecedent as shown in (d). Note that *taan* never picks out a discourse antecedent or the matrix non-subject. The same judgments are present on a verb like 'praise' shown below.

7) a. **Maaren* *tann-ei* *paarat-in-aan*
Maran.NOM taan-ACC praise-past-3sgm
For: Maran praised himself.

b. *Maaren*$_i$ *tann-ei*$_{i/\ *j}$ *paarati-**ko**-nd-aan*
Maran.NOM taan-ACC praise-**kon**-past-3sgm
For: Maran praised himself.

c. [*Maaren*$_i$ *tann-ei*$_{*i/\ j/\ *k/*m}$ *paarat-in-aan*] *enru*
Maran.NOM taan-ACC praise-past-3sgm comp
Somu$_j$ *Raman-idam*$_k$ *co-nn-aan*
Somu.NOM Raman-LOC say-past-3sgm
'Somu told Raman that Maran praised him'.

d. [*Maaren*$_i$ *tann-ei*$_{i/\ *j/\ *k/*m}$ *paarati-**ko**-nd-aan*] *enru* *Somu*$_j$
Maran.NOM taan-ACC praise-**kon**-past-3sgm comp Somu
Raman-idam$_k$ *co-nn-aan*
Raman-LOC say-past-3sgm
'Somu told Raman that Maran praised himself'.

However as shown in Chapter 1: (4), not all sentences with a nominative subject and object *taan* behave the in same way. Consider (8) and (9).

8) a. *Maaren*$_i$ *tann-ei*$_{i/ *j}$ *veru-tt-aan*
Maran.NOM taan-ACC hate-past-3sgm
'Maran hates himself'.
b. **Maaren* *tann-ei* *veruti-**ko**-nd-aan*
Maran.NOM taan-ACC hate-**kon**-past-3sgm
'Maran hates himself'.

Contrary to (5), (6) and (7), *ko(n)* is not allowed as a verbal suffix with *veru* 'hate' as seen in (8b). In simple sentences like (8a) *taan* can pick out the subject as an antecedent. We now look at embedded clauses.

9) a. [*Maaren*$_i$ *tann-ei*$_{i/ j/ *k/*m}$ *veru-tt-aan*] *enru*
Maran.NOM taan-ACC hate-past-3sgm comp
Somu$_j$ *Raman-idam*$_k$ *co-nn-aan*
Somu.NOM Raman-LOC say-past-3sgm
'Somu told Raman that Maran hates himself/ him (Somu)'.
b. *[*Maaren* *tann-ei* *veruti-**ko**-nd-aan*] *enru*
Maran.NOM taan-ACC hate-**kon**-past-3sgm comp
Somu$_j$ *Raman-idam*$_k$ *co-nn-aan*
Somu.NOM Raman-LOC say-past-3sgm
For: 'Somu told Raman that Maran hates himself/ him (Somu)'.

When the simple clause is embedded as in (9a), *taan* can pick out the matrix as well as embedded subject as an antecedent. In light of (8b), (9b) is unsurprisingly ungrammatical. When a verb like *virumbi-* 'like' is used, we see the same judgments as shown in (8) and (9) although we do not show these sentences here. We claim that the ban on *ko(n)* on *veru* and *virumbi* is a result of the verbs being psych verbs. This does have some merit to it as we will see in the next section that when the subject has dative case, *ko(n)* is also not allowed on the verb. In all instances of such sentences, the verb does seem classifiable as a psych verb.

2.1.2 Dative Subjects and non-Subject *taan*

The data which follows show *taan* as an object with a subject[4] which carries dative case.

10) a. *Maaren-iki*$_i$ *tann-ei*$_{i/ *j}$ *pidik-um*
Maran-DAT taan-ACC like-fut[5]
'Maran likes himself.'
b. **Maaren-iki* *tann-ei* *pidi-**kon**-um*
Maren-DAT taan-ACC like-**kon**-fut
For: 'Maran likes himself.'

In (10) we can see that *taan*'s distribution mirrors that of *taan* in (8). (10a) and (10b) show that *taan* can pick out the subject as its antecedent and that *ko(n)* is not allowed.

11) a. [*Maaren-iki*$_i$ *tann-ei*$_{i/ j/ *k/*m}$ *pidik-um*] *enru* *Somu*$_j$
Maran-DAT taan-ACC like-fut comp Somu.NOM
Raman-idam$_k$ *co-nn-aan*
Raman-LOC say-past-3sgm
'Somu told Raman that Maran likes himself/ him (Somu).'
b. *[*Maaren-iki* *tann-ei* *pidi-**kon**-um*] *enru* *Somu*
Maren-DAT taan-ACC like-**kon**-fut comp Somu.NOM
Raman-idam$_k$ *co-nn-aan*
Raman-LOC say-past-3sgm
For: 'Somu told Raman that Maran likes himself/ him (Somu).'

When the clause is embedded within a larger sentence as in (11a), *taan* can pick out the matrix or the embedded subject as its antecedent. (11b) is ungrammatical because of the presence of *ko(n)*on the embedded verb. Verbs such as the above in Tamil as a class tend not to allow objects let alone a self referring object (as in *pasi* 'hunger', *kobam* 'anger' and so on) but the few other verbs which do allow a dative marked subject with

[4] Although the subject in these sentences do not have nominative case like the subjects in section 2.1.1, we will keep referring to these as the subject due to the fact that in these sentence *Maran* is the most relevant entity who 'has' the state described by the verb.
[5] In this thesis, we label the suffix on dative verbs merely as 'fut' to mean 'future tense'. In certain verbs, *um* does carry the future tense interpretation, however, with verbs like 'know' and 'like' above, even with the suffix, the present tense interpretation yields. This is probably because these types of verbs are also stative verbs.

taan as an object show the same distribution. For instance, (12) shows the verb *theri-* 'know'.

12) a. *Maaren-iki*$_i$ *tann-ei*$_{i/ *j}$ *theriy-um*
Maran-DAT taan-ACC know-fut
'Maran knows himself.'

b. **Maaren-iki* *tann-ei* *theri-**kon**-um*
Maren-DAT taan-ACC know-**kon**-fut
For: 'Maran knows himself.'

c. [*Maaren-iki*$_i$ *tann-ei*$_{i/ j/ *k/*m}$ *theriy-um*] *enru* *Somu*$_j$
Maran-DAT taan-ACC know-fut comp Somu.NOM
co-nn-aan
say-past-3sgm
'Somu said that Maran knows himself/ him.'

d. *[*Maaren-iki* *tann-ei* *theri-**kon**-um*] *enru* *Somu*
Maren-DAT taan-ACC know-**kon**-fut comp Somu.NOM
co-nn-aan
say-past-3sgm
For: 'Somu said that Maran knows himself/ him.'

In (12) we see the now familiar distribution of *ko(n)* and *taan* in these sentences. While *ko(n)* is not allowed in simple or embedded clauses, *taan* can refer to the subject while in simple clauses, and the embedded as well as matrix subjects in embedded clauses. The picture of the basic facts, however, would not be complete without looking at the occurrence of *taan* as a subject as well. Thus far we have seen that regardless of whether *taan* refers to a local or long-distance antecedent, the antecedent has to be a subject. We will see that this subject orientation of *taan* continues even when *taan* itself is a subject.

2.1.3 Nominative Subject *taan*

Apart from occurring as the object, *taan* can also occur as the subject both in nominative and dative case and in this sub-section, we provide an overview of the instance when *taan* occurs in the nominative case.

13) a. **taan* *paadath-ei* *padi-tt-aan*
taan.NOM lesson-ACC study-past-3sgm
For: Self studied the lesson.

b. **taan* *paadeth-ei* *padithi-**ko**-nd-aan*
taan.NOM lesson-ACC study-**kon**-past-3sgm
For: Self studied the lesson.

(13b) shows that *ko(n)* is not allowed on these sentences. But (13a) shows that even without *ko(n)*, *taan* with nominative case cannot occur without an antecedent. The sentences are ungrammatical regardless of whether there is a *ko(n)* suffix on the verb. This is due to the ban on discourse antecedents for *taan* that we have already encountered when *taan* occurred as an object in the earlier sections. As predicted by this generalization, when (13a) is embedded, the sentence is grammatical as shown below in (14a).

14) a. [*taan* $_{i/*j/*k}$ *paadeth-ei* *padi-tt-aan*] *enru*
taan.NOM lesson-ACC study-past-3sgm comp
$Somu_i$ $Raman\text{-}idam_j$ *co-nn-aan*
Somu.NOM Raman-LOC say-past-3sg
Somu told Raman that he (Somu) studied the lesson.

b. *[*taan* *paadeth-ei* *padithi-**ko**-nd-aan*] *enru*
taan.NOM lesson-ACC study-**kon**-past-3sgm comp
Somu *Raman-idam* *co-nn-aan*
Somu.NOM Raman-LOC say-past-3sg
For: Somu told Raman that he (Somu) studied the lesson.

When the simple clause is embedded, *taan* can pick out the matrix subject as its antecedent but only if there is no *ko(n)* on the embedded verb as in (14a). Another thing to note in this sentence is that *taan* in the embedded clause can only pick out the matrix subject as the antecedent and not any other nominal element occurring in the matrix sentence, in this instance the goal *Raman.* (14b) is also ungrammatical due to the presence of *ko(n)* on the embedded verb. The behavior of *taan* is the same even if the clause contains a human object as can be seen in (15) and (16).

15) a. **taan* *Maaran-ei* *adi-tt-aan*
taan.NOM Maran-ACC beat-past-3sgm
For: Self beat Maran.

b. **taan* *Maaran-ei* *adithi-**ko**-nd-aan*
taan.NOM Maran-ACC beat-**kon**-past-3sgm
For: Self beat Maran.

(15a) and (15b) show that *taan* behaves the same way even when there is a potential antecedent in the object position. *Taan* needs an antecedent from the sentence but cannot refer to the human object even if *ko(n)* is present as in (15b). The following shows the clauses in (15) when embedded.

16) a. [*taan* $_{i/*j/*k/*m}$ *Maaran-ei*$_j$ *adi-tt-aan*] *enru*
taan.NOM Maran-ACC beat-past-3sgm comp
Somu$_i$ *Raman-idam*$_k$ *co-nn-aan*
Somu.NOM Raman-LOC say-past-3sg
Somu told Raman that he (Somu) beat Maran.

b. *[*taan* *Maaran-ei* *adithi-**ko**-nd-aan*] *enru*
taan.NOM Maran-ACC beat-**kon**-past-3sgm comp
Somu$_i$ *Raman-idam* *co-nn-aan*
Somu.NOM Raman-LOC say-past-3sg
For: Somu told Raman that he (Somu) beat Maran.

When the clause is embedded, *taan* can also only refer to the matrix subject as in (16a). (16b) shows that *ko(n)* on the embedded verb is not allowed when the embedded subject is *taan*. The generalization that we can come up with in this section is that when *taan* occurs as the subject of an embedded clause, only the matrix subject can be its antecedent. This is the subject oriented nature of *taan* that we have noted in the previous sub-sections. We will see that this generalization holds even when *taan* occurs as a dative subject.

2.1.4 Dative Subject *taan*

The data which follows show *taan* as a subject which carries dative case.

17) a. **tan-iki* *Maaran-ei* *pidi-kum*
taan-DAT Maran-ACC like-fut
For: Self like Maran.

b. **tan-iki* *Maaran-ei* *pidi-**kon**-um*
taan-DAT Maran-ACC like-**kon**-fut
For: Self like Maran.

(17a) and (17b) are ungrammatical due to the fact that there is no antecedent for *taan* within the sentence. (17b) is also ungrammatical due to *ko(n*) on the verb. This is confirmed when these clauses are embedded as shown below.

18) a. [*tan-iki*$_{i/*j/*k/*m}$ *Maaran-ei*$_j$ *pidi-kum*] *enru* *Somu*$_i$
taan-DAT Maran-ACC like-fut comp Somu.NOM
Raman-idam$_k$ *co-nn-aan*
Raman-LOC say-past-3sgm
Somu told Raman that he (Somu) likes Maran.

b. *[*tan-iki* *Maaran-ei* *pidi-**kon**-um*] *enru* *Somu*
taan-DAT Maran-ACC like-**kon**-fut comp Somu.NOM
Raman-idam *co-nn-aan*
Raman-LOC say-past-3sgm
Somu told Raman that he (Somu) likes Maran.

taan which has dative case can only pick out the matrix subject as its antecedent as seen in (18a). We can confirm that the suffix *ko(n)* makes these sentences ungrammatical as seen in (18b). In the next sub-section, we consolidate all the facts we have seen so far.

2.1.5 Consolidation and Summary

In the four sub-sections above, we have seen the behavior of *taan* in various sentences. The table below consolidates these facts.

	***ko(n)* allowed?**	**Local subject as antecedent?**	**Matrix subject as antecedent?**
***taan* as object..**			
..with NOM subject	YES (Obligatory)	YES	NO
..with DAT subject	NO	YES	YES
***taan* as subject..**			
..with NOM case	NO	n.a	YES
..with DAT case	NO	n.a	YES

Table 1: Distribution of *taan*

Table 1 shows all the distributional facts about *taan* seen thus far. Looking at the table, we can see that *taan* behaves more or less uniformly if we leave *taan* as an object with a nominative marked subject out of the picture. When *taan* occurs as an object with a nominative subject, *ko(n)* is obligatory and only the local subject can be the antecedent. Note that this excludes the instances where the verb is a psych verb such as *veru-* 'hate' but yet still has a nominative subject (see (8) and (9)). In these and all other distributions of *taan*, *ko(n)* is not allowed. Furthermore the matrix subject is always available as an antecedent.

Having presented the data set, we now turn to two major syntactic treatments of anaphors in the literature, SBT and R&R, to account for the data. We outline the key concepts in both theories and determine whether either can provide a tenable unitary account for *taan.*

2.2 Standard Binding Theory

Standard Binding Theory (SBT) outlined by Chomsky (1981) is part of the Government and Binding theory which in turn is part of the Principles and Parameters framework of language. SBT aims to explain the distribution of anaphoric and non-anaphoric elements in any given language. Crucial to this enterprise is the idea that the distribution of anaphoric and non-anaphoric elements is intrinsically tied to their inherent properties which are determined by the composition of the features [+/- pronominal] and [+/- anaphor]. Note that [+] and [-] are contradictory feature specifications, i.e. an element cannot be both [+pronominal] and [-pronominal]. With this division and the contradictory nature of the feature specifications, SBT outlines 3 different types of overt elements in natural languages shown below in the table.

	[anaphor]	[pronominal]
Anaphors (Reflexives & Reciprocals)	+	-
Pronouns	-	+
R-Expressions	-	-

Table 2: Typology of elements according to SBT

Table 2 shows that anaphors have the feature specifications [+anaphor][-pronominal], pronouns are [-anaphor][+pronominal] and that r-expressions (referring expressions like full NPs) are [-anaphor][-pronominal]. The implication of this is only apparent if we look at the binding principles which are outlined below. We highlight the version of the theory elucidated in Chomsky (1981).

Principles of SBT

Principle A: Anaphors are bound within their governing domain.

Principle B: Pronouns are free within their governing domain.

Principle C: R-expressions are free.

where 'bound' is defined as follows,
α BINDS β iff
a. α c-commands β, and
b. α and β are coindexed.

and where 'governing domain' is defined as follows,
a governing domain Z is the governing domain for X if Z is the minimal category with a subject containing X, a governor G for X, and where the binding requirements of X and G are satisfiable[6].

where 'governor' is defined as follows,
W, a head, is a governer of Y iff, a) W c-commands Y and, b) no non-IP phrasal category dominates Y but not W.

These properties and principles account for the following facts. First, it allows us to distinguish between elements like *himself* (reflexives), *him* (pronominals), and *John* (r-expression).

19) a. $John_i$ likes $himself_{i/*j}$.

[6] Here we quote the definition of governing domain given in Rizzi (1990). He uses this definition with the addition that the binding requirements of the governer must also be satisfiable within Z.

b. Tom_i said that [$_{IP}$ $John_j$ likes $him_{i/*j}$].
c. Tom told him_i that [$_{IP}$ Mary likes $John_{*i/j}$].

In (19a) *himself* which is a reflexive must be bound within its governing domain. The governing domain here is the clause. According to the definition of governing domain above, the binding requirements of *himself* can be satisfied within the clause and there is also a governor of *himself* within the clause. The governor in this case is the verb which c-commands the object and there is no XP that dominates the object but not the verb. In (19b) the governing domain is the embedded clause for the same reason as (19a) but *him* must be free within this domain. Thus *John* cannot be the antecedent of the pronoun whereas the matrix subject which is outside the embedded clause can be the antecedent. In (19c) the notion of a governing domain is not relevant as *John* which is an r-expression must be free regardless of what the domains are. *Him* in the matrix clause cannot be the antecedent of John because, if they were to corefer, then *him* would be effectively binding *John* but this is ruled out by Principle C. Furthermore, even though *him* is outside the governing domain of *John* and should co-refer as allowed by principle B, *him* c-commands *John* and not vice versa which means that *John* cannot bind *him*.

Apart from attempting a straight explanation of the facts, SBT also makes a few predictions about the distribution of anaphors and pronouns in natural languages. In particular, Principle A and Principle B both take the same governing domain as the point of reference. Since within this domain, Principle A states that anaphors are bound and Principle B states that pronouns are free, we would expect anaphors to occur in positions where pronouns cannot and vice versa. This is the complementarity of anaphors and pronouns which has been well documented in the literature, most recently by Büring (2005). This prediction does work for sentences like (19) presented below as (20).

20) a. [$_{IP}$ John$_i$ likes *himself$_i$* /**him$_i$*].
b. Tom$_i$ said that [$_{IP}$ Mary likes **himself$_i$*/ *him$_i$*].

In (20a) the reflexive *himself* can pick out the subject as the antecedent while the pronoun *him* cannot. In (20b), the situation is reversed. While the pronoun can pick out the matrix subject, the reflexive cannot. However consider the following sentences from R&R: 661.

21) a. Max saw a gun [near himself/ him].
b. Lucie counted five tourists in the room [apart from herself/ her].
c. Lucie saw a picture [of herself/ her].
d. Max likes jokes [about himself/ him].

In (21a-d), both the reflexive and pronoun can refer to the subject. The supposed complementarity of the reflexive and pronoun have apparently broken down here. However this does not necessarily mean that the SBT is wrong. For one, the theory itself has a potential explanation for the breakdown in complementarity. This lies in the manipulation of the term *governing domain*. Since the reflexive and pronoun can both occur in the sentence, if one was to stipulate that the governing domain of the pronoun is the adjunct phrase (square bracketed) while the governing domain of the reflexive is the clause, then we are able to account for the data with SBT. However we will see later that even if we were to manipulate the notion of governing domain, this explanation breaks down in Tamil. We now move on to see the Tamil data presented in Section 2.1 as explained by SBT.

2.2.1 SBT & *taan*

We will look at the Tamil data insofar as SBT applies to *taan* and we will evaluate if SBT can account for the all distribution and reference possibilities of *taan*. We will look at the situation of complementarity in Tamil and show that SBT does not seem to be able to explain what *taan* is.

Looking at the context where *taan* occurs as an object with a nominative subject and *ko(n)* on the verb, we can see that Principle A does account for this piece of data.

22) a. *Maaren*$_i$ *tann-ei*$_{i/ *j}$ *aditi-**ko**-nd-aan*
Maran.NOM taan-ACC beat-**kon**-past-3sgm
'Maran beat himself'.

b. [*Maaren*$_i$ *tann-ei*$_{i/ *j/ *k}$ *aditi-**ko**-nd-aan*] *enru* *Somu*$_j$ *co-nn-aan*
Maran.NOM taan-ACC beat-**kon**-past-3sgm comp Somu say-past-3sgm
'Somu said that Maran beat himself'.

In (22a) *taan* can only pick out the subject as the antecedent and in (22b) the antecedent can only be the embedded subject. If we take the clause to be the governing domain in both sentences and *taan* to be straightforwardly a reflexive, then Principle A does predict that only the clausemate subject can be the antecedent of *taan*. Since SBT specifies the element itself to contain properties which determine its distribution, this means that *taan* is [+anaphor][-pronominal]. However the following sentence shows that the feature specifications of *taan* alone are not responsible for the coreference facts. In the following sentence, *ko(n)* is not present on the embedded verb.

23) a. **Maaren* *tann-ei* *aditi-tt-aan*
Maran.NOM taan-ACC beat-past-3sgm
For: 'Maran beat himself'.

b. [*Maaren*$_i$ *tann-ei*$_{*i/ j/ *k}$ *aditi-tt-aan*] *enru* *Somu*$_j$ *co-nn-aan*
Maran.NOM taan-ACC beat-past-3sgm comp Somu say-past-3sgm
'Somu said that Maran beat himself'.

If *taan* is indeed [+anaphor][-pronominal], we would expect it to always behave like an anaphor as defined by SBT, i.e. obey Principle A. However in (23a) when there is no *ko(n)* on the verb, the sentence is ungrammatical. In (23b) when such a simple sentence is embedded, *taan* can only refer to the matrix subject. This of course means that in these

sentences *taan* does not behave in accordance to Principle A anymore. In fact one could claim that it is behaving more like a pronoun. Whatever the case might be, we cannot maintain that *taan* is [+anaphor][-pronominal] in (23) as we suggested it is in (22). This means that contrary to what SBT tells us, it cannot be the properties of *taan* alone which lead to the coreference facts in Tamil. The picture gets even more complicated for SBT when we consider the coreference facts of *taan* when it is in a sentence with a dative subject.

24) a. *Maaren-iki*$_i$ *tann-ei*$_{i/ *j}$ *pidik-um*
Maran-DAT taan-ACC like-FUT
'Maran likes himself.'

b. [*Maaren-iki*$_i$ *tann-ei*$_{i/ j/ *k}$ *pidik-um*] *enru* *Somu*$_j$
Maran-DAT taan-ACC like-FUT comp Somu.NOM
co-nn-aan
say-past-3sgm
'Somu said that Maran likes himself/ him.'

(24a) shows that *taan* is now behaving like a reflexive as it picks a local antecedent. However in (24b) when *taan* occurs in an embedded clause, the choice of antecedent is ambiguous. Principle A would dictate that the antecedent should be the embedded subject and while this is true, the fact that *taan* can pick out the matrix subject does contradict the same principle. Manipulating the notion of governing domain seems to be the only way out. However if we stipulate that the reason why *taan* is ambiguous in (24b) is due to the fact that *taan* can have different governing domains, then SBT would have to claim that *taan* is not unambiguously an anaphor after all but ambiguous between an anaphor and a pronoun in certain contexts. To complete the picture of the entire set of Tamil data, we will look at *taan* in sentences where it is the subject and evaluate what it tells us about the classification of *taan* with respect to Principle A.

25) a. [*taan*$_i$ *paadeth-ei* *padi-tt-aan*] *enru*
taan.NOM lesson-ACC study-past-3sgm comp
Somu$_i$ *co-nn-aan*
Somu.NOM say-past-3sgm
Somu told Raman that he (Somu) studied the lesson.

b. [*tan-iki*$_i$ *Maaran-ei* *pidi-kum*] *enru* *Somu*$_i$
taan-DAT Maran-ACC like-fut comp Somu.NOM
co-nn-aan
say-past-3sgm
Somu told Raman that he (Somu) likes Maran.

When *taan* occurs as a nominative subject as in (25a) or a dative subject as in (25b), we find that *taan* behaves with accordance to Principle A. This follows as there is no potential antecedent for *taan* in the embedded clauses, thus making the governing domain the entire sentence. This means that picking out the matrix subject does follow Principle A and *taan* behaves like an anaphor. While one could think of *taan* as a pronoun here as its antecedent is outside its domain, we know that *taan* cannot have a discourse antecedent and this must mean that *taan* is an anaphor and that Principle A is indeed active in these sentences.

In (22-25), we have seen that a unified characterization of *taan* is not possible under SBT. This arises from the fact that while (22) and (24) show us that *taan* is a reflexive as defined by SBT, (23) shows us that *taan* cannot be considered a reflexive but a pronoun as it has to be free in its governing domain now in the absence of *ko(n)*. (24) does not make the picture any clearer as *taan* can be ambiguously an anaphor or a pronoun in these sentences. Given the inability of SBT to pin down the characterization of *taan* in the various sentence types that we have seen, we are warranted in rejecting SBT as the correct way in explaining *taan*. But suppose we keep the assumption that *taan* is indeed an anaphor and that the reason why it does not behave like an anaphor in some instances is construction specific. Perhaps *ko(n)* is necessary for demarcating a governing

domain for independent reasons and and it is for this reason that *taan* is unable to pick the local antecedent in (23a). Furthermore note that (24) is very different structurally from (22), (23) and (25). Unlike the rest, (24) does not have full agreement or tense marking. Perhaps this structural difference is the reason why *taan* does not behave like a reflexive in these sentences. If these were true then we can indeed claim that *taan* is a reflexive and that certain constructions like (24) are anomalous and have to be explained by other means. However, even this approach will have problems. This has to do with the predictions that SBT makes.

Recall that SBT predicts a strict complementarity between anaphors and pronouns. This means that where Principle A does account for *taan*, we should find that pronouns cannot occur in the same contexts as *taan*. However, this is not what we find. The following data sets show this. First, we will look at sentences with a nominative subject and *ko(n)* on the verb.

26) a. *Maaren*$_i$ *tann-ei*$_{i/ *j}$*/aven-ei*$_{i/*j}$ *aditi-**ko**-nd-aan*
Maran.NOM taan-ACC/ 3sgm-ACC beat-**kon**-past-3sgm
'Maran beat himself'.

b. [*Maaren*$_i$ *tann-ei*$_{i/ *j/ *k}$ */aven-ei*$_{i/*j/*k}$ *aditi-**ko**-nd-aan*] *enru*
Maran.NOM taan-ACC/3sgm-ACC beat-**kon**-past-3sgm comp
Somu$_j$ *co-nn-aan*
Somu say-past-3sgm
'Somu said that Maran beat himself'.

In (26a) and (26b), *aven* can be just as easily be substituted where *taan* occurs and the sentence would have the same meaning. The indices show that just like *taan*, *aven* can only pick out the embedded subject as its antecedent. Recall that these are the sentences where *taan* seemingly behaved in accordance to Principle A. This suggests that SBT is not the principle which is dictating the distribution of *taan* in at least these sentences. We

shall now move on to look at the status of complementarity in sentences such as (26) without *ko(n)*.

27) a. *Maaren*$_i$ **tann-ei/ aven-ei*$_{*i/j}$ *aditi-tt-aan*
Maran.NOM taan-ACC/ 3sgm-ACC beat-past-3sgm
'Maran beat himself'.

b. [*Maaren*$_i$ *tann-ei*$_{i/ *j/ *k}$ */aven-ei*$_{*i/j/k}$ *aditi-tt-aan*] *enru*
Maran.NOM taan-ACC/3sgm-ACC beat-past-3sgm comp
Somu$_j$ *co-nn-aan*
Somu say-past-3sgm
'Somu said that Maran beat himself'.

In (27a) we find that *aven* can occur where *taan* cannot. In (27b) while both *taan* and *aven* can pick out the matrix subject as their antecedent, only *aven* can pick out a discourse antecedent. However, this cannot be established as the complementarity that SBT predicts. This is because while complementarity as predicted by SBT requires pronouns to be free in those contexts where anaphors are bound, *taan* is not even allowed in (27a). Given this, the prediction that SBT makes is irrelevant for (27a). In (27b) the complementarity seems to have truly broken down. This is because both *taan* and *aven* are free in the governing domain. Perhaps the non-existence of complementarity in these sentences is not surprising given that Principle A is not able to account for *taan* in these sentences anyway. We would then have to claim that *taan* in these contexts falls out of the purview of SBT. A similar state exists in sentences where the subject has dative case.

28) a. *Maaren-iki*$_i$ *tann-ei*$_{i/ *j}$*/aven-ei*$_{*i/j}$ *pidik-um*
Maran-DAT taan-ACC/3sgm like-FUT
'Maran likes himself/him.'

b. [*Maaren-iki*$_i$ *tann-ei*$_{i/ j/ *k}$*/aven-ei*$_{*i/j/k}$ *pidik-um*] *enru*
Maran-DAT taan-ACC/3sgm-ACC` like-FUT comp
Somu$_j$ *co-nn-aan*
Somu.NOM say-past-3sgm
'Somu said that Maran likes himself/ him.'

In (28a) while *taan* can refer to the clause subject, the pronoun cannot. In (28b) *taan* can be ambiguously bound or free in its governing domain whereas the pronoun has to be free. These can be reconciled with SBT. In (28a) where Principle A seems to determine *taan*'s distribution, the pronoun has to be free. In (28b) where Principle A does not seem to be in force, there is no complementary distribution of *taan* and *aven*. However, while SBT does not seem to be violated here, we once again find instances where *taan* would fall out of the purview of SBT. We now move on to sentences where where *taan* is a subject.

29) a. [*taan* $_{i/ *j}$/ *aven* $_{i/ j}$ *paadeth-ei* *padi-tt-aan*] *enru*
taan.NOM/ 3sgm.NOM lesson-ACC study-past-3sgm comp
Somu$_i$ *co-nn-aan*
Somu.NOM say-past-3sg
Somu told Raman that he (Somu) studied the lesson.

b. [*tan-iki*$_{i/ *j}$/ *aven-iki*$_{i/ j}$ *pasi-kum*] *enru* *Somu*$_i$
taan-DAT/ 3sgm-DAT hunger-fut comp Somu.NOM
co-nn-aan
say-past-3sgm
Somu told Raman that he (Somu) will be hungry.

In (29a) and (29b) *taan* as well as *aven* have to be free. Recall that *taan* can be characterized as a reflexive in these sentences but there is no complementary distribution. However this could be because the governing domains of *taan* and *aven* can plausibly be different. This is because *taan* does not have a possible antecedent in the embedded clause and thus has to look at the matrix clause for its antecedent as dictated by Principle A. The pronoun, on the other hand, has to be free within the embedded clause and the facts suggest that this is the case.

In terms of complementarity, this is what we have found. In sentences with *ko(n)* as in (26) and no *ko(n)* as in (27), there is no complementary distribution of *taan* and *aven* as both have to be free. In sentences where the subject is in dative case as in (28), simple clauses exhibit complementarity while embedded clauses do not. In sentences

where *taan* is a subject as in (29), there is no complementarity. However, this one instance can be still resolved within SBT. The following table summarises the findings.

	Obeys Principle A?	**Shows Expected Complementarity?**
Taan as object and *ko(n)* on verb (26)	Yes	No
Taan as object and no *ko(n)* on verb (27)	No	No
Taan as object with a Dative subject (simple clauses) (28a)	Yes	Yes
Taan as object with a Dative subject (embedded clauses) (28b)	No	No
Taan as subject (29)	Yes	n.a

Table 3 SBT applied to *taan*

The table shows that *taan* does not obey Priniciple A in sentences with an object *taan* and no *ko(n)* and in embedded clauses with a dative subject. We suggested that the reason for this was that *ko(n)* could be a domain indicator and that sentences with dative subjects could be anomalies. However looking at the prediction of complementarity that SBT makes, we find that there will be additional problems with these claims. If *ko(n)* was a domain indicator of some sort and Principle A only kicks in when *ko(n)* is present, then we would expect these sentences to exhibit complementarity. However, we have found that it is exactly in these sentences that complementarity breaks down. It is not immediately clear how SBT would be able to salvage this as this does look like a fatal problem. Furthermore we mentioned that dative subjects could be anomalous constructions. However, this does not seem possible either as simple clauses with dative subjects behave exactly as Principle A dictates. These sentences also exhibit the complementarity predicted. It turns out that simple clauses with dative subjects are not

anomalous at all with respect to SBT. It is not apparent why these clauses should suddenly fall out of the purview of SBT when embedded.

If we were to assume that *taan* has the feature specifications [+anaphor][-pronominal], we would be able to account for when *taan* occurs as an object with a nominative subject and with *ko(n)* on the verb, when object *taan* occurs with a dative subject in a simple clause and when *taan* is a subject. *Taan* in these instances does obey Principle A. However when *taan* occurs as an object with a nominative subject without *ko(n)* on the verb as well as when the subject has dative case in an embedded clause, *taan* no longer follows Principle A. This either means that its feature specifications have changed or that the constructions where it does not follow Principle A are anomalies to be accounted for using other means. The first option of changing feature specifications is surely untenable, as postulating the changing of the features of *taan* would be too ad hoc. The second option of demarcating anomalous constructions does not work either. This is because the constructions where *taan* follows Principle A are not always the constructions where *taan* occurs in a complementary distribution with the pronoun *aven* as predicted by SBT. This follows from the fact that in sentences with *taan* as an object and nominative subject with *ko(n)* on the verb, *taan* follows Principle A and does not obey complementarity. The only constructions where both Principle A and complementarity is obeyed is when *taan* is a subject and in simple clauses with a dative subject. However, given the other sentences, this looks largely circumstantial rather than indicative of any real compliance to SBT.

What can we conclude from the above? It appears that the reference of *taan* cannot be attributed to the inherent properties of *taan* alone as described by SBT. If we

were to do this, then we would have to postulate ad hoc changes in the feature specifications of *taan*. This then means that SBT's typology of anaphoric expressions cannot capture *taan* and that we have to explore the alternative that was outlined in a previous sub-section. This alternative lies in the distribution of *taan* being due to the inherent properties of the predicate. The leading theory that has proposed this is R&R's Reflexivity framework and we turn to this next.

2.3 The Reflexivity Framework

Reinhart & Reuland (1993) view the distribution of pronouns and anaphors to be determined by the property of the predicate they occur with rather than as a function of the property of the pronoun or anaphor in question. The property within predicates which determines the distribution of most anaphoric items has been identified as reflexivity. This view is not novel and has been proposed in earlier works such as Keenan (1987) and Chierchia (1989) and according to R&R, as early as Jespersen (1933) and Gleason (1965). R&R derive a different set of anaphoric elements from that present in SBT using the twin features of REF (possessing a reflexivizing function) and R (having referential independence).

	[REF]	[R]
SELF anaphors	+	-
SE anaphors	-	-
Pronouns	-	+

Table 4 Typology of Anaphors according to R&R

The feature [REF] is only relevant insofar as it affects changes on the predicate of which the SELF anaphor is a part of. We will see more of this later. The [R] feature which refers to having referential independence is pertinent to being able to pick out some referent from the discourse. The [R] feature will be further illustrated when we show how R&R

invoke the Chain Condition to rule out certain constructions not ruled out by their conditions of reflexivity. According to Table 4, SELF anaphors have the reflexivizing function and typical examples of SELF anaphors include English *himself* and Dutch *zichzelf* (Everaert 1991). Examples of pronouns are English *him* and Chinese *ta*. Examples of SE anaphors which have neither the reflexivizing function nor referential independence are Dutch *zich*. We now turn to the binding conditions outlined by R&R which explains the distribution of these anaphoric elements.

R&R define a predicate as being reflexive *iff (at least) two of its arguments are coindexed* (R&R: 662). They further claim that the reflexivity of a predicate has to be licensed in one of two ways; either by being morphologically reflexive marked or by being marked in the lexicon. They spell out their Condition A and Condition B within their framework of reflexivity.

R&R's Reflexivity Framework

Condition A: A reflexive marked predicate is reflexive.

Condition B: A reflexive predicate is reflexive marked.

where a predicate is *reflexive* iff two of its arguments are coindexed.

and where a predicate is *reflexive-marked* iff either the predicate is lexically reflexive or one of the predicate's arguments is a SELF anaphor.

R&R claim that these two conditions are not only able to account for a wide range of distributional facts but also do not make the wrong predictions that SBT makes with respect to the complementary distribution of anaphors and pronouns. First we describe the basic facts covered by their conditions.

30) a. John_i likes $\text{himself}_{i/*j}$
b. John_i likes $\text{him}_{*i/j}$

In (30a) the predicate is reflexive marked because one of the arguments is a SELF anaphor. According to Condition A the predicate has to be reflexive, meaning that the coarguments have to be coindexed. In (30b) there is no SELF anaphor and the predicate is clearly not lexically reflexive, thus the predicate is not reflexive marked. Since the predicate is not reflexive marked, according to Condition B, the predicate cannot have coarguments which are coindexed and only a SE anaphor or pronoun can occur as the object. Since English is accepted to not have a SE anaphor, the pronoun is used to illustrate this. In languages which do have a SE anaphor, the SE anaphor can be coindexed with its coargument as long as the predicate is lexically reflexive. The following Norwegian examples from Hellan (1988) illustrate this.

31) a. *Jon$_i$* *wasket* *seg$_i$*
Jon washed SE
Lit: John washed himself.
b. *Jon$_i$* *skammer* *seg$_i$*
Jon shames SE
Lit: John is ashamed.

In (31a) and (31b), the predicates are reflexive and thus according to Condition B, both should have some kind of reflexive marking. Since a non-SELF anaphor is licensed as an object in these sentences, R&R claim the reason for this as being that these verbs are intrinsically reflexive, i.e. lexically reflexive. R&R claim that lexically reflexive verbs as those found in (31) come in two types. Verbs such as *schamen* 'shame' are intrinsically reflexive in that they do not allow any object which is distinct in reference from the subject. On the hand, verbs such as *wassen* 'wash' which do allow a distinct object are listed twice in the lexicon; Once as intrinsically reflexive and once as not intrinsically reflexive. When the former is generated in the sentence, a SE anaphor is used and when the latter is generated in the sentence, a SELF anaphor is necessary. So far, R&R are able

to account for all the data that SBT can. However, R&R develop their theory to show that they do not make the same wrong predictions with respect to complementarity that SBT makes. Consider (32) from R&R.

32) a. Lucie saw a picture of herself/ her.
b. Max likes jokes about himself/ him.

In (32), the reflexive as well as the pronoun can occur in the preposition phrase while picking out the subject as their antecedents. Unlike SBT, which predicts this not to be the case (without stipulative manipulation of the governing domain), R&R argue that their theory does not make any such claim. As their conditions on reflexivity only affect coarguments, the occurrence of both the pronoun and reflexive in the same position in (32) is not surprising given that the position in question is not a coargument of the subject. In (32a) *herself*/ *her* occurs in a position which is not an argument of *saw*. The reflexive and pronoun in this sentence are instead arguments of the preposition *of*. However the subject *Lucie* is an argument of *saw*. Since *Lucie* and *herself/ her* are arguments of different predicates, R&R's theory does not say anything about how they can occur together in a sentence. The same applies to (32b). While *Max* is an argument of *like*, *himself/ him* are arguments of the the preposition *about*. They use the same line of reasoning to explain the occurrence of reflexives as so-called logophors[7]. The following data used by R&R is taken from Zribi-Hertz (1989).

33) a. 'It angered him that she … tried to attract a man like himself.'
b. *'It angered him that she tried to attract himself.'

[7] R&R use the term logophoric to refer to reflexives which are used in non-reflexive contexts. This means that the use of the reflexive in (39) would be considered logophors. There is a more technical use of logophor which refers to point of view (Hagege 1974). We deal with this definition of the term in the next chapter.

In (33a) the reflexive is allowed while in (33b) it is not. The reason for this lies in coargumenthood again. In (33a) *himself* and *she* are not coarguments as the elided text in the example introduces other predicates into the sentence. However, in (33b) where there is no other text separating *she* and *himself,* they are coarguments. Since the embedded predicate has been reflexive marked (by the SELF anaphor) in (33b), the predicate should be reflexive. However, *she* and *himself* cannot be coindexed due to feature conflicts, thus (33b) is not legitimate. On the other hand, since *she* and *himself* are not coarguments in (33a), Condition A is not violated in this sentence. While R&R can provide an explanation for a lot of data, note that what has been discussed so far does not distinguish between the grammatical (34a) and ungrammatical (34b) below.

34) a. $John_i$ likes $himself_i$.
b. *$Himself_i$ likes $John_i$.

The difference between (34a) and (34b) is one of word order. R&R's conditions by themselves do not rule out (34b) as all that is required to reflexivize a predicate is that one of its arguments be a SELF anaphor and coindexation will occur. These requirements are met in both sentences. To rule out (34b), R&R invoke their version of the Chain Condition which is spelled out below.

General Condition on A-chains (R&R: 696)

A maximal A-chain (α_1, …, α_n) contains exactly one link-α_1-that is both +R and Case-marked.

> where an A-chain is defined as one where there is a sequence of coindexation that is headed by an A-position and satisfies antecedent government.

We do not delve into the technical aspects of what constitutes an A-chain but merely note its implications for the data. In (34a) *John* and *himself* form an A-chain and *John* which

heads the chain is the only element in this chain which is [+R] as it is a full NP and case marked whereas *himself* is referentially deficient (see Table 4) even though it does have accusative case. This makes the sentence legitimate. In (34b), however, the chain is headed by *himself*. In this sentence, while there is still only one element *John* which is [+R] and case marked in this chain, it does not head the chain. This means that this sentence violates the condition on A-chains and is thus ruled out.

So far we have seen how the theory outlined in R&R explains the distribution of anaphors and pronouns. They essentially propose a move towards a predicate-centric account for reflexivity and using their conditions for reflexivity they are able to account for a lot of data and have the added advantage of not making wrong predictions when it comes to the complementary distribution of anaphors and pronouns unlike SBT. Furthermore, they are able to account for the distribution and behaviour of SE anaphors - something which SBT cannot do. By invoking their condition on A-chains, they are also able to rule out sentences headed by anaphors. However, the downside to their account seems to be that their account for logophors seems too broad, thus generalizing all instances where reflexives occur in non-coargument positions into one category. In addition, recall the SBT approach mentioned above to solving the non-complementarity problem. SBT would have to stipulate different governing domains for the reflexive and pronoun to account for their occurrence in the same position. While R&R do seem to have a more attractive approach to the problem, authors such as Safir (2004) have claimed that R&R's reflexivity framework is too powerful and has too little predictive power. Noting these objections to R&R, we now move on to how R&R would account for Tamil *taan*.

2.3.1 The Reflexivity Framework & *taan*

We will first lay out the range of data to be accounted for by R&R and provide our evaluation of the theory's scope over *taan*. First, we describe the cases when *taan* occurs as an object with a nominative subject.

35) a. *Maaren*$_i$ *tann-ei*$_{i/*j}$ *aditi-**ko**-nd-aan*
Maran.NOM taan-ACC beat-**kon**-past-3sgm
'Maran beat himself'.

b. **Maaren*$_i$ *tann-ei*$_i$ *aditi-tt-aan*
Maran.NOM taan-ACC beat-past-3sgm
For: 'Maran beat himself'.

At first glance at the data, the successful application of R&R's theory to the data seems promising. Recall that R&R's theory mentions that for coarguments to be coindexed, the predicate must be reflexive marked. Comparing (35a) and (35b), we can see that such coindexing of coarguments is possible in (35a) but not in (35b). The minimal difference between these two sentences is the occurrence of *ko(n)* on the verb. Although R&R do not explicitly consider the possibility of lexically reflexivizing the verb with overt morphology, we, for now, follow Lidz (1995) who claims that *ko(n)* does indeed do so in Kannada. If we look at *ko(n)* as a lexical reflexivizer in Tamil as well, the data in (35) makes perfect sense. In (35a) the predicate is lexically reflexive due to *ko(n)* on the verb and the coarguments *Maaren* and *taan* can be coindexed according to Condition A. When there is no such marking on the verb as in (35b), such coindexing is not allowed according to Condition B since the verb is not lexically reflexive. This would explain why a pronoun is possible in the object position of sentences such as (35) seen below in (36).

36) a. *Maaren*$_i$ *aven-ei*$_{i/*j}$ *aditi-**ko**-nd-aan*
Maran.NOM 3sgm-ACC beat-**kon**-past-3sgm
'Maran beat himself'.

b. *$Maaren_i$ *aven-ei$_i$* *aditi-tt-aan*
Maran.NOM 3sgm-ACC beat-past-3sgm
For: 'Maran beat himself'.

In (36) a pronoun is in the object position of the sentence. However, with *ko(n)* on the verb in (36a), the coarguments have to be coindexed and we find that despite the fact that the object position is filled with a pronoun, it has to be coindexed with *Maaren*. In (36b) where there is no *ko(n)* reflexivizing the verb, *aven* cannot refer to *Maran* anymore. Considering *ko(n)* a reflexive marker explains the coreference possibilites of *taan* in embedded clauses as well. Consider (37).

37) a. [*$Maaren_i$* *tann-ei$_{i/ *j/ *k}$/ aven-ei$_{i/ *j/ *k}$* *aditi-**ko**-nd-aan*] *enru*
Maran.NOM taan-ACC/ 3sgm-ACC beat-**kon**-past-3sgm comp
$Somu_j$ *co-nn-aan*
Somu say-past-3sgm
'Somu said that Maran beat himself'.

b. [*$Maaren_i$* *tann-ei $_{*i/ j/ *k}$/ aven-ei$_{*i/ j/ k}$* *aditi-tt-aan*] *enru*
Maran.NOM taan-ACC/ 3sgm-ACC beat-past-3sgm comp
$Somu_j$ *co-nn-aan*
Somu say-past-3sgm
'Somu said that Maran beat himself'.

The facts above fall out if we assume that *ko(n)* is the reflexive marker which makes the embedded verb reflexive. In (37a) the embedded predicate is reflexive as *ko(n)* reflexive marks the verb and *taan* or the pronoun *aven* can only pick out its coargument, the embedded subject as its antecedent as dictated by Condition A. When there is no *ko(n)* on the embedded verb as in (37b), the predicate is not reflexive as it is no longer reflexive marked and thus *taan* can only refer to an element which is not its coargument- in this case *Somu*. The pronoun also loses the ability to refer to the coargument which is the embedded subject. If *taan* and *aven* do refer to their coargument, *Maaren* in this sentence, note that this will be a Condition B violation. In this sentence, while R&R's theory cannot actually tell us which antecedents *taan* and *aven* will pick out, it suffices to note

that their theory is not violated. However, recall that the distribution of *taan* and *ko(n)* coincides only when *taan* is an object in a clause with a nominative subject. As *ko(n)* does not occur with the other instances of *taan*, we move to them now starting with *taan* as a subject. We find that instances where *taan* occurs as a subject can also be easily incorporated within R&R's framework. Consider the following sentences.

38) a. [*taan* $_{i/*j/*k/*m}$ *Maaran-ei*$_j$ *adi-tt-aan*] *enru*
taan.NOM Maran-ACC beat-past-3sgm comp
Somu$_i$ *co-nn-aan*
Somu.NOM say-past-3sg
Somu said that he (Somu) beat Maran.

b. [*tan-iki*$_{i/ *j/*m}$ *pasi-kum*] *enru* *Somu*$_i$ *co-nn-aan*
taan-DAT hunger-fut comp Somu.NOM say-past-3sgm
Somu said that he (Somu) will be hungry.

In (38a) *taan* is a nominative subject and in (38b) it is a dative subject. In both sentences, *taan* is coindexed with a non-coargument, thus falling out of the purview of R&R's Reflexivity framework. This means that these uses of *taan* would simply be considered logophoric. While much of the Tamil data can be explained by R&R (or at least sent to the domain of logophors), one set of data is much harder to reconcile with R&R's theory. These sentences are the ones where object *taan* has a dative subject. We produce the relevant data below.

39) a. *Maaren-iki*$_i$ *tann-ei*$_{i/ *j}$ *pidik-um*
Maran-DAT taan-ACC like-fut
'Maran likes himself.'

b. [*Maaren-iki*$_i$ *tann-ei*$_{i/ j/ *k}$ *pidik-um*] *enru* *Somu*$_j$
Maran-DAT taan-ACC like-fut comp Somu.NOM
co-nn-aan
say-past-3sgm
'Somu said that Maran likes himself/ him.'

In (39a) *taan* can pick out the subject as its antecedent even though there is no *ko(n)* on the verb. In fact recall from (10b) that *ko(n)* is not allowed on the verb in these cases. In

(39b) *taan* can pick out the embedded as well as the matrix subject. If R&R are to remain consistent with the data in (35) and (36), they would have to conclude that *taan* is not a SELF anaphor as it is unable to reflexivize the predicate in (35b) and (36b). Furthermore it cannot be a pronoun as there is evidence from the condition on A-chains against this. This is evident in the fact that *taan* cannot occur as the sole argument of a clause or head an A-chain which indicates that it is [-R]. See (40) below.

40) *aven$_i$/ *taan* *van-th-aan.*
3sgm/taan come-past-3sgm
He/ *Self came.

In (40), *taan* cannot occur as the sole argument of a clause, unlike the pronoun. According to the condition on A-chains, if *taan* was [+R], we would expect (40) with *taan* to be grammatical. However, this is not the case and we have to conclude that *taan* is [-R]. Coupled together with (35) and (36) which shows that *taan* is [-SELF], looking at table 4, we have to conclude that *taan* is a SE anaphor in R&R's framework. Going back to (39), R&R would have to explain how *taan* being a SE anaphor can be coindexed with its coargument despite the lack of *ko(n)* which we have assumed is a reflexive marker. R&R would have to say one of two things to avoid a Condition B violation in these sentences. They have to either claim that the verb *pidi-* 'like' and other similar verbs are lexically reflexive or that *Maaren* and *taan* are not coarguments in (39). If the verb is lexically reflexive in (39), then *ko(n)* is not necessary to reflexive the predicate as it is already reflexive and the coarguments *Maaren* and *taan* can be coindexed. On the other hand, if we can show that *Maaren* and *taan* are not coarguments at all, then the fact that they are coindexed will fall out of the scope of R&R's theory and will not be a violation

of Condition B. We will show that neither solution is sustainable without a lot of added stipulations. We first deal with the potentially lexically reflexive nature of the verb first.

Psych Verbs as Inherently Lexically Reflexive

Assuming that the verb in (39) is lexically reflexive means that we would have to conclude that every verb which does not allow *ko(n)* to be suffixed to it but yet allow *taan* as an object which is coindexed with the subject (presumably every psych verb) would have to be lexically reflexive. We venture that this does not empirically follow. R&R claim that intrinsically reflexive words like *schamen-* 'shame' in Norwegian shown in (31b) above do not allow a distinct object. However we know that verbs in Tamil which do not allow *ko(n)* are verbs like *pidi* 'like', *theri* 'know' and *veru* 'hate', all of which do allow a distinct object. (39) is representative of all these verbs. However R&R do identify another group of intrinsically reflexive predicates. Recall from earlier that we mentioned that R&R claim that intrinsically reflexive predicates like Norwegian *wassen* 'wash' in (31a) do allow a distinct object as they are listed twice in the lexicon; once as a intrinsically reflexive verb and once as a non-reflexive verb. When the former is selected from the lexicon, a SE anaphor is licit and when the latter is selected, a SELF anaphor is required. In the same vein, R&R would simply claim that the Tamil verbs such as *pidi* 'like', *theri* 'know' and *veru* 'hate' are all listed twice in the lexicon and thus lexically reflexive in one instantiation. This means that *taan* being a SE anaphor can be licensed as coindexed arguments to these verbs when licensed as lexically reflexive. Even if this (seemingly ad hoc) solution is accepted, note what happens when a pronoun is used as an

object in sentences like (41). *Pidi* 'like', being lexically reflexive, should allow a pronoun to occur as an object and still be reflexive. This does not fall out.

41) a. *Maaren-iki*$_i$ *aven-ei* $*_{i/j}$ *pidik-um*
Maran-DAT 3sgm-ACC like-fut
'Maran likes him.'

b. [*Maaren-iki*$_i$ *aven-ei*$_{*i/j/k}$ *pidik-um*] *enru* *Somu*$_j$
Maran-DAT 3sgm-ACC like-fut comp Somu.NOM
co-nn-aan
say-past-3sgm
'Somu said that Maran likes him.'

In (41a) the pronoun *aven* can only pick out a discourse antecedent and not the subject. In (41b) the pronoun cannot pick out the embedded subject either. If the verb *pidi* 'like' was indeed lexically reflexive, we would expect the pronoun to be able to pick out the subject in both sentences. Perhaps the reason why *aven* cannot be coindexed with the subject is due to the fact that this will actually violate the Chain Condition as this chain would contain both *Maran* and *aven,* both of which are [+R] and case marked. However, note that the occurrence of morphological reflexive marking with *ko(n)* allows both *taan* and *aven* to occur in the same contexts and pick out the same antecedents as shown in (35) and (36). Clearly in sentences such as (35) and (36) where the pronoun is allowed, the Chain Condition can be violated. For the sake of argument, if we were to assume that the Chain Condition can be violated here as well, since lexical reflexivization is just another way of marking a predicate as reflexive, we would expect a lexically reflexive predicate to 'force' a pronoun to pick out the same antecedent as *taan*. This is clearly not the case as seen in (41). Lidz (2001) actually shows that verbs that are semantically reflexive do not form a class with verbs that are reflexive marked, contrary to what would be expected to fall out from R&R's theory. This corroborates our argument here. One way to cope with the data would be for R&R to claim that the Tamil psych verbs are indeed listed

twice; once as reflexive and once as non-reflexive. However they would have to add that *taan* is licensed only when the reflexive one is used and when the non-reflexive one is used, only the pronoun is allowed. This would explain why *taan* and not the pronoun can be coindexed with the subject of the verb. However at this point, it seems that we are merely creating more stipulations to satisfy the data and in the process losing all predictive power. The reader will recall Safir (2004)'s objection to R&R as being too powerful that was alluded to earlier. Here we see a concrete example as applied to Tamil.

Taan as a Non-CoArgument

The other solution to reconcile (39) with R&R is to stipulate that the subject and object in these sentences are only apparently so and that in fact they are not coarguments at all. If this can be established, then *taan* in these sentences would be considered as falling outside of the scope of the Reflexivity Framework. There is some evidence for suggesting that structurally the dative subject and object *taan* are not coarguments. For one, unlike nominative subjects, dative subjects never trigger agreement on the verb as can be seen in all the earlier relevant examples. Perhaps this is because the subject is not in a spec-head agreement position with the verb. This does seem a promising line of enquiry. However, there are nominative subjects which occur with psych verbs which behave exactly like dative subjects. Consider the following.

42) a. *Maaren*$_i$ *tann-ei*$_{i/*j}$ *veru-tt-aan*
Maran taan-ACC hate-past-3sgm
Maran hates himself/ him.

b. **Maaren*$_i$ *tann-ei*$_{i/*j}$ *veru-**ko**-nd-aan*
Maran taan-ACC hate-**kon**-past-3sgm
Maran hates himself/ him.

In (42) the verb *veru* 'hate' licenses the occurrence of *taan* as the object which is able to pick out the subject as its antecedent. (42b) shows that *ko(n)* is in fact not allowed. Even if we could establish that in (39), the surface subject and object are not coarguments, we would not be able to explain (42). Here there seems to be the necessary spec-head agreement that was lacking before and yet this sentence behaves the same way as (39) in not allowing *ko(n)* on the verb and allowing *taan* to pick the subject as its antecedent. Ideally, we should be able to explain (42) in the same way as (39). However if we were to stipulate that there are no coarguments in (39), we do not see how that could be maintained for (42).

Furthermore an embedded clause can also occur without agreement on the verb which means that a lack of agreement cannot be pointed to as one indication that there are no coarguments in a sentence. Consider the following.

43) a. [*Maaren* *thambi-yei* *adi-tt-aan*] *enru*
Maran.NOM brother-ACC beat-past-3sgm comp
Somu *co-nn-aan*
Somu.NOM say-past-3sgm
'Somu said that Maran beat Brother.'

b. [*Maaren* *thambi-yei* *adita-taage*] *Somu* *co-nn-aan*
Maran.NOM brother-ACC beat-that Somu.NOM say-past-3sgm
'Somu said that Maran beat Brother.'

In (43a) the embedded clause has the verb *adi-* 'beat' and full agreement and tense marking. In this sentence, one could very safely assume that *Maaren* and *thambi* are indeed coarguments of the embedded verb *adi-* 'beat'. However the sentence can also be realized as shown in (43b). In this sentence, the verb is the same but there is no tense or agreement marking on the verb. The complementizer *enru* has cliticized onto the embedded verb instead. Although there is no agreement on the embedded verb in (43b), it

would still be prudent to assume that *Maaren* and *thambi* are still coarguments in this sentence given (43a).

With this we see that there is one set of Tamil data which cannot be easily reconciled with R&R as both alternatives for reconciling it have been unfruitful. However even those sentences which can be reconciled seem to have their own problems. We turn to these next.

Ko(n) as a Reflexive Marker

Earlier we said that if *ko(n)* was considered a reflexive marker, then we would be able to explain the data in (35) and (36). However, *ko(n)* cannot be simply considered a reflexive marker because of all its other uses in non-reflexive contexts. Consider the following.

44) a.	*Maaren*	*kathav-ei*	*moodi-**ko**-nd-aan*
	Maran.NOM	door-ACC	close-**kon**-past-3sgm
	Maran closed the door.		
b.	*Maaren*	*naak-ei*	*niiti-**ko**-nd-aan*
	Maran.NOM	tongue-ACC	stick.out-**kon**-past-3sgm
	Maran stuck out his tongue.		

In (44a) and (44b) even though *ko(n)* occurs on the verb, the subject and object are clearly not the same entity. Following Lidz (2001), who shows the same facts in Kannada, we claim that the licensing of *ko(n)* is not determined by semantic reflexivity as predicted by R&R. However, it could be the case that one of the functions of *ko(n)* is to act as a reflexivizer in certain situations. Even if this was the case, it would mean that R&R's theory is not able to explain when it does occur as a reflexivizer and when it does not. Furthermore their theory is also not able to explain why *ko(n)* is incompatible with psych verbs as we have already seen.

Given these problems with R&R's theory when trying to account for *taan,* it seems prudent to conclude that R&R's Reflexivity Framework does not capture the Tamil facts adequately. There are two main problems. The first is reconciling the sentences where object *taan* occurrs with a dative subject. We found that regardless of whether we stipulated that such predicates were always inherently reflexive or whether we stipulated that the surface subject and object were not really coarguments, we would run into trouble. The second problem is that *ko(n)* which would be considered a reflexive marker under R&R's framework seems to have a much broader function in Tamil. First, it can occur in sentences where reflexivization does not occur. Second, even if reflexivization was required, the use of *ko(n)* is not always legitimate.

2.4 Chapter Summary

In this chapter of the thesis, we have seen a wide range of Tamil data, and two influential theories which could account for them. The data consisted of *taan* occurring as the object with a nominative subject and with *ko(n)* on the verb, *taan* occurring as the object with a nominative subject and without *ko(n)* on the verb, *taan* as an object with a dative subject, and *taan* as subject with both nominative and dative case.

We then moved on to the first of the two major syntactic theories which have been used to explain the distribution of pronouns and anaphors cross-linguistically. This theory is the Standard Binding Theory (SBT). We described the basic facts about SBT and the principles contained therein. We then applied SBT to Tamil *taan* to see how well the theory could explain the Tamil facts. We found that *taan* behaves in accordance with Principle A when it occurs as an object with a nominative subject and *ko(n)* on the verb, in simple clauses with a dative subject as well as when *taan* occurs as a subject. However

in sentences where *taan* is an object and the subject has dative case, *taan* always violates Principle A in embedded clauses and follows it in simple sentences. Lastly *taan* violates Principle A when it occurs as an object with a nominative subject without *ko(n)* on the verb. Disregarding the instances where *taan* does not follow Principle A as anomalies in order to tidy the picture also is unsuccessful as it was found that the anomalous sentences actually satisfies the complementarity of anaphors and pronouns predicted by SBT. Furthermore not all the sentences where *taan* satisfies Principle A follows this prediction of complementarity. While *taan* as a subject does follow complementarity, sentences where *taan* occurs with a *ko(n)* on the verb does not. With that, we rejected SBT as a potential way of fully accounting for *taan* and we moved on the next syntactic approach to anaphora which was predicate-centric.

R&R's Reflexivity Framework aims to explain reflexivity and we showed the basic mechanism behind this framework. We then applied this theory to *taan* and found that this theory cannot explain the Tamil facts satisfactorily either. There were two main problems with this theory. The first was that the set of data where object *taan* and a nominative subject occurs could not be reconciled within R&R's theory without too many added stipulations. At this point, there is no more predictive power left for the theory. The second problem is that the facts show that *ko(n)* which would be considered a reflexive marker under the Reflexivity Framework is neither a necessary nor sufficient condition for reflexivization in Tamil.

Does this mean that a syntactic description of *taan* is entirely on the wrong track? We claim that the syntactic characteristics of *taan* do have to be taken into consideration. However, as we have seen in this chapter, any attempt to classify *taan* purely as a

syntactic element does seem impossible. With this in mind, we now move on to the next chapter. We will be focusing on the semantic and discourse properties of *taan* which will illuminate different properties of *taan* which have not been seen yet. Uncovering these properties will bring us closer to providing a more satisfactory characterization of *taan*.

Chapter 3 *Taan* as a Semantic/ Discourse Element

In the previous chapter we examined two influential syntactic approaches to anaphora and concluded that neither of them could satisfactorily account for all the Tamil data. Thus, in this chapter, we will be focusing on the more prominent semantic and discourse approaches to anaphora. Specifically, we will be looking at VP-ellipsis and what this tells us about *taan*. We conclude that this approach, while providing a different perspective to *taan*, is also unable to provide a characterization of *taan* that will enable us to classify it. We finally move on to *taan* as a *logophoric pronoun*. Here we do not mean *logophor* in R&R's sense, but in the sense first put forth by Hagege (1974) that refers to the reporting of a particular point-of-view. We show that analysing *taan* as a logophoric pronoun actually enables us to account for all the Tamil data. We pursue this line of inquiry by looking at Sells (1987) and we show that Sells' adapted account provides a parsimonious account for *taan* as well as *ko(n)*. We conclude that *taan* in Tamil is indeed a logophoric pronoun. This is a novel claim about *taan* and what we present here is enough to justify an investigation into other Dravidian anaphora along similar lines. We will not undertake such a comprehensive investigation here but leave it to future research. We now start with VP ellipsis.

3.1 VP ellipsis

VP ellipsis, whose observation is first credited to Sag (1977), refers to a multi clausal sequence where the VP of the following clause has been elided. Consider the following example.

45) a. John [likes Mary] but Tom doesn't [*e*].

b. John [likes Mary] but Tom doesn't [like Mary].

In (45a), the VP of the following clause is elided (here represented by *e*). Nonetheless, the second clause is interpreted as if there is a copy of the first VP. This is captured by the representation in (45b). This is a rather general example of VP ellipsis which will not concern us here. The type of VP ellipsis we are concerned with here occurs when there is an anaphoric element or a pronoun in the VP itself. Consider (46).

46) a. John [likes his mother] but Tom doesn't [*e*].
b. John [likes his mother] but Tom doesn't [like his mother].

In (46a) there is a pronoun *his* within the VP itself. The reconstruction in (46b) shows that the elided VP is the same as the lead clause VP. In (46a) there are three different interpretations that the elided pronoun can have. The elided pronoun can refer to Tom, John or some discourse antecedent like Alan. However, these different interpretations are not freely available but dependent on the interpretation of the pronoun in the lead VP. The following shows the various interpretations of the pronoun in the lead clause and the subsequently available readings of the pronoun in the following clause.

Pronoun in lead clause	**Pronoun in following clause**
Alan	Discourse Reading (Alan/ *Others)
	*Lead clause reading
	*Following clause reading
John	*Discourse Reading
	Lead clause reading (John)
	Following clause reading (Tom)

Table 5 Interpretation of pronouns in VP ellipsis

The pronoun in the lead clause can only be one of two referents, a discourse entity like *Alan* or the subject of the lead clause, *John*. When the lead pronoun has the discourse entity reading, the elided pronoun must also pick out the same discourse entity. However,

when the lead pronoun has the interpretation of *John*, the elided pronoun can have the interpretation of *John* or *Tom* but not any discourse referent. When the anaphor is replaced with *himself* in the lead clause, the availability of interpretations is greatly constrained.

47) a. John [likes himself] but Tom does not [*e*].
b. John [likes himself] but Tom does not [likes himself].

In (47a), the elided reflexive can only refer to Tom. When we look at the reflexive in the lead VP, *himself* itself can only refer to the lead clause subject. While we have only looked at a very small set of data with regards to English VP ellipsis here, these will be enough to motivate the discussion that follows. At the heart of the VP ellipsis phenomenon, there are two main questions relevant to anaphora. How do we account for the differences between the available interpretations for pronouns and reflexives? Furthermore, what leads to the dependency of interpretation that the anaphoric element in the following clause has on the interpretation of the element in the lead clause?

There have been semantic as well as syntactic approaches which have been used to account for data like (46) and (47) as pointed out by Dalrymple, Shieber and Pereira (1991) among others. We only concentrate on the semantic perspective of VP ellipsis. While this will illuminate more properties of what an anaphoric element is, we will show that even this will not allow us to nail down the classification of *taan*.

3.1.1 Anaphors as Variables

One semantic approach to VP ellipsis lies in the treatment of anaphors and pronouns as variables whose interpretation is determined by whether they are free or bound variables. Whether they are free or bound variables is indicated by the availability

of strict and sloppy readings (Heim & Kratzer 1998). The strict reading occurs when the lead and following pronoun both refer to the same entity. The sloppy reading occurs when the lead pronoun and following pronoun refer to their own clause subjects. Looking back at the data in (46) and (47), the pronoun in (46) can have both strict and sloppy readings. The strict reading yields when the lead and following pronoun refer to the same element- some discourse referent or *John*. The sloppy reading yields when the lead and following pronouns refer to different entities- such as *John* in the lead clause and *Tom* in the following clause. Note that there cannot be a sloppy reading when the lead pronoun refers to a discourse referent. Furthermore in (47) the lead and following reflexive cannot have the same interpretation and thus can only have the sloppy reading. So what is the reason behind the fact that pronouns (as in (46)) can have strict and sloppy readings while reflexives (as in (47)) can only have a sloppy reading?

The answer that we adopt here is taken from Heim & Kratzer (1998). Following them, if we think of pronouns and reflexives as variables which can be free or bound, the data in (46) and (47) can be derived without extra stipulations. Essentially this means that if a free variable with a certain interpretation occurs in the lead clause, the elided element will also be a free variable with the same interpretation. However if the lead pronoun is actually a bound variable, the elided pronoun will also be a bound variable. We will now go on to look at this mechanism in some detail.

One of the core assumptions under this approach is that every variable gets its value either from coreference or binding. Coreference occurs when a variable gets its interpretation through a function called an assignment function. This basically assigns

values to variables from some domain of individuals. This determines the interpretation of the pronoun in (48).

48) John likes his mother.

In (48) the pronoun can either mean that John likes *Alan's* mother or John likes *John's* mother[8]. In this sentence, the assignment function provides the required reading. Assume that there exists a domain of individuals from which the assignment function assigns values to variables in a sentence.

49) Domain of Individuals and their corresponding indices:
John → 1
Tom → 2
Alan → 3
Susan → 4

Given such a set of individuals, the assignment function can assign the value 1, 2 or 3 to the pronoun *his* in (49) depending on context. '4' is not a possible value because of the feature mismatch between *Susan*, an inherently female name and *his* which is a male pronoun. Thus (48) can have the following readings as shown in (50). Additional context will then filter out the incorrect readings.

50) John likes $his_{1,2,3}$ mother.

However, the assignment function is not the sole way in which a variable can get a value. Consider the following sentence with a reflexive.

51) $John_i$ likes $himself_{i/*j}$.

In (51) the reflexive *himself* can only mean *John*. In this sentence, the assignment function is not responsible for attributing a value to the variable *himself* but rather the sentence structure. This is, in fact, binding, which is the second way in which variables

[8] The latter reading can be derived from *John* binding *him* as well. We will move on to this later.

can get their meaning. Reflexives are not the only elements that get their interpretation through binding. The pronoun in (50) can also be bound by *John*. This means that the pronoun can receive the interpretation of *John* either through the assignment function or through binding. We have seen that pronouns can get their meaning in two ways while reflexives can only be bound. Now we shall move on to see how these enable us to account for the VP ellipsis facts in (46) and (47).

Recall that the pronoun in (46) can have both strict and sloppy readings. The fact that *his* can corefer and be bound does seem on the surface to be the reason behind the strict and sloppy readings of the pronoun in VP ellipsis. Let's look at the strict readings first. Recall that this means that the lead and elided pronoun have to refer to the same entity. When a pronoun is assigned a value, say *Alan* or *John,* through the assignment function, the elided pronoun has to be reconstructed as the same variable at LF. Thus when the pronoun is assigned a value of '1', the elided pronoun has to be reconstructed as the same variable with the same value. When the lead pronoun is given a value of '2', the elided pronoun has to be reconstructed as a variable with the same value. Since these variables are given a value through the assignment function, we can also refer to these variables as free variables. This accounts for the strict readings. We have seen that the strict reading can be reduced to the generalization that a pronoun can be a free variable. Turning to the sloppy reading, recall that the other way in which a pronoun can be given a value is through binding. The variables in these structures are naturally called bound variables. When the lead pronoun is a bound variable and is elided through VP ellipsis, the bound variable is reconstructed at LF. However, since the interpretation of the bound variable is dependent on the binder, the reconstructed pronoun will now have a new

binder, the following clause subject. This leads to the sloppy reading. In this way, Heim & Kratzer's mechanism is able to explain that the reason why the elided pronoun in (46) can have a strict and sloppy reading is because the pronoun is ambiguous between a free and bound variable. This line of reasoning accounts for the reflexive in (47) easily. We know from (51) that the reflexive can only be a bound variable. This means that when the elided reflexive is reconstructed in (47), the interpretation of the reconstructed reflexive will depend on the new binder, the following clause subject. This is what leads to the sloppy reading.

We have seen that the free and bound variable status of pronouns and reflexives is able to explain the strict and sloppy readings in VP ellipsis. Not only can we account for the different interpretations of pronouns and reflexives in these constructions, we can also account for why there is a dependency of the following clause interpretation on the lead clause. We now move on to apply Heim & Kratzer's approach to Tamil *taan*. We will conclude that although it does tell us more about *taan*, it cannot provide an analysis which will enable us to classify *taan* with respect to its anaphoric status.

3.1.2 *Taan* as a Variable

In this sub-section, we shall see whether *taan* behaves like a free or bound variable with respect to Heim & Kratzer (1998)'s approach to VP ellipsis. We will do this by determining whether strict or sloppy readings are available in the various constructions where *taan* can occur. What we will find is that except for the one construction where *ko(n)* occurs with *taan* as an object, *taan* yields both the strict and sloppy readings in the rest of the constructions where *taan* occurs. We shall first look at the construction where *ko(n)*can occur.

52) a. [*Maaren*$_i$ *tann-ei*$_{i/ *j}$ *aditi-**ko**-nd-aan*] *enru* *Somu*
Maran.NOM taan-ACC beat-**kon**-past-3sgm comp Somu
co-nn-aan. *Raman-num* *kooda*.
Sy-past-3sgm Raman-COOR[9] too
'Somu said that Maran beat himself. Raman did too.'
Interpretation: Raman is sad that Maran beat himself. (Sloppy)

b. [*Maaren*$_i$ *tann-ei*$_{*i / j/ *k}$ *adi-tt-aan*] *enru* *Somu*$_j$
Maran.NOM taan-ACC beat-past-3sgm comp Somu
co-nn-aan. *Raman- um* *kooda*
say-past-3sgm Raman-COOR too
'Somu said that Maran beat him. Raman did too.'
Interpretation: Raman said that Maran beat Raman. (Sloppy)
Raman said that Maran beat Somu. (Strict)

In (52a) the verb together with the object *taan* has been elided in the following clause and only the sloppy interpretation of the elided *taan* is available. This suggests that here *taan* is behaving like a bound variable. However, as alluded to above, this is the only construction where *taan* occurs with the sloppy reading. In all other constructions where *taan* occurs, both strict and sloppy readings are available. In (52b) where the embedded clause does not have *ko(n)*, *taan* picks out the matrix subject as its antecedent. The elided *taan* in the following clause here has both the sloppy and strict interpretations where *taan* can pick out the lead sentence matrix subject as well as the following clause subject as its antecedent. The availability of strict and sloppy readings with *taan* here and elsewhere suggests that in these sentences *taan* can be both a free as well as a bound variable. We shall next move on to the rest of the sentences where *taan* can occur, starting with *taan* as an object with a dative subject.

53) *Maaren-iki*$_i$ *tann-ei*$_i$ *pidik-um.* *Raman-ik-um* *kooda*
Maran-DAT taan-ACC like-fut Raman-DAT-COOR too
'Maran likes himself. Raman did too.'
Interpretation: Raman likes Raman. (Sloppy)
Raman likes Maran. (Strict)

[9] Note that COOR in these sentences refers to a coordinating suffix

In (53) where the verb and object *taan* are elided from the following clause, both strict and sloppy readings are available. The strict and sloppy readings are available even when *taan* is a subject as can be seen in (54)

54) a. [*taan* $_{i/*j/*k}$ *paadeth-ei* *padi-tt-aan*] *enru*
taan.NOM lesson-ACC study-past-3sgm comp
Somu$_i$ *co-nn-aan.* *Maran-um* *kooda*
Somu.NOM say-past-3sg Maran-COOR too
'Somu said that he (Somu) studied the lesson. Maran did too.'
Interpretation: Maran said that Maran studied. (Sloppy)
Maran said that Somu studied. (Strict)

b. [*tan-iki*$_{i/ *j/*m}$ *pasi-kum*] *enru* *Somu*$_i$
taan-DAT hunger-fut comp Somu.NOM
co-nn-aan. *Maran-um* *kooda*
say-past-3sgm Maran-COOR too
'Somu said that he (Somu) will be hungry. Maran did too.'
Interpretation: Maran said that Maran will be hungry. (Sloppy)
Maran said that Somu will be hungry. (Strict)

In (54a) *taan* is an embedded subject with nominative case. In (54b) *taan* is an embedded subject with dative case. Strict and sloppy readings are available in both sentences regardless of the case that *taan* is in.

We have seen the entire set of Tamil data with respect to VP ellipsis above. Ignoring the one instance when *ko(n)* occurs on the verb for now, we can summarise that *taan* always gives rise to strict and sloppy interpretations. Note that this behaviour is very much like the English pronoun in (53) and not the reflexive in (54). Does this mean that *taan* is a pronoun and not an anaphor? After all, authors like Amritavalli (1984) have indeed considered *taan* to be a pronoun and not an anaphor. Furthermore the one instance in which *ko(n)* occurs on the verb and *taan* only has a sloppy interpretation like the reflexive in (52) is not a strong counter-example to the fact that *taan* could be a pronoun. This is because the pronoun *aven* in Tamil yields the same pattern of strict and sloppy

readings as *taan* in the same contexts. Consider the following sentences with *ko(n)* on the verb. Note that in all the other sentences illustrated above, *aven* will give both strict and sloppy readings like *taan* and are not illustrated here for reasons of space.

55) [*Maaren*$_i$ *aven-ei*$_{i/ *j}$ *aditi-**ko**-nd-aan*] *enru* *Somu*
Maran.NOM 3sgm-ACC beat-**kon**-past-3sgm comp Somu
co-nn-aan *Raman-um* *kooda*.
Say-past-3sgm. Raman-COOR too
'Somu said that Maran beat himself. Raman did too.'
Interpretation: Raman said that Maran beat himself. (Sloppy)

In (55), *aven,* which is a normal pronoun, occurs with *ko(n)* on the verb. In these sentences, the following clause only has the sloppy interpretation. We cannot conclude that *aven* is not really a pronoun because of this one set of data. While we do not commit ourselves to anything concrete here, one could argue that *ko(n)* is merely behaving like some sort of operator which causes the anaphoric element in the clause to be bound within the scope of the operator, in this case the embedded clause. Likewise when *taan* occurs with *ko(n)* on the verb, the same thing could be happening. Because of this, we cannot reject the claim that *taan* is a pronoun merely because of one set of data where *ko(n)* occurs on the verb and only the sloppy reading is available. However, there is a crucial difference between the pronoun *aven* and *taan* which requires us to think of *taan* as different from *aven.* Consider the following set of data.

56) a. *Maaren*$_i$ *aven-ei*$_{*i/ j}$ *adi-tt-aan*. *Somu-vum* *kooda*.
Maran.NOM 3sgm-ACC beat-past-3sgm Somu-COOR too
'Maran beat him. Somu did too.'
Interpretation: Somu beat him. (Strict)

b. **Maaren*$_i$ *tann-ei*$_i$ *adi-tt-aan.* *Somu-vum* *kooda.*
Maran.NOM taan-ACC beat-past-3sgm Somu-COOR too
'Maran beat him. Somu did too.'

In (56a) *aven* occurs without *ko(n)* on the verb. Note that in this lead sentence, *aven* receives an interpretation of some salient discourse entity, say *Balan*. The elided *aven* in the following clause has to refer to the same discourse antecedent *Balan*. There is nothing unusual about this. Recall from our discussion earlier that when a free variable which has been given a particular value by the assignment function is elided, the reconstructed free variable has to have the same value. This is what is happening with *aven* in (56a). *Aven* has been assigned the value *Balan* and when this pronoun is elided and reconstructed at LF, the same pronoun with the same value is created and yields the strict reading. On the other hand, if *taan* was truly a free variable like *aven*, we would expect *taan* to behave in the same manner as *aven.* However, we find that this is not the case as can be seen in (56b). This indicates that *taan* cannot be completely free when it occurs as a free variable. While *taan* does give rise to strict readings- which suggests that the assignment function is able to provide values to it- there do seem to be restrictions on the values which can be given to *taan* (in this case a discourse referent is not allowed). Conceptually, this is not difficult to reconcile with Heim & Kratzer's theory. Recall that the assignment function cannot assign values to variables which conflict in terms of features. For example, a female value cannot be assigned to a male pronoun as we saw earlier in the discussion of (48). Along the same lines, *taan* seems to have an additional restriction that any value that is given to it has to be from some context. However, this would have to mean that we cannot consider *taan* to be a normal pronoun as pronouns cross-linguistically seem able to pick out discourse entities and making an exception in the case of *taan* does not seem warranted. This requires us to conclude that *taan* is not a pronoun after all, contra Amritavalli (1984).

On the other hand, does this mean that *taan* is an anaphor like a reflexive? In (47) we saw that the English reflexive *himself* could only have the sloppy reading as it is unambiguously a bound variable. However, the fact that *taan* has sloppy and strict readings should not cause us to reject the possibility that it is not an anaphor. Huang (2005) shows that, cross-linguistically, a wide range of languages have anaphors which can have strict as well as sloppy readings. For example, consider the following Icelandic example taken from Thrainsson (1991:60)

57) *Jon*$_i$ *sadjoi adj* [*pu hefdjir svikidj sig*$_i$] *og Petur*
John said that you had betrayed self and Peter
gerdji padj lika
did so too.
'John said that you had betrayed self and Peter said so too.'
Interpretation: Peter said you had betrayed Peter. (Sloppy)
Peter said you had betrayed John. (Strict)

In (57) *sig* considered to be an anaphor can lead to both strict and sloppy interpretations just like *taan* above. This is not an isolated occurrence cross-linguistically as Hellan (1991) claims the same for Norwegian. Furthermore even English *himself* can give rise to strict readings in particular contexts. Consider the following sentence from Hestvik (1992).

58) Fred defended himself better than his lawyer did.
Interpretation: The lawyer defended Fred. (Strict)

In (58) although the reflexive has been elided in the following clause, the strict interpretation where the lawyer defended Fred is still available.

What we have seen so far with regards to VP ellipsis and *taan* suggests that the availability of strict and sloppy readings is not a definite way of determining whether an element is a pronoun or an anaphor. As such we do not seem to have come any closer to determining the status of *taan.* However we now know that whatever account that is

given to explain *taan* has to be able to explain not only all the data seen in Chapter 2 but also the VP ellipsis data seen above.

Thus far in the thesis, we have illustrated the various characteristics of *taan* by looking at syntactic as well as semantic approches to anaphora. However we cannot pinpoint what exactly determines the distribution of *taan* or how *taan* acquires its menaing. With this in mind, we now move on to the discourse aspects of *taan*, namely the possibility of *taan* being a logophoric pronoun. We highlight Sells (1987) and show that he can account for all of the *taan* data if we were to look at *taan* as a logophoric pronoun.

3.2 Logophoricity

Hagege (1974) was the first to coin the term *logophor* in his study of African languages and his term logophor refers to *a particular category of anaphoric pronouns, personal and possessive, which refer to the author of a discourse or to a participant whose thoughts are reported* (Translated by Stirling 1993: 253). Since then, many authors have done work on the African languages to determine the scope and ways in which logophoricity is realized in these languages and two main ways have been characterized. Languages like Ewe have logophoric pronouns distinct from their normal pronouns which are cliticized to the embedded verb (Clements 1975). The following example is taken from Clements (1975: 142).

59) a. *Kofi be **ye**-dzo*
Kofi say **Log**[10]-leave
'Kofi$_i$ said that he$_i$ left.'

[10] Here *Log* and *Pro* refer to logophoric pronoun and pronoun respectively. In (60), note that Log refers to logophoric marker.

b. *Kofi* *be* ***e**-dzo*
Kofi say **Pro**-leave
'Kofi$_i$ said that s/he$_j$ left.'

In (59a) the embedded verb has the cliticized *ye* on the embedded verb. The logophoric pronoun has to take the matrix subject as its antecedent. In (59b), however, the cliticized pronoun can only refer to some other referent. On the other hand, the second type of system of logophoricity does not utilize special logophoric pronouns. In these systems, the normal pronoun is used but a verbal affix is used to indicate that the pronoun is used logophorically. The data from Gokana taken from Comrie (1983) illustrates this.

60) a. *ae* *ko* *ae* *do*
He said he fell
'He$_i$ said that he$_j$ fell.'
b. *ae* *ko* *ae* *do-**e***
He said he fell-**Log**
'He$_i$ said that he$_i$ fell.'

In (60a) the embedded pronoun subject cannot refer to the matrix subject. However, in (60b) in the presence of the verbal suffix the coreference becomes obligatory. Apart from these systems Hagege has also claimed that logophors can also be realized as long distance reflexives in languages like Japanese.

While the concept of logophors was mainly used as a descriptive term in Hagege, Sells (1987) aimed to show that the idea of logophors can be reconciled within a larger framework of anaphora more formally. We show that Sells' theory can be adapted to account for Tamil *taan*. We turn to a brief description of Sell (1987) first.

3.2.1 Sells (1987): Source, Self and Pivot

Sells (1987) stays close to the definition of logophors first coined by Hagege (1974). However he claims that the concept *logophor* is actually made up of 3 more

primitive notions (the source, self and pivot). Sells also provides a formal representation of logophoricity using Kamp (1981)'s Discourse Representation Structures. This latter aspect of his paper is less important for our purposes here and, as a result, we will not go into it.

Sells (1987) identifies 4 different discourse environments in which his notions of source, self and pivot are classified as being internal or external to the sentence. The following table taken from Sells (1987: 456) shows this.

	Direct Speech	3rd Person 'point of view'	Psych Verb	'Logophoric' Verb
SOURCE	External	external	External	Internal
SELF	External	external	Internal	Internal
PIVOT	External	internal	Internal	Internal

Table 6 Sells (1987)'s Discourse Environments

Table 6 shows how the 3 primitive notions together determine the discourse environment depending on whether they are internal or external. According to Sells, the 'source' is the intentional agent of the communication and the 'self' is the person whose mental state or attitude is described. The 'pivot' is the one with respect to whom (space-time) location is evaluated. We will go on to illustrate what each of these primitives mean with examples that Sells uses.

61) a. John said that he saw Mary.
b. That Susan likes him pleases John.
c. John's mother came to the hospital to visit him.

In (61a) the sentence contains the logophoric verb *say* and the embedded subject pronoun logophorically links to *John*. This is because *John* is the source of the sentence. In (61b) the sentence contains the psych verb *please* and the pronoun links to *John* who is the self

in the sentence. In (61c) the 3rd person point of view arises due to the construction type and the pronoun can refer to *John* which is the pivot.

Sells describes three diagnostics which show that such primitives do exist in natural languages. Specifically, Sells claims that the use of evaluatives such as 'the fool', 'mysteriously' as well as deictic words such as 'come' and 'go' can show that the primitives of source, self and pivot respectively do exist in natural languages. We will not go into these diagnostics and we refer the interested reader to Sells (1987) to see how these diagnostics work as well as the formal account of his theory. Here we will move on to the criticisms of the theory that will motivate our modification of the theory which will then be applied to Tamil.

One of the main criticisms of Sells (1987) as pointed out by Sterling (1993) is that Sells mainly focuses on those logophors which coincide with long distance reflexives. Sterling claims that Sells does not provide an analysis which can account for local logophors like those found in African languages. We agree that this is a minor shortcoming of Sells' treatment but we will show that Sells' theory can be easily used to account for certain local anaphors in Tamil sentences.

Sterling also criticizes the fact that Sells' diagnostics are not reliable. This is because Sells claims that there is an implicational relationship between the source, self and pivot in a sentence which does not empirically hold. Specifically, Sells claims that when the source is internal, the self has to be internal as well. He says that this is due to the fact that verbal communication cannot occur without the consciousness that is behind the communication. He also claims that when self is internal, the pivot is internal as well. This is because, according to him, when a particular state of mind is reported such a

reporting can only be possible if one is standing in that person's shoes as well. We, on the other hand, agree with Sterling's criticism about the unreliability of Sells' diagnostics as the existence or significance of the implicational nature of the three primitives cannot be verified empirically. Consider (61a). Sells claims that in (61a) the self and pivot are internal since the source is internal, but it is not clear if the mental state and point of view of John are reported at all. This is because the sentence is not making any statement about John's internal mental state. Furthermore there are no clues regarding whose perspective the sentence is reported from. In (61b) according to Sells, since the self is internal, the pivot is also supposed to be that of the internal protagonist as shown in Table 6. However, in reality, the point of view could be that of some external protagonist who is mistaken about John being pleased. All of this means that the purported implicational nature of self and pivot in sentences with psych verbs is not necessary.

Furthermore the proposed implicational nature of the primitives does not seem to have any significance either as a pronoun picks up an antecedent that is the most relevant primitive in a particular discourse environment. In a sentence with a logophoric verb such as 'say' or 'think', the source is the most relevant primitive. In a sentence with a psych verb, the self is the most relevant primitive. And finally, in a sentence taking a 3rd person point of view, the pivot is the most relevant. Thus in (61a), *John,* being the source in a sentence with a logophoric verb, is the antecedent for *him*. In (61b) *John* is the self in a sentence with a psych verb and is the antecedent for the pronoun. In (61c) *John* is the pivot in a sentence taking a 3rd person point of view and is thus the antecedent for the pronoun. This can be established without postulating an implicational nature between the primitives.

With these in mind, we claim that the different discourse environments do exist across languages as Sells claims but that the implicational nature of the primitives does not. As described in (61) we will assume that each discourse environment has its own relevant primitive and it is this primitive that determines which antecedent is available for a logophoric pronoun. Thus for logophoric verbs, the source is relevant. For psych verbs, the self is relevant. In sentences reporting the 3rd person point of view, the pivot is important. In what follows, we will show that logophoric verbs, psych verbs and the pivot as described by Sells can be used to account for the distribution of *taan* wherever it occurs. A crucial claim of our treatment is that *taan* is a logophoric pronoun. In addition, we will also be showing that *ko(n)* is actually a pivot marker. After that, we will provide a description of all the predictions that are made by such an analysis as well as account for the Tamil VP ellipsis facts that we saw earlier. We wrap up the thesis by invoking the Chain Condition to explain all the distributional differences between *taan* and *aven* as well as matters of word order.

3.3 Taan as a Logophoric Pronoun

In this sub-section, we claim that *taan* is a logophoric pronoun which always refers to the relevant primitive in a particular discourse environment. We will apply Sells' simplified theory to Tamil and show that it can account for a lot of the data. Sentences with logophoric verbs and psych verbs can easily be found and accounted for in Tamil as such discourse environments are licensed by the verb cross-linguistically and Tamil verbs such as *col-* 'say' and *pidi* 'like' license the respective discourse environments in Tamil as well. In such discourse environments, we will show that *taan* will pick out the source and self respectively. On the other hand, a 3rd person point of view is not licensed by a

verb but rather arises due to a particular construction type. We will show that such a construction arises when *ko(n)* occurs on the verb. We also incorporate the Chain Condition in our analysis to provide an account for *taan* as well as the pronoun *aven*. We wrap up this chapter by showing that some of the more pertinent predictions which arise due to our analysis of *taan* here do fall out.

3.3.1 Logophoric Verbs and *taan*

The discourse environment of a logophoric verb arises due to a verb such as *col-* 'say'. In our slight modification of Sells' theory, we claim that in such a discourse environment only the source in the sentence can be the antecedent of a logophoric pronoun and that the presence or nature of the other primitives, self and pivot, are not relevant. We find this to be the case in Tamil. Consider the following. The data below is reproduced from (11c) and (12c) above.

62) a. [*taan* $_{i/*j/*k/*m}$ *Maaran-ei*$_j$ *adi-tt-aan*] *enru*
taan.NOM Maran-ACC beat-past-3sgm comp
Somu$_i$ *Raman-idam*$_k$ *co-nn-aan*
Somu.NOM Raman-LOC say-past-3sg
Somu told Raman that he (Somu) beat Maran.

b. [*tan-iki*$_{i/ *j/*m}$ *pasi-kum*] *enru* *Somu*$_i$ *Raman-idam*$_j$
taan-DAT hunger-fut comp Somu.NOM Raman-LOC
co-nn-aan
say-past-3sgm
Somu told Raman that he (Somu) will be hungry.

In (62) *taan* occurs as the embedded subject with either nominative (a) or dative case (b). The clause itself is embedded by the verb *col* 'say'. We mentioned earlier that *taan* can pick out the matrix subject (*Somu*) but not the matrix object (*Raman)* as its antecedent just as we are seeing in (62). The reason for this is straightforward if we think of *col-* 'say' as licensing the discourse environment of a logophoric verb within which *taan,*

being a logophoric pronoun, can only refer to the source which in these sentences is *Somu*. While this works, we have as yet shown any evidence which suggests that a c-command relation is not responsible for *taan* picking out only the matrix subject as its antecedent. After all, only the matrix subject c-commands the embedded clause. We can effectively rule out a c-command relationship to be the reason behind the coreference possibilities in (64) with the following example.

63) [*taan*$_{i/*j}$ *migavum* *puthisaali*] *enpathu* *Raman-in*$_i$
Taan very intelligent comp Raman-Gen
nambikkei.
belief
It is Raman's belief that he (Raman) is smart.

In (63) the antecedent of the embedded subject *taan* is the possessor in the genitive NP in the matrix clause. In this sentence, *Raman* clearly does not c-command the embedded clause, yet can still be the antecedent of *taan*. This means that the actual way in which *taan* gets its antecedent in (62) is by linking to the matrix subject some other way. Here we have claimed that the mechanism responsible is logophoricity. The other instance where *taan* picks out the source of a sentence with a logophoric verb is when *taan* occurs as an object with a nominative case marked subject with no *ko(n)* on the verb. However we leave this to when we illustrate the distribution of *ko(n)* below. We now move on to psych verbs and *taan*.

3.3.2 Psych Verbs and *taan*

The next discourse environment that Sells outlines that we will be looking at is his psych verbs. Recall that in this discourse environment, the relevant primitive is the self. It appears that (65) from above is an obvious candidate to be considered as representative of this discourse environment. We reproduce it below.

64) [*taan*$_{i/*j}$ *migavum puthisaali*] *enpathu Raman-in*$_i$
Taan very intelligent comp Raman-Gen
nambikkei.
belief
It is Raman's belief that he (Raman) is smart.

In (64) there is no logophoric verb such as 'say' in the matrix clause. However it is obvious that the mental state of the internal protagonist is being reported due to the matrix predicate 'belief'. This would mean that in this sentence, *taan,* a logophoric pronoun is linking to the self within the discourse environment created by the psych verb. Although Sells does not explicitly illustrate how local logophors can be accounted for in this way, we show that such an account can also be extended to simple sentences which have a psych verb. Consider (65) taken from (7a) and (8a) above.

65) a. *Maaren*$_i$ *tann-ei*$_{i/*j}$ *veru-tt-aan*
Maran.NOM taan-ACC hate-past-3sgm
'Maran hates himself'.
b. *Maaren-iki*$_i$ *tann-ei*$_{i/*j}$ *pidik-um*
Maran-DAT taan-ACC like-fut
'Maran likes himself.'

In (65) the verbs *veru* 'hate' and *pidi* 'like' both report on the internal state of mind of the subject and license the discourse environment of a psych verb. In such a discourse environment *taan* picks out the self, *Maran,* as its antecedent. Recall that we can easily account for sentences such as (65) with our account whereas R&R's theory can only cope with this piece of data with a lot of added stipulations.

So far we have shown the following- *Taan* is a logophoric pronoun and it links to the relevant primitive in the appropriate discourse environment licensed by a particular verb. In the discourse environment of a logophoric verb licensed by verbs such as *col-* 'say', *taan* refers to the source of the sentence and in the discourse environment of a psych verb licensed by verbs such as *pidi-* 'like', *taan* refers to the self of the sentence.

With these two discourse environments, we have accounted for a lot of the Tamil basic data set. In fact we have only one other significant basic Tamil data set to account for. This is the sentences where *ko(n)* occurs on the verb. Unfortunately, accounting for these sentences is not as straightforward as what we have already seen with the rest of the data. But in explaining our treatment of such sentences, we will provide an account for the distribution of *ko(n)* as well as pronouns.

We will show in the next section that *ko(n)* is a pivot marker which creates a discourse environment where only the subject can be the pivot of the sentence. We will first start by providing a much clearer definition for what a pivot is than what Sells provides.

3.3.3 3rd Person point-of-view and *taan*

According to Sells the pivot of a sentence is the entity with respect to whom (space-time) location is evaluated. Unlike logophoric verbs and psych verbs, the 3rd person point of view, as Sterling (1993) points out, arises due to specific constructions and not due to a lexically specified verb. We claim that in Tamil such a construction arises when *ko(n)* is affixed to the verb. We will now go on to illustrate this.

Sells claims that his notion of pivot is meant to be understood in a very physical way. Thus when a particular entity, say *John,* is the pivot of a sentence, what is predicated by the verb has to occur from John's physical perspective. This is why Sells uses deictic terms to help him identify pivots. For instance, Sells claims that in sentences such as (61c), the sentence is evaluated from John's perspective and thus only the deictic term 'came' can be used. Although this piece of judgment is by no means rigorous as

indicated by other authors like Sterling (1993), for Sells, 'went' is actually ungrammatical.

We claim that in Tamil, *ko(n)* also causes the sentence to be evaluated from the perspective of the pivot. However, we will attempt to define it in a more precise way. Consider the following.

66) a. *Maaren*$_i$ *tann-ei*$_{i/*j}$ *aditi-**ko**-nd-aan*
Maran.NOM taan-ACC beat-**kon**-past-3sgm
'Maran beat himself'.

b. **Maaren* *tann-ei* *aditi-tt-aan*
Maran.NOM taan-ACC beat-past-3sgm
'Maran beat himself'.

In (66a) we are claiming that *ko(n)* introduces the 3rd person point of view to the sentence and requires the sentence to be evaluated from the subject's (*Maran*) physical perspective. We will define the pivot as well as outline the following hypothesis about *ko(n)* in Tamil as follows.

67) a. PIVOT (*first formulation*)
The pivot in Tamil is the physical-temporal space on which the activity described by the verb takes place.

b. *KO(N)* HYPOTHESIS (*first formulation*)
Ko(n) requires the subject to be the pivot of the sentence.

What (67) means is that no matter what the verb is, if *ko(n)* is present, the subject becomes the pivot and the action or activity described by the verb has to literally happen on the subject. We can illustrate this using (66). In (66a) where *ko(n)* marks the verb, the subject, *Maran,* has to be the pivot; the action described by the verb has to happen 'on' the subject. Thus the action of 'beat' has to happen on *Maran*. Note that this means that *taan* is not necessary or sufficient for the intended meaning to come across. One corollary of this can be seen in (66b) where there is no *ko(n)* on the verb. Since there is no 3rd

person point of view in this sentence, even though *taan* occurs in the sentence, the intended meaning where *taan* is coindexed with the subject is not possible. On the other hand, as long as there is *ko(n)* and the object position is filled with an element which is not referentially saturated (like r-expressions), the sentence will be licit. This explains why in sentences with *ko(n)* on the verb, *aven* is perfectly fine as shown below.

68) *Maaren*$_i$ *aven-ei*$_{i/*j}$ *aditi-**ko**-nd-aan*
Maran.NOM 3sgm-ACC beat-**kon**-past-3sgm
'Maran beat himself'.

In (68) with *ko(n)* on the verb, the subject of the sentence must be the physical-temporal location on which the activity described by the verb occurs. In this sentence the object is not an r-expression which is fully specified but a pronoun which has to be assigned a value. Since *ko(n)* requires the action to take place on the subject, the only way this can happen is if *aven* is given the value of *Maran*. If (67) is correct, then we will expect that sentences with the verb 'beat' and *ko(n)* will not allow an object which is an r-expression. This is because with an object r-expression, which is distinct from the subject, the action of beating cannot possibily be on the subject when beating someone else. This prediction is borne out.

69) a. **Maaren* *Mala-vei* *aditi-**ko**-nd-aan*
Maran.NOM Mala-ACC beat-**kon**-past-3sgm
'Maran beat Mala'.
b. *Maaren* *Mala-vei* *adi-tt-aan*
Maran.NOM Mala-ACC beat-past-3sgm
'Maran beat Mala'.

In (69a) the verb has *ko(n)* on it and the action has to take place on the subject. However, with *Mala* as the object, the act of beating has to happen on *Mala* and not the subject. There is thus a disconnect between the spatial-temporal description of the verb phrase and what *ko(n)* requires and the sentence is ungrammatical. (69b), which does not have

ko(n) is perfectly fine with *Mala* as the object as there is no requirement that the subject be the pivot. However, there are sentences where the object is a distinct r-expression with *ko(n)* on the verb but yet still grammatical. This is because even in these sentences, the pivot requirement can be maintained. We will now go on to show this. Consider the following sentences.

70) a. *Maaren* *naak-ei* *niiti-**ko**-nd-aan*
Maran.NOM tongue-ACC stick.out-**kon**-past-3sgm
Maran stuck out (*a/ his) tongue.

b. *Maaren* *naak-ei* *niiti-n-aan*
Maran.NOM tongue-ACC stick.out-past-3sgm
Maran stuck out (a/ his) tongue.

In (70a) the object position is occupied by 'tongue'. Since there is *ko(n)* on the verb, we should expect this sentence to be ungrammatical as the object is a distinct r-expression from the subject. However, this sentence is grammatical as the activity described by the verb can still occur on the subject as required by *ko(n)* as described in (67). One situation where (67) can be maintained is if the tongue is *Maran*'s tongue. This is indeed the interpretation which yields. In (70b) where there is no *ko(n)* on the verb, the tongue can be his own tongue or some other entity's. The minimal dfference between these two sentences is the presence of *ko(n)* on the verb and this does have consequences for the interpretation of the sentence. Note that (70a) behaves in accordance with (67) where we outlined what the pivot in Tamil is. The sticking out of the tongue has to happen 'on' the subject with *ko(n)* and the only way such an interpretation can be derived is if the tongue belonged to *Maran* and this is indeed the only interpretation possible. The following sentences show that some other entity's tongue cannot be explicit in (70a) but possible in (70b) which further supports our claims here.

71) a. **Maran* *pomei-yudeya* *naak-ei* *niiti-**ko**-nd-aan.*
Maran.NOM doll-GEN tongue-ACC stick.out-**kon**-past-3sgm
Maran stuck out a doll's tongue tongue.

b. *Maran* *pomei-yudeya* *naak-ei* *niiti-n-aan.*
Maran.NOM doll-GEN tongue-ACC stick.out-past-3sgm
Maran stuck out a doll's tongue.

In (71a) with *ko(n)* on the verb, the sentence is ungrammatical as the tongue is explicitly mentioned to belong to a doll. This violates (67). On the other hand in (71b) the tongue can belong to a doll in the absence of *ko(n)*. So far our first formulation of the pivot has been able to account for all the sentences with *ko(n)* thus far. However the picture is not complete. Consider the following.

72) a. *Maaren* *kathav-ei* *moodi-**ko**-nd-aan.*
Maran.NOM door-ACC close-**kon**-past-3sgm
Maran closed the door (in on himself).

b. *Maaren* *kathav-ei* *moodi-n-aan.*
Maran.NOM door-ACC close-past-3sgm
Maran closed the door.

In (72a) the object is 'door' and there is no possible interpretation in which the door can be closed 'on' the subject as the 'door' and *Maran* are distinct physical locations. However, the sentence is still grammatical. Although our first formulation of what the pivot in Tamil is does not seem able to capture (72a), we claim that only a revision to our definitions in (67) is required. This is because the translation in (72a) shows that *Maran* has to be closed in by the door whereas such an interpretation is not necessary in (72b). The following discourse scenarios will make the meaning more apparent.

73) a. *#Maaren* *katha-vei* *moodi-**ko**-nd-aan.* *Atharpiragu*
Maran.NOM door-ACC close-**kon**-past-3sgm After.that
udane *veliye* *se-ndr-aan.*
immediately outside go-past-3sgm
Maran closed the door in on himself. Immediately after that, he went out.

b. *Maaren* *katha-vei* *moodi-n-aan.* *Atharpiragu*
Maran.NOM door-ACC close-past-3sgm After.that
udane *veliye* *se-ndr-aan.*
immediately outside go-past-3sgm
Maran closed the door. Immediately after that, he went out.

(73a) is infelicitous as *Maran* has to close the door in on himself and it will not be possible for him to leave the house immediately after that. However, in (73b) where there is no *ko(n)* on the verb, there is no such restriction on *Maran's* position and thus the fact that he leaves immediately after closing the door is not infelicitous. What we have shown with (72) is that while our formulation of pivot in (67a) does not capture (72), *Maran*'s physical location is still relevant when *ko(n)* is on the verb. With this is mind, we reformulate our definitions which will incorporate all the sentences we have seen so far.

74) a. PIVOT (*second formulation*)
The pivot in Tamil is the physical-temporal *point* around which the activity described by the verb is carried out.

b. KO(N) HYPOTHESIS (*second formulation*)
Ko(n) allows only the subject-as-pivot interpretation.

In (74) we provide our formulation of the pivot in Tamil as a single discrete physical point which has to be established. Thereafter the activity described by the verb will be carried out in some way relative to this point. Furthermore we claim that what *ko(n)* does is to require the sentence to have the subject-as-pivot interpretation which merely means that only the subject can be the physical temporal point around which the activity described by the verb can occur. In (72b) where there is no *ko(n)*, both interpretations are possible. If the object-as-pivot interpretation is used, then the position of *Maran* is irrelevant. When the subject-as-pivot interpretation is used, then *Maran* has to be enclosed by the door. However, when there is *ko(n)* on the verb as in (72a), only the

subject-as-pivot interpretation is possible and thus the only possible interpretation is the one where the door encloses *Maran.*

All the sentences that we have already seen in this section can be reconciled similarly. In (66a) when there is *ko(n)*, only the subject-as-pivot interpretation is possible. Thus when there is *taan* or *aven* (68) occurring as the object, these objects have to be coindexed with the subject as the only way a felicitous interpretation obtains is if the subject is carrying out the action of beating on him or herself. (66b) is ungrammatical because the subject-as-pivot and the object-as-pivot interpretations are both not allowed. Since the sentence has no physical temporal location, the entire sentence is ungrammatical. The subject-as-pivot interpretation can potentially arise if *taan* is coindexed with the subject, however, this is not possible as *taan* is a logophoric pronoun. For *taan* to link to the subject without *ko(n)*, the subject has to be some sort of logophoric center, i.e. source or self, which it clearly is not. On the other hand, the object-as-pivot interpretation would have been possible if *taan* refers to some other extra-sentential element but this is also not possible as *taan* once again requires a logophoric center as its antecedent. When there is some logophoric center in the form of a matrx subject (source), the sentence is, as predicted, grammatical as seen in (6a). The judgements in (69) can also be easily explained using (74). In (69b) where there is no *ko(n)* on the verb, only the object-as-pivot interpretation obtains as the activity of beating is happening at the physical 'point' described by the object. In (69a) *ko(n)* requires the subject-as-pivot interpretation which is not available in the first place in such sentences. Since there is a conflict between what *ko(n)* requires and what is available in (69a), the sentence is ungrammatical.

The distribution of (70) falls out similarly. In (70b) both the subject-as-pivot and object-as-pivot interpretations are available as the tongue could plausibly belong to the subject *Maran* or a doll. However, in (70a) with *ko(n),* only the subject-as-pivot interpretation is possible and thus this is the only interpretation that is possible as seen in (71a).

With what we have seen so far, it appears that the occurrence of *ko(n)* does require the subject-as-pivot interpretation of the sentence. Note that this does not mean that a sentence can only have one single reading. Consider the following sentences.

75) a. *Maaren* *kudei-yei* *pidithi-**ko**-nd-aan*
Maran.NOM umbrella-ACC hold-**kon**-past-3sgm
Maran held the umbrella.
b. *Maaren* *kudei-yei* *pidi-tt-aan*
Maran.NOM umbrella-ACC hold-past-3sgm
Maran held the umbrella.

In (75a) *Maran* and 'umbrella' are two possible physical points but the presence of *ko(n)* means that only subject-as-pivot interpretation is allowed. This means that the space of 'umbrella' is somehow incorporated into the pivot. There are (at least) two ways in which such incorporation can occur. One way is if the umbrella is being held open over the subject and the second reading is if the umbrella is being held closed in the subject's hand. In both readings, note that only the subject-as-pivot interpretation can be maintained. Additional data does support this interpretation. Consider the following.

76) a. *#naan* *maLai-yil* *nanei-yaamel* *iruka* [*Maaren* *enakaage*
1sg rain-LOC wet-not be Maran.NOM for.me
kudei-yei *pidithi-**ko**-nd-aan*]
umbrella-ACC hold-**kon**-past-3sgm
For: Maran held the umbrella for me ao that I will not get wet in the rain.

b. *naan* *maLai-yil* *nanei-yaamel* *iruka* [*Maaren* *enakaage*
1sg rain-LOC wet-not be Maran.NOM for.me
kudei-yei *pidi-tt-aan*]
umbrella-ACC hold-past-3sgm
Maran held the umbrella for me ao that I will not get wet in the rain.

In (76a) with *ko(n)*, the reading where *Maaren* holds the umbrella as a shelter for the 1st person author is infelicitous. This is exactly what our analysis predicts. When there is *ko(n)* on the verb, only the subject-as-pivot interpretation is possible. However when the subject is holding out the umbrella away from himself for another person, the subject is no longer the point around which the activity occurs and this is not allowed in a sentence like (75a). In (75b) and (76b) where *ko(n)* does not occur on the verb, such a reading is perfectly fine.

We believe that we have shown that thinking of *ko(n)* as a pivot marker is indeed the right way to characterize Tamil *ko(n)*. *Ko(n)* does not have a syntactic or semantic character but a logophoric one, specifically one which requires a subject-as-pivot interpretation. This means that *ko(n)* can occur in a wide variety of readings, ranging from reflexive readings (66a & 68) to inalienable possession readings (70a) to location-specific readings (72a & 75a). The one commonality in all these readings is the subject-as-pivot interpretation.

With the characterization of *ko(n)* as a pivot marker, we can complete the picture of logophoricity in Tamil. In the next sub-section of this thesis which will also be the last before a summary of our findings, we will outline some of the predictions that our claim makes and show that the predictions all pan out as expected. This will further strengthen our claim that *taan* is indeed a logophoric pronoun and that the characterization that we have provided for it as well as *ko(n)* in this thesis is on the right track.

3.3.4 Wrapping Up: *taan* as a Logophoric Pronoun

In this section of the chapter, we outline the 3 data sets mentioned in this thesis but have yet to be accounted for and show that our analysis of *taan* here accounts for all of them. These 3 are embedded clauses, the inability of psych verbs to have *ko(n)* as a suffix as well as the VP-ellipsis data.

3.3.4.1 Embedded Clauses

In Chapter 2 we showed that in certain clause embeddings, *taan* can have ambiguous antecedents while in other clause embeddings, *taan* can take only one antecedent. In this sub-section, we will show how our analysis accounts for all of them. We will start with embedded clauses with psych verbs.

77) a. [*Maaren*$_i$ *tann-ei*$_{i/\,j/\,*k}$ *veru-tt-aan*] *enru*
Maran.NOM taan-ACC hate-past-3sgm comp
Somu$_j$ *co-nn-aan*
Somu.NOM say-past-3sgm
'Somu said that Maran hates himself/ him'.

b. [*Maaren-iki*$_i$ *tann-ei*$_{i/\,j/\,*k}$ *pidik-um*] *enru* *Somu*$_j$
Maran-DAT taan-ACC like-FUT comp Somu.NOM
co-nn-aan
say-past-3sgm
'Somu said that Maran likes himself/ him.'

In (77a) and (77b), the embedded clause contains a psych verb. In these sentences, the embedded object *taan* can refer to the embedded subject *Maran* or the matrix subject *Somu*. Note that in our analysis, the embedded subject is the logophoric center of 'self' and the matrix subject is the logophoric center of 'source'. This means that *taan* as a logophoric pronoun can potentially link to both the 'source' and 'self' in a sentence where both are available. We have seen that accounting for embedded psych verbs is

straightforward in our account. We now move on to embedded clauses with *ko(n)*. The following are taken from (5c) and (5d) respectively.

78) a. [*Maaren*$_i$ *tann-ei*$_{*i/\ j/*k}$ *adi-t-aan*] *enru*
Maran.NOM taan-ACC beat-past-3sgm comp
Somu$_j$ *co-nn-aan*
Somu.NOM say-past-3sgm
'Somu said that Maran beat him'

b. [*Maaren*$_i$ *tann-ei*$_{i/\ *j/\ *k}$ *aditi-**ko**-nd-aan*] *enru* *Somu*$_j$
Maran.NOM taan-ACC beat-**kon**-past-3sgm comp Somu
co-nn-aan
say-past-3sgm
'Somu said that Maran beat himself'.

In (78a) the embedded verb does not have *ko(n)* and *taan* can only refer to the matrix subject *Somu*. In (78b) when the embedded verb does have *ko(n), taan* can only refer to the embedded subject *Maran.* This set of data can also be explained straightforwardly in our account. In (78a) *taan* can only refer to the source which is the matrix subject *Somu* as there is no other logophoric center available in the embedded clause. In (78b) *ko(n)* requires the subject-as-pivot reading and as a result, *taan* must refer to the embedded subject for such a reading to arise. One might raise the question of why *taan* cannot refer to the source in (78b). Consider what will happen if such a reading was possible. *Taan* will now be referring to the matrix subject but the embedded clause which is marked with *ko(n)* will no longer have the subject-as-pivot interpretation as the activity of beating will not be at the point represented by the embedded subject. This as we have already seen is not allowed and the entire sentence will be ungrammatical. To prevent such ungrammaticality, in sentences such as (78b) only one interpretation is possible. Below in the section on VP ellipsis, we will further elaborate that the reason why *taan* cannot refer to the source in such a sentence is due to the fact that only the binding mechanism is functional in these sentences. The assignment function which would otherwise be able to

assign the value of the source to *taan* is obviated in such sentences for reasons we will illustrate below.

In this sub-section we have looked at embedded clauses in Tamil and shown that the reasons behind the fact that *taan* can sometimes refer ambiguously and sometimes not falls out straightforwardly from our account in this thesis. We shall now move on to look at the reason why *ko(n)* is not compatible with psych verbs.

3.3.4.2 Psych Verbs and *ko(n)*

Consider the following taken from above.

79) a. *Maaren-iki*$_i$ *tann-ei*$_{i/*j}$ *pidik-um*
Maran-DAT taan-ACC like-FUT
'Maran likes himself.'

b. **Maaren-iki* *tann-ei* *pidi-**kon**-um*
Maren-DAT taan-ACC like-**kon**-FUT
For: 'Maran likes himself.'

c. *Maaren*$_i$ *tann-ei*$_{i/*j}$ *veru-tt-aan*
Maran.NOM taan-ACC hate-past-3sgm
'Maran hates himself.'

d. **Maaren* *tann-ei* *veruti-**ko**-nd-aan*
Maran.NOM taan-ACC hate-**kon**-past-3sgm
For: 'Maran hates himself.'

(79) shows that *ko(n)* is not allowed on a psych verb. (79a) and (79c) follows from our our analysis of *taan* as a logophoric pronoun which links to a logophoric center, in this case, the self. However we have yet to explain why *ko(n)* is incompatible with a psych verb as shown in (79b) and (79d). The solution is simple though. It lies in the nature of the verbs in these sentences and the function that *ko(n)* has. Recall that the pivot of a sentence is the physical-temporal point around which the activity described by the verb happens and that *ko(n)* requires this point to be the subject (what we have been calling the subject-as-pivot interpretation). In the sentences where *ko(n)* occurs, the verbs have been

such that the action described can be directed towards a particular point. Such a direction can only be possible if the verb describes an action which can take place at a discrete point in time and space. Verbs such as *adi-* 'beat' fall into such a characterization. However psych verbs by their nature are stative and not directional. Although one can speak of 'hate' or 'love' being directed at someone or oneself, such a direction can only be done in the metaphorical sense. This means that the verb does not describe an action which occurs in a discrete point in time and space. This results in *ko(n)* not being compatible with verbs which do not describe an action which can be literally directed and this is what we see with sentences such as (79b) and (79d).

We now move on to our final piece of data that we will account for. This concerns the VP ellipsis data that we looked at earlier.

3.3.4.3 VP Ellipsis and *taan*

In the VP ellipsis data that we saw in Section 3.1, we found that *taan* always gave rise to strict and sloppy readings except when *taan* is an object with a nominative subject and there is *ko(n)* on the verb. In this instance, only the sloppy reading is possible. We briefly reiterate our findings about *taan* in that section before explaining the data using our account. The following are taken from above.

80) a. [*taan* $_{i/*j/*k}$ *paadeth-ei* *padi-tt-aan*] *enru*
taan.NOM lesson-ACC study-past-3sgm comp
Somu$_i$ *Raman-idam*$_j$ *co-nn-aan.* *Maran-um* *kooda.*
Somu.NOM Raman-LOC say-past-3sg Maran-COOR too
'Somu told Raman that he (Somu) studied the lesson. Maran did too.'
Interpretation: Maran told Raman that Maran studied. (Sloppy)
Maran told Raman that Somu studied. (Strict)

b. *Maaren-iki*$_i$ *tann-ei*$_i$ *pidik-um.* *Raman-ik-um* *kooda*
Maran-DAT taan-ACC like-FUT Raman-DAT-COOR too
'Maran likes himself. Raman did too.'
Interpretation: Raman likes Raman. (Sloppy)
Raman likes Maran. (Strict)

In (80a) the following clause can have the interpretation that *Maran* studied or that *Somu* studied. Similarly in (80b) the following clause can have the interpretation that *Raman* is liked or that *Maran* is liked. We find a similar scenario when *ko(n)* does not occur on the verb when *taan* occurs as an object in embedded clauses with a nominative subject.

81) [*Maaren*$_i$ *tann-ei*$_{*i\,/\,j/\,*k}$ *adi-tt-aan*] *enru* *Somu*$_j$
Maran.NOM taan-ACC beat-past-3sgm comp Somu
co-nn-aan. *Raman- um* *kooda*
say-past-3sgm Raman-COOR too
'Somu said that Maran beat him. Raman did too.'
Interpretation: Raman said that Maran beat Raman. (Sloppy)
Raman said that Maran beat Somu. (Strict)

In (81) the following clause can have the interpretation where *Raman* was beaten or *Somu* was beaten. The availability of strict and sloppy readings in this set of sentences is similar to those found in (80). However when *ko(n)* occurs on the verb, only the sloppy reading is available.

82) *Maaren*$_i$ *tann-ei*$_{i/\,*j}$ *aditi-**ko**-nd-aan.* *Somu-vum* *kooda*
Maran.NOM taan-ACC beat-**kon**-past-3sgm Somu-COOR too
'Maran beat himself. Somu did too.'
Interpretation: Somu beat himself. (Sloppy)

In (82) the following clause can only have the interpretation where *Somu* beat himself. In the earlier section when we described these data in detail, we concluded that *taan* behaves neither like a pronoun nor an anaphor because of certain standard assumptions outlined earlier. The reader is invited to review Chapter 3.1.2. However looking at (80-82), we can conclude that such a neat classification is not possible with *taan* as it is

ambiguously a free and bound variable in some contexts and only a bound variable in others.

We can now claim that the reason why is because of the fact that it is neither an anaphor nor a pronoun but a logophoric pronoun. Recall that we outlined Heim & Kratzer (1998)'s account of how a variable can be assigned a value through the assignment function or by being bound. A pronoun can be given a value through both mechanisms while an anaphor can only be bound. *Taan*, on the other hand, being a logophoric pronoun can be given a value through both mechanisms just like a pronoun but is more sensitive to whether a logophoric reading is present.

A pronoun which can be generally assigned a value from the assignment function can refer to some discourse referent. However the assignment function 'knows' that certain values are not possible even with pronouns in specific contexts. Recall our example from above.

83) Domain of Individuals and their corresponding indices:
John → 1
Tom → 2
Alan → 3
Susan → 4

84) John likes his$_{1,2,3}$ mother.

In (84), there are 4 elements in the domain of individuals and *Susan* is automatically rejected as it does not match the feature specifications of the pronoun. Out of the three choices then, only one will be possible depending on further contextual restrictions. Likewise *taan* can also be assigned a value through the assignment function which will then be restricted. However *taan* is sensitive to logophoric contexts. One possible way in which *taan* gets its value is shown below.

85) *John-iki* *tann-ei*$_{1, 2, 3, 4}$ *pidik-um*

In (85) *taan* can be assigned values from the domain of individuals shown in (83). Since there is no feature mismatch between any of the individuals and *taan,* all are possibilities. However since *taan* is a logophoric pronoun which can only refer to the self in sentences such as (85), only value '1' survives and *taan* will be given the value of *John.* Apart from the assignment function, *taan* can also be given a value through being bound by *John.* Since the logophoric context only requires *taan* to refer to the self, there is no conflict with *taan* being bound by *John*. Note that in our account, the binding will only survive if such a binding does not conflict with the logophoric reading of *taan*. In (85) a conflict does not arise and the binding goes through. These two ways in which *taan* gets its value then straightforwardly explains why there are strict and sloppy readings in (80b). When *taan* gets the value of *John* through the assignment function, the elided *taan* will also get the value of *John* when reconstructed at LF. When *taan* is bound by *John*, the reconstructed *taan* will also be bound. The presence of strict and sloppy readings in (80a) arises in the same way. *Taan* can get the value of *Somu* through the assignment function or by being bound. With the former, the strict reading arises and with the latter, the sloppy reading arises.

(82) is not so straightforward as only the sloppy reading is available. This means that for some reason the assignment function is no longer responsible for the value of *taan*. Only the binding goes through. This must mean that *taan* is unambiguously a bound variable in such sentences. Even a pronoun which is utilized in such sentences will become unambiguously a bound variable. The reason why this happens probably has to do with the occurrence of *ko(n)*. Since *ko(n)* requires the subject-as-pivot reading, by

removing the assignment function from consideration, the sentence ensures that there can be no way the assignment function will assign a value which is not the embedded subject, especially in a sentence where a source as the matrix subject is available. With just the binding mechanism, the restriction that *ko(n)* places can be satisfied everytime.

One last minor point to take care of before ending the section pertains to what we said earlier about binding 'surviving'. While binding and the assignment function are ways in which an element can get a value, it is important to note that these mechanisms are only responsible for placing an appropriate value on *taan*. In other words, since *taan* is a logophoric pronoun, only those mechanisms which lead to a logophoric reading will be utilized. An element like *taan* does not "care" about which way it gets its value as long as the appropriate logophoric reading is achieved. This was seen in all the previous VP ellipsis sentences. In (80a) and (81) the source reading obtains through binding or the assignment function. In (80b) the self reading is achieved in these two ways as well. Thus we see strict and sloppy readings in all these sentences. In (82) only the binding mechanism is utilized to achieve the subject-as-pivot interpretation and as a result only the sloppy reading is available. One prediction that such an analysis makes is that sentences where *taan* gets a logophoric reading only through the assignment function should only lead to strict readings. This is exactly what we see. Consider the following.

86)	[*taan*$_{i/*j}$	*migavum*	*puthisaali*]	*enpathu*	*Raman-in*$_i$
	Taan	very	intelligent	comp	Raman-Gen
	nambikkei.				
	belief				
	It is Raman's belief that he (Raman) is smart.				

We encountered (86) earlier when we were talking about psych verbs and in this sentence, the antecedent (*Raman*) does not c-command *taan*. This means that *taan* cannot be bound

by *Raman.* Since *Raman* is the logophoric center of the self in this sentence, *taan* should link to it through the one remaining mechanism, the assignment function. This should result in only the strict reading in VP ellipsis.

87) *Somu-vin* *nambikkei-yum* *athu-thaan*.
Somu-GEN belief-COOR that-too
Somu's belief is that too.
Interpretation: Somu's belief is that Raman is very smart. (Strict)

In (87) when the entire embedded clause is elided, as expected, only the strict reading is available. This arises from the fact that in this sentence, *taan* can get a value only through the assignment function as a binding relationship does not exist. The assignment function then assigns the same value to the reconstructed *taan* at LF thus leading to the strict reading.

With that we believe that we have shown how VP ellipsis works in Tamil sentences with *taan* under our account. It is clear that considering *taan* a logophoric pronoun allows us to account for all the facts that we have seen with respect to embedded clauses, psych verbs and *ko(n)* as well as VP ellipsis. Together with the basic facts about Tamil *taan* we hope to have presented a compelling case for our claim that *taan* indeed a logophoric pronoun. There is one last state of affairs to take care before concluding. This has to do with how pronouns fit into the picture as well as matters of word order. We will use the Chain Condition to derive the remaining facts.

3.3.5 Pronouns and the Chain Condition

In the above sections we have seen how *taan* is distributed. However we have said little about the distribution of pronouns or matters of word order. Consider the following.

88) a. *Maaren-iki$_i$* *tann-ei$_i$* *pidik-um*
Maran-DAT taan-ACC like-fut
'Maran likes himself.'

b. **Maaren-iki$_i$* *aven-ei$_i$* *pidik-um*
Maran-DAT 3sgm-ACC like-fut
'Maran likes him.'

c. **tan-iki$_i$* *Maran-ei$_i$* *pidik-um*
taan-DAT Maran-ACC like-fut
'Maran likes himself.'

(88a) shows a psych verb where *taan* refers to *Maran*. We now know that the reason for this is because *taan* is a logophoric pronoun which refers to the self (*Maran*). However nothing we have said so far will rule out (88b) where a pronoun cannot logophorically refer to the subject or (88c) where *taan* is the subject. To correctly rule these sentences out, we will claim, following R&R, that the Chain Condition is required. We will show that the Chain Condition is the final piece in the picture and with this additional independently needed principle we can account for the entire set of Tamil data that we have talked about in this thesis. We reintroduce the Chain Condition below.

General Condition on A-chains (R&R: 696)

A maximal A-chain (α_1, …, α_n) contains exactly one link-α_1-that is both +R and Case-marked.

> where an A-chain is defined as one where there is a sequence of coindexation that is headed by an A-position and satisfies antecedent government.

The Chain Condition used by R&R will be adopted here[11]. Recall that +R just means that the element has referential independence. We can now easily account for the sentences in (88). In (88a), *taan* is –R as we have already seen from its inability to take any discourse referents. Thus in (88a), the chain consisting of *Maran* and *taan* is headed by *Maran*

[11] Our rejection of R&R's Reflexivity Framework does not conflict with our use of their version of the Chain Condition as these are independently derived principles.

which is the only element which is +R and case marked. Thus the sentence does not violate the Chain Condition. On the other hand, in (88b) the pronoun occurring in the object position is +R as well as case marked. If it formed an A-chain with *Maran*, then there would be two elements in the A-chain which would be +R and case marked. This is clearly not allowed and thus the sentence is ruled out. (88c) is also easily ruled out. In this sentence, the A-chain is headed by *taan* which is –R. Thus even though the A-chain only consists of one element which is +R and case marked, *Maran*, it does not head the chain and the sentence is correctly ruled out.

The set of data where *taan*/ *aven* occurs as a subject can also be easily accounted for as it falls out of the scope of the Chain Condition. This is because in such sentences, even if an embedded pronoun subject is coindexed with a matrix subject, they do not form an A-chain as they are arguments of different verbs and as such are not dictated by the Chain Condition. This is illustrated below.

89) a. [*taan* $_{*i/\ j/*k}$ *Maaran-ei*$_i$ *adi-tt-aan*] *enru*
taan.NOM Maran-ACC beat-past-3sgm comp
Somu$_j$ *Raman-idam*$_k$ *co-nn-aan*
Somu.NOM Raman-LOC say-past-3sg
Somu told Raman that he (Somu) beat Maran.

b. [*aven* $_{*i/\ j/\ k}$ *Maaran-ei*$_i$ *adi-tt-aan*] *enru*
3sgm.NOM Maran-ACC beat-past-3sgm comp
Somu$_j$ *Raman-idam*$_k$ *co-nn-aan*
Somu.NOM Raman-LOC say-past-3sg
Somu told Raman that he (Somu or someone else) beat Maran.

In (89a) *taan* cannot refer to *Maran* for the same reason why (88c) is bad. *Taan* is –R whereas *Maran* is +R and thus *taan* cannot be the head of the chain. Since *col-* 'say' licenses the discourse environment of a logophoric verb, *taan* can only refer to the source *Somu* and not *Raman*. In (89b) *aven* cannot refer to *Maran* as then there will be two +R and case marked elements in the A-chain. Since this is the only syntactic restriction on

how pronouns refer, *aven* can refer to both *Somu* and *Raman* in the matrix clause and is to be disambiguated by context.

The final set of Tamil data concerns the sentences with *ko(n)*. Consider the following.

90) a. *Maaren*$_i$ *tann-ei*$_{i/ *j}$ *aditi-**ko**-nd-aan*
Maran.NOM taan-ACC beat-**kon**-past-3sgm
'Maran beat himself'.

b. **taan*$_i$ *Maran-ei*$_i$ *aditi-**ko**-nd-aan*
taan.NOM Maran-ACC beat-**kon**-past-3sgm
'Maran beat himself'.

91) a. *Maaren*$_i$ *aven-ei*$_{i/ *j}$ *aditi-**ko**-nd-aan*
Maran.NOM 3sgm-ACC beat-**kon**-past-3sgm
'Maran beat himself'.

b. **aven*$_i$ *Maran-ei*$_i$ *aditi-**ko**-nd-aan*
3sgm.NOM Maran-ACC beat-**kon**-past-3sgm
'Maran beat himself'.

In (90a) *ko(n)* requires the subject-as-pivot reading and *taan* must be coindexed with *Maran* for this reading to obtain. Since *taan* is –R and forms an A-chain with *Maran*, this sentence is legitimate as there is only one element (*Maran*), the head, that is that both +R and case-marked. (90b) is ruled out as the head is now *taan*. On the face of it (91a) is problematic as *aven* should technically be a +R element. This means that there are two elements in the A-chain which are +R and case-marked contrary to the Chain Condition. However we have already mentioned earlier that in such sentences, *taan* and *aven* are unambiguously bound variables. This means that contrary to appearances, *aven* is actually –R in sentences such as (91a). There is independent evidence which tells us that *aven* in (91b) is actually –R.

Reinhart in her correspondence with Lidz claims that being +R is a syntactic notion and that referential independence is an entailment of being +R. If it is true that an

element which is +R is referentially independent, it also logically follows that an element which is not referentially independent cannot be +R. And showing that the pronoun in (91a) is not referentially independent is relatively easy. Contrast (91a) with (92) below.

92) *Maaren*$_i$ *aven-ei*$_{*i/j}$ *aditi-tt-aan*
Maran.NOM 3sgm-ACC beat-past-3sgm
'Maran beat him'.

In (92) there is no *ko(n)* on the verb and the pronoun can refer to some discourse entity which is indicative of its referential independence. The pronoun cannot refer to the subject as this would violate the Chain Condition. However when there is *ko(n)* on the verb as seen in (91a), *aven* cannot refer any such discourse referent which means that it is not +R anymore. Further evidence that *aven* in these sentences is actually –R comes from embedding (91b). If the pronoun in these sentences was in fact +R, we would expect it to take a matrix entity as its antecedent. We find this to not be the case.

93) *[*aven*$_i$ *Maran-ei* *aditi-**ko**-nd-aan*] *enru* *Somu*$_i$
3sgm.NOM Maran-ACC beat-**kon**-past-3sgm comp Somu.NOM
co-nn-aan
say-past-3sgm
For: 'Somu said that Maran beat him'.

In (93) if *aven* was +R, we would expect it to be able to refer to the matrix subject. However we find that the sentence is ungrammatical with this reading as well. This must mean that we are correct in claiming that *aven* in these sentences with *ko(n)* is not +R anymore but –R. Since the pronoun in (91a) is actually –R, it then does not represent a violation of the Chain condition anymore as there is only one +R element in the A-chain, *Maran* which is also the head. This must also mean that the reason why (91b) is ungrammatical is because the A-chain is headed by an element which is –R.

With that we have reconciled (91) with our account as well and explained the distribution of *aven* as well as the word order facts. This wraps up the thesis. We believe that we have provided a strong case for concluding that *taan* is indeed a logophoric pronoun in Tamil and that the distribution of *ko(n)* follows from an analysis along similar lines. We now move on to the thesis summary where we outline the basic findings of this thesis.

3.4 Thesis Summary

In Chapter 1 we briefly illustrated the basic objectives of this thesis. We showed that the characterization of *taan* in the literature is inconsistent and that much of what has been said about *taan* does not include recent developments in the theory of anaphora. We also provided a brief overview of the status of Tamil and its linguistic background. We also decided upon the high variety of Tamil as the version that was to be accounted for in this thesis. The simple reason for this is that the high variety of Tamil is more consistent among the versions of Tamil spoken in different parts of the world as the higher variety is also written. This means that the findings of this thesis can also be better generalized over the different versions of Tamil found in the world.

In Chapter 2 we illustrated the data that we aim to cover in this thesis. The data set was split into three main groups. They are 1) *taan* as an object with a nominative subject, 2) *taan* as an object with a dative subject and 3) *taan* as a subject. We also reviewed two influential syntactic approaches to anaphors in this chapter and showed that neither could account for *taan* satisfactorily. The first approach is SBT and we found that the Tamil data either followed Principle A or the prediction of complementarity that SBT makes but

not both. This was enough grounds to reject SBT as a viable explanation of the distribution of *taan*.

The other theory of anaphora that we reviewed was Reinhart & Reuland (1993)'s Reflexivity Framework and found that this explanation also falls short of accounting for all the Tamil data. While R&R could account for much of the data, their theory could not account for the instances where *taan* occurs as an object with a dative subject. However even those sentences which can be reconciled in their theory ran into its own problems as we found that contrary to their predictions, *ko(n)* is neither a necessary nor sufficient reason for reflexivization in Tamil.

In Chapter 3, we moved on to looking at the semantic and discourse properties of *taan*. In particular we looked at VP ellipsis data in Tamil and what it tells us about *taan*. We reviewed Heim & Kratzer (1998)'s account of VP ellipsis data through free and bound variables. We found that while it provided definitive mechanisms to provide a variable with values, simply understanding anaphors as bound variables and pronouns as free variables does not enable us to determine what *taan* is. This is because although *taan* behaves as a free and bound variable, there are constructions where it can only be a bound variable.

We then moved on to considering logophoricity as a potential way of explaining *taan*. We reviewed Sells (1987)'s account of logophoricity with the breakdown of this phenomenon into three primitives; source, self and pivot. After reviewing his theory, we applied it to Tamil and found that it had quite a lot of explanatory power. However we did have to modify his theory as we found that it had elements which we showed to be superfluous or poorly motivated. With our modified version of his theory, we were able

to account for all of the data. We showed that considering *taan* to be a logophoric pronoun enables us to explain its distribution when it occurs as a subject as well as when it occurs as an object. We showed that *taan* links to the source or self depending on the discourse environment licensed by the verb in the sentence. We also showed that *ko(n)* requires a sentence to have the subject-as-pivot reading in Tamil and that for this to obtain *taan* is neither necessary nor sufficient. This means that *aven* can also be possibly used in such sentences. This analysis also works as we are able to explain embedded clauses with and without *ko(n)*, the incompatibility of psych verbs and *ko(n)* as well as the VP ellipsis data that we grappled with earlier. We concluded the thesis with a section on the Chain Condition which explains the distribution of *aven* as well as certain word order facts.

Our analysis of *taan* as a logophoric pronoun and *ko(n)* as a sort of pivot marker does have a lot of explanatory power as we are able to account for all of the Tamil data that we have introduced in this thesis. However we have not attempted to generalize these findings to other Dravidian languages. This was not possible due to space considerations. We leave this task to future research but we believe that what we have outlined here is a solid basis from which to work.

References

Amritavalli R. 1984. Anaphorization in Dravidian. *CIEFL Working Papers 1.1*. 1-31.

Annamalai E**.** 1999. Lexical Anaphors and pronouns in Tamil. *Lexical pronouns and anaphors in selected South Asian languages: a principled typology*. eds. Lust B, Wali K, Gair J and Subburao K V. Berlin: Mouton De Gruyter.

Asher R.E. 1985. *Croom Helm Descriptive Grammars: Tamil*. Routeledge Publishers, London.

Belletti, A., and L. Rizzi (1988) "Psych Verbs and Θ-theory," *Natural Language and Linguistic Theory 6*, 291–352.

Britto F. 1987. *Diglossia: A study of the theory with application to Tamil.* Georgetown University Press, Washington.

Büring D. 2005. *Binding Theory*. Cambridge University Press: UK.

Caldwell R. 1956. *A Comparative Grammar of Dravidian Languages*. Madras. 3rd Edition.

Chierchia, Gennaro. 1989. Anaphora and attitudes de se. In *Semantics and contextual expression*, ed. Renate Bartsch, Johan van Benthem, and Peter van Emde Boas. Dordrecht: Foris.

Chomsky N. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky N. 1986. *Barriers*. Linguistic Inquiry Monograph 13, MIT Press, Cambridge, MA.

Clements G N. 1975. "The logophoric pronoun in Ewe: its role in discourse," *Journal of West African Languages* 10.2, 141-177.

Comrie, Bernard. 1983. Switch-reference in Huichol: A Typological Study. In *Switch-Reference and Universal Grammar*, eds. John Haiman and Pamela Munro. Amsterdam: John Benjamins.

Dalrymple Mary, Shieber Stuart and Pereira Fernando. 1991. Ellipsis and Higher-Order Unification. *Linguistics and Philosophy* 14: 399-452.

Emeneau M B. 1967. The South Dravidian Languages. *Journal of the American Oriental Society 87*. pp 365-413.

Everaert M, 1991, Contextual Determination of the anaphor/ pronominal distinction, in *Long Distance Anaphora,* Cambridge University Press, pp 77-118.

Ferguson C A. 1959. Diglossia. *Word 15*. pp 325-340.

Gleason, H. 1965. *Linguistics and English grammar*. New York: Holt, Rinehart and Winston.

Grimshaw Jane. 1990. *Argument Structure***.** MIT Press. Cambridge, MA.

Hagege C, 1974, Les pronoms logophoriques, *Bulletin de la Societe de Linguistique de Paris* 69: 287-310.

Heim Irene & Kratzer Angelica. 1998. *Semantics in Generative Grammer.* Oxford: Blackwell.

Hellan, Lars. 1988. *Anaphora in Norwegian and the theory of grammar*. Dordrecht: Foris.

Hellan, Lars. 1991. Containment and Correctedness Anaphors, in *Long Distance Anaphora,* eds. Koster & Reuland. Cambridge Uni Press.

Hestvik, A., 1992. LF movement of pronouns and antisubject orientation. *Linguistic Inquiry* 23, 557–594.

Huang J & Liu L. 2001. Logophoricity, Attitudes and ziji at the Interface. *Syntax and Semantics 33.* Academic Press.

Huang Y. 1994. *The Syntax and Pragmatics of Anaphora: A Study with special reference to Chinese.* Cambridge University Press. Cambridge.

Huang Y. 2000. *Anaphora.* Oxford University Press. New York.

Jespersen, Otto. 1933. *Essentials of English grammar*. London: Allen and Unwin (1983).

Kamp J A W. 1981. A Theory of Truth and Semantic Representation", in *Formal Methods in the Study of Language*. (eds) J. Groenendijk, Th. Janssen & M. Stokhof (eds.), Mathematisch Centrum, Amsterdam, pp. 277–322.

Keenan, Edward. 1987. On semantics and the binding theory. To appear in *Explaining language universals*, ed. John Hawkins. Oxford: Blackwell.

Kitagawa, Yoshihisa. 1991. Copying Identity. *Natural Language and Linguistic Theory* 9: 497-536.

Lehmann, T. 1989. *A Grammar of Modern Tamil*. Pondicherry: Pondicherry Publications.

Lidz J. 1995. Morphological Reflexive Marking: Evidence from Kannada. *Linguistic Inquiry* **26**, 705–710.

Lidz J. 2001. Condition R. *Linguistic Inquiry* **32**, 123-140.

Lidz J. 2001. The Argument Structure of Verbal Reflexives, *Natural language & Linguistic Theory 19:* 311-353.

Lust B. 1986. (ed) *Studies in the Acquisition of Anaphora.* Vol 2. Dordrecht: D. Reidel.

Masica C P. 1976. *Defining A Linguistic Area: South Asia*. University of Chicago Press. Chicago.

Mohanan K P. 1981. *Grammatical Relations and Anaphora in Malayalam*. M. Sc. Thesis, MIT.

Partee B & Bach E. 1981. *Truth, interpretation and information: Selected papers from the Third Amsterdam Colloquium*. ed. Groenendijk J, Janssen T, and Stokhof M, 1984. Dordrecht: Foris.

Pollard C & Sag L. 1994. *Head-Driven Phrase Structure Grammar*. Stanford: CSLI, and Chicago: The University of Chicago Press.

Pollard C & Sag L. 1992. Anaphors in English and the Scope of Binding Theory, *Linguistic Inquiry* 23: 261-303.

Reinhart T & Reuland E. 1993. Reflexivity. *Linguistic Inquiry* 24. pp 657–720.

Reinhart, T. 1976. *The Syntactic Domain of Anaphora*. Ph.D. Dissertation. Cambridge: MIT.

Rizzi L. 1990. On the Anaphor-Agreement Effect. Rivista di Linguistica 2: 27–42.

Safir Ken. 2004. The Syntax of Anaphora, Oxford University Press.

Sag, I. 1977. Deletion and *Logical Form*. Doctoral Dissertation. MIT, Cambridge, MA.

Sells Peter. 1987. Aspects of Logophoricity. *Linguistic Inquiry* 18: 445-479.

Stirling L. 1993. Switch Reference and Discourse Representation. In *Cambridge Study of Linguistics.* Cambridge University Press.

Thrainsson, Hoskulder. (1991). Long Distance Reflexives and the Typology of NPs, in *Long Distance Anaphora,* eds. Koster & Reuland. Cambridge Uni Press.

Yadurajan K S. 1987. The Reflexive Construction in Dravidian. *Studies in Dravidian and General Linguistics.* eds. Lakshmi Bai B and Ramakrishna Reddy B. Osmania University Publications in Linguistics 6. 233-244.

Zribi-Hertz, Anne. 1989. A-type binding and narrative point of view. *Language* 65:695-727.