Newcastle University

The Syntax of Pro-drop in Thai

On-Usa Phimsawat

B.A. Burapha 2000

M.A. NIDA 2003

Thesis submitted in partial fulfilment of the requirements of the regulations for the degree of Doctor of Philosophy

at

NEWCASTLE UNIVERSITY

SCHOOL OF ENGLISH LITERATURE, LANGUAGE, & LINGUISTICS

SEPTEMBER 2011

Table of Contents

Abstract

Thai is a discourse pro-drop language (also called 'radical pro-drop' language), since it exhibits highly frequent use of null pronouns without the involvement of agreement morphology. The descriptive goal of this thesis is to describe the different syntactic contexts in Thai where null pronouns occur and where they do not occur. The theoretical goal is to explain why pronouns sometimes have to be pronounced, sometimes may, but need not be, and sometimes cannot be pronounced, in Thai, and how this relates to pro-drop as found in other languages. The thesis will thereby hopefully contribute to the theory of pronouns, pronominal reference, and pro-drop.

After an introduction (Chapter 1), the distribution of null/overt pronouns and the restrictions on their occurrences are discussed (Chapter 2). It is demonstrated that even though a null pronoun can appear in any argument position, it cannot take the position of a prepositional complement or occur in a conjoined NPs construction. It is also demonstrated that a null pronoun/argument in Thai always looks for an antecedent for a referential reading (Chapter 5). If there is no controlling linguistic antecedent for a null pronoun in a higher clause, then a discourse topic will be the antecedent. If there is no discourse topic either, the speaker is always available as a local antecedent of the null pronoun. The default referential reading of a null pronoun is therefore first person singular. It appears that this generalisation holds true in other discourse pro-drop languages as well (Chapter 3). Correspondingly, a pronoun without an antecedent must be overt. Pronouns with a generic or arbitrary reading are a special case of antecedentless pronouns, which can be null but only when they have generic inclusive reading (Chapter 4).

I propose that null pronouns in Thai, and discourse pro-drop languages generally, have no φ-features, except for an unvalued referential feature [uR] and a general nominal feature [N]. These features are sufficient for the pronoun to function as an argument, being assigned a θ-role. They are dependent on being bound or controlled by *one* of the following: (i) a locally c-commanding referential NP, (ii) a null topic operator, which itself is linked to a referential NP in the discourse, (iii) 'the speaker', as default, (iv) a generic operator, or (v) a higher generic argument (Chapter 6).

That an inclusive generic pronoun is not pronounced (as opposed to other generic/ arbitrary pronouns, which are pronounced) is explained by the presence of a generic operator. The operator behaves just as an adverb that quantifies over the null arguments, i.e. 'It is *generally* true for *x*.' The [uR] feature of the pronoun is probed by this generic operator. As the pronoun has no φ-features, it gets a referentially unrestricted reading. This is the inclusive reading of a null generic pronoun. It includes the speaker, the

addressee, and any other people. In other words, the generality of the inclusive reading is from the ϕ-featureless pronoun, in which case it has no restricted reference. Since exclusive generic/arbitrary and quasi-inclusive generic pronouns have more restricted reference, i.e. third person plural and first person plural, respectively, they must be overt when bound by the generic operator. If they were null, they would be indistinguishable from the unrestricted inclusive generic pronoun and the referential first person pronoun 'I', since they have no antecedent providing them with features.

This means Thai can have relatively unrestricted use of referential third person/impersonal null subjects and a null inclusive generic subject pronoun. I show that this pattern is also found in other discourse pro-drop languages, and is restricted to languages where agreement is not part of sentential syntax. A null argument of the type [uR, N] in turn constitutes a new category in the typology of null arguments. It explains why it cannot function as a prepositional complement, which requires a complement with ϕ-features. This implies that referential null pronominal arguments do not inherit any ϕ-features from their antecedents. The only thing they inherit is the referential index. To summarise, null pronominal arguments do not have ϕ-features, and thus are not pronounced. Pronominal arguments with ϕ-features can be pronounced. In this, and in several other respects, null pronouns in Thai and other discourse pro-drop languages are similar to PRO in languages like English.

Acknowledgements

I have been looking forward to writing these acknowledgements, knowing that by now this thesis will have been completed and hoping that it will be in good enough shape not to need any major adjustments. It means that nothing else is left for me to do and the time is right for submission.

There have been several sources of inspiration that have encouraged me to complete this thesis. I am much indebted to my theoretical linguistics teacher, who also became the supervisor of this PhD thesis, Professor Anders Holmberg. He has done everything humanly possible to assist the completion of this work in a timely manner. Despite his heavy duties, he has always had the time for in-depth discussions with me, and the development of this thesis benefited significantly from his comments and suggestions at every single stage. His enthusiasm about syntactic research and his encyclopaedic knowledge of linguistics have greatly impressed me. In addition, he was also instrumental in helping me solve funding problems in the latter stages of my studies, for which I am extremely grateful.

I am also truly grateful to Dr William van der Wurff, who has spent time meticulously reading my chapters, providing insightful comments and constructive criticism on my work, and on whose support I still very much rely. I also would like to thank Professor Maggie Tallerman, my internal examiner, who suggested I should take her Cross-linguistic Syntax module in my first year. This module really inspired me to start my independent research in syntax. My special thanks go to Professor Noel Burton-Roberts, whose module was the first syntax module I took at Newcastle University. I also benefited greatly from auditing his modules in Semantics and Pragmatics and Grammar, Language and Mind. My thanks too to Dr Dimitra Kolliakou for her half year of supervision. Her passion for and knowledge of linguistics have always inspired me to aim for the best.

Special thanks go to several individuals who have contributed to my postgraduate studies in various ways, including Professor Karen Corrigan, Professor Vivian Cook, Dr SJ Hannahs, Dr Martha Young-Scholten, Dr Isabelle Buchstaller, Professor Manfred Pienemann, Dr Geoffrey Poole, Dr Clare Wright, Dr Anthony JR Waterston and Dr Elizabeth Waterston, Assistant Prof. Areeluck Harnmontree, Assistant Prof. Dr. Saksit Saengboon, Dr. Suchada Rattanawanitpun, Dr. Ubon Dhanesschaiyakupa, Orisa Yoo-Udomsook, Assistant Prof. Dr. Arnon Chaisuriya, Jim Siman, Thomas Edward Rushlow JR, Caroline Mason, the management of Burapha University who granted me a study leave and a partial scholarship, and my postgraduate classmates, especially Ali Algryani.

Finally, my deepest gratitude goes to my parents, my sister, my aunt, my cousin and my husband, without whose tremendous support and encouragement none of my academic goals would have been achieved. Their faith and belief in me, and my love for them, has allowed me to dedicate myself to my studies during my four years in the UK.

List of Tables and Figures

List of Abbreviations and Keywords

Agr	agreement
Adj	adjective
Adv	adverb
arb	arbitrary
A-movement	Argument movement
A-position	Argument position
ASP	aspect
A'-binding	A-bar binding
A'-movement	non-argument movement (A-bar)
A'-position	A-bar position (a null operator)
CLS	classifier
COMP, CP	complementizer phrase
compl	complement
CONJ	conjunction
D	definiteness feature
DIR	direction
DEM	demonstrative
e/ Ø	empty node
EPP	Extended Projection Principle
Fin	finite
FocP	focus phrase
FM	focus marker
FUT	future tense
GCR	Generalised Control Rule
GEN	genitive
gn	generic
G-pronouns	generic pronouns
iff	if and only if
LF	Logical Form
Lit.	literal translation
N	nominal feature
NEG	negative (negation particle)
NP	noun phrase

NSL	null subject language
OC	obligatory control
Op	operator
P	phonological feature
PASS	passive
PERF	perfective
PF	Phonetic Form
PL	plural
PRES CONT	present continuous tense
pro-drop	pronoun drop/ null pronoun
PRT	(pragmatic) particle
Q	question marker
ref	referential
R-expression	referential expression
SG	singular
Spec	specifier
SVCs	serial verb constructions
t	trace
TM	topic marker
TopP	Topic phrase
TP	tense phrase
U	utterance
uD	unvalued definiteness feature
uR	unvalued referential feature
uφ	unvalued phi-features
VP	verb phrase
vP	light verb phrase
θ-role	theta-role
φ-features	phi-features
(*he)	obligatorily null 'he'
*(he)	obligatorily overt 'he'
(he)	optionally null 'he'
#	pragmatically, semantically odd

agreement pro-drop
anaphor
binding theory
bound variable
c-command
control
discourse functions
discourse (radical) pro-drop
ellipsis
expletive
extraposition

focus
genericity
information structure
left dislocation
locality
null arguments
overt pronoun
pro-drop
referentiality
topic (prominence)
topicalization

Chapter 1: Introduction

The Null Subject Parameter, also called the Pro-drop Parameter, determines among other things whether or not a language requires pronunciation of subject pronouns. As discussed by Biberauer et al. (2010: i), it is "one of the best known and widely discussed examples of a parameter". It entails that there are two types of languages, pro-drop languages and non-pro-drop languages. Pro-drop languages, also known as null subject languages (NSLs) allow pronominal subjects to be null in finite clauses, as can be seen in languages like Italian. Non-pro-drop languages, also known as non-null subject languages (non-NSLs), forbid null subjects in finite clauses, as seen in English, for example. The standard view of the parameter, formally expressed in Rizzi (1982, 1986), is that pro-drop is dependent on rich agreement. In a non-pro-drop language like English, a subject pronoun cannot be omitted, since the agreement marking is too meagre to sufficiently determine its content. More recently the importance of rich agreement for pro-drop has been challenged (see Holmberg (2005), Frascarelli (2007), Modesto (2008), among others). Even more strikingly, discourse pro-drop languages (Chinese, Thai and others) exhibit highly frequent use of null pronouns where agreement marking is not involved at all. It was recognised very early, for example, in Rizzi (1986) and C.-T. J. Huang (1984), that the absence of agreement marking may play a role in accommodating the availability of null subjects in the discourse pro-drop languages. Chinese has been extensively studied (C.-T. J. Huang (1984) and subsequent; Y. Huang (1994, 2000); Yang (1994); Xu (1986, 2003); Qinan (2008), among others) but Thai has not yet received the attention it deserves (though see Aroonmanakun (1999)).[1]

This thesis will hopefully fill the gap by:

(i)   investigating the nature of discourse pro-drop in Thai and its properties, both with and without contexts and describing and identifying the different syntactic contexts where

---

[1] Aroonmanakun studied pro-drop in terms of null pronoun resolution for computer processing systems as in machine translation. His focus is on translating from pro-drop sentences in Thai into English. He assumes that the resolution of zero pronouns can be done at two levels: the sentence level and the discourse level. On the one hand, resolution at the sentence level can be implemented in accordance with the government and binding theory. On the other hand, zero pronouns that cannot be resolved by the government and binding theory must be resolved by discourse principles. The referents of the zero pronouns are expected to be the entity on which most attention is focused, or the backward centre of an utterance, according to Centering Theory (Grosz, Joshi & Weinstein 1995; Walker et al. 1998).

null pronouns occur, where they need not occur, and where they cannot occur.

(ii)   explaining why pronouns sometimes have to be pronounced, sometimes may, but need not be, and sometimes cannot be pronounced, in Thai, and how this relates to pro-drop as found in other languages, thereby contributing to the theory of pronouns, pronominal reference, and pro-drop, *and* positioning Thai in the typology of pro-drop languages, and more specifically, in that of discourse pro-drop languages.

1.1 Nature and Distribution of Overt and Null Pronouns in Thai

1.1.1 Overt Pronouns in Thai

In spite of the fact that the Thai pronouns are morphologically simple, and not inflected for case, the personal pronoun system of Thai appears complicated and laden with honorifics. This is because Thai grammar recognises several levels of respect, politeness, and honorification, all of which are reflected in its personal pronoun system. Table 1 illustrates common pronominal forms, although there are many other personal pronouns also in regular use.

Table 1: *The Thai pronouns paradigm*

First person					
SG (I, me)	Feminine	Masculine	Neutral	PL (we, us)	Level of speech[2]
chán	✓			phûak-chán	informal, spoken, confrontational
chǎn			✓	phûak-chǎn	neutral, spoken
phǒm		✓		phûak-phǒm	fairly formal, to older persons or superiors
khâphàchâw			✓	n/a	very formal mainly in writings, official ceremonies
kràphǒm		✓		n/a	very formal, humble, to superiors, old-fashioned
dìchǎn	✓			n/a	formal, to superiors
nǔu			✓	phûak-nǔu	neutral, humble, to older persons or superiors

---

[2] The levels of speech have been agreed by five native speakers of Thai, two of whom are linguists.

	Feminine	Masculine	Neutral		
raw			✓	phûak-raw	neutral, spoken
kháw	✓			phûak-kháw	informal, spoken
kuu			✓	phûak-kuu	very informal, offensive, confrontational

<div align="center">

Second person

</div>

SG (you)	Feminine	Masculine	Neutral	PL (you)	Level of speech
naay		✓		phûak-naay	informal, spoken
lòn	✓			phûak-lòn	informal, spoken, sarcastic
thân			✓	phûak-thân	formal, to superiors/ ones in a religious community, respect implied
thəə			✓	phûak-thəə	informal, spoken
muŋ			✓	phûak-muŋ	very informal, offensive
kɛɛ			✓	phûak-kɛɛ	very informal, spoken
nǔu			✓	phûak-nǔu	neutral, to younger persons

<div align="center">

Third person

</div>

SG (he, him, she, her, it)	Feminine	Masculine	Neutral	PL (they, them)	Level of speech
khǎw			✓	phûak-khǎw	informal
lòn	✓			phûak-lòn	informal, mainly in novels
man			✓	phûak-man	very informal, contempt implied
kɛɛ			✓	n/a	informal, to elderly persons, disrespect implied
thân			✓	n/a	formal, to superiors/ ones in a religious community, respect implied

According to the paradigm, pronominal forms in Thai vary according to gender and level of formality and politeness. In each social setting, several pronominal forms can be selected for use. For instance, both the first person pronouns for masculine, i.e. *phǒm* and *kràphǒm* 'I' can be selected for use in formal settings. Essentially, according to Iwasaki and Ingkaphirom (2009: 52), first and second pronouns form natural reciprocal pairs depending on the social

setting in which a conversation occurs. For instance, in a classroom setting, a female speaker may refer to herself as *chăn* 'I', and her listener(s) may refer to the speaker as *thəə* 'you'. Therefore, the reciprocal pronoun pair in this sense is comprised of *chăn* and *thəə*. Note that the third person *lòn* 'she' is generally used in fiction, as can be seen in the constructed example:

(1.1) 'khray yùu troŋnán' lòn tàkon phlaaŋ tòkcay
      who BE   there   she shout CONJ startle
      'Who's there?' She shouted, startled.

Furthermore, some pronouns are versatile in that one pronoun can be used interchangeably between the first, second and third persons. Moreover, pronoun selection is subject to social circumstances or attitudes towards the speaker, listener, and a third person who is referred to. For example, *thân* 'you/ (s)he' can be used to refer to the second and third person in a formal setting. Therefore, *thân* is a so-called versatile pronoun.

Thai pronouns can also have connotations attached to them. For instance, the second person *lòn* 'you' connotes sarcastic, hostile, and unfriendly elements. An example to illustrate the point is a satire on a colleague who was half an hour late for a meeting:

(1.2) mɛ̌ɛ   lòn maa tɔɔnníi maa pìd   kaanpràchum lə̌ə
      EXCLAMATION you come now   come close meeting   Q
      'Oh, you've come this time to close the meeting, haven't you?'

As can be seen, the pronominal system in Thai is complicated pragmatically, rather than morphologically. That is, distinctions for case are not found at all in the Thai pronominal system. For instance, *chăn* 'I' and *raw* 'we' remain the same whether they are the subject or object of the sentence. Furthermore, Thai does not differentiate between the attributive form 'my' and the predicative form 'mine'. To express possession, postmodification with *khɔ̌ɔŋ* 'of' in Thai is used instead. For instance, 'my friend/a friend of mine' would be literally 'a friend of I'.

Although there are a great many pronominal options in Thai, in spoken Thai, speech participants frequently drop pronominal arguments, where the reference of null pronouns is understood between or among interlocutors (in ways to be described in this work). Besides, proper nouns can be used to replace pronouns (Campbell (1969)). For instance, when addressing someone in a higher position or an honoured person or referring to them in the third person, common nouns representing certain social statuses are used instead, for example, *khunmɔ̌ɔ* 'doctor', *khunkhruu* 'teacher', *pɔ̂ɔ* 'father'. An example can be drawn from a commercial context where a customer uses a kinship term to address an elderly female fishmonger who is not related to him:

(1.3) Customer: wannîi *pâa* mii   plaadùk   máy

                       today   aunty have catfish   QM

                 'Have you got catfish today?'

1.1.2 Pro-drop in Thai

Like many other discourse pro-drop languages, Thai has widespread distribution of null pronouns used in actual discourse. Null pronouns can appear in almost all possible environments, that is as subject of a main clause, subject of an embedded clause, object of a main clause, and object of an embedded clause, but not as complement of a preposition. The examples below illustrate contrasting cases in Thai of overt pronoun and null pronoun. (1.4) and (1.5) exemplify an optional null pronominal subject in the main clause and the embedded clause respectively. The optional spell-out of (1.4) means that a null pronoun need not be pronounced, since it refers to the speech participant, who is either the speaker or the addressee. On the other hand, the use of an embedded third person pronominal subject in (1.5) is optional, due to the locally c-commanding antecedent. (1.6) exemplifies an obligatory null pronominal object coreferent with a topic. (1.7) exemplifies an obligatory overt pronominal object of a preposition. Where a pronominal argument has its antecedent in the preceding utterance, the use of the overt pronoun is acceptable only in cases of emphasis. Otherwise, an overt pronoun without an antecedent indicates the case of generic use (see Chapter 4).

(1.4) (chán) phêŋ sɔ̀ɔb sèd
       I   just   exam finish
    'I've just finished the exam.'

(1.5) kim₁   bɔ̀ɔk   wâa   (khǎw₁) dâay ŋaan lɛ̂ɛw
    Kim   say   that   (s)he   get   job PERF
    'Kim said that (s)he has just got a job.'

(1.6) [...] (chán) sòŋ   (*man)³ paylɛ̂ɛw mûawaan
          I   submit   it   PERF   yesterday
    'I submitted it [the report]⁴ yesterday.'

(1.7) naan   tâŋcay   rian   phɯ̂a *(khǎw)
    Nan   pay attention study for   him/her
    'Nan paid attention to her studies for him/ her.'

Null pronouns in Thai, according to Aroonmanakun (1999), can be used in four different ways: their usage can be deictic, anaphoric, generic, or discourse deictic. To begin with, the deictic use of a pronoun expresses nearness to the speaker or the addressee. Example (1.8) shows that the null pronoun directly refers to the addressee.

(1.8) wannîi   (thəə) tɛ̀ɛŋtua sǔay
    today   you   dress   beautiful
    'You are beautifully dressed today!'

Such a pronominal subject in this type of sentence is always optional. That is, it can optionally be null, since it refers to the addressee, as in (1.8). Furthermore, null deictic

---

³ The usual rules are not applicable when *man* 'it' is used to refer to a non-human animate entity, i.e. animals, where it cannot be omitted at all in finite clauses, for instance:

(i) thâa mǎa₁ cəə kràdùuk₂ *(man₁) cà   khíew (*man₂) pay   khrûŋ chûamooŋ
   if dog find bone   it   FUT chew it   PERF half   hour
   Lit. 'If a dog/ dogs find(s) a bone/ bones, it/ they will chew it/them in half an hour.'
  = 'If a dog/ dogs find(s) a bone/ bones, it/ they will finish chewing it in half an hour.

⁴ Here and throughout the thesis, items in square brackets are those which are absent in the source language but have been added to clarify the translation.

pronouns can be used to denote a third person only in cases where a referent is present or can be inferred in the current circumstance. Example (1.9) exemplifies a situation where the departure of a manager who has been supervising new staff very closely makes them feel relieved. The staff then talk behind the boss's back:

(1.9) (*kháw)   pay   sáthii!
　　　　 he　　　 go　　 PRT
　　　　 'He's gone!'

The second usage of null pronouns is anaphoric, as can be seen in example (1.5). To illustrate, in a sentence where absolutely no context is given, the embedded null subject is understood as being controlled by the higher argument, the matrix subject in the case of (1.5). Therefore, the interpretation of the embedded null subject relies on this.

Aroonmanakun's third kind of usage, the 'generic', refers to a null pronoun whose referent is unidentifiable or insignificant. He gives the following example (1.10):

(1.10) Ø pramaan wâa ratthàbaan fáràŋsèet cà   sáamàat rûabruam ŋen   càak
　　　　　 estimate that government French will can   gather   money from
　　　kaanprɛɛrûub   ráthàwísǎahàkìt [...]
　　　transformation   state-enterprise
　　　'It is estimated that the government would earn money from transforming state
　　　 enterprises [...].'   Aroonmanakun (1999: 19)

Although it seems that the null subject has no antecedent, it does not appear in the environment of a generic sentence. In fact, the sentence cannot be an instance of a generic usage, since the null subject does not refer to a quasi-universal set of individuals, roughly equivalent to people in general, everyone or anyone. To illustrate, the verb *pramaan* 'estimate' requires a thematic subject and the Thai sentence is not in the passive construction. One cannot start a conversation with such a sentence. In fact, out of the context, the sentence cannot be well-formed. This suggests that it is more like a simple active statement where a referential antecedent is understood and must be established in the preceding discourse than a sentence with an arbitrary subject in a passive. For these reasons, the term 'generic' is potentially confusing here; the translation is also misleading. Therefore,

the null subject in (1.10) should be treated as having a referential interpretation, contra Aroonmanakun.

Fourthly, null pronouns can be used in discourse deixis. Aroonmanakun states that a null pronoun can take an antecedent which is a non-NP constituent, for instance, a VP or a sentence. Example (1.11) illustrates a case of discourse deixis in which the null pronoun refers back to the prior portion of the discourse, i.e. to 'a serious matter':

(1.11) [mûawaannîi kə̀əd rûaŋ   khrîad]₁ dǐaw (chán) lâw (*man)₁ hây (thəə) faŋ

          yesterday   occur matter serious shortly I   tell   it   give you listen

          'Yesterday something serious occurred. I'll tell you about it shortly.'

The use of null pronouns in Thai appears preferable to most native Thai speakers in many cases, as they occur very frequently. Thus, accurate interpretation among interlocutors is vital. However, in Aroonmanakun's classification, some of the uses of Thai null pronouns appear to overlap, and the distinctions he makes can even be misleading. Take the case of anaphoric usage; he claims that it should cover not only the case of null pronouns with local linguistic antecedents, but also those with discourse antecedents. Given the fact that null pronouns observe Principle B, calling this the *anaphoric* use can be potentially misleading, as it may be taken to be a case of control of PRO. I shall refer to anaphoric usage of pronouns as referential use of null pronouns, including any referential null pronouns whose interpretations rely on definite antecedents which can either be a discourse or a close-enough (c-commanding) antecedent. With this in mind, the discourse deictic use should be subsumed under the referential uses of null pronouns.

There is another use of null pronouns where there is no thematic role, i.e. the expletive use. The expletive in Thai is not pronounced, and indeed I argue for the stronger claim that Thai does not have an expletive. The evidence comes from sentences with a so-called weather verb, in which case the subject must obligatorily be null (cf. 1.10):

(1.12) (*man) fǒntɔ̀k lɛ́ɛw

          it   rain   PERF

          'It has rained.'

The subject has neither semantic nor phonological content. Consider (1.13):

(1.13) (*man) thiâŋkhɯɯn lɛ̂ɛw　pay　nɔɔn
　　　　　it　　midnight　PERF　go　sleep
　　　　'It's midnight already. Go to sleep!'

(1.12) and (1.13) show that Thai need not have a pronounced subject for a well-formed sentence, and therefore Thai does not observe the EPP condition. There is an exception in the case of an extrapositive *it*,[5] where the subject can optionally be null. When overt, the expletive pronoun is *man*, equivalent to English 'it' whose properties are third person, singular, and gender-neutral:

(1.14) (man) cà　　dee　mâak thâa thəə　maa　troŋ weelaa
　　　　it　　FUT　nice　very　if　you　come on　time
　　　　'It would be very nice if you could come on time.'

As we have seen, the clausal subject is displaced to the end of the sentence, whilst the extrapositive *it* is in the canonical subject position.

I shall informally characterise the null pronouns illustrated so far as *pro*. The distribution of *pro* in Thai as illustrated so far can be summarised in the following table:

Table 2: *The distribution of pro-drop in Thai*

	1/ 2 *pro*	3 *pro*[6]	Expletive *pro*	Extrapositive *it*
Obligatorily null?	N	Y	Y	N

---

[5] Alternatively, the sentence-final clause is moved to the front of the sentence, eliminating *it*. The construction obtained is still grammatical in Thai, as seen in the following:
(i) thâa thəə　maa troŋ weelaa cà　　dee mâak
　　if　you　come on　time　FUT　nice very
　　'If you could come on time, it would be very nice.'

[6] That a referential third person *pro* is obligatorily null applies to cases where it is co-referent with a discourse topic, rather than a c-commanding antecedent in a higher clause, in which case can optionally be null.

Instances of *pro* shown in the table occupy any argument positions, whilst a null expletive and an extrapositive *it* occur in subject position. Deictic *pro* and referential *pro* have a full θ-role; the expletive *pro* and the extrapositive *it* are non-referential and have no θ-role, according to Rizzi (1986), cited in Ackema (2006: 12). I will argue in Chapter 6 that the deictic *pro* and referential *pro* when null do not have any ϕ-features, at least for Thai. It follows that the term *pro* will have to be abandoned in Chapter 6.

The relationship between the two linguistic elements, namely pronouns and their antecedents, is understood in this study as one of referentiality. In this thesis, my investigation is restricted to definite, referential pronouns in finite and non-finite contexts, and generic pronouns with the following properties: (i) a pronoun which refers to what its antecedent refers to (Lyons (1977), cited in Y. Huang (1994:15)), and (ii) a pronoun which has a c-commanding antecedent in a higher clause, or has a discourse antecedent, or none at all. With regard to referential pronouns, I shall concentrate on *null* referential pronouns in particular.

1.2 Overview of Thai Syntax

Thai is considered to be an analytic language, i.e. one in which inflectional morphology is not found in (pro)nouns or verbs. For this reason, verbs are marked for neither tense nor agreement. A sentence containing an active verb, without explicit time reference such as a time adverb, can be *vague*, although the interpretation of such a verb is in the past tense by default:

(1.16) (chăn) pay tàlàad

      I   go   market

    'I went/ go [habitually] / will go/ have been to the market.'

As an analytic language, Thai has whole, uninflected words, and the verb in (1.16) exemplifies this. When the sentence appears in isolation, four different temporal and aspectual interpretations are possible, according to Koenig & Muansuwan (2005). To resolve the vagueness due to lack of tense and agreement markings, the use of adverbials and discourse contexts is essential for retrieving tense, person, gender and number. Due to the fact that the verb itself is uninflected, aspect, mood and particles positioned pre-verbally and

clause finally are employed instead of inflections to express what in English is represented by tense. For this reason, I assume that verbs and other aspectual elements can be fitted into a dedicated tense phrase projection (TP).

(1.17) (chăn) pay tàlàad   *thùk wan*

      I   go   market   every day

     'I go to the market every day.'

(1.18) (chăn) *cà*   pay tàlàad

      I   FUT go   market

     'I'll go to the market.'

As can be seen, adding a temporal adjunct results in only one intended reading. Nonetheless, even without an adjunct, such vagueness may be diminished through a discourse context:

(1.19) A: (thəə) pay năy   maa

        you go   where PERF

      'Where've you been?'

    B: (chăn) pay tàlàad

       I   go   market

     'I've been to the market.'

The example indicates that Thai has a positive value for the discourse-orientation parameter (C.-T. J. Huang (1984)) as well as the pro-drop parameter. In terms of word order, Thai exemplifies the basic pattern of SVO, where word order appears to be rigid. Since Thai has neither agreement morphology nor case, scrambling is not allowed unless the object is topicalized. When the object is topicalized, it is moved to sentence initial position.

(1.20) năŋ₁   rûaŋ nîi (chăn) duu *t*₁ lɛ̂ɛw

     film   CLS this   I   see   PERF

     'This film, I have watched.'

The sentence in turn exemplifies a marked order where the direct object: *năŋ* 'film' is given a particular emphasis. In unmarked cases, not only the constituent order of SVO, but also

aspect, particles and discourse contexts play a crucial role in compensating for the lack of morphological inflection for number, gender, tense and agreement.

1.2.1 The head is initial

Thai is consistently right-branching, where the head (shown in italic) always precedes its modifier (an adverb (Adv)/ adjective (Adj)) and its complement (Compl):

(1.21) a. [ᵥₚ *tham* [COMP kaanbâan]]

       [ᵥₚ do   homework]

       'do home work'

    b. [ᵥₚ *tham* [Adv ciŋ]]

       [ᵥₚ do   actually]

       'actually do'

(1.22) a. [NP *phâakwíchaa* [COMP phaasǎa   tàwanʔɔ̀ɔk]]

       [NP department   language eastern]

       'Department of Eastern Languages'

    b. [NP *dèk* [Adj dii]]

       [NP child   good]

       'a good child'

(1.23) a. [AP *klua* [COMP phǐi]]

       [AP afraid   ghost]

       'afraid of ghosts'

    b. [AP *ruay* [Adv mâak]]

       [AP rich   very]

       'very rich'

(1.24) a. [PP *bon* [COMP lǎŋkhaa]

       [PP on   roof]

       'on the roof'

    b. [PP *bon* [Adv sùd]]

       [PP on   very end]

       'on top'

As has been seen, the head of every phrase precedes both its complement and its modifiers. Noun phrase structures in Thai can become complicated due to the presence of a numeral-classifier and/ or a demonstrative among their constituents, and such structures can be classified into two patterns. The following are examples of word order in NPs in Thai, more precisely, in numeral-classifier phrases, according to Juntanamalaga (1988):

Pattern One: Noun + Numeral/ Quantifier + Classifier

(1.25) khuàd sɔ̌ɔŋ bay

   bottle two CLS

   'two bottles'

(1.26) khon baaŋ khon

   person some CLS

    'some people'

Pattern Two: Noun + (Numeral) + (Classifier) + demonstrative

(1.27) rôm   sǎam khan níi

   umbrella three CLS this

    'these three umbrellas'

(1.28) rôm   làwníi

   umbrella these

   'these umbrellas'

Numeral and *Classifier* in Pattern Two are optional, since they merely function as a modifier. In Pattern One, both elements are required, since they function as a complement. Pattern One has an indefinite reading, whilst Pattern Two has a definite reading due to the presence of a demonstrative. All elements (if present) constitute a full DP. Nonetheless, the head nouns are obligatory and are usually initial. The classifier without a numeral is possible only when indicating a definite single entity, for example 'umbrella CLS this' which is equivalent to 'this umbrella'. Note that in spoken contexts, the classifier in Pattern One may be omitted and the number will precede the head noun, as can be seen from the following:

(1.29) baaŋ khon

   some person

    'some people'

1.2.2 Serial verb constructions

In Crystal's (2008) words, a serial verb is "a type of construction for a sequence of verbs or verb phrases within a clause (or a sequence of clauses) in which the syntactic relationship between the items is left unmarked. The verbs share a semantic argument, but there is no conjunction or inflection to mark co-ordination or subordination." In the literature on Thai serial verbs, the subject of the second/ third verbs in the sequence is analysed as PRO, thereby being anaphoric to the subject of the first serial verb (Thepkanjana (1986); Aroonmanakun (1999); Sudmuk (2005)). The example illustrates a case of the serial verb construction in Thai in which a null argument in the position of the second/third serial verb is considered as an empty category of the type PRO:

(1.30) cim   chɔ̂ɔb   PRO/ *cim pay PRO/ *cim ʔɔ̀ɔkamlaŋ

     Jim   like   go   exercise

     'Jim likes to go to exercise.'

As the complements of the transitive verb *chɔ̂ɔb* 'like', the second and the third verbs in the serial verb string – *pay* 'go' and *ʔɔ̀ɔkamlaŋ* 'exercise', are analysed as non-finite elements, since these serial verbs cannot have any modals or aspects appearing in between. All the verbs in the sequence, however, must share at least the same argument, namely *cim* 'Jim', and this shows that the relation between PRO and the antecedent appears to be one of obligatory control (more discussion of PRO in Thai can be found in Chapter 6, Section 6.3.1.3).

1.2.3 Topic prominence

Thai has a positive value for the topic prominence parameter (Modesto (2008)), leading to widespread pro-drop patterns. Such pro-drop can be found at subject and object positions, despite complete lack of morphological agreement:

(1.31) khuntaa$_1$ bɔ̀n   (wâa)   (thân$_1$) pùat lǎŋ

     grandfather complain COMP   he   ache   back

     'My grandfather complained of backache.'

Without an explicit context, *khuntaa* 'grandfather' is construed as the antecedent of the null embedded pronominal subject. On the other hand, if the context is provided, the context may dictate that a more distant NP or the discourse topic is the antecedent of the null embedded subject. In terms of sociolinguistic usage, the use of null pronominal form is preferred in informal settings, whereas the overt pronominal form is preferred in a more formal setting.

With regard to topic prominence in Thai, topics are generally marked through sentence initial position, with a demonstrative or an optional particle functioning like a topic marker, unlike in other topic-prominent languages, like Japanese, where a topic marker is obligatory. In the following example, the context is explicit, and the subject *ráan naaykɯ́ak* 'Naaykuak Restaurant' is the topic:

(1.32) ráan   naaykɯ́ak$_1$ (nía) khǎay sùkî$_2$ pen câwrɛ́ɛk thîinîi (*khǎw$_1$) khǎay
        restaurant naaykuak PRT sell suki BE first   here   they   sell
        (*man$_2$) dii   dûay ná
          it   good too   PRT
        (Lit.) 'Naaykuak Restaurant is the first sukiyaki restaurant there. They sell it well.'
         = ...They sell a lot of it (it = the sukiyaki).

The topic *ráan naaykɯ́ak* 'Naaykuak Restaurant' is made prominent, thereby resulting in the pro-drop in the subsequent utterance. Although subjects are often found as topics of sentences, subjects do not have to be topics. A grammatical object can be a topic through the use of special syntactic patterns such as topicalization, as seen in the following example:

(1.33) [phítsâa tîi   yùu bon   tó]$_1$ thom kin (*man$_1$)[7] pay lɛ́ɛw
         pizza   which BE   on   table Tom   eat   it   go PERF
        '[As for] the pizza which was on the table, Tom ate it.'

Once the topic is introduced into the discourse, *phítsâa tîi yùu bon tó* 'pizza which was on the table' is interpreted as the antecedent of the null object in the following sentence. It follows

---

[7] On the issue of obligatorily null object *man* 'it' apart from the fact that it is coreferent with the topic, see Chapter 5, Section 5.4

that the null object is obligatory (see Chapter 5 for obligatorily null pronunciation of referential pronouns):

(1.34) [phítsâa thîi yùu bon tó]₁ thom kin t_1 pay lɛ́ɛw (thəə) mày tôŋ   hǎa
         pizza which BE on table Tom eat   go PERF you NEG need search
       (*man₁) là
          it   PRT
      'As for the pizza which was on the table, Tom ate it. You don't need to look for it.'

1.2.4 Particles

Since Thai exhibits no inflectional morphology, particles are used extensively, mainly in spoken Thai to show tense, aspect, mood, etc. In addition, other significant functions of particles include signalling question types, politeness, commands, types of conversational or situational response, and addresser-addressee relationships (Iwasaki & Ingkaphirom (2009)). Some commonly used particles are illustrated in the table below, according to their semantic or pragmatic functions:

Table 3: *Commonly used particles in Thai*

Particles	**Functions**	**Examples**	**Meanings**
sí, sî, sì	commands	maa nîi *sí*	Come here!
	persuasion	lɔɔŋ thaan *e* duu *sì*	Try to eat (it).
ná	signalling questions/	*e* ʔûan khûn *ná*	You are getting fat, aren't you?
	commands/ requests	klàb maa *ná*	Come back, please.
	indicating desire for response	duu lɛɛ tuaʔeeŋ *ná*	Take care of yourself, OK?
máy	questions	chɔ̀ɔb *e* *máy*	Do you like it?
lá (lɛ́ɛw)	conclusion	pay sá *lá*	Gone!
	signalling complete activities	sèt *lá*	Finished!

As can be seen, lack of inflectional morphology is compensated for by means of particles. One particle form can have several variants in tone, each with a different (or perhaps similar) pragmatic function. The position of such particles is generally clause final. There is sometimes code-switching between Thai and English where Thai particles are attached clause finally. An example '*See you soon, na*' is drawn from a text message sent between Thais who speak English as a foreign language. The sentence not only reveals the code-switching between the two languages but also encodes an element of intimacy.

1.3 Outline of the Thesis

Chapter 2 concentrates on data relating to null pronouns in Thai. The data presented relate to null pronouns that appear controversial or problematic in the literature. Environments where pronouns are obligatorily overt or null will be highlighted so that generalisations can be drawn and then analysed in later chapters. The chapter also addresses the central theme of the thesis which will be a working hypothesis.

In Chapter 3, pro-drop in Thai will be discussed in terms of a Government and Binding theory approach, since it is informally characterised as falling within empty categories of the type *pro*. The characteristics of discourse pro-drop languages will be discussed first. Then, the distribution of referential *pro* in Thai at sentence and discourse levels will be elaborated in comparison with its overt counterpart. Big PRO will also be briefly discussed. Next, the occurrence of *pro* in Thai will be compared and contrasted with other discourse pro-drop languages, like Chinese and Korean, in the light of the literature on discourse pro-drop languages (C.-T.J. Huang (1984 and subsequent); Xu (1986, 2003); Tomioka (2003); Saito (2007); and Speas (2001)).

Chapter 4 is devoted to non-referential pronouns in Thai. That is to say, generic pronouns (G-pronouns) in Thai will be defined, classified, and discussed, according to their internal features/properties. Generic pronouns behave similarly to referential pronouns in that their realisation as overt or null is rule-governed. Generic pronouns are discussed in the light of a generalisation put forward by Holmberg (2010a,b) according to which there is complementary distribution between referential and generic null pronouns. An explanation is provided as to why the generalisation does not hold for Thai or other discourse pro-drop

languages. The chapter finishes with a discussion of indefinite pronouns, focusing on their function as topics.

In Chapter 5, attention is shifted to the theory of discourse functions in general, i.e. information structure and discourse functions (Vallduví (1990); Lambrecht (1994); Rizzi (1997), among others). The distinction between elements functioning as the topic and comment of utterances will be discussed, with particular attention to pronominal elements, both null and overt. The chapter then moves on to a discussion of referential null argument, interpreted as coreferent with a linguistic (c-commanding) antecedent or the discourse topic (Frascarelli & Hinterhölzl (2007); Frascarelli (2007); C.-T. J. Huang (1984 and subsequent)). Certain types of topics that correlate with the occurrence of obligatorily null pronouns will also be identified.

Chapter 6 is an analysis and discussion chapter in which null arguments, both referential and generic pronouns will be dissected and discussed in terms of their obligatory and optional spell-out. The analysis will concentrate primarily on whether a referential and generic null pronoun is bound/ controlled, and why a null pronoun is obligatory or optional. The analysis will reveal the internal structures and properties of pronominal arguments in Thai, which in turn may challenge the classic Government and Binding theory as well as the Generalised Control Rule proposed by C.-T.J. Huang (1984 and subsequent). Control of null arguments in finite and non-finite clauses will also be discussed. Lastly, the thesis ends with the proposition of a new typology of null nominal categories for Thai, that of syntactically projected null arguments, which differs from the old typology of null nominal categories in the literature based on Chomsky's (1982) theory of empty NPs. This is followed by a discussion about whether these categories are compatible with discourse pro-drop languages in general, and an attempt to place Thai within a typology of pro-drop languages.

Chapter 2: Typology of Thai Null Pronouns

In Chapter 1, a great range of available choices of personal pronouns in Thai have been described. There are many pronouns in Thai, but contrary to what one might expect, they are often dropped. That is to say, it appears that pronominal subjects and objects in a finite clause are dropped freely, in spite of the lack of agreement, and so a description and brief explanation of the distribution of pro-drop has been provided. The fact that a sentence with a missing pronominal argument means the same as one with an overt pronominal argument in non-pro-drop languages indicates that Thai has mechanisms for identifying the missing information. For this reason, C.-T. J. Huang (1984) classified Chinese, a discourse pro-drop language like Thai, as a 'cool' language.[8] One important question which arises regarding null pronouns is how the features of person, gender, and number of a syntactic gap are determined and whether the gap has internal structure. In this chapter, crucial data in relation to pronominal elements found in different types of sentence structures will be illustrated so that they can be examined and analysed in later chapters. Therefore, this is a data-oriented, descriptive chapter. An overview of the binding principles of Thai NPs will be presented, and the issue of null pronominal argument distribution will be revisited in more detail, in light of data that appear problematic in the literature and of the restrictions on the occurrence of a null pronoun. Obligatorily null and optionally null pronouns, as opposed to overt pronouns, will be identified.

2.1 R-expressions, Pronouns, Anaphors and the Binding Rules in Thai

With respect to the binding principles in Thai, nominal categories have a salient feature of R-expressions that can be used pronominally, referred to by Hoonchamlong (1991: 45) as *pronominal R-expressions*. Broadly speaking, these include names, titles, and definite NPs. It follows that they behave like pronouns in that they must be free within their minimal TPs,

---

[8] 'Cool' languages, according to C.-T. J. Huang (1984: 531), are those which require interlocutors to rely on contexts to recover certain anaphoric elements which are not overtly presented or heard in utterances. This differs from 'hot' languages, in which anaphoric elements are more explicit, since they cannot be deleted from grammatical sentences, as seen, for example, in English.

and may be bound outside their minimal TPs. As for the Thai pronouns and anaphors,[9] they conform to most of the binding principles. The distribution of R-expressions and pronouns in Thai is similar to that in English, except for anaphors, which can be bound outside their minimal TP domains, as in (2.1):

(2.1) pim$_1$ chûa   wâa   *(tuaeeŋ$_1$) rɔ́ɔŋphleeŋ phrɔ́
      pim   believe COMP   oneself   sing   beautifully
      'Pim believes that she sings beautifully.'

(2.2) (thəə) chɔ̌ɔb wɛ̌ɛn woŋ nîi   ríplàw
      you   like   ring CLS DEM Q
      'Do you like this ring?'

(2.3) [...] *(ceen$_1$) chɔ̌ɔb (*khǎw$_2$)[10] mâak
            Jane   like   (s)he   really
      '[...] Jane really likes him/ her.'

(2.4) *(ceen$_1$) rák   ceen$_{*1/\,2}$/ tuaʔeeŋ$_1$
      Jane   love Jane/   oneself
      'Jane loves herself.'

In (2.1), the θ-role that the verb *chûa* 'believe' assigns to the subject *Pim* is the role of Experiencer, in which case a reflexive pronoun can be used in Thai. Such a reflexive pronoun does not have its antecedent in the same minimal TP, but the sentence is still well-formed, provided that the reflexive pronoun is overt. This indicates that Thai does not abide by binding Principle A, especially in this context.

When these minimal TPs constitute independent clauses, the pro-drop distribution pattern is similar to the case where such minimal TPs are embedded in the matrix sentence. First person pronouns and second person pronouns can optionally be null in any finite clause, as in (2.2), unless they associate with focus readings. Third person pronouns must obligatorily be null,

---

[9] Only reflexives are discussed in this study, and although I recognise that there is another type of anaphor, namely reciprocals (i.e. 'each other'), they are beyond the scope of the study.

[10] The pronoun given in the brackets indicates that the context excludes the focus reading, so the pronoun must obligatorily be null.

given that each of them is coreferent with a discourse topic, as in (2.3). Otherwise, it is overt when receiving a phonological stress for the purpose of focus. Given a minimal TP, R-expressions cannot be repeated to have a reflexive reading in Thai, as in (2.4). If an R-expression is repeated, then the two occurrences are construed as representing different referents. That is, the use of an R-expression in place of a reflexive pronoun will result in its being disjoint in reference with the c-commanding antecedent in the same minimal TP. For an anaphoric reading, a reflexive pronoun is used, and this also applies in English.

However, Lasnik (1986), discussing Chomsky's Principle C in Thai, claimed that Thai does not obey the binding principle. He proposed a parameterised Principle C to account for Thai, as Thai is not subject to Principle C. He observed that in Thai, names can be bound by identical names. His example was (2.5):

(2.5) cɔɔn_1 chɔ̂ɔp cɔɔn_1
     John   like   John
     'John likes John.'   Lasnik (1986: 13)

(2.6) cɔɔn_1 chɔ̂ɔp tuaʔeeŋ_1
     John   like   oneself
     'John likes himself.'

In fact, the sentence in (2.5) yields a very unnatural reading (cf. example (2.4)). As Hoonchamlong points out R-expressions in Thai have a pronominal function, which rules out the object *cɔɔn* 'John' in (2.5) because it violates Principle B, in that it is bound locally in its minimal TP. Therefore, the R-expression in (2.5) should be replaced by a reflexive pronoun, as in (2.6). As R-expressions can be used pronominally, names in (2.7) and titles in (2.8) can be bound non-locally outside their minimal TPs. It appears that Thai violates Principle C in this context:

(2.7) kim_1 khít   wâa $(\text{kim}_1/\,\text{khǎw}_1)$ chɔ̌ɔb dam
     Kim   think   COMP   Kim/ she   like   Dam
     'Kim thinks that she likes Dam.'

(2.8) mɔ̌ɔ   kim₁ bɔ̀ɔk wâa   phrûŋníi   (mɔ̌ɔ   kim₁/ khǎw₁) mây   wâaŋ
      doctor Kim said COMP tomorrow   doctor Kim/ she   NEG   free
      'Doctor Kim said that she wouldn't be free tomorrow.'

The embedded null subject in (2.7) and (2.8) can optionally be null. This is due to the locally accessible antecedent in the higher clause.

2.2 Null Pronouns and their Antecedents in Thai

We have seen that Thai permits a null pronoun in subject and object positions. The types of their antecedents may range from c-commanding antecedents to discourse antecedents. This section illustrates the occurrence of null pronouns in Thai with regard to their antecedents.

2.2.1 C-commanding antecedents

A null subject or object pronoun in an embedded clause can have a c-commanding antecedent in a higher clause. The anaphoric relation in this case is an instance of control of a null pronominal argument into the finite clause. In Thai, there are two main types of control verbs taking subject control, namely *say/ expect*-type verbs[11] and *promise*-type verbs. The *promise*-type verbs are different from the former in the sense that the matrix subject of a *promise*-type verb never directly controls the embedded null object. (2.9), (2.10), (2.11) and (2.12) exemplify complements of *say/ expect*-type verbs. (2.13) exemplifies complements of *promise*-type verbs. All pronominal arguments can optionally be null:

---

[11] Thai appears to be similar to partial null-subject languages in term of control of null subjects into finite clauses. Holmberg et al. (2009) discuss the properties of Brazilian Portuguese, Finnish, and Marathi in which they allow a null subject in a restricted condition, i.e. in an environment where there is a controlling antecedent (including a c-commanding antecedent, too). The data in Thai indicates that it behaves like the partial null-subject languages in this respect as well:

(i) John₁ khush hota   karan   (tya-la)₁ pushkar bheti milyala
    John happy be-pst3sm because (he-acc) very   gifts receive-pst3plf
    'John was happy because he received many gifts.   Marathi (Holmberg et al. 2009:81)

(2.9) dam₁   bɔ̀ɔk   wâa   (khǎw₁) thùuk   huǎy
       Dam   say   COMP   he   win   lottery
       'Dam said that he had won the lottery.'

(2.10) dam₁   khâat   wâa   (khǎw₁) cà   thùuk   huǎy
        Dam   expect COMP   he   FUT   win   lottery
        'Dam expected that he would win the lottery.'

(2.11) lòn₁   khít   wâa   cim   chɔ̂ɔb   (lòn₁)
        she   think COMP   Jim   like   she
        'She thinks that Jim likes her.'

(2.12) lòn₁   khâat   wâa   cim   cà   chɔ̂ɔb   (lòn₁)
        she   expect COMP   Jim   FUT   like   she
        'She thinks that Jim will like her.'

(2.13) khǎw₁   sǎnyaa   wâa   (khǎw₁)   cà   maa   trɔŋwɛɛlaa
        (s)he   promise COMP   (s)he   FUT   come   on time
        '(S)he promised that (s)he would come on time.'

As stated above, the embedded argument in each sentence can be null, since it is controlled by a locally c-commanding linguistic antecedent in the higher clause. The object null argument can directly be controlled by the matrix subject, given that it is a complement of a *say/ expect*-type verb. However, subject control into complements of certain types of verbs may be impossible. (2.14) and (2.15) exemplify a null argument clause as the complement of *order*-type verbs:

(2.14) dam₁   sàŋ   (wâa) $e_{*1/2/3}$ tôŋ   thamŋaan   sèt   kɔ̀ɔn   thîaŋ
        Dam   order that   must work   finish before noon
        (Lit.) 'Dam ordered that (s)he/ I/ you must finish work before noon time.'
        = 'Dam ordered him/ her/ me/ you to finish...'

(2.15) dam₁ sàŋ kim₂ wâa $e_{*1/2}$ tôŋ thamŋaan sèt kɔ̀ɔn thîaŋ

Dam order Kim that   must work   finish before noon

(Lit.) 'Dam ordered Kim that she must finish work before noon time.'

= 'Dam ordered Kim to finish ...'

The presence of the matrix object in (2.15) resolves the ambiguity, as it can be seen that the interpretation of the null embedded object is restricted to 'Kim'. That is, in (2.15), object control, not subject control, is allowed into the complement of *sàŋ*, 'order'. The verb *sàŋ* 'order' in (2.14) requires a complement in which the embedded subject pronoun must refer to any referent which must not be *dam* 'Dam', the matrix subject. These examples show that whether subject control is possible depends on the type of verb involved.

In a case where there are several TPs, the chance is that the null subject would be understood as being bound by the nearest nominal argument, provided absolutely no context is given. A distant antecedent is judged as only marginally acceptable:

(2.16) khǎw₁ bɔ̀ɔk wâa thəə₂ bɔ̀ɔk wâa $e_{\#1/2}$ thùuk huǎy

(s)he   say   that you   say   that   win   lottery

'(S)he said that you said that you had won the lottery.'

2.2.2 Discourse (non c-commanding) antecedents

This class includes all other types of antecedents than c-commanding ones. Crucially, pronouns in this class are not directly controlled by antecedents. C-command is not presupposed. Rather, the relation of a null pronominal argument and its antecedent is achieved through coreference.   Discourse antecedents can be further classified into three major types, as seen below.

2.2.2.1 Antecedents in a preceding sentence

Typically, this long-distance antecedent is allowed in discourse-oriented languages. The sentence below is an example of the null subject referring to the topic of the discourse:

(2.17) kim₁ pay ŋaantàŋ (khɔ̌ɔŋ) níd₂ maa (*kháw₁) cǝǝ phûan kàw jɤ́jɛ́

     Kim go wedding of-GEN Nid PERF she meet friend old many

     'Kim went to Nid's wedding. She met many old friends.'

It appears that the interpretation of the null subject is dependent on the discourse properties of the topic constituent, namely *kim* 'Kim' in the opening discourse. Once the topic is established, null pronominal arguments are interpreted as coreferent with it. In this sense, the pronominal arguments are obligatorily null.

2.2.2.2 Understood antecedents

This subtype of antecedent is not present in the spoken discourse; it is understood in an utterance or situation. The antecedent type concerns particularly person deixis, i.e. the speaker and the addressee. Unlike third person pronouns, first and second person pronouns can optionally be null in any environment. Referential third person pronouns can and sometimes must be null, unless they carry stress for a focus reading.

(2.18) (khun₁) hǎay   pay nǎj   maa mûakhɯɯn (khun₁) màay klàb bâan

    you   disappear go where PERF last night you   NEG go home

    'Where have you been? You didn't come home last night.'

In the current discourse, the use of null pronouns referring to a third person or thing is possible despite no overt antecedent being found in such discourse. However, it is essential that the null pronoun has an antecedent established in the previous discourse, and the interlocutors need to have the same background knowledge; this precisely reflects what C.-T. J. Huang (1984) termed 'cool'-type languages. The example below illustrates a situation where B has been searching for a necklace, which A knows, and they meet a little later:

(2.19) A: (khun₁) hǎa (*man₂) cǝǝ yaŋ

       you   seek   it   find Q

       'Have you found it?'

    B: (chǎn₁) yaŋ hǎa (*man₂) màay cǝǝ (chan₁) màay rúu (chǎn₁) sày (*man₂)

       I   still search it   NEG find I   NEG know I   put   it

wây   thîinǎy

PERF   where

'I still can't find it. I don't know where I put it.'

As can be seen, discourse antecedents do not involve a c-commanding relation. Nonetheless, the null object is associated with a discourse antecedent, and thus is understood between the interlocutors.

It appears that the interpretation of null pronouns is straightforward, since it relies on syntactic information, such as a c-commanding antecedent, or the presence of a discourse topic. Note, however, that the overt form of the commonly used Thai pronoun *khǎw* '(s)he' (as seen in the gloss e.g. the one in (2.16), it could refer to either a male/ female referent without a sufficient context) merely encodes the features *third person*, (maybe) *singular*[12] but no gender. Thus, it too cannot unambiguously restrict the choice of the potential antecedent. As for a null pronoun, it does not even provide phonological clues to restrict the choice of referent, so interpretation of the null pronoun relies even more heavily on an appropriate context.

2.3 Types of Null Pronouns in Thai

In the last section, only one type of null pronouns which are referential has been discussed. As revealed in the data, there are, in fact, two main types of null pronouns found in Thai, i.e. a definite, referential null pronoun and an inclusive generic null pronoun, i.e. a generic null pronoun corresponding to generic *one* in English ('inclusive' because it includes the speaker and the addressee in its reference). The former is distinguished from the latter by the presence of an antecedent. That inclusive generic null pronouns should exist in Thai is surprising, given the generalisation that pronouns can be null only if they have an antecedent providing interpretation for them. With regard to the generalisation of antecedents in Thai, the definite, referential pronouns can be further divided into two sub-types, based on their antecedents, as discussed in Section 2.2. Firstly, a null pronoun may have its antecedent in the next clause up

---

[12] The pronoun *khǎw* can be an abbreviated version of the third person plural (*phûak*)-*khǎw*. In addition, it is a gender-neutral pronoun. It can thus be equivalent to any of the English third person pronouns: *he, she, they, him, her,* and *them*, depending on the context.

(see Section 2.2.1). Secondly, a null pronoun can be coreferent with a discourse antecedent, which may, for instance, be in a preceding sentence:

(2.20) lûuksăaw$_1$ (khɔ̌ɔŋ) ceen$_2$ kèŋ   khamnuan khruu$_3$ bɔ̀ɔk wâa   (*khăw$_1$)
       daughter of-GEN Jane good at calculation teacher say   COMP   she
       sɔ̀ɔb lêek   dây khánɛɛn súuŋsùd
        exam maths get mark   highest
        (Lit.) 'Jane's daughter is good at calculations. The teacher said that she got top marks
        for maths.'

The context in (2.20) is explicit that *lûuksăaw khɔ̌ɔŋ ceen* 'Jane's daughter' is the topic of the discourse. When the topic is established in the discourse, the interpretation of a null pronoun in the following/ subsequent sentences is dependent on such a topic, unless a new topic is introduced. In this situation, the pronoun can and must be null.

According to C.-T. J. Huang (1984), the topic of a sentence can be null when it is identified with the topic of the preceding sentence. This is a so-called *topic chain*, where null topics and the original referents are connected:

(2.21) [ród khan nán$_1$] (*man$_1$) mây suǎy   (*man$_1$) kinnáamman (*man$_1$) rakha kôo
        car CLS DEM   it NEG beautiful   it   consume oil   it   price also
        phɛɛŋ (chǎn) mây síi   (*man$_1$) nɛ̂ɛ
        expensive I   NEG buy   it   sure
        'That car, it's not beautiful; it consumes a lot of petrol; it's also expensive. Surely,
        I'm not going to buy it.'

The example illustrates an instance of a topic chain. Thai appears to allow topic deletion if the topic is identical to the overt topic established in the opening discourse. That is to say, the overt topic, namely *ród khan nán* 'that car' is found in the chain initial position, followed by three clauses whose topics are missing, and such clauses, therefore, represent links in a topic chain. However, further data suggest that null subjects are not necessarily coreferent with the discourse topic:

(2.22) U1 dam (na) khàb ród pay ráb   kim thîi thamŋaan (*khɔ̌ɔŋ thəə) thúkwan

           Dam TM drive car go pick up Kim at office   of-GEN she everyday

     U2 (*khunthəə) sàbaay   cinŋcinŋ

               she   fortunate really

   '[As for] Dam, he drives to pick up Kim at her office every day. She's so fortunate.'

The example shows that a null subject need not be correlated with the topic[13] in Thai, given that it is coreferential with an antecedent which is already established in the discourse. Note that possessors can optionally be null in Thai, given that they (the possessors) are referring to the speech participants: either the speaker or the addressee. On the other hand, the third person possessor referring to the topic of the discourse is obligatorily null. In terms of discourse functions in Thai, some discourse requires obligatory overt pronouns, for instance, when the pronoun receives stress for the purpose of focus.

Referential null arguments (of the non-controlled type) which are not coreferent with the discourse topic appear to be a marginal phenomenon. In most cases, such null arguments appear to be interpreted as coreferent with the topic discourse antecedent. Consider this example:

(2.23) [ŋantɛ̀ɛŋŋan khɔ̌ɔŋ   pim]$_1$ càd   sǔay   mâak thúkkhon chɔ̂ɔb (*man$_1$)

         wedding   of-GEN Pim organise beautiful very everybody enjoy   it

   'Pim's wedding was very beautiful. Everybody enjoyed it.'

That a null pronoun is obligatory in this context is in line with Iatridou & Embick (1997). They demonstrate that a null pronoun, or *pro*, must be used to refer to states of affairs (SOAs) when the linguistic entity through which the SOA is expressed is an NP. In this example, the SOA is referred to by the NP, namely *ŋantɛ̀ɛŋŋan khɔ̌ŋ pim* 'Pim's wedding', and so *pro* has to be used. Having said that, in Thai *pro*/null pronoun can also take a C/TP as its antecedent, as sentences having such a relationship of coreference appear to be accepted as grammatical.

---

[13] In fact there are different kinds of topics, and distinctions between them are crucial to the proper characterisation of antecedence for null pronouns. These matters will be discussed in detail in Chapter 5.

(2.24) [mûawaannîi naay₁ rîak (chǎn₂) pay phób]₃ dǐaw (chǎn₂) lâw (*man₃) hây (khun₄)

yesterday   boss call I   go see   shortly I   tell   it   give you

faŋ

listen

'Yesterday the boss called me in to see her. I'll tell you about it shortly'

Unlike the data in (2.23), the null object in (2.24) does not take the NP, namely 'boss', as the antecedent. Rather, the whole TP where the SOA is fully expressed is the linguistic antecedent of the null direct object. In other words, such a null object in (2.24) stands for a category other than NP, and is assumed to refer to the whole of the preceding sentence. This pronoun is obligatorily null. Only the deictic null pronouns/ *pro*s referring to the speech participants are optional.

As for impersonal pronouns, most pronouns having generic or arbitrary readings have to be overt in Thai, for instance:

(2.25) *(kháw) bɔ̀ɔk wâa   khâŋen dɔllâa cà khɛ̌ɛŋtua khûnʔìik

they   say   COMP currency dollar FUT strengthen further

'They said that the dollar would further be strengthened.'

(2.26) *(raw) tɔ̂ŋ   kin

we   have to eat

'We have to eat.'

(2.25) has an arbitrary reading; (2.26) has a generic quasi-inclusive reading. Because their reference is not provided with the help of an antecedent, they have to be overt. However, there is another type of generic pronoun, whose reference is also not derived from an antecedent, but which nevertheless can be, and even must be null. This is the inclusive generic pronoun corresponding to English *one* or generic *you*:

(2.27) kruu   sɔ̌ɔn wâa   Ø cong   pen khon   dii

teacher teach COMP   should BE person good

'Teachers teach that you/ ones should be the good.'

Discourse pro-drop languages, like Thai and Chinese,[14] do have a null pronoun with a generic inclusive reading. This will be discussed in more detail in Chapter 4. Types of null pronouns found so far are summarised in Table 4.

Table 4: *Different types of null pronouns in Thai*

	Deictic	**Referential**	**Expletive**	**Inclusive generic**
Overt	[±]	[±]	[−]	[−]
	[1/ 2]	[3]	[3]	
Understood features	[± singular]	[± singular]	[+singular]	[+plural]

2.4 Restrictions on the Occurrence of Null Pronouns in Thai

Although pronominal arguments can be omitted in Thai quite freely, there are some positions in which pronouns cannot be null at all. Such positions include the complement to a preposition and when the null pronoun is part of a conjoined NP. (2.28) exemplifies a case of a prepositional complement; (2.29) and (2.30) exemplify object and subject arguments in conjoined NPs respectively:

(2.28) cɔɔn yùu kàp *(*chán*) talɔ̀ɔdweelaa

John stay with   I   all the time

'John stays with me all the time.'

(2.29) nók₁ bɔ̀ɔk wâa   pim₃ hěn tháŋ *(*khǎw₁/₂*) lɛ́   ceen₄

Nok say COMP Pim see both   (s)he   and Jane

(Lit.) 'Nok said that Pim saw both him/ her and Jane.'

(2.30) nǔu₁ khit waa   *(*nǔu₁*) kàp nɔ́ɔŋsǎaw₂   ca   pay hɔ̂ŋsàmùd

I   think COMP   I   and younger sister FUT go   library

'I think that I and my younger sister will go to the library.'

---

[14] In Chinese, an arbitrary (generic inclusive) reading of a null pronoun can also be constructed:
  (i) fuqin   shuo *e* yao   weirenzhengzhi
      father   say   should upright
      'Father says that (one) should be upright.'   (Chinese: Y. Huang 1994: 36)

That the null pronouns are not permitted in such environments is problematic, given that Thai has a positive value for the pro-drop parameter. Such constraints on the occurrence of null pronouns do not exist in all other discourse pro-drop. (2.31) exemplifies the occurrence of a null pronominal object in a conjoined NP in Imbabura Quechua:

(3.31) Juan$_1$ yayan chay runa (*pay-ta$_1$*) Marai-wan rikushka-ta

     Juan think that man he-acc Mary-and saw

     'Juan thinks that man saw him and Mary.'   Imbabura Quechua: Cole (1987: 602)

Under the coordinate NP structure, the object pronoun is optionally unpronounced in Imbabura Quechua. Cole (1987) argues that the null object which appears in an island structure should be *pro*. Since it cannot move out of the island, it cannot be an instance of a topic variable or a trace of any kind. Nonetheless, Thai does not behave like Imbabura Quechua. Postulating a null object as being of the type *pro* is, therefore, not sufficient to explain why such a position cannot be null in Thai. If such a sentence is translated into Thai, a pronominal object must be overt. The fact that it must obligatorily be overt leads to a hypothesis that there are some syntactic constraints on null pronouns in such positions. Another hypothesis regarding the conjoined NPs is that one pronoun cannot be null because a null pronoun is too different in terms of feature content to be conjoined with an overt NP (see Chapter 6, Section 6.5 for the analysis of a pronoun as a prepositional complement and as part of a coordinate NP structure).

The data so far suggest that apart from the above syntactic constraints, Thai has unrestricted use of referential null pronominal arguments and inclusive generic null subject pronouns, despite non-involvement of agreement morphology. Also, the features encoded on the pronominal overt form in Thai are language-specific. For example, the most commonly used Thai third person pronoun, namely *khǎw* '(s)he' does not encode gender. Nonetheless, phonologically null pronominal forms are available in the grammar, i.e. Thai has referential null subjects and objects, provided there are antecedents available, and also has inclusive generic null pronouns. Major types of pro-drop in Thai have the following properties, based on Cabredo Hofherr's (2006) pronominal paradigm: the pronouns of first and second person singular and plural are deictic, containing reference to the speaker and the hearer; the

distinction between anaphoric and non-anaphoric pronouns or generic pronouns applies largely to third person pronouns. This is summarised as follows:

Major types of pro-drop in Thai:

(1) Deictic null pronouns marked [+speaker]/ [+hearer]; the null pronouns are optional.

(2) Referential null pronouns: definite third person null pronouns which take a discourse referent previously introduced in the discourse or a higher c-commanding NP as their antecedent; the null pronouns are obligatory and optional respectively.

(3) Expletive: third person null pronouns, being non-thematic, do not take up a discourse referent previously introduced in the discourse; expletive *it* is obligatorily null.

(4) Inclusive generic null pronouns: null pronouns which do not take up a discourse referent previously introduced in the discourse; the pronouns are obligatory null.

It appears that the use of null pronouns is the preferred option as long as no constraints apply. This leads to a working hypothesis that null pronouns in Thai are preferred, and that overt pronouns are used when there are restrictions on the realisation of null pronouns. Such restrictions include syntactic constraints and some discourse functional constraints.

2.5 Summary

Although Thai is a discourse pro-drop language, grammatical operations and principles are different from those in other pro-drop languages, and limited to the parameterised aspect of Thai grammar. As can be seen, Thai allows both third person referential null pronouns and inclusive generic null pronouns. However, it is not the case that these can appear in any argument positions. A null pronoun is not syntactically permitted in a prepositional complement or coordinate NP structures. When a subject pronoun is null, an object may be either null or overt. In other words, object *pro*s are independent from subject *pro*s and vice versa. There is also a significant correlation between a discourse topic and the occurrence of obligatorily null pronouns in Thai.

On the basis of the discussion thus far, it is clear that there are three major issues to explore in the next chapters: (i) the root causes of optionality and obligatoriness as regards the pronunciation of pronouns, (iii) definite, referential and non-referential (generic) pronouns, both null and overt, and (iii) the internal structure of null pronominal arguments in Thai.

Chapter 3: Empty Categories in Discourse Pro-drop Languages

Chomsky's (1981 and subsequent) principles of Government and Binding Theory (GB) that are relevant for the treatment of empty categories (ECs) include Government, Binding, and Case. Government is concerned with the relation between the head of a construction and a category dependent on it; Binding deals with the set of conditions on the co-occurrence of different types of NPs, both overt and null; Case deals with the assignment of abstract Case and its morphological realization (Chomsky 1981: 5). In a wide range of studies of discourse pro-drop languages, it appears that Chinese (C.-T.J. Huang (1984 and subsequent); Xu (1986, 2003); Y. Huang (1994, 2000); Modesto (2008, among others)) has received greater attention than other discourse pro-drop languages, for instance, Thai (Aroonmanakun (1999); Hoonchamlong (1991); Speas (2001)), Korean (Choo (1994); Han (2003)), and Japanese (Saito (2007); Tomioka (2003)).

In this chapter, the ECs in GB Theory will be illustrated with particular attention to Binding Theory. Next, the theory of discourse pronominal anaphora proposed by C.-T. J. Huang (1984, 1989, 1991) will be the principal focus of my critique; arguments will be presented both in support of and against his theory, since his analysis concerning variables bound by a topic operator will be the main framework for my analysis of referential null pronominal arguments in Chapter 6. Different proposals made by Xu (1986, 2003), Tomioka (2003), Saito (2007) and Speas (2001) in respect of null arguments in discourse pro-drop languages will also be reviewed and discussed. The occurrence of *pro* in Thai will be compared and contrasted with other discourse pro-drop languages, namely Chinese, Korean, and Japanese.

3.1 Empty Categories (ECs)

In the GB framework of syntax (Chomsky (1981, 1982, 1986)), a typology of NPs is proposed which comprises overt (lexically realised) and empty (i.e. null/ covert) NPs. According to this typology, nominal empty categories and their overt counterparts can be interpreted as a combination of two binary features as follows:

Table 5: *Chomsky's typology of NPs*

	Overt	**Empty**	**Overt**	**Empty**
+ anaphoric	_	PRO	Anaphors	NP-trace
– anaphoric	**Pronouns**	*pro*	R-expression	variable
	+pronominal		**– pronominal**	

The four types of empty categories, namely NP-*trace*, *pro*, PRO, and variable fall under different principles with respect to antecedent assignment. The following examples illustrate the contrasting cases of overt and empty NPs:

PRO

(3.1) John$_1$ would like PRO$_1$ to join the club.

(3.2) John$_1$ would like us$_2$ PRO$_{2/*1}$ to join the club.

Pronoun and *pro* (Thai)

(3.3) John$_1$ thought that he$_1$ had been a fool.

(3.4) phûak-khǎw$_1$ bɔ̀ɔk wâa   *pro*$_1$ dâay ŋaan lɛ́ɛw

         they   say COMP   get job PERF

         'They said that they have got a job.'

Anaphor and NP-*t*race

(3.5) The students$_1$ cook for themselves$_1$/ each other$_1$.

(3.6) The children$_2$ seem t_2 to enjoy the show.

R-expression and variable/ wh-*t*race

(3.7) *He$_1$ loves John$_1$.

(3.8) Who$_2$ did Jane see t_2?

In terms of ECs, NP-trace, *pro*, and variable are subject to Binding Principles A, B, and C respectively, whilst PRO complies with separate principles of the theory of control. The theory of GB puts an emphasis on the referential properties of the nominal and covert elements. Based on the referential properties, the Binding Principles were formulated as follows.

3.2 The Binding Principles

Principle A: an anaphor must be bound in its governing category where

> An element α is bound by β if β c-commands α, and α and β are coindexed;
>
> and the governing category of α is the minimal XP containing α, a governor,
>
> and a subject accessible to α.

Principle B: A pronoun must be free in its governing category where free means not bound.

Principle C: R-expression must be free (everywhere).

Following the Binding Principles, the identity of an EC is determined on the grounds of its relation to an antecedent. Chomsky (1981: 330) proposed that the identity of an EC is functionally determined, according to the following principles:

a. An EC is a pronominal iff it is free or locally bound by an element with an independent thematic role, and a non-pronominal otherwise.

b. A non-pronominal EC is an anaphor iff it is locally A-bound, and a variable if locally A'-bound.

3.2.1 Binding of ECs

NP-trace

NP-trace is an instance of EC that is left behind by NP-movement to an A-position, according to Chomsky (1981). In (3.6) the subject *the children* is base-generated as the subject of *enjoy* where the θ-role is assigned, and moves to the subject argument position of *seem* to get Nominative case. Put another way, the NP trace and its antecedent *the children* form a single chain, i.e. the NP-trace is in A-position (subject of infinitive) and the nearest controller is *the children*, which c-commands the trace, and the trace is coindexed with it. Therefore, the NP trace is A-bound, and it is a null counterpart of an anaphor.

PRO

From the binding theoretical perspective, PRO is expected to satisfy both binding conditions A and B. That is to say, PRO is anaphoric in the sense that it is dependent on another NP for its interpretation, thus observing Principle A, as in (3.1) and (3.2). Correspondingly, it is pronominal when it is taken to refer to a specific referent, for instance PRO refers to the referent *John* in (3.1). It is nonetheless, neither purely pronominal nor anaphoric, as it cannot appear in a governed position. From this conflicting requirement, the PRO theorem is derived:

The PRO theorem:

PRO is ungoverned.

Given the theorem, PRO cannot have a governor. The controller of PRO may be a subject (3.1) or an object (3.2), or it can have a generic (arbitrary) interpretation. In (3.1) *John* and PRO are coindexed. On the one hand, *John* is in an A-position and has a θ-role (i.e. experiencer of *would like*) which is independent from that of PRO. As for PRO, it appears in a non Case-marked position, as the subject of the infinitival clause (the agent of *join*). Thus, the two arguments belong to separate chains. On the other hand, the reference of PRO is determined by the coindexed antecedent *John*, which thus controls PRO. In (3.2), PRO is controlled by the complement of the finite verb, namely *us*.

Pro

Since *pro* is found in a governed position (Spec, TP), it can be replaced with an overt pronoun in some contexts, unlike PRO, which is ungoverned and must obligatorily be null in all cases. For instance, in (3.4), *pro* (like any overt pronominal) is A-bound by the matrix subject, namely *phûak-khǎw* 'they'. Therefore, *pro* is considered as a null counterpart of a lexical pronoun. Some discourse pro-drop languages are claimed to allow subject *pro*s, but not object *pro*s, due to certain restrictions on the binding requirements (C.-T. J. Huang (1984 and subsequent)). The issue of asymmetries between them will be discussed in detail in the next section.

Variables

A variable or a wh-trace is an EC left behind by the movement of a wh-expression, whose landing site is Spec, CP, in an A'-position, unlike NP-trace. Such a movement creates an A'-chain where a variable is A'-bound. The trace is considered as a phonetically null R-expression. Based on the property of R-expression, a variable is free everywhere, in the sense 'A-free'. Therefore, in (3.8), the trace of *who* cannot be A-bound by *Jane*, but can be A'-bound by *who*, which is not an argument, but an operator in Spec, CP.

In Chomsky's minimalist framework (1995, 2001), the theory of empty categories is no longer an object of interest. Instead, the copy theory of movement is adopted. For instance, wh-traces are construed as deleted copies of A'-movement, not as –anaphoric and – pronominal empty categories, i.e. variables, as in the previous GB framework. NP-traces are construed as deleted copies of A-movement, not as +anaphoric and –pronominal empty categories, i.e. traditional anaphors. Therefore, the classification of empty categories on the basis of ± pronominal/ ± anaphor features is invalidated. Example (3.9) involves a trace as a copy of a moved element that is deleted in the phonological component:

(3.9) Which picture of himself$_{1/*2}$ did John$_1$ like?

The reflexive pronoun *himself* can only have *John* as its antecedent due to Principle A. This is accounted for if, in the derivation of the sentence, the wh-phrase, namely *which picture of himself* moves to Spec, CP, while leaving a copy behind. The reflexive pronoun is then bound by *John* by virtue of the copy. The structure after movement is therefore the following:

(3.10) [$_{CP}$ which picture of himself [$_C$ did [$_{TP}$ John like [~~which picture of himself~~]]]]

Wh-movement is triggered by the need to check a [-interpretable] wh-feature in CP-domain. In early Chomskyan minimalist feature theory (Chomsky 1995), [-interpretable] features have to be checked and deleted before LF; only [+interpretable] features are allowed at LF. In versions prior to Chomsky (2000), the movement effected the checking. In later versions, the checking is does at a distance (by Agree), while the movement is triggered by an EPP-feature. The lower copy is deleted in the derivation of PF.

3.3 Null Arguments in Discourse Pro-drop Languages

It has been recognised since the earliest studies, for example in Rizzi (1986), that discourse pro-drop languages allow empty pronouns from finite clauses in spite of there being no inflectional system in the grammar, and that therefore such pro-drop is not dependent upon agreement (Agr). Rizzi's 'pro module', which states that subject pros are identified through verbal inflections, does not account for the Chinese pro, which is found to be in the subject position of a finite clause. As Rizzi argued, since there is no Agr specification in Chinese, there is no need for an AgrP projection, and Agr is not involved in the licensing of null arguments. Thus, these pro-drop languages are exempted from the licensing and identification conditions he proposed. This led C.-T. J. Huang (1984 and subsequent) to develop an alternative pro-drop principle. *Pro*s in discourse pro-drop languages are identified by discourse, with the help of topic-prominent structure. That is, the interpretation of *pro* relies on a topic of the discourse. Thereby, in discourse pro-drop languages *pro* can be exhaustively identified by an argument NP outside the sentence but within the discourse. In fact, recent research has shown that the idea that null arguments are ever identified by Agr, in Rizzi's (1986) sense, is mistaken. Instead, referential null arguments (most clearly in the case of third person arguments) are always identified, i.e. assigned reference, by an antecedent NP; see Holmberg (2005) among others.[15]

3.3.1 C.-T. J. Huang (1984, 1989)

Two main issues have attracted C.-T. J. Huang's attention, i.e. topic-prominence in Chinese-type languages and the pronominal aspect of PRO and *pro*. As for the former, he observed pronoun behaviours in different argument positions in Chinese, and found that

---

[15] The classical view of subject *pro* in which *pro* is an inherently unspecified nominal whose features are supplied by Agr can no longer be maintained, in particular if the feature theory of Chomsky (1995, 2001) is adopted. Since, by Chomsky's definition, the ϕ-features in T are uninterpretable, it is not possible for the unspecified pronoun to be specified by the ϕ-features in T. Holmberg (2005), discussing the existence of *pro* in Finnish, argues that null arguments are derived by PF deletion of a pronoun, rather than insertion of a silent lexical item, *pro*. In this sense, no *pro* exists. This is so-called Hypothesis B and is supported by the case of Finnish. Hypothesis B (Holmberg 2005: 538) states that "The null subject is specified for interpretable ϕ-features, values the uninterpretable features in Agr, and moves to Spec, TP, just like any other subject. This implies that the nullness is a phonological matter: the null subject is a pronoun that is not pronounced". In later works, Holmberg (2010a) and Roberts (2010a,b) argue that in agreement pro-drop languages, as in Italian for instance, a null subject in a finite clause is permitted, since the ϕ-features in T are so rich that all the features of the pronominal subject are represented in T, after valuation. In that situation, the pronominal subject will be a copy of T, and as such can be deleted. If they are right, null subjects in agreement pro-drop languages in general are interpreted as deleted copies of pronouns incorporated in (forming a chain with) T, rather than *pro*.

structurally overt arguments are not obligatory, despite lack of verbal agreement features. A null argument in discourse pro-drop languages can be interpreted as a deleted topic. That is, the null argument is coreferent with the discourse topic. Thus, he proposes the existence of *empty topics* in Chinese. The structure where the null argument is allowed is illustrated in (3.11) (see Section 3.3.1.1 for a discussion on topic variables). Such a null argument is a variable A'-bound by a topic NP:

(3.11) [$_{CP}$ topic$_1$ [$_{TP}$ . . . t_1. . .]]

Second, due to the very nature of PRO, which is in part pronominal, he integrated PRO and *pro* into one single, pronominal EC type, thereby deviating from Chomsky's category of null pronominals. C.-T. J. Huang argues that PRO and *pro* are really variants of the same entity[16] and are subject to the same control rule. For instance, *pro* in Chinese may be free, apart from being controlled, and this is a property shared by PRO:

(3.12) Zhangsan shuo   [*pro* hen xihuan Lisi].

        Zhangsan say   very like   Lisi

        'Zhangsan said that *e* liked Lisi.'

(3.13) [PRO xiyan] you hai.

           smoke has harm

        'Smoking is harmful.'   Chinese: C.-T. J. Huang (1989: 193)

According to C.-T. J. Huang, the embedded null subject in (3.12) can be free in its reference (co-indexed-free: given the right context, the null argument can refer to other people than the matrix subject, or to the addressee). The null subject in (3.13) has an arbitrary (generic) interpretation. (3.12) and (3.13) also entail another similarity of *pro* and PRO, in that both can occur only in subject position, particularly in Chinese.[17] For these reasons, C.-T. J. Huang proposes another typology for Chinese-type languages:

---

[16] I will argue in Chapter 6 that *pro* and PRO in discourse pro-drop languages are the same null nominal category, occurring in the same environments, i.e. where there is no agreement, and thus confirming that the formulation of C.-T. J. Huang's typology of ECs is correct (see Table 6).

[17] I will argue in Section 3.3.1.1 that the analysis that *pro* is allowed only in subject position cannot hold true for Thai.

Table 6: *C.-T. J. Huang's typology of ECs*

	Overt	Empty	Overt	Empty
+anaphoric	–	–	anaphors	NP-trace
–anaphoric	pronoun	*pro/* PRO	R-expression	variable
	+pronominal		**–pronominal**	

As PRO and *pro* can be collapsed into one, C.-T. J. Huang proposes a pro-drop principle, extended from Chomsky's (1981) rule of control to cover both PRO and *pro*. He then maintains that a null argument in subject position can be treated either as a variable or as *pro* if the clause is finite (see (3. 21) and (3.20) respectively). Alternatively, it is a PRO if the clause is non-finite. A null argument in the object position can only be treated as a variable (see (3.26) in Section 3.3.1.1 below). (3.14) is an instance of PRO in Chinese:

(3.14) Lao Wang qing   Xiao Li *e* lai.

Lao Wang invite Xiao Li   come

'Wang invites Li to come.'   Y. Huang (1994: 25)

The null argument is interpreted as PRO, and is obligatorily null, since the clause containing it is non-finite, i.e. the position of the subject of the second serial verb. According to C.-T. J. Huang's pro-drop principle, an empty pronominal takes the closest potential NP as its antecedent, unless it violates the principle of Disjoint Reference (DJR: a pronoun must be free in its governing category, which is synonymous with Chomsky's Binding Principle B). *Pro* in discourse pro-drop languages can be licensed, despite lack of rich agreement features, through the condition of index assignment from the antecedent:

(3.15) *Coindex an empty pronominal with the closest nominal element.*

C.-T. J. Huang's control rule, however, cannot account for data such as:

(3.16) Mama shuo yihuir   *e*   yao xiayu le.

mum   say in a moment   will rain ASP

'Mum says that it is going to rain soon.'

(3.17) Lao Wang$_1$ tongzhi Lao Li$_2$ $e_{1+2/3+4}$ xiawu   yiqi   qu kaihui.

   Lao Wang inform Lao Li   afternoon together go have a meeting

   'Wang informs Li that they will go to a meeting together this afternoon.'

<div align="right">Y. Huang (1994: 37)</div>

The embedded null subject in (3.16) cannot get a thematic role, as it is a quasi-argument, having no ϕ-features. It is, then, expletive *pro*. Such a null subject, therefore, cannot be coindexed with the closest matrix subject. The embedded null subject in (3.17) requires split antecedents, which can be either the matrix arguments or other referents fixed outside the sentence.

C.-T. J. Huang (1989), then, proposes revising control theory. His revised theory is known as the Generalised Control Rule (GCR) and has the following formulation:

(3.18) C.-T. J. Huang's (1989) Generalised Control Rule (GCR)

   A pro/ PRO is controlled in its control domain (if it has one).[18]

The significant refinement is that the requirement of coindexation with the closest nominal element is abandoned. In terms of control, it is predicted that in case *pro*/ PRO has a control domain, it is syntactically controlled in that domain, while the antecedent has to be local, unique, non-split, and non-arbitrary. In discourse pro-drop languages, *pro*/ PRO is controlled by a controller from outside its minimal TP. Put another way, lack of Agr leads such a controlee to be identified with a nominal element from a higher domain. Thus, *pro*/ PRO in such a language type is identified by control. In case *pro*/ PRO does not have a control domain, it may take a remote antecedent, it may receive an arbitrary interpretation, and it may have split antecedents, according to C.-T. J. Huang. In the following examples, in Thai and Chinese respectively, the EC is taken to be a subject *pro* in a finite clause:

---

[18] *Control Domain*:

α is the control domain for β if it is the minimal category satisfying (i) and (ii):

(i) α is the lowest S or NP containing

     (a) β, or

     (b) the minimal maximum category containing β;

(ii) α contains a SUBJECT accessible to β.

(3.19a) ...[TP *pro* kin lέεw]

                 eat PERF   Thai

(3.19b) ...[TP *pro* lai le]

                 come PERF   Chinese: C.-T. J. Huang (1989)

In both examples, the control domain for *pro* cannot be the minimal TP, as there is no subject accessible to it. If the minimal TP in (3.19a, b) constitutes an independent clause, the null subject has no control domain, and thus its antecedent cannot be determined. It follows that the null arguments cannot be interpreted as referential third person *pro*s. The GCR would predict the null subjects to be topic variables or have an arbitrary (generic) interpretation. In Thai, I claim that neither a topic variable nor a generic pronoun can fill the subject position in such cases. If the sentence (3.19a) is uttered out of the blue, the only possible interpretation of the null subject is referring to the speaker by default (see Chapter 6, Section 6.3.3.1). In (3.19b), *pro* can be interpreted to be third person iff there is an understood topic; and this applies to Thai as well. When *pro* has a control domain, for example when the minimal TP is a complement embedded under control verbs, such as *say/ promise*-type verbs, according to the GCR, it is coindexed with the matrix subject, as can be seen in Thai:

(3.20) [CP [TP dam$_1$ bɔ̂ɔk [CP wâa [TP e_1 dâayyin (lέεw)]]]]

          Dam say COMP   hear   PERF

     'Dam said that he could hear.'

C.-T. J. Huang (1989) argues that the embedded null subject is controlled in the control domain that contains the accessible subject – the closest nominal element, namely *dam* 'Dam'. Therefore, *pro* can take the embedded subject position iff it is controlled by a higher argument.[19] The embedded subject position can be occupied not only by *pro* but also by a variable. If it is a variable, then it is A'-bound by a null topic:

---

[19] In this respect, Thai and Chinese pattern like partial pro-drop languages, e.g. Marathi, Finnish, Brazilian Portuguese, in relation to control into finite clauses, as investigated in Holmberg et al. (2009) and Holmberg and Sheehan (2010). That is to say, the only exception where the null subject can have a referential reading (*pro*) is where they are bound by a c-commanding higher referential DP. This is because the partial pro-drop languages lack D-feature [D] in T (see also Chapter 6, Section 6.7).

(3.21) [$_{CP}$ TOP$_2$ [$_{TP}$ cim$_1$ bɔ̂ɔk [$_{CP}$ wâa [$_{TP}$ t_2 maa   (lɛ́ɛw)]]]]

                Jim   say   COMP   come   PERF

      'Jim said that (s)he came.'

The embedded null subject refers to somebody else introduced in the previous discourse. C.-T. J. Huang's claim that an embedded null subject can be a topic variable in turn poses a problem for the GCR. In (3.21), although the matrix clause contains a subject accessible to the embedded null subject, the matrix subject is not actually the antecedent of the embedded null subject, and this contradicts both versions of C.-T. J. Huang's control rules.

Y. Huang (1994, 2000) observes that C.-T. J. Huang's control rules cannot account for Chinese with regard to several interpretations of *pro*s, as shown in (3.22). Also, data in Korean and Thai appear to contradict C.-T. J. Huang's (1984, 1989) rules, as can be seen from the following:

(3.22) fuqin   shuo   *e*   yao   weirenzhengzhi

      father   say   should upright

     'Father says that one should be upright.'   Chinese: Y. Huang (1994: 36)

(3.23) mɛ̂ɛ$_1$ (khɔ̌ɔŋ chǎn) hěn   khǎw$_2$ kòon (khǎw$_{\#1/2}$) pay hǎa   mɔ̌ɔ

      mother of-GEN I   see   (s)he   before (s)he   go   see   doctor

     'My mother saw him/ her before (s)he went to see the doctor.'   Thai

(3.24) *e*$_1$ John-i$_2$ *e*$_{2/1}$ Mary-lul   poassta-ko malhayssta

      TOP John-Nm   Mary-Acc saw-Comp said

     'John said that he saw Mary.'   Korean: Choo (1994: 18)

(3.25) Xiaohong$_1$ de   meimei$_2$ shuo (ta$_{1/2/3}$) xihuan tan   gangqin

      Xiaohong GEN younger sister say   she   like   play   piano

     'Xiaohong's younger sister says that (she) likes to play piano.' Chinese: Y. Huang

                                                      (2000: 66)

In (3.22), *pro* can have an arbitrary interpretation (in fact, generic: see Chapter 4 for a distinction between an arbitrary reading and a generic reading of pronouns), rather than being

controlled by the accessible subject *fuqin* 'father', according to Y. Huang. In (3.23), either subject control or object control is allowed in the sentence. In other words, not only the subject, but also the object of a main clause may be an antecedent of the null subject,[20] thereby resulting in an ambiguous sentence if it appears in isolation. Having said that, the embedded null subject in (3.23) tends to be interpreted as coreferent with the matrix object. That is to say, as indicated by the hash, the preferred reading is that the embedded null subject takes the closest NP as its antecedent. In (3.24) and (3.25), however, the matrix clauses need not be the control domain for *pro*s, given that the sentences appear in isolation. It is possible to construe *pro* as a controlee of the matrix subject or as a variable A'-bound by a null topic or a long-distance (understood) antecedent.

3.3.1.1 Topic variables and object ECs

C.-T. J. Huang (1984, 1989, 1991) hypothesised that when there is an (understood) discourse topic, null objects will be interpreted as coreferent with the topic, thereby forming a topic chain. In his analysis, a variable results from a movement to a null topic operator in A'-position of a base-generated null object. The canonical object position is therefore a trace of the moved topic variable. A null object *pro* is ruled out by the conflicting requirements of the GCR and Binding Principle C. When there is a null object, the matrix subject is the potential NP that the GCR requires a null pronoun to be coindexed with, but Binding Principle C requires the null object to be disjoint from the subject. If the null object cannot be a pronoun, it must be a variable. As such, it can only have its reference fixed in non-argument position outside the sentence. Therefore, for C.-T. J. Huang, null objects are more restricted than null subjects in that they cannot be an instance of *pro*. This is illustrated in (3.26):

(3.26) Zhangsan$_1$ xiwang Lisi$_2$ keyi kanjian $e_{*1/ *2/ 3}$

      Zhangsan hope   Lisi can   see

      'Zhangsan$_1$ hopes that Lisi can see him.'   Chinese: C.-T. J. Huang (1984: 538)

The coindexation of the null object and the matrix subject, namely *Zhangsan*, is predicted to be ungrammatical. Given the GCR, the null object is assumed to be a topic variable. As a topic variable, it cannot be A-bound by the matrix subject owing to the binding Principle C.

---

[20] For the analysis of sentences, like (3.24) see Chapter 6, Section 6.6.1.

Thus, the coindexation between the null object and a null topic in A'-position helps the sentence escape from Condition C violation. A Sentence like (3.26) has a representation as the following:

(3.27) [$_{CP}$ Op$_3$ [$_{TP}$ Zhangsan$_1$ hopes [$_{CP}$ that [$_{TP}$ Lisi$_2$ can see [t_3]]]]]

However, naturally occurring data in discourse pro-drop languages indicates that C.-T. J. Huang's analysis of null objects does not hold true across the board. In Thai, a null object is not restricted to a topic variable, but it may be coreferential with the matrix subject as well. In other words, the subject argument of the matrix clause can serve as the antecedent of the null object. This is to be expected if the null complement object is an instance of *pro*:

(3.28) (chăn$_1$) rúusùk wâa   mii khɔn   mooŋ   (chăn$_{1/3}$) yùu

        I   feel   COMP have person look   I   PRES CONT

        'I feel that somebody is looking at me/ him/ her.'

Assuming that the sentence appears in isolation, it is ambiguous. In this sentence, the null object can be either antecedent-linked, given a possibility that the matrix subject can be the antecedent of the null object, or topic-linked when null (despite being a less preferable reading). Based on the fact that the embedded subject can optionally be null, it is assumed to be directly controlled by the matrix subject, contra C.-T. J. Huang's analysis. The matrix subject itself can optionally be null, since it refers to the dominant speech participant. A similar reading can be constructed in Chinese as well. In fact, in (3.29) the interpretation of the null object is restricted to the matrix subject, for pragmatic reasons. More specifically, the verb *kanjian* 'see' in the embedded clause rules out a topic-variable reading, according to Xu (1986):

(3.29) xiaotou$_1$ yiwei mei ren   kanjian $e_{1/*2}$

        thief   think no   man   see

        'The thief thought that nobody saw him.'   Chinese: Xu (1986: 78)

(3.28) and (3.29) show that the null object can be coreferential with the matrix subject, and as such this will suffice to indicate that the null object need not be an instance of a topic variable, at least in these examples. Another example, from Thai, indicates that the matrix

subject as the antecedent is the only option. Again, this is due to pragmatic reasons. The object can optionally be null, since it is directly controlled by the local antecedent:

(3.30) Ɂɛɛn₁ pay khruŋthêep phûahây   lûuklûuk (khɔ̌ɔŋ thəə)₂ dâay cəə (thəə₁/*₃)

        Ann   go Bangkok   in order to children of-GEN her   able see her

    'Ann went to Bangkok so that her children could see her.'

There are several linguists who provide counter-examples to the claim that a null object is a topic variable by arguing that embedded null objects in such discourse pro-drop languages as Korean, Thai, and even Chinese are instances of *pro*s, since they are allowed to be coreferential with the matrix subject (Xu (1986); Cole (1987); Aroonmanakoon (1999); Han (2006), among others). Thus, it is likely that the reference of a null object is pragmatically inferred. As a result, it can be identified through the matrix subject (the higher NP as its antecedent) or a discourse topic. The null pronominal objects in the following examples are considered well-formed. Analysing all null objects as topic variables is therefore too restricted:

(3.31) Toli₁-ka   Swuni-ka₂ $e_{1/3}$ kwoylophi-ess-ta-ko   malha-ess-ta.

    Toli-Nom Swuni-Nom (OBJ) tease-Past-Dec-Quote   say-Past-Dec.

    'Toli said that Swuni teased him.'   Korean: Han (2006: 13)

(3.32) kim₁ bɔ̀ɔk wâa dam₂ chɔ̂ɔb (khǎw₁,₃)

    Kim   say   that Dam   like   he

    'Kim said that Dam likes him.'   Thai: Aroonmanakoon (1999: 33)

If the context is insufficient, the interpretation of the null object in these examples can be ambiguous. Such a null object has two options in terms of its interpretation – it can be linked[21] to either a (linguistic) antecedent or a distinct topic, except for (3.30) where the null object is only antecedent-linked. In (3.32) the embedded object can optionally be null, as it is directly controlled by the matrix subject – the local antecedent of the null object.

---

[21] The topic-linking and antecedent-linking of null arguments are two types of context-linking. Linking of first and second person null arguments to the speaker and hearer is also context-linking of a sort, according to Sigurðsson (2004).

To summarise, in C.-T. J. Huang's (1984 and subsequent) proposals, there are two parameters involved in the discourse pro-drop languages: the discourse-orientation parameter and the pro-drop parameter. In other words, the major distribution of null arguments in discourse pro-drop languages within C.-T. J. Huang's approach depends on discourse aspects. Although a null object in Thai and Chinese can be directly controlled by the local matrix argument, the analysis that null arguments are identified by a null topic, which in turn is linked to a discourse topic, is crucial, and will be pursued in the analysis chapter (Chapter 6). Nonetheless, to state that a null object cannot be an instance of *pro* appears too strong. I have shown that a null argument, for instance, in object position can be coreferent with the matrix subject, and as such is an instance of *pro* in terms of GB theory. The coreference between null arguments and their antecedents is subject to discourse rules of grammar. However, the fact that a null argument can be controlled by a c-commanding matrix argument or bound by a discourse topic indicates that it should belong to another syntactically projected null argument, other than *pro* or a topic variable (in C.-T. J. Huang's sense, the distribution of *pro* in particular is too restrictive). There is a similarity in the interpretations of the null arguments in both cases (c-commanding antecedent and discourse antecedents): at least a referential third person null argument must have an antecedent which is third person, regardless of whether there is a control domain. The absence of control domain poses no problem for a null argument in searching for its antecedent, since it is the context that dictates which is the antecedent of the null element. This suggests that the analysis of a null argument as *pro* in GB terms may be irrelevant at this point (See Chapter 6, Section 6.1 for the analysis of null arguments in Thai and discourse pro-drop languages).

3.3.2 Xu (1986, 2003)

In contrast to C.-T. J. Huang, Xu (1986), discussing the obligatorily/ optionally overt pronouns in Chinese, abandons the notion of *pro*. Xu proposes that Chinese has a Free Empty Category (FEC), which is emptier than *pro* in the sense that it is not specified for pronominal or anaphoric features. In a later work, Xu (2003) indicates that the choice of an overt or null form of a pronominal argument is determined by three factors: whether it is lexically controlled or uncontrolled, grammatically determined or undetermined, and contextually sensitive or free. These factors are represented by binary combinations: [± L], [± G], and [± C] respectively.

The first factor [±L] is based on the nature of the matrix verbs in Chinese. Xu claims that a strong transitive verb, like 'take', must be followed by an object, but the object may not be pronounced. A weak transitive verb, like 'eat', need not take an object; whether the complement has to be null or overt is not what matters. The null object, therefore, is lexically controlled by the verb type. On the other hand, a Chinese matrix subject is [–L], since it is lexically uncontrolled by the matrix verb.

Second, Chinese is always grammatically undetermined [–G], given the presence of a certain type of matrix verbs, such as *advise*-type verbs, according to Xu. (3.33) exemplifies an optionally embedded null subject which is dependent on the matrix verb *quan* 'advise':

(3.33) wo *quan* Zhangsan mingtian wanshang ba dianzhong zhiqian (*ta*) bu yao qu

        I   advise Zhangsan tomorrow evening eight o'clock   before   he not will go

        'I advised Zhangsan not to go before eight o'clock tomorrow evening.'

<div align="right">Xu: Chinese (2003: 91)</div>

Xu claims that Chinese does not make a distinction between finite and non-finite elements, unlike English, in which a non-finite clause must follow the verb 'advise'. The choice of English overt/ null pronouns is [+L] and [+G], whilst in Chinese it is [+L] and [–G]. That is to say, in Chinese, any type of clause can follow the matrix verb *quan* 'advise', and therefore an embedded subject can optionally be pronounced. Thus, whether an argument can be pronounced tends to be contingent on lexical control, rather than gramatical determination. However, such a conclusion concerning a lack of grammatical determination [–G] cannot account for the occurrence of overt/ null pronouns in discourse pro-drop languages in general. In fact, both grammatically determined and undetermined factors play a vital role in the occurrence of a null/ overt pronoun. Take the case of embedded subjects in Thai, which are characterised as either [+ or –G]. Thai is considered to be [+G] when a null embedded subject is controlled by a local c-commanding antecedent. On the other hand, Thai is considered to be [–G] when an embedded subject is controlled via a null topic chain, or when it is associated with a discourse function, hence overt (see Chapter 5). Therefore, in Chinese-type languages, [–G] is rather unlikely to be the case across the board.

Contextual sensitivity is Xu's last factor accounting for the occurrence of overt/ null pronouns. The fact that Chinese-type languages exhibit highly frequent pro-drop despite the

absence of verbal agreement means that the recovery of the antecedent relies on the context. It follows that such languages exemplify [+C].

It appears that Xu's so-called binary features are purely descriptive, in that the occurrence of overt/ null pronouns is described as governed by lexical, grammatical, or contextual factors which are themselves in need of explanation. In fact, they cannot provide an exhaustive account for the root cause of the null argument problem: sometimes these pronouns have to be pronounced, sometimes they may, but need not be, and sometimes they cannot be pronounced. For this reason, Xu's account is untenable, and will not be pursued here.

3.3.3 Tomioka (2003)

Tomioka (2003) point outs that discourse pro-drop languages do not have obligatory marking of (in)definite determiners, nor plural morphology on NPs. The languages also reveal the obligatory use of a classifier in the presence of a numeral. Thus, bare NPs, which belong to a minimally specified nominal category displaying no determiners, can have a wide array of interpretations. He sets out to study discourse pro-drop languages with regard to semantic aspects, and proposes that null pronouns in these languages are synonymous with bare NP arguments. Due to similar interpretations across such languages, Tomioka (2003:336) puts forward a hypothesis which he calls the Discourse Pro-drop Generalisation:

(3.34) *All languages which allow discourse pro-drop allow (robust) bare NP arguments.*

In this hypothesis, bare NP arguments and null pronouns exhibit the same property, thus leading to the conclusion that null arguments in Japanese are instances of *pro*.

To see whether discourse pro-drop languages allow bare NP arguments in the syntax, Tomioka observed Japanese, Chinese and Korean, and found that these languages have the same range of interpretational possibilities. Such possibilities include definite pronouns of sloppy identity, indefinite pronouns of sloppy identity, and quantifier stranding. These are exemplified below in (3.35), (3.36), and (3.37) respectively:

(3.35) Zhangsan kanjian ziji de   xiong le   Lisi ye kanjian *e* le.

   Zhangsan see   self GEN bear   PERF Lisi also see   PERF

   'Zhangsan saw self's bear. Lisi saw self's bear.'[22]   Chinese: Tomioka (2003: 333f)

(3.36) Zhangsan kanjian xiong le.   Lisi ye   kanjian *e* le.

   Zhangsan see   bear PERF Lisi also see   PERF

   'Zhangsan saw a bear. Lisi saw a bear.'

(3.37) Zhangsan kanjian 2-zhi xiong le.   Lisi kanjian 3-zhi *e* le.

   Zhangsan see   2-CL bear   PERF Lisi see   3-CL   PERF

   'Zhangsan saw two bears. Lisi saw three bears.'

Different interpretations of *pro* other than pronominal in Chinese are possible, according to Tomioka. This generalisation also holds true for Thai. (3.38) and (3.39) exemplify cases of definite and indefinite pronouns of sloppy identity respectively. (3.40) illustrates a case of quantifier stranding:

(3.38) pim khǎay thorásàp mɯɯthɯ̌ɯ khɔ̌ɔŋ   tuaʔeeŋ pay nók kôo khǎay *e* mɯ̌ɯnkan

   Pim sell phone   mobile   of-GEN self   go Nok also sell   as well

   'Pim's sold her mobile phone. Nok's also sold her mobile phone.'

(3.39) pim klua phǐi nok   kôo klua *e* mɯ̌ɯnkan

   Pim fear ghost Nok also fear   as well

   (Lit.) 'Pim fears ghosts. Nok also fears ghosts.'

   = 'Pim is afraid of ghosts and so is Nok.'

(3.40) pim sɯ́ɯ sɯ̂ɯ   sɔ̌ŋ-tua nán hây cim nók sɯ́ɯ *e* sǎm-tua níi   hây dam

   Pim buy shirt 2-CL   DEM give Jim Nok buy   3-CL   DEM give Dam

   'Pim bought those two shirts for Jim. Nok bought these three shirts for Dam.'

The null object in (3.35) and (3.38) illustrates a pronoun with a pronoun-containing antecedent, according to Tomioka. The null object cannot share the same index as the one in the antecedent clause. If they do, the sentences will be ruled out as semantically and

---

[22] Note that this is Tomioka's (literal) translation.

pragmatically ill-formed due to a referential reading. Thus, *pro* in this case must be disjoint from the object NP in the antecedent clause, hence a sloppy interpretation. Thus a strict reading is ruled out for pragmatic reasons. The null object in (3.36) and (3.39) is interpreted as indefinite. As Tomioka points out, it appears that such a construction does not exist in English. If one wishes to express a similar reading like (3.36), one should use a construction known as N'-Deletion: 'Jane found a dog, and Jim found *one*, too.'. The null object in (3.37) and (3.40) illustrates that Thai and Chinese allow a quantificational interpretation. For instance, in (3.40), the shirts *Nok* bought can be, and must be, different from those *Pim* bought.

These examples show that null NPs can have different interpretations. Tomioka's analysis is that null pronouns in discourse pro-drop languages are simply the result of deleted N' /NP without determiner stranding possible because nominal arguments in these languages do not necessarily have a determiner. Therefore, the null arguments in the above examples have a syntactic representation as follows:

(3.41)   NP
                 |
                 Ø   Tomioka (2003: 336)

Tomioka's analysis of null objects is *not* entirely different from the theory that will be put forward in Chapters 4, 5, and 6, where I will argue that null arguments are null nouns. However, Tomioka's (2003) theory accounts only for a narrow range of null arguments (mainly null objects). It does not generalise in any straightforward fashion, for example, to cases where subjects are optionally null, as when they refer to the speaker (see Chapter 2, Section 2.2.2.2), or when they are controlled by a higher argument, as discussed in Section 3.3.1 above.

3.3.4 Saito (2007)

Saito studies the interpretation of argument ellipsis,[23] building on Oku (1998) and Saito

---

[23] Oku (1998) was among the very first to propose 'argument ellipsis' in order to explain the null argument phenomenon in sloppy identities and bound variable readings, particularly in Japanese and Korean. That is to say, null arguments stem from elision of full-fledged structures based on the fact that nominal arguments in such languages are allowed to undergo ellipsis; see also Takahashi (to appear) for a similar idea.

(2003).[24] He analyses it as involving covert LF copying. (3.42) shows the availability of argument ellipsis embedded in the VP; (3.43) exemplifies Saito's analysis of (3.42):

(3.42) Hanako-wa [CP [DP zibun-no teian -ga] [T' saiyoosareru] to] omotte iru
        Hanako-TOP   self -GEN proposal-NOM accepted-be that think
        'Hanako₁ thinks that her₁ proposal will be accepted.'
        Taroo-mo [CP [DP e] [T' saiyoosareru] to] omotte iru
        Taroo-also   accepted-be that think
        'Taroo₂ also thinks that his proposal₂ will be accepted.'

(3.43) Taroo-mo [CP [DP *zibun-no teian -ga*] [T' saiyoosareru] to] omotte iru
        Taroo-also   accepted-be that think

                                                Japanese: Saito (2007: 7)

In (3.42) the embedded subject of the second sentence is null in the overt syntax (hence at PF). *Zibun-no teian-ga* 'self's proposal' in the antecedent clause is then copied and merged with the predicate in the embedded clause in (3.43) in the covert syntax. At LF, the possessive reflexive *zibun* can then be bound by the subject, which yields the sloppy identity reading. Thai appears to pattern similarly to Japanese with regard to sloppy identity:

(3.44) pim rák bâan khɔ̌ɔŋ khǎw kim kɔ̂ rák *e* mǔankan
        Pim love house of-GEN her Kim also love   also
        'Pim loves her house. Kim also loves her house.'
        = Pim loves her house and so does Kim.

(3.45) pim₁ rák [NP bâan khɔ̌ɔŋ khǎw₁] kim₂ kɔ̂ rák [NP *bâan khɔ̌ɔŋ khǎw₂*] mǔankan
        Pim love   house of-GEN her Kim then love house of-GEN her   also
        'Pim loves her house. Kim also loves her house.'

According to Saito's argument-ellipsis analysis, the sentences in (3.44) have the representations in (3.45) in covert syntax. After spell-out, on the LF-branch of the derivation

---

[24] Oku (1998) and Saito (2003) propose that the availability of scrambling, a movement operation responsible for the free word order phenomenon, in Japanese is correlated with the possibility of NP-ellipsis.

(in a model as in Chomsky (1993, 1995)), the antecedent NP is copied onto the ellipsis site in (3.44), resulting in the LF representation in (3.45), and having a sloppy reading, regardless of whether the pronoun is overt or null. For this reason, this analysis is known as covert LF copying of an NP argument. If the NP is copied with the index, a strict reading will be obtained, if it is copied without the index, it gets re-bound, and a sloppy reading will be obtained. This is the mechanism of VP-ellipsis, too.

Argument ellipsis and radical pro-drop (equivalent to discourse pro-drop), according to Saito, originate from the same source, as he points out:

> The distributions of argument ellipsis and radical pro-drop seem identical if we abstract away from the differences that arise from the properties of ellipsis and pronouns; the former requires a linguistic antecedent and allows sloppy interpretation whereas the latter receive definite interpretation. (2007: 18)

He then concludes that radical pro-drop is a kind of argument ellipsis. In fact, Saito's theory is based on the idea that what is crucial about radical pro-drop languages is that they have no [uϕ] features, hence no agreement. If a language has [uϕ] features, arguments have to be merged in overt syntax, because otherwise the features will not get valued before they get spelled out at PF (i.e. agreement will fail). But if a language has no [uϕ] features, arguments can be merged in covert syntax, i.e. they can be null. This is an interesting idea. However, as Saito is not explicit about how this theory would account for the more detailed properties of the variety of null arguments discussed in this thesis, I will from now on disregard Saito's theory.

3.3.5 Speas's (2001) Optimality Theory (OT) for the syntax of null pronouns

Speas (2001) develops an OT-based model of null pronouns in an attempt to explain null argument phenomena cross-linguistically. In search of the optimal *candidate* (a null argument sentence) that is syntactically well-formed, she proposes a set of constraints in the light of Cole (1987) and C.-T. J. Huang's (1989) GCR to filter out unqualified candidates. The constraints are comprised of CONTROL, BINDING PRINCIPLE B, and MAX (*Pro*). The most crucial constraint is MAX (*pro*): If *pro* occurs in the input, then its output correspondent is *pro*. Each asterisk represents one violation (Dekkers et al. (2000)). Speas also notes that in

Thai, the constraint CONTROL is not as crucial as the other constraints – MAX (*Pro*) and BINDING PRINCIPLE B. This is based on C.-T. J. Huang's idea that the null embedded subject can be a topic variable, being disjoint in reference from a higher matrix argument. If the constraint CONTROL is violated, the candidate still has a chance to be optimal. Correspondingly, if the constraint, either MAX (*Pro*) or PRINCIPLE B, is violated, such a candidate will be filtered out.

Table 7: *Speas's (2001:409) tableau of Thai embedded subjects*

Input: Nít bɔ̀ɔk wâa *pro* hěn Nɔ́ɔy	PRINCIPLE B	MAX (*Pro*)	CONTROL
☞(a) Nít bɔ̀ɔk wâa [*pro* hěn Nɔ́ɔy] Nit said that saw Noy			*
☞(b) Nít₁ bɔ̀ɔk wâa [*pro*₁ hěn Nɔ́ɔy]			*
(c) Nít bɔ̀ɔk wâa khǎw hěn Nɔ́ɔy Nit said that he saw Noy		*!	

As the constraints are set up, the candidates (a) and (b) are optimal, since each respects PRINCIPLE B and contains an embedded subject *pro*, which satisfies MAX (*pro*). As for (c), it violates MAX (*pro*), given that pro-drop is allowed in this context. If we take the line that the asterisk indicates a certain violation, it is not, however, clear how the candidate (b) violates CONTROL, since the embedded null subject is essentially controlled by the c-commanding matrix argument, and thus satisfies CONTROL. In this respect, it is likely that (b) is better than (a) (hence the optimal candidate), as all the constraints are respected.

In its present form, the theory accounts for only a small fragment of the cases where null arguments occur in Thai. It is therefore not comparable to the theory put forward in the next chapters. Given that the determination of this approach to select the optimal candidate is influenced by the binding principles and C.-T. J. Huang's GCR, it cannot account for the environments where pronouns in Thai are obligatorily overt or optionally null, or even the distribution of quasi-inclusive generic pronouns and exclusive generic pronouns in Thai, which cannot be null (see Chapter 4, Sections 4.5.2 and 4.5.3).

3.4 Summary

Even though the GB theory of empty categories has for the most part lost its relevance, the

question remains whether null arguments in discourse pro-drop languages are *pro* or PRO or variables (traces of null topics), or possibly deleted N' or NP. Discourse pro-drop languages have thematic null subjects in finite clauses, but no identifying agreement morphology. The classical view of pro-drop is that the null pronoun is identified by agreement, and that only discourse pro-drop languages rely on the context for identification. As mentioned, this view has recently become obsolete; all languages rely on the context for assignment of reference to null arguments. Even so, the presence or absence of agreement may play a crucial role. Several theories concur that the absence of agreement in discourse pro-drop languages is a crucial feature which allows the relatively free pro-drop. Nonetheless, they treat ECs differently, and thus diverge markedly from each other. C.-T. J. Huang treats null arguments as *pro*s when controlled by a c-commanding NP in a higher clause. If there is no controlling higher argument, they will be variables A'-bound by a discourse-topic operator in the matrix Spec, CP, which itself is linked to a referential NP in the discourse. Xu claims that the occurrence of null/ overt pronouns in Chinese is determined by three binary features, i.e. ± grammatically determined; ± lexically controlled; ± contextually sensitive. Tomioka argues that (a certain type of) null arguments are derived by NP-deletion, leaving a null argument iff the argument has no determiners. Saito diverges from both Huang and Tomioka, and hypothesises that a null object is grammatical when the referential NP in the antecedent clause is copied and merged with the predicate after spell-out in the derivation of LF. After the copying operation, the VP will have a licit transitive configuration, and the missing NP can thereby have a referential reading (where an anaphor in the copied NP can be bound by antecedent in the second conjunct, to yield sloppy identity). Speas (2001) is an attempt to formalise a part of the syntax of null pronouns in terms of Optimality Theory. The basic idea is that there is a set of constraints, with language-specific grammars having different rankings of these constraints.

The theories focus on different aspects of the theory of pro-drop in discourse pro-drop languages. None of them present an exhaustive account (although Huang (1989) comes closest). Taking an important step towards reconciling the proposals discussed above, I shall investigate in more detail how discourse influences the occurrence of referential null arguments in Chapter 5. The different theories, nonetheless, provide a useful baseline for analysing null arguments in Thai.

Chapter 4: Impersonals, Genericity and Indefinite Topics

Thai exhibits extensive use of pronouns to refer to impersonal referents, yielding a wide range of interpretations. This chapter explores the occurrence of different types of impersonal pronouns, particularly the correlation between generic or arbitrary readings of personal pronouns and their appearance (null *vs* overt), together with their internal properties. Relations between the first and the second occurrences of impersonal expressions in the same sentence are also discussed, whether the relation is operator binding or co-reference. The chapter ends with a discussion of null indefinite argument topics, their properties and interpretation, in Thai.

4.1 Some Definitions

Jensen (2009: 85) defines what he calls impersonal pronouns as being characterised by having referents that are human and generalised:

> ...the descriptive reference may include the speaker, the addressee or some specific third party, but it always goes beyond that in an unspecified way (though the context of use often delimits the extension to some degree). The pronoun refers to a generalized person, and what is predicated about this referent is asserted to hold for every instantiation of the type.

This is in fact a definition of generic pronouns, or G-pronouns, rather than impersonal pronouns. Under this definition, G-pronouns have generic inclusive readings. *Inclusive* here means reference which includes the speaker and the addressee. The definition, nonetheless, fails to capture the variety of readings that impersonal pronouns may have. Impersonal pronouns in this thesis are understood as referring to one or more persons, but no specific person is picked out in contrast to the personal pronouns. Thus impersonal pronouns typically have no antecedent. They can be further classified into two main types: pronouns that have generic readings and those with arbitrary readings. The former type always has a plural reading; the interpretation of the latter is roughly paraphrased by existential quantification, like *someone*, and not necessarily including the speaker or addressee.

According to Egerland (2003) and Sigurðsson & Egerland (2009), there are three possible readings that an impersonal argument, more precisely an impersonal subject, can have, i.e. either generic, specific, or arbitrary. An example of an impersonal subject having a generic reading is (4.1).

(4.1) *Man*   måste arbeta till 65.

     one   must   work until 65

     'You must work until you are 65.'   Swedish: Egerland (2003: 7)

The sentence has a generic reading. Not only the reference of the pronoun but also the time reference must be generalised. The subject *man* refers to a quasi-universal set of individuals, roughly equivalent to people in general in a non-restricted sense with unspecified time reference, according to Egerland (2003).

When an impersonal subject is used for a specific reading, it also has a plural interpretation, just as when the impersonal subject has a generic reading. The French impersonal pronoun *on* in (4.2) is used for a specific reading:

(4.2) Hier   soir   *on* a été congédié.

    yesterday evening ON has been fired

    'We were fired yesterday evening.'   French: Sigurðsson & Egerland (2009: 163)

The pronoun *on* is in fact equivalent to an impersonal use of the English personal pronoun *we*, which has a generic quasi-inclusive reading (cf. (4.13) and (4.14)). It does not necessarily include the addressee. The term 'specific reading' of an impersonal pronoun is potentially confusing, as it can be taken to refer to the referential reading of personal *we*. For this reason, I shall refer to an impersonal subject having a specific reading as a *quasi-inclusive* G-pronoun.

Egerland (2003) goes on to point out that an impersonal subject can also have an arbitrary interpretation, the speaker and the addressee excluded. The following example shows a further usage of the Swedish pronoun *man* taking an arbitrary reading:

(4.3) *Man* arbetade i tre   månader för att lösa   problemet. Swedish: Egerland   (2003: 73)
　　 man worked for three months   to   solve the problem
　　 'People/ They have worked for three months to resolve the problem.'

According to Egerland, the tense reference is episodic and the subject takes an existential reading. In this case, the subject refers to an unspecified, 'arbitrary' set of people, not including the speaker or the addressee. The exclusive reading of an arbitrary pronoun is therefore similar to that of an exclusive G-pronoun.

4.2 The English System of Impersonals

As English is a non-pro-drop language, English impersonal pronouns are obligatorily overt. Rizzi (1986) does not distinguish between arbitrary and generic readings of pronouns. He calls all of them 'arbitrary'. When used impersonally, the term 'arbitrary' should, however, be reserved for the meaning 'some arbitrary person(s)', and thus should be distinguished from the term 'generic', since the two readings seem to diverge markedly from each other. Taking this into account, the system of impersonals in English consists of one purely generic pronoun, i.e. purely generic *one*, and a set of pronouns which can have referential or generic/ arbitrary meaning, consisting of *you*, *we*, and *they*. *You* and *one* can be used interchangeably when they have the generic inclusive reading. (4.4) is an example of a generic inclusive reading:

(4.4) One/You has/ have to pay more for organic products.

Such pronominals stand in contrast to *they* in terms of clusivity. *They* has a personal, referential reading and two impersonal readings, i.e. a generic exclusive reading, as in (4.5), and an arbitrary or quasi-existential[25] reading in an episodic sentence such as (4.6):

(4.5) They go to the temple in Thailand on New Year's Day.
　　 = Some people (not someone) go to the temple in Thailand on New Year's Day.

---

[25] Cinque (1988) calls this reading quasi-existential, as it is compatible with an existential quantifier.

(4.6) They stole my car yesterday.

    = Someone stole my car yesterday.

As the term *genericity* implies, the generic exclusive reading of *they* in (4.5) always has a plural reading. The arbitrary *they* in (4.6) may be used to refer to unspecified 'they' or 'someone'. Therefore, it does not necessarily have a plural reading. Due to its compatibility with an existential quantifier, the arbitrary *they* allows a singular reading, while the generic exclusive *they* does not.

Another prominent feature of a generic exclusive reading of *they* is that it typically co-occurs with a locative, unlike arbitrary *they*, which does not require a locative to be present. Compare:

(4.7) They eat snails *(in France).

(4.8) They said there was a fuel shortage (in France).

Without the locative 'in France' *they* in (4.7) would lose the generic exclusive reading (marked by the asterisk), and the interpretation would be referential. In (4.8) the sentence is ambiguous between a referential reading and an arbitrary reading. Assuming an arbitrary reading of *they* in (4.8), the absence of the same locative expression does not affect the arbitrary reading of *they* at all. The presence of such an expression simply makes the sentence more informative.

Arbitrary *they*, according to Cabredo Hofherr (2003), can be further classified into four main sub-types, i.e. specific existential, vague existential, inferred existential, and corporate readings. (4.9) illustrates a case of a specific existential reading of *they*:

(4.9) They are fixing your car now.

They in (4.9) is specific existential in the sense that it is anchored to a specified point of time. The sentence is also considered more informative than the sentence in (4.10) in that it can be a reply to questions: *what* and *when*. If the pronoun is not anchored to a particular point of time, the reading will be vague existential:

(4.10) They have found my dog.

The sentence focuses on the current status of the found dog, yet when it was found is not focused on. Thus this type of sentences is considered vague existential. (4.11) shows an inferred existential reading of *they*, where the sentence uttered is inferred from the visible result:

(4.11) They have cleaned the public toilet.

The sentence illustrates that the most informative part of the sentence is the predicate, rather than the subject. Arbitrary *they*, by definition, refers to some arbitrary persons, and thus is not the focus of the sentence at all. In fact, the predicate shows that the speaker infers from the cleanliness of the toilet that some cleaning activities have taken place, without paying any attention to the person who cleaned it.

Cabredo Hofherr's fourth possible reading of arbitrary *they*, i.e. a corporate reading is shown in (4.12). It is used to refer to a particular group, in this case *postmen*:

(4.12) They delivered the parcel yesterday.

The *postmen* is a designated group of people who actually carried out the parcel delivery. As such, it cannot be *anybody* or *someone*, but is restricted to that group of people. As stated above, the existential readings can be paraphrased by an existential quantifier without any syntactic and semantic ill-formedness. That is, *they* in (4.8), (4.9), (4.10), and (4.11) may be replaced by 'someone' and a similar meaning is still retained. The corporate reading of *they* is, however, an exception, since the predicate pragmatically restricts the designated subject. Therefore, it cannot be replaced by an existential quantifier. Whether the arbitrary *they* can be paraphrased by an existential quantifier depends on what reading type it belongs. Nonetheless, the case of the corporate reading of *they* is quite tricky to handle, since it may be interpreted to be arbitrary, existential, or indefinite, or overlap between two of those readings. As it is not purely existential, it cannot be paraphrased by existential quantification.

The aforementioned examples show that the personal pronouns *you* and *they* can have an impersonal use. In fact, the pronoun *we* can also be used in the same generic sense as the pronouns *you* and *they*, according to Kamio (2001). The impersonal reading of personal pronoun *we* would be generic quasi-inclusive, referring to people in general, including the

speaker in its reference. This leaves the generic quasi-inclusive *we* in between the generic inclusive reading and the generic exclusive reading in the continuum of genericity.

(4.13) We all have to eat.

(4.14) We have the Glastonbury Festival in England.

The personal pronoun *we* in these examples is used generically. The sentence in (4.13) has the reading that includes any human beings, and as such *we all* could be replaced by *one*. In Holmberg's (2010b: 208) words, "it follows from the inclusive property of the null G-pronoun that it has human reference". Humans must then be included in the reference. The latter example refers to the local people who live in England. It can only be truthfully uttered by someone who lives in England, and if the addressee does not also live in England, it will not include the addressee in its reference. It is interesting to note that the generic quasi-inclusive *we* is similar to the generic exclusive *they* in the sense that both are anchored to a scene-setting adverbial. The difference is that the impersonal use of the pronoun *we* has a generic quasi-inclusive reading. Put differently, the unspecified people referred to by *we*, according to Kamio (2001: 1116), must fall into the speaker's 'territory of reference' in that *I* is the core member. The speaker considers himself as part of the group, but at the same time the sentence needs to be generalised, contrasting generic *we* with referential *we*.

4.3 Generalisations of Impersonal Pronouns in Thai

In Thai, just as in English, impersonal pronominal meanings are expressed through the use of personal pronouns. That is to say, a handful of existing definite pronouns are used for generic[26] reference (generic *inclusive* readings excluded). For example, the first person plural

---

[26] In a case like (i), it may seem that the pronoun *khǎw* 'they' is generic, referring to people in general:
(i) sìŋ   thîi   nayók càtôŋ tɔ̀ɔbkhamthǎam khɔn   thay dâay kɔ̂ɔ khʉa nayók chʉ̂a
    matter COMP PM   must answer   people Thai able is that PM   believe
    lúamây wâa   khrooŋkaan nay láksànà níi   cà   sáamâat chûay (*phuâk-khǎw) dâay ciŋ
    Q   COMP project   in kind DEM FUT capable help   them   able really
    'The Prime Minister must answer the Thai people as to whether or not he believes that kind of project can really help them.'
However, the fact that it is controlled by the definite antecedent in the main clause prevents it from having generic reference. Therefore, the interpretation is *khǎw/ khɔn thay* 'them/ the Thai people', not people in general. Cases like (i) should be distinguished from purely generic *one* (see Section 4.3.1 below), given the fact that a generic pronoun in Thai does not have an antecedent. The pronominal object in (i) is treated as a

pronoun (*phûak*)-*raw* 'we' is used for the generic quasi-inclusive interpretation; the third person plural pronoun (*phûak*)-*khǎw*[27] 'they' is used for the generic exclusive and arbitrary readings. Furthermore, generic NPs like *pkɔn* 'people' may be used interchangeably with *khǎw* 'they' to convey the generic exclusive meaning. It is significant for the G-pronouns that they do not have antecedents. Due to the absence of an antecedent, most G-pronouns have to be overt. This falls under the generalisation that null pronouns have to have an antecedent. As discussed, in the case of referential pronouns which are null, there is either a c-commanding antecedent or a discourse antecedent available for a referential interpretation of the null pronoun. On the other hand, to express a generic inclusive meaning, a null spell-out is obligatory (see Section 4.3.1 below), as a purely inclusive generic overt pronoun does not exist in Thai. The pronoun *khun* 'you' cannot be generic either, but is restricted to referential contexts. The impersonal uses of personal pronouns in Thai are summarised below:

Table 8: *The impersonal uses of personal pronouns in Thai*

Personal Pronouns	Types of Readings				
	referential	generic inclusive	generic quasi-inclusive	generic exclusive	arbitrary
khun 'you'	✓ (2)				
khǎw 'they'	✓ (1)			✓ (2)	✓ (3)
raw 'we'	✓ (2)		✓ (3)		

Note: 1 = obligatorily null; 2 = either null or overt; 3 = obligatorily overt

Kitagawa & Lehrer (1990: 753) claim that "the extension of the second person pronoun to an impersonal is possible only in languages with small, closed pronoun sets."[28] As stated above, it appears that Thai exhibits a restriction on the impersonal use of the second person pronoun in a similar way to Japanese and Korean (which are discussed by Kitagawa and Lehrer). In particular, the Thai pronouns are open-class items because (i) proper names and titles can be

---

referential third person null pronoun, and the reason why it is obligatorily null will be discussed in detail in Chapter 5.

[27] *Khǎw* is an abbreviated version of *phûak-khǎw* 'they' which is commonly used in spoken Thai. A context is then essential to determine whether *khǎw* is really 3SG or 3PL.

[28] However, Cabredo Hofherr (2010: 4) shows that Kitagawa and Lehrer's claim cannot account for languages like Farsi where pronouns are closed-class items, yet, no impersonal use of the second person pronoun is found.

used with a pronominal function, and (ii) the existing referential pronouns encode a variety of features relating to social and formality factors. If Kitagawa and Lehrer are right, this richness of pragmatic information could be what blocks the use of the second person pronoun as a generic pronoun. The following example illustrates that the only possible interpretation is with a referential reading, even if this interpretation is pragmatically odd:

(4.15) díawníi   ŋaan hǎa yâak   mâak thâa khun mây cɔ̀b   trii
        nowadays job   seek difficult very   if   you   NEG finish BA
     (Lit.) 'Nowadays to seek a job is difficult if you (the addressee) haven't finished a BA,'
      = 'It's difficult to find a job these days if you don't have a BA'
       Note: not '...if one doesn't finish a BA.'

Based on the distribution of personal pronouns for an impersonal use, I shall discuss each of them in terms of their readings, namely the generic inclusive, generic quasi-inclusive, generic exclusive, and arbitrary readings.

4.3.1 Generic inclusive and generic quasi-inclusive readings

In Thai, I claim that the generic inclusive reading has a null pronominal form as the default, which is roughly equivalent to the English generic inclusive *one*. When an overt pronoun appears, the interpretation shifts to referential. Thus if one wishes, for instance, to express a generic inclusive meaning, the pronominal argument has to be left unpronounced. Concerning the meaning of the inclusive G-pronoun, Moltmann (2006: 258) writes: "... a speaker draws a generalisation by applying the predicate to any (human) being as if that human being was himself, making both generic statements and, in a way, reference to himself." For instance, (4.16) implies that the speaker might be experiencing a difficulty in finding a job. However, the same reading implies that the inclusion of the speaker is not valid for (4.15), since the reading is made explicit in the definite, referential pronoun *you*.

(4.16) díawníi   ŋaan hǎa yâak   mâak thâa Ø mây cɔ̀b   trii
        nowadays job   seek difficult very   if   NEG finish BA
     (Lit.) 'Nowadays to seek a job is difficult if *one* has not finished a BA.'
      = 'It's difficult to find a job these days if you don't/ one doesn't have a BA.'

(4.17) bâanmɯaŋ cà   nâayùu   Ø tɔ̂ŋ   chûay ráksǎa khwaamsàʔàad

      city   FUT liveable   must help keep cleanliness

      'The city will be liveable if *one* keeps it clean.'

Generic inclusive null subject pronouns in Thai, labelled with Ø in the examples, represent a quasi-universal set of human beings, corresponding to the readings of English *one* or *anyone*. The generic reading in each example appears non-restricted, in that it can refer to anyone or people in general and has inclusive [human] and plural interpretation. (4.18) and (4.19) exemplify that apart from the fact that a null inclusive G-pronoun takes subject position, possessive pronouns and object pronouns, too can be null with a generic inclusive reading.

(4.18) kaankoohòk sǎamâat nam paysùu kaansǐa chûɯsǐaŋ   Ø

      lie   can   bring DIR   harm   reputation of-GEN one

      'Lies can harm *one*'s reputation.'

(4.19) pleeŋ khláassìk chûayhây Ø phɔ̀nkhlaay

      music classic   help   relax

      'Classical music helps *one* relax.'

As the literal translation indicates, the null pronouns are generic inclusive. The sentence in (4.19) could therefore be continued e.g. by a Thai phrase equivalent to the English phrases...*and for this reason I want to buy Beethoven's collected works* or *... so I guess that's why you have a lot of Mozart's music*. The missing object can be understood as a pronominal element having minimally the speaker and the addressee included, and thus yielding a plural reading.

A quasi-inclusive G-pronoun in Thai, just like that in English, does not necessarily include the addressee. Therefore, it has a narrower scope of interpretation than do the null pronouns that have the generic inclusive reading.

(4.20) *(raw) phûut phaasǎathìn

       we speak dialect

       'We speak dialect.'

(4.21) *(raw) kin   cee   nay dɯan   tùlaakhɔm

      we   have vegetarian food in   month October

   'We have vegetarian food in October.'

(4.22) *(raw) rák   naylɯ̆aŋ

      we   love king

   'We love the king.'

These examples exemplify instances of a generic quasi-inclusive reading of first person plural pronouns, corresponding to the quasi-inclusive G-pronoun 'we' in English. It appears that the pronouns are obligatorily overt. A null subject in these examples would result in a referential reading, which in the absence of a context would specifically be 'I'.

4.3.2 Generic exclusive and arbitrary readings

The interpretation of generic exclusivity excludes the speaker and the addressee. It exhibits third-person orientation, and such a third-person orientation applies to arbitrary referents as well. In Thai, an exclusive G-pronoun tends to be found in the active voice in sentences containing a scene-setting adverbial, as seen in the following examples:

(4.23) *thîi mùubâan níi*   (khǎw) mây kin nɯ́a kanleey

      at   village DEM   they NEG eat meat at all

   'In this village they don't eat meat at all.'

(4.24) *bon kɔ̀   níi* sùanyài (khǎw) plùuk chaa khǎay

     on island DEM mostly   they grow tea   sell

   'On this island they mostly grow and sell tea.'

The generic exclusive and arbitrary readings both exclude the speaker and addressee. A difference is that an exclusive G-pronoun having a universal reading must be licensed by a locative, similar to the situation in English, as discussed in Cabredo Hofherr (2003). This marks a contrast with the arbitrary reading, which is independent of a locative. The fact that the generic exclusive reading of the third person pronoun (*phûak-*)*khǎw* (PL) 'they' must be

anchored to a specified place explains why it is not obligatorily overt, but can optionally be null. If the locative expression is removed, the sentences in (4.23) and (4.24) will instead be understood as having referential readings.

When the personal pronoun *kháw* 'they' is used for an arbitrary interpretation, it can have two major readings, i.e. a vague existential reading, exemplified below in (4.25), and a corporate reading, exemplified below in (4.29).

(4.25) *(kháw) bɔ̀ɔk wâa   kruŋthêep náam cà   thûam
        They say   COMP Bangkok   water FUT flood
    'They said that Bangkok would be hit by floods.'

(4.26) *(mii khɔn) bɔ̀ɔk wâa   kruŋthêep náam   cà   thûam
        has person say COMP Bangkok   water   FUT flood
      'They said that Bangkok would be hit by floods.'

(4.27) *(mii khɔn) khɔ́   pràtuu yùu
        has person knock door   now
      'Someone's knocking at the door now.'

(4.28) kháw_{ref/ *arb} khɔ́   pràtuu yùu
        they   knock door   now
    'They're knocking at the door now.' (well-formed for referential readings only)
     Note: ≠ someone's knocking at the door.

(4.29) *(kháw) khûn raakhaa khày ʔìiklɛ́ɛw
        they   raise price   egg again
    'They have raised egg prices again.'

(4.25) and (4.26) show that the sentences have exactly the same vague existential arbitrary readings. (4.26) also shows that Thai uses a construction of *has + generic NP* to express an existential arbitrary reading with or without temporal anchoring, which can be paraphrased by an expression with the quantifier 'someone'. One difference between Thai and English arbitrary *they* is that Thai uses the personal pronoun *kháw* 'they' in the specific existential sense sparingly, mostly when it is the subject of the *say*-verb types, as seen in (4.25). Rather,

the personal pronoun *kháw* 'they' tends to be reserved for a referential use. This explains why (4.28) is not well-formed for an arbitrary reading. For an existential arbitrary reading, *has + generic NP* construction, as in (4.27), is used instead. (4.29) shows another sub-type of arbitrary *kháw* 'they', which has a corporate reading. Such a corporate reading necessarily appears with a predicate that presupposes a designated group performing a typical action, according to Cabredo Hofherr (2003). That is, in (4.29) the activity of selling eggs is understood as being carried out by merchants and the like. Thus the predicate behaves like a collocation, and it is then understood that *kháw* 'they' has a corporate reading, rather than a referential one.

Arbitrary readings are also typically found in the passive voice with a necessarily null subject. In the Thai passive construction, the referent of a null arbitrary pronoun can be a voluntary agent. The passive construction undergoes demotion of the subject which in turn makes the subject syntactically null. Nonetheless, not all passives have a null subject in the syntax,[29] as in (4.30). In a passive sentence where the null pronominal subject is arbitrary, it will be equivalent to an active sentence where the subject is generic NP, third person, as in (4.31):

(4.30) thîi naa   díawníi   thùkk khǎay yé
        field paddy nowadays PASS sell   a lot
        (Lit.) 'Paddy fields nowadays are sold a lot.'
       ='Nowadays paddy fields are being sold by many people'

(4.31) *(khɔn/ kháw) kháay thîi naa   (*khɔ̌ɔŋ   khǎw) yé   díawnîi
         People/ they sell   field paddy of-GEN they   a lot nowadays
         (Lit.) 'People/ They sell their paddy fields a lot nowadays.'
        = 'Nowadays many people are selling their paddy fields'

---

[29] An immediate test to verify the presence of a null agent is to add an agent-oriented adverbial to a passive. In (4.30) we get an ill-formed sentence when such an adverbial is inserted sentence-finally. Thus, there is no syntactically active agent in (4.30). (i) illustrates this point:
  (i) * thîi naa   díawníi   thùkk khǎay yé   yàaŋmâytemcay
        field paddy nowadays PASS sell   a lot unwillingly
        Lit. 'Nowadays many paddy fields are sold unwillingly.'

In a passive sentence containing an agent-oriented constituent with generic time-reference, a generic quasi-inclusive reading is also involved. In such a case, a passive may be ambiguous between arbitrary and generic quasi-inclusive readings. For example, (4.32) is ambiguous. Under one interpretation it is a statement about the properties of a new car, excluding any views on the part of the speaker and the addressee, which is equivalent to the English sentences 'It is a pleasure to drive a new car.' or 'Driving a new car is a pleasure.' The other interpretation assumes that the null subject is a null-experiencer argument, which includes the speaker in its reference. It also implies that the speaker has experience of driving a new car:

(4.32) rɔ́d màymày dâay   khàp lɛ́ɛw dii   ciŋciŋ
        car new   PASS   drive then pleasant   very
        'It's very pleasant when one is driving a new car.

Here, the Thai agent-oriented adjective, namely *dii ciŋciŋ* 'very pleasant', indicates that there is a null subject in the syntax. Such an adjective is assumed to introduce an understood null experiencer argument, as it can be continued with a phrase, like *rɔ́d màymày dâay khàp lɛ́ɛw dii ciŋciŋ sámlàb chăn* 'It's a pleasure for me that a new car is driven' (equivalent to 'It's a pleasure for me to drive a new car.') I do not pursue this matter further here, leaving detailed discussions to future work.

The arbitrary readings can also be implied in a sentence in which the described event is *hypothetical*, apart from a generic inclusive reading, as observed in Icelandic zero impersonals by Sigurðsson and Egerland (2009). It may then be ambiguous between the arbitrary and the generic inclusive reading:

(4.33) Á tunglinu væri ferðast á báti.
        on moon.the were traveled on boat
        'One/ they would travel on a boat on the moon' Icelandic: Sigurðsson & Egerland
                                                                    (2009: 175)

In Thai, it is not ambiguous, as the difference between an arbitrary reading and a generic inclusive reading of a pronoun is clearly distinguished by whether it is spelled out. To describe a hypothetical event in Thai, there are two options. The overt subject exemplifies a

generic exclusive/ arbitrary reading. The null subject exemplifies a generic inclusive reading only; it cannot have any other impersonal reading:

(4.34) thâa náamthûam lôok   Ø/ khɔn   khɔ̂ɔkhɔŋ nǐipay yùu bɔn yɔ̂ɔd khǎw

    if   flood   earth   people would   flee   live on top mountain

    'If there is a great flood inflicted upon earth, one/ people/ they will flee to mountain tops.'

4.4 Agreement and the Occurrence of Overt Pronouns

Since Thai allows pro-drop, a null pronominal argument may have either a referential reading or a generic inclusive reading, for instance:

(4.35) tham dii   dâay dii   (referential or generic inclusive)

    do   good get   glory

    (Lit.) 'One/ I/ you$_{ref}$/ (s)he do(es) good deeds and get(s) glory [in return].'

(4.36) cim bɔ̂ɔk wâa   tham dii   dâay dii

    Jim say   COMP do   good get   glory

    (Lit.) 'Jim says that he/ one does good deeds and gets glory [in return].'

(4.36) has the matrix subject to filter out a number of potential antecedents, unlike (4.35). If (4.35) appears in isolation, there are more possible pronominal subjects each of which could equally well be an antecedent. Having said that, the embedded subject in (4.36) is still ambiguous between referential reading and generic inclusive reading. This is due to the fact that the verbal expression is not inflected for person and number. The examples also reveal that complementary distribution of third person definite, referential null subject and generic null subject is not found in Thai and other discourse pro-drop languages in general (see also Examples (4.37) and (4.38)). Even though there is a third person matrix subject in both (4.37) and (4.38), the interpretation of the embedded null subject is ambiguous. The intended reading then mainly depends on the context and the interpretation of the pronominal argument. This indicates that Thai does not fit into the generalisation proposed by Holmberg (2010a,b), according to which what he calls consistent pro-drop languages never have a null

inclusive G-pronoun (see also Chapter 6, Section 6.6). Other discourse pro-drop languages, like Japanese, Korean and Chinese also behave like Thai, for instance:

(4.37) Ah John waa hai Yinggwok yiu   gong Yingman   Cantonese

     prt John say in   England   need speak English

    'John says that one/he needs to speak English in England.'

(4.38) John-wa kono beddo-de-wa   yoku nemu-reru-to-iu.   Japanese

    John-top this   bed-in-top well sleep-can-comp-say

    'John says that one/he can sleep well in this bed.'   Holmberg et al. (2009: 79)

The possibility of a generic inclusive reading of a null pronominal subject in these languages suggests that the available generic operator (see Section 4.5 below) is able to bind such pronominals. The fact that a generic inclusive reading is allowed in discourse pro-drop languages, but disallowed in consistent pro-drop languages also suggests that it is the rich referential verbal agreement found in the latter that blocks the binding by the generic operator of a pronominal element. On the contrary, nothing prevents such an operator from binding a null pronominal entity in the former.

4.5 Properties and Internal Structures of G-pronouns

We have seen that Thai allows null pronouns in both subject and object position, but not in all contexts. Basically, Thai allows null pronouns because there are antecedents with which they are able to be coindexed. In the case of G-pronouns, there are no antecedents. Therefore, they have to be overt generally. This does not include an inclusive G-pronoun in Thai which is null, despite having no antecedents. This section will be a discussion and explanation of the reasons why inclusive G-pronouns in Thai are not overt.

To account for the fact that a null inclusive G-pronoun is obligatory, I propose that there is a generic operator in sentence-initial position available to bind any type of G-pronouns. According to Moltmann (2006), the generic operator is the carrier of generic force, which is maximally general: the speaker, the addressee, and other people inclusive. In other words, the operator does not provide any features other than a generic feature (value). Carrying a

feature of genericity, the operator is called a generic operator in the CP-domain. I do not need to make a claim as strong as that. Instead of postulating that the operator has a generic inclusive feature, the generality of the inclusive reading comes from the absence of restriction on the φ-features of a pronoun. That is to say, the inclusive reading does not come from the operator but from the absence of restriction. This means that if a generic-operator bound pronoun has no φ-features, it will have a generic inclusive reading, including the speaker and the addressee. It follows that it will not be pronounced, as it is φ-featureless. Correspondingly, if a generic-operator bound pronoun has φ-features, for instance, [3PL] the reading of the pronoun will be more restricted: generic exclusive, and it has to be pronounced.

To show that a null pronoun without an antecedent is bound by a generic operator, I illustrate the case of two null pronouns appearing in the same sentence. Since they are in a general statement, these two pronouns are understood as being generic inclusive, referring to a quasi-universal set of individuals:

(4.39) thâa Ø tham khɔ̂ɔsɔ̀ɔb thùuk thúk khɔ̂ɔ dâay Ø kɔ̂ɔ kèŋ là

      if do exam correct every item able then smart PRT

    'One is very smart if one can answer every exam question correctly.'

The example does not have the reading, 'People in general are very smart if people in general can answer every exam question correctly'. Instead, it tends to have the following reading:

(4.40) thâa khɔn tham khɔ̂ɔsɔ̀ɔb thùuk thúk khɔ̂ɔ dâay khǎw kɔ̂ɔ kèŋ là

      if people do exam correct every item able they then smart PRT

    'People are very smart if they can answer every exam question correctly.'

The bound variable reading in (4.40) would have the analysis in (4.41):

(4.41) For every x who is a member of people in general (x is smart if x can answer every exam question correctly).

This exemplifies an instance of two variables (which do not c-command each other) bound by the same operator,[30] rather than coreference between two pronouns. I will argue that null pronouns in Thai do not refer by themselves. They are φ-featureless, carrying just a general nominal feature [N] and an unvalued reference feature [uR].[31] The null pronoun in the main clause, then, cannot be an instance of referential antecedent argument. Therefore, there are no coreference effects, but there are instances of binding effects. By analogy to Diesing's (1992) theory of the quantificational variability of indefinites,[32] a generic operator behaves similarly to a quantificational adverb 'generally'. The variables introduced by the null pronominal elements are bound by an implicit quantificational adverb serving as the operator. In other words, according to Diesing, adverbs are able to bind variables introduced by indefinite arguments. The representation of quantificational adverb binding variables is shown in (4.42). (4.43) illustrates that the adverb 'generally' does not quantify over events or acts, but over the null inclusive G-pronouns:

(4.42)

Generally x

Ø

$[x]$ ...   Ø

$[x]$
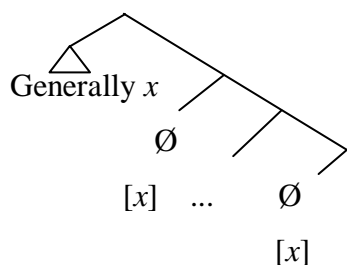
(4.43) GENERALLY$_x$ [x kôɔ kèŋ   là   thâa x tham khôɔsɔ̀ɔb thùuk thúk khôɔ dâay]
                    then smart PRT if   do   exam   correct every item able

The two null pronouns, labelled with x, are interpreted as bound, since there is an implicit binder in Spec, CP, i.e. the generic operator. After the binding, they would have generic inclusive readings with the interpretation that 'it is generally true for x'.

---

[30] I will argue in Chapter 6 that the second null argument is essentially bound by a higher generic argument (see Chapter 6, Section 6.3.1.1) due to the fact that the second generic NP is obligatorily null regardless of its reading.

[31] See Chapter 6, Section 6.1 for a discussion of the internal structure of null pronominal arguments in Thai and a definition and characterisation of [uR] feature.

[32] Diesing develops the theory, following Kamp's (1981) and Heim's (1982) theory of NP interpretation, based on the idea that an indefinite, like *a book*, is not a quantified expression, and has no quantificational force of its own. Instead, it introduces a variable in the structure and the variable is bound by a quantificational adverb as the operator, which in turn gives it quantificational force. The LF of (i) is represented in (ii):
(i) A cellist seldom plays out of tune. (ii) Seldom$_x$ [x is a cellist] x plays out of tune.   Diesing (1992: 8)

In what follows, I shall discuss in detail the internal properties of null pronouns with a generic inclusive reading, followed by those of null pronouns with generic quasi-inclusive reading, exclusive generic and arbitrary reading of pronouns respectively.

4.5.1 Inclusive G-pronouns

A Thai inclusive G-pronoun is informally characterised as a null pronominal of the type *pro*. The main questions are: what features does it have? and why can it not be overt like other G-pronouns? The null inclusive G-pronoun tends to have the following properties: (i) when it is generic operator-bound, it has the most *general* reading: the speaker, the addressee, and other people inclusive, (ii) it has no topic antecedent reading, and (iii) it cannot be replaced by an overt pronominal, due to lacking a lexical content (i.e. no overt pronoun in Thai corresponds to the English generic pronoun 'one'). As far as these properties are concerned, it does not belong to the type *pro* in the binding theory sense, since it cannot refer independently. Being a minimal nominal category, made up of just [uR, N] features, the null pronominal entity must be locally bound by a generic operator in the CP domain to receive the generic reading. Having no φ-features, the null pronoun has no restriction on the interpretation, and it necessarily ends up having a generic inclusive reading after being probed by the generic operator. It therefore refers to people in general, necessarily including the speaker, addressee, and anyone else. The representation of the features of Thai inclusive G-pronouns is schematised in (4.44). (4.45) and (4.46) represent the derivation of operator-binding.

(4.44)   *pro*   (where *pro* is simply the label for this set of features)

$$\begin{bmatrix} uR \\ N \end{bmatrix}$$

(4.45) *Before generic operator-binding*
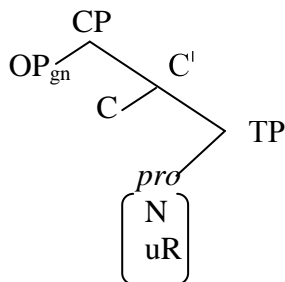
(4.46) *After generic operator-binding*
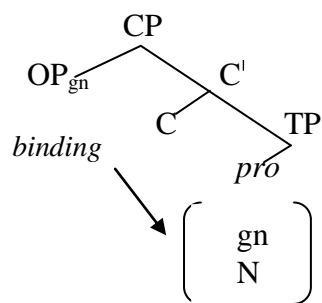
The above representations illustrate the generic operator-binding hypothesis. The null argument labelled *pro* has the feature [uR], and can then be bound by a generic operator.[33] Bound by the operator, it obtains the most general (unrestricted) reading, which is generic inclusive due to there being no restriction on the reference. For sentences like (4.17) and (4.19), a representation would therefore be as in (4.47) and (4.48) respectively:

(4.47) $[_{CP}[OP_{gn}]_1 [_C [_{TP}$ bâanmɯaŋ cà nâayùu [pro_{gn}]$_1$ tôŋ chûay ráksǎa khwaamsàʔàad]]]

               city   FUT liveable   must help   keep   cleanliness

    'The city will be liveable if one keeps it clean.'

(4.48) $[_{CP}[OP_{gn}]_1 [_C [_{TP}$ pleeŋ khláassìk chûayhây [pro_{gn}]$_1$ phɔ̀nkhlaay]]]

               music classic   help   relax

    'Classical music helps one relax.'

Syntactically, an inclusive G-pronoun and a bound variable behave similarly in that both are bound pronouns. Thus, the interpretation of the pronominal subject in the complement clause in (4.49) depends on the antecedent in the matrix clause:

(4.49) mɔ̌ɔ$_1$ thúk khɔn thít wâa *(khǎw$_1$)[34] chàlàat

    doctor every CLS think COMP (s)he   smart

    'Every doctor thinks that *(s)he* is smart.'

---

[33] Since the null argument has no referential index, it is referentially defective, and thus has a generic inclusive reading after being probed by the generic operator.

[34] Bound pronouns are not obligatorily null in Thai unless they have an inclusive generic reading. In fact, the pronominal *khǎw* '(s)he' in (4.49) can be, and must be, overt. It can even be replaced with the reflexive pronoun *tuaʔeeŋ* 'oneself' in the embedded subject position. This appears to contradict Montalbetti's (1984) claim that bound pronouns are obligatorily null in pro-drop languages generally.

Just like bound variables, inclusive G-pronouns in Thai have to be bound. This seems to contradict Prince (2006) and Cabredo Hofherr (2010) who argue that a sentence exemplifying two inclusive generic *one*s can have a coreferent reading:

(4.50) When one₁ isn't careful, one₁ /*he₁ can catch a cold.   (Cabredo Hofherr 2010: 9)

According to Prince and Cabredo Hofherr, the coreferent reading applies when the generic *one* is repeated in the main clause. The ungrammaticality of *he* indicates the impossibility of coreference. This analysis implies that the coreference is possible iff there is a repetition of the generic *one*.

A similar analysis is done in Korean. Han (2006) analyses indefinite personal zero anaphors referring to people in general in Korean (equivalent to null G-pronouns in our terms), and concludes that if there are two or more null G-pronouns in one sentence, the relation between/ among them tends to be one of coreference, for example:

(4.51) U1: Ø holangi-lul cap-ulye-myen Ø san-ey   ka-ya-ha-nta.
              tiger-Acc   catch-Intend-If   mountain-Des go-must-PresDec.
          'If one wishes to catch a tiger, one must go to the mountain.'

      U2: kukes-to antoy-myen, Ø tongmwulwen-ey-lato ka-eyaha-nta.
          it-also   not-do-if   zoo-to-even   go-must-PresDec.
          'If that won't do, one must still go to a zoo.'   (Han 2006: 64)

It appears that G-pronouns in (4.51) are obligatorily null, corresponding to the English *one*. However, null arguments here cannot have coreferent reading. I claim that the null pronouns in discourse pro-drop languages start with [uR, N] in the syntax, and having no referential index/ generic value, they are therefore incapable of coreferencing (see also Chapter 6, Section 6.6). Interestingly, the null spell-out of the pronominal in (4.51) indicates that Korean has no overt inclusive G-pronoun equivalent to English generic *one*, supporting the above claim. That is to say, to express an inclusive generic reading, a subject pronoun must not be pronounced (hence a null pronoun: a zero pronoun, to use Han's term). As indicated by the translation, the interpretation of the null pronouns is restricted to a generic inclusive reading. This confirms that the null pronouns do not have a coreference relation, since coreference

allows pronouns to choose their (overt) reference from the discourse. The null pronouns in (4.51) do not show this effect. This means that Korean is typologically similar to Thai in that null pronouns in both languages cannot induce the generality of the inclusive reading on their own. Rather, they need to be bound by a generic operator in the CP-domain of the sentence containing it for a generic inclusive interpretation. The null pronoun, therefore, can be null because it is unrestricted.

4.5.2 Quasi-inclusive G-pronouns

Surprisingly, the last section shows that an inclusive G-pronoun must obligatorily be null, despite the fact that there is no antecedent. This phenomenon can be explained through the generic-operator binding together with the absence of ϕ-features of the pronoun, giving the unrestricted, generic inclusive reading. Therefore, it is not overt. In what follows, I discuss the quasi-inclusive G-pronouns and exclusive G-pronouns in terms of their internal properties to see why they are overt, based on the binding of the generic operator.

The pronominals *raw* 'we' and *khǎw* 'they' in Thai can be used either referentially or generically. When they are used generically, I propose that they behave like an inclusive G-pronoun in being bound by the same generic operator. Broadly speaking the generalisation is that Thai G-pronouns are generic operator-bound. Quasi-inclusive G-pronouns, exclusive G-pronouns and arbitrary pronouns are not as empty as inclusive G-pronouns, as they are restricted by [number] and [person] features. In other words, they have ϕ-features.

As for a quasi-inclusive G-pronoun, after being probed by the generic operator, it refers to people in general, but excludes the addressee. It also follows that such pronouns must be overt, since (i) they have no antecedent providing them with a referential index (except a generic reading as a result of generic-operator binding), and (ii) if they were null, they would be indistinguishable from the unrestricted inclusive G-pronoun or ambiguous with the referential 'I'. The internal features of pronominal *raw* 'we' before binding is the following:

raw 'we'
$$\begin{pmatrix} \text{uR} \\ \text{N} \\ \text{1PL} \end{pmatrix}$$

Consider this example:

(4.52) *(raw) rák lûuk   (khɔ̌ɔŋ   raw) sàmə̌ə

          we love children of-GEN   we   always

          'We always love our children.'

This example distinguishes two major readings that *we* may have, i.e. a referential reading or a bound variable reading. Note that a grammatical possessive NP is comprised of a head noun plus a possessive form with no inflection. The pronominal *raw* then displays the same variant in the possessive pronoun and pronominal argument, thereby being an invariant word. With a null subject pronoun, (4.52) will have the following referential reading:

(4.53) I/ You love my/ your children always.

As mentioned, the null subject has the potential for ambiguity if it appears in isolation – it may even have a generic inclusive reading. If it is uttered out of the blue, the preferred interpretation of such a null pronoun will be deictic, referring to a discourse participant, particularly the speaker. With an overt subject pronoun, (4.52) has two possible readings. (4.54) exemplifies a referential reading; (4.55) exemplifies a bound variable reading:

(4.54) #All of us love all of our children always.

(4.55) For each x (including the speaker): x loves x's child(ren) always.[35]

However, my informants and I agree that (4.55) is the preferred option. I assume this is due to pragmatic reasons that block the coreference reading. (4.56) also shows that a possessive pronoun can optionally be null. In (4.56) such a possessive pronoun in direct object position which has been topicalized is optionally null, whereas the canonical generic subject is obligatorily overt.

(4.56) sǒmbàt (khɔ̌ɔŋ   raw) *(raw) kôo hǔaŋ   pen thammàdaa

          treasure of-GEN we   we then possessive BE normal

          (Lit.) 'Normally, treasures of us, we are possessive of.'

---

[35] In accordance with Diesing 1992, the reading can have the analysis in (i):

(i)   ALWAYS$_x$, if x has children, x loves x's children.

In other words, if there are two occurrences of a quasi-inclusive G-pronoun, both pronoun and possessive pronoun in the same sentence, then what is expected is that the quasi-inclusive G-pronoun in the canonical subject position must be overt. The possessor in object position which is topicalized is optionally null, since the quasi-inclusive generic reading is already established by the pronominal canonical subject. In fact, generic pronouns (G-pronouns) in Thai are compatible with bound possessives. However, there are languages where G-pronouns are not allowed to be possessive pronouns. Take the case of Finnish, where a null inclusive G-pronoun cannot be possessor of NP, as discussed by Holmberg (2010b). He concludes that the ungrammaticality of the possessive G-pronoun lies in the fact that it violates Principle A.

(4.57) a. *Ø lapsensa   tuottaa aina   huolia.

        children-PX cause   always worry

    Intended reading: 'One's children are always a cause of worry.'

   b. Hänen   lapsensa   tuottaa aina   huolia.

    he-GEN children-PX cause   always worry

    'His children are always a cause of worry.'   Finnish: Holmberg (2010b: 224)

A grammatical possessive NP, according to Holmberg, consists of a head noun, namely *lapsi* 'children' with the possessive suffix –*nsa* (PX, a third person anaphor), and an overt possessor pronoun with GEN case, as can be seen in the anaphoric reading in (4.57b). If, for some reason, the null generic pronoun is not licensed in the possessor position, the possessive anaphor violates Principle A. Why would it not be licensed there, though? I propose, instead, that the explanation is that the possessive suffix functions like agreement: It has [uϕ] features in need of valuation.[36] But if the null inclusive generic pronoun is ϕ-featureless in Finnish just like in Thai, it cannot value the suffix. Consequently the possessive suffix is left with unvalued ϕ-features, violating Full Interpretation at LF (Chomsky 1995, 2001).[37]

---

[36] The possessive suffix is inflected for person and number (Holmberg, p.c.): minun lapse-*ni* 'my child-1SG', sinun lapse-*si* 'your child-2SG', etc. for all persons and numbers.

[37] Chomsky's principle of full interpretation, cited in Crystal (2008), states that "...there should be no redundant or superfluous elements in the presentation of sentence structure: each element must play a role and must be interpreted."

This generalisation, however, cannot hold true for Thai, as Thai pronouns do not inflect for case. More specifically, Thai does not have any possessive agreement element with unvalued φ-features, as shown by the Thai counterpart below:

(4.58) lûuk   mák   pen tôntɔɔ khɔ̌ɔŋ hùaŋhɛ̀ɛŋthùk
      children always BE   cause of   worry
      'One's/ my/ your children are always a cause of worry.'

Given the sentence appears in isolation, it is ambiguous, as indicated by the translation. Due to the analysis that a generic operator binds pronouns in Thai without constraints on agreement, I assume that it can similarly bind a possessor. The example below is also ambiguous between a generic inclusive reading and a referential one:

(4.59) ŋen   thɔɔŋ khɯa khɔ̌ɔŋ nɔ̂ɔk   kaay
      money gold BE   thing outside body
      'One's/my/your money and gold are not as important as one's/my/your life.'

The possibility of a generic inclusive reading indicates that a null pronoun in Thai, both in pronominal and possessive forms, is bound by a generic operator. Another possibility is a referential reading in which the null argument and the null possessor are understood as referring to a speech participant, either the speaker or the addressee. Therefore, Thai allows a null G-pronoun to be a possessor of NP in parallel with a referential possessive pronoun.

4.5.3 Exclusive G-pronouns and arbitrary pronouns

The personal pronoun (*phûak-*)*kháw* 'they' can have either a generic exclusive or an arbitrary reading after being probed by a generic operator. The generic exclusive reading of *kháw* refers to people in general, typically in some location, but not including the speaker or addressee. A generic exclusive reading of *kháw* 'they', therefore, typically has locative anchoring. A stronger claim is that in the absence of a locative adverb, the pronoun cannot have a generic exclusive reading. Instead, it would have an ill-formed referential third person reading when overt due to having no antecedent. Correspondingly, in the presence of the locative, the pronoun is not pronounced. A null exclusive G-pronoun and an overt counterpart

are then accounted for by some different internal mechanisms. In fact there are three possible options. First, when an exclusive G-pronoun is overt, it behaves similarly to a quasi-inclusive G-pronoun in that it is bound by a generic operator in Spec, CP. It has more restricted reference, thereby appearing overt. The internal features of pronominal *kháw* 'they' before generic-operator binding would be:

$$\text{kháw 'they'} \begin{bmatrix} \text{uR} \\ \text{N} \\ \text{3PL} \end{bmatrix}$$

Bound by the generic operator, the pronoun therefore has a generic exclusive reading, and is pronounced, since it has ϕ-features:

(4.60a) $[_{CP}[\text{OP}_{gn}]_1$ $[_C$ $[_{TP}$ bon kɔ̀   níi   sùanyài *(kháw$_{gn}$) plùuk chaa khǎay]]]

                      on island DEM mostly   they   grow tea   sell

     'On this island they mostly grow and sell tea.'

Second, if it is null,[38] I assume that the locative adverbial introduces an implicit argument 'people' or human entities, depending on the context. Thus *bon kɔ̀ níi* 'on this island' means (or can mean) 'people on this island', and this phrase behaves just like an antecedent. Therefore, the null pronoun essentially has its antecedent. The interpretation of the null subject can only be a third person plural reading (see Brody (2011) for a similar idea).

(4.60b) (khɔn$_1$) bon kɔ̀   níi   sùanyài e_1 plùuk chaa khǎay

     people on island DEM mostly   grow tea sell

     'On this island they mostly grow and sell tea.'

Third, an overt generic NP can take the subject position. If this is the case, then the locative adverbial no longer encodes the interpretation 'people on this island'. The reading is still generic exclusive:

---

[38] The fact that an exclusive G-pronoun can be null applies to Korean as well, for instance:
(i) tongmwulwon-eyse-nun Ø   halu-ey twu-kki meki-lul   cwu-unta
     zoo-at-Top   (Subj) day-in two-class fodder-Acc give-PresDec
     'In zoos (they, e.g. zoo-keepers) give fodder twice a day.'   Han (2006:61)
The fact that the subject pronoun is not pronounced means that the sentence in (i) resorts to Option 2.

(4.60c) khɔn bon kɔ̀ níi sùanyài plùuk chaa khǎay

    people on island DEM mostly grow tea sell

    'On this island they mostly grow and sell tea.'

As for a pronoun having an arbitrary reading, it must obligatorily be pronounced. In fact, I assume that a corporate reading of an *arbitrary* pronoun presupposes generic-operator binding, since it yields a non-referential, plural reading. The generic operator does not account for other readings of arbitrary pronouns where a quasi-existential (singular) reading is allowed (cf. (4.25)). (4.61) is an example of arbitrary *kháw* 'they' having a corporate reading, and (4.62) shows its logical form:

(4.61) A: tɔɔní námman phɛɛŋ mâak

    now oil   expensive very

    'Now oil is very expensive.'

 B: hén *(kháw) bɔ̀ɔk wâa (*kháw) cà phɔ̂əm kaanphàlìt námman (*khɔ̌ɔŋ kháw)

   hear they say COMP they FUT increase production oil   of-GEN they

   'I heard that they said they would increase their oil production.'

(4.62) [$_{CP}$ [gn] [$_C$[$_{TP}$ ..kháw$_{gn}$ bɔ̀ɔk wâa kháw$_{gn}$ cà phɔ̂əm kaanphàlìt námman

   khɔ̌ɔŋkháw$_{gn}$]]

The predicate here restricts the reference choice to a corporate arbitrary reading. The designated subject would be such particular groups as *oil companies*. The fact that the subject pronoun *kháw* 'they' is obligatorily overt is because of its restricted φ-features and to prevent it from having a generic inclusive reading. On the other hand, the second and third occurrences of the pronominals can be, and must be, null, since they are understood as sharing the same interpretation as the pronominal subject which is already introduced.

Having discussed the internal properties of different types of non-referential pronouns, we can conclude that an operator-bound inclusive G-pronoun in Thai is obligatorily null due to the fact that it has no restriction on the interpretation at all, so it receives the default reading of inclusive genericity. Nonetheless, all generic pronouns have one thing in common in that they are operator bound. The generic operator is the same operator in all different types of generic expressions. If there is no overtly expressed restriction on the operator, the default restriction is humans and if the pronoun has no overtly marked person feature, the pronoun

has unrestricted clusivity, i.e. it is inclusive. Other non-inclusive G-pronouns, including quasi-inclusive G-pronouns, exclusive G-pronouns, and arbitrary pronouns having a corporate reading, cannot be null, given that they are probed by a generic operator. Instead, they have to be overt, since they have more restricted reference. That is to say, they have ɸ-features. This is contrary to Barbosa's (2011: 577) claim that "there is a correlation between lack of full ɸ-feature specifications and the availability of a generic null subject". The availability of a generic reading in Thai appears to be contingent on the presence of an implicit generic operator, regardless of whether a pronoun has fully fledged ɸ-features.

4.6 Indefinite Topics

This section deals with a type of inter-sentential anaphora involving null indefinite topics in Thai. It entails neither first-person nor second-person orientation, so sentences with an indefinite topic can only be interpreted as having exclusive readings.

This type of null topic, which is equivalent to the English 'such people', is special. Although it is used to refer to a class of people, it is not definite. Besides, it is anaphoric, and it is not quantificational. In addition, it behaves more like an R-expression than a pronoun, given the binding principles. Lastly, it typically refers to a group of human entities that have been already mentioned. It is not clear, though, that 'such people' is indefinite in the same sense as pure indefinites, like *a book*, *people*, and so on. Despite being neither specific nor indefinite, I will continue to refer to 'such people' as an indefinite topic. Consider this example:

(4.63) [khon thîi   pen choléstoorân sǔuŋ]$_1$ khuan duulɛɛ   rûaŋ   ʔahǎan (*khɔ̌ɔŋ ton)
         person COMP have cholesterol   high   should take care matter food of- GEN self
         mûa (*phûak-khǎw$_1$) mii weelaa (*phûak-khǎw$_1$) tɔ̂ŋ pháyayaam ʔɔ̀ɔkkamlaŋ bâaŋ
         when they   have time   they   must try   exercise   sometimes
         'Anyone whose cholesterol level is high should take care of their food intake. When such people have free time they should try to exercise.'

The subjects in the second sentence are obligatorily null. They are understood as being coindexed with the indefinite discourse topic *khon thîi pen choléstoorân sǔuŋ* 'anyone whose cholesterol level is high'. In fact, the expression 'such people' in (4.63) can be more

complicated, as it may be reconstructed to be a predicate, namely 'they are such people'. Such a predicate indicates that a Thai null spell-out of *such people* is referring to an indefinite topic. It appears then that the non-definite subject in the opening discourse forms a natural topic. This reconstructed constituent is optionally pronounced:

(4.64) khon   thîi   pen choléstoorân sǔuŋ khuan duulɛɛ   rûaŋ   ʔahǎan (*khɔ̌ɔŋ ton)
       person COMP have cholesterol high should take care matter food of- GEN self
       (*khon   pen bɛ̀ɛbníi*)   mûa (*phûak-khǎw) mii weelaa (*phûak-khǎw) tɔ̂ɔŋ
       person BE such person when they   have time   they   must
       pháyayaam ʔɔ̀ɔkkamlaŋ bâaŋ
       try   exercise   sometimes
       (Lit.) 'Anyone whose cholesterol level is high should take care of their food intake.
       They are such people; when they have time, they should try to exercise.'

Although I recognise the complication, I will not take the whole VP *khon pen bɛ̀ɛbníi* 'they are such people' into an analysis, as it is beyond the scope of the present study. Rather, I focus here on the null indefinite topic NP, as it constitutes a part of discourse pro-drop in Thai. The first attempt at the analysis is to challenge Moltmann's (2006) bound variable hypothesis. By analogy with the sentential generic operator, the null possessor in the first sentence is assumed to be bound by *khon* 'person', which behaves like an overt operator. The null pronominal subjects in the subsequent sentences, however, cannot be bound in the same way as the null possessor. We have (4.65) as the LF of (4.64):

(4.65) [$_{CP}$ khon$_1$ [$_{TP}$ [$_{NP}$ t_1] thîi pen choléstoorân sǔuŋ] khuan duulɛɛ rûaŋ ʔahǎan [pro_1]]]
       [$_{CP}$ [OP ??] [$_{TP}$ mûa [pro] mii weelaa [$_{CP}$ [OP ??] [$_{TP}$ [pro] tɔ̂ɔŋ pháyaam ʔɔ̀ɔkkamlaŋ bâaŋ]]]

The binding[39] scope of *khɔn* 'person' is its c-command domain. Thus, the widest scope must be within a sentence. The null arguments in the second sentence cannot be bound by the same operator in the Spec, CP of the first sentence. This phenomenon poses a problem for any theory which excludes inter-sentential binding.

---

[39] The relation between the antecedent and the anaphor cannot be an instance of coreference, since the antecedent is not definite, and thus lacks the specific values (person and gender) required for referentiality. Consequently, the null subjects cannot be interpreted referentially.

An alternative is to postulate that the obligatorily null subject is a copy of the indefinite topic operator in Spec, CP. (4.67) is the analysis of (4.66):

(4.66) nɔ́ɔŋ   thîi   tâŋcay   thamŋaan cà pen thîirák   khɔ̌ɔŋ phûanrûamŋaan
       person COMP conscientious work   FUT BE well-respected of-GEN colleague
       (*phûak-khǎw) dâay lûankhân   rew (*phûak-khǎw) mii ʔookàat thîi   diikwàa
       they   able get promoted fast they   have chance COMP better
       'People who are conscientious will be well-respected by their colleagues. Such people will get promoted fast. They will have a better chance.'

(4.67) [nɔ́ɔŋ thîi tâŋcay thamŋaan, i] ...[$_{CP}$ [i [N]] [C [$_{TP}$ T [$_{VP}$ [$\overline{i\,[N]}$] v VP...]]]]

Here the indefinite *nɔ́ɔŋ thîi tâŋcay thamŋaan* 'people who are conscientious' is the discourse topic, qualifying the discourse antecedent. The fact that the subjects cannot be pronounced with a pronominal form indicates that it is controlled via a null topic chain in the same way as a referential null topic (see Chapter 6, Section 6.3.1.2 for a detailed analysis of obligatorily null subjects as well as the derivation of [i [N]]). The topic chain itself is linked to the antecedent established in the opening discourse. Therefore, c-command is not presupposed. Following C.-T. J. Huang's (1984 and subsequent) proposal, the null subjects in the canonical positions are copies of the moved indefinite topic (or topic variables, to use C.-T. J. Huang's term) that moves to an abstract topic operator in Spec, CP, thereby referring to "the conscientious working people", and not to the colleagues nor any other topic.

4.7 Summary

In this chapter, I have offered an analysis of generic pronouns in Thai which explains why some of them are obligatorily overt while others are obligatorily null. Whether a generic pronoun can be overt is determined by the presence of ϕ-features. Bound by a generic-operator, a pronoun without ϕ-features will have a generic inclusive reading and is null, since there is no ϕ-features restriction on the generic operator. If it were overtly expressed using the personal pronoun *khun* 'you', it would lose its generic inclusive reading and become referential. On the other hand, quasi-inclusive G-pronouns, arbitrary pronouns with corporate readings, and exclusive G-pronouns are always overt when bound by a generic operator. This

falls under the generalisation that specified ϕ-features are always pronounced, to be discussed in Chapter 6. Null counterparts of the non-inclusive generic and arbitrary pronouns would be ambiguous between generic inclusive readings and referential readings, which in the absence of context would mean referring to the speaker. Since the quasi-inclusive G-pronouns, arbitrary pronouns under a corporate reading, and exclusive G-pronouns have ϕ-features, this results in the interaction between the overtly marked ϕ-features of a pronoun and a generic operator, as both account for what reading a pronominal entity could have and whether it can be pronounced. The phenomenon that the generic operator is one and the same for all generic pronouns, null or overt, which is language-specific, cannot support the idea that the operator-bound generic pronouns are null. As for a null subject having an indefinite topic reading, it is *prima facie* plausible to say that this is an instance of a copy of the moved indefinite topic in Spec, CP.

Chapter 5: The Interpretation of Referential Null Pronouns and Discourse Functions in Thai

We have seen in the earlier chapters that there are two main possibilities for a null pronoun in Thai in terms of its interpretation. First, it can be referential (deictic), referring to the speaker by default as a local antecedent. Second, it can be generic inclusive, i.e. a null pronoun bound by a generic operator. In this chapter, we will look at another interpretation that null pronouns can have, in particular referential third person null pronouns. I assume that the antecedent of a referential third person null pronominal argument is an Aboutness Topic of the discourse (see Section 5.2.2 below). Once such a topic is established in the discourse, a null subject or object in subsequent sentences will be interpreted as being coreferential with it. Another possible antecedent type of a referential third person null pronoun is an NP argument in a higher clause. The locality of the antecedents, i.e. c-commanding and discourse topic antecedents, will also be discussed, as will the question of whether they involve movement. This chapter shows that the occurrence of a null pronoun and its well-formedness (in terms of pragmatic felicity) involves the discourse functions of the elements in the sentence. Thus, this chapter will investigate the major characteristics of discourse functions in Thai. In doing so, the notions of topic and focus are investigated, by looking at how these discourse functions trigger the availability of *pro* and its overt counterpart.

5.1 Some Definitions of Topic and Focus

In this section, the conventional notions of topic and focus in terms of old/ given information and new information are elaborated first, followed by a discussion of why the notion of old and new information might be misleading. It is shown that a topic is not necessarily associated with old information. For instance, a topic can receive the salient focus of the utterance. This occurs when the topic is part of the background (old information) but is focalised. In addition, such terms as 'discourse-oldness' will be put forward. The different statuses of referents in relation to discourse-oldness and the speech participants will also be reclassified, resulting in a categorisation that goes beyond simply discourse-old and hearer-new, but also includes, for instance, discourse-new/ hearer-new (Prince (1981); Casielles-Suarez (2004)). These referential statuses will be explained using Lambrecht's (1994) system of discourse states of referents viewed in terms of identifiability.

Reinhart (1981) views topics as old information. Similarly, Erteschik-Shir (1997) proposes that all topics are old or presupposed. Vallduví (1990) restricts topics, which he calls links, to the sentence-initial position. Topics in Lambrecht's (1994) sense are defined as sentence-initial elements, often encoded in pronominal forms. He restricts topics to discourse referents and proposes the following definition: "A referent is interpreted as the topic of a proposition if in a given situation the proposition is construed as being about this referent i.e. as expressing information which is relevant to and which increases the addressee's knowledge of this referent." (Lambrecht (1994:131)). The general characterisation, according to these scholars, is that a topic is what the sentence is about, a different linguistic constituent from the focus, and often expressed by a pronoun (Vallduví (1990); Lambrecht (1994); Reinhart (1983); Erteschik-Shir (1997)). Its position is typically clause-initial.

The focus of the sentence or utterance is traditionally taken to be the new information, which is the centre of the communication interest. In most languages, foci are explicitly marked phonologically. To illustrate, consider a simple sequence like 'It wasn't me; SHE did it.' (adapted from Lambrecht (1994:48)). Here, SHE is the focus and receives contrastive stress. It packages significant linguistic means for conveying information structure that is beyond the literal meaning. It also conveys old information through the use of the anaphoric pronominal form *she* and the position of this pronominal subject, which is clause-initial. Also, the predicate *did it* refers to a portion of the preceding discourse. The sentence, nonetheless, expresses new information to the addressee through the prosodic focus of the subject *SHE*, signalling a contrast between *me* and *she*. Associated with new information, the focus moves the utterance forwards. For instance, in the reply to a wh-question a focused constituent represents a new item of information ('new information' in Lambrecht's term). The listener may then extend the conversation based on the focused constituent, thereby making it a topic in a subsequent utterance. To summarise, new information can give either totally new knowledge or modify the old information presupposed in a previous utterance. A topical constituent, on the other hand, connects the sentence to the previous discourse.

Related to the topic-focus dichotomy are two major analyses of information structure in the clause, i.e. Theme-Rheme (Halliday (1967); Calabrese (1986)) and Topic-Comment (Reinhart (1981); Gundel (1988)). Obviously, there are differences in how these authors define the notion of theme or topic, but their definitions, nonetheless, share the common

ground that a sentence is always split up into a topic or theme element and the rest of the sentence. Consider example (5.1):

(5.1) A: ceen pay nǎy

      Jane go   where

     'Where did Jane go?'

  B: *e*   pay tàlàad

    she go   market

   'She went to the market.'

The reply *e   pay tàlàad* 'She went to the market.'exhibits the typical information structure of a Thai sentence, which can be summarised as in the following table:

Table 9: *Typical information structure of a Thai sentence*

GIVEN/OLD		NEW
BACKGROUND/PRESUPPOSITION		FOCUS
e she	pay went to	tàlàad the market
TOPIC/ THEME	COMMENT/ RHEME	

The null subject pronoun is construed as coreferential with the discourse element *ceen* 'Jane' introduced in the question. It signals what the sentence is about, and it is within the domain of presupposition, i.e. knowledge that is presupposed. The presupposed knowledge is that there is a person called Jane who went somewhere. Thus, both the null subject referring to *Jane* and the main verb represent background information in the reply. The new information in the reply is the focused constituent *tàlàad* 'the market' (and a reply consisting of just this constituent would also be felicitous). This focused constituent, which is the directional complement, is an instance of narrow or information focus. Narrow focus, in Frascarelli's (2010: 2122) words means "focus on a single word or constituent". This is equivalent to Kiss's (1998) information focus, where new information is conveyed. On the other hand, broad focus is focus on a more complex constituent (VP, TP or CP), in which all parts are

introduced into the discourse at the moment of utterance. Typically, this would be the case in a reply to a question like 'What happened?' or 'What're you doing?' (Féry (2007), among others). Overall, the subject and the predicate in Table 9 are presented from given to new, and the order, with the topic found sentence-initially preceding the focus, is as expected. In (5.1), the subject pronoun in the reply is obligatorily null, given that it is coreferent with the discourse topic.[40]

However, the idea that a topic is associated with old information and a focus with new information is not always adequate to explain the referential status of elements in discourse in more complicated cases. An additional issue, as Casielles-Suarez (2004) notes, is that it is not obvious that old information can always be interpreted as context-old and new information as context-new. For this reason, to account for the old/ presupposed *vs* new parts of the discourse, Prince's (1981) terms are adopted here, i.e. discourse-old and discourse-new. Discourse-oldness may be considered to be more equivalent to context-oldness than old information. In other words, topics are not necessarily associated with discourse-oldness, and thus the discourse-old/new status of the referent of an expression should be distinguished from its information status, as topical or focal, according to Casielles-Suarez (2004). To illustrate, Vallduví (1990) and others have pointed out that in some cases, discourse-old elements can constitute the focus of the utterance. An example can be seen in the constructed Thai exchange in (5.2):

(5.2) A: [...] thəə wâa khray cà   mây maabâaŋ
            you   think who FUT NEG come
      '[...] Who do you think is not coming?'

   B: chán wâa *KHǍW* tɔ̂ŋ   mây maa nɛ̂ɛ
       I   think (s)he must NEG come surely
      (Lit.) 'I think (S)HE is not coming, surely.'

[40] Given that proper names in Thai can have a pronominal function, the proper name may be spelled out as an overt subject, although this is judged as only marginally acceptable (due to the requirements of economy). That is, *ceen* 'Jane' may be pronounced as the subject pronoun in the reply in (5.1). For the issue of the obligatorily null pronominal subject, see the discussion in Chapter 6, Section 6.3.1.2.

A: chây (chǎn) khít yàaŋnán mǔankan

   yes   I   think so   as well

   'Yes, I think so, too.'

The embedded pronominal subject *KHǍW* '(s)he' in B's utterance must be interpreted as coreferent with a topic, a discourse-old element, in the preceding discourse. It is now introduced as the centre of the communication interest in the reply. This overt active topic element is also interpreted as part of the focus or contrastive topical element. Although this use of a pronoun in Thai is marginal and infrequent (in contrast to English pronouns, where overt pronouns frequently occur as active topical phrases), it shows that a pronoun as a discourse-old element can be focused, in spite of being coreferent with the discourse topic.

The discourse-old/new status of referents can be further elaborated in terms of their identifiability. Identifiability is concerned with the presence of knowledge and a shared representation in the mind of the speaker and the addressee at the time of utterance. Arguments with identifiable referents are highly eligible to constitute topical constituents. The well-formedness of a topic, then, depends on whether an identifiable referent can be activated. Unidentifiability results when the necessary knowledge and representations are available only in the speaker's, not the hearer's mind. To account for this, Lambrecht (1994: 109) proposes the model shown in Figure 1.

Figure 1: *Lambrecht's (1994) categorisation of discourse states of referents*
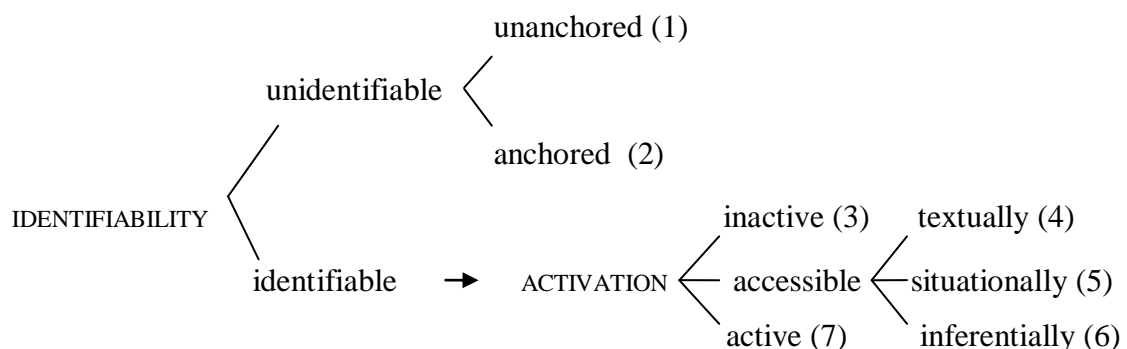
(1) unidentifiable/ brand-new   (2) unidentifiable anchored/ brand-new anchored
(3) inactive/ unused   (4) textually accessible
(5) situationally accessible   (6) inferentially accessible
(7) active/ given

The discourse states of *identifiable* and *unidentifiable* are further classified into sub-types. Since each category in the sub-types represents a particular, unique discourse state, I assume that each category is separate, independent, and not linked to any other. Unidentifiable referents are not identifiable to the addressee or hearer. In a sentences such as "I like *some film genres*", the referent of *some film genres* is not identifiable in the sense that the hearer is presumably unable to identify the specific genres the speaker has in the mind, and so it cannot constitute a well-formed topic (Casielles-Suarez (2004: 29)). However, according to Casielles-Suarez, an unidentifiable referent may be anchored when it is linked to some other discourse entity by means of another NP, properly contained in it. For example, an expression such as *a house* would be unanchored but *a girl I like* would be anchored.

If a referent is identifiable, one of the three states: active, semi-active (or accessible) and inactive, will be activated.[41] To illustrate, consider this Thai example:

(5.3) *mûawaanníi* (*chǎn*) cəə *kǐm$_2$* wà e_2 dəən yùu   kà *fɛɛn* e_2

yesterday   I   see kim PRT she walk PRES CONT with boyfriend of-GEN she

'Yesterday I saw Kim. She was walking with her boyfriend.'

The first sentence has a null deictic pronominal subject, *chǎn* 'I', which is fully active and represents part of the text-external world, i.e. a speech participant. It can optionally be null, since the CP-domain of the clause contains a feature representing the speaker (Sigurðsson (2004)). Similarly, the null anaphoric pronominal subject in the following sentence is active because of its anaphoric status in the text-internal world. Its representation, namely *Kim*, has been evoked, and hence it is a well-formed active topical element. An active referent, therefore, is one which is currently highlighted and identifiable, being capable of constituting a topical element. It cannot be overt, though, since it is coreferent with the discourse topic, unless it is part of the focus or a contrastive topic (see Chapter 6, Section 6.3.1.2 for a syntactic explanation of this). Among the italicised referents in (5.3), *fɛɛn* 'boyfriend' is brand new but anchored (even though the possessor is not spelled out as it is understood). The time expression *mûawaanníi* 'yesterday' is considered to be peripherally active, i.e.

---

[41] This is not the case for pronouns having a deictic or a generic reading, according to Lambrecht. They do not have to be activated, since they have no antecedent.

situationally accessible (associated with the non-linguistic environment) but not being focused on at this point in the conversation.

When an identifiable referent is inactive, the use of a pronominal element may lead to ambiguity. An example is:

(5.4) A: (chǎn) mây yàak hěn nâa (*khɔ̌ɔŋ kháw*) ʔìiklɛ́ɛw

      I   NEG want see face of-GEN (s)he again

   'I don't want to see his/ her face again!'

  B: (thəə) phùut thǔŋ   khray

     you   talk about who

   'Whom did you talk about?' or 'Who are you talking about'?

The possessor in A's utterance is either null or overt. When null, the speaker assumes that the possessor is identifiable and accessible to the addressee, since coreference between the possessor and the discourse topic is presupposed (hence topic-linked). As indicated by the reply, it turns out that the speaker's assumption is wrong. The addressee's response suggests that the possessor, namely *khɔ̌ɔŋ kháw* 'his/ her', is not at all activated even if it is pronounced (hence focused). A stronger claim is that it may be interpreted from the reply that the referent is unidentifiable to the addressee. Also, when the possessive pronoun in (5.4A) is not pronounced, it adds a further degree of unidentifiability, since it refers to a person whose referent is not the current focus of interest of the addressee. To disambiguate the sentence, the inactive referent needs to be encoded as a definite lexical NP, rather than a pronoun, and perhaps receiving prominent stress so that it can be activated. An example is:

(5.5) (thəə) cam   *phûuyǐng chúd dam* thîi khuy yùu   kap ceen

   you   remember woman   dress black COMP talk PRES CONT with Jane

   mûawaan dâay máy [...]

   yesterday able Q

   'Can you remember the woman in black who was talking to Jane yesterday? [...]'

The example shows a straightforward method (i.e. the use of an R-expression) for activating a referent which has been unused/ inactive or unidentifiable.

Brand-new referents, by definition, denote discourse-new/ hearer-new entities. A brand-new referent may be made a topic by means of a topicalizing strategy:

(5.6) *This car*, Jane has just bought (but the other one I bought.)

A discourse-new element can be a contrastive topic even though this is not a preferred option in English. The object is moved to the front of the sentence, being promoted to topic, while leaving a trace/ copy in the canonical position.[42] Some discourse-new constituents in certain contexts, nonetheless, do not undergo movement or dislocation in order to acquire a topic function. In fact, they are promoted to topic status in a presentational clause of their own (Lambrecht (1994)). This is typical for the opening discourse in tales and the like. Consider the following contrast:

(5.7) kaanlakhráŋnùŋ yaŋ   mii khɔnkhrέɛ ʔaasǎy yùu   bɔn yɔ̂ɔd khǎw [...]
      once   PAST have dwarf   live PRES CONT on top mountain
      'Once there was a dwarf living on the top of a mountain [...]'

(5.8) # khɔnkhrέɛ khǎw ʔaasǎy yùu bɔn yɔ̂ɔd khǎw   kaanlakhráŋnùŋ [...]
       dwarf   he   live   BE on top mountain once
       '#A dwarf, he lived on top of a mountain once [...]'

The Thai example in (5.7) is well-formed, and appears similar to the English counterpart where the topic is introduced into the discourse without the use of a topicalizing device. In contrast, the sentence in (5.8) does not follow the fairy-tale genre's conventional way of introducing the topic in a presentational clause, and thus it is judged as very odd pragmatically. If the adverb *kaanlakráŋnùŋ* 'once' is removed, the sentence is not odd anymore. However, it would then be an instance of a simple statement, rather than fitting in

---

[42] This is typical of English topicalization. The presence of null spell-out of the canonical obbject marks a difference between topicalization and left-dislocation. In Thai, matters are not as straightforward as in English, for instance:

(i) rɔd khan níi   pim phêŋ sɯ́ɯ maa mày
    car CLS DEM Pim just buy PERF new
    'This car, Pim's just bought.'

It is difficult to tell which topicalizing device is being used in the sentence, since Thai allows object drop. See more discussion on this issue in Section 5.4 below.

with a fairy-tale genre. Another context where a brand-new referent may be found as a topic is in newspaper headlines. Examples like the following have clear topic-predicate articulation:

(5.9) Anderson beat 19 other finalists to be named this year's MasterChef.

(Online BBC news: 28.4.11)

Although English has well-known devices such as topicalization and left-dislocation to signal a topic, these are not used in (5.9). Obviously, the subject is not left-dislocated, since no resumptive element is found. As for topicalization, according to Lambrecht (1994) in English a constituent is typically topicalized when it has a topic-shifting role, or when the topic status of the constituent is not very clear. But neither of these is the case when the subject is a discourse topic, as in (5.9). It is, then, not topicalized either. In the context of a newspaper headline, the sentence-initial argument tends to be topical. It is therefore natural to construe the referent of *Anderson* as having a topic function, since it fulfils the sentence-initial requirement, despite being brand new. Hence, there is a correlation between sentence-initial position and topicality.

The referential categories of discourse elements have been shown to consist of more than discourse-old and hearer-new, and therefore topics need not be anchored to discourse-oldness and foci need not be equated with hearer-newness. In fact, according to Lambrecht (1994) and Casielles-Suarez (2004), there are three major states that referents can have: discourse-new/hearer-new (brand-new), discourse-new/hearer-old (unused) and discourse-old/hearer-old (evoked). A topic expression may be discourse-new/hearer-new or discourse-old/hearer-old. The former can be construed as a shifting topic, and the latter can be interpreted as a familiar topic. This brings us to the next section, on different types of referential topics.

5.2 Referential Topics

Frascarelli and Hinterhölzl (2007) divide different types of topics into the following categories, on the basis of their formal and discourse properties: Aboutness-shift (A-shift Topics), Contrastive (C-Topics), and Given/ Familiar (F-Topics) topics; in addition, there are Hanging Topics (H-Topics). Using a Rizzi-type of approach to clausal stucture, each of these

topics can be fitted into the specifier of a dedicated Top projection (TopP) in the CP zone. The central idea of the left periphery of the clause is explored by Rizzi (1997) for data from Italian. His proposal is that a complete clause (CP) can be split into more than one type of projection. Specifically, it includes two further preverbal projections, i.e. a Focus Phrase (FocP) and a Topic Phrase (TopP):
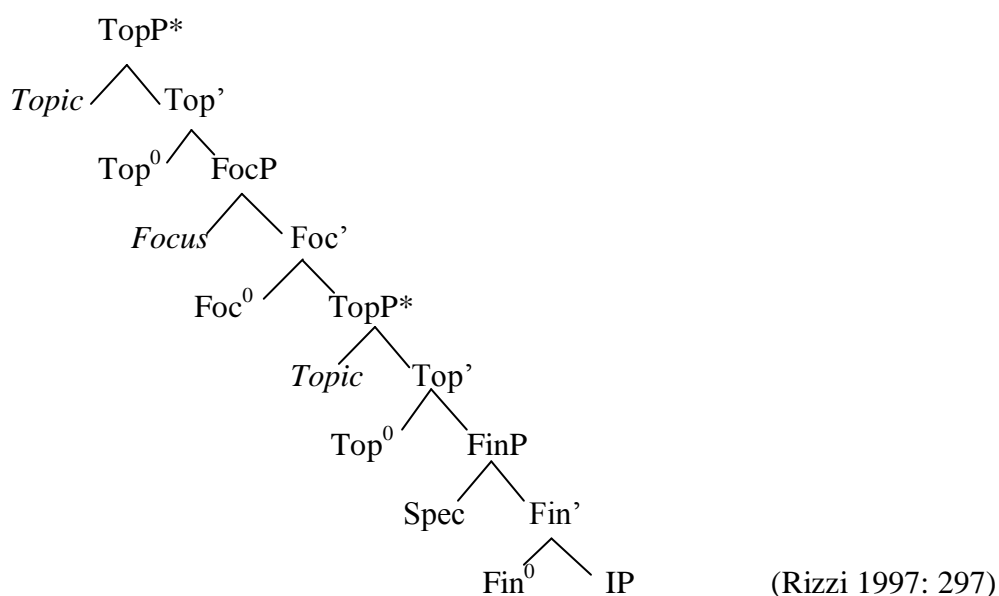
(Rizzi 1997: 297)

Rizzi (1997) thus assumes that topics and foci are realised within a Topic-Focus system at the left periphery. Since multiple topics are attested in most languages, the iteration of TopP (indicated by the * symbol) is crucial. As a result, the TopPs located between CP and TP (IP) are cross-linguistically found both higher and lower in the tree than FocP.

In Thai, an H-Topic (if available) is ranked on top of the hierarchy of topic constituents; A-shift Topics are often the next highest, while C-Topics and F-Topics may each precede the other. However, the data show that multiple topics are found only very marginally in Thai clauses, and I therefore assume that it is the Spec, CP position which can host a topic, apart from other functions. In what follows, I discuss each topic type in terms of its characteristics, to see whether or not it is obligatorily resumed by a null pronominal argument.[43]

---

[43] It should be noted that some constituents can appear in leftmost position but *cannot* be topics. Examples are given in (i), (ii) and (iii), which feature a quantified NP, a question word, and a sentence adverbial respectively:
(i) *năŋ baaŋ rŭaŋ* chăn mây chɔ̂ɔb
    film some CLS I   NEG like
    'Some films I don't like.'

5.2.1 Hanging topics (H-Topics)

Typically, H-Topics in Thai correspond to the English phrase *as for X,* placed on top of the topic hierarchy in the left peripheral position. They are externally merged with CP, and not moved:

(5.10) sâmlàb kim mâykheey mii khray thamhây *(thəə) phɔɔcay dây leey

       as for Kim never   have who make   she please able at all

       'As for Kim, no one is able to please her.'

(5.11) sâmlàb kim *(thəə) mây tôŋkaan pen lûukcâaŋ tàlɔ̀ɔdpay

       as for Kim   she NEG want   BE employee forever

       'As for Kim, she doesn't want to be an employee forever.'

The word *sâmlàb* 'as for' is a preposition. Like other prepositions in Thai, it needs an obligatorily overt phrase as its complement. The following TP must contain a resumptive pronoun. The word *sâmlàb* 'as for' can be left out, but this would turn the example into an instance of a left-dislocated construction. Since both H-Topics and left-dislocates (see Section 5.4 below) are base-generated, not moved, in the left periphery, they can be equivalent. This explains why they are called H-Topics. The canonical position of the H-

---

(ii) # *KHRAY* ceen chɔ̂ɔb
     who   Jane like
     'Whom does Jane like?'

(iii) *phûud taamtrɔŋ*   khǎw mây mɔ̀   pen ʔɛɛ
     honestly/ speak frankly she-TOP NEG   appropriate BE air hostess
     'Honestly/ Frankly speaking, she's not fit to be an air hostess.'

These constituents are not referential/ definite, and are not old information either. A quantified NP in (i) does not constitute a topic, as it cannot denote a referent whose representation could be activated, and this applies to the other examples as well. In (ii) the fact that the questioned constituent receives stress indicates that it has a focus function, rather than a topic role. The hash indicates that it sounds very odd when *KHRAY* 'who' takes subject position, given that Thai is a *wh*-in-situ language. In (iii), the sentence adverbial is base-generated external to the maximal verbal projection. As it does not refer to any entity, it cannot be a topic. These examples show that a well-formed topic needs to be definite, referential, being capable of being identifiable. Note that the relative order of a topic and a sentence adverbial is flexible in Thai. However, referential locative and temporal adverbials may function as optional arguments binding traces in the *v*P and hence they can be topics in the left periphery, according to Kiss (2008). An example is:

(iv) thîi phuukèt khɔn pay thîew yɔ́ mâak
     at Phuket people go visit a lot very
     'Phuket, people visit quite a lot.'

Topic cannot be filled by a null pronoun (unlike left-dislocates, and this is how an H-Topic is distinguished from a left dislocate) as seen in both (5.10) and (5.11); nor can it serve as the antecedent for a null pronoun in subsequence utterances. Also, H-Topics do not signal a topic shift:

(5.12) phrɔ́ *khàyǎai khǝǝfiw* chûaŋníi *e* ʔɔ̀ɔk càak bâan lǎng sìithûm mây
because extend curfew these days go out from home after 22.00 NEG
dâay wáy (chán) thoo nád ʔìikthii
able let I phone appoint later
'These days one/ I can't go out after 22.00 because of the curfew extension. Let me phone and make an appointment later.'

The phrase *khàyǎai khǝǝfiw* 'the curfew extension' in the first sentence is an H-topic. Clearly, the null subject is not coreferent with it. Instead, it is ambiguous between two readings: it can have a generic inclusive reading, or it can refer to the speaker. This is, however, not the case for the null subject of the second sentence. Such a second sentence neither has a plural reading nor generic inclusive reading. For this reason, the predicate dictates that this null subject can only be interpreted as referring to the speaker.

5.2.2 Aboutness topics (A-Topics) *vs* Aboutness-shift topics (A-shift Topics)

A-Topics, as the name indicates, represent what the sentence is about (Reinhart (1981); Lambrecht (1994)), whilst an A-shift Topic introduces a shift of topic or reintroduces a topic in the discourse (Givón (1983)). In Thai, when there is a shift to a new topic,[44] the use of an overt nominal constituent is required. The overt nominal acts as a signal to the addressee or listener that a topic is being shifted. Once the A-shift topic is established, it becomes an A-Topic, which is resumed by null arguments, unless there is a shifting topic. Thai therefore appears to have a different way of introducing topic shifting from, for instance, Italian, as can be seen from Italian (5.13) in comparison to the translation into Thai, given in (5.14):

---

[44] In Thai, an overt NP is obligatorily used if the shifting topic and the current A-Topic are third person. If the shifting topic is first or second person, the use of an overt title or name is optional.

(5.13) domani devo   andare con mio fratello$_1$ e   mia cognata$_2$ a   comprare
       tomorrow I have to go   with my brother and my sister-in-law for buy
       le fedi […] e_{1+2} restano qui alla Garbatella per il momento − comunque lei$_2$
       the rings   they stay here at Garbatella for a moment   she
       mi ha   detto che appena e_2 può e_2 se ne va   perché non per la zona
       me have told that when she can she there move because it is for the zone
       credo perché è   la casa dove lei_2 è   cresciuta per cui bene o male
       I think because it is the house where she was grown up so that good or bad
       la casa sì   qualcosa l'hanno cambiata quando i genitori sono andati via
       the house yeah something was   changed when the parents were left   away
       però   lei$_2$ dice cioè […]
       however she says that […]   Italian: Frascarelli (2007: 712)

'Tomorrow I must go with my brother$_1$ and my sister-in-law$_2$ to buy the wedding rings […]
They$_{1+2}$ are staying here at the Garbatella [a quarter in Rome] for the moment − anyway she$_2$
told me that as soon as *pro*$_2$ can *pro*$_2$ moves because, not for the zone, I think because this is
the house where she$_2$ grew up so that, yeah, more or less, something was changed in the
house when her parents left however she$_2$ says […]"

(5.14) phrûŋníi (chǎn) tôŋ pay súɯ wɛ̌ɛn tèŋŋaan kàb phîichaay$_1$ (khɔ̌ɔŋ chǎn) kà
       tomorrow I   must go buy ring wedding with brother   of-GEN I   and
       phîisàpháy$_2$ (khɔ̌ɔŋ chǎn) [...] e_{1+2} kɔ̀ɔ khɔŋ   cà phák thîi kaabaatelâa sàk
       sister-in-law of-GEN I   they then probably FUT stay at Garbatella for
       phák   tɛ̀ɛ   phîisàpháy$_2$ (khɔ̌ɔŋ chǎn) kheey bɔ̀ɔk (chǎn) wâa
       moment anyway sister-in-law of-GEN I   ever tell   I   COMP
       e_2 cà   rîab yáayʔɔ̀ɔk hâyrewthîisud thâwthîi e_2 cà   thamdâay
       she FUT hurry move out soonest   as   she FUT can
       mâychây wâa   khèet thîiyùu nán mây dii tɛ̀ɛ (chǎn) khít wâa
       NEG   COMP zone living DEM NEG good but   I   think COMP
       [...] yàaŋnɔ́ɔy sǎmlàb bâan lǎŋ níi baaŋyàaŋ dâay plìanpay tâŋtɛ̀ɛ phɔ̂ɔmɛ̂ɛ
             at least as for house CLS DEM something PERF change since parents
       càakpay tɛ̀ɛ phîi(sàpháy)$_2$ kôo bɔ̀ɔk [...]
       leave   but sister-in-law   then say [...]

(Lit.) 'Tomorrow I must go to buy the wedding rings with my brother$_1$ and my sister-in-law$_2$.
[...] They$_{1+2}$ are staying here at the Garbatella [a quarter in Rome] for the moment. Anyway
my sister-in-law$_2$ has ever told me that as soon as e_2 can e_2 moves because, not for the zone, I

think [...] at least, as for this house, something has changed in the house since her parents left; but, sister-in-law$_2$ says [...]"

In Italian, *mia cognata* 'my sister-in-law' is proposed as an A-shift Topic by using the strong pronoun *lei*. Thai, on the contrary, uses the full NP *phîisàpháy* 'sister-in-law' to introduce the topic shift, despite its being established as a marginalised constituent in the opening discourse. Being introduced into the discourse, the A-shift Topic now becomes and A-Topic coded with null arguments in the phrase: *phîisàpháy$_2$ (khɔ̌ɔŋ chǎn) kheey bɔ̀ɔk (chǎn) wâa e$_2$ cà rîab yáayʔɔ̀ɔk hâyrewthîisud thâwthîi e$_2$ cà thamdâay* 'my sister-in-law$_2$ told me that as soon as *pro$_2$* can *pro$_2$* moves [...]'. In other words, the following null pronouns will be interpreted as coreferential with the established A-Topic.[45] The speaker, then, introduces an H-Topic 'the house', and the topic is shifted again to 'my sister-in-law'. Since *phîisàpháy* 'my sister-in-law' qualifies as a shifting topic in the last clause, it is coded with an overt R-expression in Thai. Some similarity is revealed between the two languages, however, in that the use of null pronouns is related to a parameter involving topic prominence. That is, a third person null pronominal subject is understood as coindexed with the antecedent which is the current A-Topic.

In Thai, neither A-Topics nor A-shift Topics are preceded by a preposition unless they serve as a prepositional complement. The following example is another instance of topic shifting which supports the above claim that an A-shift Topic in Thai takes the form of an overt NP:

(5.15) A: ceen$_2$ kàb nɔn$_3$ yàaŋthîi pim$_1$ lâw [...] (*khǎw$_{2+3}$) pen khɔn dii

       Jane and Non as   Pim say   they   BE person good

       mâak   (*khǎw$_{2+3}$) ʔawcaysày lûuknɔ́ɔŋ (*khɔ́ɔŋ khǎw$_{2+3}$) dii thúkkhɔn

       very   they   take care   staff   of-GEN they   good everyone

       tɛ̀ɛ nɔn$_3$ diikwàa nítnuŋ (*khǎw$_3$) chɔ̂ɔb phaa (chǎn) paykinkhâaw

       but Non better slightly   he   like   take   I   dine out

       lɛ́ɛw (*khǎw$_3$) yaŋ càay ŋen thúkthii [...]

       and   he   also pay money every time

---

[45] This mainly applies to colloquial spoken Thai. In written Thai, a following coreferential argument may be overt, even if it is not associated with focus.

'Jane and Non, as Pim said [...] are very good people. They take good care of all their staff. But Non is slightly better. He likes to take me out to dinner. He also pays for the meal every time.'

After its introduction in the discourse, *nɔn* 'Non' receives an A-shift Topic function by means of the overt R-expression. Part of the reason for the use of an R-expression rather than an overt pronoun is that names can be used pronominally in Thai, and also the fact that the third person pronoun *khǎw* is not distinguished in terms of gender. (5.15) shows that null subject is coreferent with the current A-Topic which is not too far away; such a topic is in the immediately preceding sentence. There are cases where there are several possible antecedents available for the interpretation of a null argument. In such a case, it is the context that dictates what should be a well-formed antecedent of a null argument and whether or not such an antecedent has a topic role:

(5.16) ceen$_1$ (nà) t_1 khàyǎn mâak cim bɔ̀ɔk wâa   (*khǎw$_1$) pay hôŋsàmùt tɛ̀ɛ   cháaw

        Jane TM   studious very Jim say COMP   she   go library   early morning

        'Jane, she is very studious. Jim says that she goes to the library in the early mornings. '

In this example, the context appears to favour the more distant antecedent, i.e. the A-Topic *ceen* 'Jane' established in the opening discourse. The null embedded subject in the second sentence is coindexed with it, rather than with the matrix subject. In this regard, a null argument is allowed, given that its interpretation can be traceable according to the context (Minegishi (2011)).When coreferent with the A-Topic, the null subject must obligatorily be null.

However, it is also possible for the null embedded subject to corefer with the matrix subject even if such a subject does not have a topic role:

(5.17) A: câwtùub$_1$ hǎay   pay sɔ̌ɔŋ wan lɛ́ɛw

          Ajax   disappear DIR two day PERF

          'My dog has been gone for two days!'

B: cim$_2$ bɔ̀ɔk waa (khǎw$_2$) cəə *(man$_1$)^{46} lɛ́ɛw [...]

   Jim say COMP he   find it   PERF

   'Jim said that he has found it [...]'

The context makes clear that the A-Topic of the discourse is *câwtùub* 'my dog', rather than *cim* 'Jim'. Therefore, *cim* 'Jim' qualifies as a linguistic antecedent of the embedded null subject. *cim* 'Jim' may, but need not be, the A-shift Topic. The embedded subject can optionally be null, since it is identified by a local c-commanding antecedent without an explicit topic function. The following examples also show a null embedded argument that is coreferential with an element in the higher clause. (5.18) exemplifies a coreferential null embedded subject; (5.19) exemplifies a coreferential null embedded object:

(5.18) ran$_1$ bɔ̀ɔk wâa   (khǎw$_1$) pay hǎa mɔ̌ɔ mûawaan

     Ran say COMP he   go see doctor yesterday

     'Ran said that he went to see the doctor yesterday.'

(5.19) ceen$_1$ bɔ̀ɔk wâa   khunphɔ̂ɔ (*khɔ̌ɔŋ thəə) súɯ rɔ́d hây (khǎw$_1$) mûawaan

     Jane say COMP father   of -GEN she buy car give she   yesterday

     'Jane said that her father bought her a car yesterday.'

The matrix subject in both examples constitutes the linguistic antecedent of the null embedded argument. Since the antecedent is in a locally c-commanding position, the argument can optionally be null. Thus, a null embedded argument can be directly controlled by a higher argument, as in (5.17), (5.18), and (5.19). In other words, Thai allows subject control into complements of *say/ expect* -type verbs. Alternatively, a null embedded argument can be controlled via a null topic chain which is itself linked to the A-Topic of the discourse, not the matrix subject, as in (5.16). The issue of control will be discussed in detail in Chapter 6, Section 6.4.

---

^{46} For the obligatory overt pronoun referring to animals, see Section 5.4 below.

5.2.3 Familiar topics (F-Topics)

F-Topics are discourse-linked constituents used to resume background and given information for the sake of topic continuity, according to Frascarelli and Hinterhölzl (2007). As a shared representation already exists in the interlocutors' minds, the referent of an F-Topic is identifiable. As a result, it is resumed by a null pronoun, just like an A-Topic. What distinguishes an F-Topic from an A-Topic is that the former is generally found in post-verbal position, whilst the latter tends to be found in pre-verbal position. Example (5.20) illustrates the case of F-Topics and the interaction between an F-Topic and other topics:

(5.20) U1: rin_1 duu   keemchoo thaaŋ thiiwii

       Rin   watch game show on   TV

     'Rin watched a game show on TV.'

  U2: khànà mii khoodsànaa kôo   (*$khaw_1$) thoo háa pam_2

    during has commercial then   she   call   to Pam

    'During a commercial break, she called Pam.'

  U3: (*$khǎw_1$)   chuan (*$khǎw_2$) maa kin khâaw thîi bâan (*khɔ̌ɔŋ $khǎw_1$)

     she   invite   her   come eat meal   at   house of-GEN she

    'She invited her to have a meal at her house.'

  U4: pam_2 bɔ̀ɔk   (*$khǎw_1$) tòkloŋ

    Pam   answer   her   accept

    'Pam accepted her invitation.'

  U5 tɛ̀ɛ   phɔɔ thǔŋ wan nád   (*$khǎw_2$) kə̀əd   pùay kôo yóklə̀ək (*$khǎw_1$)

    but when reach day appointment   she   happen sick then cancel   her

    (Lit.) 'On the appointment day, she happened to be sick. She then cancelled her for the dinner at her house.'

As the context makes clear, the subject in the opening discourse represents a topic element, which is first introduced as an A-Topic, then resumed by a null pronominal subject in U2 and U3. The null subject in U2 and U3 represents the pragmatic relation of aboutness in Reinhart's (1981) sense. It is finally repeated as an F-Topic realised as a right-hand topic in the post-verbal position, in U4 and U5. This is because *Pam* in U4, now overt, marks a shift

in the conversation, assuming the status of A-shift Topic. This example shows that the F-Topic, resulting from topic shifting, is a former A-Topic.

5.2.4 Contrastive topics (C-Topics)

C-Topics should be distinguished from Contrastive Foci (CF), since they do not behave in an exactly similar way. To begin with, C-Topics have the discourse function of focus by generating oppositional pairs in relation to other topics (Lambrecht (1994)). Unlike other types of topic, C-Topics in Thai are preferably coded with an R-expression (cf. (5.2)), since each topic referent is represented as a member in an entire set in the discourse. To contrast one topic out of the set, a full lexical NP is normally employed:

(5.21) A: lùukmǎa thîi   (khun) líaŋ   wáy   pen ŋaybâaŋ

          puppy   COMP you   look after PERF BE Q

      'How are the puppies that you're taking care of?'

    B: câwtualék thîisùd   taay paylɛ́ɛw [...]

       puppy   smallest   die PERF

      'The smallest puppy has died [...]'

The entire available topic referent set *lùukmǎa* 'puppies' was activated by Speaker A, and thus is active. In the reply to the question, *câwtualék thîisùd* 'the smallest puppy' as a topic is being contrasted with the remaining topics (the other members of the set). This leaves the other puppies' status as unknown. Nonetheless, it can be assumed that they are in contrast with the C-Topic in the reply. That is, the other puppies are alive. The example in (5.22) patterns in the same way; regardless of how many topics there are in the question, one of the topics must be pronounced in the reply, coded with an R-expression:

(5.22) A: (thəə) cəə dam kàp kim thîi rooŋrian lǔaplàaw

       You   meet Dam and Kim at   school   Q

      'Did you meet Dam and Kim at school?'

    B: (chǎn) cəə   tɛ̀ɛ kim

     I   meet just Kim

     (Lit.) 'I met just Kim [but not Dam]'

The example exhibits an instance of C-Topic where the referents have already been introduced in the discourse. In the reply, the topic *kim* 'Kim' is being contrasted with the other topic. In Thai, in situ focus is used in C-Topic constructions, rather than initial focus as in the English cleft construction. As discussed in Chapter 2, Thai has a common gender third person pronoun *khǎw* '(s)he', and an R-expression can be used pronominally; this explains why use of the pronoun is not the preferred option in the answer to (5.22B).

CF, on the other hand, has no topic function. It is also known as 'identificational focus' (Kiss (1998)), a different category from 'information focus' (see Section 5.1 above). What is normally found in contrastive focus contexts in Thai is that CF is generated in corrections, with the adverbial *tàhàak* 'instead' typically being used. The adverbial is roughly equivalent to the English cleft construction. The presence of this adverbial makes a C-Topic sentence ungrammatical. Compare the CF in (5.23) with the C-Topic in (5.24):

(5.23) A: Kim pay bâan phɔ̂ɔ (khɔ̌ɔŋ) phɯ̂an (*khɔ̌ɔŋ thəə) thîi   pen mɔ̌ɔ
           Kim go   house father of-GEN friend   of-GEN she COMP BE doctor
           'Kim's gone to the house of her father's friend, who is a doctor.'
     B1: mâychây (*khǎw) pen *KHRUU tàhàak*
           no   he   BE   teacher instead
           'No, he's a teacher [not a doctor].'
     B2: mâychây (*khǎw) pay *BÂAN (khɔ̌ɔŋ)   CIM tàhàak*
           no   she   go   house of-GEN Jim instead
           'No, it's JIM'S HOUSE she's gone to, [not her father's friend's].'

(5.24) A: lùukmǎa thîi   (khun) líaŋ   wáy   too mɔ̀d lɛ́ɛw châymǎy
           puppy   COMP you   look after PERF grow all ASP Q
           'All the puppies you look after have grown up, haven't they?'
     B: * câwtualék thîisùd   tàhàak   taay paylɛ́ɛw [...]
           puppy   smallest instead die PERF
           'The smallest puppy is dead [...]'

Since the adverbial *tàhàak* 'instead' is restricted to a CF construction, (5.24B) is ill-formed. The word *tàhàak* 'instead' cannot be used in an information Focus/ narrow Focus construction either. (5.25) is an example:

(5.25) A: kim pay náy

         Kim go   where

       'Where did Kim go?'

     B: * *e* pay tàlàad   tàhàak

         She go   market instead

        'It's the market that she went to.'

The examples given so far reveal that the interpretation of a null pronoun in Thai is dependent on an A-Topic or an F-Topic. In what follows, the properties of such referential null topics are discussed in order to determine which particular types of topics significantly correlate with the occurrence of referential null arguments.

5.3 Properties and Internal Structures of Referential Null Pronouns

I will argue that a referential third person null pronoun in Thai is distinctive in the following way: it has no ϕ-features, and thus is a 'minimal noun'. If there is no discourse-based antecedent or c-commanding antecedent in a higher clause, a null argument cannot have a referential third person interpretation, but will be interpreted as first person by default. Thus to have a referential third person reading, the null argument must look for an antecedent, which is a referential third person. Consider this example:

(5.26) U1: [...] yàaŋthîi pim$_1$ lâw *câwnaay khɔ̌ɔŋ chǎn$_2$* pen khon   dii   mâak

           as   Pim say boss   of–GEN I   BE person good very

   U2: (*khǎw$_2$) Ɂawcaysày lûuknɔ́ɔŋ (*khɔ̌ɔŋ khǎw$_2$) thúkkhon

          he   take care staff   of-GEN he   every

   U3: thɛ̌ɛm (*khǎw$_2$) chɔ̂ɔb phaa (chǎn) paykinkhâaw dûay

         plus   he   like   take   I   dine out   too

   'My boss, as Pim said [...], is a very good person. He takes care of all the staff. Also, he likes to take me out to dinner.'

The entire example could be a reply to a question like "*What about x?*", or a command like "*Tell me something about x*". If someone asks: *câwnaay mày   penŋay bâaŋ* 'What about your new boss?', then there are three options although the null subject is the most natural

reply to that question: the use of an overt R-expression *câwnaay* as the subject in the reply, or an overt pronoun, or a null element. If the subject is pronounced, a focus or contrastive interpretation will be implied. In (U1), *câwnaay khɔ̌ɔŋ chǎn* 'my boss' functions as an A-shift Topic, and the whole discourse mainly conveys information about that topic. The interpretation of the null subjects in the example requires the presence of such a topic so that they can be coindexed with it. Without the presence of a topic, the null subjects would lose their third person referential reading.

The relation between the topic and the null arguments in the subsequent utterances can be analysed as an instance of indirect control via a null topic chain, requiring no c-command (and thus not being a case of binding). Assuming that a null argument in Thai is minimal and bears no specifications other than [uR, N],[47] it will be dependent upon an antecedent for a referential interpretation. Hence it needs to be controlled if it is to be referentially interpretable. Resulting from A'-(topic) movement, a null argument, of the type [uR, N], is a copy of the A-Topic operator in the CP-domain where there is referential index-copying. The referential index is also transmitted to the copy in the A-position (see Chapter 6, Section 6.3.1.2 for a detailed analysis of the derivation of null topics). Therefore, the copy of the A-Topic operator has a referential reading, and is not pronounced, due to its having no φ-features. The index-copying relation exists between Topics, i.e. referential NPs, in Spec, CP across sentences. This suggests that φ-features are not a prime requirement for the interpretation of a null argument in Thai, as long as it has a referential index assigned. The following example illustrates that a null pronominal argument in a complement clause can be controlled by a discourse topic, rather than by the matrix subject:

(5.27) A: kim_1 pen ŋaybâaŋ

      Kim BE Q

      'How is Kim?'

   B: $ceen_2$ bɔ̀ɔk wâa   (*$khǎw_1$) dâay ŋaan tham lɛ́ɛw

      Jane say COMP   she get job do   PERF

      (Lit.) 'Jane said that she's got a job already.'

---

[47] For the function and properties of the [uR] feature, see Chapter 6, Section 6.1.

If B's utterance were an out-of-the-blue sentence, its null subject would be interpreted as being directly controlled by the matrix subject *ceen* 'Jane'. In (5.27), this is not an option. The antecedent can only be the A-Topic introduced in the question. As a result, there can be no co-indexation of the embedded null subject in the reply and the c-commanding argument. Put differently, the fact that coindexation of the embedded null subject and the matrix subject is not possible suggests that the null subject represents a null A-Topic of the discourse.

As for F-Topics, the examples above illustrate that they are a result of topic continuation, so they are typically a repetition of a previously established A-Topic. An F-Topic in Thai is typically established as a right-hand topic, as can be seen in the following:

(5.28) U1: ceen$_2$ kàb nɔn$_3$ yàaŋthîi pim$_1$ lâw [...] (*khǎw$_{2+3}$) pen khon   dii
         Jane and Non as   Pim say   they   BE person good
      mâak   (*khǎw$_{2+3}$) ʔawcaysày lûuknɔ́ɔŋ (*khɔ́ɔŋ khǎw$_{2+3}$) dii   thúkkhon
      very   they   take care   staff   of-GEN they   good everyone
   U2: tὲὲ nɔn$_3$ diikwàa   nítnuŋ   (*khǎw$_3$) chɔ̂ɔb phaa (chǎn$_1$) paykinkhâaw
       but Non better   slightly   he   like take I   dine out
      lέεw (*khǎw$_3$) yaŋ càay ŋen   thúkthii
      and   he   also pay money every time
   U3: (chǎn$_1$) rúusùk kreeŋcay (*khǎw$_3$) thǔɯwâa (*khǎw$_3$) pen khon thîi
        I   feel   considerate he   all in all   he   BE person COMP
      (chǎn$_1$) nábtɯ̌ɯ mâak khon nùŋ
        I   respect very CLS one
      'Jane and Non, as Pim said [...] are very good people. They take good care of all
      their staff. But Non is slightly better. He likes to take me out to dinner. He also
      pays for the meal every time. I feel this is very considerate of him. All in all he is
      the one I respect [as a boss].'

The first sentence in U3 reveals that *nɔn,* the speaker's boss, becomes a right-hand topic, since the topic is shifted to the speaker and then shifted again to the boss in the second sentence. Note that the Thai word *kreeŋcay* 'considerate' can be a transitive verb. In U3, the null object of the verb *kreeŋcay* represents F-Topic, whist it retains the coreference with the previously established A-Topic subject. The fact that the third person referential null subject as the null F-Topic in U3 does not obviate coreference with respect to the A-Topic suggests

that it is controlled by the A-Topic, which means that there must be a null A-Topic operator in Spec, CP of the sentences in U3. In other words, the null object in U3 is a copy of the A-Topic operator as well.

Summing up, the interpretation of referential null pronouns in Thai is mainly dependent on two topic types, i.e. A-Topics and F-Topics. A-Topics have the following characteristics: (i) they correlate with sentence-initial position; (ii) they do not necessarily involve discourse-old and hearer-old elements; and (iii) they can be discourse referents or referents in a higher clause (in case the context makes explicit that the c-commanding referential NP is the topic). F-Topics, on the other hand, have the following properties: (i) they generally correlate with sentence-final position; (ii) they necessarily associate with discourse-old and hearer-old information; and (iii) they are restricted to discourse-referents. Significantly, both topic types are controlled by the same null A-Topic operator in Spec, CP, since they exhibit the same referential index. The operator itself is linked to a referential topic NP in the discourse. It is then sufficient to postulate that the interpretation of referential third person null pronouns in Thai depends on the referential index of a null A-Topic operator in the CP-domain as a result of referential index transmission from the referential topic antecedent of the discourse. Alternatively, a referential third person null pronoun can be controlled by a matrix NP in a higher clause. A generalisation which can be drawn is that Thai avoids an overt pronoun when there is an established topic (A-Topic or F-Topic) in the preceding sentence or in a higher clause.[48] When a pronoun has no special discourse function, such as Focus, the pronoun is obligatorily null.[49] This, however, excludes a pronoun that has no antecedent at all, as discussed in Chapter 4, in which case it needs to be overt in order to have a generic/ arbitrary reading. Therefore, the generalisation accounts for referential pronouns. The following preliminary formulation for pronominal realisation in Thai in relation to discourse function is proposed:

---

[48] For English-type languages, it is obviously not the case that a referential third person pronoun is obligatorily null, even when there is an established topic in the discourse. English is known to be strict about the phonological EPP-condition. That a Spec, TP is occupied by a syntactic subject is insufficient; it needs to be spelled out (Holmberg 2005). For further discussion of the issue of obligatorily null pronouns, see Chapter 6.

[49] This is in line with the Avoid Pronoun Principle, which Chomsky (1981: 65) formulates as a constraint on overt pronouns. He establishes that a null pronoun is always preferable due to the requirements of economy. This principle is considered a well-formed condition that is operative in pro-drop languages. However, it is arguable where exactly this condition operates. If it is a principle for discourse functions for the purpose of coherent information structure, then it can be argued to hold true for Thai as well.

(i)　Avoid overt pronouns, whenever a predicational sentence contains either an A-Topic or F-Topic unless a pronoun (i) associates with certain discourse functions such as Focus; (ii) is controlled by a linguistic c-commanding antecedent in a higher clause without an explicit topic function or (iii) has a generic quasi-inclusive reading, a generic exclusive reading or an arbitrary reading.

(ii)　A null argument in Thai has no ϕ-features. To achieve a thematic referential third person reading as a null topic, it undergoes A'-movement from its canonical position (if it can) to the A-Topic operator in the CP-domain where a referential index from the discourse antecedent can be assigned. The null argument in the canonical position would then be a copy of the A-Topic operator. Alternatively, it can be directly c-commanded by a higher referential NP, without undergoing movement.

The implication of (i) is that, on the one hand, a referential third person *null* pronoun in Thai must have a topic antecedent. That is, active topical elements are expressed by null pronouns. In the absence of an antecedent, the interpretation shifts to the first person. *Overt* pronouns, on the other hand, are not necessarily associated with a topic. When it is overt, a pronoun can have an antecedentless reading (hence be an instance of a quasi-inclusive G-pronoun, an exclusive G-pronoun or an arbitrary pronoun), or a referential reading having a local (close-enough) antecedent in a higher clause. The relation of the c-commanding antecedent in the matrix clause and the null embedded subject can be an instance of direct control, rather than being mediated via a null topic chain (see Chapter 6, Section 6.4 on the issue of control).

5.4　Some Criteria for Distinguishing Resumptive Pronouns and Traces/ Copies in Thai

Thus far, the distinction between a trace/ copy and a resumptive pronoun in topicalization and left dislocation respectively has not been mentioned. Topic structures are grammatical configurations, consisting of the topic (either a canonical subject or object argument), normally found in the left periphery, and the comment. There are two possible analyses for a topic in an A'-position in the literature. In a movement analysis, an argument is assumed to be preposed from the maximally extended verb phrase into the A'-position, with a trace/copy in the *v*P (except in the case of H-Topics which are externally merged, not moved). This is

known as topicalization, a cyclic rule that observes Subjacency.[50] The second analysis is left dislocation, where the left-dislocated element is base-generated in the A'-position and is therefore not constrained by conditions associated with islands and movement rules. (5.29) shows the logical form of the analysis of topicalization and left dislocation respectively:

(5.29) a. $[_{CP}$ khǎw$_i$] $[_{TP}...t_i...]$

  b. $[_{CP}$ khǎw$_i$] $[_{TP}...$pronominal$_i...]$

In (5.29a) the landing site for the topicalized element is not Spec, TP, as the examples given below show the presence of a canonical subject, occupying Spec, TP. For this reason, I assume that the Spec, CP position is available as a landing site for topicalization, and thereby *khǎw* '(s)he' undergoes A'-type movement. On the other hand, (5.29b) does not show this effect. The left-dislocated element is base-generated in Spec, CP, whilst the canonical position is occupied by a coindexed resumptive pronominal element. I will show below that a resumptive pronoun in Thai can be, and must be, null in certain cases.

If there is movement, as in (5.29a), the gap within the canonical position is not pronominal. Left dislocation, then, should be distinguished from topicalization. Given the fact that Thai allows pro-drop, it is not straightforward to identify whether a canonical position found in a sentence with a nominal constituent in the left periphery is occupied by a pronoun or a trace/ copy. Nevertheless, as I shall show, criteria can be established for determining whether a null element is a trace or a resumptive pronoun.

Firstly, in Thai, movement is constrained by certain configurations, namely islands and stranding of a preposition. Movement of an argument out of the canonical position in these constructions affects the grammaticality of the sentence. Each of the two cases is exemplified below:

(5.30) phûuyǐŋ khɔn níi$_1$ chǎn rúucàk phûchaay chɯ̂ɯ cɔɔn thîi   chɔ̂ɔb __ mâak
    woman CLS DEM I   know man   name John COMP like   very
    'This woman, I know a man called John that likes __ very much.'

---

[50] The subjacency condition requires that a constituent cannot be moved (in any single application) across more than one bounding node, according to Crystal (2008).

(5.31) yɔ́kwén kim nɔ̀ɔknán (chǎn) mây rúucàk khray ləəy

       except Kim other than   I   NEG know   who   at all

       'I know no one except Kim.'

In (5.30), the gap in the comment part of the clause is not compatible with the topicalization, and thus no movement out of the comment part of the clause is possible. Use of an overt pronoun will make the sentence grammatical, as shown in (5.32), which would be a case of left dislocation:

(5.32) [$_{CP}$ phûuyǐŋ khɔn níi$_1$ [$_{TP}$ chǎn rúucàk [$_{NP}$ phûchaay chɯ̂ɯ cɔɔn [$_{CP}$ thîi [$_{TP}$ chɔ̂ɔb

        woman CLS DEM   I   know   man   name John   COMP like

        thəə$_1$ mâak]]]]]

         she   very

     'This woman, I know a man called John that likes her very much.'

Given the island, *phûuyǐŋ khɔn níi* 'this woman' in (5.32) is base-generated in sentence-initial position, while its canonical position is filled by an overt resumptive pronoun, which is anaphorically related to the left dislocate. The fact that the pronoun in the island is pronounced exemplifies the second criterion distinguishing between a trace and a resumptive element: a trace cannot be pronounced. As such, the position of *thəə* in (5.32) cannot be a trace of the left-peripheral constituent. Islands, therefore, are a good diagnostic test for movement. (5.33) illustrates the subject island test:

(5.33) * phɔ̂ɔmɛ́ɛ [kaanduulɛɛ *t*] thɯ̌ɯ   pen sìŋ Ø phuŋ kràtham

       parents   taking care   consider BE what   should do

       'Parents, taking care of __, is what one should do.'

The ungrammaticality of a trace inside the subject island in (5.33) suggests that there is no movement either, but a dislocated NP at the left periphery. Therefore, the gap in the complex NP island should be treated as a resumptive pronoun that can optionally be pronounced. Nonetheless, the resumptive pronoun in (5.32) tends to be overt, rather than null. I assume that the distance between the left dislocate and the pronoun is important here. In (5.32), the two coindexed elements are separated by a number of intervening nodes: NP and CP. Since there is long-distance binding, the resumptive pronoun is preferably overt. This is not the case

in (5.33), since there is only one intervening node between the left dislocate and the resumptive pronoun. (5.33) therefore shows a short-distance anaphor and is bound by its antecedent, i.e. the left dislocate. Due to the short distance, the resumptive pronoun need not be pronounced. This effect of the spell-out condition of a resumptive pronoun wears off when the left-dislocated topic is inanimate:

(5.34) [CP bâan láŋ   níi [TP chǎn rúucàk [NP câwkhɔ̌ɔŋ chŵ̵u cɔɔn [CP thîi

       house CLS DEM I   know   owner   name John   COMP

     [TP kamlaŋcà   khǎay (#*man*)]]]]]]

       FUT PRES CONT   sell   it

    'This house, I know the owner called John who is going to sell it.'

Regardless of the distance of the left dislocate, the resumptive pronominal object is preferably null when the referent of the left dislocate is an inanimate entity.

As for a non-human animate left-dislocate, after a non-human animate referent is introduced into the discourse, such a referent must be resumed b an overt pronominal *man* 'it' only (and nothing else) can be used:[51]

(5.35) mǎa tua nán (chǎn) hěn khɔn thîi pay tə̀ *(*man*) thùuk *(*man*) kàd

     dog CLS DEM I   see person who PERF kick it   PASS   it   bite

    (Lit.) 'That dog, I saw a man who kicked it, and he was bitten by it.'

(5.36) thâa mǎa₁ cəə kràdùuk₂ *(*man₁*) cà   khíew (*man₂) pay   khrûŋ chûamooŋ

     if dog find bone   it   FUT chew   it   PERF half   hour

    (Lit.) 'If a dog/ dogs find(s) a bone/ bones, it/ they will chew it/them in half an hour.'

     = 'If a dog/ dogs find(s) a bone/ bones, it/ they will finish chewing it in half an hour.

As shown in these examples, where the pronoun refers to an animal (*mǎa* 'a dog') it must obligatorily be overt; where the pronoun refers to a thing (*kràdùuk* 'bone(s)') it is obligatorily null. Also, the resumptive pronoun in (5.33b), referring to humans, can optionally be overt. This applies to other pronouns, too. The fact that a resumptive pronoun in Thai cannot be pronounced is unexpected, since its position inside an island shows that it cannot be a trace.

---

[51] This excludes the case where the pronoun *man* 'it' is used to refer to a third person in an insulting or unfriendly manner, see Chapter 1, Table 1.

Probably the person and animacy of the left-dislocated NP triggers the spell-out of resumptive pronominal elements in Thai. I shall leave this complication for future research.

Finally, consider examples (5.31), repeated here as (5.37a):

(5.37a) yɔ́kwén   kim   nɔ̀ɔknán   (chǎn)   mây   rúucàk   khray   ləəy

       except   Kim   other than   I   NEG   know   who   at all

       'I know no one except Kim.'

(5.37b) * kim [PrepP yɔ́kwén *t*] nɔ̀ɔknán   (chǎn)   mây   rúucàk   khray   ləəy

       Kim   except   other than   I   NEG   know   who   at all

       'Kim, I know no one except __.'

(5.37b) shows that stranding of a preposition by topicalization is not possible either. Instead, an overt pronoun inside the prepositional phrase is required:

(5.38) kim1 [PrepP yɔ́kwén khǎw1 nîilà] nɔ̀ɔknán   (chǎn)   mây   rúucàk   khray   ləəy

       Kim   except   she   PRT   other than   I   NEG   know   who   at all

       'Kim, I know no one except her.'

That a trace is not possible is not surprising, since Thai does not allow a null prepositional complement. The well-formed overt pronoun which can be coindexed with the left-dislocated element entails that the referent *Kim* in (5.38) is base-generated in the left periphery. These examples show that the topic constituent takes the topic position in the CP-domain, and its canonical position is occupied by a resumptive pronominal element with the same reference. The topic then, can bind an anaphorically canonical argument in the TP across any number of intervening NP and CP nodes. Whether the resumptive pronoun is null or overt depends on the person and animacy of the topic NP, the distance between the left-dislocated phrase and the resumptive pronoun, and such configurations as prepositional stranding. It can be concluded from the examples that the Thai structure allows the dislocation of a nominal phrase so that a topic can be prominently established. Conditions on movement are irrelevant for the coindexation/ binding of the resumptive pronouns, for instance, in complex islands with the topic in the left-dislocated structure. Thus, the relation between the resumptive pronoun and the left dislocate is not sensitive to island constraints. In this respect, Thai appears to be similar to English. On the basis of the examples with a complex island or

stranding of a preposition, it can be concluded that topicalization is not interchangeable with left dislocation. I assume that, in the absence of the aforementioned configurations, a sentence with a left peripheral nominal constituent can be analysed as an instance of topicalization. This applies to, for instance, the referent *rɔ́d khan níi* 'this car' found in (i), footnoot 42. I shall leave for future study the issue of whether the left peripheral nominal constituent is actually topicalized in such a position.

5.5 Summary

Null pronouns in Thai are interpreted as active topical elements or non-topical elements, being controlled by a higher argument NP, whilst overt pronouns are interpreted as observing a discourse function requirement or having a generic/ arbitrary reading. That is to say, in the absence of agreement morphology, the interpretation of a referential null pronoun depends exclusively on the discourse context. Whether a pronoun can be null or not is determined by particular types of topic: A-Topic, F-Topic, or a local linguistic antecedent in a higher clause. I have argued that when a referential third person null pronoun appears, there must be a discourse third person topic whose referent is already active at this point in the discourse, which can be one of the above topics. If there is a locally c-commanding referential NP in the matrix clause apart from a discourse topic, it is the context that dictates the interpretation of a null argument. If the context favours the matrix subject as the antecedent of the null embedded argument, such a matrix subject behaves like a controller of PRO. However, a null embedded third person subject must be disjoint in reference with the matrix subject if the context favours a more distant antecedent which is an A-Topic of the discourse. Therefore, a referential null pronoun necessarily has an antecedent NP which appears either (i) earlier higher up in the sentence (not necessarily denoting a topic role), or (ii) earlier in the discourse (necessarily denoting a topic role). If a pronoun is overt, there are two possibilities. One is that the pronoun has an antecedent, but is overt since it is part of the focus or perhaps is a contrastive topical element, or it is controlled by a higher argument (which may also be null). Another possibility is that the pronoun does not have an antecedent at all; the reading is then generic or arbitrary. In addition, referential null pronouns in Thai can also occur in syntactic islands, in which case they are resumptive pronouns linked to a left dislocated element.

Chapter 6: The Structure of Null Arguments in Thai

6.1 The Properties and Internal Structure of Null Arguments in Thai

In investigating the phenomenon of Thai null pronominal arguments in this thesis, there are two main goals. Descriptively, to elicit and illustrate the contexts where overt and null pronouns can and cannot occur. Theoretically, to explain the distributional patterns of pronouns – when they have to be pronounced, when they may, but need not be, and when they cannot be pronounced – and see how the findings contribute to a better understanding of pro-drop in Thai and other pro-drop languages.

In the previous chapters, I have investigated and discussed the constraints on the occurrence of null pronouns, and found that constraints are associated with the syntactic context and/ or discourse functions of sentences. The significant environments where pronominal elements are found establish the relation between null pronouns with antecedents on the one hand, and overt pronouns without antecedents on the other. Based on the data and findings thus far, there are two major types of null pronominal arguments in Thai:

(i)   Referential null pronouns: if there is a topic embedded/ understood in the context, such null pronouns will be coindexed with the topic; alternatively, in the absence of context, the null pronouns will be interpreted as deictic, referring to the dominant discourse participant, i.e. the speaker. In other words, without a topic antecedent, the speaker is the default antecedent. These null pronouns are in contrast with overt referential pronouns, which have their ϕ-features together with their own referential index, i.e. [i [ϕ [N]]], and are therefore pronounced.

(ii) Null inclusive G-pronouns: even though they are generic pronouns with no antecedents, they are not overt. This is rather surprising, given that a G-pronoun without a linguistic antecedent or discourse antecedent is expected to be overt, and indeed is overt in the case of a non-inclusive G-pronoun. Saying that the Thai lexicon happens not to have an overt G-pronoun like English *one* is not an adequate explanation for this. Other languages make use of the second person singular pronoun as a G-pronoun, or use an indefinite pronoun (see Holmberg 2010b). I have argued in Chapter 4 that the use of the second person singular

pronoun *khun* 'you' for a generic reading is not a possible option, since its reading is necessarily referential. Even an indefinite pronoun or an indeterminate pronoun is not exactly synonymous with a purely inclusive G-pronoun[52] for Thai. Therefore, inclusive G-pronouns in Thai can only be null. I have proposed in Chapter 4 that they are bound by a generic operator. The pronoun has no restricted reference, so the reading is maximally general, including the speaker and the addressee and anyone else, thereby appearing null. Other G-pronouns in Thai are overt, as expected. Assuming the presence of the above generic operator, the other G-pronouns are restricted due to their having more features, and thereby appear as overt.

Here we come to a point of central theoretical interest and importance. Null pronouns in Thai in any position belong to a class of null arguments whose properties do not fit into Chomsky's (1982) theory of empty NPs. In fact, there are no inherent or lexical differences between them. They are all made up of the features [uR, N].[53] Thus, they do not refer by themselves. Given that the [uR] feature is accessed (valued) from outside, I assume that [uR] is the head. The [uR] feature is merged with the [N] feature, making up a nominal category. These features are sufficient for the category to function as an argument, capable of receiving a θ-role. Being φ-featureless, referential null pronouns and inclusive G-pronouns in Thai rely on an antecedent/ operator. They must be in the control domain of a (local) antecedent or operator, which determine their values, generic or referential. In other words, the [uR] feature of a null pronoun has to be valued by a referential index or generic operator. For instance, in order for the [uR, N] to have a referential reading, it has to be assigned a referential index *i,*

---

[52] Compare the reading of the inclusive G-pronoun in (i) and those of the indefinite pronoun or indeterminate pronoun in (ii):
(i) bâanmɯaŋ cà   nâayùu   Ø   tôŋ   chûay ráksǎa khwaamsàʔàad
     city   FUT liveable   must help   keep   cleanliness
     'The city will be liveable if one keeps it clean.'
(ii) bâanmɯaŋ cà   nâayùu *khɔn/ #khraykhray* tôŋ chûay ráksǎa khwaamsàʔàad
      city   FUT liveable people/ who   must help   keep   cleanliness
      'The city will be liveable if people/ they/ #one keep(s) it clean.'
The reading in (ii) lacks the generic inclusiveness while that of (i) does not. As for the indeterminate pronoun *khraykhray* 'who', it sounds very odd in such a position, as shown by the hash. It, too, makes the following sentence degraded when taking object position, given an intended generic inclusive reading (cf. 4.18):
(iii) kaankoohòk sǎamâat nam *#khraykhray* paysùu kaansǐa chûɯsǐaŋ
       lie   can   bring who   DIR   harm   reputation
       'Lies can harm *#one*'s reputation.'
Therefore, this suffices to support my claim that neither does an overt inclusive G-pronoun exist nor can Thai resort to other strategies, i.e. a second person singular pronoun, an indefinite pronoun, or an indeterminate pronoun to express a generic inclusive reading. The only option to do so is by means of null spell-out of an argument.
[53] [uR] is equivalent to Holmberg's (2010a,b) unvalued definiteness feature [uD].

coming from an available referential, definite antecedent. The index is the value of the [uR] feature, thereby resulting in [i[N]]. This null pronoun is not pronounced, as it does not have ϕ-features. The dependence of the null pronouns on the antecedents and operators for their interpretations means that they are in part an anaphor and variable, subject to a (c-command) control requirement.   In other words, they are used anaphorically and are dependent on operators for their interpretation. Nonetheless, they do not behave exactly like pure anaphors, as they are bound outside their minimal TPs. None of these properties indicate that they are pronouns that are not pronounced. They cannot therefore be *pro*. In fact, in terms of feature composition, they are emptier than *pro*s or even phi-phrases (ϕPs, Holmberg's (2005) term). The null arguments are not pure PRO in GB's sense either, since they can appear in governed positions, as subjects of finite verbs, and objects of verbs, as well as in the position of subject of the first serial verb (equivalent to English infinitives). What makes them look like PROs is the property of control: they are either controlled or they have a generic interpretation.   The upshot is another type of syntactically projected null argument found only in Thai and (possibly) in other discourse pro-drop languages where agreement morphology is absent from sentential syntax. Consider these examples:

(6.1) cɔɔn$_1$ bɔ̂ɔk wâa   *e*$_1$ pay yîipùn maa

　　John   say COMP   go Japan PERF

　　'John said that he had been to Japan.'

(6.2) Ø tham dii   Ø dâay dii

　　　do   good   get glory

　　(Lit.) 'One does good deeds and gets glory [in return].'

 (6.3) kim$_1$ rúusùk wâa   mii khɔn   mɔɔŋ *e*$_1$ yùu

　　 kim feel   COMP have person stare   PRES CONT

　　'Kim feels that someone is staring at her.'

(6.4) pleeŋ khláassìk chûayhây Ø phɔ̀nkhlaay

　　music classic   help   relax

　　'Classical music helps one relax.'

(6.1) and (6.2) exemplify null subjects with a referential reading and a generic reading, respectively. (6.3) and (6.4) exemplify null objects with a referential reading and a generic

reading, respectively. The obvious difference at the surface is that the examples having referential readings contain an antecedent, whilst the examples having generic ones do not. Beyond the surface similarity lies the fact that null arguments in both cases are controlled, either by the c-commanding matrix subject or by a generic operator (I will return to the issue on controllers in Section 6.2 below). The null pronouns in these examples appear in argument positions of TP projections (the second TP in the case of the embedded null argument).

Referential null pronouns in Thai, then, are always controlled. Correspondingly, null arguments without a controlling antecedent are not allowed to have a referential reading in Thai. I will return to this point in Section 6.4 below, by looking in more detail at control of null arguments in governed positions in finite clauses.

6.2 Types of Controllers and Null Arguments

We have established that a Thai null argument can be one of the following: it can be topic-linked, if there is an A-Topic in the opening discourse or in the preceding clause, as in (6.5); it can have a matrix antecedent, as in (6.6); it can be generic-operator bound, as in (6.7); it can be deictic-related, as in (6.8); it can be elided under identity with the object in the antecedent clause, as in (6.9); it can function like PRO, as in (6.10), or it can be a resumptive pronoun, as in (6.11):

(6.5) ceen$_1$ khàyǎn mâak cim bɔ̀ɔk wâa $e_{1/*2}$ pay hɔ̂ŋsàmùt tɛ̀ɛ   cháaw   thúkwan
     Jane studious very Jim say COMP   go   library   early morning everyday
     'Jane is very studious. Jim says that she goes to the library in the early morning every day.'

(6.6) cɔɔn$_1$ bɔ̀ɔk wâa $e_{1/\#2}$ pay yîipùn maa
     John   say COMP   go Japan PERF
     'John said that he had been to Japan.'

(6.7) Ø tham chûa   Ø dâay dii   mii thɔ́mpay
        do   evil   get glory has a lot
     'One does evil deeds and gets glory [in return]; there are a lot of such people.'

(6.8) (chǎn) dùam nɔm thúk wan

     I   drink milk every day

   'I drink milk every day.'

(6.9) pim klàbpay yîam phɔ̂ɔmɛ̂ɛ (*khɔ̌ɔŋ thəə) ceen kɔ̂ɔ klàbpay yîam e mɯ́ankan

     Pim go   visit parents   of-GEN she Jane then go   visit   also

  'Pim went to visit her parents. Jane went to visit her [own] parents as well.'

(6.10) cim$_1$ yàak $e_{1/*2}$ kin kɛɛŋ

     Jim   want   eat curry

   'Jim wants to eat some curry.'

(6.11) phûuyǐŋ khɔn níi$_1$ chǎn rúucàk phûuchaay chɯ̂ɯ cɔɔn thîi chɔ̂ɔb e_1 mâak

     woman CLS DEM I   know man   name John COMP like   very

   'This woman, I know a man called John that likes __ very much.'

These findings illustrate that third person null pronouns in Thai always have antecedents. (6.7) and (6.8) are similar, in that they both have a syntactic antecedent, i.e. a generic operator and a logophoric agent respectively in the CP-domain. The difference is that the subject in (6.7) must obligatorily be null, whilst the one in (6.8) can optionally be overt. I will argue in Section 6.3.3.1 that the speaker is always available as a local antecedent, and as such the subject need not be null. When a subject as in (6.8) is null, it is argued to be ɸ-featureless, being [uR, N], just like an inclusive G-pronoun and other referential null pronouns. (6.5) exemplifies an instance of a discourse antecedent, i.e. *ceen* 'Jane'. (6.5) is different from (6.6) and (6.10), since the antecedent in (6.5) does not actually c-command the embedded null subject, but the null subject is controlled via a null topic chain. In (6.6.) and (6.10), the antecedent argument directly controls the null subject. (6.9) illustrates a null object construction where the null object in the second clause contains an anaphor as the possessor of the object (Takahashi, to appear). In this case, (6.9) cannot have the reading that 'Jane visited Pim's parents', if the sentences are uttered out of the blue. Therefore, (6.9) shows an instance of obligatorily sloppy interpretation. (6.11) is an instance where a null realisation of a resumptive pronoun is allowed (despite being less preferred than its overt counterpart, given the long distance between the binder and the resumptive) when it is bound by a human left

dislocate. The relation of null arguments and antecedents is summed up in the following table:

Table 10: *Null arguments in Thai distinguished by their binder and controller*

	Types of Controllers/ Binders				
	C-commanding referential NP	**Discourse topic**	**Generic operator**	**Deixis** (*speaker by default*)	**NP in the antecedent clause**
Obligatorily null?	N	Y	Y[54]	N	Y
Antecedent control?	Y	Y	Y	Y	N[55]

The controller types which are the object of our interest here are c-commanding referential NPs, discourse topics, and generic operators, since they provide either a referential value or a generic value to null arguments. Both c-commanding referential NPs and discourse topics associate with referential interpretations. The former can directly control the embedded null subject and are thus overt controllers. Note that a c-commanding referential NP can control both resumptive and other pronouns. In contrast with these overt controllers, it is well known that other controllers may remain implicit (Higginbotham (1999); Landau (2000)), as in the case of an A-Topic operator and a generic operator. We know that such operators exist, since the interpretation of a null argument with and without them is not the same. Both of them are in an A'-position in the CP-domain of the clause containing the null argument. These operators, then, bind any co-indexed null arguments in their c-command domain.

In what follows, I discuss each type of null argument in terms of its obligatory or optional pronunciation. Since most null arguments in Thai are controlled, I will discuss in Section 6.4 what type of control relation exists between a null argument and its controller.

---

[54] The reading can only be generic inclusive.

[55] This does not include the case where strict identity is allowed, for instance,
(i) kim rák khruu (*khɔ̌ɔŋ khǎw) pim kôo rák *e* mɯ́ankan
    Kim love teacher of-GEN she   Pim also love   too
    'Kim loves her teacher. Pim loves her teacher, too.'
The example shows a potential for semantic ambiguity. It can have a reading either 'Pim loves her [own] teacher.' (sloppy reading), or 'Pim loves Kim's teacher.' (strict reading).

6.3 Phonological Realisation of Pronominal Arguments in Thai

The major controllers and the relation between them and their associated null arguments in Thai are the following:

(6.12) A null argument must be controlled or bound by *one* of the following:

   (i) a generic operator in Spec, CP, or

   (ii) a higher generic argument, or

   (iii) an A-Topic operator in Spec, CP, which itself is linked to a referential NP in
         the discourse, or

   (iv) 'the speaker', as default in Spec, CP, or

   (v) a c-commanding referential NP in A-position, or

   (vi) a c-commanding referential NP in A'-position without a θ-role (left dislocation).

According to this, a null argument (single *pro*, to use a GB term) can have only one referential index or it has a generic interpretation. Therefore, it is not possible that a null argument is bound/ controlled by multiple antecedents/ phrases.

6.3.1 Obligatorily null pronominal arguments

Pronominal arguments must be covertly realised when bound either by a generic operator (for a generic inclusive reading) or controlled via a null topic chain. As seen previously, the referential third person null pronoun in (6.5) and the inclusive G-pronoun in (6.7) are accounted for by the same operation of operator structures.

6.3.1.1 Inclusive G-pronouns

Inclusive G-pronouns in Thai are always null. The example below has a generic inclusive reading:

(6.13) hàak Ø wáicay khɔn   Ø cà   cɔncay ʔeeŋ[56]

        if   trust   person   FUT regret by one self

        'One will regret [later] if one trusts acquaintances.'

The sentence has the following reading, exemplifying a form of quantificational adverb binding (Diesing (1992)):

(6.14) GENERALLY$_x$ [x will regret it later if x trusts acquaintances]

As discussed in Chapter 4, the adverb quantifies over the null arguments, rather than over the event. Such an adverb appears to behave like an operator. The null arguments in turn are roughly equivalent to bound variables. Since the matrix null argument has the [uR] feature and is ϕ-featureless, the generic inclusive reading come from its not being restricted. It includes the speaker, the addressee, and any other people. As for the embedded null argument, also necessarily generic; two options are assumed. In the first one, it is controlled by the higher generic argument. That is, after the matrix null argument is probed by the generic operator, becoming [gn [N]], it is then able to control the lower generic argument. The second option is that both the null arguments are bound simultaneously by the generic operator, located in the matrix CP-domain. They start with [uR, N] in the syntax, and will then be generic-operator bound at the end of derivation. Having no restriction in reference, they therefore have a generic inclusive reading. (6.15) exemplifies that the lower argument takes the higher generic argument as its antecedent, as in the first option. (6.16) exemplifies the simultaneous binding of two null subjects by the generic operator:

(6.15a) [$_{CP}$ [$_C$ GN [$_{TP}$ [gn, N]   [$_{CP}$ [$_{TP}$ [gn, N]]]]]]

(6.15b) [$_{CP}$ [$_C$ GN [$_{TP}$ [gn, N]   [$_{CP}$ [$_{TP}$ [gn, N]]]]]]

To see which option works for Thai, consider this example:

---

[56] This is adapted from a well known Thai saying. The default interpretation would be generic. Given a proper context, it could have a referential reading, referring to the addressee.

(6.16) *(raw) phûut   phaasǎathìn thâa (*raw) phûut   dâay

       we   speak dialect   if   we   speak able

       'We speak dialect if we can.'

The second generic *raw* 'we' must obligatorily be null. This indicates that the generic operator only binds the higher argument. The fact that the lower argument is not pronounced means that it is directly controlled by the higher one as its antecedent, hence evidence in favour of *Option 1:* direct control (see Section 6.4). This is also the case with several occurrences of inclusive generics and arbitrary pronouns with a corporate reading. To illustrate, as long as the embedded quasi-inclusive generic is controlled by the higher quasi-inclusive generic which has ϕ-features, as in (6.16), it does not need (even cannot have) ϕ-features of its own. For this reason, it gets the interpretation of its antecedent, which is generic (because of the operator), but restricted generic because of the ϕ-features.

6.3.1.2 Null topic operators

According to Frascarelli (2007), an A-Topic operator should be analysed as occupying Spec, CP, which itself is linked to a referential NP in the discourse of the sentence containing a null argument. Probed by the A-Topic operator, the null argument can therefore have a referential reading. This presupposes that a topic chain is formed between the A-Topic operator and an established topic in the preceding/ opening sentence – the 'ultimate antecedent' of the null pronoun. The topic chain, exemplifying the inter-sentential relation between a null argument and its ultimate antecedent, does not involve a binding relation, as it does not presuppose c-command. Instead, I claim that there is a distinct, primitive, index-copying relationship between Topics (i.e. referential NPs in Spec, CP-position) across sentences in a discourse. Consider example (6.17):

(6.17) ceen$_1$ (nà) t_1 khàyǎn mâak cim bɔ̀ɔk wâa $e_{1/*2}$ pay hɔ̂ŋsàmùt tɛ̀ɛ cháaw

      Jane TM   studious very Jim say COMP   go library   early morning

      'Jane, she is very studious. Jim says that she goes to the library in the early mornings.'

The topic-chain formed between the topic operator and the A-Topic of the discourse can be schematised as below:

Figure 2: *Topic-chain relation between a referential 3rd person null argument and its ultimate antecedent*
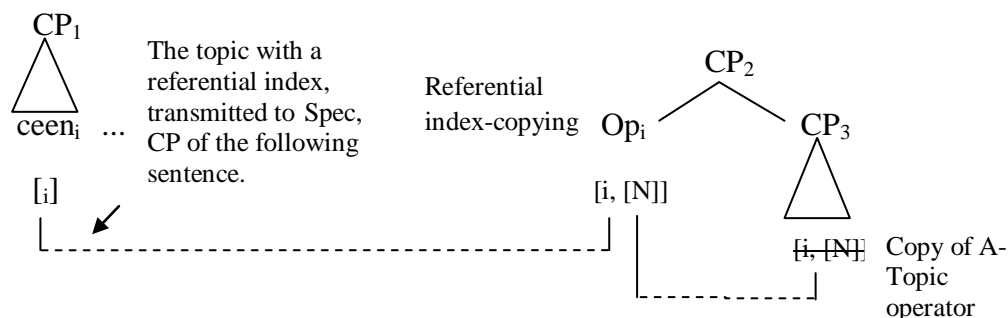
Suppose that CP_1 is the first sentence in the opening discourse. *Ceen* 'Jane' qualifies as A-Topic, the ultimate antecedent, assuming that it undergoes topicalization, marked by an optional topic marker. The topic, labelled with *i*, is syntactically represented in the designated A-Topic position, i.e. in Spec, CP_1. I assume that the second sentence in (6.17) has an indexless null topic in the CP-domain, resulting from A'-movement from its canonical position, following C.-T. J. Huang (1984 and subsequent), if there is no syntactic constraint on the movement as in the case of a complex island. The absence of an index is essentially like an unvalued feature, which must be assigned a value, i.e. an index; this is the [uR] feature. In other words, before valuation in Spec, CP_2, the null topic is simply a feature composition, making up a null argument, type [uR, N]. The referential index *i* (hence its content) is then abstractly transmitted from the ultimate antecedent to the Spec, CP_2 where the indexless null argument located. The index is the value assigned to [uR], represented in Figure 2 as Op_i. That is, the [uR] feature of the null argument is assigned the referential index *i* in the matrix Spec, CP_2 by means of index-copying, thereby becoming [i [N]] – a controlled, null topic with a referential value. This value is also transmitted to its copy (or 'trace') in the A-position, i.e. the embedded subject position in (6.17), which is in Spec, TP of CP_3 in Figure 2. Although the embedded subject has a referential index, and is therefore referential, it is not spelled out due to lack of ɸ-features. Overall, the relation in Thai between a null pronoun and its ultimate antecedent is direct, and not mediated by agreement, as in agreement pro-drop languages.[57] A null argument will be interpreted in relation to the closest available A-Topic. The null A-Topic analysis can be summarised as below:

---

[57] Pro-identification in agreement pro-drop languages is traditionally analysed as being dependent on the rich agreement element, where the content of a null subject can be identified by such an element (Rizzi 1986, among others). Recent analyses of Italian (Grimshaw & Samec-Lodovici (1998); Frascarelli (2007)) show that instead of pro-identification by I/Agr, third person subject drop of the Italian type has to link to a discourse topic,

(6.18)   the value, A-Topic operator

[*ceen*, i]...[$_{CP}$ [i [N]] [C [$_{TP}$ T [$_{VP}$ [i [N]] *v* VP]]]]

the ultimate antecedent   a copy acquiring the value through the chain

As a copy of the null topic without φ-features, it must obligatorily be null. In contrast, a topical constituent that is contrastive must be pronounced.

(6.19) ceen$_1$ khàyǎn mâak cim bɔ̀ɔk wâa *LòN$_1$* pay hɔ̂ŋsàmùt tɛ̀ɛ   cháaw   sùan

    Jane studious very Jim say COMP SHE go library   early morning while

    KHĂW yaŋ mây tùɯn thâa mây thîaŋ [...]

     HE   yet NEG get up if   NEG noon

   'Jane is very studious. Jim says that SHE goes to the library in the early morning, while

   HE does not get up before noon [...].

In this case, given the analysis that φ-features are the condition of overt spell-out, the pronouns have φ-features, so they are pronounced. This makes it possible for them to carry stress at PF, and to receive a contrastive interpretation at LF.

The availability of the A-Topic is crucial. If the complement clause in (6.17) constitutes an independent clause, and is uttered out of the blue, as in (6.20), it cannot be construed as having a third person subject.

(6.20) *e* pay hɔ̂ŋsàmùt tɛ̀ɛ   cháaw

     go library   early morning

   'I go to the library in the early morning.'

   (NOT: *'(S)he/ they go(es) to the library...'*)

---

usually the closest one (see also Sigurðsson (2011) for a similar idea). That is, a referential third person null pronoun is identified through *Agree* with the local A-Topic (Aboutness-shift Topic, in Frascarelli's term). The interpretation of a referential third person null pronoun, then, relies on the A-Topic. Holmberg et al. (2009), assuming Frascarelli's (2007) proposal, argue that the value of D is a referential index. If they are right, the relation between the null subject and the A-Topic will be indirect, crucially involving I/Agr: the null subject enters a chain with T containing a rich set of φ-features. In other words, the D-less pronoun is incorporated in T.

In the opening discourse, either the arguments have φ-features, for example, [3SG.F] and a referential index, as in (6.19) which yields a pronounced form, or they do not, for instance, as in (6.20), in which case they will be assigned the index of the speaker, by default.

6.3.1.3 Big PRO

I claim that null arguments in Thai are subject to control (see Section 6.4 for the issue of Control), and exclusively identified by antecedents/ operators. Without them, they cannot have referential readings, except for a first person referential reading. This is also true for the case of G-pronouns. They need to be identified through a generic operator, or they will be interpreted to be sloppily referential due to no available antecedent. As controlees, they are analogous to controlled PRO in that they require a controlling antecedent. Traditionally, two types of PRO are recognised: controlled PRO and arbitrary PRO (which typically means that it has a generic interpretation). Control in the literature is further classified into optional and obligatory control. The examples below exemplify the case of obligatory control and optional control respectively:

(6.21) Jane$_1$ tried [[PRO$_1$ to behave *oneself/ herself.]]

(6.22) Tim wondered [how [PRO to behave oneself/ himself.]]   Haegeman (1995: 277)

PRO in (6.21) is obligatorily controlled by an NP, namely *Jane*, and PRO cannot have an arbitrary (generic) reading. (6.22) shows two alternative readings: PRO can have a referential reading which is controlled by the matrix subject *Tim*, or it may have a (generic) arbitrary reading. A characteristic property of all instances of PRO in English is that it does not have a pronounced counterpart. This is explained by the fact that PRO must occur in 'ungoverned positions', where it will not be assigned Case. If we take this as a defining criterion, then the closest counterpart to PRO in Thai is found in Serial Verb Constructions (SVCs) where the Thai counterpart to PRO is used referentially.

In comparison with agreement pro-drop languages, Thai is crucially different with regard to [uφ]-features (i.e. agreement), provided that there is a relation between φ-features and (Abstract) Case, according to Chomsky (1995, 2001). If Case is the 'flip-side of agreement',

assigned as part of [uφ]-valuation, then I claim that Thai does not have Case either, or at least not Nom/Acc type Case, since Thai does not have any agreement.

Since Thai lacks agreement, the only difference between PRO and a null referential pronominal argument is that the former is found in non-finite (ungoverned) contexts, and the latter is found in finite (governed) contexts. It follows that PRO is only possible in SVCs, in the position of subject of the second/ third serial verb, and it is identified by an antecedent NP in contexts like the following:

(6.23) nɔ́ɔy   hǔŋ   khâaw PRO kin
        Noy   cook rice   eat
        'Noy cooked rice to eat.'

The verb *kin* 'eat' has a covert subject, assigned the Agent role. This subject is PRO, which is controlled by the antecedent, the subject argument of *hǔŋ* 'cook', thereby resulting in a control structure, involving two argument chains:

(6.24a) [$_{TP}$ nɔ́ɔy$_1$ [$_{VP}$ hǔŋ khâaw [$_{TP}$ PRO$_1$ [$_{VP}$ kin]]]]

The fact that PRO inherits the referential index from its antecedent argument indicates the similarity between PRO and a null referential pronominal argument in a finite clause. Put differently, null arguments in Thai have the same range of interpretations as PRO has: they are either controlled or generic. For these reasons, I propose that PRO and a null argument should be collapsed into a unified, single minimal null argument type, namely [uR, N]. The following exemplifies such an analysis, where the null argument of the type [uR] is controlled by a higher NP in an A-position:

(6.24b) [$_{CP}$ C [$_{TP}$ nɔ́ɔy$_i$ [$_{VP}$ hǔŋ khâaw [$_{TP}$ [i [N]] [$_{VP}$ kin]]]]]

That the two null nominal categories are essentially the same supports C.-T. J. Huang's (1984) typology of ECs, concerning *pro*/ PRO (cf. Table 6). Since I do not make a distinction between big PRO and null pronominal arguments in finite clauses, all null arguments in Thai have essentially the same feature composition, that of the null nominal category, i.e. [uR, N]

without ɸ-features. Thus, [uR, N] can occur in either A-positions or A'-positions (see also Section 6.4 below). This includes the case of resumptive pronouns when null. Such a null argument type [uR, N] is therefore dependent on binding or control that makes a difference in terms of the interpretation (see Section 6.4).

6.3.1.4 Null object arguments

An argument which is elided under identity with an object in the antecedent clause is not pronounced, as in (6.25):

(6.25) pim$_1$ khǎay thoorásàp mɯthɯ̌ɯ (*khɔ̌ɔŋ tuaʔeeŋ$_1$)] pay nók$_2$ kɔ̂ɔ khǎay $e_{*1/2}$
        Pim   sell   phone   mobile of-GEN self   go Nok also sell

        mɯ̌ɯnkan
        as well

        (Lit.) 'Pim's sold her mobile phone. Nok's also sold her [own] mobile phone.'

The fact that a strict reading is not allowed in (6.25) is due to pragmatic reasons. The possessor must obligatorily be null, since it is coindexed with the topic.

The occurrence of a sloppy reading is standardly explained by the assumption that ellipsis requires identity between the elided XP and the antecedent XP 'up to existential closure' (Merchant (2001)). That is, they count as identical as long as they have existentially bound variables in the same places, even if the binders are not identical. Since Merchant (2001), licensing conditions on XP-ellipsis have been understood as in part semantic in nature, in that the existential closures of the 'would-be elided clause' and the antecedent clause have the same truth conditions.

I need not make such a strong claim. Rather, I assume that the null object has no ɸ-features, and is made up of just the features [uR, N]. Note that XP-ellipsis here is understood as NP-ellipsis, rather than VP-ellipsis, following Otani and Whitman (1991), based on the fact that adjuncts are not susceptible to elision (see also Oku (1998); Takahashi, to appear), for instance:

(6.26) dam thamkhwaamsàʔàad bâan *yàaŋmɔ̀dcɔ̀d* ʔaathítlákhráŋ kim mây dâay tham *e*

Dam clean house thoroughly once a week Kim NEG able clean

(Lit.) 'Dam cleans the house thoroughly once a week. Kim does not clean.'

The first sentence contains the adjunct *yàaŋmɔ̀dcɔ̀d* 'thoroughly'; however, the second sentence does not have the reading with such an adjunct. It simply has the reading 'Kim does not clean the house.', not 'Kim does not clean the house thoroughly.'.

According to Takahashi (to appear), such a construction does not involve VP-ellipsis. If it did, it would seem that the interpretation in the second sentence should include the adverb. I therefore assume that V-stranding VP-ellipsis is not available in Thai at least in this example. I shall leave for future research the issue of whether corresponding sentences without such an adjunct are instances of VP-ellipsis.

(6.27a) exemplifies an instance of a null object construction, too, with an analysis as in (6.27b):

(6.27a) pim cəə [mɛ̂ɛ (*khɔ̌ɔŋ khǎw)] mûawaan nɔ́ɔy kɔ̂ɔ cəə [mɛ̂ɛ *e* mǔankan]

Pim meet mother of-GEN she yesterday Noy also meet mother as well

'Pim met her mother, and Noy also met her mother.'

(6.27b) pim$_i$ cəə mɛ̂ɛ (*khɔ̌ɔŋ khǎw$_i$) mûawaan nɔ́ɔy$_j$ kɔ̂ɔ cəə [uR, N]

Pim meet mother of-GEN she yesterday Noy also meet

'Pim met her mother yesterday, and Noy met her mother, too.'

The second sentence can mean 'Noy met Pim's mother or Noy met her [own] mother.' although the latter reading is preferable. A unified explanation can be achieved if we say that the elided NP *mɛ̂ɛ khɔ̌ɔŋ khǎw* 'her mother' where the reference for '*khɔ̌ɔŋ khǎw*' 'her' is not yet fixed (valued). It can be fixed through referential-index copying, either from 'Pim', i.e. [i[N]] or from 'Noy', i.e.[j[N]] depending on the intended reading and the context. Note that the third person possessive pronoun in the antecedent clause must be null since it is coreferent with the third person matrix subject.

It seems that the A-Topic in (6.27b) is 'the mother', rather than 'Pim' which is typically an A-Topic due to its initial position. Under the present theory, a null element is predicted to represent a null topic. Turning the other way round, the object 'the mother' appears to be the A-Topic assuming from its null pronunciation. In cases like this, it is obvious that there is A-Topic of the discourse. However, the subject is not the topic; the subject 'Pim' has switched to 'Noy'. Therefore, the objects in (6.27b) qualifies as the A-Topic.

6.3.2 Obligatorily overt pronominal arguments

As discussed in Section 6.3.1.2, one context where pronouns have to be overt is when they have referential readings associated with some discourse functions other than A-Topic. One such context is contrast. The impossibility of a contrastive null pronoun is due to the fact that a null pronoun cannot carry stress. On the other hand, being overt at PF means having ϕ-features in the syntax. For this reason, only an overt pronoun can convey the phonological expression of contrast. Viewed slightly differently, a pronominal argument must sometimes be null in order to avoid unwanted contrastive connotations (Minegishi (2011)).

Another context where a pronoun must be overtly realised is in the absence of discourse or linguistic antecedents. In pro-drop languages, there are basically two such cases. One is when a new argument (hence A-shift-Topic) is introduced in the discourse. The other is the case of a generic/ arbitrary pronoun. The former case in fact applies to agreement pro-drop languages, like Italian, where an overt pronoun (with a rising tone; a stressed strong pronoun in Frascarelli's term) has to be used to start a new topic chain, according to Frascarelli (2007). I have shown in Chapter 5 that in Thai, in such a case, a full NP (R-expression) is used instead. Therefore, an overt pronoun in Thai is used when it serves as focus or contrast or when it is directly controlled by a c-commanding antecedent, or when it has no antecedent. Pronouns without an antecedent will be interpreted to be generic quasi-inclusive, generic exclusive, or arbitrary (corporate readings), as discussed in Chapter 4. These must be overt, since they have restricted reference: first person and Plural, and third person and Plural, respectively. In (6.28) and (6.29) I give an analysis of pronouns before generic-operator valuation (to have generic quasi-inclusive, exclusive generic and (corporate) arbitrary readings at the end of derivation respectively):

(6.28) $[_{CP} [_{C} \text{GEN}_x [_{TP} \text{T} [_{vP} [\text{uR, 1PL, N}] v \text{VP}]]]]$ (*raw* 'we')

(6.29) [$_{CP}$ [$_C$ GEN$_x$ [$_{TP}$ T [$_{vP}$ [uR, 3PL, N] v VP]]]]] (*khǎw* 'they')

The pronouns have all the ϕ-features: [uR [ϕ [N]]]. The [uR] feature must then be bound by an implicit quantificational adverb serving as the (generic) operator. Since they have ϕ-features, they are more restricted than an inclusive G-pronoun and as such they are necessarily pronounced, thereby all yielding the following structure: [gn [ϕ [N]]], i.e. a quasi-inclusive G-pronoun, an exclusive G-pronoun, and an arbitrary pronoun having a corporate reading.

6.3.3 Optionally null pronominal arguments

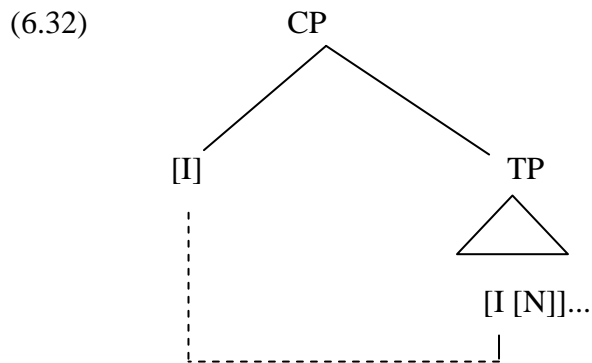### 6.3.3.1. The speaker as the default antecedent

First and second person null subjects are probably the most frequent kind of null arguments in spoken discourse. There are two possible variants of such subjects, in that they can optionally be spelled out. As for the first person null argument, it is found in the data that if an argument is dropped in a sentence (hence a non-generic statement) uttered out of the blue, as seen in (6.30) and (6.31), such a null argument will be construed as referring specifically to the speaker, even if the interpretation is pragmatically odd.

(6.30) *e*   ráb   thàay phâab nɔ̀ɔk sàthǎanthîi

     offer shoot photo out place

    'I/ *one/ *(s)he offer on-location photo shoot.'

(6.31) phleeŋ rɔ́kʔɛ̀ɛnroo thamhây *e* khlaay khrîad

    music   rock 'n roll make   lessen stress

    'Rock 'n roll music makes me/ *one/ *him/ *her feel less stressed.'

I propose that it is possible to give a unified account of such first person null pronouns and the third person null pronouns discussed above, which depend on binding by a topic or a c-commanding matrix constituent. This can be done by assuming that in cases like (6.30) and (6.31) there is an abstract local antecedent available as a controller of the null pronoun. For concreteness, I adopt the idea, articulated in Sigurðsson (2004), that every sentence has a logophoric agent represented as a feature – a so-called speech feature – in the CP-domain. In

the case of independent sentences, as in (6.30) and (6.31), the logophoric agent is always the speaker. And this agent is available as a local antecedent. The following illustrates a null pronoun having a first person reading.

(6.32)

```
              CP
            /    \
         [I]      TP
          ┊       /\
          ┊      /  \
          ┊     /____\
          ┊    [I [N]]...
          ┊       ┊
          └───────┘
```

For a second person reading of a null argument, more context is required, for instance:

(6.33) A: (thəə/ *chan) phûut   phaasǎathìn lǒə

       You/ I   speak   dialect   Q

       '(You/ *I/ *One) speak dialect, don't you?'

    B: mânlέεw

       yes

       'Yes, I do.'

(6.34) A: phleeŋ rɔ́kʔὲεnroo thamhây *e* khlaay khrîad dâay nîaʔná mâynâa chûa

       music   rock 'n roll make   lessen stress able PRT NEG believable

       'Rock' n' roll music makes (you/ *me/ *one) feel less stressed. It's unbelievable!'

    B: ciŋciŋ

       absolute

       'Absolutely.'

In these cases, the context disfavours a first person reading, for pragmatic reasons. In such a case, the addressee can be a licit antecedent of a null argument. Again, I rely on Sigurðsson (2004), who argues that the logophoric patient (i.e. the addressee) is represented as a speech-feature in the CP-domain of every sentence as well, and is thus available to serve as a local antecedent of a null argument. It may be noted here that Sigurðsson's theory is a modern version of the Performance Hypothesis of Ross (1970), according to which every declarative

sentence is embedded in a higher sentence meaning 'I say to you that __'. (See Section 6.5.2 below for an argument that null first/second person pronouns are ϕ-featureless, on a par with referential third person null pronouns.)

I assume that the fact that the first/second person pronoun is optionally spelled out is due to the abstract speech feature in the CP-domain being available as a local antecedent of *every* sentence, unlike the case of referential third person null pronouns, which depend on argument antecedents or an A-Topic operator. The fact that a first/ second person pronoun can be pronounced entails that there are several lexical options. If the option with ϕ-features (first or second person singular) is selected, the pronoun is pronounced. If the ϕ-featureless option is selected, the pronoun will be null. This is possible, since it can inherit a referential index from the logophoric agent or patient locally available in the Spec, CP of the sentence containing the null pronoun.

Having elaborated obligatory and optional null and overt pronouns, I next examine the core relation between antecedents and referential third person null pronouns, addressing in particular the question, 'What are the properties of control?'.

6.4 On the Issue of Control

Referential null pronouns in finite clauses in Thai have some properties in common with obligatory control (OC) of PRO as found in non-finite clausal complements in many languages (cf. (6.23) for controlled PRO). In Section 6.3.1.3, I have argued that PRO and a null pronoun in a finite clause are exactly the same, having only [uR, N] features. In other words, when a null argument of the type [uR, N] is in A-position, it can be directly controlled by a higher antecedent argument.

There is another type of control. Section 6.3.1.2 has shown that definite, referential third person null arguments are indirectly controlled via a null-topic chain. These null arguments undergo A'-movement to A-Topic operator, leaving behind a copy in their canonical positions (hence a null topic copy). In what follows, indirect control is discussed first, followed by direct control by an antecedent argument in a higher clause.

6.4.1 Indirect control

When a topic is explicitly established in the discourse, the context appears to favour a reading where such a topic is a null argument's antecedent. In this section, the embedded null subjects coindexed with the ultimate antecedents are primarily focused. As such, the analysis in Section 6.3.1.2 applies. Adopting a topic-movement account, as proposed by C.-T. J. Huang (1984 and subsequent)), a copy/ trace (a topic variable, to use C.-T. J. Huang's term) results from A'-movement of a null argument to A-Topic operator for a referential index valuation. The copy in the A-position then has merely a referential index transmitted from the operator, and is thus not pronounced owing to the complete lack of ϕ-features. As will be discussed in Section 6.5.2, this explains why a null referential argument cannot be a prepositional complement. The ultimate antecedent is thus an A-Topic in the discourse (see Figure 2). This is what I refer to as indirect control.

A discourse topic established in the preceding sentence, however, cannot serve as the antecedent of a null argument in an island if there is a matrix subject disjoint in reference from the discourse topic:

(6.35) ceen$_1$ t_1 (nà) wâadrùub kèŋ dam$_2$ bɔ̀ɔk waa nɔ́ɔŋsǎaw (*khɔ̌ɔŋ khǎw$_2$) kôo
Jane TM drawing good Dam say COMP sister of-GEN he then
wàad kèŋ chênkan
drawing good as well
'Jane, she is very good at drawing. Dam says that his/ *her sister is very good at drawing as well.'

(6.36) kim$_1$ bɔ̀ɔk nɔ́ɔy$_2$ wâa khɔ̌ɔŋ thîi cà hây e_2 thamhǎay paylέɛw
Kim tell Noy COMP thing COMP FUT give lose PERF
'Kim told Noy that she had lost the thing she was going to give her.'

In order for a discourse topic to be the antecedent, the null possessor in (6.35) and the null object inside the relative clause in (6.36) would have to undergo A'-movement to the Spec, CP of the matrix clause. However, due to the island constraint, they cannot move out of the NPs. This, then, is an argument that A'-movement is crucial in indirect control. In the preferred reading, the matrix subject in (6.35) is the antecedent of the null possessor; the

matrix object in (6.36) is the antecedent of the null object in the complex NP. In this regard, they are treated as an instance of direct control. (6.35) and (6.36) also show that the closest antecedent NP may, but need not, be a topic in order to control a null argument.

If there is no island, the reading that an embedded null subject is coreferent with the A-Topic (ultimate antecedent) is favoured, given that such an A-Topic is established explicitly in the discourse. In the following example, I show that even if the matrix subject in the second sentence is made a focal constituent by means of a cleft construction, the embedded null subject is still understood as being coreferent with the already-established A-Topic, namely *ceen* 'Jane', this being the ultimate antecedent:

(6.37) $ceen_1$ (ná) t_1 khàyǎn mâak cim_2 ŋay thîi   bɔ̀ɔk wâa $e_{1/ *2}$ pay hɔ̂ŋsàmùt
       Jane TM   studious very Jim FM COMP say COMP   go library
       tɛ̀ɛ cháaw   thúkwan
       early morning everyday
       'Jane, she is very studious. It's Jim that says that she goes to the library in the early morning every day. '

This example shows that a subject-control reading of the embedded null argument is ruled out if there is an explicit discourse topic available. When split antecedents are explicitly made topics, they can control the interpretations of null arguments as well. In the following example, the matrix subject: *cim* 'Jim' does not directly control the embedded null argument. Instead, the null embedded subject is a copy of the A-Topic operator. The structure of (6.38) is given in (6.39).

(6.38) $ceen_i$ (ná) t_i khàyǎn mâak pim_j (nà) t_j kôo khàyǎn mǔankan cim_k bɔ̀ɔk wâa
       Jane TM   studious very Pim TM   then studious too   Jim say COMP
       $e_{i+j/ *k}$ pay hɔ̂ŋsàmùt tɛ̀ɛ   cháaw
              go   library   early morning
       'Jane, she is very studious. And Pim, she is studious, too. Jim says that they go to the library in the early mornings.'

(6.39) [ceen, i] [pim, j]...[$_{CP}$ [$_{A\text{-}ToP}$ [i+j [N]] [C [$_{TP}$ [k [φ [N]] T [$_{CP}$ [$_{TP}$ T [$_{VP}$ [i̶+̶j̶ [̶N̶]̶] v
       VP]]]]

My informants and I agree that the null embedded subject acquires the reading 'they', so the two referential indices are construed here as just one unit as a group, i.e. 'they', rather than a conjunction of indices, i.e. *i*, *j*, *k*, etc. That is to say, the antecedents 'ceen' and 'pim' which form a single unit, i.e. 'they' simultaneously enter a topic-chain relation with the null topic. They are then transmitted to the topic operator of the sentence containing the null argument. This shows that a null argument in Thai can get the interpretation of 'they'. In this respect, the referential index is crucial in an interpretation of a referential third person null pronoun. Nonetheless, the embedded null subject is not pronounced, due to the absence of ϕ-features. This is a case of indirect control, which shows that split antecedents are possible.

On the other hand, split antecedents are sometimes not allowed when one of the arguments is not explicitly made a topic. (6.40) exemplifies such a case: the null embedded subject is preferably interpreted as coreferent with the matrix subject *nɔn* 'Non':

(6.40) nɔn$_1$ (ná) t_1 kin khâawcháaw kâp pim$_2$ kɔ̀ɔn $e_{1/??1+2}$ pay rooŋrian

      Non TM   eat breakfast   with Pim before   go school

      'Non, he ate breakfast with Pim before he went to school.'

All informants agree with me in preferring the reading where the matrix subject is the antecedent, i.e. '[...] before he (Non) went to school'. The split antecedent reading is judged as only very marginally acceptable. That the subject forms a natural topic explains, in part, why the matrix subject is the preferred antecedent. I assume that the two antecedents cannot enter a topic-chain relation simultaneously, since *pim* 'Pim' is not marked as a topic. *Nɔn* 'Non' then has separately to enter a topic-chain relation with the null topic. The split antecedents reading in (6.40) is therefore (almost) impossible. Alternatively, given that the topic is located in the same sentence as the null argument, to be precise in the next clause up, it could be that the null embedded subject is directly controlled by the matrix topic. In other words, the context dictates that the matrix subject in (6.40) is the topic, and thus an instance of subject control, blocking object control.

These examples show that an embedded null subject can be a copy of an A-Topic operator, assuming that there is a movement to the A-Topic-operator position, which in turn is linked to the topic established in the discourse (unless there are syntactic constraints on indirect

control). This indicates that there is a difference between the topic-chain case and the direct control case (even if the direct controller is a Topic, as in (6.40)).

Another relevant case of an A'-chain is resumptive pronouns. Coindexed resumptive pronouns often co-occur with left dislocations. According to Adger (2007), a resumptive pronoun can appear at an unbounded distance from a left-dislocated phrase, and there can be islands intervening between them. The relation between the resumptive pronoun and the left-dislocated phrase does not look like the kind of local control (via a null operator in Spec, CP) postulated for other referential null arguments, given the island boundaries. The locality in this study is understood as being the closest c-commanding NP without a θ-role, but the binder – the left- dislocated element – needs a θ-role from the resumptive pronoun. Otherwise, the sentence is not grammatical, owing to a θ-criterion violation. Given the fact that resumptive elements in Thai can be null, as exemplified in Chapter 5, I claim that when they have ϕ-features, they are pronounced. When they lack ϕ-features, they are null, just like other referential null arguments. The point is that they have to be bound[58] by the left-dislocated phrase, since they do not have a referential index, unlike referential *overt* pronouns where the ϕ-features and the referential index cling together. They then have to be A'-bound by the binder, i.e. the left-dislocated NP, base-generated in the CP-domain. The relation between the binder and the resumptive pronoun is not coreference. Rather, they form an A'-chain, since the left-dislocated NP gets a θ-role from the resumptive pronoun. In turn, the resumptive pronoun receives the referential index from the left-dislocated NP. (6.41) exemplifies an instance of a resumptive pronoun, with a structure as in (6.42):

---

[58] A bound resumptive pronoun is not similar to a quantifier-bound pronoun, since the former is referential, and thus requires the referential index from a left-dislocated NP. The quantifier-bound pronoun, which is null, is not referential, and is incapable of referring independently. It is a pure variable. However, a null argument as a quantifier-bound pronoun, (as in (i)), is applicable to other pro-drop languages than Thai, since quantifier-bound pronouns in Thai can be, and must be, overt (see also footnote 34). Applying the analysis, i.e. a null argument of the type [uR, N], I assume that the internal structure of a quantifier-bound pronoun after valuation for other pro-drop languages would be:

(i) Every doctor$_x$ [thinks that [x, N] is smart.]]

Note that x denotes a variable, not a referential index. I do not pursue this matter further here, primarily because a quantifier-bound pronoun in Thai is overt. In fact, Thai resorts to reflexives to obtain a quantifier-bound reading. That is, the variable in the embedded clause in (i) is replaced by an overt reflexive pronoun.

(6.41) [CP phûuyǐŋ khɔn níii [TP chǎn rúucàk [NP phûchaay chûɯ cɔɔn [CP thîi [TP chɔ̂ɔb

           woman CLS DEM I know man name John COMP like

           thəəi mâak]]

           she very

       'This woman, I know a man called John who likes her very much.'

the resumptive provides a θ-role for the left-dislocated NP (binder)

(6.42) [CP [phûuyǐŋ khɔn níi, i] [C [TP T [VP [i [ɸ [N]]] *v* VP]]]]

the binder provides the resumptive pronoun with an index through A'-binding

6.4.2 Direct control

As for direct control, a null pronominal argument controlled by an antecedent *argument* in a higher clause exhibits an OC-like structure. This happens when the embedded null subject is directly controlled by the antecedent argument from an A-position in the next clause up, instead of being linked to an A-Topic in the discourse via a Topic-operator in Spec, CP. Since the embedded pronominal subject is directly controlled by a locally higher argument, it can optionally be null. The fact that it can be pronounced entails lexical choices, just as in the case of first and second person pronouns. If the option with ɸ-features (third person singular/ plural) is chosen, the pronoun is spelled out. If the ɸ-featureless pronoun is chosen, it will not be spelled out. Locality is crucial for direct control and the spell-out of a pronoun. Consider example (5.17), repeated here as (6.43):

(6.43) A: câwtùub1 hǎay   pay sɔ̌ɔŋ wan lɛ́ɛw

       Ajax   disappear DIR two day PERF

       'My dog has been gone for two days!'

     B: cim2 bɔ̀ɔk waa (*khǎw2*) cəə *(man1) lɛ́ɛw [...]

       Jim say COMP he find it PERF

      'Jim said that he has found it [...]'

It is clear from (6.43) that the matrix subject need not be a topic in order to control the null embedded subject. The control relation here is direct, respecting locality. The matrix subject

and the controlled embedded null subject constitute two argument chains, bearing two separate θ-roles (they are agents of *bɔ̀ɔk* 'say' and *cəə* 'find', respectively), similar to the case of control of PRO. Locally A-bound from a θ-position, the null argument refers to the same entity as the controller: the matrix subject.

C-command is crucial for direct control of null arguments into finite clauses in Thai, as seen in the following example:

(6.44) nɔ́ɔŋ₁ (khɔ̌ɔŋ) ceen₂ tὗὗntên mâak cɔn $e_{1/\,*2}$ lɯam   thúkyàaŋ

　　　 sister of-GEN Jane nervous very   that   forget everything

　　　 'Jane's sister was so nervous that she forgot everything.'

The embedded null subject can only take a c-commanding argument as its antecedent, namely *nɔ́ɔŋ₁ khɔ̌ɔŋ ceen* 'Jane's sister'. For this reason, *ceen* cannot be a grammatical antecedent.[59] A strong argument that this is a case of direct control is the fact that split antecedents are not possible, as shown in (6.45) in comparison with (6.46):[60]

(6.45) nɔn₁ bɔ̀ɔk pim₂ wâa   *(phûak-khǎw₁₊₂) tɔ̂ŋ   thamŋaan hâyset wanníi

　　　 Non tell Pim COMP   they   must work   finish today

　　　 'Non told Pim that they must finish the job by today.'

If the embedded pronominal subject is overt, it can unproblematically can take the split NPs as its antecedent. But if the embedded subject is null, the preferred reading is the one where the antecedent of the embedded null subject is the matrix object, which is the closest. This supports the theory that control of null pronouns behaves like control of big PRO (see footnote 60). An alternative reading where the antecedent is the matrix subject is judged as

---

[59] Consider this example:

(i) *mɛ̂ɛ   khɔ̌ɔŋ   cɔɔn₁ rák e_1

　　 mother of-GEN John love

　　 'John's mother loves John.'

The fact that (i) is ungrammatical shows that binding without c-command is not possible; therefore, c-command is necessary (cf. Footnote 67).

[60] (6.45) and (6.46) can be compared with (i), a standard case of control, in English, where split antecedents are also impossible:

(i) John told Mary [PRO to finish the job] . ≠ 'John told Mary that they should finish the job.'

only very marginally acceptable, unless the matrix subject 'nɔn' is made a topic explicitly. This is due to locality (the 'minimal distance principle,[61] see Rosenbaum (1967)).

(6.46) Nɔn₁ bɔ̀ɔk pim₂ wâa   (khǎw₂/ *₁+₂) tɔ̂ŋ   thamŋaan hâyset wanníi

      Non   tell   Pim   COMP   she   must   work   finish today

      'Non told Pim that she must finish the job by today.'

Another possible interpretation of these structures is one where the null subject of the complement clause is generic, as in (6.47). Note that the controller of the embedded null subject may also be the matrix subject, in cases where the pragmatics appear to override locality:

(6.47) cim₁ bɔ̂ɔk nɔn₂ wâa $\emptyset_{GEN/ *1+2/ 1}$ tham dii   dâay dii

      Jim   tell   Non   COMP   do   good   get   glory

      'Jim told Non that one/ he does good deeds and gets glory [in return].'

The generic reading is possible because Spec, CP of the embedded clause can host a generic operator. As shown, the null subject argument in the complement clause cannot take split antecedents. If a split antecedent reading were allowed in (6.47), the relation of split antecedents and the embedded null subject would be an instance of coreference. This shows that we are dealing with direct control, rather than pragmatically governed coreference.

These cases show that not all referential null pronouns are controlled via a null topic chain. In other words, if there is no explicit topic or some strong pragmatic constraints, locality is crucial in determining which argument can control a null embedded subject. When there is a matrix object, it is the default controller. Otherwise, it is the closest subject:

(6.48) nɔn₁ bɔ̀ɔk   wâa ceen₂ bɔ̀ɔk   wâa   (khǎw₂) thùuk   huǎy

      Non   say   that Jane   say   that   she   win   lottery

      'Non said that Jane said that she had won the lottery.'

---

[61] Rosenbaum (1967), cited in Culicover & Jackendoff (2001) proposes that, "the position of the controller is determined by a "Minimal Distance Principle" (MDP): the NP closest in the tree to the infinitival is the controller. When there is no direct object, a complement infinitival is closest to the subject; when there is a direct object, a complement infinitival is closer to it than to the subject." This accounts directly for the possible and impossible readings in the following:
(i) John₁ got to shave himself₁/*oneself. (ii) John₁ got Fred₂ to shave himself₂,*₁/*oneself.

The preferred reading would be that the embedded null pronominal subject takes the closest argument, *ceen*, as its antecedent, rather than the matrix subject *nɔn*. The pragmatics, nonetheless, may favour the more distant subject if the matrix subject *nɔn* 'Non' is turned into an explicit topic, for instance by means of topicalization. The most deeply embedded null subject is then interpreted as coreferent with the matrix subject, even if it is not the closest controller. If this is the case, then the choice of controller is not determined by a formal locality condition, but for a pragmatic reason.

This means that the crucial measure of locality is 'the closest argument'. Another context where locality is not respected is when an intervening clause has no argument. In cases like that, there can be control by a distant antecedent which is more than one clause away:[62]

(6.49) cim bɔ̀ɔk wâa   chôokráay maak thîi   (khǎw) thùuk ród chɔn

     Jim say COMP unlucky   very COMP he   PASS car hit

    'Jim said that it was unlucky that he had been hit by a car.'

All of the examples discussed so far have shown that both types of control, direct control and indirect control via A-Topic operator, require the existence of matrix arguments and discourse topics, respectively. It is the context and certain syntactic/ pragmatic constraints that dictate which will be the antecedent of a null argument. We may also wonder what the interpretation of a null embedded subject is if a discourse topic, for instance, the one in (6.37) is removed, yielding a sentence like (6.50):

(6.50) cim_1 bɔ̀ɔk wâa $e_{1/*2}$ pay hɔ̂ŋsàmùt tɛ̀ɛ   cháaw

    Jim say COMP   go library   early morning

    'Jim says that he goes to the library in the early mornings.'

---

[62] Thai appears to be similar to Finnish and Brazilian Portuguese (BP) in this case. That is, both tolerate control of the null subject of a finite clause by a subject which is more than one clause away (Holmberg et al. 2009).

(i) A $Maria_1$ disse [que é verdade [que $(ela)_1$ entornou   o copo]].

   det Maria said   that is true   that she knocked.over the glass

   'Maria said it's true that she knocked over the glass.'   BP (Holmberg et al. 2009: 82)

(ii) $Jukka_1$ sanoi [että oli onni   [että $(hän)_1$ oli   voittanut arpajaisissa]].

   Jukka said   that was fortune that   he   had-3sg won   lottery-in

   'Jukka said that it was fortunate that he had won in the lottery.'   Finnish.

Suppose that the sentence is uttered out of the blue. Unlike the reading in (6.37), the only interpretation of the null embedded subject is that it is strictly coreferent with the matrix subject, as an instance of direct control. If the matrix subject is removed, then the null subject will be understood as referring to the speaker. The same principle applies to (6.49). When the matrix clause is removed, the sentence will be interpreted as 'It was unlucky that I was hit by a car.'. These facts show that control of null subjects in finite clauses relies on the availability of antecedent arguments or Topics. This applies to control of a null subject in non-finite clauses, too, in SVCs where an antecedent argument is required. Control, therefore, is a crucial requirement for referential readings of null arguments in Thai.

6.5 Other Syntactic Constraints on the Occurrences of Null Arguments

6.5.1 Conjoined NPs

In a conjoined NP, one pronoun cannot be null if the other one is overt. This, I assume, is due to a general requirement on a coordinate structure. The coordinator requires the conjoining of two syntactically equivalent elements.

(6.51) *(nǔu) kàp nɔ́ɔŋ   (*khɔ̌ɔŋ nǔu) khɔ̌ɔ pay duu năŋ nákhá
        I   and younger sister   of-GEN I   request go   see film PRT
     'May I and my younger sister go to see the film?'

(6.52) * dɛɛn kàp   *e* rák   kan
         Dan   and   I   love each other
      'Dan and I love each other.'

The internal structure of the referential first person null pronoun in (6.52), i.e. unpronounced *nǔu*, in comparison with the other conjunct, overt *dɛɛn*, is shown in Table 11:

Table 11: *The internal structure of the 1st and 3rd person pronouns in (6.52)*

	nǔu	**dɛɛn**
ϕ-feature encoded	N	Y
phonologically overt	N	Y

Coordinate structure generally requires parallel syntactic structure between the two coordinated NPs. The ungrammaticality of (6.51) and (6.52) indicates that the parallel relation of the coordinate structure is not respected. One of the conjoined NPs does not have ϕ-features, thereby being covert/ unpronounced at PF. For this reason, the sentence cannot satisfy the symmetry required by coordination (Kehler (1996)).

6.5.2 Prepositional complements

In Chapter 2, it is demonstrated that a null argument in Thai can appear in any environment, except as a complement to a preposition. I argue that this constraint derives from properties of the preposition itself:

(6.53) * naan tâŋcay   rian   phûa   *e*
         Nan pay attention study for
       * 'Nan paid attention to her studies for.'

(6.54) # cim phùut thúŋ   thîiwâa (phûak)-khǎw cà   maa cəə ceen
          Jim talk about COMP   they   FUT come meet Jane
        * 'Jim talked about that they would come to meet Jane.'

In Pesetsky's (2011) words, a preposition is "picky about the nature of its complement". For example, in English, it requires a complement with ϕ-features; a CP-complement will not do.[63] As shown by (6.54), the same constraint holds in Thai as well, in which case the sentence sounds very odd, given the presence of the CP-complement. I have argued that null pronouns in Thai are 'minimal nouns', consisting of only a nominal feature and a [uR]

---

[63] McCloskey (1991), cited in Pesetsky (2011) presents some evidence that a declarative CP lacks ϕ-features, as can be seen in the following English example:
(i) That the position will be funded and that Mary will be hired now seems/ *seem likely.
Although the subject consists of two declarative CPs, the main verb is still inflected for third person singular (with -*s*). For this reason, according to Pesetsky (2011), the construction *[P CP] is ill-formed in English. Next, consider a Thai example:
(ii) phôɔ   yûŋ mɛ́ɛ   kɔ̂ɔ mây yùu nân/ *làwnàn mǎaykhwaam
     father of-Gen I busy mother of-Gen I then NEG here that/ those   mean
       wâa chán tɔ̂ŋ   tham thúkyàaŋ eeŋmɔ̀d
     COMP I   must do   everything self
'My father is busy. My mother is not here. That means/*Those mean I will have to do everything by myself.'
Since Thai lacks agreement, the argument cannot be duplicated, as seen that 'that' is used as default. Due to no agreement in the grammatical system, there is no independent support that CPs lack ϕ-features in Thai.

feature. Having no φ-features they cannot be pronounced. I have also argued for the stronger claim that pronouns with φ-features must be pronounced. One obvious case was that of the generic pronouns, for instance, an exclusive G-pronoun which must be pronounced. The ungrammaticality of (6.53) and (6.54) follows, without any further assumptions: they are ruled out because the complement of the preposition lacks φ-features.

With regard to null arguments, the claim that they are φ-featureless is sufficient to explain why a null pronoun as a prepositional complement is ungrammatical. If a preposition is removed, the object can be, and must be, null, given that it is controlled via a null topic chain:

(6.55) [...] naan tâŋcay pay phɔ́b   (*khǎw)

        Nan   intend   go   see   him/ her

    ' [...] Nan intended to go to see him/ her.'

The generalisation that null arguments in Thai are φ-featureless is thus supported by the fact that they cannot serve as prepositional complements.

Note that a preposition cannot even have a null first or second person pronoun as its complement. This supports the claim made earlier that all null pronouns in Thai are the same, carrying only [uR, N] and being φ-featureless. The ungrammaticality of a referential null pronoun after the preposition means that the null pronoun does not receive any φ-features from a controller. If it did, (6.53) would be well-formed. In fact, all that a null pronoun inherits from its antecedent is the referential index. This is all that is required for the interpretation: the null pronoun has the *same* reference as the antecedent.

6.6 The applicability of the φ-Featureless Null Noun Account to Other   Discourse Pro-drop Languages

Wang et al. (1992) summarise the three principal approaches to the null subject problem that have been proposed in the literature:

(6.56) a. [$_{CP}$ *pro*$_i$ [$_{Infl}$ Agr/Tense] ...]

     (identification by Agr, Italian) Rizzi, (1986); Jaeggli & Safir (1989)

b. Discourse Topic$_i$ [$_{CP}$ topic$_i$ [$_{TP}$ t_i ...]]

 (identification by a discourse topic, Chinese) C.-T. J. Huang (1984, 1989)

c. Subject$_i$ verb [$_{CP}$ e_i VP]

 (identification by a c-commanding NP, Chinese) C.-T. J. Huang (1984, 1989)

According to Wang et al., there are crucial differences between the grammars of agreement pro-drop languages on the one hand, and discourse pro-drop languages on the other hand. This results in two distinct methods of identification of null arguments. However, recent research has cast doubt on this classification. Thus, it appears that option (6.56a) does not really exist (see footnote 57), since agreement pro-drop languages (both the consistent and the partial type: see Holmberg (2005, 2010a); footnote 15), have *pro*-identification via a null topic strategy, especially in the case of third person *pro*, as argued by Samek-Lodovici (1996), Frascarelli (2007) and Modesto (2008), among others.

Modesto (2008) argues that Brazilian Portuguese, Finnish and Chinese are similar in that they license and identify null subjects regardless of active verbal agreement. He proposes that referential null subjects are actually elided topic (PF-deleted elements), being in topic position as a result of abstract topicalization. This is in line with Grimshaw & Samek-Lodovici (1998) and Frascarelli (2007). They argue that a third person null subject in Italian refers to an entity introduced as a topic of the discourse. Frascarelli proposes that the A-shift Topic is involved in the interpretation of null subjects, in the light of the theory of the left periphery in Frascarelli & Hinterhölzl (2007). The function of the A-shift topic is to (re)introduce a topic in the discourse, and it should be kept distinct from other types of topics such as F-Topics, C-Topics, and H-Topics, as in principle, they behave markedly differently from one another, as discussed in Chapter 5.

Options (6.56b, c) exist in Thai; the former has been discussed above as an instance of indirect control, and the latter as an instance of direct control. As also discussed in Chapter 3, C.-T. J. Huang (1989) was the first to propose the identification of an empty category by means of the closest nominal element in its control domain, if it has one. This was reformulated later as the Generalised Control Rule (GCR).[64]

---

[64] C.-T. J. Huang's GCR states that a pro/ PRO is controlled in its control domain (if it has one):
Control Domain:
α is the control domain for β if it is the minimal category satisfying (i) and (ii):

If *pro* has a control domain, it must have a local, unique, referential antecedent, otherwise it will have a non-referential reading, which may be generic or arbitrary. This basically means that a null argument is always controlled, by virtue of coindexing with a referential NP wherever possible. Having no agreement morphology, Thai has the control domain in the higher category, since the minimal TP containing a null pronoun is not the control domain. For this reason, Barbosa (2011) adopts C.-T. J. Huang's (1989) idea that *pro* can only take an embedded subject position where it can be directly controlled by a higher argument, and is a variable bound by a null topic elsewhere. The GCR nicely captures the case of direct control of a null subject into a finite clause where the matrix clause contains only a subject argument. But when the matrix clause contains an object apart from the subject, the GCR fails to explain the well-formedness of the reading where either a matrix subject or a discourse topic, but not the matrix object, is a licit antecedent of the null subject. My informants also note that a null embedded subject in Chinese can even take a split antecedent, although it is less acceptable than taking the matrix subject or a discourse topic. If the sentence appears in isolation, it will be ambiguous, as in (6.57):

(6.57) Meilin$_1$ gaosu Lisi$_2$ $e_{1/3/\,1+2}$ bu   neng jintian gongzuo

　　　 Meilin tell   Lisi   NEG can   today   work

　　　 'Melin told Lisi that she/ they can't work today.'

This means that Chinese is not as strict as Thai in terms of the locality condition. There are more examples that cannot be explained in the light of C.-T. J. Huang's (1989) GCR:

(6.58) kim$_1$ bɔ̀ɔk wâa $e_{1/\,2}$ khɔŋcà   mây maa

　　　 Kim say COMP   probably NEG come

　　　 'Kim said that probably (s)he won't come.'

(6.59) nít$_1$ khít wâa   mɛ̂ɛ$_2$ (*khɔ̌ɔŋ khǎw) mây rák e_1

　　　 Nit think COMP mother of-GEN her   NEG love

　　　 'Nit thinks that her mother doesn't love her.'

(i) α is the lowest S or NP containing (a) β, or (b) the minimal maximum category containing β;
(ii) α contains a SUBJECT accessible to β.

(6.60) rák dèkdèk

love children

'*One/ I love(s) children.'

(6.61) thîi phâaknŭa *(raw) phûut phaasăathìn

at north we speak dialect

'In the north, we speak dialect.'

(6.62) phûuyĭŋ khon níi$_i$ chǎn rúucàk phûchaay chŵɯ cɔɔn thîi chɔ̂ɔb e_i mâak

woman CLS DEM I know man name John COMP like very

'This woman, I know a man called John who likes her very much.'

Given the intended reading, the GCR does not account for the fact that antecedent assignment to a null pronoun is dictated by the context. To illustrate, the closest nominal element for the null subjects in (6.58) and (6.59) is the matrix subject, namely *kim* 'Kim' and *nít* 'Nit' respectively. 'Kim' constitutes an accessible subject located in the matrix clause, which is predicted by the GCR to be the control domain of the null subject. Nonetheless, the possible antecedent of the null embedded subject in (6.58) is not restricted to the matrix subject but can be a discourse topic. On the other hand, the closest nominal element to the null object in (6.59) is the embedded subject, which cannot be the antecedent, due to the GCR and the binding principles. However, the null object is not exclusively a topic variable, as C.-T. J. Huang claims for Chinese.[65] In fact, the matrix subject *nít* 'Nit' is a licit antecedent, exhibiting a direct control, ruling out the topic-variable reading without any movement. In (6.60), the preferred interpretation of the null subject is that it is referring to the speaker. It cannot have a generic inclusive reading, as the context is referential, rather than generic. In (6.61), the context dictates that the sentence has a generic quasi-inclusive reading due to the

---

[65] A null object must be a topic variable, according to C.-T. J. Huang. When there is a null object, GCR requires the embedded subject to be the antecedent of the null object. In fact, such a subject cannot be its antecedent, since Binding Principle B requires the null object to be disjoint in reference from its antecedent. Principle C also disallows coreference between the matrix subject and the embedded object. Without the grammatical control domain, they can only have references fixed in non-argument positions outside the sentence. However, as noted by my informant, a similar reading as in (6.59) can be constructed even in Chinese where the null object can be co-indexed with the matrix subject, contra C.-T. J. Huang's claim:

(i) Meilin$_1$ jue de youren zai kan (ta)$_1$

Meilin think somebody at look her

'Meilin thinks that somebody is looking at her'

presence of a locative adverbial. Based on the data, it is clear that null arguments in Thai can be controlled inside or outside their control domains. C.-T. J. Huang's proposal also cannot account for a null object case as in (6.62) where the null object is not exclusively subject to the same constraint as movement (Y. Huang (2000)). In fact, in (6.62), the null object is a resumptive pronoun, rather than a topic variable, due to island movement constraints.

Crucially, the well-formedness of the null pronominal object in (6.59) entails that Thai null arguments do not fit in the restrictive framework of Binding Theory nor the GCR. Instead of postulating the possibility of a null variable object in the presence of a topic operator and the impossibility of an object *pro*, I assume that a null object as well as a null subject exhibit a feature composition of [uR, N] without ϕ-features. They can be either directly controlled (although control of a null object is not found as frequently as control of null subject) or they can be resumptives which are dependent on a left-dislocated NP, or a copy (or trace) of a moved null topic.

I claim that null pronouns in discourse pro-drop languages generally are also ϕ-featureless, made up of only [uR, N]. They are therefore dependent on being controlled either by a locally c-commanding referential NP or by 'the speaker', as default; see Sigurðsson (2004). Alternatively, a null subject can be interpreted as a copy of a null topic. Controlling by a higher generic argument is also a possibility. In what follows, I discuss how this claim can account for other discourse pro-drop languages.

6.6.1 Null arguments in other discourse pro-drop languages

The examples below, taken from Chapter 3, exemplify the distribution and interpretation of null arguments in Korean and Chinese:

(6.63) John-i$_1$ e_1 Mary-lul   poassta-ko malhayssta

     John-Nm   Mary-Acc saw-Comp said

     'John said that (he) saw Mary.'   Korean: Choo (1994: 18)

(6.64) Xiaohong$_1$ de   meimei$_2$ shuo $e_{1/2}$ xihuan tan   gangqin

     Xiaohong GEN younger sister say   like   play piano

     'Xiaohong's younger sister says that (she) likes to play piano.' Chinese: Y. Huang

                                            (2000: 66)

Assuming that all null arguments are ϕ-featureless, they must rely on their antecedents to acquire a designated interpretation. In (6.63), there is no doubt that the embedded null subject is directly controlled by a higher argument in the matrix clause, and as such, it can optionally be null.

For (6.64), Y. Huang (2000: 66) claims that there are a number of potential antecedents for the embedded null subject, namely *Xiaohong* or *meimei* 'sister' or even a third person, given the right context. If the sentence is uttered out of the blue, then the interpretation is that *Xiaohong de meimei* 'Xiaohong's sister' is the preferred antecedent of the embedded null argument.[66] The translation into Thai[67] has a similar reading, with the c-commanding subject being the antecedent. Assuming my analysis is right, the ϕ-featureless option is chosen, and the embedded null subject in (6.63) is not pronounced. It is still interpretable, as it inherits the referential index from the c-commanding antecedent. This shows an instance of direct control of the null subject into a finite clause.

(6.65) Meilin$_1$ gaosu Lisi$_2$ $e_{1/*2}$ bu   neng jintian gongzuo

     Meilin tell   Lisi   NEG can   today work

     'Meilin told Lisi that she can't work today.'

(6.66) Meilin$_1$ gaosu Lisi$_2$ $e_{*1/2}$ bu   yao   jintian gongzuo

     Meilin   gaosu Lisi   NEG have to today work   Chinese: C.-T. J. Huang

     'Meilin told Lisi that she doesn't have to work today.'   (1989:21)

In both (6.65) and (6.66), there are two higher potential c-commanding controllers, the matrix subject and matrix object. It is not possible, though, for the embedded null subject to take split antecedents, provided that the locality requirement is respected. If one assumes direct control of the null subject into the finite clause in (6.66), then *Lisi* qualifies as the antecedent due to locality. Nonetheless, the ungrammaticality of *Lisi* as the antecedent of the null embedded subject in (6.65) entails that the context forces the reading where the matrix

---

[66] This was checked and confirmed by several informants.

[67] nɔ́ɔŋ$_1$ khɔ̌ɔŋ   sìawmɛ̌y$_2$ bɔ̀ɔk wâa $e_{1/\#2}$ chɔ̂ɔb lên pianoo
  sister of-GEN Xiaomei   say COMP   like play piano
  'Xiaohong's younger sister says that (she) likes to play piano'

subject is the antecedent; the use of *neng* 'can' means that the speaker is talking about herself. In other words, the modal rules out the object control reading. In (6.66), the informants agree that the matrix subject is a possible antecedent of the null embedded subject when *Meilin* is the explicit topic, even if the interpretation may be odd. Therefore, there are various ways in which the context appears to dictate whether the subject or the object of the matrix clause is a likely antecedent.

Next, consider the following examples:

(6.67) Toli$_1$-ka   Swuni-ka e_1 kwoylophi-ess-ta-ko   malha-ess-ta.

Toli-Nom   Swuni-Nom   (OBJ)   tease-Past-Dec-Quote   say-Past-Dec.

'Toli said (that) Swuni teased (him).'   Korean: Han (2006: 13)

(6.68) mɛ̂ɛ   (khɔ̌ɔŋ chǎn)   hěn   pim$_2$ kɔ̀ɔn   (khǎw$_{\#1/2}$) pay   hǎa   mɔ̌ɔ

mother   of-GEN I   see   Pim   before   (s)he   go   see   doctor

'My mother saw Pim before she went to see the doctor.'   Thai

(6.69) huandi$_1$ yao   tade chezi   chaojian $e_{1/*2}$

king   want   his   ministers   see

'The king wanted his ministers to see him.'   Chinese: Xu (2003: 87)

The Korean example in (6.67) is straightforward. The embedded null object looks for its antecedent, which of course cannot be the embedded subject due to Principle B. It then ends up inheriting the referential index from the matrix subject in the next clause up, thereby observing locality. The relation between the embedded null subject and the antecedent is direct subject control. The same explanation applies to Chinese, as in (6.69) as well. Pragmatically, Xu (2003) notes that the verb *chaojian* 'see' is reserved for the context of the monarchy, thereby supporting that (6.69) exemplifies an instance of subject control.

In (6.68), the clause *kɔ̀ɔn* (*khǎw$_{\#1/2}$*) *pay hǎa mɔ̌ɔ* 'before (s)he went to see the doctor' contains the adverbial subordinator *kɔ̀ɔn* 'before', so the whole clause constitutes a sentential

temporal adjunct adverbial.[68]   Being in a c-command position, the matrix object is a licit antecedent of the null embedded subject, respecting locality requirement. The matrix subject can be the antecedent if it is explicitly made a topic. This example shows that Thai allows a null subject in a finite clause directly controlled by the nearest NP argument which is in a higher clause. The null embedded subject can optionally be pronounced due to the local (close-enough) antecedent.

(6.70) fuqin   shuo   *e* yao   weiren zhengzhi
      father   say   should upright
      'Father says that one should be upright.'   Chinese: Y. Huang (1994: 36)

(6.71) *e* pam-uy   etwum   sok-eyse kil-ul   ilh-l   ttay cili-nun
        night-Gen darkness midst-at   way-Acc   lose-RelEnd when geography-Top
      maywu   phenliha-ta.
      very   convenient-Dec.
      'When one gets lost in the darkness of the night, geography comes in very handy.'
                                        Korean: Han (2006: 72)

(6.72) John-wa kono beddo-de-wa yoku nemu-reru-to   iu.
      John-top this bed-in-top   well sleep-can-comp say
      'John says that one/he can sleep well in this bed.' Japanese: Holmberg et al. (2009: 79)

Based on the fact that there is no available antecedent and that the pronouns having a generic inclusive reading in (6.70) and (6.71) are not pronounced, I assume that there is an implicit quantificational adverb 'generally' serving as a generic operator available to bind the [uR] feature of the null arguments in such examples. After the binding, they would have generic inclusive readings due to no restriction on the reference. The null argument in each example, therefore, becomes [gn [N]] at the end of derivation. In (6.72), the sentence is ambiguous between a referential reading and a generic inclusive reading, in spite of the presence of a locally c-commanding referential NP. The ambiguity comes from the insufficiency of the

---

[68] Adverbial clauses marked by an initial subordinator in Thai commonly occur in both sentence-initial and sentence-final position, although Thai is rather rigidly VO. This supports Diessel (2001) who studied the ordering distribution of main and adverbial clauses, and found that if adverbial clauses in language *x* are marked by initial subordinators, then the chance is that such clauses can either precede or follow the main clause. A null subject tends to be found in such adverbial clauses, rather than in the main clause.

context. If the context makes clear that the embedded subject is generic inclusive, then I claim that there is an implicit generic operator located in the embedded CP-domain of the clause containing the null argument giving it a generic reading. The inclusiveness comes from the unrestricted reference of the null argument. Interestingly, these discourse pro-drop languages, including Thai, share the property that a generic inclusive reading is expressed by a null argument. Bound by the generic operator, the null pronoun therefore is able to refer to people in general, including the speaker, the addressee, and other people. I shall discuss further in Section 6.7 the generalisation of this unrestricted use of referential third person null subjects *and* a null generic inclusive subject pronoun.

(6.73) Zhangsan$_k$ shuo ziji$_k$ de haizi xihuan Xiaohong. Lisi$_i$ shuo *e* xihuan Xiaoli.

     Zhangsan say self of child like   Xiaohong Lisi say   like   Xiaoli

     (Lit.) 'Zhangsan said his child liked Xiaohong. Lisi said *e* liked Xiaoli.'

                                                     Chinese: Takahashi (to appear: 36)

(6.74) kim$_i$ bɔ̀ɔk wâa   lûuklûuk (*khɔ̌ŋ thəə$_i$) chɔ̂ɔb pay sǔansàt nɔ́k$_j$ bɔ̀ɔk wâa *e*

     Kim say COMP children of-GEN she like   go   zoo   Nok say COMP

     chɔ̂ɔb pay (*sǔansàt) mǔankan

     like   go   zoo   as well

     (Lit.) 'Kim said that her children liked to go to the zoo. Nok said that *e* did, too.'

As argued by C.-T. J. Huang (1984), cited in Takahashi (to appear), a similar sentence construction in Chinese, as in (6.73), does not allow the sloppy identity. The pronominal interpretation of the embedded subject in the second sentence is the only option. In other words, the embedded null subject in (6.73) is not derived by ellipsis. Assuming that the embedded null argument in both examples is ϕ-featureless, it must be directly controlled by the matrix antecedent NP, namely *Lisi* and *Nok*, yielding [i[N]] and [j[N]] respectively and supporting C.-T. J. Huang's claim. The examples show both Thai and Chinese are perfectly fine with the interpretation of the null argument of the type [uR, N] in which the control relation respects locality.[69]

---

[69] Things are not straightforward in Japanese. Oku (1998), cited in Takahashi (to appear), argues that null subjects can yield sloppy interpretation in Japanese, contra Thai and Chinese:

(i) a. Taroo-wa [zibun-no kodomo-ga eigo-o   hanasu to] omotteiru.

     Taro-TOP self-GEN child-NOM English-ACC speak that think

     (Lit.) 'Taroo thinks that self's child speaks English.'

When there is an appropriate context, ambiguities are resolved, as can be seen in Korean:

(6.75) U1: Cijo$_2$-nun chaykpelley-i-ta.

        Cijo-Top bookworm-Cop-Dec.

        'Cijo is a bookworm.'

    U2: ecey-to e_1 e_2 secem-eyse manna-ess-nuntey,

      yesterday-also   bookstore-at meet-Past-And

      '(I) met (him) at the bookstore yesterday.'

    U3: $e_{*1/2}$ chayk-man ilk-ko   iss-ess-ta.

         book-only   read-AuxEnd is-Past-Dec.

      '(*I/he) was just reading books.'   Kim (2003), cited in Han (2006: 95)

The topic, namely *Cijo*, in U1 is transmitted to Op in Spec, CP of the clause in U2 with a referential index. The null object in (U2) is, then, a copy of the A-Topic operator. The only reading that the null object can have is that it refers to the A-Topic of the discourse, i.e. the ultimate antecedent, and this applies to the null subject in (U3) as well. Therefore, the null object in (U2) and the null subject in (U3) represent the null A-Topic. The null subject in U2 is understood as referring to the speaker, as the context makes clear that the speaker is talking about a third person. The interpretation of the referential null pronominal subject as referring to the speaker, as indicated by the translation, shows that the speaker is available as a local antecedent in the CP-domain of the clause, in accordance with Sigurðsson (2004). It appears then that discourse pro-drop languages also pattern similarly with regard to the correlation of a first person reading and a null pronominal subject. That is, it shows that *only* the subject, but not the object, can receive the default first person reading. When a sentence with a null object is uttered out of the blue, it need not have a first person reading.

---

    b. Ken-wa [*e* furansugo-o   hanasu to] omotteiru.

     Ken-TOP   French-ACC speak that think

     (Lit.) 'Ken thinks that self's child speaks French.'

The present theory cannot account for the data like the embedded null subject in (i) which can have the reading that 'Ken thinks his [own] child speaks French.'(hence sloppy identity). In this case, Takahashi analyses the embedded null subject as genuinely derived by ellipsis. One way to see this issue is that control does *not* operate the same way in Japanese as in Thai and Chinese. Where control, under locality, is the default in Thai and Chinese (it applies where it can), that is not the case in Japanese. What the analysis is, exactly, of the Japanese sentence we can only speculate. Is there a noun 'child' which is deleted under identity, leaving it a pragmatic matter whose child? I shall leave this complication to future research.

Although there is a need for more data, the sample sentences in discourse pro-drop languages analysed above illustrates that Chinese, Japanese, and Korean tend to pattern like Thai in relation to the interpretation of null arguments. That is to say, a null argument composed of [uR, N] features requires either an antecedent or a generic operator in order to provide it with a referential index or a generic reading, i.e. 'it is genrally true for *x*'. As for referential null pronouns, I have argued that they do not need ϕ-features as long as they inherit a referential index, thus supporting that the working hypothesis is on the right track. None of them allows a third person referential null argument without a local linguistic antecedent or clause-external discourse topic. For instance, in (6.63), if the matrix subject is removed, the null pronoun of the clause will be understood as referring to the speaker. Therefore, null arguments in discourse pro-drop languages are not of the type *pro*, but belong to a different empty category, ϕ-featureless null arguments.

6.7 The Status of Thai in the Typology of Pro-drop Languages

Holmberg (2005, 2010b) observes that null-subject languages have either a null third person referential pronoun, *or* a null third person inclusive generic/ impersonal pronoun. The Romance and Slavic null-subject languages belong to the former class. That is to say, they have an overt inclusive generic pronoun. In particular, they resort to the overt reflexive morpheme *se:*

(6.76) Aqui não *se* pode nadar.

'One can't swim here.'   European Portuguese: Holmberg (2010b: 228)

Hausa, a consistent pro-drop language, also patterns similarly to the above languages. It uses an overt impersonal pronoun to expresses a generic inclusive reading. According to Jaggar (2001: 207f.), cited in Holmberg (2010b: 228), *á* (with several allomorphs) is the special impersonal pronoun in Hausa, referred to as the fourth person plural pronoun (4PL pronoun) in grammatical studies of this language. Brazilian Portuguese, Finnish, Marathi, Hebrew, and probably many other Indo-Aryan languages belong to the latter type. That is to say, they have restricted use of referential null subjects, but they have a null, inclusive G-pronoun:

(6.77) Hya khurchi- war aaramani   bushushakto.

      this chair   -on comfort-with sit-PRES.3SG

      'One can sit comfortably in this chair.'   Marathi: Holmberg (2010b: 93)

The following table summarises referential third person pronouns and inclusive G-pronouns in the languages illustrated above:

Table 12: *Referential 3rd person pronouns and G-pronouns*

	Agreement pro-drop languages				**Discourse pro-drop languages**	
	Consistent NSLs		**Partial NSLs**			
	Hausa	Italian	Finnish	Brazilian Portuguese	Thai	Chinese
Impersonal (expletive) 3rd person null pronoun	Y	Y	Y	Y	Y	Y
Referential 3rd person null pronoun	Y	Y	Y (by a higher NP only)	Y (by a higher NP only)	Y	Y
Generic inclusive overt pronoun	*à*.PL	*si*.SG	N	N	N	N
Generic inclusive null pronoun	N	N	Y	Y	Y	Y

Although discourse pro-drop languages appear to pattern more like the Romance and Slavic null-subject languages, e.g. Portuguese, Greek, and Italian (in the sense that both types of languages allows referential third person null subjects) rather than like Finnish, Marathi, Icelandic, etc., such discourse pro-drop languages have no referential agreement, whilst consistent NSLs do. According to Holmberg, "the presence of a D(efinite)-feature in T means that a null ϕP that enters into an Agree relation with T can be interpreted as definite, referring to an individual or a group" (2005: 555). Since consistent NSLs have 'referential agreement', a third person null pronoun is interpreted as referential. They have to have an overt indefinite pronoun to serve as a G-pronoun, thus there are no null G-pronouns in such languages. Nonetheless, if consistent pro-drop languages really have [D] in T, then an impersonal (expletive/ locative) third person null pronoun is predicted to be spelled out, on a par with an overt G-pronoun. In fact, such an impersonal (expletive/ locative) third person null pronoun) is null, as shown in the table. Holmberg (2010a: 115) solves this problem in the following manner: consistent pro-drop languages do not have a 'phonological EPP', formally a feature

[P] in T. For this reason, they can have a non-thematic null subject. On the other hand, Sheehan (2006), cited in Biberauer (2007:9), formalises the null subject parameter as presence *vs* absence of T [uD], instead of an interpretable D-feature (T[D]). An expletive can then be considered as suitably impoverished pronouns, which can be deleted under identity with the feature located on T.

On the other hand, partial NSLs have agreement, but it is not referential – there is no Definiteness feature [D] in T; the only way they can have a referential third person null pronoun is by means of being bound (controlled) by a higher referential NP. Without controlling antecedents, null third person pronouns will have to be interpreted as generic (the third person case) or expletive, according to Holmberg (2010a).

In contrast with these agreement pro-drop languages, Thai, Chinese, and Korean have no agreement. They depend on the discourse context for the interpretation of a null argument. It seems that the agreement is crucial in determining if a language can have either a null third person referential pronoun, or a null third person inclusive generic pronoun. In fact, the agreement neither helps the identification of a null argument nor licenses *pro* or anything else (see Section 6.6); instead the interpretation of a third person null subject relies on the antecedent (similarly to discourse pro-drop languages), as argued by Samek-Lodovici (1996); Frascarelli (2007), and Modesto (2008), among others.

As seen in Table 12, the case of Thai and Chinese shows that the complementary distribution found in agreement pro-drop languages is not a universal but is subject to parametric variation. That is to say, consistent pro-drop languages have [uD] feature in T; partial pro-drop languages lack [uD] in T. Discourse pro-drop languages, however, neither have agreement nor any features in T – there are no uϕ-features in T. For this reason, discourse pro-drop languages exhibit highly unrestricted use of referential third person null subjects and a null inclusive generic subject pronoun. Therefore, a controlling antecedent/ operator is crucial for the identification of null arguments in Thai and other discourse pro-drop languages.

Based on the fact that the interpretation of null arguments in discourse pro-drop languages requires the presence of an antecedent/ generic operator, it appears that these languages belong to a class of languages that exhibits the following traits:

(i)   There can be a proper null argument which is ϕ-featureless in Spec, TP or VP, or a copy of the moved null topic (copy of null Op) which is ϕ-featureless.

(ii) Being ϕ-featureless nominals, they have to receive a referential index by control or from the left dislocate, or by a topic chain, or they receive inclusive generic interpretation from a generic operator.

Chapter 7: Conclusion

Rizzi's (1986: 545) speculation that "Universal Grammar offers the option of using φ-features, and some grammatical systems take it, whereas others do not..." is based on the observation that discourse pro-drop languages such as Chinese, which apparently belong to the latter group, allow very frequent use of pro-drop despite no φ-features in the grammatical system. Therefore, they are assumed to rely on a totally different way of licensing null pronouns from agreement pro-drop languages. In this thesis, I have shown exactly how and in what way they are different from agreement pro-drop languages, as well as the effects of the presence and the absence of φ-features in the pronominal structure on the spell-out of pronouns.

The thesis has discussed the syntax-discourse interface of pro-drop in Thai. It started out with a working hypothesis that null pronouns in Thai are preferred, and that overt pronouns are used when there are restrictions on the realisation of null pronouns. Such restrictions include syntactic constraints and some discourse functions.

The thesis has, it is hoped, made a contribution to the literature on the problem of null arguments in Thai. To begin with, a syntactic analysis of the phenomenon of null arguments in Thai as a whole has been offered – both referential and non-referential. Descriptively, Thai has a positive value for the pro-drop parameter. In essence, null arguments in Thai are dependent on either an antecedent or operator; otherwise, they are interpreted as referring to the speaker as default. A critique of existing theories of null pronominal arguments in discourse pro-drop languages has then been provided, by demonstrating the limitations of competing theories. Most theories attempt to account for null argument phenomena by focusing on null arguments found in particular sentence constructions, and thus the analyses risk being incompatible with the interpretation of null arguments found in other sentence types. Another limitation is the analysis of a null argument in terms of a government-and-binding-based typology of empty categories. The alternative theory I propose is that pronominal arguments are φ-featureless and, as such, are not pronounced, consisting only of [uR, N] features, where the [uR] feature is assumed to be the head, being valued externally. Put differently, [uR] is assumed to be the head of [uR, N], the way D is taken to be the head of [D NP]. Together, the [N] and [uR] features can function as an argument, but they do not

refer independently. Therefore, the null arguments have to either receive a referential index by control or by a topic chain, or they have to receive generic reference from a generic operator. This is a postulate, basic to the understanding of null argument phenomena found in any type of sentence.

The previous chapters have argued in favour of the idea that null pronominal arguments in Thai have no ϕ-features, therefore are not pronounced, and rely on their antecedents for their interpretation. This is a referential reading, based on the fact that a referential null pronoun cannot have a third person reading unless there is a third person antecedent available (first and sometimes second person reading is possible, but that is because there is always a first and typically also a second person antecedent). To have a third person reading, the features [uR, N] must be controlled by either a c-commanding referential NP in a higher clause or via a null topic chain. Such a null pronominal argument does not pick up the ϕ-features of its antecedent. Instead, it is the referential index that the null pronoun requires from the antecedent. If the null pronoun inherited the ϕ-features, too, it would be able to be a complement of a preposition, which demands ϕ-features of its complement. ϕ-features, therefore, do not seem to play any role for the null pronouns in Thai (see also Hoonchamlong (1991) for a similar idea. The referential third person null pronoun after valuation is just [i [N]] (where i is a referential index), and is not pronounced. This also provides an answer to a key issue in Chapter 2, which states "One important question which arises regarding null pronouns is how the features of person, gender, and number of a syntactic gap are determined and whether the gap has internal structure."

The gap, or null pronominal argument, does have internal structure, i.e. [uR, N]. Significantly, a referential index is all that is required for the interpretation of a null pronoun: having the same reference as the antecedent without involvement of ϕ-features. Assuming the present account is right, then a genuinely null pronoun in Thai does not exist. A null argument is therefore an instance of a radically null noun which is just used pronominally. Null pronominal arguments in Thai share the same range of interpretation as PRO in GB theory, either controlled, generic, or arbitrary. In fact, they are the same thing as PRO, and it is clear why: they occur in environments where there are no [uϕ]-features, i.e. no agreement.

In the case where a pronominal argument can optionally be null, it is logophorically linked to a c-commanding matrix argument, and as such will be directly controlled, given that the context does not mark the matrix argument as an explicit topic. In other words, the [uR] feature of a null argument needs a referential index from its antecedent which need not be a topic. Having said that, it is not the case that an embedded pronominal argument is always simply coreferent with a higher argument in the matrix clause. It is the context that determines the interpretation of a null pronominal argument. The context may favour a more distant antecedent, an A-Topic of the discourse, which can be either a definite topic or an indefinite topic. If this is the case, such a null argument will be a copy of a null A-Topic operator, assigned a referential index via a topic chain. Hence, a null argument which is coreferential with the A-Topic of the discourse does not depend on the c-command relation and must be null.

Resumptive pronouns occur in connection with a left dislocation operation. The resumptive pronoun is in the core part of the clause, but is linked to an element in the CP-domain in the left periphery. When overt, these are bound, pronounced pronouns without the referential index. Thus, they must be bound by the left-dislocated NP. Therefore, this constitutes another type of overt pronoun, separate from coreferential overt pronouns in which the ϕ-features and the referential index cling together. When null, they are analysed as different from other null, controlled pronominal arguments, since the resumptive pronoun provides a θ-role for the binder, which is not the case for other null, controlled pronominal arguments nor copies of null topics. This applies to resumptive pronominal arguments in Thai. In English, however, a resumptive pronoun cannot be null.

English does not have [uR, N] in subject position of finite clauses, nor in object position. That the [uR, N] is not allowed in subject position of a finite clause is straightforward: T requires ϕ-feature values. In fact, neither a null resumptive pronoun nor a copy of a null topic is allowed in object position; this suggests that v has uϕ-features as well and therefore requires an object with ϕ-features. According to Chomsky (2001), Accusative case is assigned when the uϕ-features of v are valued. On the other hand, Thai allows object pro-drop, which is either a null (resumptive) pronominal argument or a copy of a null topic. This implies that the ϕ-feature valuation of v may be absent in Thai (see also Takahashi (to appear) for a similar idea with regard to the absence of uϕ-features of T and v in Japanese). I shall leave this issue open for future research.

In contrast to referential null pronouns, the quasi-inclusive G-pronouns, exclusive G-pronouns, and arbitrary pronouns with a corporate reading in Thai have antecedentless readings. They are claimed to be composed of [uR [φ [N]]], and as such they must be bound by a generic operator. They are referentially defective, since they have no index. Having the [uR] feature, they can be bound (valued) by a generic operator. At the end of derivation, they are necessarily pronounced, due to a more restricted reading, yielding [gn [φ [N]]]. The corresponding referential pronouns have their own referential index [i [φ [N]]], and they are also pronounced. On the other hand, the argument without φ-features, would get the generic inclusive reading, since this is the most general unrestricted reading, including the speaker, the addressee, and other people inclusive.

To summarise, Thai, and probably Korean, Chinese, and Japanese too, belong to a class of languages which have no uφ-features in T, whilst agreement pro-drop languages like Italian, Turkish, Finnish, and Marathi, or non-pro-drop languages like English and Germanic languages, and many others, do. Therefore, the subject (in finite clauses) in agreement pro-drop languages has to have φ-features. In Thai, it can do without these features, if it can get a referential index from an antecedent, or be generic-bound.

Since the subject in languages with agreement has to have φ-features, it can be null only if the uφ-features of T are so rich that all the features of the pronoun are represented in T, after valuation. In that situation, the subject pronoun will be a copy of T, and can therefore be deleted, according to Roberts (2010a,b) and Holmberg (2010a). A new typology of null nominal categories is therefore proposed as follows:

(i)   Null argument = [uR, N] in A-position (PRO in GB terms)

(ii)   Null operator = [uR, N] in A'-position

(iii)   Deleted copies, including the deleted copy of incorporation (by Agree) in T (*pro*)

      a.   Deleted copy of A'-movement   (wh-trace)

      b.   Deleted copy of A-movement   (NP-trace)

Note that (iii) is the case in Italian-type languages and other agreement pro-drop languages, but not in English. In English, agreement is weak; the subject pronoun cannot be a copy of T. For this reason, it must be pronounced. This provides an explanation for C.-T. J. Huang's

(1989) observation that languages with rich agreement and languages with no agreement both allow pro-drop but languages with weak agreement do not.

Bibliography

Ackema, P. et al. 2006. 'The role of agreement in the expression of arguments'. In P.Ackema et al. (eds.) *Arguments and Agreement*. Oxford: Oxford University Press. 1–34.

Adger, D. 2007. 'Agreement and Resumption.' <http://www.llf.cnrs.fr/Autres/Resumptives/AgreementResumption.pdf.> (accessed 6 June 2011).

Aroonmanakun, W. 1999. *Extending Focusing for Zero Pronoun Resolution in Thai*. PhD, Georgetown University.

Barbosa, P. 2011. Pro-drop and Theories of pro in the Minimalist Program Part 2: Pronoun Deletion Analyses of Null Subjects and Partial, Discourse and Semi pro-drop. *Language and Linguistics Compass* 5/8 (2011): 571–587.

Biberauer, T 2007. Handout to 'New approaches to null subject phenomena – romance style null subject languages' (lecture given on 28 July 2007: Summer school of generative grammar EGG 2007, Masaryk University). <http://www.phil.muni.cz/~docekal/egg/handouts/EGG%20Null%20Subjects%20Lecture%203%20(ho).pdf> (accessed 13 January 2011).

Biberauer, T. et al.2010. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge, UK: Cambridge University Press.

Brody, M. 2011. On the 'universal-impersonal' readings of personal pronouns. Talk presented at ICSH 10, Lund University.

Cabredo Hofherr, P. 2003. 'Arbitrary readings of 3pl pronominals'. In M. Weisgerber (ed.) *Proceedings of the Conference "sub7— Sinn und Bedeutung',* 7th Annual Meeting of the Gesellschaft für Semantik'. Arbeitspapier Nr. 114, FB Sprachwissenschaft, Konstanz University.

2006. 'Arbitrary *pro* and the theory of pro-drop'. In P. Ackema, P. Brandt, M. Schoorlemmer and F. Weerman (eds.) *Arguments and Agreement*. Oxford: Oxford University Press. 230–260.

2010. 'Binding properties of impersonal human pronouns in generic and episodic contexts'. Paper presented at *Workshop on impersonal human pronouns* UMR 7023 Structures formelles du langage CNRS /Université Paris-8, 28th May 2010.

Calabrese, A. 1986. 'Pronomina. Some properties of the Italian pronominal system'. In N. Fukui, T. R. Rapoport and E. Sagey (eds.) *Papers in Theoretical Linguistics*. 1– 46. Cambridge, MA: MIT Press.

Campbell, R. 1969. *Noun Substitutes in Modern Thai: A Study in Pronominality*. The Hague: Mouton.

Casielles-Suarez, E. 2004. *The Syntax-Information Structure Interface: Evidence from Spanish and English*. New York: Routledge.

Cinque, G. 1988. On si constructions and the theory of ARB. *Linguistic Inquiry*19: 521–81.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

1982. *Some Concepts and Consequences of the Theory of Government and Binding* Cambridge, MA: MIT Press.

1986. *Knowledge of Language: Its Nature, Origin, and Use*. Praeger: New York.

1993. A minimalist program for linguistic theory. In K. Hale and J. Keyser (eds) *The view from Building 20: essays in Linguistics in honor of Sylvain Bromberger*. Cambridge, MA: MIT Press, 1–52.

1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

2000. Minimalist inquiries: the framework. In R. Martin, D. Michaels and J. Uriagereka (eds.) *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: MIT Press. 89–156.

2001. 'Derivation by phase'. In M. Kenstowicz (ed.) *Ken Hale: A Life in Language*. Cambridge, MA: MIT Press. 1–52.

2008. On phases. In R. Freidin, C. Otero and M-L. Zubizarreta (eds.) *Foundational issues in linguistic theory*. Cambridge, MA: MIT Press, 133–66.

Choo, M. 1994. *A Unified Account of Null Pronouns in Korean*. PhD, Universality of Hawaii.

Cole, P. 1987. 'Null objects in universal grammar'. *Linguistic Inquiry* 18: 597–612.

Crystal, D. 2008. *A Dictionary of Linguistics and Phonetics*. Malden, MA: Blackwell.

Culicover, P. and R. Jackendoff. 2001. Control is not movement. *Linguistic Inquiry* 32: 493–512.

Dekkers, J. et al. 2000. *Optimality Theory: Phonology, Syntax, and Acquisition*. Oxford: Oxford University Press.

Diesing, M. 1992. *Indefinites*. Cambridge, MA: MIT Press.

Diessel, H., 2001. The ordering distribution of main and adverbial clauses. A typological study. *Language* 77: 433–455.

Egerland, V. 2003. 'Impersonal pronouns in Scandinavian and Romance'. *Working*

Papers in Scandinavian Syntax 71: 75–102.

Erteschik-Shir, N. 1997. *The Dynamics of Focus Structure*. Cambridge, UK: Cambridge University Press.

Féry, C. 2007. 'Information structural notions and the fallacy of invariant correlates'. In C. Féry, G. Fanselow and M. Krifka (eds.) *Interdisciplinary Studies on Information Structure*. Vol. 6. ISIS, Working Papers of the SFB 632, University of Potsdam. 161–184.

Frascarelli, M. 2007. 'Subjects, topics and the interpretation of referential pro: An interface approach to the linking of (null) pronouns'. *Natural Language and Linguistic Theory* 25: 691–734.

2010. 'Narrow Focus, clefting and predicate inversion'. *Lingua* 120.9: 2121–2147. September 2010.

Frascarelli, M. and R.Hinterhölzl. 2007. 'Types of Topics in German and Italian'. In S. Winkler and K. Schwabe (eds.), *On Information Structure, Meaning and Form*. Amsterdam/ Philadelphia: John Benjamins. 87–116.

Givón, T. 1983. 'Topic continuity in discourse: A quantitative cross-language study'. *Typological Studies in Language*. Vol. 3. Amsterdam: John Benjamins.

Grimshaw, J. and V. Samek-Lodovici. 1998. 'Optimal subjects and subject universals'. In P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis and D. Pesetsky (eds.) *Optimality and Competition in Syntax*. Cambridge, MA: MIT Press.193–219.

Grosz, J., A.K. Joshi and S.Weinstein. 1995. 'Centering: A framework for modelling the local coherence of discourse'. *Computational Linguistics* 21:2: 203–225.

Gundel, J. 1988. *The Role of Topic and Comment in Linguistic Theory.* New York: Garland.

Haegeman, L. 1995. *Introduction to Government and Binding Theory*. Cambridge, MA: Blackwell.

Halliday, M.A.K. 1967. 'Notes on transitivity and theme in English'. *Journal of Linguistics* 3: 199–244.

Han, N. 2006. *Korean Zero Pronouns: Analysis and Resolution*. PhD, University of Pennsylvania.

Heim, I. 1982. *The Semantics and Definite and Indefinite Noun Phrases*. PhD, University of Massachusetts at Amherst.

Higginbotham, J. 1999. A plea for implicit anaphora. In H. Bennis, P. Pica, and J. Rooryck (eds.) *Atomism and Binding*. Dordrecht: Foris. 183-203.

Holmberg, A. 2005. 'Is there a little pro? Evidence from Finnish'. *Linguistic Inquiry* 36: 533–564.

    2010a. 'Null subject parameter'. In T. Biberauer et al. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge, UK: Cambridge University Press. 88–124.

    2010b. 'The null generic subject pronoun in Finnish: a case of incorporation in T'. In T. Biberauer et al. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge, UK: Cambridge University Press. 200–230.

Holmberg, A. et al. 2009. 'Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish, and Marathi'. *Studia Linguistica* 63:59–97.

Holmberg, A. and M. Sheehan. 2010. 'Control into finite clauses in partial null-subject languages'. In T. Biberauer et al. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge, UK: Cambridge University Press. 125–152.

Hoonchamlong, Y. 1991. *Some Issues in Thai Anaphora: A Government and Binding Approach*. PhD, University of Wisconsin-Madison.

Hornstein, N., J. Nunes and K.K. Grohmann. 2005. *Understanding Minimalism*. Cambridge Textbooks in Linguistics. Cambridge, UK: Cambridge University Press.

Huang C.-T.J. 1984. 'On the distribution and reference of empty pronouns'. *Linguistic Inquiry* 15: 531–574.

    1989. 'Pro-drop in Chinese: a generalized control theory'. In O. Jaeggli and K. Safir (eds.) *The Null Subject Parameter*. Dordrecht: Kluwer Academic Publishers. 185–214.

    1991. 'Remarks on the status of the null object' In R. Freidin (ed.) *Principles and Parameters in Comparative Grammar*. Cambridge, MA: MIT Press. 56–76.

Huang, Y. 1994. *The Syntax and Pragmatics of Anaphora: A Study with Special Reference to Chinese*. New York: Cambridge University Press.

    2000. *Anaphora: A Cross-Linguistic Study*. Oxford: Oxford University Press.

Iatridou, S. and D. Embick. 1997. 'A propos pro'. *Language* 73:58–80.

Iwasaki, S. and P. Ingkaphirom. 2009. *A Reference Grammar of Thai*. Cambridge: Cambridge University Press.

Jaeggli, O. and K. Safir. 1989. *The Null Subject Parameter*. Dordrecht: Foris.

Jaggar, P. 2001. *Hausa*. Amsterdam: Benjamins.

Jensen, T. J. 2009 'Generic variation? Developments in use of generic pronouns in late 20th century spoken Danish'. *Acta Linguistica Hafniensia* 41.1: 83–115.

Juntanamalaga, P. 1988. 'Social issues in Thai classifier usage'. *Language Sciences* 10. 2: 313–330

Kamio, A. 2001. 'English generic we, you, and they: An analysis in terms of territory of information'. *Journal of Pragmatics* 33: 1111–1124.

Kamp, H. 1981. 'A theory of truth and semantic representation'. In J. Groenendijk, T. Janssen and M. Stokhof (eds.) *Formal Methods in the Study of Languages*. Amsterdam: Mathematisch Centrum. 277–322.

Kehler, A. 1996. 'Coherence and the coordinate structure constraint'. In *Proceedings of the 22nd Annual Meeting of the Berkeley Linguistics Society*: 220–231.

Kiss, É. K. 1998. 'Identificational focus versus information focus'. *Language* 74: 245–273. 2008. 'Topic and focus: Two structural positions associated with logical functions in the left periphery of the Hungarian sentence'. In C. Féry, G. Fanselow, and M. Krifka (eds.) *The Notions of Information Structure. Interdisciplinary Studies in Information Structure.* Volume 6. Potsdam: Universitätsverlag.

Kitagawa, C., and A. Lehrer. 1990. 'Impersonal uses of personal pronouns'. *Journal of Pragmatics* 14: 739–59.

Koenig, J.-P. and N. Muansuwan. 2005. 'The syntax of aspect in Thai'. In *Natural Language and Linguistic Theory* 23.2: 335–380.

Lambrecht, K. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge Studies in Linguistics. Cambridge, UK: Cambridge University Press.

Landau, I. 2000. *Elements of Control: Structure and Meaning in Infinitival Constructions*. Dordrecht: Kluwer Academic Publishers. 2003. 'Movement out of control'. *Linguistic Inquiry* 34: 471– 498.

Lasnik, H. 1986. 'On the necessity of binding conditions'. In H. Lasnik *Essays on Anaphora*. Dordrecht: Kluwer Academic Publishers (1989). 149–167. [Also published in R. Freidin (ed.) *Principles and Parameters in Comparative Grammar*. Cambridge, MA: MIT Press (1991). 7–28].

Lyons, J. 1977. *Semantics.* 2 vols. Cambridge, UK: Cambridge University Press.

McCloskey, J. 1991. 'Clause structure, ellipsis and proper government in Irish'. *Lingua* 85: 259–302.

Merchant, J. 2001. *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford: Oxford University Press.

Minegishi, M. 2011. Description of Thai as an isolating language. *Social Science Information* 50.1: 62–80.

Modesto, M. 2008. 'Topic prominence and null subjects'. In T. Biberauer (ed.) *The Limits of Syntactic Variation*. Amsterdam: John Benjamins. 375–410.

Moltmann, F. 2006. 'Generic one, arbitrary PRO, and the first person'. *Natural Language Semantics* 14: 257–281.

Montalbetti, M. 1984. *After Binding: On the Interpretation of Pronouns*. PhD, MIT.

Oku, S. 1998. *A Theory of Selection and Reconstruction in the Minimalist Perspective*. PhD, University of Connecticut.

Otani, K. and J. Whitman. 1991. V-raising and VP-ellipsis. *Linguistic Inquiry* 22: 345-358.

Pesetsky, D. 2011. Handout to 'Islands, case and licensing: the neglected role of the attractor' (lecture given on 28 February 2011, Newcastle University).

Prince, E. F. 1981. 'Toward a taxonomy of given/new information'. In P. Cole (ed.) *Radical Pragmatics*. New York: Academic Press. 223–255.

2006. Impersonal pronouns in French and Yiddish: semantic reference vs. discourse reference. In B. Birner and G.Ward (eds.) *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn*. Philadelphia/Amsterdam: John Benjamins. 295–315.

Qinan, H. 2008. *A Corpus-Based Study on Zero Anaphora Resolution in Chinese Discourse*. PhD, City University of Hong Kong.

Reinhart, T. 1981. 'Definite NP anaphora and c-command domains'. *Linguistic Inquiry* 12: 605–635. <http://www.jstor.org/stable/4178245>. (Accessed 27 June 2009).

1983. *Anaphora and Semantic Interpretation*. Croom Helm linguistics series. London: Croom Helm.

Rizzi, L. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

1986. 'Null objects in Italian and the theory of pro'. *Linguistic Inquiry* 17: 501–57.

1997. 'The fine structure of the left periphery'. In L. Haegeman (ed.) *Elements of Grammar. Handbook of Generative Syntax*. Dordrecht: Kluwer Academic Publishers. 281– 337.

Roberts, I. 2010a. *Agreement and Head Movement: Clitics, Incorporation and Defective Goals*. Cambridge, Mass.: MIT Press.

2010b. 'A deletion analysis of null subjects'. In T. Biberauer et al. *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge, UK: Cambridge University Press. 58–87.

Rosenbaum, P. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: MIT Press.

Ross, J. 1970. 'On declarative sentences'. In R. Jacobs and P. Rosenbaum (eds.) *Readings in English Transformational Grammar.* Waltham, MA: Ginn.

Saito, M. 2007. *Notes on East Asian Argument Ellipsis*. Unpublished manuscript. Nanzan University and University of Connecticut. <http://www.ic.nanzan-u.ac.jp/LINGUISTICS/staff/saito_mamoru/pdf/saito-2007-LangResearch.pdf> (accessed 1 February 2010).

Samek-Lodovici, V. 1996. *Constraints on Subjects: An Optimality Theoretic Analysis*. PhD, Rutgers University.

Sheehan, M. 2006. *The EPP and Null Subjects in Romance*. PhD, Newcastle University.

Sigurðsson, H. A. 2011. Conditions on argument drop. *In Linguistic Inquiry* 42.2: 267−304.

2004. 'The syntax of person, tense, and speech features'. *Rivista di Linguistica /Italian Journal of Linguistics* 16: 219–251.

Sigurðsson, H. A. and V. Egerland. 2009. 'Impersonal null subjects in Icelandic and elsewhere'. *Studia Linguistica* 63. Forthcoming.

Speas, M. 2001. 'Constraints on null pronouns'. In G. Legendre, J.B. Grimshaw and S. Vikner *Optimality-Theoretic Syntax. Language, Speech, and Communication*. Cambridge, MA: MIT Press. 393–425.

Sudmuk, C. 2003. *Semantic Merger in Serial Verb Constructions: Evidence from Sequential Serial Verb Constructions in Thai*. Unpublished manuscript.

Takahashi, D. Forthcoming. *Argument Ellipsis, Anti-agreement, and Scrambling*.

Thepkanjana, K. 1986. *Serial Verb Constructions in Thai*. PhD, University of Michigan.

Tomioka, S. 2003. 'The semantics of Japanese null pronouns and its cross-linguistic implications'. In K. Schwabe and S. Winkler (eds.) *The Interfaces: Deriving and Interpreting Omitted Structures.* Amsterdam: John Benjamins. 321– 40.

Vallduví, E. 1990. *The Informational Component*. PhD, University of Pennsylvania.

Walker, M. A., A. K. Joshi and E.F. Prince. 1998. *Centering Theory in Discourse*. Oxford: Clarendon Press.

Wang, Q. et al. 1992. 'Null subject versus null object: some evidence from the acquisition of Chinese and English'. *Language Acquisition* 2: 221–54.

Xu, L. 1986. 'Free empty category'. *Linguistic Inquiry* 17: 75–93.

2003. 'Choice between the overt and the covert'. *Transactions of the Philological Society* 101: 81–107.

Yang, S. 1994. 'Null pronouns in Chinese'. In P. Koskinen (ed.) *Proceedings of the 1994 Annual Conference of the Canadian Linguistics Association.* 645–656.