Southern Illinois University Carbondale

OpenSIUC

5-1-2023

APPEARANCES ARE DECEIVING: LONG-DISTANCE SUBJECT ANAPHORS AND PHASAL BINDING DOMAINS

Fahad A. Almalki
Southern Illinois University Carbondale, ifahadalmalki2020@gmail.com

Follow this and additional works at: https://opensiuc.lib.siu.edu/theses

Recommended Citation

APPEARANCES ARE DECEIVING: LONG-DISTANCE SUBJECT ANAPHORS AND

PHASAL BINDING DOMAINS

by

Fahad A. Almalki

B.A., Umm Al-Qura University, 2017

A Thesis
Submitted in Partial Fulfillment of the Requirements for the
Master of Arts Degree

School of Languages and Linguistics
in the Graduate School
Southern Illinois University Carbondale
May 2023

THESIS APPROVAL

APPEARANCES ARE DECEIVING: LONG-DISTANCE SUBJECT ANAPHORS AND

PHASAL BINDING DOMAINS

by

Fahad A. Almalki

A Thesis Submitted in Partial

Fulfillment of the Requirements

for the Degree of

Master of Arts

in the field of Linguistics

Approved by:

Jeffrey Punske, Chair

Karen Baertsch

Katherine I. Martin

Michael T. Putnam

Graduate School
Southern Illinois University Carbondale
April 5, 2023

AN ABSTRACT OF THE THESIS OF

Fahad A. Almalki, for the Master of Arts degree in Linguistics, presented on April 5, 2023, at Southern Illinois University Carbondale.

TITLE: APPEARANCES ARE DECEIVING: LONG-DISTANCE SUBJECT ANAPHORS AND PHASAL BINDING DOMAINS

MAJOR PROFESSOR: Dr. Jeffrey Punske

An unusual behavior of anaphors is to occur in embedded subject positions and be bound across a finite clause boundary by a matrix subject. This thesis, however, demonstrates that such constructions exist in Malki Arabic, besides other languages. First, this thesis shows that the clause size of the embedded clause in which subject anaphors are allowed is CP and not always a TP. Second, in light of current reductionist approaches to binding domains of the classical binding theory to phase theory, a cross-clausal binding relation bears issues to those approaches, as a long-distance antecedence relation crosses a phase boundary. Taking long-distance bound subject anaphors as the main empirical focus in this thesis, I show that the cross-clausal binding relation in Malki Arabic is not bona fide evidence against reducing binding domains to phases. Following Wurmbrand (2019) and Lohninger et al. (2022), I propose that constructions with long-distance bound subject anaphors theoretically resemble cross-clausal A-dependencies, like hyperraising and long-distance agreement, for undergoing movement to a position in the edge of the embedded clause and showing similar properties. Third, I show that reducing binding domains to whole phases is plausible, but taking spell-out domains as binding domains is untenable. Finally, the

proposal suggested in this thesis also sheds lights on the possibility of the anaphor agreement effect as an interface condition, in addition to highlighting an account for the accusative-marked embedded subject in Modern Standard Arabic.

DEDICATION

To my father and mother, who do not know what this thesis is about but raised me to be the person who I am today, have the language faculty and ability to use the language which has become my own!

To my lovely and inspiring wife, son, and daughter on whose shoulders I have shined and completed this work!

To my friends Khalid, Yousef, and Hassan whose friendship since college has changed and motivated me to be the linguist I aspire to be!

To everyone who believes and disbelieves in me, you have been my drive and inspiration both ways!

Thank you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

Agr	Agreement head
v	Voice/Verbal Phrase head
vP	Voice/Verbal Phrase
1	First person
2	Second person
3	Third person
A	Argument
A′	Non-argument
ABS	Absolutive
ACC	Accusative
ANAPH	Anaphor
ASP	Aspect
AUG	Augment
C	Complementizer/Complementizer phrase head
CLF	Classifier
COM	Comitative
COMP	Complementizer
COP	Copula
CP	Complementizer Phrase
CVB	Converb

DAT	Dative
DEF	Definite
DP	Determiner Phrase
ECM	Exceptional Case Marking
EPP	Extended Projection Principle
ERG	Ergative
F	Feminine
FUT	Future
GEN	Genitive
IND	Indicative
INFL	Inflectional head
INTR	Intransitive
IP	Inflectional Phrase
IPFV	Imperfective
IR	Irrealis
LF	Logical Form/Semantic Interface
M	Masculine
MSA	Modern Standard Arabic
N.PST	Non-past
NEG	Negation marker
NMLZ	Nominalizer
NOM	Nominative

NP	Noun Phrase
NPI	Negative Polarity Item
NPL	Nonplural
O	Object agreement
OBJ	Object
OBL	Oblique
OSV	Object-Subject-Verb
OVS	Object-Verb-Subject
PASS	Passive
PF	Phonetic Form/Phonology Interface
PL	Plural
PREP	Preposition
PRF	Perfect
PRN	Pronoun
PROG	Progressive aspect
PRS	Present
PST	Past
PTCP	Participle
REFL	Reflexive
RL	Realis
RP	Resumptive Pronoun
S	Argument of intransitive verb

SA   Subject agreement

SBJ   Subject

SBJV   Subjunctive

SG   Singular

SOV   Subject-Object-Verb

T   Tense head

TOP   Topic

TP   Tense Phrase

TR   Transitive

VOS   Verb-Object-Subject

VP   Verb Phrase

VSO   Verb-Subject-Object

CHAPTER 1

INTRODUCTION

This introductory chapter serves to provide essentials for the points addressed in this thesis. Section 1.1 provides an overview of the problem tackled in this thesis along with the contributions the thesis makes to the syntactic analyses of anaphors, the internal structure of embedded clauses, anaphor agreement effect, and the accusative embedded subject in Arabic. In Section 1.2, a necessary grammatical background on Malki Arabic is included. Section 1.3 provides a summary of each chapter.

1.1 OVERVIEW

Anaphors behave differently in being locally and/or long-distance bound by their antecedents. The distribution of long-distance bound anaphors varies cross-linguistically. In embedded clauses, some languages allow anaphors to occur in the object position and/or the subject position. Chinese, for instance, allows long-distance anaphors in the object (1a) or the subject (1b) position in the embedded clause.

(1)   Mandarin Chinese

   a. Zhangsan$_i$ renwei [Lisi$_j$ hai-le   ziji$_{i/j}$].
      Zhangsan think [Lisi hurt-ASP self   ]
      'Zhangsan$_i$ thought that Lisi$_j$ hurt himself$_{i/j}$'   (Huang & Tang 1991:263)

   b. Zhangsan$_i$ shuo [ziji$_i$ kanjian-le Lisi].
      Zhangsan say   [self see-PERF Lisi]
      'Zhangsan$_i$ said that he$_i$ saw Lisi.'   (Huang & Liu 2001:168)

Since the introduction of the Minimalist Program (Chomsky 1995 *et seq.*), alternative proposals to account for binding adhering to the principles of the Minimalist Program and dispensing with the vestige of the Government and Binding Theory (indices

and government) have been suggested. Among these proposals are the movement analysis (Hornstein 2001, Drummond, Kush & Hornstein 2011, *inter alia*), Agree-based analysis (Kratzer 2009, Hicks 2009, Reuland 2005, 2001, 2011, Rooryck & Vanden Wyngaerd 2011, among others) and the phase-based analysis (Quicoli 2008, Lee-Schoenfeld 2008, Despić 2015, Charnavel & Sportiche 2016). Approaches taking a phase-theoretic view on locality have aimed at reducing binding domains to whole phases (Lee-Schoenfeld 2008, Quicoli 2008) or spell-out domains (Charnavel & Sportiche 2016). Taking long-distance bound subject anaphors into account, the cross-clausal binding relation between an anaphor and its antecedent might seem to be problematic to the reduction of binding domains to phases. This thesis however shows that such constructions with long-distance bound subject anaphors are not in opposition to the phasal approach to binding domains. Similar to configurations involving cross-clausal dependencies, long-distance binding involving subject anaphors in embedded clauses involves a movement to the edge of the embedded clause, following the analyses of Wurmbrand (2019) and Lohninger et al. (2022) which aim at explaining cross-clausal A-phenomena.

This thesis, therefore, contributes with the addition of novel empirical data on the existence of subject anaphors in Malki Arabic to the set of other languages that allow anaphors in the subject position. Taking long-distance bound subject anaphors, like (2), as the main empirical focus, the thesis explores the explanatory power of the current analyses which adopt phase theory to account for the domains of Binding Theory. Within such an approach, the thesis also offers an analysis that shows that assuming phases as binding domains is on the right track and long-distance subject anaphors are derived from

movement.

(2)   Malki Arabic
        Ɛabiir$_i$ ti-ħassib   [(inn) nafs-aha$_{i/*j}$ bi-ta-rooh   al-maṭaar   ].
        Abeer 3SG.F-believe [COMP self-her   FUT-3SG.F-go the-airport]
        'Lit. Abeer believes that herself will go to the airport.'

A question that might arise is whether subject anaphors bound cross-clausally are possible to be in the same phasal binding domain with its antecedent. I argue that subject anaphors can be bound by their antecedents in the same binding domain via movement to the edge of the embedded clause. As a consequence, subject anaphors become accessible and bound by their antecedent in the same phasal binding domain. In the light of two views on when spell-out is triggered, taking whole phases rather than spell-out domains as binding domains comes from the result of the moved subject anaphor not being contained in the same spell-out domain of the antecedent. The anaphor and its antecedent can however be in the same phase. I furthermore show that the complement clause of the verb *ðana* 'believe' in Arabic is a Complementizer Phrase (CP) and not only a Tense Phrase (TP). In contrast to the usual view on the CP-domain in Arabic as an A′-domain, I provide evidence showing that the CP-domain can include A-properties which in turn indicate the presence of a landing site for arguments in the left periphery of the embedded CP clause. Following Wurmbrand (2019) and based on the argument for the A-behavior of the CP-domain in Arabic, I argue that the long-distance behavior of the bound subject anaphors is reduced locally because the configuration with subject anaphors bound cross-clausally is similar to a cross-clausal A-dependency in which movement to a position transparent to an element in another domain is required. In addition, two theoretical

implications are entertained in this thesis: showing that the anaphor agreement effect can be an interface condition and that embedded accusative-marked subjects in Modern Standard Arabic (MSA) can also be a case of movement to the same position to which the embedded subject anaphor moves for accessibility reasons.

1.2 MALKI ARABIC: GRAMMATICAL SKETCH

The data in this thesis is mainly drawn from Malki Arabic which is a spoken variety of Arabic and is spoken in the southwestern area of Saudi Arabia with approximately 20,000 speakers from the Bani Malik tribe. This particular dialect is an undocumented and unstudied dialect of Arabic in the literature[1], in contrast to Najdi Arabic, Hijazi Arabic and other Arabic dialects which are to some extent (well-)studied in the syntax literature.

Like most Arabic varieties, Malki Arabic has a rich morphological system. Verbs in Malki Arabic can be inflected for tense, aspect, subject-agreement in number, person and gender, and voice. Similar to MSA and other Arabic vernaculars, Malki Arabic is also a pro-drop language in which subjects can be omitted and their lexical content can be understood from the inflectional agreement on the verb, as can be seen from the optionality of the subject in (3).

(3)   Malki Arabic
       (huuh) ʃaaf   al-bint fii as-souq.
       he   saw.3SG.M the-girl in the-mall
       'He saw the girl in the mall.'

---

[1]As a native speaker of Malki Arabic, the judgements on the data in this thesis are my own. With some instances of data, a confirmation with other native speakers of the dialects has been sought.

Like Jordanian Arabic (Jarrah 2019), among others, Malki Arabic takes the SVO as the unmarked word order. Nevertheless, like MSA (Mohammad 2000), Malki Arabic allows all the six logical word orders, as can be seen in (4).

(4)   Malki Arabic

    a. (ʔħmad) ikil   at-tuffaħa.   SVO
        Ahmed   ate.3SG.M the-apple
        'Ahmad ate the apple'

    b. ikil   (ʔħmad) at-tuffaħa.   VSO
        ate.3SG.M Ahmed   the-apple
        Ahmad ate the apple'

    c. ikil   at-tuffaħa (ʔħmad) .   VOS
        ate.3SG.M the-apple Ahmed
        'Ahmad ate the apple'

    d. ?(ʔħmad) at-tuffaħa ikil-aha.   SOV
        Ahmed   the-apple ate.3SG.M-3SG.F
        'Ahmad ate the apple'

    e. ?at-tuffaħa ikil-aha   (ʔħmad).   OVS
        the-apple ate.3SG.M-3SG.F Ahmed
        'Ahmad ate the apple'

    f. ?at-tuffaħa (ʔħmad) ikil-aha.   OSV
        the-apple Ahmed   ate.3SG.M -3SG.F
        'Ahmad ate the apple'

Note that the SOV, OVS, and OSV word orders are acceptable in the context where the subject in (4d) is left-dislocated and the objects in (4e) and (4f) are fronted-focused DPs.

Embedded finite clauses in Malki Arabic can be headed by an optional overt complementizer *inn* 'that'. Although all word orders are possible in single clauses in Malki Arabic, the word order in embedded clauses is constrained. As can be seen in (5), only the SVO order can be possible without resumption nor the attachment of a clitic to

the complementizer. Embedded clauses with sentences of the VSO word order, as in (5b),

are possible upon the presence of a clitic attached to the complementizer. Embedded

clauses with sentences of VOS and SOV are not allowed, as shown by the

ungrammaticality of (5c) and (5d), respectively. The word orders that utilize resumption

are the OVS and OSV. As sentences (5e) and (5f) show, the OVS and OSV sentence

structures in embedded clauses are cases of left-dislocation and can only be allowed if the

object has a resumptive pronoun attached to the embedded verb.

(5)   Malki Arabic

a. fahad ya-guul   (inn) ?ħmad ikil   at-tuffaħa.   SVO ✓
    Fahad 3SG.M-say COMP Ahmed ate.3SG.M the-apple
    'Fahad says that Ahmad ate the apple'

b. fahad ya-guul   inn-*(uh)   ikil   ?ħmad at-tuffaħa.   VSO✗ / C-PRN VSO✓
    Fahad 3SG.M-say COMP-3SG.M ate.3SG.M Ahmed the-apple
    'Fahad says that Ahmad ate the apple'

c. *fahad ya-guul   (inn) ikil   at-tuffaħa ?ħmad .   VOS ✗
    Fahad 3SG.M-say COMP ate.3SG.M the-apple Ahmed
    'Fahad says that Ahmad ate the apple'

d. *fahad ya-guul   (inn) ?ħmad at-tuffaħa ikil-(aha).   SOV ✗
    Fahad 3SG.M-say COMP Ahmed the-apple ate.3SG.M-3SG.F
    'Fahad says that Ahmad ate the apple'

e. fahad ya-guul   (inn) at-tuffaħa ikil-*(aha)   ?ħmad.   OV-RP S ✓ / OV-ØS ✗
    Fahad 3SG.M-say COMP the-apple ate.3SG.M-3SG.F Ahmed
    'Fahad says that Ahmad ate the apple'

f. fahad ya-guul   (inn) at-tuffaħa ?ħmad ikil-*(aha).   OSV-RP✓ / OSV-Ø✗
    Fahad 3SG.M-say COMP the-apple Ahmed ate.3SG.M   -3SG.F
    'Fahad says that Ahmad ate the apple'

MSA exhibits different patterns of agreement on the verb in SVO and VSO word

orders. With SVO order (6a), the verb must always show full agreement in number,

person, and gender. The verb in VSO order, on the other hand, can only show a partial

agreement with the postverbal subject in person and gender, as shown in (6b).

(6)   MSA

    a. ʔal-ʔawlaad-u qaraʔa-uu   ʔal-dars-a.
        the-boys-NOM read-3PL.M the-lesson-ACC
        'The boys read the lesson'

    b. qaraʔa   ʔal-ʔawlaad-u ʔal-dars-a.
        read.3SG.M the-boys-NOM the-lesson-ACC
        'The boys read the lesson'   (Soltan 2006)

As sentences (7) show, in Malki Arabic the verb must always exhibit full agreement with

the subject, be it preverbal (7a) or postverbal (7b).

(7)   Malki Arabic

    a. al-ʕiyyal ikil-uu/*Ø   at-tuffaħa.
        the-boys ate-3PL.M/*3SG.M the-apple
        'The boys ate the apple'

    b. ikil-uu/*Ø   al-ʕiyyal at-tuffaħa.
        ate-3PL.M/*3SG.M the-boys the-apple
        'The boys ate the apple'

The uniformity of full agreement with both SVO and VSO sentence structures has also

been observed in Jordanian Arabic (Jarrah 2017), Moroccan Arabic (8), and Lebanese

Arabic (9), as pointed out in Aoun et al. (1994).

(8)   Moroccan Arabic (Aoun et al. 1994)

    a. *nʕas/nʕas-u   lə-wlaad.
       slept.3SG.M/slept-3PL.M the-children
       'The children slept.'

    b. lə-wlaad   *nʕas/nʕas-u.
       the-children slept.3SG.M/slept-3PL.M
       'The children slept.'

(9)   Lebanese Arabic (Aoun et al. 1994)

   a. *neem/neem-o   lə-wlaad.
     slept.3SG.M/slept-3PL.M the-children
     'The children slept.'

   b. lə-wlaad   *neem/neem-o.
     the-children slept.3SG.M/slept-3PL.M
     'The children slept.'

   Another property that sets Malki Arabic apart from MSA is the absence of overt grammatical case marking. Malki Arabic does not realize morphological case on nouns, adjectives, or some adverbs. The loss of overt case realization can also be noticed in most spoken varieties of Arabic. Case is realized, however, on pronouns in being either morphologically free or cliticized. When pronouns are marked with nominative case, they stand alone as a free morpheme (10a). Pronouns marked with accusative or genitive case become cliticized to a complementizer for the former (10b) or to preposition for the latter (10c).

(10)   Malki Arabic

   a. ħassaan   gaal   inn   **hiih**   saafar-at   al-ʒanauub.
     Hassan.NOM said.3SG.M COMP **she.NOM** traveled-3SG.F the-south
     'Hassan said that she travelled to the south.'

   b. ħassaan   gaal   inn-**ha**   saafar-at   al-ʒanauub.
     Hassan.NOM said.3SG.M COMP-**3SG.F.ACC** traveled-3SG.F the-south
     'Hassan said that she travelled to the south.'

   c. χuluud   ta-laʕab   muʕ-**uh**/*muʕ **huuh**.
     Kholoud 3SG.F-play with-**3SG.M.GEN**/*with **him.NOM**
     'Kholoud plays with him.'

Another instance of morphological realization of case in Malki Arabic can be found with negative polarity items (NPI) like *ʔabad* 'never'. As illustrated in (11), the NPI is marked with accusative case by the suffix *-aa*.

(11) Malki Arabic
      al-walad maa ya-laʕab   baraa   ʔabad-aa.
      the-boy NEG 3SG.M-play outside never-ACC
      'The boy never plays outside.'

Malki Arabic therefore shares multiple similarities with the other spoken varieties of

Arabic in showing insensitivity of verbal agreement to whether the subject follows or

precedes the verb in the sentence structure. Furthermore, the loss of overt case marking

seems to be another property of Malki Arabic that puts this dialect among the other

varieties of Arabic and separates it from MSA which shows morphological case and

different agreement patterns that are sensitive to the position of the subject with respect

to the verb position in the sentence structure. For the structure of the embedded finite

clause, Malki Arabic allows sentences with SVO word order but not with VOS and SOV.

Embedded clauses with VSO, OVS, and OSV in Malki Arabic are constrained by the

presence of a clitic on the complementizer for the former and resumptive pronouns in the

latter word orders.

1.2.1   Reflexives

Reflexivity in Malki Arabic can be marked on the following nouns: *nfs* 'soul', *rooh*

'soul', and *ʕumr* 'age'. This group of nouns is marked for reflexivity with a pronominal

suffix which agrees with the noun it refers to in gender, number, and person, as can be

seen in Table 1.1. Thus, the nominal structure of reflexives in Malki Arabic, similar to

English and MSA, is complex, meaning that the reflexive is comprised of a particular set

of nouns, plus an attached pronoun (as stated in Reinhart & Reuland 1993). Malki

Arabic does not have simplex reflexives such as *ziji* in Mandarin Chinese, *zich* in Dutch,

and *sig* in Icelandic which hold a default agreement, regardless of the number, gender,

and person properties of the referring expression.

The paradigm in Table 1.1 also indicates special properties of reflexives in Malki Arabic related to case and gender. Reflexives in Malki Arabic are not marked for case, as shown by having the same form of the reflexive with the nominative, accusative, and genitive cases. That reflexives do not carry a morphological case is not surprising because nominals in Malki Arabic do not show overt case marking. For gender, the reflexive *nfs* in the first person is syncretic for gender. Specifically, the singular form *nafs-i* 'myself' and plural form *nafs-anaa* 'ourselves' of the reflexive can be used in the same morphological form for the feminine and masculine genders in the first person.

Person	Case	Singular		Plural	
		MASC.	FEM.	MASC.	FEM.
1	NOM	*nafs-i*		*nafs-anaa*	
	ACC	*nafs-i*		*nafs-anaa*	
	GEN	*nafs-i*		*nafs-anaa*	
2	NOM	*nafs-ak*	*nafs-ik*	*nafs-akum*	*nafs-skunnah*
	ACC	*nafs-ak*	*nafs-ik*	*nafs-akum*	*nafs-akunnah*
	GEN	*nafs-ak*	*nafs-ik*	*nafs-akum*	*nafs-akunnah*
3	NOM	*nafs-uh*	*nafs-aha*	*nafs-ahum*	*nafs-ahunnah*
	ACC	*nafs-uh*	*nafs-aha*	*nafs-ahum*	*nafs-ahunnah*
	GEN	*nafs-uh*	*nafs-aha*	*nafs-ahum*	*nafs-ahunnah*

Table 1.1. Grammatical Paradigm for Reflexive *nfs* in Malki Arabic

Similar to reflexives in English and other languages, the reflexive in Malki Arabic can take any antecedents whose ɸ-content (person, number, and gender) agrees with the one of the reflexives. As can be seen in (12), the reflexive has total agreement with its antecedent; otherwise, the antecedence relation does not hold.

(12)   Malki Arabic

    a. **juusif**   ya-ħub   nafs-**uh**.
        yousef.NOM 3SG.M-love self-**3SG.M**
        'Yousef loves himself.'

    b. **ʕabiir**   ti-ħub   nafs-**aha**.
        Abeer.NOM 3SG.F-love slef-**3SG.F**
        'Abeer love herself.'

    c. **ar-riʒaal**   ya-ħub-uun nafs-**ahum**.
        the-men.NOM 3M-love-PL self-**3PL.M**
        'The men love themselves.'

    d. **al-bas**   ya-ħub   nafs-**uh**.
        the-cat.NOM 3SG.M-love self-**3SG.M**
        'The cat loves its(him/her)self.'

    e. **ʔana** ʔa-ħub   nafs-**i**.
        I.NOM 1SG-love self-**1SG**
        'I love myself.'

    f. **nuħun** na-ħub   nafs-**anaa**.
        we.NOM 1PL-love self-**1PL**
        'We love ourselves.'

    g. **ʔaint**   ta-ħub   nafs-**ak**.
        you.M.NOM 2SG.M-love slef-**2SG.M**
        'You(M) love yourself(M).'

    h. **ʔaint-i**   ta–ħub-iin nafs-**ik**.
        you.NOM-**F** 2SG-love-F self-**2SG.F**
        'You(F) love yourself(F).'

With regards to the syntactic distribution of the reflexive *nfs* in Malki Arabic, the reflexive can occur in object and subject positions. In single clauses, the reflexive can only be in the object position. Compared to the sentence in (12a) above, the sentence in (13) is ill-formed because reflexives cannot be the subject of a monoclause, but rather must be the object in the clause. This property of reflexives in single clauses follows from the restriction that reflexives need be preceded by their antecedents.

11

(13)   Malki Arabic
    *nafs-uh   ya-ħub   juusif.
    self-3SG.M 3SG.M-love yousef.NOM
    Lit:'Himself loves Yousef.'

In verbless sentences (without main verbs), Malki Arabic allows reflexives to follow

adjectives, as can be seen in (14) in which the reflexive is in a position after the adjective

ð̣aalim 'unjust'.

(14)   Malki Arabic
    al-walad kaan ð̣aalim   nafs-uh.
    the-boy was unjsut.3SG.M self-3SG.M
    'The boy was unjust to himself.'

Besides their availability in the object position of monoclausal sentences, Anaphors in

Malki Arabic can also appear in the object position of embedded clauses. In bi-sentential

clauses, the reflexive can be found in the object position. The reflexive moreover can only

be co-referring with the closest agreeing-nominal in the clause, as illustrated in (15):

(15)   Malki Arabic
    ʕabiir$_i$ ti-ħassib   [(inn) mayyayam$_j$ ti-ħub   nafs-aha$_{*i/j}$].
    Abeer 3SG.F-believe [COMP Mary   3SG.F-love self-her   ]
    'Lit. Abeer believes that Mary loves herself.'

Different from MSA and similar to languages like Tamil (Sundaresan 2012), reflexives in

Malki Arabic can also be seen in the subject position of the embedded clause with verbs

of knowledge and belief, like *ħassib* 'believe', *ʃuuf* 'see', and *ħuus* 'feel', as demonstrated

in (16), but not with other verbs like 'say' in (16d).

(16) a. ʕabiir$_i$ **ti-ħassib**   [(inn) nafs-aha$_{i/*j}$ bi-ta-rooh   al-maṭaar].
    Abeer 3SG.F-believe [COMP self-her   FUT-3SG.F-go the-airport]
    'Lit. Abeer believes that herself will go to the airport.'

b. ʕabiir$_i$ ta-**ʃauuf** [(inn) nafs-aha$_{i/*j}$ bi-ta-drus   al-ħaasib   ].
   Abeer 3SG.F-see [COMP self-her   FUT-3SG.F-study the-computer science]
   'Lit. Abeer sees that herself will study computer science.'

c. ʕabiir$_i$ ta-**ħauus** [(inn) nafs-aha$_{i/*j}$ bi-ta-fuuz   bi-l-ʒaaʔizah   ].
   Abeer 3SG.F-feel [COMP self-her   FUT-3SG.F-win PREP-the-prize]
   'Lit. Abeer feels that herself will win the prize.'

d. *ʕabiir$_i$ ta-**quul**   [(inn) nafs-aha$_{i/*j}$ bi-ta-drus   al-ħaasib   ].
   Abeer 3SG.F-say [COMP self-her   FUT-3SG.F-study the-computer science]
   'Lit. Abeer says that herself will study computer science.'

Similar to its behavior in the embedded object position, the reflexive in the embedded subject position can only be anteceded by the closest agreeing-noun. When the reflexive is embedded under multiple clauses with subjects with the same ϕ-content, the reflexive can only co-refer with the subject of the verb 'believe', as can be seen in (17):

(17)   fahad$_m$ ya-guul   [inn   muħammad$_k$ ya-ʕalim   li-juusif$_j$ [inn   ʕali$_i$
        Fahad   3SG.M-say [COMP Mohammad   3SG.M-tell to-Yousef [COMP Ali
        ya-ħassib   [(inn) nafs-uh$_{i/*j/*k/*m}$ bi-ya-rooh   al-maṭaar]]].
        3SG.M-believe [COMP self-3SG.M   FUT-3SG.M-go the-airport]]]
        'Fahad says that Mohammad tells Yousef that Ali believes that himself will go to the airport.'

The description of reflexives in Malki Arabic shows an interesting phenomenon of anaphors occurring in a subject position and anteceded across a finite clause. The long-distance antecedence relation in Malki Arabic raises questions and challenges to current syntactic treatments of locality, following minimalist constructs like phases. The constructions with long-distance bound subject anaphors are the main empirical focus of this thesis because they appear to bear a challenge to hypotheses of reducing binding domains (especially of Binding Condition A) to phase theory.

1.3   THESIS ORGANIZATION

The thesis is organized into five chapters. Chapter 2 provides an extensive background on anaphors[2], and their long-distance behavior cross-linguistically. This chapter also includes the previous theoretical work on anaphors and introduces the puzzle of this thesis. Furthermore, I discuss the implication of long-distance subject anaphors to analyses proposing phases as binding domains and whether the definitions of exemption postulated in these approaches apply to the kind of subject anaphors in Malki Arabic. Another point discussed in chapter 2 is that long-distance subject anaphors do not incur violation to the anaphor agreement effect. Chapter 2 ends with a preview of the proposal I suggest to explain how the long-distance antecedence relation between the subject anaphor in the embedded clause and the antecedent in the matrix clause holds, assuming phasal binding domains.

The focus of chapter 3 is the clause size of the complement clause of the verbs that allow long-distance anaphors. I show that the complement of the verbs of knowledge and believe, especially *ħassib*, is not always a TP but can rather be a CP with an overt complementizer. I also argue that the clause where we find the subject anaphor is not an Exceptional Case Marking (ECM) construction based on evidence from Binding Condition B, among others.

In chapter 4, I show that long-distance binding in Malki Arabic is not a counterexample to the phasal approach to binding domains even though it superficially seems to be. I thus suggest a cross-clausal analysis for the long-distance anaphoric

---

[2]I use the terms reflexives and anaphors interchangeably. I do not include reciprocals in the term, however.

relation, in line with previous analyses of long-distance phenomena, like agreement and hyperraising. I also provide explanations of why subject anaphors can occur with a subset of verbs and not with others. I show that the suggested proposal does not only account for the long-distance binding relation, but it also provides insights on the prospect of the anaphor agreement effect as an interface (post-spell out) condition and it might be able to explain the accusative-marked embedded subject in MSA. Finally, in chapter 5, a conclusion of the arguments suggested in this thesis is provided.

CHAPTER 2

THE PUZZLE

In this chapter, I introduce anaphors and their types based on their distributional properties. I particularly address their properties through the Classical Binding Theory (Chomsky 1981, 1986), the co-argumenthood theory of binding (Reinhart & Reuland 1993, Reuland 2001, Pollard & Sag 1992, among others), and the phase-theoretic approaches to the binding domains (Lee-Schoenfeld 2008, Charnavel & Sportiche 2016, among others). I also touch upon the implication of the long-distance subject anaphors on current conditions and definitions of exemption of anaphors from Condition A of the Binding Theory, especially based on the phasal analyses of binding domains. Furthermore, I show subject anaphors do not violate the anaphor agreement effect (Rizzi 1990) and that Malki Arabic does not utilize any of the strategies that other languages employ to obey such effect. Finally, a preview of the solution for the puzzle, along with theoretical goals of the thesis, is illustrated at the end of this chapter.

2.1   ANAPHORS

Per the Classical Binding Theory (Chomsky 1980, 1981, 1986), anaphors, such as reflexives and reciprocals, are linguistic elements that typically seek a referent denoted by an antecedent which is already part of the prior information given in the sentence. The antecedence (dependency) relation between an anaphor and its antecedent is not established in a vacuum. Rather, this relation is regulated by distributional properties of anaphors concluded in the Binding Condition A of the Classical Binding Theory: this

16

condition states that an anaphor must be bound in its governing category (Chomsky 1981)/local domain (Chomsky 1986).[3] In order for the anaphor to be bound, it must be c-commanded by and co-indexed with its antecedent. An antecedent c-commands an anaphor if the antecedent does not dominate the anaphor and every category that dominates the antecedent dominates the anaphor (Chomsky & Lasnik 1993:518). Here 'dominate' refers to the relation in which both the antecedent and the anaphor are hierarchically under the non-branching node (the maximal projection) that includes the antecedent and the anaphor. The c-command requirement holds when the antecedent cannot be the sister of the anaphor since c-command requires the former be higher hierarchically and asymmetrically c-command its sister which includes the latter, the anaphor. The sentences in (18) illustrate the c-command requirement for the Binding Condition A. In (18a), the anaphor *herself* is structurally higher than (thus c-commands) its antecedent *Mary*, hence this sentence is ungrammatical because it violates Condition A of Binding Theory. This is however not the case in (18b) where the antecedent *Mary* c-commands the anaphor *herself*, abiding by Condition A of the Classical Binding Theory.

(18) a. *Herself$_i$ likes Mary$_i$.

    b. Mary$_i$ likes herself$_i$.

For the requirement of the anaphor to occur in the local domain of its antecedent, the governing category is the minimal projection that includes the anaphor, the governor of

---

[3]The Classical Binding Theory also includes pronouns, empty categories, PRO/*pro*, and referring expressions in the following conditions:

(i) (B) A pronominal must be free in its governing category.

    (C) A Referring expression must be free.

  Because anaphors are the main focus of the thesis, the main discussion will only involve the Binding Condition A.

that anaphor, and accessible SUBJECT (Chomsky 1981:211). The governing projections might be a TP, small clause, infinitival clause, or Determiner Phrase (DP). Possible accessible SUBJECTs could be the subject in the specifier of TP, a possessor in the specifier of DP/NP, or Agr. For instance, in constructions like (19) the binding domain for the reflexive *himself* is TP_2 or CP_2, not the higher TP_1 or CP_1.

(19) $[_{CP_1} [_{TP_1}$ Yousef$_i$ said $[_{CP_2}$ that $[_{TP_2}$ Fahad$_j$ likes himself$_{i*/j}]]]$.

Thus, only *Fahad* is the legitimate binder and the accessible SUBJECT in the binding domain (here TP_2). The anaphor cannot be bound by *Yousef* because it does not occur in the same domain of the reflexive. Moreover, the subject in the specifier of TP_2, *Fahad*, is the only accessible SUBJECT in that binding domain.

2.1.1   Long-distance Anaphors

However, local anaphors are not always the norm across languages, even in English. Languages like Dutch, Icelandic and Chinese have long-distance anaphors which do not obey the locality requirement of the Binding Condition A and can be anteceded by elements occurring outside their local domain, thus so-called long-distance or exempt (as called in Pollard & Sag 1992) for obviating Binding Condition A,[4] as can be seen in Icelandic (20), Mandarin Chinese (21), and English (22), translations are mine for the sentence from Icelandic.

---

[4] Not all long-distance anaphors have an antecedent, however. Languages like English and other languages include logophors, which are anaphors that are boundless: they do not require an antecedent in the sentence or an antecedent understood in the discourse (i).

(i)   People$_i$ like yourself$_j$ are crazy about analyzing data from different languages.

The logophoric behavior of anaphors was noted in Ross (1970), Cantrall (1974), and Kuno (1972, 1975). Similarly, as noted in Postal (1970), reflexives with the form *picture+NP* can occur in a position not c-commanded by any antecedent (ii).

(ii)   A picture of myself would look nice on that wall.   (Reuland 2011: 44)

(20)   Icelandic (Thráinsson 1976)

Joón$_i$ segir [að   Billi [að   Mariía$_j$ heimsæki telji   [að   Haraldur$_k$
John says [that Bill [that Mary   visit.SBJV believe.SBJV [that Harold
vilji   sig$_{i/j/k}$]]].
want.SBJV REFL
'John says that Bill whom Mary visits believes that Harold wants himself/herself.'

(21)   Mandarin Chinese (Huang & Tang 1991:263,(1a))

Zhangsan$_i$ renwei [Lisi$_j$ hai-le   ziji$_{i/j}$].
Zhangsan think Lisi   hurt-ASP self
'Zhangsan$_i$ thought that Lisi$_j$ hurt him/herself$_{i/j}$'

(22)   English (Ross 1970:227, (17b))

Tom$_i$ thinks that I tried to get Mary to make you say that the paper had been written by Ann and himself$_i$.

In Icelandic, as shown in (20), the anaphor *sig* can be bound by three possible antecedents: 'John' which is the highest subject, 'Mary' as the subject of the relative clause, and 'Harold' the local subject in the domain of the anaphor. Similarly, the anaphor *ziji* in (21) can take either the noun 'Lisi' in its local domain (namely the embedded clause) or the matrix subject 'Zhangsan' as its antecedent. As can been seen in (22), although the domain of anaphor *himself* in English is the lowest embedded clause, it is bound by the antecedent *Tom*, the subject of the highest clause. The difference between the long-distance bound anaphors in English, on one hand, and Icelandic and Mandarin Chinese, on the other, is that the anaphors in the latter languages are monomorphemic and unmarked for gender, hence allowing multiple potential antecedents, despite the difference between 'John' and 'Mary' in gender. In English, on the other hand, a pronoun is attached to the reflexive and it indicates which nominal should be targeted as a potential antecedent, matching for person, number, and gender. The terminological use of monomorphemic anaphors and reflexives like those in English will be

further defined when the predicate-based approaches to binding are addressed, especially

Reinhart & Reuland (1993), among others.

2.1.2 Long-distance Subject Anaphors

As we have seen earlier, anaphors like *sig* in Icelandic, *zibun* in Japanese, and *ziji* in

Mandarin can be found in the object position in the embedded clause and be bound

cross-clausally by the matrix subject. An empirical question that can arise is whether

there are languages that allow anaphors in the embedded subject position and to have

matrix subjects as their antecedent. The answer to this question is positive. Subject

anaphors are attested in Mandarin, Tamil, and Icelandic in (23), (24), and (25),

respectively.

(23)   Mandarin Chinese (Huang & Liu 2001:168)
        Zhangsan$_i$ shuo [ziji$_i$ kanjian-le Lisi].
        Zhangsan say   [self see-PRF   Lisi]
        'Zhangsan$_i$ said that he$_i$ saw Lisi.'

(24)   Tamil (Sundaresan 2012)
        Raman$_i$ [taan$_{i,j}$ paris-æ   dʒejkka-poo-r-aan-nnŭ   ]
        Raman.NOM [ANAPH.NOM prize-ACC win-go-PRS-3SG.M-COMP]
        kaɳɖupiɖi-tt-aan.
        find.out-PST-3SG.M
        'Raman found out that he was going to win the prize.'

(25)   Icelandic (Maling 1984)
        Hún sagði [að   sér   þætti   vænt um mig].
        she$_i$ said [that self$_i$.DAT was.SBJV fond of me]
        'She said that she was fond of me.'

In all the sentences above, the embedded anaphors are bound across a finite clause

boundary, which is in fact their local binding domain. The anaphors *ziji, taan,* and *sér*

are all in antecedence relations with subjects in the higher clause, namely *Zhangsan,*

Raman, and the pronoun 'she', respectively.

Unlike Tamil and the other languages above which allow the simplex anaphors (in the sense of Reinhart & Reuland 1993) in the languages above, as we have seen in chapter 1, Malki Arabic allows complex anaphors in the embedded subject position and to have an antecedence relation with an element in the matrix clause. As illustrated in (26a), the complex anaphor *nafs-aha* in the embedded subject position can only be bound across a finite clause by the matrix subject 'Abeer'. It is not, however, possible for the same anaphor to have the matrix subject as its antecedent when the anaphor is in the object position. As shown by the ungrammaticality of (26b), when the anaphor 'herself' is in the object position of the embedded clause, it cannot be bound by the subject 'Abeer' in the matrix clause.

(26)   Malki Arabic
   a. ʕabiir$_i$ ti-ħassib   [(inn) nafs-aha$_{i/*j}$ bi-ta-rooh   al-maṭaar].
       Abeer 3SG.F-believe [COMP self-her   FUT-3SG.F-go the-airport]
       'Lit. Abeer believes that herself will go to the airport.'

   b. ʕabiir$_i$ ti-ħassib   [(inn) mayyayam$_j$ ti-ħaub   nafs-aha$_{*i/j}$].
       Abeer 3SG.F-believe [COMP Mary   3SG.F-love self-her   ]
       'Lit. Abeer believes that Mary loves herself.'

From these items of data, two remarks are worth mentioning. First, that the anaphor must be anteceded indicates that the logophoric interpretation is barred in this language (See footnote 4). Another important remark to point out is that the long-distance bound anaphor in (26a) contradicts the observation of Pica (1987) that long-distance anaphors tend to be monomorphemic, as we have seen above with *taan* in Tamil,[5] *sig* in Icelandic,

---

[5]The anaphor *taan* in Tamil can be marked for number. See Sundaresan (2012) for more discussion about the monomorphemicity of anaphors.

zibun in Japanese, and *ziji* in Mandarin. Although the anaphor in (26a) shows full agreement with its antecedent (complex), it is bound cross-clausally by the matrix subject (long-distance binding).

2.2 LONG-DISTANCE ANAPHORS: THEORETICAL BACKGROUND

A brief typological survey shows that there exist languages with anaphors that can be bound outside their local domains, contra Condition A of the Classical Binding Theory. In the literature, several analyses within the Government and Binding Theory (GB), predicate-based theory of binding, and analyses built on minimalist concepts, like phases, have been suggested to account for the properties of long-distance (bound) anaphors, e.g., Charnavel & Sportiche (2016) and Charnavel (2019). I am going to address these analyses below.

2.2.1 The Classical Binding Theory

In the Classical Binding Theory, especially the Binding Condition A, long-distance binding of anaphors is reduced to local anaphors via covert movement of anaphors to the inflectional domain (INFL) (Lebeaux 1983, 1984, Chomsky 1986). Another approach in this vein is treating anaphors as adjuncts to their antecedents via head movement from INFL to another INFL in the matrix clause (Pica 1987, Cole et al. 1990, Cole & Sung 1994, Cole & Wang 1996, among others). A further analysis by Huang & Tang (1991) suggests an XP-movement besides the IP-adjunction of the anaphor beneath the antecedent.[6]

Because long-distance anaphors are typically anteceded by subjects, as noted in

---

[6]Because some long-distance anaphors adjoining to the embedded IP can be c-commanded by an object in the matrix clause, Cole et al. (2006) propose that the VP should be the landing site for the anaphor instead of the IP. See Charnavel (2019) for a detailed discussion.

Icelandic and Italian (Koster & Reuland 1991, Giorgi 2006), it was proposed that long-distance (monomorphemic) anaphors move from INFL to another INFL covertly at LF. In examples like (27), the long-distance anaphor *each other* moves at LF from the embedded nominal in the subject position to adjoin INFL via head movement. In this position, the anaphor becomes local to its antecedent, as illustrated in (27b) where the anaphor adjoins INFL that has the licit antecedent. The interpretation thereby holds between the antecedent and its trace in the canonical position in the embedded clause (a government relation per GB terms, see Chomsky (1986:175) for more details).

(27) a. The children thought that [$_S$ [$_{IP}$ pictures of each other] were on sale].

b. The children *each other*$_i$-INFL thought that [$_S$ [$_{IP}$ pictures of e_i] were on sale].
$$\text{(Chomsky 1986)}$$

In the XP-movement approach to long-distance binding, Huang & Tang (1991) suggest the covert phrasal movement to an adjunct position in IP. This approach is motivated by sentences like (28a) where multiple antecedence interpretations are possible with the anaphor *ziji* in Mandarin which can have *Wangwu* in (28b), *Lisi* in (28c), or *Zhangsan* in (28d) as its binder. As can be seen in (28), the landing position of the anaphor with each possible anaphoric interpretation is the specifier of IP of the clause which contains the binder.

(28)   Mandarin Chinese (Huang & Tang 1991:273)

a. Zhangsan$_k$ manyuan Lisi$_j$ chang shuo Wangwu$_i$ bu xihuan ziji$_{i/j/k}$.
Zhangsan   complain Lisi often say   Wangwu not like   self
'Zhangsan complained that Lisi often said that Wangwu does not like self'

b. Zhangsan$_k$ manyuan [Lisi$_j$ chang shuo [Wangwu$_i$ bu xihuan **ziji**$_i$]]

c. Zhangsan$_k$ manyuan [Lisi$_j$ chang shuo [**ziji**$_j$ [Wangwu$_i$ bu xihuan t_j]]

d. Zhangsan$_k$ manyuan [**ziji**$_k$ [Lisi$_j$ chang shuo [t_k [Wangwui bu xihuan t_k]]]]

Although the approaches of reducing long-distance anaphors to local anaphors had success in accounting for why long-distance anaphors are usually monomorphemic and subject-oriented, besides being parsimonious in nature in reducing two categories of anaphors into one category, these LF-movement approaches encounter problems with anaphors which can be long-distance bound outside islands (see Huang & Tang 1991 and Cole et al. 2006). Moreover, a problem to the head movement explanation is the violation of minimality, a condition which states that movement should not move across potential landing sites. The movement of the anaphor from one INFL to another is just a perfect obviation case for the minimality condition. Namely, it is not clear why head movement should proceed from INFL to another despite the availability of positions like the heads v and C (in some languages that allow I/T-to-C movement).[7]

2.2.2 Co-argumenthood Approaches to Binding Theory

In the co-argumenthood (also called predicate-based) approaches to Binding Theory (Reinhart & Reuland 1993, Reuland 2001, 2011, *et seq.*), anaphors are partitioned morphologically into two types: simplex or complex. Simplex anaphors have a default form and do not show agreement in number, gender, and person with their antecedents. Instances of simplex anaphors are *zich* in Dutch, *sig* in Icelandic, *ziji* in Mandarin Chinese, and *zibun* in Japanese. Complex anaphors, in contrast, show agreement in gender, number, and person with their antecedents. Some of complex anaphors are the *pronoun-self* in English and *nfs-pronoun* in Arabic.

---

[7]See Charnavel (2019) and Sundaresan (2012) for more discussion about why movement approaches to long-distance binding are untenable.

In the co-argumenthood approach to anaphors, the Binding Condition A is redefined based on whether an anaphor is a co-argument of a predicate, rather than the distribution of the anaphors and their antecedents, as the case with the Classical Binding Theory. Anaphors must have co-arguments of the same predicates in order for them to be bound. Otherwise, Condition A does not apply to anaphors which occur as an argument of a different predicate other than the predicate of the external argument: these anaphors are thus called 'exempt anaphors' as they are not included under Condition A. Exempt anaphors can be found in English (29), where *himself* is not an argument of *put* whose external argument *John* is the antecedent of the complex anaphor.

(29) $John_i$ put the book behind $himself_{i/*j}$.

Per the coargumenthood-based analyses of binding, exemption applies to anaphors which occur as the only argument of a syntactic predicate and cannot be anteceded by a co-argument, as with the case in English with the form *picture of SELF* where reflexives are considered the sole argument of the predicate *picture* (Pollard & Sag 1992, Reinhart & Reuland 1993, Reuland 2011, a.o.). A further scenario in which anaphors are exempt is when they occur in a conjoining phrase embedded in the predicate argument, as shown in (30).

(30) a. Max_i boasted that the queen invited Lucie and **$himself_i$** for a drink.

   b. It angered him_i that she invited a man like **$himself_i$**.

(Reinhart & Reuland 1993:670)

For the predicate-based approaches to binding, exempt anaphors are not co-arguments with possible antecedents while true anaphors are (Pollard & Sag 1992, Reinhart &

Reuland 1993, Reuland 2011). Exempt anaphors are only defined based on their distributional properties with respect to a predicate and argued to be influenced by discourse-related constraints, such as perspective (Pollard & Sag 1992) and/or focus and logophoricity (Reinhart & Reuland 1993).

In sum, the predicate-based theory of binding does not attempt to reduce long-distance anaphors into local anaphors. Exempt anaphors are rather of special type of anaphors. This theory thus does not explain how long-distance anaphors can be bound by an antecedent.[8] As far as long-distance subject anaphors in Malki Arabic and other languages are concerned, the Binding Condition A of the co-argumenthood approaches does not apply to these anaphors because they are not co-arguments with their antecedents and of the same predicate. Subject anaphors are thereby exempt from Condition A.

2.2.3   Minimalist Analyses of Anaphors

Since the introduction of the Minimalist Program (Chomsky 1995 *et seq.*), alternative proposals to account for binding adhering to the principles of the Minimalist Program have been suggested. Among these proposals are the movement analysis (Hornstein 2001, Drummond et al. 2011, among others) and Agree-based analysis (Kratzer 2009, Hicks 2009, Reuland 2005, 2001, 2011, Rooryck & Vanden Wyngaerd 2011, a.o.). Another group of studies has suggested that binding domains can be reduced to the phase theory of locality (Quicoli 2008, Lee-Schoenfeld 2008, Despić 2015, Charnavel & Sportiche 2016).

The focus of the rest of this section is the phasal analyses of anaphors, as the problem this thesis investigates is related to the binding domains. Moreover, the other

---

[8]See Charnavel (2019:21-22) for problems with the predicate-based analyses of binding.

minimalist approaches utilizing movement and Agree assumed a similar version of phases to account for binding domains. Before addressing the phasal approaches to binding, a background overview of phase theory is provided.

2.2.3.1 Theoretical Background: Phases

Chomsky (2000, 2001) proposes that the derivation proceeds in stages, so-called phases. A phase is a full propositional unit and "semantically and phonologically coherent and independent" (Chomsky 2004:124).[9] Structurally, a phase is composed of two elements: the edge (its head and specifiers) and spell-out domain (its complement). The definition of phase above implies that not any group of merged lexical items can constitute a phase and that only CP and vP are phases.[10] What is special about these phrases is that they carry the information structure of the discourse (event and tense structure) for the CP at its edge and a complete argument structure (a predicate and its external and internal arguments) for the vP (Chomsky 2001, 2008).

Another important aspect of phase theory is the point at which a phasal complement becomes opaque to syntactic operations like movement and agreement. Based on

---

[9]The opposite view is that the status of phasehood can be dynamic as a result of head movement (den Dikken 2007, Gallego 2010). In other words, a phrasal category can be a phase at a point in the derivation, and then that same phrase, when it undergoes head movement, can lose its phasehood status. Such approach defines phases not on the basis of centrality of uninterpretable features (Chomsky 2008) or being propositional or the argument structure (Chomsky 2000, 2001), but rather based on the configuration of *subject* and *predicate*.

(i) a. $[_{XP}$ X $[_{HP}$ ZP $[_{H'}$ H YP]]]
    b. $[_{XP}$ WP $[_{X'}$ **X** $[_{HP}$ $[_{H'}$ H YP]]]]

Briefly, in (i) the head H is a phase in (ia) and YP is the spell-out domain, but when H moves to X, H is not the phase anymore. The phasehood status transitions to X (ib) and the spell-out domain becomes HP and its elements.

[10]Nevertheless, previous studies have shown that other categories are possible to be phasal heads. Some of these categories are DP (Chomsky 2004, 2005, 2007, Matushansky 2005, Svenonius 2004, Marantz 2007, Ott 2008, Hiraiwa 2005), PPs with functional prepositional heads (Abels 2003, Bošković 2013, Drummond et al. 2010, Kayne 1999, 2004, Citko 2014, *inter alia*), and Predication Phrases (Citko 2014).

Chomsky (2000, 2001), two views on when the spell-out/transfer of the phasal complement is triggered have been suggested in two versions of the Phase Impenetrability Condition (PIC): strong PIC (PIC$_1$) and weak PIC (dubbed PIC$_2$, following Müller 2004). One view is derived from PIC$_1$ in (31) where spell-out occurs as soon as the phase is complete, meaning once all elements (specifiers, head, and complement) are merged. Per PIC$_2$ in (32), the instant the head of the higher phase is merged, the complement of the phase is transferred to the interfaces and the elements become inaccessible for any further computational operations.

(31)   **Phase Impenetrability Condition (PIC$_1$)** (Chomsky 2000:108):

In a phase α with head H, the domain of H is not accessible to operations outside α, only H and its edge are accessible to such operations.

(32)   **Phase Impenetrability Condition (PIC$_2$)** (Chomsky 2001:14):

Given the structure [$_{ZP}$ Z ... [$_{HP}$ α [$_{H'}$ H YP]]], where H and Z are phase heads, the domain of H is not accessible to operations at ZP; only H and its edge are accessible

Another restrictive proposal for the timing of spell-out of the phasal complements is suggested by Chomsky (2008) who argues that the valuation of ɸ-features on phasal heads is what triggers Transfer. In particular, it is Agree that triggers the transfer of the spell-out domain to the interface: as soon as the ɸ-features of a phase head are valued, the complement of that phase is transferred to the interfaces.

2.2.3.2 Phase-based accounts of binding domains

Since the introduction of phases in Chomsky (2000, 2001), different views have been proposed to reduce binding domains to phase theory. The phase-theoretic concept of locality has been specifically employed to the Binding Condition A.[11] In these approaches, the binding domains are argued to be either whole phases (Lee-Schoenfeld 2008, Quicoli 2008) or the smallest spell-out domains (Charnavel & Sportiche 2016). Both views agree on the cyclicity of Binding Conditions and on that the anaphor and its antecedent need to be in the same domain. The domain for the whole-phase approach is the same phase that 'contains' the anaphor and its binder: e.g., vP and CP. The binding domain for the other phase approach, on the other hand, is the 'Spell-Out domain' of a phasal head, such as a tensed TP and DP. Consider (33).

(33) a. Abeer$_i$ says that Mary$_j$ hates herself$_*i/j$.

    b. [$_{CP}$ Abeer$_i$ says [$_{CP}$ that [$_{TP}$ Mary$_j$ [$_{vP}$ <Mary$_j$> hates herself$_*i/j$]]]].

The anaphor *herself* can only have the embedded subject *Mary* as its possible antecedent, but not the higher subject *Abeer*. Under the assumptions of the whole-phase approach, this is expected because the anaphor *herself* must be bound in the embedded phase vP (the legitimate binding domain). At a previous stage in the deviation, the licit antecedent *Mary* is merged at its agent thematic position at the specifier of vP, as shown in (33b), where both the anaphor and its binder are contained in the same phase. Thus, the higher antecedent is barred from binding the anaphor since it does not occur in the

---

[11]Furthermore, several studies provided analyses suggesting phasal domains for the Binding Condition B (Lee-Schoenfeld 2008, Quicoli 2008) and other studies argued for the elimination of the Condition altogether, claiming that Binding Condition B can be derived from economy principles like last resort (Hornstein 2001, Hicks 2009).

binding domain of the anaphor, here the vP phase. According the proposal that binding domains can be reduced to spell-out domains (Charnavel & Sportiche 2016), the reason why the anaphor *herself* in (33a) cannot co-refer with subject of the main clause, here *Abeer*, is that the antecedent and the anaphor are not in the same TP. The only possible anaphoric interpretation, therefore, is with the subject of the embedded clause, *Mary* since they are both in the same spell-out domain, namely the embedded TP.

According to the phasal approaches to binding, anaphors are constrained by the following conditions: the whole-phase approach (34a) and the spell-out domain approach (34b).

(34)   Phasal Binding Condition A

   a. A reflexive must be bound in its phase.   (Lee-Schoenfeld 2008:291)

   b. A plain anaphor must be bound within the Spell-Out domain containing it.
                                                              (Charnavel & Sportiche 2016:71)

According to these analyses, constructions wherein an anaphor and its antecedent occur in different domains/phases are not possible, like (35a) in English and schematized in (35b).

(35) a. *Mary believes that herself knows Bill.   (Charnavel & Sportiche 2016:72)

     b. *Mary believes [$_{CP}$ that [$_{TP}$ herself knows Bill]].

For the whole-phase approach, (35b) is ungrammatical because the anaphor and its antecedent are not contained in the same phase, namely the phase CP. Per the analysis of Charnavel & Sportiche (2016), the embedded tensed TP, the smallest spell-out domain, only contains the anaphor but not the binder *Mary*, hence the sentence is barred.

Although the examples in (33a) and (35a) appear to be innocuous to both phasal

approaches to the binding domain of the Condition A, taking long-distance anaphors into consideration, however, bears issues to these approaches. As we have seen above, the distributional properties of anaphors vary from one language to another. In some languages, anaphors can only be locally bound, while others allow cross-clausal binding relations. One of latter languages is Malki Arabic, along with Tamil, Mandarin Chinese, and Icelandic, as we have seen earlier in Section 2.1.2.[12] In these languages, subject anaphors can be anteceded across a finite clause boundary, as can be seen in Malki Arabic (26a), repeated below as (36a).

(36)   Malki Arabic
    a. ʕabiir$_i$ ti-ħassib   [(inn) [nafs-aha$_{i/*j}$ bi-ta-rooh al-maṭaar].
       Abeer 3SG.F-believe [COMP self-her FUT-3SG.F-go the-airport]
       'Lit. Abeer believes that herself will go to the airport.'

    b. ʕabiir$_i$ ti-ħassib [$_{CP}$ (inn) [$_{TP}$ nafs-aha$_{i/*j}$ [$_{vP}$ <nafs-aha> bi-ta-rooh al-maṭaar]]].

If the locality of binding to is be reduced to phasal domains, sentence (36a) is predicted incorrectly to be impermissible, as the anaphor is not bound in its phase, namely the lower vP or CP, or the smallest Spell-Out domain, here tensed TP, as illustrated in (36b)

Although long-distance anaphors are not discussed in the whole-phase approach, anaphors in the embedded subject position (as in (36a)) are treated in the Spell-Out domain approach as exempt from the phasal Binding Condition A in (34b), given the exemption definition of Charnavel & Sportiche (2016) in (37), taken from Charnavel (2019).

---

[12]In this thesis, the arguments that follow will be built based on data from Malki Arabic.

(37)   Definition of exempt anaphors (Charnavel & Sportiche 2016)

An . . . anaphor is exempt if it is not bound or if its binder is outside the smallest

spell-out domain containing it.   (Charnavel 2019)

However, such a definition does not explain how the subject anaphor can be bound across

a finite clause boundary, i.e., outside its phasal/spell-out domain. One proposal for

explaining the bound (and unbound) interpretations of long-distance anaphors suggests

that every spell-out domain contains a logophoric center (LogP), headed by a silent

operator Log_{Op} (Charnavel 2019). Such logophoric center regulates the binding variation

among long-distance anaphors.[13] However, as to be shown in the current proposal in this

thesis, such stipulated null operators are not necessary, at least with subject anaphors (in

Malki Arabic). Therefore, I suggest that long-distance bound embedded subject anaphors

are not exempt based on at least a structural, phasal-based definition of exempt

anaphors, similar to (37).

2.2.4   Anaphor Agreement Effect: Embedded Subject Anaphors

A pivotal topic related to subject anaphors is the anaphor agreement effect (Rizzi

1990). As noted in Reuland (2011), no condition (in the predicate-based theories of

binding) can prevent complex-anaphors from occurring in a subject position in a finite

clause since they are fully specified for their ɸ-features. However, this does not predict

why complex anaphors in English (38) cannot occur in a subject position. Reuland

concludes, based on the argument of Anagnostopoulou & Everaert (1999), that this issue

is a result of the morphosyntax of anaphors.

---

[13]Similar proposal is suggested in Sundaresan (2012) and her subsequent work. Sundaresan argues that long-distance anaphors are evidence of a perspectival phrase with a silent operator that co-varies for ɸ-agreement and regulates the bound and free interpretations of long-distance binding.

(38)   *Himself/John plays football.

Anagnostopoulou & Everaert (1999) argue that the possibility of subject anaphors in Modern Greek are due to the similarity of their internal and external syntactic structure to the structure of a possessive NP. When the anaphor *ton eafto tu* occurs in the object position (39a), it is assigned an accusative case. When it occurs in a subject position (39b), however, it is assigned a nominative case. The case assigner is the article which is specified for ɸ-features.

(39)   Modern Greek (Anagnostopoulou & Everaert 1999:106-8)

    a. O Jannis$_i$ ton$_i$ agapai ton eafto tu$_i$
       the John.NOM CLF.ACC loves   himself.ACC
       'John loves himself.'

    b. [O   eaftos   tu]$_i$ tu$_i$ aresi   [tu   Petru]$_i$
       the.NOM self.NOM his.GEN CLF.DAT like.3SG   the.DAT Peter.DAT
       'Peter pleases himself.'

Thus, in English, anaphors cannot be present in a subject position because there is no nominative form of *himself* as *heself*. Although the gap in the case paradigm of reflexives might explain why anaphors cannot occur in the subject position in English, other languages exhibit reflexives with a nominative case and, however, are barred from subject positions. This observation coupled with empirical evidence from Italian (40) motivated Rizzi (1990) to suggest the anaphor agreement effect which states that anaphors cannot be in subject positions because they (anaphors) occur in triggering positions for agreement, like the subject position in languages with subject-verb agreement. Thus, (40a) is grammatical because the anaphor (the theme of the verb *importa*) with the dative case does not trigger agreement, while the anaphor in (40b) is in

the nominative case and agrees with the verb *interessano* triggering agreement in person

and number, hence the ungrammatical status of (40b).

(40)   Italian (Rizzi 1990)
   a. A loro   importa   solo di se stessi.
       to them.DAT matters.3SG only of themselves.GEN
       'They$_i$ only matter to themselves$_i$.'

   b. *A loro   interessano solo se stessi.
       to them.DAT interest.3PL only themselves.NOM
       'They$_i$ are interested only in themselves$_i$.'

Similarly, dative anaphors in Icelandic can be subjects of finite clauses because they do

not trigger agreement, as shown in (41a). It is not possible for nominative subjects, on

the other hand, to be anaphors because nominative subjects trigger agreement, as

illustrated in (41b).

(41)   Icelandic
   a. Hún sagði að sér   þætti   vænt um mig.
       she$_i$ said that self$_i$.DAT was.SBJV fond of me
       'She said that she was fond of me.'   (Maling 1984,ex.(8b))

   b. *Jón segir að sig   elski   Maria.
       Jon$_i$ says that self$_i$.NOM loves.SBJV Maria
       'John says that he loves Maria.'   (Rizzi 1990,ex.(15b))

The generalization of Rizzi has been tested and explored further across other

languages (Woolford 1999, Tucker 2011). That multiple obviating cases to the anaphor

agreement effect seem to be counter-evidence to the anaphor agreement effect is not in

fact accurate, as demonstrated in Woolford (1999). One superficial violation case of the

anaphor agreement effect is noticed by Woolford in languages with object agreement. In

these languages, anaphors can be found in agreeing positions. The type of agreement

verbs show with these anaphors is of a special kind, so-called anaphoric agreement

(Woolford 1999), which is found in Swahili (42). The difference between such type of agreement and the canonical agreement between a nominal and a verb is that the latter typically involves a ɸ-agreement and the former is depicted in the realization of a special morpheme on verbs to signal reflexivity.

(42)   Swahili (Vitale 1981:137)
   a. Ahmed a-na-**ji**-penda   mwenyewe.
      Ahmed 3.SBJ-PRS-REFL-love himself
      'Ahmed loves himself.' (Empthatic reflexive)

   b. Ahmed a-na-**ji**-penda.
      Ahmed 3.SBJ-PRES-REFL-love
      'Ahmed loves himself.'

As can seen in (42a), the agreement on the verb 'love' with the anaphor is a reflexive infix *-ji-* and not only agreement in ɸ-features. This is also evident in (42b) where the verb can be rendered intransitive and the anaphoric interpretation holds via the anaphoric morpheme on the verb.

Further insights into the anaphor agreement effect are provided in Woolford (1999) and Tucker (2011). Several languages utilize different strategies to avoid the violation of the anaphor agreement effect. Languages like Swahili, as we have seen above, exhibit an anaphoric agreement in which there is no agreement between anaphor and the verb. Another type of languages deletes the anaphor in the agreeing position and leaves an intransitive structure with an anaphoric interpretation. One of these languages is Inuit, an object-verb agreement language. In Inuit, as described in Woolford (1999), when the object of the verb is an anaphor, the verb shows no regular agreement (43a) with the anaphor. The construction, rather, either turns intransitive (43b) or the anaphor carries

an oblique case (43c), as the case in Italian and Icelandic (see above).

(43)   Inuit (Bok-Bennema 1991)

a. *Hansiup   immi   asap-puq.
   Hansi.ERG himself.ABS wash-IND.3SG.3SG
   'Hansi washed himself.'

b. Asap-puq.
    wash-IND.3SG
    'He washed himself.'

c. Angut immi-nut   taku-vuq.
    man   himself-DAT see-IND.3SG
    'The man sees himself.'

Other languages, like Italian and Icelandic, show default agreement when an anaphor

occurs in the agreement-triggering position, in addition to the obliqueness of case.

Another group of languages (like Greek and Selayarese) protects anaphors in agreement

positions by embedding them inside a possesser DP or PP, thus with no agreement, as

indicated in Woolford (1999).

In the case of subject anaphors in Malki Arabic, the situation in Malki Arabic is not

similar to Italian, despite case differences (that Malki Arabic does not have dative

subjects). In Italian, unlike Malki Arabic, anaphors can occur in an agreement position

but with a default agreement on the verb. In Malki Arabic, in contrast, the anaphor

occurs in an agreement subject position and the embedded verb shows a co-varying

agreement with the subject anaphor in person, number, and gender, evident by the

sentences in (44).

(44)   Malki Arabic

a. juusif   ya-ħassib   inn   nafs-uh   bi-ya-rooh   as-souq.
    yousef.NOM 3SG.M-believe COMP self-him.ACC FUT-**3SG.M**-go the-mall.ACC
    'Yousef believes that he will go to the mall.'

b. ʕabiir   ti-ħassib   inn nafs-aha   bi-ta-rooh   as-souq.
Abeer.NOM 3SG.F-believe COMP self-her.ACC FUT-3**SG.F**-go the-mall.ACC
'Abeer believes that she will go to the mall.'

c. ar-rijaal   ya-ħassib-uun inn nafs-ahum   bi-ya-rooh-uun as-souq.
the-men.NOM 3PL.M-believe COMP self-their.ACC FUT-**3.M**-go-**PL** the-mall.ACC
'The men believe that they will go to the mall.'

d. al-bas   ya-ħassib   inn nafs-uh   bi-ya-ṣiid
the-cat.NOM 3SG.M-believe COMP self-him/it.ACC FUT-**3SG.M**-catch
al-faar   bisuuhulaah
the-mouse.ACC easily
'The cat believes that he/it will catch to the mouse easily.'

e. ʔana ʔa-ħassib inn nafs-i   ba-ʔa-rrkab   maʕa maaʒid.
I.NOM 1SG-believe COMP self-my.ACC FUT-**1SG**-ride with Majid
'I believe that I will ride with Majid.'

f. ʔaint   ta-ħassib   inn nafs-ak   ba-ta-rrkab   maʕa maaʒid.
you.M.NOM 2SG.M-believe COMP self-your.M.ACC FUT-**2SG.M**-ride with Majid
'You.M believe that you.M will ride with Majid.'

g. ʔaint-i   ti-ħassib-iin inn nafs-ik   ba-ti-rrkab-iin maʕa maaʒid.
you.NOM-F 2SG-believe-F COMP self-your.F.ACC FUT-**2SG**-ride-**F** with Majid
'You.F believe that you.F will ride with Majid.'

Moreover, as evident by the data above, Malki Arabic does not bear a resemblance to

Greek nor Selayarese in which anaphors can occur in an agreement position only when

they are embedded in a possesser DP or PP positions. The subject anaphor, furthermore,

is neither like Italian nor Icelandic which mark reflexives with oblique case, like dative.

As shown in the sentences in (44) and argued in the syntax of case in Arabic (Fassi Fehri

1993, Benmamoun 2000, Aoun et al. 2010), Arabic has a nominative-accusative case

system which does not include oblique case variants. What the data above can also

indicate is that the subject anaphor cannot be explained like Swahili which marks the

verb with an anaphoric agreement. The anaphor in the embedded subject position does

exhibit ɸ-agreement with the verb, similar to the agreement of other nominals with verbs.

The above observations indicate that constructions with subject anaphors in Malki Arabic seem to be a counterexample to the anaphor agreement effect. However, in line with Woolford (1999), among others, I show in this thesis that constructions with embedded subject anaphors (in Malki Arabic) do not constitute a bona fide violation case of Rizzi's generalization.

2.3 TOWARDS A SOLUTION

As shown above, subject anaphors seem to be problematic to the phase-based analyses of the Binding Condition A because the anaphor can be bound outside its phasal and spell-out domain. Both phasal analyses of binding would explain the cross-clausal binding relation in Malki Arabic either as that the construction where subject anaphors are allowed is of an ECM type (a TP clause) or that the complement is not a CP but a TP as non-phase, as per the whole-phase approach. As I will show in chapter 3, however, the complement clause which contains the long-distance anaphor is a CP based on evidence from the morphosyntactic and semantic properties of embedded CP clauses.

Furthermore, the main goal of this thesis is to show that configurations with long-distance bound subject anaphors are not counter-evidence against the phasal analyses of binding domains. Therefore, I suggest that the availability of long-distance subject anaphors is dependent on the internal structure of the complement clause. Malki Arabic only allows long-distance anaphors in the external argument position of finite embedded clauses and not the subject position in single clauses (45) below.

(45)   Malki Arabic
        *nafs-uh   ya-rooh   as-souq biduun abu-uh
        self-him.NOM 3SG.M-go the-mall without father-his.GEN
        'Lit. Himself goes to the mall without his father.'

According to this empirical observation, besides others to come later in the thesis, I propose that the cross-clausal binding relation in Malki Arabic is similar to other cross-clausal phenomena that constitute an A-dependency with elements in another clause by a movement to the edge of the CP phase. Thus, the movement to the left periphery of the embedded clause allows the embedded subject anaphor and its antecedent to be in the same binding domain. This hypothesis comes in line with Huang & Tang (1991) in following a phrasal movement and with Lebeaux (1984) and Chomsky (1986) in assuming a (covert) movement of anaphors. The difference between the proposal in this thesis and other anaphor movement approaches resides in that the landing site of the phrasal movement of long-distance anaphors is a position which is theoretically and empirically justified, rather than an adjunction to INFL.

Another point the thesis seeks to show is that the whole-phase approach is a better account to explain how the binding relation between the subject anaphor and its antecedent in the matrix clause in Malki Arabic can hold. The current proposal, however, is in contention with the spell-out domain approach. Taking spell-out domains as binding domains is an untenable approach. Testing such an approach with various proposals on the spell-out timing of phasal complements, I show that containing long-distance subject anaphors and their antecedents in a spell-out domain is not possible.

In addition, the approach in this thesis also shows that explaining long-distance anaphors bound across phases by a covert logophoricity operator (Charnavel 2019, 2020)

is not needed. Instead, I argue that subject anaphors can be regulated by essential syntactic operations like movement and properties of complementation. Furthermore, contra Charnavel's (2019) proposal, the hypothesis suggested in this thesis is theoretically desirable, as it is simpler and derived from basic operations and resorting to silent operators like Log_{Op} is not necessary, at least in Malki Arabic.

Moreover, identifying an anaphor as exempt from a particular condition does not explain the binding relation between the anaphor and its antecedent, as suggested in Charnavel & Sportiche (2016). In fact, as far as long-distance subject anaphors are concerned, the diagnostics of Charnavel & Sportiche's (2016) for exemption in (37) above is inconsistent with the agreement-based condition on anaphors in (46). Although the embedded subject anaphor is not bound in its phasal binding domain and occurs in an agreement position, it does not trigger exemption, contra the phasal definition of exemption in (37). As a contribution to such a contradiction, I show that the generalization which states that anaphors occurring in an agreement position do not trigger exemption from Condition A is on the right track, hence conforming to the agreement-based condition on (exempt) anaphors (Charnavel 2019) in (46).

(46)   **Agreement-based condition on (exempt) anaphors** (Charnavel 2019:76)

An anaphor cannot be exempt if it occurs in a syntactic position construed with agreement (e.g. nominative position in languages with subject agreement).

The confirmation to the generalization in (46) comes from predictions the current hypothesis in thesis makes on the anaphor agreement effect. Following from the proposal suggested in this thesis, I show that although embedded subject anaphors in Malki Arabic

initially seem to violate the anaphor agreement effect of Rizzi (1990), the anaphor does not 'stay' in its apparent agreement position in the embedded clause, but rather moves out of the agreement position to a higher position in the same clause. I conclude that languages like Malki Arabic employ a process through which the anaphor moves to a non-agreeing position to obey the anaphor agreement effect, and that this effect is possibly a condition that operates at the interface, probably at the PF. In sum, not only does the movement analysis proposed in this thesis support the condition that anaphors cannot be exempt in syntactic positions that trigger agreement, but it also contributes with new insights on the anaphor agreement effect as an interface condition.

CHAPTER 3

THE COMPLEMENT OF *ðANA/ħASSIB*: THE CLAUSE STRUCTURE

Due to the optionality of an overt complementizer in the complement clause of the verb *ðana/ħassib* 'believe' in Arabic, several proposals have been suggested to determine the size of the complement clause. One view suggests that the complement clause of the verb 'believe' lacks a CP layer and is therefore a TP (Soltan 2007). In another study, the complement clause is suggested to be a CP at first and then the complementizer undergoes truncation (Fassi Fehri 2012). A further suggestion put forward by Al-Balushi (2016) is that the complement clause is a CP, despite the apparent possibility of constructions with the verb *ðana* and its complement clause without an overt complementizer. In line with the latter approach, I argue that the clause size of complement clause of the verb 'believe' is a CP, based on the cross-linguistic (morpho)syntactic and semantic properties of CP complements which are different from those of other types of complements, like TP clause size.[14] As we will see below, I am going to show that the complement of the verb *ðana/ħassib* 'believe' is an embedded CP complement that has an optional finite complementizer, necessitates the presence of an embedded subject, can include sentential negation, has similar or different voice values from the matrix predicate, and allows a similar instance of the adverb that has another instance in the matrix clause. Moreover, it disallows focus movement to the matrix clause, has simultaneous or independent tense value with/from the tense value of the matrix verb, and can be evaluated for truth and allow indexical shift.

---

[14]Although the data is mainly from Malki Arabic, the same applies to MSA, unless I note otherwise.

3.1 PRESENCE OF THE COMPLEMENTIZER

One piece of evidence of embedded CP clauses is the presence of a finite clause introducer. In Malki Arabic, the presence of the finite complementizer *inn* is optional, as can be seen with the attitude predicates 'say' in (47a) and 'believe' in (47b). As I will show later in this chapter, the absence of the complementizer is not necessarily an indication of a non-finite CP clause.

(47)   Malki Arabic

    a. fahad ya-guul   (inn) juusif ʔaχaða   al-kuurah.
        Fahad 3SG.M-say COMP Yousef took.3SG.M the-ball
        'Fahad says that Yousef took the ball.'

    b. fahad ya-ħassib   (inn) juusif ʔaχaða   al-kuurah.
        Fahad 3SG.M-believe COMP Yousef took.3SG.M the-ball
        'Fahad believes that Yousef took the ball.'

With the complement of the non-finite verb 'want', the presence of a complementizer is prohibited (48), unlike its counterpart in MSA in which the presence of the non-finite clause introducer is obligatory, as can seen in (49) below.

(48)   Malki Arabic

    al-walad ya-bɣii   (*inn) ya-ʃuut   al-kurah.
    the-kid   3SG.M-want (that) 3SG.M-hit the-ball
    'The kid wants to hit the ball.'

(49)   MSA

    yu-riidu   al-walad-u ʔan/*Ø ya-rkul   al-kurat-a.
    3SG.M-want the-kid-NOM COMP/*Ø 3SG.M-hit the-ball-ACC
    'The kid wants to hit the ball.'

Although the presence of a complementizer might be controversial as there appear to be two complementizers, the non-finite complementizer *ʔan* in MSA is argued not to be a real complementizer but only a subjunctive mood assigner. See Albaty (2019), among

43

others, for a relevant discussion.

3.2   OBLIGATORY EMBEDDED SUBJECTS

Another property signaling a CP clause is the obligatoriness of hosting subjects. In

Malki Arabic, the complements of the verbs 'say' and 'believe' must include an embedded

subject, as shown in (50) and (51), respectively. Otherwise, the absence of the embedded

subject in the complement clauses yields the ungrammaticality of the sentences in (50b)

and (51b).

(50)   Malki Arabic

    a. fahad ya-guul   inn juusif ʔaχaða   al-kuurah.
       Fahad 3SG.M-say that Yousef took.3SG.M the-ball
       'Fahad says that Yousef took the ball.'

    b. *fahad ya-guul   inn Ø ʔaχaða   al-kuurah.
       Fahad 3SG.M-say that Ø took.3SG.M the-ball
       '*Fahad says that took the ball.'

(51)   Malki Arabic

    a. fahad ya-ħassib   (inn) juusif ʔaχaða   al-kuurah.
       Fahad 3SG.M-believe COMP Yousef took.3SG.M the-ball
       'Fahad believes that Yousef took the ball.'

    b. *fahad ya-ħassib   (inn) Ø ʔaχaða   al-kuurah.
       Fahad 3SG.M-believe COMP Ø took.3SG.M the-ball
       '*Fahad believes that took the ball.'

This can also be observed in MSA, as can be seen in (52b) for the speech predicate *guul*

'say' and (53b) for the predicate *ḍan* 'believe'.

(52)   MSA

    a. al-walad ya-guulu ʔanna muħammad-aa rakala   al-kurah.
       the-kid 3SG.M-say that Mohammad-ACC hit.3SG.M the-ball
       'The kind says that Mohammad hit the ball.'

b. *al-walad ya-guulu ʔanna Ø rakala   al-kurah.
   the-kid 3SG.M-say that   Ø hit.3SG.M the-ball
   '*The kid says that hit the ball.'

(53)   MSA

a. al-walad ya-ðun   ʔanna muħammad-aa   rakala   al-kurah.
   the-kid 3SG.M-believe that   Mohammad-ACC hit.3SG.M the-ball
   'The kid believes that Mohammad hit the ball.'

b. *al-walad ya-ðun   ʔanna Ø rakala   al-kurah.
   the-kid 3SG.M-believe that   Ø hit.3SG.M the-ball
   '*The kid believes that hit the ball.'

With non-finite clauses, in contrast, the embedded subject is obligatorily prohibited in

the complement of the non-finite verb 'want' (54a) in MSA. Similar observations arise

with the non-finite verb 'want' in Malki Arabic. Even if we enforce a complementizer with

the non-finite complement clause, embedding a subject in that clause is not possible, as

can be seen by the ungrammaticality of (55).

(54)   MSA

a. *al-walad yu-riid   ʔan   muħammad-aa   ya-rkul   al-kurah.
   the-kid 3SG.M-want that/to Mohammad-ACC 3SG.M-hit the-ball
   '*The kid wants that Mohammad hit the ball.'

b. al-walad yu-riid   ʔan   Ø ya-rkul   al-kurah.
   the-kid 3SG.M-want that/to Ø 3SG.M-hit the-ball
   'The kid wants to hit the ball.'

(55)   Malki Arabic

   *al-walad ya-bɣii   inn ʕali ya-ʃuut   al-kurah.
   the-kid 3SG.M-want that Ali 3SG.M-hit the-ball
   'Lit. The kid wants that Ali hits the ball.'

3.3   SENTENTIAL NEGATION AND NPI LICENSING

A further property of CP complements is their ability to host a sentential negation

and disallow cross-clausal NPI licensing. In Buryat, for instance, the embedded clause

includes the negation affix marker *güj* on the verb 'paint' (56), and negation only takes

scope in the embedded clause.

(56)   Buryat (Wurmbrand & Lohninger In Press)
[badm-in   xarʃə ʃɜrdə-xə-güj-ə-n'   ] dugər   mɜd-ɜ.
[Bandma-GEN fence paint-FUT-NEG-ACC-3SG] Dugar.NOM know-PST
'Dugar found out that Badma won't paint the fence.' (T. Bondarenko, p.c.)

The complement clauses of the verbs 'believe' and 'say' in Malki Arabic can also

include a sentential negation, marked by the negation marker *maa* as shown in (57a) and

(58a) respectively. Even if the matrix predicate is negated, the scope of that negation

does not affect the interpretation of the embedded clause (57c), as both the matrix and

embedded clauses have their own negation scope and because both clauses can have

sentential negation markers in the same construction, as shown in (57b) and (58b).

(57)   Malki Arabic
  a. fahad ya-ħassib   (inn) juusif maa ya-ʃuut   al-kurah baʕiid.
     Fahad 3SG.M-believe COMP Yousef NEG 3SG.M-hit the-ball faraway
     'Fahad believes that Yousef does not hit the ball faraway.'

  b. fahad maa ya-ħassib   (inn) juusif maa ya-ʃuut   al-kurah baʕiid.
     Fahad NEG 3SG.M-believe COMP Yousef NEG 3SG.M-hit the-ball faraway
     'Fahad does not believe that Yousef does not hit the ball faraway.'

  c. fahad maa ya-ħassib   (inn) juusif ya-ʃuut   al-kurah baʕiid.
     Fahad NEG 3SG.M-believe COMP Yousef 3SG.M-hit the-ball faraway
     'Fahad does not believe that Yousef hits the ball faraway.'

(58)   Malki Arabic
  a. fahad ya-guul   inn   juusif maa ya-ʃuut   al-kurah baʕiid.
     Fahad 3SG.M-say COMP Yousef NEG 3SG.M-hit the-ball faraway
     'Fahad says that Yousef does not hit the ball faraway.'

  b. fahad maa ya-guul   inn   juusif maa ya-ʃuut   al-kurah baʕiid.
     Fahad NEG 3SG.M-say COMP Yousef NEG 3SG.M-hit the-ball faraway
     'Fahad does not say that Yousef does not hit the ball faraway.'

46

Another property of CPs in Arabic is that matrix negation cannot license an NPI in the embedded clause, as shown in the data form Malki Arabic below with the verb 'say' and 'believe'. As the example in (59a) indicates, the NPI *ʔabad* in the embedded clause cannot be licensed because the negation marker *maa* is located in the higher clause. Because cross-clausal NPI licensing is barred, this indicates that the embedded clause is of the CP clause size, as CP clauses constitute an opaque domain for cross-clausal NPI licensing.

(59)   Malki Arabic

    a. ***maa** ya-quul   al-walad [inn   al-mabnaa iʃtaaɣaal   **ʔabad-aa**].
       NEG .3SG.M-say the-boy [COMP the-building work.PRFV.3SG.M never-ACC]
       'The boy does not say that the building has never been working.'

    b. ya-quul   al-walad [(inn) al-mabnaa   **maa** iʃtaaɣaal   **ʔabad-aa**].
      .3SG.M-say the-boy [COMP the-building NEG work.PRFV.3SG.M never-ACC]
      'The boy says that the building has never been working.'

(60)   Malki Arabic

    a. ***maa** ya-ħassib   al-walad [inn   al-mabnaa   iʃtaaɣaal   **ʔabad-aa**
       NEG 3SG.M-believe the-boy [COMP the-building work.PRFV.3SG.M never-ACC
      ].
      ]
      'The boy does not believe that the building has never been working.'

    b. ya-ħassib   al-walad [(inn) al-mabnaa   **maa** iʃtaaɣaal   **ʔabad-aa**
      .3SG.M-believe the-boy [COMP the-building NEG work.PRFV.3SG.M never-ACC
      ].
      ]
      'The boy believes that the building has never been working.'

3.4   VOICE MISMATCHING

Although voice matching is usually taken to be an accurate diagnostic for restructuring (Chung 2004, Wu 2013, Wurmbrand & Shimamura 2017), voice (mis)matching can also be an adequate test to show whether the embedded clause is

opaque or transparent to the higher voice of the matrix clause (Wurmbrand 2016). In other words, voice mismatching shows whether the embedded predicate is independent of and not predetermined by the voice of the matrix predicate. Consider (61) from Chamorro.

(61) Chamorro (Chung 2004)

a. pära tafan-ma-chägi ma-na'fanätuk ni lalahi siha
FUT 1PL.IR.INTR-[PASS] -try NPL.RL.INTR.[PASS] -hide OBL men PL
'The men will try to hide all of us.'

b. *Tinituhun ha-lalatdi si Dolores i famagu'un.
NPL.RL.INTR.[PASS] .begin 3SG.RL.TR-scold Dolores the children
'Dolores began to scold the children.'

Because both sentences above are monoclausal and the complements of the verbs 'try' and 'begin' are of the TP type, the embedded verb can only show the same voice of the matrix predicate. As demonstrated in (61a), if the matrix verb carries a passive voice, then the embedded verb must be marked for passive, as well. Similarly, the embedded verb must be in the active form if the main verb is in the active form, as can be seen by the ungrammaticality of the sentence in (61b).

Per the data in (62) and (63), Malki Arabic is similar to Chamorro in that the matrix and embedded verbs can be in the active or the passive voice form.

(62) Malki Arabic

a. fahad ya-guul (inn) juusif ʃaat al-kurah ʕalaa biit
Fahad 3SG.M-say.[ACT] COMP Yousef threw.3SG.M.[ACT] the-ball PREP house
al-ʒiiraan.
the-neighbors
'Fahad says that Yousef threw the ball in the neighbors' house.'

b. in-gaal inn al-kurah in-ʃaat-at ʕalaa biit al-ʒiiraan.
[PASS] -said COMP the-ball [PASS] -thrown-3SG.F PREP house the-neighbors
'It was said that the ball was thrown in the neighbors' house.'

c. in-gaal inn juusif ʃaat al-kurah ʕalaa biit al-ʒiiraan.
   [PASS] -said COMP Yousef threw.3SG.M.[ACT] the-ball PREP house the-neighbors
   'It was said that Yousef threw the ball in the neighbors' house.'

d. fahad ya-guul inn al-kurah in-ʃaat-at ʕalaa biit
   Fahad 3SG.M-say.[ACT] COMP the-ball [PASS] -thrown-3SG.F PREP house
   al-ʒiiraan.
   the-neighbors
   'Fahad says that the ball was thrown on the neighbors' house.'

Voice mismatching can also occur with the verb *ħassib*, as indicated by the data in (63).

The verbs in the embedded and matrix clauses can show all the four possible

combinations of voice forms: active-active (63a), passive-passive (63b), passive-active

(63c), or active-passive (63d).

(63)   Malki Arabic
   a. fahad ya-ħassib inn juusif ʃaat al-kurah ʕalaa
      Fahad 3SG.M-believe.[ACT] COMP Yousef threw.3SG.M.[ACT] the-ball PREP
      biit al-ʒiiraan.
      house the-neighbors
      'Fahad believes that Yousef threw the ball in the neighbors' house.'

   b. ħussib inn al-kurah in-ʃaat-at ʕalaa biit al-ʒiiraan.
      believed.[PASS] COMP the-ball [PASS] -thrown-3SG.F PREP house the-neighbors
      'It was believed that the ball was thrown in the neighbors' house.'

   c. ħussib inn juusif ʃaat al-kurah ʕalaa biit
      believed.[PASS] COMP Yousef threw.3SG.M.[ACT] the-ball PREP house
      al-ʒiiraan.
      the-neighbors
      'It was believed that Yousef threw the ball in the neighbors' house.'

   d. fahad ya-ħassib (inn) al-kurah in-ʃaat-at ʕalaa biit
      Fahad 3SG.M-believe.[ACT] COMP the-ball [PASS] -thrown-3SG.F PREP house
      al-ʒiiraan.
      the-neighbors
      'Fahad believes that the ball was thrown on the neighbors' house.'

However, Malki Arabic is similar to the behavior of voice in Chamorro with non-finite

verbs. The similarity stems from the fact that both the main and second embedded verbs

must be in the active voice (64a) or the passive (64b) with the non-finite verb 'want'.

(64) Malki Arabic

   a. al-walad ya-bɣii   ya-ʃuut   al-kurah.
        the-kid 3SG.M-want.⟨ACT⟩ 3SG.M-hit.⟨ACT⟩ the-ball
        'The kid wants to hit the ball.'

   b. in-bɣaa-at   al-kurah ta-n-ʃaat.
        ⟨PASS⟩ -wanted.3SG.F the-ball ⟨PASS⟩ -thrown-3SG.F
        'It was wanted for the ball to be thrown.'

   c. *al-walad ya-bɣii   in-ʃaat-at   al-kurah.
        the-kid 3SG.M-want.⟨ACT⟩ ⟨PASS⟩ -thrown-3SG.F the-ball
        'The kid wants the ball to be thrown.'

As argued in Albaty & Ouali (2018), Moroccan and Najdi Arabic dialects show

similar empirical facts, as shown by the non-finite constructions in (65) and (66)

respectively. Since non-finite clauses in most Arabic dialects are not CPs, the voice of the

embedded predicate must match that of the matrix predicate, as in (65c) for Moroccan

and (66c) for Najdi Arabic. Thus, as can be seen in (65a) and (65b), voice mismatching

between the embedded and matrix clause in Moroccan Arabic is not possible as the

embedded predicate is transparent to the higher predicate. Nor is the mismatch between

the voice aspects of the embedded and the matrix clause possible in Najdi Arabic, as

illustrated in (66a) and (66b).

(65) Moroccan Arabic (Albaty & Ouali 2018:188-189)

   a. *tənsa   j-ʒiib   ʕali l-ktaab.
        forgot.⟨PASS⟩ 3SG.M-bring.⟨ACT⟩ Ali the-book

   b. *nsa   ʕali j-tʒaab   l-ktaab.
        forgot.⟨ACT⟩ .3SG.M Ali 3SG.M-bring.⟨PASS⟩ the-book

c. tənsa   j-tʒaab   l-ktaab.
    forget.[PASS] .3SG.M 3SG.M-bring.[PASS] the-book
    'The book was forgotten to be brought.'

(66)   Najdi Arabic (Albaty & Ouali 2018:188-189)

a. *insii   j-dʒiib   ʕali al-ktaab.
   forgot.[PASS] 3SG.M-bring.[ACT] Ali the-book

b. *nisa   ʕali j-dʒaab   al-ktaab.
   forgot.3SG.M.ACT Ali 3SG.M-bring.[PASS] the-book

c. insii   j-dʒaab   l-ktaab.
    forgot.[PASS] .3SG.M 3SG.M-bring.[PASS] the-book
    'The book was forgotten to be brought.'

The conclusion from the facts presented above with regards to voice (mis)matching is that

if embedded clauses are CPs, then it should be possible to have different or similar voice

values on both the matrix and embedded predicates, unlike the situation with non-finite

embedded verbs which must show the same voice value of the matrix (or main) verb.

3.5   ADVERB CO-OCCURRENCE

A bi-sentential construction can have two instances of the same adverb, while a

monoclausal sentence can have only one instance of that adverb (Cinque 2006). In

Italian, for example, the adverb *giá* in (67a) can occur twice as the sentence is composed

of two clauses. However, evident by clitic climbing is the mono-clausality of (67b) in

which the co-occurrence of two instances of the same adverb is barred.

(67)   Italian (Cinque 2006:17)

a. Maria vorrebbe **giá** aver**lo giá** lasciato.

b. *Maria **lo** vorrebbe **giá** aver **giá** lasciato.
    'Mary would already want to have already left him.'

In sentences with the verb 'believe', Malki Arabic allows adverb co-occurrence, as shown

in (68a). The adverb 'quickly' can occur twice: one modifying the matrix predicate *ħassib*

'believe' and the other instance modifying the embedded predicate *raah* 'go'. This is, in

contrast, not the case with the non-finite (monoclausal) verb 'want' in which the

occurrence of two instances of the same adverb is not allowed (68b).

(68)   Malki Arabic

    a. fahad ya-ħassib   bi-suraʕah   inn juusif bi-ya-rooh   as-souq
        Fahad 3SG.M-believe PREP-quickly that Yousef FUT-3SG.M-go the-mall
        bi-suraʕah.
        PREP-quickly
        'Fahad believes quickly that Yousef will go to the mall quickly.'

    b. *fahad ya-bɣii   bi-suraʕah   ya-ʃuut   al-kurah bi-suraʕah.
        Fahad 3SG.M-want PREP-quickly 3SG.M-hit the-ball PREP-quickly
        'Fahad wants quickly to quickly hit the ball.'

This is not only with the verb 'believe'. The propositional predicate 'say' which selects a

CP complement shows a similar behavior with adverb co-occurrence, as can be seen in

(69) below.

(69)   Malki Arabic

    fahad daaʔim-an ya-guul   inn juusif bi-ya-rooh   as-souq daaʔim-an.
    Fahad always-ACC 3SG.M-say that Yousef FUT-3SG.M-go the-mall always-ACC
    'Fahad says always that Yousef will go to the mall always.'

Besides Italian and Malki Arabic, the same observation is noticed in MSA with the

auxiliary verb *kaan* which is usually monoclausal. Evident by the example in (70), only

one occurrence of the adverb 'always' is possible (70a), and therefore two instances of the

same adverb cannot occur in one clause (70b).

(70)   MSA (Albaty 2019:58)

    a. kaana   aħmad-u   ya-drusu   daaʔiman.
        was.3SG.M Ahmad-NOM 3SG.M-study always
        'Ahmad was always studying.'

b. *kaana   daaʔiman aħmad-u   ya-drusu   daaʔiman.
   was.3SG.M always   Ahmad-NOM 3SG.M-study always
   'lit. Ahmad was always studying always.'

3.6 FOCUS MOVEMENT

Another test that shows that the embedded clause of the verb 'believe' is a CP is

focus movement. Focus movement in embedded CP clauses is argued to be clause-bound,

as observed in Italian (Rizzi 2001) as well as MSA (Aoun et al. 2010, Ouhalla 1994,

Albaty 2019, among others) and Malki Arabic. The landing site of the focused object

resides in the embedded clause, and thus the foci cannot move to the left periphery of the

matrix clause. As can be seen in the examples from Italian below, the embedded direct

object 'this' can only move below the complementizer 'that' (71a) and not above it (71b).

(71)   Italian (Rizzi 2001)
   a. Credo che QUESTO avreste dovuto dirgli (non qualcos'altro)
       'I believe that THIS you should have said to him, not something else.'

   b. *Credo QUESTO che avreste dovuto dirgli (non qualcos'altro)
       'I believe THIS that you should have said to him, not something else'

Similarly, in MSA with the verb 'believe', the foci is not allowed to move cross-clausally

(72b) nor above the embedded complementizer (72c). Only below the complementizer of

the embedded clause is the legitimate landing position for the foci, as shown in (72d).

(72)   MSA (Albaty 2019)
   a. ʔaðunn-u ʔanna zajd-an   kasara   al-baab-a.
       think-1SG that   Zayd-ACC broke1SG.M the-door-ACC
       'I think that Zaid broke the door'

   b.??/*AL-BAAB-A ʔaðunn-u ʔanna zajd-an   kasara.
       the-door-ACC think-1SG that   Zayd-ACC broke1SG.M
       'I think that Zayd broke the door.'

c. *ʔaðunn-u AL-BAAB-A ʔanna zajd-an kasara.
   think-1SG the-door-ACC that Zayd-ACC broke1SG.M
   'I think that Zayd broke the door.'

d. ʔaðunn-u ʔanna AL-BAAB-A$_i$ kasara zajd-un t_i.
   think-1SG that the-door-ACC broke.1SG.M Zayd-NOM t
   'I think it is the door that Zayd broke.'

This is similar to the behavior of the propositional (speech) predicate 'say', as can be seen in (73) below. Most notably, the extraction of the foci (here the embedded object) to a position outside the embedded clause is not permissible, as shown in (73c).

(73) MSA

a. ya-guulu al-walad-u ʔanna muħammad-an rakala al-kurat-a.
   3SG.M-say the-kid-NOM that Mohammad-ACC hit.3SG.M the-ball-ACC
   'The kid says that Mohammad hit the ball.'

b. ya-guulu al-walad-u ʔanna AL-KURAT-A rakala muħammad-un t.
   3SG.M-say the-kid-NOM that the-ball-ACC hit.3SG.M Mohammad-NOM t
   'The kid says that Mohammad hit the ball.'

c. *ya-guulu al-walad-u AL-KURAT-A ʔanna muħammad-an rakal t.
   3SG.M-say the-kid-NOM the-ball-ACC that Mohammad-ACC hit.3SG.M t
   'The kid says that Mohammad hit the ball.'

However, that focus movement must be clause-bound is not the case with non-finite verbs like 'want'. The focused element 'the ball' cannot move within the embedded non-finite clause since that clause is not a CP, despite the presence of what looks like a complementizer (see Albaty 2019 for discussions on *ʔan* as a subjunctive assigner). As shown in (74c), the only landing site for the foci is the left periphery of the whole clause, which is a CP.

(74) MSA

a. yu-riidu al-walad-u ʔan ya-rkul al-kurat-a.
   3SG.M-want the-kid-NOM that/to 3SG.M-hit the-ball-ACC
   'The kid wants to hit the ball.'

54

b. *yu-riidu   al-walad-u   ʔan   AL-KURAT-A ya-rkul   t.
   3SG.M-want the-kid-NOM that/to the-ball-ACC 3SG.M-hit t
   'The kid wants to hit the ball.'

c. AL-KURAT-A, yu-riidu   al-walad-u   ʔan   ya-rkul   t.
   the-ball-ACC 3SG.M-want the-kid-NOM that/to 3SG.M-hit t
   'The kid wants to hit the ball.'

Focus fronting in Malki Arabic shows a similar behavior of focus movement noticed in

Italian and MSA. The foci 'the ball' cannot move to either the beginning of the sentence

(75b) nor immediately above the complementizer *inn* (75d). What this indicates is that

focus movement is bound in the embedded clause, a CP, and cannot move across the

embedded clause to the matrix clause.

(75)   Malki Arabic

   a. fahad ya-guul   inn juusif ʔaχaða   al-kuurah.
      Fahad 3SG.M-say that Yousef took.3SG.M the-ball
      'Fahad says that Yousef took the ball.'

   b. *AL-KUURAH fahad ya-guul   inn   juusif ʔaχað   t.
      the-ball   Fahad 3SG.M-say COMP Yousef took.3SG.M t
      Intended:'Fahad says that it is the ball Yousef took.'

   c. fahad ya-guul   inn   AL-KUURAH juusif ʔaχað   t.
      Fahad 3SG.M-say COMP the-ball   Yousef took.3SG.M t
      'Fahad says that it is the ball Yousef took.'

   d. *fahad ya-guul   AL-KUURAH inn   juusif ʔaχað   t.
      Fahad 3SG.M-say the-ball   COMP Yousef took.3SG.M t
      Intended:'Fahad says that it is the ball Yousef took.'

The verb 'believe' behaves similarly to the speech predicate 'say'. The foci cannot cross

the boundary of the embedded clause to the matrix clause (76b) and cannot land above

the complementizer (76d).

(76)   Malki Arabic

    a. fahad ya-ħassib   inn   juusif ʔaχaða   al-kuurah.
       Fahad 3SG.M-believe COMP Yousef took.3SG.M the-ball
       'Fahad believes that Yousef took the ball.'

    b. *AL-KUURAH fahad ya-ħassib   inn   juusif ʔaχað   t.
       the-ball   Fahad 3SG.M-believe COMP Yousef took.3SG.M *t*
       Intended:'Fahad believes that that it is the ball Yousef took.'

    c. fahad ya-ħassib   inn   AL-KUURAH juusif ʔaχað   t.
       Fahad 3SG.M-believe COMP the-ball   Yousef took.3SG.M *t*
       'Fahad believes that it is the ball Yousef took.'

    d. *fahad ya-ħassib   AL-KUURAH inn   juusif ʔaχað   t.
       Fahad 3SG.M-believe the-ball   COMP Yousef took.3SG.M *t*
       Intended:'Fahad believes that it is the ball Yousef took.'

This indicates that the verb 'believe' behaves similar to speech predicates, like the verb 'say'. Thus, the complement of these verbs is not a TP or a *v*P but a CP that precludes focus movement beyond its boundary and hosts left-peripheral properties like focus.

3.7   TENSE-MODAL-ASPECT (TMA) ELEMENTS AND TEMPORAL IN-DEPENDENCE

Other properties that determine whether the complement of the verb 'believe' is a CP are related to tense and its markers. Specifically, CP clauses can have TMA elements, like modals, and the tense of the clause is not predetermined by the predicate in the matrix clause, thus the temporality of the embedded predicate is independent of the tense of the matrix predicate.

Besides the property of including TMA elements, CP complements can either have a simultaneous temporal interpretation of matrix clause or the time can shift to the past or the future from the perspective of the matrix temporal property (Wurmbrand &

Lohninger In Press). Consider the sentences in (77) and (78). In (77) from English indicates, the embedded clause can be interpreted with a tense value independent of its counterpart of the matrix clause, as can be seen in (77a). Furthermore, both clauses in English can have similar tense values, as demonstrated in (77b).

(77) a. Clara believes/claims that Danny ate salad.

b. Nova claims/knows that she {stole/ is stealing/ will steal} Danny's salad {yesterday/ now/ tomorrow}. (Modified from Wurmbrand & Lohninger In Press)

Similarly, in Buryat (78), the temporal interpretation of the embedded clause can be simultaneous with the time reference of the matrix predicate (78a) or can shift to the future tense interpretation (78b). In (78a), both predicates have the past temporal value. However, in (78b), the temporal modifier 'tomorrow' modifies the embedded clause while 'yesterday' modifies the embedded clause, ending up with different temporal interpretations from each other. An important point worth remark is that the embedded clause has one of the TMA elements, namely the future suffix -xə in (78b), indicating that the embedded clause is a CP because CP clauses typically host TMA elements.

(78)   Buryat
a. sajənə [bi   tɜrgə ɜmdəl-ɜ-b   gɜžə] mɜd-ɜ.
    Sajana [1SG.NOM cart   break-PST-1SG COMP] know-PST
    'Sajana_i found out that she_i/I broke the cart. '

                                                    (Bondarenko 2017, Wurmbrand 2019)

b. [**üglödər**   badmə   xarš šɜrdə-**xə** gɜžə] **üstər**   dugər
    [**tomorrow** Badma-NOM fence paint-**FUT** COMP] **yesterday** Dugar.NOM
    mɜd-**ɜ**.
    know-**PST**
    'Yesterday Dugar found out that Badma will paint the fence tomorrow.'

                                                    (Modified from Wurmbrand & Lohninger In Press)

Embedded clauses in Malki Arabic also show a temporal independency of the tense of the matrix clause with the verb 'believe'. As shown in (79) below, the embedded verb can host the past (79a), present (79b) or the future (79c) temporal interpretations while the matrix verb can have the present, past, or the future tense values.

(79)   Malki Arabic

a. fahad ħassab   / ya-ħassib   / by-ya-ħassib   (inn)
    Fahad believed.**PST**.3SG.M / 3SG.M-believe.**PRS** / **FUT**-3SG.M-believe COMP
    juusif ʃaat   al-kurah ʕalaa biit   al-ʒiiraan   ʔams.
    Yousef threw.**PST**.3SG.M the-ball PREP house the-neighbors yesterday
    'Fahad believed/believes/will believe that Yousef threw the ball in the neighbors'
    house yesterday.'

b. fahad ħassab   / ya-ħassib   / by-ya-ħassib   (inn)
    Fahad believed.**PST**.3SG.M / 3SG.M-believe.**PRS** / **FUT**-3SG.M-believe COMP
    juusif ya-ʃuut   al-kurah ʕalaa biit   al-ʒiiraan   kul yauum.
    Yousef 3SG.M-throw.**PRS** the-ball PREP house the-neighbors every day
    'Fahad believed/believes/will believe that Yousef throws the ball in the neighbors'
    house every day.'

c. fahad ħassab   / ya-ħassib   / by-ya-ħassib   (inn)
    Fahad believed.**PST**.3SG.M / 3SG.M-believe.**PRS** / **FUT**-3SG.M-believe COMP
    juusif bi-ya-ʃuut   al-kurah ʕalaa biit   al-ʒiiraan   bukrah.
    Yousef **FUT**-3SG.M-throw the-ball PREP house the-neighbors tomorrow
    'Fahad believed/believes/will believe that Yousef will throw the ball in the
    neighbors' house tomorrow.'

d. fahad ħassab   / ya-ħassib   / by-ya-ħassib   (inn)
    Fahad believed.**PST**.3SG.M / 3SG.M-believe.**PRS** / **FUT**-3SG.M-believe COMP
    juusif gaaʕyad ya-ʃuut   al-kurah ʕalaa biit   al-ʒiiraan   ðalħiin.
    Yousef PROG   3SG.M-throw.**PRS** the-ball PREP house the-neighbors now
    'Fahad believed/believes/will believe that Yousef is throwing the ball in the
    neighbors' house now.'

In addition to the temporal independency of the embedded clause from the matrix clause, as example (79d) shows the embedded clause can also include a TMA element, namely the progressive aspectual marker *gaaʕyad*, which is a property of propositional complements.

The complement of the verb 'believe' is not the only complement that shows the aforementioned temporal and TMA properties. The complement of the speech predicate 'say' can also be temporally independent from the matrix clause, as can be seen in (80) in which the embedded verb is in the past (80a), present (80b) or the future (80c) and the matrix verb can be in the past, present or the future. In line with the complement of the verb 'believe', the complement of the verb 'say' can also have the progressive aspect marker *gʕd*, as can be seen in (80d) below.

(80)   Malki Arabic
  a. fahad gaal   / ya-guul   / bi-ya-guul   (inn) juusif
      Fahad said.**PST**.3SG.M / 3SG.M-say.**PRS** / **FUT**-3SG.M-say COMP Yousef
      ʔams   ʃaat   al-kurah ʕalaa biit   al-ʒiiraan.
      yesterday threw.**PST**.3SG.M the-ball PREP house the-neighbors
      'Fahad said/says/will say that Yousef threw the ball in the neighbors' house yesterday.'

  b. fahad gaal   / ya-guul   / bi-ya-guul   (inn) juusif
      Fahad said.**PST**.3SG.M / 3SG.M-say.**PRS** / **FUT**-3SG.M-say COMP Yousef
      ya-ʃuut   al-kurah ʕalaa biit   al-ʒiiraan   kul yauum.
      3SG.M-throw.**PRS** the-ball PREP house the-neighbors every day
      'Fahad said/says/will say that Yousef throws the ball in the neighbors' house every day.'

  c. fahad gaal   / ya-guul   / bi-ya-guul   (inn) juusif
      Fahad said.**PST**.3SG.M / 3SG.M-say.**PRS** / **FUT**-3SG.M-say COMP Yousef
      bi-ya-ʃuut   al-kurah ʕalaa biit   al-ʒiiraan   bukrah.
      FUT-3SG.M-throw the-ball PREP house the-neighbors tomorrow
      'Fahad said/says/will say that Yousef will throw the ball in the neighbors' house tomorrow.'

  d. fahad ya-guul   (inn) juusif gaaʕyad ya-ʃuut   al-kurah ʕalaa
      Fahad 3SG.M-say.**PRS** COMP Yousef PROG   3SG.M-throw.**PRS** the-ball PREP
      biit   al-ʒiiraan   ðalħiin.
      house the-neighbors now
      'Fahad says that Yousef is throwing the ball in the neighbors' house now.'

What indicates the autonomy of the embedded clause from the matrix clause and reveals

the category of embedded clause is that each one of these clauses can have its own

progressive marker without any effect on the sentence judgment, as illustrated in (81).

(81)　　Malki Arabic

fahad gaaʕyad ya-guul　　　(inn) juusif gaaʕyad ya-ʃuut　　　al-kurah
Fahad PROG　　3SG.M-say.PRS COMP Yousef PROG　　3SG.M-throw.PRS the-ball
ʕalaa biit　al-ʒiiraan　　ðalħiin.
PREP house the-neighbors now
'Fahad is saying that Yousef is throwing the ball in the neighbors' house now.'

A piece of evidence that the covertness of the complementizer does not indicate the lack

of the CP layer comes from the fact that embedded (without an overt complementizer)

and matrix clauses in MSA can show different and similar tense values, as shown in (82).

The matrix clause can have the perfective/past interpretation while the embedded clause

with the 'writing' event is interpreted in the future denoted by the future modifier

'tomorrow'.[15]

(82)　　MSA (Fassi Fehri 2012:249, (58))

ʔamsi　　ħasib-tu r-rajul-a　　ya-ktub-u　　r-risaalat-a　　ɣadan.
yesterday thought-I the-man-ACC 3SG.M-write-IND the-letter-ACC tomorrow
'Yesterday, I thought the man will write the letter tomorrow.'

In addition, the embedded clause can have the progressive aspectuality by the modifier

'now' and the present temporal interpretation while the matrix clause is interpreted in

the past (83). Empirically, this is unlike Malki Arabic in which the progressive aspect is

marked with the modifier *gʕd*.

---

[15]Look at Al-Balushi (2016) for arguments that the embedded subject is in a CP rather than a TP
clause, despite the lack of an overt complementizer.

(83)   MSA (Al-Balushi 2016:12)

ðanna-Ø   ʕaliyy-un ṭ-ṭullaab-a   ya-lʕab-uu-na
PST.believe.3SG.M-IND Ali-NOM   the-students-ACC IMPF.PRS-play.3-PM-IND
l-ʔaan-a.
the-now-ACC
'Ali believed the students are playing now.'

Despite the lack of an overt complementizer, the matrix and embedded clauses can also

have a simultaneous temporal interpretation, as can be seen in (84) in which both the

state of 'Ali thinking' and the event of 'the students sleeping' occurred in the past.

(84)   MSA (Al-Balushi 2016:10, (35))

ðanna-Ø   ʕaliyy-un ṭ-ṭullaab-a   bi-l-ʔams-i
PST.believe.3SM-IND Ali-NOM   the-students-ACC in-the-yesterday-GEN
naam-uu-Ø.
PST.sleep.3-PM-IND
'Ali believed the students slept yesterday.'

The data on temporal independency above indicates that there is no difference between

clauses with an overt or covert complementizer. The absence of the complementizer in

some constructions, therefore, does not provide compelling evidence for non-CP clauses.

3.8   CPS CAN BE TRUE OR FALSE

Another property of CP clauses is that they have truth values, unlike other types of

clauses, like TP or *v*P, which cannot be evaluated for truth. The embedded clause of the

verb 'believe' (85a) as well as the speech predicate 'say' (85b) can be evaluated for truth.

In contrast, TP complements cannot be evaluated for truth, as shown by the sentence in

(85c) where 'buying a ball for Yousef' is just an event that cannot be true or false, as

forcing truth evaluation is infelicitous (not meaningful), as indicated in (85d).

(85) Malki Arabic

    a. fahad ya-ħassib   (inn) juusif raah   al-madiinah, wa haaðaa
       Fahad 3SG.M-believe COMP Yousef went.PST.3SG.M the-Medina   and this
       ṣaaħ.
       right
       'Fahad believes that Yousef went to Medina, and this is right.'

    b. fahad ya-guul   (inn) juusif raah   al-madiinah, wa haaðaa ṣaaħ.
       Fahad 3SG.M-say COMP Yousef went.PST.3SG.M the-Medina   and this   right
       'Fahad says that Yousef went to Medina, and this is right.'

    c. fahad ya-bɣaa-ni   ʕa-ʃtari kurah li-juusif,   wa haaðaa ʃaii
       Fahad 3SG.M-want-1SG 1SG-buy ball PREP-Yousef and this   action
       mumtaaz ni-sawi-ih li-juusif   lama na-χaruʒ   al-ħadiigah.
       good   1PL-do-it PREP-Yousef when 1PL-went-out the-park
       'Fahad wants me to buy a ball for Yousef, and that is a good thing to do for
       Yousef when we go out to the park.'

    d. fahad ya-bɣaa-ni   ʕa-ʃtari kurah li-juusif,   #wa haaðaa ṣaaħ.
       Fahad 3SG.M-want-1SG 1SG-buy ball PREP-Yousef and this   right
       'Fahad wants me to buy a ball for Yousef, and #this is right.'

3.9 INDEXICAL SHIFT

Indexical shift is one of the CP properties (Sudo 2012, Anand 2006, Anand & Nevins 2004, Sundaresan 2018, 2012, Messick 2016, Deal 2017, among others). In fact, it is one of the finest distinguishing properties between a propositional complement (CP) and other types of complements, as only the former allows shifted indexicals (Wurmbrand & Lohninger In Press). Indexicals, unlike definite descriptions, refer to elements whose interpretation depends on utterance-related or contextual properties. The pronouns *I* and *you* and the adverbials *now* and *here* are some of these indexicals. The first and second person pronouns can refer to a speaker/addressee in the context or an antecedent in the sentence in the matrix CP. This process is called indexical shift. For example, languages like Buryat only allow first/second person indexicals to optionally shift when embedded in

the nominative subject position (86).

(86)   Buryat (Modifed from Wurmbrand 2019)
        sajənə [bi   tɜrgə ɜmdəl-ɜ-b   gɜžə   ] mɜd-ɜ.
        Sajana [1SG.NOM cart   break-PST-1SG COMP] know-PST
        'Sajana$_i$ found out that she$_i$ broke the cart. '   (Bondarenko 2017:19, (83))
        'Sajana found out that I broke the cart.'   (T. Bondarenko, p.c.)

As the (non)shifted readings of (86) indicate, the first person pronoun can either refer to

the speaker of the utterance or the subject *Sajana* of the matrix clause. However, it is not

possible for the first person pronoun to have the first reading as the pronoun obligatorily

cannot shift and refer to the matrix subject when they are in the embedded accusative

subject position (87).

(87)   Buryat (Bondarenko 2017:19, (82))
        sajənə [naməjə   tɜrgə ɜmdəl-ə(*-b)   gɜžə   ] mɜd-ɜ.
        Sajana [1SG.ACC cart   break-PST(*-1SG) COMP] know-PST
        *'Sajana$_i$ found out that she$_i$ broke the cart.'
        'Sajana found out that I broke the cart.'

However, the behavior of indexical shift in Buryat is not exactly the same in Malki

Arabic. First, in most languages indexical shift is most often noticeable with

speech/attitude report predicates (Anand 2006, Shklovsky & Sudo 2014, among others),

but in Malki Arabic indexical shift is possible not only with the speech predicates like

'say' (88) and (90) but also with the intentional predicate 'believe' (89), similar to Uyghur

(Shklovsky & Sudo 2014, Sudo 2012). Secondly, as shown in the examples from Buryat

above, indexical shift is sensitive to the case of the embedded subject: indexical shift is

only possible with the nominative and not the accusative embedded subject position.

Malki Arabic, in contrast, behaves differently in that indexicals can shift in the accusative

embedded subject position. As shown in (88) and (89), the first person indexical can shift even though it occurs in the accusative embedded subject position.

(88)   Malki Arabic

fahad   qaal   inn-i   lagii-t   al-walad fii   al-mazraaʕah.
Fahad said.PST.3SG.M COMP-1SG.ACC found-1SG the-boy PREP the-farm
'Fahad said that I found the boy in the farm.'

   (i)   Fahad$_i$ said that he$_i$ found the boy in the farm.

   (ii)   Fahad said that I (SPEAKER) found the boy in the farm.

(89)   Malki Arabic

fahad ya-ħassib   inn-i   lagii-t   al-walad fii   al-mazraaʕah.
Fahad 3SG.M-believe COMP-1SG.ACC found-1SG the-boy PREP the-farm
'Fahad believes that I found the boy in the farm.'

   (i)   Fahad$_i$ believes that he$_i$ found the boy in the farm.

   (ii)   Fahad believes that I (SPEAKER) found the boy in the farm.

Similar observations obtain with the second person pronoun, as can be seen in (90) in which both the shifted reading with the object (90i) and the unshifted reading for the hearer (90ii) are possible.

(90)   Malki Arabic

fahad ʕaalam   saʕad$_i$ inn-ak$_{i/j}$ lagii-t   al-walad fii
Fahad told.PST.3SG.M Saad COMP-2SG.M.ACC found-1SG the-boy PREP
al-mazraaʕah.
the-farm
'Fahad told Saad that you found the boy in the farm.'

   (i)   Fahad told Saad$_i$ that you$_i$ found the boy in the farm.

   (ii)   Fahad told Saad that you (ADDRESSEE/HEARER) found the boy in the farm.

3.10   ECM VS. CP

A common debate in the literature of Arabic syntax on the verb 'believe' is whether it has an ECM-like complement clause. This comes as a result of the accusative-marked subject in the embedded clause in Arabic. Contra arguments presented in Soltan (2007)

and elsewhere, I argue that the complement of the attitude predicate 'believe' is a full clause and not of an ECM type based on effects shown by Condition B of the Binding Theory, long-distance passivization, and the possibility of focused and clitic-left dislocated (CLLD) elements to to precede the embedded subject.

Binding Condition B has been commonly used to test whether an element occurs in one clause or a different clause. In other words, a pronoun occurring in a monoclausal construction cannot be anteceded by a nominal in the same clause. When that pronoun however occurs as the embedded subject in a sentence with two clauses, the pronoun can take the matrix subject as its antecedent. Such behavior of pronouns follows from the Condition B of the Classical Binding Theory: a pronoun must be free in its binding domain. Condition B of the Binding Theory is used as a test, for instance, to see whether the accusative DP is in the matrix or the embedded clause in Janitzio P'urhepecha (Zyman 2018). By the same token, Binding Condition B can be used to test if the complement of the verb 'believe' is an ECM or a CP clause. A possibility of a pronominal occurring in the embedded subject position to co-refer with the matrix subject means that the embedded subject (the pronoun) occurs in a non-local domain (another full clause, aka CP). If that bound interpretation does not hold between the pronoun and the nominal, as the case with ECM clauses (91a), it indicates that the pronoun (embedded subject) is in the same clause of the higher subject.

(91)   Malki Arabic
    a. χuluud$_i$ ti-ħassib-aha$_{*i/j}$ ta-laʕab   bi-sayyarat maaʒid.
       Kholoud 3SG.F-believe-3SG.F 3SG.F-play PREP-car   Majid
       'Kholoud$_i$ believes she$_{*i/j}$ plays with Majid's car.'   ECM

b. χuluud$_i$ ti-ħassib inn-aha$_{i/j}$ ta-laʕab bi-sayyarat maaʒid.
   Kholoud 3SG.F-believe COMP-3SG.F 3SG.F-play PREP-car Majid
   'Kholoud$_i$ believes that she$_{i/j}$ plays with Majid's car.'   Full Clause

Furthermore, ECM clauses allow raising of the embedded subject to the matrix subject

position when the matrix verb is passivized (92a), while this is not the case with an overt

complementizer introducing a CP complement and preceding the embedded subject, as

can be seen in (92b) where the embedded subject 'Abeer' cannot raise to the higher

subject position, but rather stays in its embedded position (92c).

(92)   Malki Arabic

   a. ħussib-at ʕabiir ti-saafir baaryys ðalħiin.
      PASS.PST.believed-3SG.F Abeer 3SG.F.travel Paris now
      'Abeer was believed to be travelling to Paris now.'

   b. *ħussib-at ʕabiir inn t ti-saafir baaryys ðalħiin.
      PASS.PST.believed-3SG.F Abeer COMP t 3SG.F.travel Paris now
      'Abeer was believed to be travelling to Paris now.'

   c. ħussib inn ʕabiir ti-saafir baaryys ðalħiin.
      PASS.PST.believed-3SG.M COMP Abeer 3SG.F.travel Paris now
      'It was believed that Abeer is travelling to Paris now.'

Another test to distinguish ECM clauses from CP clauses is the (im)possibility of a

focused or CLLD-DP to precede the embedded subject. ECM subjects cannot be

preceded by a focused (93a) nor CLLDed (93b) DP. However, it is possible for these

elements to precede the subject of an embedded CP with an overt complementizer for the

focused DP (93c) and for the CLLDed DP, as in (93d).

(93)   Malki Arabic

   a. *fahad ħassab KUURAT AL-GAADAM maaʒid ya-laʕab t, maa huu
      Fahad believed.PST.3SG.M ball the-foot Majid 3SG.M-play t NEG it
      kurrat al-saalah.
      ball the-basket
      'Fahad believed that football Majid plays, not basketball.'   Focus without C

b. *fahad ħassab **[kuurat al-gaadam]**$_i$ maaʒid ya-laʕab-**aha**$_i$.
Fahad believed.PST.3SG.M ball the-foot Majid 3SG.M-play-it
'Fahad believed that football Majid plays, not basketball.' CLLD without C

c. fahad ħassab [inn KUURAT AL-GAADAM maaʒid ya-laʕab t,
Fahad believed.PST.3SG.M [COMP ball the-foot Majid 3SG.M-play t
maa huu kurrat al-saalah].
NEG it ball the-basket]
'Fahad believed that football Majid plays, not basketball.' Focus with C

d. fahad ħassab [inn **[kuurat al-gaadam]**$_i$ maaʒid ya-laʕab-**aha**$_i$
Fahad believed.PST.3SG.M [COMP ball the-foot Majid 3SG.M-play-it
].
]
'Fahad believed that football Majid plays, not basketball.' CLLD with C

In CP embedded clauses, as shown in (94a), the embedded subject can show agreement with a complementizer, as the subject is not an argument of the matrix verb. In contrast, ECM embedded subjects can be realized as an encliticized pronominal attached to the matrix predicate (94b), and thus the occurrence of overt argument (a pronoun (94c) or nominal (94d)) after the pronominal clitic is not possible.

(94) Malki Arabic

a. maaʒid ya-ħassib inn-aha hiih/ʔamzaad ta-laʕab bi-al-kuurah
Majid 3SG.M.believe COMP-3SG.F she/Amjad 3SG.F-play PREP-ball
baaraa.
outside
'Majid believes that she/Amjad plays with the ball outside.'

b. maaʒid ya-ħassib-aha ta-laʕab bi-al-kuurah baaraa.
Majid 3SG.M.believe-3SG.F.OBJ 3SG.F-play PREP-ball outside
'Majid believes that she plays with the ball outside.'

c. *maaʒid ya-ħassib-aha hiih ta-laʕab bi-al-kuurah baaraa.
Majid 3SG.M.believe-3SG.F she 3SG.F-play PREP-ball outside
Intended:'Majid believes that she plays with the ball outside.'

d. *maaʒid ya-ħassib-aha   ʔamzaad ta-laʕab   bi-al-kurah baaraa.
   Majid   3SG.M.believe-3SG.F Amjad   3SG.F-play PREP-ball   outside
   Intended:'Majid believes that Amjad plays with the ball outside.'

In conclusion, some tests seem to show that the verb 'believe' selects two complement clause sizes: TP and CP. Another set of other diagnostics, on other hand, shows that the complement clause is only a CP. What these tests do not show is that the type of complement clause of the verb *ħassib/ðana* is always a TP, as argued in Soltan (2007). Rather, when the complement clause is introduced with an overt complementizer, that clause shows all the properties of a CP clause size. Crucial to this thesis is that the embedded clause with an overt complementizer is in fact a CP and not of a smaller size of clauses (TP/vP). Long-distance embedded subject anaphors in Malki Arabic therefore occur in an embedded CP clause.

CHAPTER 4

PROPOSAL

In this chapter, I argue that the cross-clausal binding relation in Malki Arabic is not a counterexample to the phasal analysis of the Binding Condition A. To show that this is the case, I adopt the analyses of Wurmbrand (2019) and Lohninger et al. (2022) which suggest that cross-clausal A-phenomena (across finite clauses) are a result of a (A) movement to an A-position in the edge of the embedded clause. By the same token, I suggest that the long-distance bound subject anaphor in the embedded clause features another case of a cross-clausal A-dependency that involves a (covert) movement to the embedded Spec-CP, yielding accessibility between the anaphor and its antecedent and taking a whole phase, rather than a spell-out domain, as the binding domain. I eventually show that having an A-position in the CP-domain provides further support that phases can account for the locality of binding domains, even if a binding relation spans across finite clauses. I provide evidence that the A-position does not appear with every verb; rather, the selectional properties of matrix verbs are determinate of whether the complement clause has such position in its edge. In this chapter, I also demonstrate the implications of the suggested proposal on the anaphor agreement effect and the possibility to account for the accusative-marked embedded subjects in MSA.

In Section 4.1, an overview of phenomena spanning across finite (full) clauses cross-linguistically is provided, followed by the theoretical analysis adopted in the thesis. Section 4.2 furnishes the proposal I suggest to account for the cross-clausal binding data

from Malki Arabic. In Section 4.3, I also touch upon the implication of the suggested proposal of the long-distance subject anaphors on agreement-based definition of anaphors and the phasal definition of exempt anaphors. Furthermore, in a subsequent section I show subject anaphors do not violate the anaphor agreement effect (Rizzi 1990) and that Malki Arabic does not utilize any of the strategies that other languages employ to obey such an effect and I discuss the prospect of this effect being an interface condition. Finally, in Section 4.3.2 I discuss the current theoretical status of the embedded accusative-marked subject and whether the proposal put forward in this thesis might contribute to the understanding of this problem in case theory.

4.1 CROSS-CLAUSAL A-DEPENDENCIES

In this section, I first introduce constructions and phenomena which involve linguistic elements establishing a cross-clausal relation with an element in another clause. Secondly, the theoretical analyses, especially those utilizing constructs mainly from the minimalist framework, of cross-clausal configurations are addressed. In particular, I give an overview of the analysis of Wurmbrand (2019) and its recent developments in Lohninger et al. (2022). In a subsequent section, I give a sum of most (if not all) phenomena that span a finite clause boundary from Malki Arabic along with MSA and other spoken varieties.

4.1.1 Empirical background

A-phenomena, like agreement and Case, and A-dependencies, like raising to object (RtO) or subject (RtS), between two elements in either the same clause or different clauses are empirically attested across languages. Of the latter type of dependencies are constructions in languages like Tsez and Hindi which show long-distance agreement

between a syntactic item in one clause with another in another clause. For example, in Tsez the matrix verb exhibits a noun-class agreement with the absolutive noun with class III in the complement clause, as shown in (95).

(95)   Tsez (Polinsky & Potsdam 2001:606)

Eni-r   [už-ā   magalu   b-āc'ru-ħi   ] b-iy-xo.
mother-DAT [boy-ERG bread.**III**.ABS III-eat-PST.PTCP.NMLZ] **III**-know-PRS
'The mother knows that the boy ate the bread.'

Other phenomena involving properties resulted from operations across finite clauses are hyperraising to subject (Hyper-RtS), hyperraising to object (Hyper-RtO) and Hyper-ECM. Such phenomena constitute the group dubbed Cross-Clausal A-dependencies (CCA, Wurmbrand 2019, Lohninger 2021). In languages with Hyper-RtS, like Brazilian Portuguese, Kipsigis, and Lubukusu, the embedded subject moves from its base-generation position in the embedded finite clause to the matrix subject position, as illustrated in (96). The argument *Babaandu* moves from the embedded clause with the overt complementizer *mbo* to the external argument position above the verb *lolekhana*.

(96)   Lubukusu (Carstens & Diercks 2013:2)

Babaandu ba-lolekhana [CP (mbo) *t* ba-kwa   ].
2.people   2.SA-seem   [CP COMP *t* 2.SA.PST-fall]
'The people seem like they fell/The people seem to have fallen.'

Hyper-RtS differs from RtS in that the former applies to subjects in finite clauses, whereas the latter is similar to the case with the raising verb *seem* in English. Raising out of embedded non-finite clauses is possible in English (97a), but raising a subject out of a finite clause (i.e., Hyper-RtS) is not allowed (97b). Then, the difference between Hyper-RtS and RtS is related to the type of the embedded clause. Hyper-RtS involves a raised subject out of the embedded finite clause (CP), while RtS includes a raised subject

from a non-finite clause (TP).

(97) a. Abeer seems to t have been all over the world.

b. *Abeer seems [that t were all over the world].

ECM involves constructions in which the subject of an embedded non-finite clause is assigned accusative case. An example of ECM construction comes from English (98a) where the embedded subject *her* is assigned an accusative case in the non-finite clause and not a nominative case. The accusative case assignment is disallowed, however, when the subject occurs in a finite clause (98b).

(98) a. John believes [$_{TP}$ **her** to be the best in her class].

b. *John believes [$_{CP}$ that **her** is the best in her class].

Hyper-ECM, however, is not totally similar to a regular ECM configuration. Hyper-ECM involves a case marking across a finite clause boundary, unlike regular ECM configurations where the embedded subject resides in an embedded non-finite clause, as illustrated in (98a) in English. Instances of Hyper-ECM can be found in other languages which allow ECM across a full embedded clause. Some of these languages are Mongolian (Fong 2019) and Buryat (Bondarenko 2017), among others. As you can see in the data from Mongolian (99) and Buryat (100), the embedded subject can be assigned an accusative case across a finite clause boundary by, presumably, the matrix functional head v.

(99)   Mongolian (Fong 2019:2)
       Bat [margaash Dulmaa-Ø/g   nom unsh-n   gej   ] khel-sen.
       Bat [tomorrow Dulmaa-NOM/ACC book read.N.PST COMP] say.PST
       'Bat said that Dulmaa will read a book tomorrow.'

(100)   Buryat (Bondarenko 2017:1)
bi [sajən-Ø/ijə   du: du:lə-žə bɜ:-ga gɜžə] du:l-a-b.
1SG [Sajana-NOM/ACC song sing-CVB be-PST COMP] hear-PST-1SG
'I heard that Sajana sang a song.'

A further cross-clausal A-phenomenon is Hyper-RtO. Hyper-RtO is different from
Hyper-RtS in terms of which a landing syntactic position the raised element occupies.
The position for Hyper-RtO is the object position of the higher clause while it is the
matrix subject position for Hyper-RtS. What makes these types of hyperraising distinct
from Hyper-ECM theoretically is that overt movement to a higher clause occurs with
Hyper-RtO and Hyper-RtS but not with Hyper-ECM. In Hyper-ECM configurations, on
the other hand, the embedded subject covertly moves to any position below the verbal
head in the matrix clause and still can be marked an accusative case.
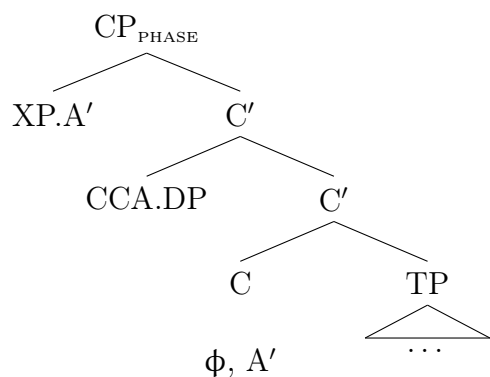
4.1.2 Theoretical background

Despite the differences between the aforementioned cross-clausal constructions, a
unification of all A-dependencies (including the types of hyperraising along with
long-distance agreement) under a single empirical domain involving arguments–called
empirical domain A–has been suggested initially in Wurmbrand (2019) and developed
further in Lohninger et al. (2022). The empirical domain A includes any configuration in
which an argument that occupies a position in an embedded finite clause has a
dependency relation with another argument in the matrix clause (Lohninger et al. 2022).

Based on an extensive typological survey subsumed under the empirical domain A,
Wurmbrand (2019) and Lohninger et al. (2022) summarized the common properties of
CCA-constructions. One property of CCA-configurations is their optionality across
languages. Hyper-ECM is optional in Buryat, for instance, where the embedded subject

73

can surface with either a nominative or accusative case. Another unique observation of CCA-constructions is their occurrence restriction with a special kind of verbs, like verbs of knowledge and belief. Then, the other property of the CCA-configurations is that their unique structure comes from the selectional properties of the matrix verbs (Lohninger et al. 2022). Some verbs allow dependencies across clauses, while others do not. As I will show below, Malki Arabic is in line with this observation, but allowing such dependencies with only a group of verbs is not always right with respect to Hyper-ECM where the embedded subject is always marked with an accusative case with all matrix predicates in MSA.

Based on the shared syntactic behavior among all CCA-configurations (occurring across full clauses), Wurmbrand (2019) suggests a theoretical analysis to unify all instances of A-dependencies across a finite clause boundary. Wurmbrand suggests that CCA-constructions are evidence for an A-position in the edge of CP whose head has A and A′ properties. Wurmbrand's proposal is schematized in (101) below.

(101)

```
                CP_PHASE
               /        \
          XP.A′          C′
                        /    \
                  CCA.DP      C′
                             /    \
                           C        TP
                                   /  \
                         ϕ, A′    ⟨...⟩
```

(Wurmbrand 2019:598)

Two vital ingredients of this analysis are the A-position in Spec-CP and ϕ-features on C. The suggestion for an A-position in Spec-CP has been called for in different studies

on Buryat (Bondarenko 2017) and Mongolian (Fong 2019). Both Bondarenko and Fong suggest that the Hyper-ECM constructions in Buryat and Mongolian are a result of the external argument of the embedded clause undergoing an A-dependency with a matrix case assigner. Therefore, the external argument moves to Spec-CP because the CP-domain has an A-feature and yet (for Fong) preserves its 'A′-quality'. The most compelling and extensive proposal for CCA-constructions is of Wurmbrand (2019). Wurmbrand's CCA analysis diverges from the (almost) similar proposals of Bondarenko and Fong in that only the analysis of Wurmbrand sheds light onto the combined A and A′ features of the C domain, preserving the A and A′ properties of the CP-domain and accounting for the cross-linguistic variation in CCA-configurations among languages. In Wurmbrand's analysis, positions above the A-position in the edge of CP have their typical status of A′-positions, as illustrated in (101). The CCA.DP (the element undergoing an A-dependency relation with another element in the matrix clause) occupies an A-position and the XP projection above it has A′-status, a position that is for focus, *wh*-phrases, and so on.

The φ-features on the head C are motivated for a combined probe C. The position in the left periphery of the CP is essentially motivated by such features on the head C, A-features (like φ and/or θ) and A′-features (like TOP, FOC, among others). Following van Urk (2015), such a probe with a mix of A and A′ features on C is dubbed 'composite probe'. The combination of two types of features under one head is not new. The TP domain is suggested to have tense and agreement features bundled under the head T (see Bobaljik & Thráinsson 1998 for discussion). The A-position and hence A-features are also

75

motivated semantically. As discussed in Lohninger et al. (2022), the origin of A-features on C is a result of merging a predicational head R of a R(elator)P (den Dikken 2006, 2017) with a CP, yielding CP.R type of clause. The function of RP is to constitute a relation of predication between its specifier and complement. Here, the specifier is the CCA.DP (the element undergoing A-dependency with a matrix element) and the complement is CP (semantically, the predicate of C.R).[16]

Following a position-based approach to the distinction between A and A′ positions raises issues with the improper A-after-A′ constraint (which its roots come from Chomsky 1973) which states that an A-movement cannot follow an A′-movement, thus precluding cross-clausal operations. Per this condition, the CP-domain can only include A′-positions. A theoretical advantage of a featural definition of the A/A′ distinction (van Urk 2015), however, arises in the overcoming of the improper A-after-A′ constraint. Specifically, no violation to the improper A-after-A′ constraint is incurred as long as the position in Spec-CP is a result of mixed A/A′-features on C in an embedded clause.

According to Lohninger et al. (2022), the cross-linguistic variation among languages that allow CCA-constructions and those that do not, like English, is derived from whether a language has the CP type of complement clause with at least A-features on the head C which triggers movement and agreement[17] to an A-position for the element (CCA.DP/DP.A) involved in an A-dependency with a cross-clausal element.

What these analyses indicate is that languages with cross-clausal configurations

---

[16]In this thesis, I ignore the semantic justification in this proposal because, as I will show in a moment, a piece of evidence for A-features on C comes from complementizer agreement where C includes ɸ-features.

[17]For arguments and evidence for composite probing/A-features triggering agreement and movement, see Coon et al. (2021), Aldridge (2004, 2008, 2016), Legate (2014), van Urk (2015), and Bossi & Diercks (2019)

include a type of clause structure that involves an A-position triggered by A-features on the head of the complementizer phrase. Languages allowing cross-clausal configurations therefore must have A-properties on the head C to attract the movement of the element constituting a dependency with a matrix element to the A-position.

4.1.3 Cross-Clausal A-dependencies in Arabic

A CCA phenomenon in Arabic and some of its varieties is Hyper-RtS. Cases of Hyper-RtS are found in Moroccan Arabic (Ura 1994) and Jordanian Arabic (Farghal 2020). The subjects in both (102) and (103) move from their base-positions (marked with a trace for convenience) to the higher subject position (in boldface). Note also that in the following examples the raising of the subjects is out of full CP clauses.

(102)   Jordanian Arabic (Farghal 2020:43)
Mariam ʃikil-ha   [CP ***t*** bitħib   il-bu:za   ]
Mariam   appearance-3SG.F [CP *t* like-3SG.F-PRS the-ice-cream]
'It seems that Mariam likes ice-cream.'

(103)   Moroccan Arabic (Ura 1994)
Ttshab-et-li **mmi**   [CP beli   ʒat   ***t***]
seem-3SG.F mother-my [CP COMP come-3SG.F-PST *t*]
'It seems that my mother came.'

In addition, MSA exhibits a Hyper-ECM construction. As shown in (104), MSA is nevertheless distinct from other Hyper-ECM languages with respect to the optionality of the accusative case marking on the embedded subject. In Mongolian and Buryat, as shown above in (99) and (100), the embedded subject can alternate for a nominative or accusative case, whereas MSA only shows an accusative case on the embedded subject.

(104) MSA

ʕali-uu gaala   ʔanna fatimat-**aa/*uu** ðahaba-t   maʕ ʔab-i-aha.
Ali-NOM said.3SG.M COMP Fatima-ACC/*NOM went.3SG-F with father-GEN-her
'Ali said that Fatima went with her father.'

Like the Jordanian and Moroccan Arabic languages, Hyper-RtS can also be found in

Malki Arabic with the raising verb *ʃakl* 'seem'. In (105), the embedded subject of a finite

clause moves to the matrix subject position. A constellation of Hyper-RtS is evident by

the agreement on the raising verb which shows agreement with the raised subject 'Yousef'

from the embedded clause.

(105) Malki Arabic

   a. ʃakl [(inn) **juusif** maa ya-laʕab   bi-as-sayyarah].
      seem [COMP Yousef NEG 3SG.M-play PREP-the-car]
      'It seems that Yousef is not playing with the car.'

   b. **juusif** ʃakl-**uh**   [(inn) ***t*** maa ya-laʕab   bi-as-sayyarah].
      Yousef seem-3SG.M [COMP ***t*** NEG 3SG.M-play PREP-the-car]
      'Yousef seems to not be playing with the car.'

Arabic appears to belong to the group of languages that prohibits Hyper-RtO. As

shown by the ungrammaticality of sentence (106) from Malki Arabic and (107) from

MSA, the embedded subject cannot raise to the object position following the verb in the

matrix clause.

(106) Malki Arabic

   *ʔal-ʔumm ti-fikiir   **al-walaad** [inn ***t*** ya-laʕab   baraa].
   the-mother 3SG.F-think the-kid   [that ***t*** 3SG.M-play outside]
   'The mother thinks that the kid is playing outside.'

(107) MSA

   *al-walad-u   ya-ðunu   **ʕali-aa** [ʔanna ***t*** rakala   al-kurat-aa].
   the-kid-NOM 3.SG.M-believe Ali-ACC [that   ***t*** hit.3SG.M the-ball-ACC]
   'The kid believes that Ali hit the ball.'

78

However, as I show below, empirical data from Malki Arabic on cross-clausal binding with a subset of verbs manifests a possibility of a (covert) Hyper-RtO in the sense of Deal (2017) where the surface realization of the embedded subject is just a PF-realization of a lower copy of the embedded subject which raises covertly to a position at the edge of the embedded clause.

4.2 CROSS-CLAUSAL BINDING AS CCA-CONSTRUCTION

As I have shown in chapter 2, Malki Arabic seems to allow a long-distance antecedence relation between an embedded subject anaphor and its antecedent in the subject position of the matrix clause. As I will show in the current chapter, the long-distance binding relation in Malki Arabic in (26a), repeated below as (108), is not a genuine counterexample to the phasal analyses of the binding domain of Condition A, despite that the binding relation spans across a finite clause (a phase). Bearing this idea in mind, I show that the long-distance subject anaphor enters into a local antecedence relation with its antecedent in the higher clause via (A-)movement to the A-position in the embedded clause above the phase head C, similar to CCA-constructions.

(108) Malki Arabic
Ɛabiir$_i$ ti-ħassib [(inn) nafs-aha$_{i/*j}$ bi-ta-rooh al-maṭaar].
Abeer 3SG.F-believe [COMP self-her FUT-3SG.F-go the-airport]
'Lit. Abeer believes that herself will go to the airport.'

Recall that languages with CCA-constructions, like Buryat and Mongolian, are characterized by the optionality of the cross-clausal dependency, having a CP structure with an A-position in its edge, and constructions involving CCA-configurations are constrained by the selectional properties of the matrix predicate (Wurmbrand 2019,

Lohninger 2021, Lohninger et al. 2022). The binding data of long-distance subject anaphors from Malki Arabic appear to hold the same characteristics as CCA-constructions, which motivates treating the long-distance binding relation in Malki Arabic as a CCA-construction.

First, the fact that the anaphoric relation alternates with a pronominal binding, as shown in (109), indicates its optionality behavior, which is a hallmark of a CCA-configuration. Although the anaphoric dependency seems to be optional in Malki Arabic, the subject anaphor must be anteceded by an antecedent in the matrix clause. This follows from the observation of Wurmbrand and Lohninger et al.: once A-dependency across clauses is established, it becomes obligatory.

(109)   Malki Arabic
   a. ʕabiir$_i$ ti-ḥassib   [inn-ha$_{i/j}$ bi-ta-rooh   al-maṭaar]
       Abeer 3SG.F-believe [COMP-3SG.F.SBJ FUT-3SG.F-go the-airport]
       'Lit. Abeer believes that she will go to the airport.'

   b. ʕabiir$_i$ ti-ḥassib   [inn   nafs-aha$_{i/*j}$ bi-ta-rooh   al-maṭaar]
       Abeer 3SG.F-believe [COMP self-her   FUT-3SG.F-go the-airport]
       'Lit. Abeer believes that herself will go to the airport.'

Second, languages with CCA-constructions share a structural A-position in the edge of the embedded CP above the head C. As noted in Wurmbrand (2019) and Lohninger et al. (2022) and assuming a non-structural (featural) approach to the A/A′ distinction (following van Urk 2015), the existence of the A-position is motivated based on the A-features on the head C of the embedded clause. The connection between the presence of CCA-configurations and A-features on C (plus its typical A′ features) has been previously established in other studies (Coon et al. 2021, Aldridge 2004, 2008, 2016,

Legate 2014, van Urk 2015, Bossi & Diercks 2019, among others). The combined features on the probe C are therefore what attracts the movement of the CCA element in the embedded clause to that position. In order to provide an explanation to the long-distance anaphoric relation in Malki Arabic, it is important to see whether Malki Arabic has a probe with a combination of both A and A′ properties (or at least the former), thus providing evidence for an A-position in Spec-CP.

Besides the A′ properties of the CP-domain, I provide evidence that the embedded CP in Malki Arabic has an A-position based on complementizer agreement, depicted in conjunct agreement, showing that C is a separate probe and has A-features. As can be seen in (110a) and (110b), the complementizer can show first conjunct agreement (3SG.F) with 'Abeer', whole conjunct agreement (PL) with 'Abeer and Fahad', or no agreement at all.

(110)   Malki Arabic
   a. ħassab-t   **inna-ha**/**hum**/**Ø**   ʕabiir wa fahad bi-saafr-uun
       believed.PST-1SG COMP-**3SG.F**/**3.PL**/Ø Abeer and Fahad FUT-travel-3.PL
       turkayya.
       Turkey
       'I believed that Abeer and Fahad will travel to Turkey.'

   b. guul-t   **inna-ha**/**hum**/**Ø**   ʕabiir wa fahad bi-saafr-uun
       said.PST-1SG   COMP-**3SG.F**/**3.PL**/Ø Abeer and Fahad FUT-travel-3.PL
       turkayya.
       Turkey
       'I said that Abeer and Fahad will travel to Turkey.'

Complementizer agreement in MSA behaves differently from Malki Arabic. In MSA, the complementizer cannot show an agreement in ɸ-features with the embedded subject (111b) when the embedded clause is of the SVO type, unlike Malki Arabic where the

complementizer can show agreement with the embedded preverbal subject.

(111)   MSA

    a. qaala   maaʒid-uu ʔanna-Ø χuluud-aa   ðahab-at   maʕa ʔab-i-aha.
       say.3SG.M Majid-NOM COMP   Kholoud-ACC went-3SG.F with father-GEN-her
       'Majid said that Kholoud went with her father.'

    b. *qaala   maaʒid-uu ʔanna-ha   χuluud-aa   ðahab-at   maʕa
       say.3SG.M Majid-NOM COMP-3SG.F Kholoud-ACC went-3SG.F with
       ʔab-i-aha.
       father-GEN-her
       'Majid said that Kholoud went with her father.'

Similar observations of ɸ-agreement on complementizers in Malki Arabic also hold in Sason Arabic (Akkuş 2021) and Jordanian Arabic (Jarrah 2019). As can be seen in (112) and (113), the complementizer and the head T can show different realizations of agreement and not necessarily the same agreement on both heads. The complementizer can exhibit a ɸ-agreement with the first element of the conjunct (first conjunct agreement) while T shows an agreement with the whole conjunct. Similar to Malki Arabic, the complementizer can optionally show a morphological agreement in ɸ-features with the whole conjunct, only with the first conjunct, or the lack thereof.

(112)   Sason Arabic (Akkuş 2021:5)

    qul-tu   {le   / le-na   / le-nen} [bɪnt-ma u   sabi-ma] ayal-o
    said-1SG {that.Ø / that-3SG.F / that-3PL} [girl.F-a and boy-a]PL ate-3PL
    anzarut.
    corn
    'I said that a girl and a boy ate the corn.'

(113)   Jordanian Arabic (Jarrah 2019:8)

    ʔabuu-j   fakkar   ʔinn-ha   ʔil-binit ʔaχazˤ-t ʔil-mafatiiħ.
    father-my believed.3SG.M COMP-3SG.F DEF-girl.F took-3SG.F DEF-keys.M
    'My father believed that the girl took the keys.'

These examples above show that the head C can exhibit an A-property, namely

φ-features. I conclude that the head C is a legitimate composite probe and thus signals the presence of an A-position in Spec-CP. As we will see, the lack of complementizer agreement provides an explanation of the difference between Malki Arabic and MSA with respect to the availability of long-distance subject anaphors in former but not the latter.

Another piece of evidence for considering the long-distance bound subject anaphor as a CCA-configuration in Malki Arabic is that subject anaphors in embedded clauses are not attested with all verbs, rather mostly with knowledge and belief verbs. This follows from the property of CCA-configurations: they are restricted by the selectional properties of the matrix verbs. The long-distance binding is only possible, for instance, with verbs like *ʃauuf* 'see' (114a) and *ħauus* 'feel' (114b), for example, but not with the speech verb *quul* 'say' (114c).[18]
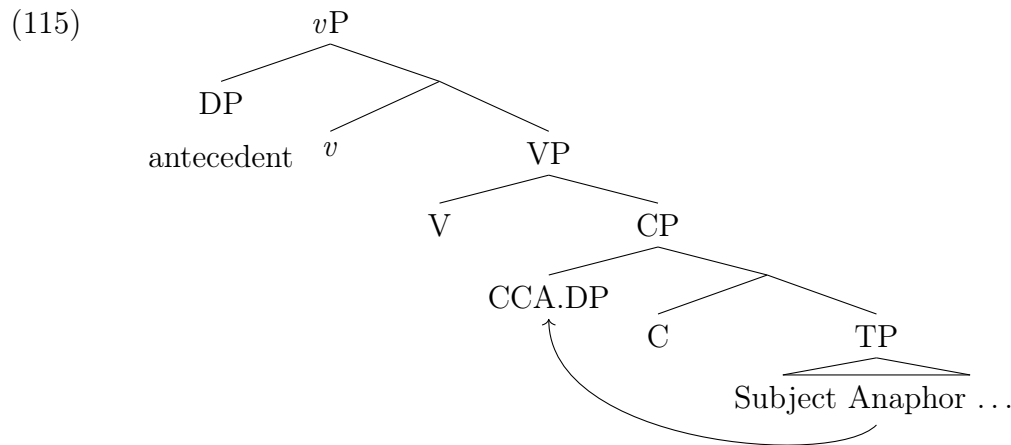
(114)   Malki Arabic

    a. ʕabiir$_i$ ta-**ʃauuf** [inn   nafs-aha$_{i/*j}$ bi-ta-drus   al-ħaasib   ].
        Abeer 3SG.F-see [COMP self-her   FUT-3SG.F-study the-computer science]
        'Lit. Abeer sees that herself will study computer science.'

    b. ʕabiir$_i$ ta-**ħauus** [inn   nafs-aha$_{i/*j}$ bi-ta-fuuz   bi-l-ʒaaʔizah].
        Abeer 3SG.F-feel [COMP self-her   FUT-3SG.F-win PREP-the-prize]
        'Lit. Abeer feels that herself will win the prize.'

    c. *ʕabiir$_i$ ta-**quul** [inn   nafs-aha$_{i/*j}$ bi-ta-drus   al-ħaasib   ].
        Abeer 3SG.F-say [COMP self-her   FUT-3SG.F-study the-computer science]
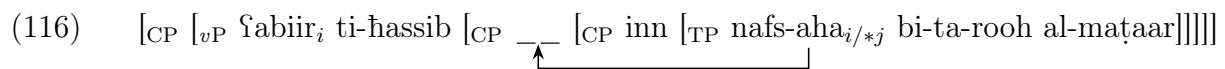   'Lit. Abeer says that herself will study computer science.'

The analysis I suggest for the long-distance bound subject anaphors is schematized in (115). Omitted for space reasons, the anaphor is initially merged in the lower clause,

---

[18]The sentence with the verb 'say' is only possible with the interpretation in which the reflexive form denotes a great desire of the soul for something. This is a result of the form *nafs* which has multiple interpretations with different syntactic categories. In the case where the sentence in (114c) can be possible, the reflexive can be a verb.

presumably in Spec-vP, and then moves to Spec-TP for EPP feature reasons. Because the head C is assumed to carry a composite probe, the anaphor is triggered to move to CCA.DP, where it becomes accessible and situated within the domain of the antecedent in the edge of the matrix phasal head v, as shown in the diagram below.

(115)

```
                    vP
              ╱        ╲
          DP              ╲
       antecedent   v        VP
                         ╱        ╲
                       V            CP
                            ╱            ╲
                       CCA.DP    C           TP
                           ↑          ╱        ╲
                            ╲    Subject Anaphor . . .
                             ╲_____╱
```

Thus, the subject anaphor comes to be locally bound and in the same phasal binding domain, here the phase vP in the upper clause. According to this proposal, sentence like (108) is schematized as follows:

(116)   [$_{CP}$ [$_{vP}$ ʕabiir$_i$ ti-ħassib [$_{CP}$ __ [$_{CP}$ inn [$_{TP}$ nafs-aha$_{i/*j}$ bi-ta-rooh al-maṭaar]]]]]

The proposal in (115) thus accounts for how the subject anaphor can be in the same phasal binding domain of its antecedent. This proposal however raises several questions about the motivation for such movement because it seems that the movement of the subject anaphor to the position above C is only for binding, casting doubt on its stipulative nature. Another concern this proposal bears is the difference in word order. If the subject anaphor moves above an overt complementizer, then sentences like (117) are predicted to be grammatical in Malki Arabic, in contrast to their actual ungrammaticality status.

(117)   Malki Arabic
        *χuluud$_i$ ti-ħassib **nafs-aha**$_{i/*j}$ [inn   *t* ta-laʕab   bi-sayyarat maaʒid].
        Kholoud 3SG.F-believe **self-her**   [COMP *t* 3SG.F-play PREP-car   Majid   ]
        'Lit. Kholoud$_i$ believes herself$_{i/*j}$ that plays with Majid's car.'

A further point that this proposal has to deal with is the lack of long-distance

subject anaphors in other Arabic varieties, like MSA. As suggested in Wurmbrand (2019)

and subsequent works of Lohninger et al. (2022), the lack of CCA-constructions is due to

whether a language has a composite probe (or A-features at least) on the head C. As

indicated by the data on complementizer agreement in (111b) and with the verb 'believe'

in (118), no ɸ-agreement on C holds and this predicts why long-distance anaphors are not

present in MSA. Although the lack of ɸ-agreement on C might explain why MSA does

not allow subject anaphors in the embedded subject position, Hyper-ECM seems to hold

across all verbs selecting a CP complement. This leads us to the question of why MSA

allows accusative case assignment on the embedded subject but not subject anaphors, a

question to be tackled in Section 4.3.2.

(118)   MSA
        *ð̣anna   maaʒid-uu [ʔanna-ha   χuluud-aa   ðahab-at   maʕa
        believed.3SG.M Majid-NOM [COMP-3SG.F Kholoud-ACC went-3SG.F with
        ʔab-i-aha   ].
        father-GEN-her]
        'Majid believed that Kholoud went with her father.'

As convincingly argued in Lohninger et al. (2022) and based on cross-linguistic data

from other languages, not only Malki Arabic, the movement to the edge of the embedded

CP is far from being stipulative but rather a necessary part of syntax to explain

dependencies spanning across full clause boundaries. Besides dependent elements like

subject anaphors in Malki Arabic, other syntactic phenomena like agreement and Case

seem to require a syntactic operation like movement to follow some conditions of locality.

This is the same in the case of long-distance binding in Malki Arabic. The subject

anaphor follows the locality condition of being contained in the same locality domain

(phase) of its antecedent. In order for it to obey the locality requirement, a movement to

an escape hatch where both the anaphor and its antecedent become in the same domain

is warranted. Similarly, when a language shows an agreement between a matrix verb and

an embedded nominal, that nominal should be in the closest local domain of the verb; the

nominal needs to move to a position local to the verb in the matrix clause, similar to

agreement in Zulu (119) and accusative case assignment in Korean (120) via Hyper-RtO.

(119)   Zulu (Halpert & Zeller 2015:476)
ngi-ya-m-funa **uSipho**   [(ukuthi) apheke   iqanda   ].
1SG-**1.O**-want **AUG.1.Sipho** [(that)   1.SBJ.cook **AUG**.5.egg]
'I want Sipho to cook an egg.'

(120)   Korean (Yoon 2007:647)
Na-nun **yeki-pwuthe-lul**$_i$ [t_i nay ttang-ila-ko   ] mitnunta.
I-TOP   **here-from-ACC**$_i$ [t_i my land-COP-COM] believe
'I believe my land begins from here.'

The predicted word order is also not problematic to the current proposal. Assuming

a copy theory of movement (Chomsky 1993), the surface phonological realization of the

anaphor in the embedded subject is a result of the pronunciation of the left (lower) copy

of the moved anaphor, a process carried out in the PF interface. The higher copy carries

out the anaphoric interpretation in the semantics interface. Thus, the subject anaphor

undergoes a covert hyperraising to the position in the embedded CP edge. That

CCA-constructions can result from covert cross-clausal operation is not new. As

demonstrated in Deal (2017), languages like Nez Perce show a configuration in which the

matrix predicate exhibits agreement with the embedded subject 'children' in the

embedded clause, as can be seen in (121).

(121)   Nez Perce (Deal 2017:6)
    ′Aayat-onm hi-**nees**-nek-se   [watiisx   **mamay′ac**
    woman-ERG 3.SBJ-**O.PL**-think-IPFV [1.day.away **children.NOM**
    hi-pa-paay-no′   ]
    3.SBJ-S.PL-arrive-FUT]
    'The woman thinks the children will arrive tomorrow.'

An emerging conceptual point the proposal in this thesis bears on the hypothesis of

taking spell-out domains as the Binding Condition A domain is the spell-out timing of a

phase complement containing the anaphor and its antecedent (Charnavel & Sportiche

2016). Some suggestions have been brought forward regarding when phase complements

become opaque to movement and agreement. As we have seen in chapter 2, the spell-out

domain of a phase is triggered either by the completeness of the phase, per PIC_1 (31),

repeated as (122) below, or by the merge of the next higher phase head, following PIC_2 in

(32), repeated below as (123).

(122)   **Phase Impenetrability Condition (PIC$_1$)** (Chomsky 2000:108):

    In a phase α with head H, the domain of H is not accessible to operations

    outside α, only H and its edge are accessible to such operations.

(123)   **Phase Impenetrability Condition (PIC$_2$)** (Chomsky 2001:14):

    Given the structure [$_{ZP}$ Z ... [$_{HP}$ α [$_{H'}$ H YP]]], where H and Z are phase heads,

    the domain of H is not accessible to operations at ZP; only H and its edge are

    accessible to such operations.

Based on the definition of spell-out in PIC_1, the phasal analysis of binding domains in

Charnavel & Sportiche (2016) cannot account for the long-distance binding relation, even along the current proposal. As illustrated in (124), assuming an initial movement of the subject anaphor from Spec-vP to Spec-TP for EPP, the subject anaphor then moves to the A-position in Spec-CP (124b) (spell-out domains are boldfaced). The movement to Spec-CP is permitted as far as the TP does not undergo spell-out, given that the CP phase is still not completed, per PIC_1. After the internal merge of the subject anaphor to Spec-CP, the embedded TP is spelt-out and the VP merges, followed by the external merge of the phase head v, as illustrated in (124c). Assuming that agentive arguments merge in Spec-vP, the matrix subject merges in that position, yielding the spell-out of VP (the complement of the phase vP). As shown in (124d), therefore, the subject anaphor in the edge of the embedded CP phase is not contained in the same spell-out domain of the antecedent, contra the prediction of Charnavel & Sportiche (2016).

(124)   **Spell-out domains as binding domains by PIC_1**

   a. [$_{CP}$ C [$_{TP}$ Subject Anaphor T …]]

   b. [$_{CP}$ Subject Anaphor C [$_{TP}$ **<Subject Anaphor> T …**]]

   c. [$_{vP}$ v [$_{VP}$ V [$_{CP}$ Subject Anaphor C [$_{TP}$ **<Subject Anaphor> T …**]]]]

   d. [$_{vP}$ antecedent v [$_{VP}$ **V** [$_{CP}$ **Subject Anaphor C**[$_{TP}$ **<Subject Anaphor>**
      T …]]]]

   Similarly, with PIC_2, the spell-out domains do not seem to be viable binding domains for the long-distance subject anaphors. The diagram in (125) shows that even if the spell-out is triggered by the merge of the highest phase head, here the matrix C, the same problem arises. The subject anaphor is in the VP complement of the phase vP and is not

contained within the same spell-out domain (TP) of the antecedent, as shown in (125e).

(125)   **Spell-out domains as binding domains by PIC$_2$**

    a. $[_{CP}$ C $[_{TP}$ Subject Anaphor T ...$]]$

    b. $[_{CP}$ Subject Anaphor C $[_{TP}$ <Subject Anaphor> T ...$]]$

    c. $[_{vP}$ v $[_{VP}$ V $[_{CP}$ Subject Anaphor C $[_{TP}$ <Subject Anaphor> T ... $]]]]$

    d. $[_{vP}$ antecedent v $[_{VP}$ V $[_{CP}$ Subject Anaphor C $[_{TP}$ <Subject Anaphor> T

        ... $]]]]$

    e. $[_{CP}$ C $[TP$ T $[_{vP}$ antecedent v $[_{VP}$ V $[_{CP}$ Subject Anaphor C $[_{TP}$ <Subject

        Anaphor> T ... $]]]$

Even if the hypothesis that spell-out is triggered by the valuation of the unvalued

ϕ-features on phasal heads (Chomsky 2008) is adopted, the subject anaphor would be

trapped in the embedded CP clause at the time the ϕ-features on the phasal head v are

valued. Hence, the spell-out domain, here VP, would contain only the subject anaphor

but not the antecedent.

What the above diagrams show is that the binding domains should rather be whole

phases, as suggested in Lee-Schoenfeld (2008) and Quicoli (2008). In (124) and (125), the

subject anaphor and its antecedent with both versions of PIC are in the same phase, the

matrix vP, but not the spell-out domain. The whole-phase approach to binding domains

is therefore superior to the spell-out domains in accounting for the long-distance binding

relation in Malki Arabic with the suggested CCA analysis.

4.3   DISCUSSION

One advantage of the CCA analysis of long-distance subject anaphors presented in this thesis is that the explanation of long-distance subject anaphors can be derived from basic syntactic operations, like internal merge, and based on the internal structure of the embedded CP, the complement of matrix verbs. Moreover, the current proposal is in line with the view of reducing long-distance anaphors to local binding, similar to Charnavel's works, among others.

Furthermore, the current analysis supports the generalization of Charnavel (2019) in (46), repeated as (126) below. The movement of the anaphor to a local domain (the phase vP) is an escape from the exemption from Binding Condition A. Thus, based on at least a phasal-based definition of exempt anaphors this analysis predicts that subject anaphors are plain anaphors since they are structurally in the same binding domain. Treating long-distance subject anaphors as a CCA-configuration is in support of the approaches of reducing long-distance binding to local binding, like the movement approaches and the analysis of Charnavel (2019).

(126)   **Agreement-based condition on (exempt) anaphors** (Charnavel 2019:76)

An anaphor cannot be exempt if it occurs in a syntactic position construed with agreement (e.g. nominative position in languages with subject agreement).

However, a point of contention to Charnavel's account stems from its conceptual aspect. Among several analyses of long-distance binding is the hypothesis of reducing long-distance anaphors to local ones via a logophoric center projection headed by a Log_{Op} operator in spell-out domains (Charnavel 2019, *et seq.*). As far as embedded subject

anaphors are concerned, such a hypothesis is not necessary. First, the hypothesis is conceptually untenable, as it comes in opposition to a minimalist view on language. A tenant of the minimalist framework is economy, meaning that derivations should be as economical as possible and unnecessary formatives (not part of the numeration) like null operators, traces, indices, etc, are prohibited to enter the derivation, except if they are necessary (Chomsky 1993, Hornstein 2001). Assuming a Log_{Op} operator, therefore, seems to be an unneeded construct in explaining the long-distance binding of subject anaphors across finite clauses, at least in Malki Arabic. Such cross-clausal binding can receive an explanation based on basic (and already suggested) syntactic notions like movement and complementation properties of verbs.

4.3.1 Anaphor Agreement Effect

Another advantage of this analysis is its contribution to the understanding of the anaphor agreement effect. Although Malki Arabic shows an anaphor occurring in a subject position that shows agreement with the embedded verb, the CCA analysis shows that such configuration is not a bona fide violation case of the anaphor agreement effect. The anaphor does not stay in an agreement-triggering position at the time of spell-out. Besides the strategies of detransitivization, argument deletion, anaphoric agreement that other languages use to obey the anaphor agreement effect, I suggest that movement out of a syntactic position construed for agreement before spell-out is another strategy languages like Malki Arabic use to obey the anaphor agreement effect. A presupposition of this strategy is that the anaphor agreement effect is an interface condition, a new observation about this effect, to the best of my knowledge. However, considering the

anaphor agreement effect as an interface condition raises the question of which interface (PF or LF) imposes such an effect. Let's take the modified versions of the anaphor agreement effect of Rizzi (1990), Woolford (1999), and Tucker (2011).

(127)   Anaphor Agreement Effect: Version 1

Anaphors do not occur in syntactic positions construed with agreement. (Rizzi 1990)

(128)   Anaphor Agreement Effect: Version 2

Anaphors do not occur in syntactic positions construed with agreement, unless the agreement is anaphoric. (Woolford 1999)

(129)   Anaphor Agreement Effect: Version 3

Anaphors do not occur in syntactic positions construed with covarying ϕ-morphology. (Tucker 2011)

Because the anaphor agreement effect involves (morpho)syntactic constructs such as agreement and a morphological realization of ϕ-features, suggesting that the anaphor agreement effect is an LF condition is not the right path to follow. Such constructs then indicate that the anaphor agreement effect is better treated as a PF condition, especially because the effect includes morphological agreement (in the version of Tucker 2011). Thus, the anaphor agreement effect can be modified as in (130).

(130)   Anaphor Agreement Effect (current version)

Anaphors do not occur in syntactic positions construed with covarying morphological ϕ-agreement at the interface.

One piece of evidence for the hypothesis that the anaphor agreement effect is an

interface condition at PF is the realization of the lower copy of the embedded subject as
pronoun, as shown in (131).[19] The sentence in (131) demonstrates that the embedded
subject position, an agreement position with the embedded verb, does not include an
anaphor when realized at the PF.

(131)   Malki Arabic
        ʕabiir$_i$ ta-ʃauuf **nafs-aha**$_{i/*j}$ [inn-**aha**   bi-ta-drus
        Abeer 3SG.F-see self-her   [COMP-3SG.F FUT-3SG.F-study
        al-ħaasib   ].
        the-computer science]
        'Lit. Abeer sees herself that she will study computer science.'

What still remains is the possibility of the anaphor to occur in an agreement-triggering
position even when the derivation is transferred to the interface. I leave this for further
research, but the prospect of treating the anaphor agreement effect as an interface
condition seems to be promising. More work needs to be done however on which parts of
the effect are related to the interpretation at the LF interface. Moreover, a deeper look is
sought into the question of whether agreement is not only a PF property and whether it
can have effects on interpretation, contra conventional thinking that agreement does not
contribute to the semantics interface.

4.3.2   Other Issues: Accusative-marked subjects in MSA

A prediction of the hypothesis that the embedded subject moves to a CCA-position
above C is that the embedded subject is marked with an accusative case. In Malki
Arabic, it is difficult to argue for whether the embedded subject exhibits a nominative or
accusative case because Case is not marked overtly on nominals in Malki Arabic, like

---

[19]This is similar to the phenomenon of 'copy raising' where a raised subject leaves a pronominal copy in
its canonical position. See Asudeh & Toivonen 2012 for an overview and comprehensive background on its
theoretical aspect.

other modern Arabic dialects (Aoun et al. 2010). Case in MSA, on the other hand, is overtly realized on nouns, adjectives, adverbs and pronouns. As we have seen above, anaphors are not allowed to occur in the embedded subject position in the embedded clause of the similar verb 'believe' in MSA. An explanation suggested in Section 4.2 is that MSA has no A-position above C because C does not show ɸ-agreement with the embedded subject, a characteristic of non-CCA configurations. A problem for such an observation is the existence of a Hyper-ECM construction in MSA, despite the unavailability of long-distance binding. As shown in (132a), the embedded subject χuluud is marked with accusative case although the same verb does not allow a long-distance subject anaphor in the embedded subject position (132b).

(132)   MSA

    a. ðana   maaʒid-uu ʔanna-Ø χuluud-aa ðahab-at   maʕa
        believed.3SG.M Majid-NOM COMP   Kholoud-ACC went-3SG.F with
        ʔab-i-aha.
        father-GEN-her
        'Majid believed that Kholoud went with her father.'

    b. *ðana   maaʒid-uu ʔanna-Ø nafs-a-hu ðahaba   maʕa
        believed.3SG.M Majid-NOM COMP   self-ACC-him went.3SG.M with
        ʔab-i-ah.
        father-GEN-his
        'Majid believed that himself went with his father.'

Different views have been proposed to account for the accusative-marked embedded subject with the 'believe' predicate in MSA (Soltan 2007, Al-Balushi 2011, 2016). One view suggests that the embedded subject undergoes raising to the object position in the higher clause (Fassi Fehri 2012:249). For the complementizer C, it undergoes truncation, as supported by the presence of pronominal clitic as the embedded subject and the

absence of the complementizer, as shown in (133).

(133)   MSA (Fassi Fehri 2012:249)
          ħasib-tu-hu   daxal-a   l-qaaʕat-a
          believed-1SG-him entered-3SG.M the-hall-ACC
          'I believed he entered the hall.'

Another analysis treats the embedded subject in Spec-TP as a left peripheral position, and the accusative case is assigned by the head v in the matrix clause via the operation Agree (Soltan 2007). A similar analysis in this vein with respect to the positioning of the embedded external argument in the left periphery is of Al-Balushi (2011) where the subject in the embedded clause occupies Spec-TopP. For this account, case is not structural, but rather a lexical in nature. In contrast to Fassi Fehri (2012), the analyses of Soltan and Al-Balushi have in common that the embedded subject stays in its embedded clause and is not raised to the matrix clause. The difference between the two proposals, however, is in the type of case assigned to the subject: structural (Soltan 2007) vs. lexical (Alboiu & Hill 2016).

The RtO approach to the case of the embedded subject (Fassi Fehri 2012) is untenable because the embedded subject cannot occupy a position in the higher clause, for it is an argument of the embedded predicate and thus it can only be in the embedded clause. Otherwise, a mismatch in the thematic mapping is expected to occur once the derivational constructs are transferred to the interfaces. Furthermore, evidence from the subject placement before adverbs and intervention effects supports the non-movement of the accusative subject from the embedded clause to the matrix clause (Soltan 2007). Explaining the accusative case assignment on embedded subjects being due to the

embedded clause of the 'believe' predicate not forming a full CP is not on right track, as many pieces of evidence show the contrary: the embedded clause of the 'believe'-predicate (and other types of predicates) is a CP, as I have shown in chapter 3.

A further note that seems to be ignored in the previous accounts of the long-distance case assignment in Arabic, except Soltan (2007) and Al-Balushi (2016), is to deal with the "double case assignment", a problem also marginally addressed in Aoun et al. (2010:15, fn.5). The embedded subject can have two cases: one from T in the embedded clause and another from either the C (following a linguistic thinking of the traditional Arabic grammarians) or the matrix verb (structurally by v). One way proposed in Soltan (2007) and Al-Balushi (2011, 2016) to overcome the double-case assignment problem is to assume a null *pro* element in Spec-vP as a resumptive null pronoun A′-bound by the embedded subject in Spec-TP/TopP. The advantage of the *pro* element is to receive the nominative case assigned by T while the Case feature on the embedded subject is left unvalued until the matrix v is merged. Such an analysis, however, only applies to a TP-complement without a CP layer. In a CP-complement, assuming a *pro* as the external argument of the embedded verb would not provide the same solution as the embedded subject would still be inaccessible to the matrix v, considering a phase-theoretic concept of locality. A desirable analysis of the accusative-marked subject in an embedded clause should then account for constructions where the complementizer is present and the embedded subject is marked with an accusative case, given that the verb *ðan* can select for a TP or CP complement.[20]

---

[20]If we assume, similar to Fassi Fehri (2012), that the complement starts as a CP and ends as a TP, the truncation of a C head could only be a process at the PF interface, not syntactic.

The contradictory behavior of case and binding in MSA, unlike the situation in Malki Arabic, takes us far away from assuming the accusative-marked subject as a CCA-configuration where the subject can move to a position accessible to the higher head *v*. Besides the lack of A-features on C in MSA, following a CCA analysis makes inaccurate predictions, one of which is allowing a long-distance binding which is not possible in MSA. However, ignoring such possibility of long-distance subject anaphors, the current proposal offers an explanation of how the embedded subject in MSA is marked with accusative case.

4.4 CONCLUSION

In this chapter, I provided evidence that long-distance bound subject anaphors in Malki Arabic occur in a CCA-configuration based on the properties of optionality, selection restrictions by the matrix verb, and availability of A-features on the head C. Thus, anaphors of this type are not counter-evidence to the phasal analysis of anaphoric binding domains. Within the current proposal, further support is provided to whole phases as binding domains, rather than spell-out domains (Charnavel & Sportiche 2016). I have also shown that, at least for subject anaphors in Malki Arabic, long-distance subject anaphors can be accounted for with phasal binding domains and movement, and resort to silent operators is not needed. In this chapter, support to the hypothesis that the anaphor occurring in an agreement position does not trigger exemption from Condition A is provided, hence confirming to the agreement condition on (exempt) anaphors (Charnavel 2019). I discussed the possibility of the anaphor agreement effect as a condition imposed by the external systems (LF and PF) and the idea of whether the

current proposal can extend to the accusative subject in embedded clauses in MSA.

CHAPTER 5

CONCLUSION

The goal of this thesis was to show that long-distance binding with subject anaphors spanning across finite clauses does not bear evidence against the reduction of binding domains to phase theory. This thesis is in line with previous analyses which aim at reducing long-distance binding to local antecedence relations. In pursuit of this endeavor, based on data from Malki Arabic, I have shown that long-distance subject anaphors constitute a binding relation similar to a CCA-configuration, like hyperraising and long-distance agreement. The similarity lies in the movement of the syntactic item having a dependency with another element in another clause to a position at the edge of the embedded clause, following Wurmbrand (2019) and Lohninger et al. (2022). In line of this hypothesis, I argued that long-distance binding in Malki Arabic is a CCA-construction because the former shows the same properties of the latter. Like CCA-constructions, I have shown that constructions with long-distance bound subject anaphors are optional, restricted by the selection properties of matrix verbs, and include a C head that can host a probe with A and A′ features. Assuming a featural perspective on the A/A′ distinction, I have suggested that the edge of the embedded clause hosts a position with A-properties, given evidence from complementizer agreement which indicates that the phasal head C can be a composite probe for hosting ɸ-features, in addition to the regular A′-features.

Furthermore, the analysis proposed in this thesis indicates that reducing binding domains to spell-out domains is unwanted and thus taking whole phases as binding

domains is theoretically desirable. Spell-out binding domains require a containment of both the anaphor and the antecedent. In the context of the current proposal, I pointed out that containing both the long-distance subject anaphor and its antecedent in one spell-out domains is not possible, despite all the possible proposals on when phasal complements are sent to the interface. I conclude therefore that binding domains are better reduced to whole phases and not spell-out domains. Another point this thesis shows is that the current proposal also supports the definition of exemption proposed in Charnavel (2019).

Moreover, in this thesis, I have shown that the properties of the complement clause of the attitude predicate *ðana/ħassib* match the properties of a CP clause and not only a TP, contra arguments taking the complement clause as only a TP (Soltan 2007). Based on the cross-linguistic (morpho)syntactic and semantic properties of CP clauses, I have shown the complement clause of the attitude verb can be introduced by an overt complementizer, requires an embedded subject, can host negation markers, allows voice mismatching, and has a similar instance of the adverb in the main clause. The complement clause is also shown to be opaque to focus movement, have simultaneous or independent tense interpretations from the temporal interpretation of the matrix clause, be evaluated for truth, and allow shifted indexicals. Based on some of these CP properties, it appears that the complement of the attitude verb can also be a TP with the absence of the complementizer. Important to the proposal in this thesis is that I have shown that the binding relation is long-distance and spans across a finite clause boundary because the embedded clause which includes the subject anaphor and which is introduced

by an overt complementizer is a CP clause and not a TP/ECM.

Following the proposal in this thesis, I have shown the prospect of treating the anaphor agreement effect as an interface condition. This suggestion is supported by the movement of the subject anaphor from its canonical position of agreement before spell-out takes place. This thesis also highlighted a promising account of the accusative-marked embedded subject in MSA, following the proposal suggested in this thesis.

REFERENCES

Abels, Klaus. 2003. *Successive cyclicity, antilocality and adposition stranding.* Storrs, CT: University of Connecticut, Storrs PhD dissertation.

Akkuş, Faruk. 2021. On the locus of A′-agreement. In *Move and Agree Forum,* The University of British Columbia and McGill University.

Al-Balushi, Rashid. 2016. The 'believe'-construction in Standard Arabic. *Journal of Linguistics* 52(1). 1–36.

Al-Balushi, Rashid Ali. 2011. *Case in Standard Arabic: The untraveled paths.* Toronto, Ontario: University of Toronto Doctoral dissertation.

Albaty, Yasser. 2019. *Modality, control and restructuring in Arabic.* Milwaukee, WI: University of Wisconsin–Milwaukee Doctoral dissertation.

Albaty, Yasser & Hamid Ouali. 2018. Restructuring and control in Arabic. *Brill's Journal of Afroasiatic Languages and Linguistics* 10(2). 169 – 199.

Alboiu, Gabriela & Virginia Hill. 2016. Evidentiality and raising to object as A′-movement: A Romanian case study. *Syntax* 19(3). 256–285.

Aldridge, Edith. 2004. *Ergativity and word order in Austronesian languages.* Ithaca, NY: Cornell University PhD dissertation.

Aldridge, Edith. 2008. Generative approaches to ergativity. *Language and Linguistics Compass* 2(5). 966–995.

Aldridge, Edith. 2016. ϕ-feature competition: A unified approach to the Austronesian extraction restriction. In *Proceedings of the 52nd meeting of the Chicago linguistic*

society (CLS), vol. 52, 21–23. Chicago, IL.

Anagnostopoulou, Elena & Martin Everaert. 1999. Toward a more complete typology of anaphoric expressions. *Linguistic Inquiry* 30(1). 97–119.

Anand, Pranav. 2006. *De de se*. Cambridge, MA: Massachusetts Institute of Technology PhD dissertation.

Anand, Pranav & Andrew Nevins. 2004. Shifty operators in changing contexts. In Robert B. Young (ed.), *Proceedings of semantics and linguistic theory (SALT14) XIV*, 20–37. Ithaca, NY: Cornell University, CLC Publications.

Aoun, Joseph, Elabbas Benmamoun & Dominique Sportiche. 1994. Agreement, word order, and conjunction in some varieties of Arabic. *Linguistic Inquiry* 25(2). 195–220.

Aoun, Joseph E, Elabbas Benmamoun & Lina Choueiri. 2010. *The syntax of Arabic*. Cambridge: Cambridge University Press.

Asudeh, Ash & Ida Toivonen. 2012. Copy raising and perception. *Natural Language & Linguistic Theory* 30(2). 321–380.

Benmamoun, Elabbas. 2000. *The feature structure of functional categories: A comparative study of Arabic dialects*. Oxford University Press.

Bobaljik, Jonathan David & Höskuldur Thráinsson. 1998. Two heads aren't always better than one. *Syntax* 1(1). 37–71.

Bok-Bennema, Reineke. 1991. *Case and Agreement in Inuit*. Dordrecht: De Gruyter Mouton.

Bondarenko, Tatiana. 2017. ECM in Buryat and the optionality of movement. In *The 12th workshop on altaic formal linguistics (WAFL 12)* (MIT Working Papers in

Linguistics 83), 31–42. MIT.

Bošković, Željko. 2013. Phases beyond clauses. In Lilia Schürcks, Anastasia Giannakidou & Urtzi Etxeberria (eds.), *The nominal structure in Slavic and beyond*, 75–128. Berlin: De Gruyter Mouton.

Bossi, Madeline & Michael Diercks. 2019. V1 in Kipsigis: Head movement and discourse-based scrambling. *Glossa: A Journal of General Linguistics* 4(1).

Cantrall, WR. 1974. Viewpoint, reflexives, and the nature of noun phrases. *Foundations of Language* 14(4).

Carstens, Vicki & Michael Diercks. 2013. Parameterizing case and activity: Hyper-raising in Bantu. In Claire Moore-Cantwell Seda Kan & Robert Staubs (eds.), *Proceedings of the 40th annual meeting of the north east linguistic society (NELS40)*, 99–118. Amherst, MA: University of Massachusetts, Graduate Linguistic Student Association.

Charnavel, Isabelle. 2019. *Locality and logophoricity: A theory of exempt anaphora.* New York, NY: Oxford University Press.

Charnavel, Isabelle. 2020. Logophoricity and locality: A view from French anaphors. *Linguistic Inquiry* 51(4). 671–723.

Charnavel, Isabelle & Dominique Sportiche. 2016. Anaphor binding: What French inanimate anaphors show. *Linguistic Inquiry* 47(1). 35–87.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson & Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, 232–286. New York, NY: Academic Press.

Chomsky, Noam. 1980. On binding. *Linguistic Inquiry* 11(1). 1–46.

Chomsky, Noam. 1981. *Lectures on government and binding: The Pisa lectures.* De Gruyter Mouton.

Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use.* New York: Praeger.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In K Hale & S Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 1–52. MIT press.

Chomsky, Noam. 1995. *The minimalist program.* Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. MA: Cambridge: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. The MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and beyond: The cartography of syntactic structures* (Oxford Studies in Comparative Syntax 3), 104–131. Oxford: Oxford University Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + recursion = language?: Chomsky's minimalism and the view from syntax-semantics*, 1–30. De Gruyter Mouton.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory*, 132–166. The MIT Press.

Chomsky, Noam & Howard Lasnik. 1993. The theory of principles and parameters. *Syntax: An International Handbook of Contemporary Research* 1. 506–569.

Chung, Sandra. 2004. Restructuring and verb-initial order in Chamorro. *Syntax* 7(3). 199–233.

Cinque, Guglielmo. 2006. *Restructuring and functional heads: The cartography of syntactic structures* (Oxford Studies in Comparative Syntax 4). Oxford: Oxford University Press.

Citko, Barbara. 2014. *Phase theory: An introduction.* Cambridge University Press.

Cole, Peter, Gabriella Hermon & C.-T. James Huang. 2006. Long-distance binding in Asian languages. In *The blackwell companion to syntax*, 21–84. John Wiley & Sons, Ltd.

Cole, Peter, Gabriella Hermon & Li-May Sung. 1990. Principles and parameters of long-distance reflexives. *Linguistic Inquiry* 21(1). 1–22.

Cole, Peter & Li-May Sung. 1994. Head movement and long-distance reflexives. *Linguistic Inquiry* 25(3). 355–406.

Cole, Peter & Chengchi Wang. 1996. Antecedents and blockers of long-distance reflexives: The case of Chinese *Ziji. Linguistic Inquiry* 27(3). 357–390.

Coon, Jessica, Nico Baier & Theodore Levin. 2021. Mayan agent focus and the ergative extraction constraint: Facts and fictions revisited. *Language* 97(2). 269–332.

Deal, Amy Rose. 2017. Shifty asymmetries: Universals and variation in shifty indexicality. Ms., University of California, Berkeley.

Despić, Miloje. 2015. Phases, reflexives, and definiteness. *Syntax* 18(3). 201–234.

den Dikken, Marcel. 2006. *Relators and Linkers: The Syntax of Predication, Predicate Inversion, and Copulas.* The MIT Press.

den Dikken, Marcel. 2007. Phase Extension Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33(1). 1–41.

den Dikken, Marcel. 2017. Predication in the syntax of hyperraising and copy raising. *Acta Linguistica Academica Acta Linguistica Hungarica* 64(1). 3 – 43.

Drummond, Alex, Norbert Hornstein & Howard Lasnik. 2010. A puzzle about P-stranding and a possible solution. *Linguistic Inquiry* 41(4). 689–692.

Drummond, Alex, Dave Kush & Norbert Hornstein. 2011. Minimalist construal: Two approaches to A and B. In Cedric Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism*, 396–426. Oxford University Press.

Farghal, Tariq Mohammed. 2020. *Hyperraising in Jordanian Arabic.* Lansing, MI: Michigan State University MA thesis.

Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words.* Dordrecht; London: Kluwer Academic.

Fassi Fehri, Abdelkader. 2012. *Key features and parameters in Arabic grammar.* Amsterdam: John Benjamins.

Fong, Suzana. 2019. Proper movement through Spec-CP: An argument from hyperraising in Mongolian. *Glossa: A Journal of General Linguistics* 4(1).

Gallego, Ángel J. 2010. *Phase Theory.* Amsterdam: John Benjamins.

Giorgi, Alessandra. 2006. From temporal anchoring to long distance anaphors. *Natural Language & Linguistic Theory* 24(4). 1009–1047.

Halpert, Claire & Jochen Zeller. 2015. Right dislocation and raising-to-object in Zulu. *The Linguistic Review* 32(3). 475–513.

Hicks, Glyn. 2009. *The derivation of anaphoric relations.* Amsterdam: John Benjamins.

Hiraiwa, Ken. 2005. *Dimensions of symmetry in syntax: Agreement and clausal architecture*: Massachusetts Institute of Technology Doctoral Dissertation.

Hornstein, Norbert. 2001. *Move! A minimalist theory of construal.* Malden, MA: Blackwell.

Huang, C.-T. James & C.-S. Luther Liu. 2001. Logophoricity, attitudes, and *ziji* at the interface. In Gabriella Hermon Peter Cole & C.-T. James Huang (eds.), *Syntax and Semantics 33: Long-distance reflexives*, 141–195. New York: Academic Press.

Huang, C.-T. James & C.-C. Jane Tang. 1991. The local nature of the long-distance reflexive in Chinese. In Jan Koster & Eric J. Reuland (eds.), *Long-distance anaphora*, 263–282. Cambridge, England; New York: Cambridge University Press.

Jarrah, Marwan. 2017. A criterial freezing approach to subject extraction in Jordanian Arabic. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 62(3). 411–448.

Jarrah, Marwan. 2019. Complementizer agreement and the T^0–ϕ Parameter in Jordanian Arabic. *Studia Linguistica* 74(1). 139–164.

Kayne, Richard. 1999. Prepositional complementizers as attractors. *Probus* 11(1). 39–74.

Kayne, Richard. 2004. Prepositions as probes. In Adriana Belletti (ed.), *Structures and beyond: The cartography of syntactic structures* (Oxford Studies in Comparative Syntax 3), 192—212. Oxford: Oxford University Press.

Koster, Jan & Eric J. Reuland (eds.). 1991. *Long-distance anaphora*. Cambridge: Cambridge University Press.

Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40(2). 187–237.

Kuno, Susumu. 1972. Functional sentence perspective: A case study from Japanese and English. *Linguistic inquiry* 3(3). 269–320.

Kuno, Susumu. 1975. Three perspectives in the functional approach to syntax. In *Proceedings of the Chicago Linguistic Society (CLS)*, 276–336. Chicago: University of Chicago: Chicago Linguistic Society.

Lebeaux, David. 1983. A distributional difference between reciprocals and reflexives. *Linguistic Inquiry* 14(4). 723–730.

Lebeaux, David. 1984. Locality and anaphoric binding. *The Linguistic Review* 4. 343–363.

Lee-Schoenfeld, Vera. 2008. Binding, phases, and locality. *Syntax* 11(3). 281–298.

Legate, Julie Anne. 2014. *Voice and v: Lessons from Acehnese* (Linguistic Inquiry Monographs). Cambridge, MA: MIT Press.

Lohninger, Magdalena. 2021. Hyper hyper! On cross-clausal A-dependencies and composite probes. In *Move and Agree forum*, The University of British Columbia and McGill University.

Lohninger, Magdalena, Iva Kovač & Susanne Wurmbrand. 2022. From prolepsis to hyperraising. *Philosophies* 7(2).

Maling, Joan. 1984. Non-clause bounded reflexives in Icelandic. *Linguistics and Philosophy* 7. 211–241.

Marantz, Alec. 2007. Phases and words. In S.H. Choe (ed.), *Phases in the theory of grammar*, 191–222. Seoul: Dong-In Publishing Co.

Matushansky, Ora. 2005. Going through a phase. In Martha McGinnis & Norvin Richards (eds.), *Perspectives on Phases* (MIT Working Papers in Linguistics 49), 157–181. Cambridge, MA: Department of Linguistics, MIT.

Messick, Troy. 2016. *The morphosyntax of self-ascription: A cross-linguistic study*. Storrs, CT: University of Connecticut Doctoral dissertation.

Mohammad, Mohammad A. 2000. *Word order, agreement and pronominalization in Standard and Palestinian Arabic* (Current Issues in Linguistic Theory 181). Amsterdam: John Benjamins.

Müller, Gereon. 2004. Phrase impenetrability and wh-intervention. In Arthur Stepanov, Gisbert Fanselow & Ralf Vogel (eds.), *Minimality effects in syntax* (Studies in Generative Grammar 70), 289–326. Berlin: De Gruyter Mouton.

Ott, Dennis. 2008. Notes on noun ph(r)ases. Ms. Hravard University.

Ouhalla, Jamal. 1994. Focus in Standard Arabic. *Linguistics in Potsdam* 1. 65–92.

Pica, Pierre. 1987. On the nature of the reflexivization cycle. In *Proceedings of the north east linguistic society (NELS 17)* 2, 483–500.

Polinsky, Maria & Eric Potsdam. 2001. Long-distance agreement and topic in Tsez. *Natural Language & Linguistic Theory* 19(3). 583–646.

Pollard, Carl & Ivan A. Sag. 1992. Anaphors in English and the scope of binding theory. *Linguistic Inquiry* 23(2). 261–303.

Postal, Paul. 1970. On so-called "pronouns" in English. In David Reibel & Sanford Schane

(eds.), *Modern studies in English*, 201–224. Englewood Cliffs, NJ: Prentice Hall.

Quicoli, A. Carlos. 2008. Anaphora by phase. *Syntax* 11(3). 299–329.

Reinhart, Tanya & Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24(4). 657–720.

Reuland, Eric. 2001. Primitives of binding. *Linguistic Inquiry* 32(3). 439–492.

Reuland, Eric. 2005. Agreeing to bind. In Hans Broekhuis, Norbert Corver, Riny
Huijbregts, Ursula Kleinhenz & Jan Koster (eds.), *Organizing grammar: Studies in
honor of Henk van Riemsdijk* (Studies in Generative Grammar), 505–513. De Gruyter
Mouton.

Reuland, Eric. 2011. *Anaphora and language design.* MIT press.

Rizzi, Luigi. 1990. On the anaphor agreement effect. *Rivista di Linguistica* 2. 27–42.

Rizzi, Luigi. 2001. On the position "Int(errogative)" in the left periphery of the clause. In
Guglielmo Cinque & Giampaolo Salvi (eds.), *Current Studies in Italian Syntax* (North
Holland Linguistic Series: Linguistic Variations 59), 287–296. Brill.

Rooryck, Johan & Guido J. Vanden Wyngaerd. 2011. *Dissolving binding theory.* Oxford;
New York: Oxford University Press.

Ross, John. 1970. On declarative sentences. In Roderick A. Jacobs & Peter S. Rosenbaum
(eds.), *Readings in English transformational grammar*, 222–277. Ginn and Company.

Shklovsky, Kirill & Yasutada Sudo. 2014. The syntax of monsters. *Linguistic Inquiry*
45(3). 381–402.

Soltan, Usama. 2006. Standard Arabic subject-verb agreement asymmetry revisited in an
Agree-based minimalist syntax. In Cedric Boeckx (ed.), *Agreement systems*, vol. 92,
239–265. Amsterdam: John Benjamins.

Soltan, Usama. 2007. *On formal feature licensing in minimalism: Aspects of Standard Arabic morphosyntax.* College Park, MD: University of Maryland Doctoral Dissertation.

Sudo, Yasutada. 2012. *On the semantics of phi features on pronouns.* Cambridge, MA: Massachusetts Institute of Technology PhD dissertation.

Sundaresan, Sandhya. 2012. *Context and (co)reference in the syntax and its interfaces.* Tromsø: University of Tromsø/University of Stuttgart Doctoral dissertation.

Sundaresan, Sandhya. 2018. An alternative model of indexical shift: Variation and selection without context-overwriting. Ms., University of Leipzig.

Svenonius, Peter. 2004. On the edge. In David Adger, Cécile De Cat & George Tsoulas (eds.), *Peripheries: Syntactic edges and their effects* (Studies in Natural Language and Linguistic Theory 59), 259–287. Dordrecht: Springer.

Thráinsson, Höskuldur. 1976. Reflexives and subjunctives in Icelandic. In *Proceedings of the sixth annual meeting of the north eastern linguistics society (NELS6)*, 225–239.

Tucker, Matthew A. 2011. Even more on the anaphor agreement effect: When binding does not agree. Ms, University of California at Santa Cruz.

Ura, Hiroyuki. 1994. Hyperraising and the theory of *pro*. In C. Giordano & D. Ardron (eds.), *Proceedings of the 6th student conference in linguistics (SCIL 6)* (MIT Working Papers in Linguistics 23), 297–316. Cambridge, MA: MIT.

van Urk, Coppe. 2015. *A uniform syntax for phrasal movement: A case study of Dinka Bor.* Cambridge, MA: Massachusetts Institute of Technology doctoral dissertation.

Vitale, Anthony J. 1981. *Swahili syntax.* De Gruyter Mouton.

Woolford, Ellen. 1999. More on the anaphor agreement effect. *Linguistic Inquiry* 30(2). 257–287.

Wu, Hsiao-hung Iris. 2013. Restructuring and clause structure in Isbukun Bunun. *Oceanic Linguistics* 52(1). 69–85.

Wurmbrand, Susanne. 2016. Complex predicate formation via voice incorporation. In Léa Nash & Pollet Samvelian (eds.), *Approaches to complex predicates* (Syntax and Semantics 41), 248–290. Leiden: Brill.

Wurmbrand, Susanne & Magdalena Lohninger. In Press. An implicational universal in complementation–theoretical insights and empirical progress. In Jutta M. Hartmann & Angelika Wöllstein (eds.), *Propositional arguments in cross-linguistic research: Theoretical and empirical issues*, Berlin: De Gruyter Mouton.

Wurmbrand, Susi. 2019. Cross-clausal A-dependencies. In Laura Stigliano Eszter Ronai & Yenan Sun (eds.), *Proceedings of the fifty-fourth annual meeting of the Chicago linguistic society (CLS54)*, 585–604. Chicago, IL: University of Chicago, Chicago Linguistic Society.

Wurmbrand, Susi & Koji Shimamura. 2017. *The features of the voice domain: Actives, passives, and restructuring* 179–204. New York: Oxford University Press.

Yoon, James H. 2007. Raising of major arguments in korean and japanese. *Natural Language & Linguistic Theory* 25(3). 615–653.

Zyman, Erik. 2018. *On the driving force for syntactic movement*. Santa Cruz, CA: University of California, Santa Cruz PhD dissertation.

VITA

Graduate School
Southern Illinois University Carbondale

Fahad A. Almalki

faamalki@uqu.edu.sa
ifahadalmalki2020@gmail.com

Umm Al-Qura University
Bachelor of Arts, English Language, 2017

Thesis Title:
    Appearances ARE deceiving: Long-distance Subject Anaphors and Phasal Binding
    Domains

Major Professor: Dr. Jeffrey Punske