Differential Object Marking in Levantine Arabic

Thesis submitted in partial fulfillment

of the requirements for the degree of

"DOCTOR OF PHILOSOPHY" in Cotutelle in Linguistics

in Ben-Gurion University of the Negev and McMaster University

by

Aya Zarka

Submitted to the Senate of Ben-Gurion University

of the Negev and to the School of Graduate Studies at McMaster University

26.12.2023

Beer-Sheva

Differential Object Marking in Levantine Arabic

Thesis submitted in partial fulfillment

of the requirements for the degree of

"DOCTOR OF PHILOSOPHY" in Cotutelle in Linguistics

in Ben-Gurion University of the Negev and McMaster University

by

Aya Zarka

Submitted to the Senate of Ben-Gurion University

of the Negev and to the School of Graduate Studies at McMaster University

26.12.2023

Beer-Sheva

This work was carried out under the supervision of

Dr. Tova Rapoport

In the Department Foreign Literatures and Linguistics, Ben-Gurion University

Dr. Ivona Kučerová

In the Department Linguistics & Languages, McMaster University

Faculty of Humanities and Social Sciences

Acknowledgments

This Ph.D. was completed as a co-shared thesis ("Cotutelle") between McMaster University, Ontario, Canada (host institution) and Ben-Gurion University of the Negev, Israel (home institution). My work was supported by Ph.D. grants from the faculty of humanities and from the Department of Foreign Literatures and Linguistics in Ben-Gurion University and by the Department of Linguistics and Languages in McMaster University.

My foremost thanks goes to my committee: Ivona Kučerová, Tova Rapoport, Idan Landau, Olga Kagan, and Jon Sprouse. It was a pleasure and a privilege to work with them. I thank them for helping me to create something that I can be proud of and for encouraging me every step of the way. I learned a lot from them, and the pages of this dissertation have benefited tremendously from their knowledge and feedback.

I would like to express my sincere gratitude to my primary supervisor, Ivona Kučerová, for her invaluable advice and guidance throughout the research phase of my PhD study at McMaster University. Her constant encouragement, unfailing support, and articulate criticisms have played a significant role in producing this work in its present form. She has had a tremendous influence on all of my work and on how I approach linguistics. I spent a long time working on topics that did not make it into this thesis. Ivona was patient with me while always gently reminding me that I had a thesis to write. I thank her for teaching me so much, for always pushing me to be more explicit about the details. She is a true example of what a good supervisor should be.

Additionally, I would like to thank my second supervisor, Tova Rapoport, for her feedback and encouragement. She constantly forced me to think about my dissertation work in the broader context, and helped me work through the information-structure research. Tova has also been tremendously supportive on both a personal and professional level, especially at times when I really needed it. I would like to thank her for giving me so much of her time and energy.

Idan needs no introduction. He is encouraging and inspirational, but at the same time tough and critical. His criticism greatly improved this dissertation, in particular chapters three and four. Idan always saw and solved my problems before I realized they existed. I'm also grateful for his outstanding workshop in my second year.

Olga deserves special thanks as she introduced me to the topic of differential object marking when she sent me a descriptive paper asking about data in Arabic. She has a particular way of asking those big questions and encouraging me to find the big picture in lots of messy data. Her comments have helped to shape this thesis into its current form, and it would not be the same without her feedback.

Although Jon and I did not have the chance to meet regularly while I was writing this dissertation, I'm grateful that he agreed to be a member at my committee and that he pressed me on the more philosophical and higher-level thinking at experimental work.

Much of the dissertation was presented at conferences and other linguistic events. I thank the audiences and reviewers for their feedback. These are the following: CLA (York University, May 31- June 2, 2023), ASAL 36 (University of Massachusetts Amherst, March 24-26, 2023), PLC 47 (University of Pennsylvania, March 18-19, 2023), MOTH (Université du Québec à Montréal, April 22-23, 2022), IATL 36 (Bar Ilan University, Israel, July 11-13, 2021, presented online), CLA (online, June 2-7, 2021), MOTH (McMaster University, April 19-20, 2021). I am particularly grateful to the following linguists from the above events for their comments and feedback: Aviya Hacohen, Zarina Levy-Forsythe, Aynat Rubinstein, Gabi Danon, David Erschler, Susana Béjar, Ali Alzayid, Arsalan Kahnemuyipour, Faruk Akkuş, John Kingston, Monica-Alexandrina Irimia, Roey Gafter, Dorit Ben-Shalom, Jordan River, Natasha Kasher, Martin Salzmann, Richard Compton, Ilena Paul, Keir Moulton, Eric Mathiew, Frederico Prado, Betsy Ritter, Patricia Schneider-Zioga, Aviv Schoenfeld, María Cristina Cuervo, Nora Boneh, Noa Bassel, Alexander Hamo, Thomas Leu, Dana Plaut Forkosh, Martin Renard.

I owe thanks to several other faculty members at McMaster: Victor Kuperman, Lili Service, Daniel Schmidtke and Alison Biggs. In particular, I would like to thank Victor and Lili for supervising my comprehensive paper during my second year. Completing the comprehensive paper could not been done without the help of Daniel on statistics in R. The members of the Syntax Lab (directed by Ivona Kučerová), both current and alumna, have

greatly influenced this work: Cassandra Chapman, Tess Hudson, Frederico Prado, Meagan Hann-Leger, Kurtis Commanda, Ethan Stollar, Audrey Ho, William Jarvis-Frrain, Victor Carri and Alina Dochu. Thanks for listening to messy practice talks and poorly worked out ideas. Your feedback has improved this thesis.

On a personal level, thank you to my family, especially my parents, Hussein and Ina'am; my parents-in-law, Slieman and Nayfa; and my sister Dalia. Their love and support have kept me going all these years, despite being so far from home. Last but most definitely not least, thanks to my wonderful husband Nabeh, who supported my PhD studies and accompanied me on this journey. There is no exaggeration when I say that none of this could have happened without him.

Abstract

This dissertation investigates the phenomenon of Differential Object Marking (DOM) in Northern Galilee Arabic (NGA). DOM is a widespread linguistic phenomenon in which certain objects of verbs appear in a different form from the expected one depending on various factors. While DOM has been well studied cross-linguistically, it has been less investigated in Arabic and in particular NGA. The thesis provides a detailed investigation of DOM in NGA which is complemented by novel arguments about the syntax, semantics and pragmatics of DOM.

I demonstrate that DOM in NGA has the following properties: (i) the object needs to be an individuated definite DP, (ii) the DP is marked by a prepositional dative, (iii) the dative marked DP is accompanied by clitic doubling i.e., the clitic and the marked DP co-refer, and (iv) the marked DP is interpreted as an aboutness topic.

I argue that DOM is derived by rightward A-movement to the edge of vP, specifically to Spec-ApplP. I provide three pieces of evidence for this argument: (i) island tests, (ii) binding, and (iii) adverbial placement tests. The empirical motivation for movement to ApplP is primarily for case checking. The movement is accompanied by clitic doubling, where the DP object and the corresponding clitic start the derivation by forming a big-DP. The DP object is differentially marked by movement to Spec-ApplP where the Appl head assigns dative case to it.

Another novel motivation for the DOM movement to ApplP comes from aboutness topicality. I argue that the property that distinguishes DOM structures in NGA from their non-DOM counterparts is aboutness topicality (Reinhart 1981). I argue that the obligatory movement of the differentially marked object to the phase edge of *v*P allows its anchoring to a referential address, making it an aboutness topic. I further observe that only nominals that can be mapped onto a referential address (Endriss 2009) can be differentially marked.

Consequently, quantifier phrases can also be DOM but only if they can be mapped onto a minimal witness set.

This dissertation also explores which types of nominals can participate in DOM. Based on Zarka and Hacohen's (2023) experimental work, I show that only highly atomic nominals in Grimm's (2012) sense can be differentially marked. I further demonstrate that these nominals are mappable onto a referential address and are able to hold an anaphoric relation with a clitic. The big-DP, i.e., the DP and the corresponding clitic which form a single unit, accurately predicts that elements that are mappable onto a referential address are able to be linked to a pronoun. I argue that the DOM structure is generated only if the nominal is merged as part of the big-DP.

However, the syntactic analysis does not fully account for the overall distribution of DOM. I demonstrate that DOM has certain pragmatic properties, which raises the question of how they arise. Adopting Kučerová and Zarka (in prep), I argue that DOM in NGA functions as an illocutionary marker that grammatically marks the asserted proposition as a non-default Discourse Commitment (Gunlogson 2001; Farkas & Bruce 2010) giving rise to a range of speech acts including emotive content. It is suggested that the obligatory illocutionary properties of DOM stem from structural economy. Since DOM involves an additional structure (clitic doubling and an applicative projection), which is absent in non-DOM counterparts, this additional structure triggers interpretive effects that would not be available otherwise (e.g., Fox 2000; Sichel and Wiltschko 2021).

The analysis put forward in this thesis is specific to NGA. It is an open question whether it extends to other languages with DOM. Similar to NGA, DOM in languages like Romanian (e.g., Hill & Tasmowski 2008) and Catalan (e.g., Escandell-Vidal 2009) is associated with discourse pragmatic effects, but further research is required for other languages. The proposed analysis of DOM raises the question of whether a uniform analysis of DOM is possible by unifying the different DOM systems as instantiations of economy-driven interpretive effects.

Keywords: Differential Object Marking, syntax, A-movement, applicative structure, information structure, aboutness topicality, anchoring, pragmatics, non-default Discourse Commitments, individuation, atomicity, referentiality, North Galilee Arabic

Table of Contents

List of Tables

List of Figures

Abbreviations

1	first person
2	second person
3	third person
ACC	accusative
ADV	adverb
APPL	applicative
APPLP	applicative phrase
ATB	across-the-board
AUX	auxiliary
BP	broken plural
CLLD	clitic left dislocation
CLRD	clitic right dislocation
C	complementizer
CL	clitic
CD	clitic doubling
CD	coreferential dative construction
CG	common ground
COLL	collective
CP	complementizer phrase
CSC	coordinate structure constraint
DOM	differential object marking
D	determiner

DP	determiner phrase
DAT	dative
DEF	definite
ERG	ergative
FUT	future tense
F	feminine
HTLD	hanging topic left dislocation
IO	indirect object
IMP	imperative
IMPF	imperfective
MSA	Modern Standard Arabic
M	masculine
NOM	nominative
NEG	negative marker
NGA	North Galilee Arabic
NP	noun phrase
OVS	object-verb-subject
OSV	object-subject-verb
OBJ	object
OM	object mass
PP	prepositional phrase
PL	plural
PERF	perfective
PST	past tense
PROG	progressive
PRS	present tense
PF	phonological form
P	phrase

QP quantifier phrase

RRC right roof constraint

RNR right node raising

SG singular

SP sound plural

SVO subject-verb-object

SM substance mass

SPEC specifier

TP tense phrase

VP verb phrase

Chapter 1

Introduction

1.1 The phenomenon

Differential Object Marking (DOM) is a widespread linguistic phenomenon in which certain objects of verbs appear in a different form from the expected one depending on various factors. It has been hypothesized that DOM arises because of certain semantic factors such as specificity, animacy, definiteness, and topicality (Silverstein 1976; Comrie 1989; Bossong 1991; Aissen 2003; García-García 2005; López 2012; Kagan 2020; Tal et al. 2022; Van der Wal 2022, among many others). Turkish is a characteristic and well studied example of a language with DOM (e.g., Enç 1991; Kornfilt & von Heusinger 2009; de Hoop & de Swart 2009). In Turkish, the morphological marking on the direct object depends on the semantic-pragmatic factor of specificity as demonstrated in (1).

(1) *Turkish* (Enç 1991: 5 (4-5))

a. Ali bir kitab- *(ı) aldi
Ali one book-ACC bought
'A book is such that Ali bought it.' DOM

b. Ali bir kitap- (*ı) aldi
Ali one book- ACC bought
'Ali bought some book or other.' NON-DOM

In other languages DOM is associated with different properties. For example, in Standard Spanish, DOM is obligatory with animate definites (2a) and animate specific indefinites (2b), but not with inanimates irrespective if they are definite or not (e.g., Irimia 2021).

(2) *Standard Spanish* (Irimia 2021: 489(1))

a. Busco ***(a)** la niña/**(*a)** la casa
look for.1SG DAT=DOM the.F.SG girl/ DAT=DOM the.F.SG house
'I'm looking for the girl/the house.'

b. Busco una niña/ **a** una niña/**(*a)** una casa
look for.1SG a.F.SG girl/ DAT=DOM a.F.SG girl/ DAT=DOM a.F.SG house
'I'm looking for a (random) girl/a specific girl/a house

Another semantic factor that has been widely discussed in the DOM literature is definiteness (e.g., Næss 2004; de Swart 2007). In some languages, such as in Modern Hebrew, DOM is sensitive to the general distinction between definite and indefinite nominals: only definite nouns are marked with the accusative marker *et* (Givón 1978; Wintner 2000; Danon 2001, 2008; Kagan 2020, among many others).

(3) *Hebrew* (Kagan 2020: 147(1))

a. raiti ***(et)** ha-yeled
I.saw ACC the-boy
'I saw the boy.' DOM

b. raiti **(*et)** yeled

I.saw ACC boy

'I saw a boy.' NON-DOM

Languages that only take animacy or definiteness as the factor that controls DOM have 'one-dimensional' DOM systems. For instance, as shown for Hebrew, definite direct objects take obligatory case marking regardless of their level of animacy. However, languages where animacy and definiteness intersect and both dimensions are taken as determining factors for the overt marking of DOM, such as in Romanian (Chiriacescu 2014), Hindi (McGregor 1972; Mohanan 1990; Verbeke & Ponnet 2022), Spanish (von Heusinger & Kaiser 2003; Leonetti 2004; de Swart and de Hoop 2007), have 'two-dimensional' DOM systems (Krause and von Heusinger 2019).

While DOM has been thoroughly studied cross-linguistically, the literature on DOM in Arabic is rather scarce. DOM is only attested in some Arabic dialects but is not attested in Modern Standard Arabic and other dialects. One of the dialects that allow DOM is North Galilee Arabic (NGA). DOM in NGA is realized as a dative case-marking. In NGA, and other closely related Arabic dialects e.g., Syrian Arabic and Lebanese Arabic, there is no overt accusative marking. DOM corresponds to the alternation between no overt marking (4a) and overt dative marking (4b) of the direct object, as shown in the following examples.

(4) *North Galilee Arabic (NGA)*

a. šar-et **l-ktaab** BASELINE

bought-1SG the-book

'I bought the book.'

b. šar-et-o **la-l-ktaab** DOM

bought-1SG-3M.SG.OBJ DAT(DOM)-the-book

'I bought the book.'

In addition to dative marking, the DOM example in (4b) also involves clitic doubling. Further, as one can notice, DOM is not conditioned by definiteness. These facts will be investigated more closely throughout the thesis.

The main aim of this thesis is to provide a detailed investigation of DOM in NGA which is complemented by novel arguments about the syntax, semantics and pragmatics of DOM.

1.2 North Galilee Arabic (NGA): A brief background

The Arabic language is a collection of variants among which the standard variety, Modern Standard Arabic (MSA), has a special status, while the other varieties are considered colloquial dialects (Bateson 1967; Holes 2004; Habash 2010). MSA is the official language of public speaking and news broadcasts on radio and television (Ryding 2005). Dialects, on the other hand, are used informally in daily life and are widespread and divided by geographic region.

There are many Arabic dialects that are used throughout the Arab world. However, there are five major dialects that are widely spread: the Gulf dialect, the Egyptian dialect, the Maghrebi dialect, the Levantine dialect, and the Iraqi dialect (Altamimi et al. 2018).

North Galilee Arabic (NGA), the dialect under investigation, belongs to the Levantine dialectal group. Levantine Arabic is primarily spoken in Syria, Lebanon, Jordan, Israel, and Palestine. NGA is spoken in the northern Galilee region of Israel. The data reported in this thesis is produced by Druze informants, a minority group in Israel,[1,2] who speak NGA. The informants live in the town of Peki'in. The dialect is closely related to and mutually intelligible with the Druze dialect in southern Lebanon, southwestern Syria, and with that of other (not necessarily Druze) informants living in the same region.

Similar to other Arabic dialects, but unlike MSA, NGA has lost its overt morphological case. The only exception is pronouns which display overt morphological

[1] There are 145,000 Druze in Israel, which constitutes approximately 1.6% of Israel's total population (Kheir 2023). Druze people in Israel have their own distinct sector, separate from that of the Arabs (Firro 2001; Halabi 2006; Kheir 2023).

[2] The Arabic dialect spoken by Druze (in Israel) has been previously studied (Isleem 2016; Kheir 2019). Recent work has focused on the influence of language contact in the Druze dialect (Kheir 2019, 2023). Previous studies also indicate that Druze dialects in Israel must be distinguished from other Arab dialects with respect to their phonological properties (Blanc 1953; Habib 2005).

case.[3] NGA is also similar to most other Arabic dialects with respect to word order in which SVO is the unmarked order. (See Aoun et al. 1994 for a similar view for Lebanese Arabic, Shlonsky 1997 and Mohammad 2000 for Palestinian Arabic, Benmamoun 2000 and Edwards 2006 for Egyptian Arabic).[4]

Lastly, like other Arabic dialects, NGA allows full subject-verb agreement irrespective of the word order used. This property is not relevant for this thesis.[5]

This thesis is dedicated to exploring the DOM construction (4b) with a focus on its interpretive, syntactic, and distributional properties.

1.3 Roadmap

- Chapter 2 "*Properties of DOM in Arabic and crosslinguistically*" first presents the DOM patterns in NGA and other closely Arabic dialects, and then discusses the Arabic facts in relation to the existing DOM literature in general.

- Chapter 3 "*The clausal syntax of DOM in NGA*" establishes an empirical generalization that DOM is an instantiation of dislocation and proposes a novel syntactic analysis of DOM. The chapter argues that DOM is derived by movement, and the directionality of the movement is rightward. The chapter then examines the height of the DOM landing position, arguing that DOM targets a clause-internal position at the edge of *v*P phase.

- Chapter 4 "*The landing position of DOM*" concerns the identity of the DOM landing position. I present three empirical motivations for ApplP as the target for DOM and the predictions that the proposal makes. The motivations are (i) checking case, (ii)

[3] The reasons for the loss of morphological case in Arabic dialects are described in Owens (2006).

[4] NGA (like other dialects) allows VSO but DOM is not preferred in this order. I leave this limitation of word order with DOM to be investigated in future work.

[5] See Mohammad (1990, 2000), Fassi Fehri (1993, 2005, 2012), Benmamoun (2000), Benmamoun and Lorimor (2006) and Soltan (2007) on this property.

locality conditions of the DOM movement, and (iii) aboutness topicality (which will be argued in detail in Chapter 5).

- Chapter 5 "*DOM and information structure in NGA*" argues that motivation for the DOM movement to ApplP is aboutness topicality. The chapter argues that the defining property underlying the interpretation of DOM is aboutness topicality. It also excludes information-structural properties such as referentiality, specificity and givenness which have been previously discussed in the DOM literature. I argue that the obligatory movement of DOM to the edge of the *v*P phase is a direct consequence of topics requiring association with a referential address.

- Chapter 6 "*DOM as a non-default Discourse Commitment Device*" argues that aboutness topicality (argued in Chapter 5) cannot fully explain the distribution of DOM in NGA. I examine cases of DOM in NGA where using DOM, as opposed to its non-differentially marked counterpart, triggers certain pragmatic effects including signaling the speaker's emotive content, a correction, an accommodation, or activating a parallel Question Under Discussion. The chapter then addresses the question of how these effects arise by adopting recent work by Kučerová and Zarka (in prep) which argues that DOM in NGA behaves as an illocutionary marker and that it grammatically marks the asserted proposition as a non-default Discourse Commitment (Gunlogson 2001; Farkas & Bruce 2010; Rett 2021).

- Chapter 7 "*Individuation: Countability or atomicity?*" explores which types of nominals may be differentially marked. I demonstrate that only individuated nominals can participate in DOM. The chapter argues that individuation in the sense of atomicity is the right notion for describing the nominal distribution of DOM in NGA. Importantly, I follow Grimm (2012) that atomicity is viewed as a gradable phenomenon.

- Chapter 8 "*Explaining the individuation restriction*" develops a semantic analysis in which aboutness topicality provides an explanation for the restriction presented in Chapter 7 that certain nominals that are high on the atomicity scale appear with DOM.

- Chapter 9 concludes the thesis and discusses some open questions that emerge from the proposals developed in this thesis.

Chapter 2

Properties of DOM in Arabic and crosslinguistically

2.1 The DOM patterns in Arabic

This chapter first presents the DOM patterns in North Galilee Arabic (NGA) and other closely related Arabic dialects, then discusses the Arabic facts in relation to the existing DOM literature in general.

DOM is a widespread linguistic phenomenon in which certain objects of verbs appear in a different case from the expected one depending on various factors. This section presents the patterns of DOM in several Arabic dialects. DOM is not attested in all Arabic dialects (Brustad 2000; Döhla 2016). The dialects that exhibit DOM are Levantine, Syrian, Lebanese, Iraqi and Maltese etc. I discuss here only a subset of these dialects, namely the Lebanese, Syrian, and Levantine dialects which are more closely related to NGA than the Iraqi and Maltese dialects. The terms dialect and variety will be used interchangeably. Most

of the data reported in this thesis are from NGA, as described in the first chapter, and data is produced by the researcher (being a native speaker of the language), unless stated otherwise.

In this section I concentrate on the basic structural properties of the DOM construction in NGA. The properties are: (i) dative case marking on the direct object, (ii) clitic doubling, (iii) definiteness, (iv) right dislocation, and (v) individuation. Let me briefly expand on each one of the properties. The object receives a prepositional dative and the proposition is *la-*. The construction also necessarily involves clitic doubling. That is, the DOM structure contains a clitic and a doubled DP, and the clitic attaches to the verbal host and both, the clitic and the doubled DP, must co-refer (Aoun 1999; Hallman and Al-Balushi 2022). In this section and throughout the thesis, I demonstrate that not every nominal can be differentially marked. Only definite DPs can be marked. Brustad (2000) observes that DOM in Syrian Arabic involves right dislocation as will be shown later in this section and analyzed more closely in Chapter 3. Lastly, DOM only occurs with individuated nominals (See Chapter 7 on the relevant notion of individuation).

In NGA and other closely Arabic varieties, including Lebanese, Syrian, Levantine, DOM is obtained by the alternation between accusative marking and overt dative case marking (Levin 1987; Aoun 1999; Brustad 2000, 2008; Döhla 2016; Sterian 2016). In these dialects, accusative marking does not have an overt inflection. In the following examples, DOM corresponds to the alternation between accusative marking (5a) and overt dative marking (5b).

(5) a. šof-t **sˁ-sˁabeyy-e**
saw-1SG the-young_lady-F.SG
'I saw the young lady.' NON-DOM

b. šof-t-a **la-sˁ-sˁabeyy-e**
saw-1SG-3F.SG.OBJ DOM(DAT)-the-young_lady-F.SG
'I saw the young lady.' DOM

In (5), the same definite argument *sˁ-sˁabeyye* 'the young lady' can be either zero-marked or overtly marked with case. In (5a), the definite argument is accusative (unmarked), while

in (5b), the definite object is obligatorily marked with a dative marker *la-* (as assumed by Aoun 1999; Hallman 2018; Hallman and Al-Balushi 2022).[6]

In NGA, the exponents of DOM and dative case are morphologically identical. The following examples (6)-(7) demonstrate that DOM indeed looks like the dative in places where, structurally, we expect dative marking. We find that the same dative case marker *la-* marks locational DPs (6a), indirect objects of a ditransitive construction (i.e., recipient) as in (6b), or benefactives as in (6c).

(6) a. roḫ-et la-l-modiir
went-1SG DAT-the-principal.M.SG
'I went to the principal.'

b. aʕṭ-it al-ktaab la-sˁ-sˁabeyy-e
gave-1SG the-book DAT-the-young_lady-F.SG
'I gave the book to the young lady.'

c. xabaz-et al-kaʕke la-sˁ-sˁabeyy-e
baked-1SG the-cake DAT-the-young_lady-F.SG
'I baked the cake for the young lady.'

Further, the dative case marking appears in possessive constructions as demonstrated in (7).

(7) *Syrian Arabic*

a. hay al-bloz-a la-sara
this.F.SG the-shirt-F.SG DAT-Sara
'This is Sara's shirt.'

b. laʔe-na ktaab-o la-xaalid
found-1PL book-GEN.3M.SG DAT-khalid
'We found Khalid's book.' (Hallman & Balushi 2022: 6(5b))

[6] Hallman and Al-Balushi (2022) pointed out that the possibility of clitic doubling with a *la*-phrase in Damascus Arabic, and other Levantine varieties in general, is thought to be a contact feature borrowed into Levantine Arabic from Aramaic, which had close contact with the Levantine varieties (Hallman and Al-Balusihi 2022: p.1298, fn.4, Hallman and Al-Balushi cited Coghill 2014; Contini 1999; Feghali 1928: 362; Lentin 2018, among others). See also Döhla (2016: 165) on the emergence of DOM in Levant dialects.

In (7), the possessor arguments, *sara* and *xaalid*, are marked with the dative case marker. Thus, DOM in NGA gets realized as dative since it appears in configurations where dative is expected such as those illustrated in (6)-(7).

The DOM structure employs the dative case marker *la-* and obligatorily includes a pronominal element which attaches to the verbal host and co-refers with the subsequent definite differentially marked object (Levin 1987; Aoun 1981, 1999; Shlonsky 1997; Brustad 2000; Döhla 2016; Sterian 2016; Hallman and Al-Balushi 2022).[7]

Existing proposals classify this construction as Clitic Doubling (CD) (Aoun 1999).[8] In CD, a phonologically bound morpheme (the clitic) expresses the agreement features φ of a full nominal phrase i.e., the associate (Harizanov 2014: 1034).[9] However, I call this construction DOM. The following examples in (8) demonstrate that the accusative clitic appears as a suffix to the verbal host and its associate (the differentially marked DP) must co-refer.[10,11]

---

[7] Maltese Arabic variety marks certain objects with dative (preposition *lil-*) but unlike other Arabic varieties, does not use a pronominal element (Rubin 2005: 106).

[8] CD is not a characteristic of Arabic in general as pointed by Sterian (2016). CD is attested only in certain Arabic dialects such as Levantine, Lebanese, Syrian and Northern Mesopotamian dialects spoken in Northern Iraq and Turkey. Contrary to these dialects, CD is not attested in Modern Standard Arabic (Alzayid 2022). This is illustrated by the contrast between Levantine Arabic (i) and MSA (ii).

(i) sami šaf-o la-jawad
Sami see.3M.SG.PST-him DOM-Jawad
'Sami saw Jawad.'

(ii) *sami-un raʔa-o la-jawad-an
Sami-NOM see.3M.SG.PST-him DOM-Jawad-ACC
('Sami saw Jawad.')

Furthermore, there are differences even within Arabic varieties with clitic doubling. Unlike NGA, Lebanese, and Syrian (where the preposition *la-* is used), an Arabic variety in Çukurova (Turkey) uses a different preposition *ʕala* 'on' (Sterian 2016: 73(3.43)).

(i) aḥmad yḥibb-u ʕala abu-hu
Ahmad loves-him on father-his
'Ahmad loves [him] his father.'

[9] See section 3.6 on comparing DOM to other similar constructions such as CD.

[10] Similarly, when the marked object is a broken plural as in (8c), the clitic is also in a plural form (no gender is specified). See more data in Chapter 7.

[11] The examples in (8) include human referents. As will be shown in (24), DOM is insensitive to animacy, thus inanimate nominals are equally fine with DOM.

(8) *DOM in NGA obligatorily involves CD*

a. dalia šaf-at-*(o) la-l-walad
Dalia saw-3F.SG-3M.SG.OBJ DOM-the-boy.M.SG
'Dalia saw the boy.'

b. dalia šaf-at-*(a) la-sˁ-sˁabeyy-e
Dalia saw-3F.SG-3F.SG.OBJ DOM-the-young_lady-F.SG
'Dalia saw the young lady.'

c. dalia šaf-at-*(on) la-sˁ-sˁabaya/wlaad
Dalia saw-3F.SG-3PL.OBJ DOM-the-young_ladies.BP/ boys.BP
'Dalia saw the young ladies/boys.'

While examples in (8) are transitive and include DP arguments, DOM is also grammatical with PP complements. The base structure in (9a), for example, has the DOM counterpart in (9b).

(9) a. ʔnti sˁarraxet ʕala yosef
you shouted at Yousef
'You shouted at Yousef.' NON-DOM

b. ʔnti sˁarraxet ʕal-ih la-yosef
you shouted at-3M.SG.OBJ DOM-Yousef
'You shouted at Yousef.' DOM

Similar to PP complements, DOM is also compatible with PP non-complements (10b).

(10) a. mšit maʕ yosef
walked.1SG with Yousef
'I walked with Yousef.' NON-DOM

b. mšit maʕ-o la-yosef
walked.1SG with-3M.SG.OBJ DOM-Yousef
'I walked with Yousef.' DOM

DOM is also attested with indirect objects in addition to direct objects (Aoun 1981, 1999).

(11) a. karim ḥaka-l-o **la-sami** qosˁsˁ-a
Karim told-to-3M.SG.IO DOM-Sami story-F.SG
'Karim told Sami a story.'

b. karim baʕath-l-o **la-sami** al-makatiib
Karim sent-to-3M.SG.IO DOM-Sami the-letters.BP
'Karim sent the letters to Sami.'

The indirect object 'Sami' can be differentially marked using the accusative clitic *-o* 'him.' The indirect object is not cliticized to the verb, but rather to a preposition because verbs such as *ḥaka* 'to say a story, to talk' or *baʕath* 'send' in Arabic c-select for a PP complement.

Although certain objects can differentially marked, subjects cannot.[12]

(12) (*la) sara$_i$ qarat-a$_i$ l-ktaab
DOM Sara read-3F.SG the-book.M.SG
('Sara read the book.')

Following Hallman and Al-Balushi (2022), I assume that the rich agreement morphology on the verb in Arabic is present irrespective of the definiteness of the subject or whether the subject is present at all. However, subjects may never occur in a *la*-phrase. I take this to mean that subjects may not be differentially marked.[13]

When the clause contains the DOM marker, CD is obligatory.[14] Consider the following examples.

[12] Aoun (1999: 712) notes that both preverbal and postverbal subjects cannot be clitic doubled.

[13] Subjects can be clitic doubled in ECM configuration as noted by Aoun (1999: 712(9)):

(i) badde-yaa yruuḥ la-kariim
want.I-him.ACC go.3PRS to-Kariim
'I want Kariim to go.'

[14] This directionality is important. NGA allows clitic doubling in another configuration such as CLLD. In CLLD, CD is obligatory, but DOM is not allowed. See examples on CLLD in section 4.2.

(13) a. dalia šaf-at-*(a) la-sˁ-sˁabeyy-e
Dalia saw-3F.SG-3F.SG.OBJ DOM-the-young_lady-F.SG
'Dalia saw the young lady.'
b. ʔnti sˁaraxet ʕl-*(ih) la-yosef
you shouted at-3M.SG.OBJ DOM-Yousef
'You shouted at Yousef.'

An additional property of DOM in NGA is definiteness. Abu-Haidar (1979) was the first to observe that, like NGA, the DOM structure in Lebanese Arabic only occurs with definite nominals. The NGA example below demonstrates that only definite objects license DOM.[15]

(14) *DOM must appear with definite objects*
šof-t-a la-***(l)**-mʕalm-e
saw-1SG-3F.SG.OBJ DOM-the-teacher-F.SG
'I saw the female teacher.'

Further, DOM in NGA can appear with proper names as illustrated in (15).

(15) *DOM is compatible with proper names*
sara šaf-at-o la-sami
Sara saw-3F.SG-3M.SG.OBJ DOM-Sami
'Sara saw Sami.'

While DOM is allowed with DPs and proper names, it is not allowed with pronouns.[16,17]

[15] Definiteness in Levantine has conventionally been depicted in binary terms. For instance, Cowell (1964) defines it as the presence or absence of a definite article *al-*. Specifically, when a common noun is modified with an article *il-* or *hal*, it is marked as definite, and when any noun is not modified as such, it is considered indefinite. While *il-* is equal to the definite article in other Arabic dialects; *hal-* is described by Cowell (1964: 556) as a prefixed or proclitic reduction of the proximal demonstrative combined with the article.

[16] Akkuş (2021: 20), unlike Aoun (1999), notes in footnote 11 that pronouns are not preferred to be doubled. Also, Sterian (2016: 18) observes that some Arabic dialects do not allow clitic doubling with pronouns.

[17] See footnote 133 for my very preliminary explanation of why DOM is illicit with pronouns in NGA.

(16) *DOM is incompatible with pronouns*

a. * sami šaf-o la-ʔlo
Sami saw.3M.SG-3M.SG.OBJ DOM-3M.SG
('Sami saw him.')

b. * sami šaf-ak la-ʔlak
Sami saw.3M.SG-2M.SG.OBJ DOM-2M.SG
('Sami saw you.')

All types of pronouns are unacceptable with DOM in NGA; there is no difference in judgments between 1st, 2nd and 3rd person pronouns (singular and plural).

Unlike in Turkish wherein specificity plays a role for DOM (shown in Chapter 1), indefinite specific nominals in NGA are prohibited from marking. In NGA, and other close Levantine dialects, the article *ši-* appears with indefinite specific nominals (in Enç's 1991 sense) where *ši-* is used to provide a referent semantic specificity, while at the same time, indicating that the speaker is unable to recognize it.[18] The following example, taken from Brustad (2000: 27) demonstrates that speakers use *ši* to indicate that they have a particular type of entity in mind.

(17) *Syrian Arabic*

laazim nʕmil-l-u ši muqaddime la-ḥatta maa-
necessary make.IMPF.1PL-DAT-3M.SG some preparation so-PURP NEG-
yinsʿidim
be.shocked.IMPF.3M.SG
'We need to arrange some sort of preparation for him so he won't be shocked.'

(Brustad 2000: 27)

Even though the speaker cannot recognize the nature of the *muqaddime* 'preparation' explicitly (i.e., the speaker does not have a concrete preparation in their mind), it is

[18] Based on Dryer's (2014) Reference Hierarchy model, Turner (2018) classifies *ši-* as a Pragmatically Non-Specific Indefinite article (PNI). Turner (2018: 35) summarizes PNI as describing "a unique entity, known to neither the speaker nor the listener, about which the speaker believes further information is available but is not able to provide personally."

semantically specific in the sense that the speaker can identify a particular purpose for which it can be used (in this case, to keep the subject of the conversation from being surprised).

The example below illustrates that indefinite-specific nominals, marked with the article *ši,* are not allowed with DOM.

(18) *Specific indefinites cannot be DOM*

* šof-t-a la-ši sˤabeyy-e

saw-1SG-3F.SG.OBJ DOM-some young_lady-F.SG

('I saw some young lady.')

Hallman and Al-Balushi (2022) note that even adding additional modificational material, which could be expected to support a specific reading of an object, fails to license DOM of an indefinite object.

(19) a. * šof-na-ha la-ʔwarib ʔšriat ʔlwan fatḥa

saw-1PL-3F.SG.OBJ DOM-boats sails colors bright

('We saw boats with brightly colored sails.')

b. * šof-na-ha la-waḥde min l-banaat bi-s-suʔ

saw-1PL-3F.SG.OBJ DOM-one of the-girls at-the-market

('We saw one of the girls at the market.')

(adapted from Hallman & Al-Balushi 2022: 1299)

Further, I build on Brustad's (2000) observation that DOM in NGA involves right dislocation. While degree adverbs such as 'very' intervene between the verb and the differentially marked object, they cannot follow the differentially marked object, as exemplified in (20).

(20) a. ma-g'dert-š afham-o$_i$ ktiir la-aḥmad$_i$

NEG-able-NEG understand.1SG-3M.SG.OBJ very DOM-Ahmad

'I can't completely understand Ahmad.'

b. ma-g'dert-š afham-o$_i$ la-aḥmad$_i$ *ktiir
NEG-able-NEG understand.1SG-3M.SG.OBJ DOM-Ahmad very
'I can't completely understand Ahmad.'

Similar to degree adverbs, other adverbials such as locatives can only intervene between the verb and the differentially marked object.

(21) a. šoft-o$_i$ honak la-aḥmad$_i$
saw.1SG-3M.SG.OBJ there DOM-Ahmad
'I saw Ahmad there.'

b. šoft-o$_i$ la-aḥmad$_i$ */??honak
saw.1SG-3M.SG.OBJ DOM-Ahmad there
'I saw Ahmad there.'

The distribution of DOM with adverbs in (20)-(21) suggests that differentially marked objects in NGA undergo obligatory dislocation. I will discuss these facts in more detail in Chapter 3.

The last property of DOM that will be briefly discussed here is individuation. It has been argued that individuation is correlated with DOM licensing in Arabic (Khan 1984; Brustad 2000, 2008). For these scholars, individuation is an umbrella term which combines many factors such as definiteness, animacy, specificity, and quantification. I follow their view that individuation is the key for licensing DOM but a narrower understating is needed to capture the DOM facts in NGA.

In Zarka (2021), I build on Brustad's (2000) original proposal that individuation is the key factor in licensing DOM in NGA. I focus on one feature, *quantification*, out of a cluster of features. I observe that a type of nouns such as the broken plurals (BPs) allow both kind and unit readings.[19] In (22), *karaasi* 'chairs' can be interpreted as either 'kind' or 'unit.'

[19] It has been noted for Modern Standard Arabic that certain patterns of broken plurals are connected to different meanings, paucity vs. abundance (Wright 1898, 1:233–4; Fischer 2002, 53–64).

(22) sara šara-at arbaʕ karaasi
Sara bough-3F.SG four chairs.BP
'Sara bought four chairs.'

→ Sara bought exactly four individual chairs (4 total)
→ Sara bought exactly 4 kinds of chairs (e.g., if 2 of each kind, then 8 chairs total)

Interestingly, DOM causes the BP to have a unit interpretation, but blocks the kind interpretation, as shown in (23).

(23) sara šara-at-on la-l-arbaʕ karaasi
Sara bought-3F.SG-3PL.OBJ DOM-the-four chairs.BP
'Sara bought the four chairs.'

→ Sara bought exactly four individual chairs (4 total)
↛ Sara bought exactly 4 kinds of chairs (e.g., if 2 of each kind, then 8 chairs total)

The BP data suggest that DOM is possible when the BP denotes a unit reading but is impossible with a kind reading. This fact will be closely examined in Chapter 7 where I discuss an experimental study on DOM (Zarka & Hacohen 2023) that tests which notion of individuation best describes the distribution of nominals with DOM. I show that the results of the experimental study suggest that individuation under the notion of atomicity is the better characterization than my original generalization in Zarka (2021).

Lastly, one feature that does not play a role in DOM in NGA is animacy. While animacy is associated with DOM in Standard Spanish (shown in Chapter 1), DOM in NGA

is insensitive to animacy.[20] Examples in (24) reveal that animate objects, including human (24a) and animals (24b), and inanimate ones can freely appear with DOM (24c).[21]

(24) *Animacy does not play a role in NGA's DOM*

a. sara šaf-at-o la-sami
Sara saw-3F.SG-3M.SG.OBJ DOM-Sami
'Sara saw Sami.' HUMAN

b. sara šaf-at-o la-l-kalb
Sara saw-3F.SG-3M.SG.OBJ DOM-the-dog
'Sara saw the dog.' ANIMATE

c. sara šaf-at-o la-l-qasˁer
Sara saw-3F.SG-3M.SG.OBJ DOM-the-palace
'Sara saw the palace.' INANIMATE

Regardless of humanness or animacy, the sentences in (25) are ungrammatical when the differentially marked objects are indefinite.

(25) a. * sara šaf-at-a la-sˁabeyye
Sara saw-3F.SG-3F.SG.OBJ DOM-young_lady
('Sara saw young lady.')

b. * sara šaf-at-o la-kalb
Sara saw-3F.SG-3M.SG.OBJ DOM-dog

[20] Unlike NGA, an animacy-sensitive DOM language is Yiddish. In Yiddish, DOM only marks common nouns that are human and worthy of respect (Aissen 2003). Nouns denoting animals are also marked, however, they do not behave uniformly with respect to case-marking. The cut-off point for the sub-categories of the animal group seems to be located on the person boundary. The class of animals is rather frequently split into two subclasses according to the "personness" (in the sense of human resemblance) of the animal. Evidently, horses have much more personness than do ants or flies, but there may be intermediate cases in which usage varies according to context (Bossong 1983). In some languages, the animals are split into categories according to their size; the big animals belong to the high category and the small ones to the low (Krause and von Heusinger 2019). In NGA, however, there is no restriction in the class of animals; both nominals denoting animals that are high in "personness" as well low are freely marked.

[21] Iemmolo (2010: 83) draws an incorrect conclusion based on a partial profile data from Koutsoudas (1967) in Lebanese Arabic that DOM system clearly based on animacy: the DOM marker is found only with definite animate nouns.

('Sara saw a dog.')

c. * sara šaf-at-o la-qasˁer
Sara saw-3F.SG-3M.SG.OBJ DOM-palace
('Sara saw a palace.')

The list of distributional properties of DOM considered in this section thus far is summarized in (26).

(26) *Differentially marked objects in NGA are…*
a. marked with dative case
b. obligatorily clitic doubled
c. definite
d. right-dislocated
e. individuated
f. insensitive to animacy

This section has reviewed the basic properties of DOM in NGA and briefly in other closely related Arabic dialects. It has covered structural and morphosyntactic characteristics of the DOM construction in Arabic dialects. The following section discusses these characteristics in the context of the ongoing DOM literature.

2.2 Properties of DOM in other languages

This section discusses the empirical descriptions of DOM in Arabic. In particular, it addresses the properties of definiteness and animacy, dative marking, and clitic doubling in relation to the existing cross-linguistic DOM literature.

2.2.1 Definiteness and animacy

As mentioned in Chapter 1, DOM is associated with different properties such as definiteness and animacy. It has been viewed that animacy and definiteness are both scalar dimensions (e.g., Silverstein 1976; Aissen 2003). I explain below how the scalar approaches of definiteness and animacy proposed by Aissen (2003) fail to predict the DOM patterns in NGA.

Functional and typological studies argue that animacy and definiteness of direct objects are the most basic semantic-pragmatic factors that condition DOM and that the occurrence of DOM correlates with the degree of these notions: the higher the animacy and/or definiteness, the higher the chances the object to be marked (e.g., Aissen 2003: 436).[22] In some languages, such as in Hebrew, DOM is sensitive to the general distinction between definite and indefinite nominals: only definite nominals are marked with the accusative marker *et* (Givón 1978; Wintner 2000; Danon 2001, 2008; Kagan 2020, among many others).

(27) *Hebrew* (Kagan 2020: 147(1))

a. raiti *(et) ha-yeled
I.saw ACC the-boy
'I saw the boy.' DOM

b. raiti (*et) yeled
I.saw ACC boy
'I saw a boy.' Non-DOM

While in Hebrew all definite nominals get overt case marking, in NGA only a subset of definite DPs can be differentially marked. Similar to NGA, in certain languages discussed by Aissen (2003), particularly Catalan and Pitjantjatjara, more specific distinctions need to be addressed within the definiteness dimension in order to account for case-marking

[22] See, however, Sinnemäki (2014) on the typological evidence that there is no universal preference for object case marking to be driven by definiteness and animacy properties. However, data from 744 languages provide statistical evidence that restricted case marking of the object was preferred over non-restricted marking. It was concluded that languages tend to restrict object case marking in some way, but the details of this variation differ cross-linguistically more than has so far been assumed.

behaviors. In these languages, certain definite objects are case-marked, while others are not.[23]

Aissen (2003) argues that DOM is sensitive to the 'Definiteness scale': the higher an object is on each scale, the higher the differential marking probability. The definiteness hierarchy is represented as follows (Silverstein 1976; Croft 1988; Comrie 1989; Aissen 2003).

(28) Definiteness scale: Pronoun > Proper noun > Definite > Indefinite Specific > Non-Specific

The higher an object is on the definiteness scale, the more prominent it is. Hence, pronouns are more prominent than definite nominals, and in turn they are more prominent than their indefinite counterparts. Within the pronoun category, 1st and 2nd person pronouns can arguably be classified higher on the definiteness scale than 3rd person ones (Aissen 2003; Fernández & Rezac 2016; von Heusinger et al. 2018).[24]

The scalar approach proposed by Aissen cannot be applied to the DOM facts in NGA. Aissen's scale predicts that pronouns are highly likely to be marked; however, as shown in the previous section, DOM is illicit with pronouns, thus the scalar approach makes incorrect predictions for DOM in NGA.

Moreover, Khan (1984) demonstrates that in several Semitic languages, definiteness is not a sufficient condition for licensing object marking, that is, object marking does not occur with all definite nominals. He argues that in some contexts definiteness is not even a necessary condition for object marking, in that an object marking may also occur with indefinite nominals. Khan's generalization does not hold true for NGA since DOM must appear with a subset of definite nominals (as shown briefly in the previous section).

[23] In Catalan, the marker *a* obligatorily precedes strong personal pronouns that appear in the object position, however, it does not precede other kinds of definite objects, such as proper names and definite descriptions (Aissen 2003).

[24] Most Basque DOM dialects show that DOM is compatible with 1st and 2nd person pronouns, while only some exhibit DOM with 3rd person objects. In many Basque varieties, DOM is only obligatory or optional for 1st and 2nd object pronouns, but completely excluded for 3rd person objects (Fernández & Rezac 2016).

Besides definiteness, animacy is another factor often associated with DOM (e.g., Silverstein 1976; Aissen 2003; Krause and von Heusinger 2019). Like definiteness, animacy is also formed on a scale. Aissen (2003: 438) provides the following animacy scale:

(29) Animacy scale: human > animate > inanimate[25] (based on Croft 1988)

The scale cannot hold for NGA because animate and inanimate objects in NGA can be equally differentially marked (as shown in section 2.1).

Other scholars claim for Semitic languages in general and for Arabic specifically that animate nominals are more likely to receive differential case marking than inanimate ones (Khan 1984; Brustad 2000). This claim cannot be true for NGA as the NGA facts play out differently.

As mentioned in Chapter 1, languages that are determined by animacy or definiteness as the factor that licenses DOM have 'one-dimensional' DOM systems. For instance, as shown for Hebrew, definite direct objects receive obligatory case marking irrespective of their level of animacy. However, languages where animacy and definiteness overlap and both dimensions are taken as determining factors for the overt marking of DOM, such as in Hindi (McGregor 1972; Mohanan 1990), Spanish (e.g., von Heusinger & Kaiser 2003; de Swart and de Hoop 2007)[26] have 'two-dimensional' DOM systems. I characterize NGA as a 'two-dimensional' DOM system since DOM appears only with a subset of definite objects. This will be more investigated in Chapters 5-7.

Unlike in NGA, animacy-sensitive DOM languages distinguish animate from inanimate direct objects and human from non-human direct objects. For instance, in

[25] In the same vein, Caro Reina (2020) has observed that the proper names are not a homogeneous group and proposed a new version of the extended animacy scale. For NGA, only pronouns within this version are not allowed with DOM.

(i) Extended animacy hierarchy: first/second-person pronoun > third person pronoun > deity name > personal/kinship name > animal name > place name > human common noun> non-human animate common noun > inanimate common noun

[26] According to Heusinger (2008), more than two factors affect the appearance of DOM in Spanish. These factors include topicality, animacy, and definiteness, thus, based on Heusinger's discussion, Spanish cannot be classified under a two-dimensional system.

Malayalam, only direct objects with animate referents are marked by the accusative case, but direct objects with inanimate referents are caseless (Asher and Kumari 1997; de Swart 2007; de Swart and de Hoop 2007).

(30) *Malayalam* (Asher and Kumari 1997: 203)

a. Avan oru paʃuvin-e vaɲɲi
he a cow-ACC buy.PST
'He bought a cow.' DOM

b. ɲaan teeɲɲa vaɲɲi
I coconut buy.PST
'I bought a coconut.' Non-DOM

The contrast in (30) illustrates that the direct object 'cow' in (30a) bears accusative case whereas the inanimate direct object 'coconut' in (30b) does not.[27] Thus, in Malayalam only objects above the animate-inanimate cut-off point are marked with accusative case. In the following section I will discuss another attested property of DOM in NGA with relation to other languages.

2.2.2 Dative marking

As shown in section 2.1, differentially marked objects in NGA are marked with dative. We find the property of differential objects marked with dative in other Semitic languages (Bossong 1991; Döhla 2016), as well as non-Semitic languages such as Hindi (Butt 1993; Bhatt 2007; Manzini & Franco 2016; Bárány 2018; Manzini et al. 2020), Spanish (Bleam 2005; Ormazabal and Romero 2007, 2013; Rodríguez-Mondoñedo 2007; López 2012; Fábregas 2013), varieties of Basque (Odria 2014; von Heusinger et al. 2018), Guaraní (Zubizarreta & Pancheva 2017). Bárány (2018) includes the following examples from Spanish showing that the exponents of DOM and dative case are identical. The

[27] de Swart (2007: 88f.) shows that, only in certain contexts, direct objects with inanimate referents can receive accusative case in Malayalam in order to avoid misinterpretations.

example in (31a) shows a transitive clause with a morphologically unmarked direct object, the definite inanimate DP *el libro* 'the book.' In (31b), the definite animate direct object *la mujer* 'the woman' triggers the appearance of the DOM marker *a*. As (31c) demonstrates, a homophonous marker appears with the indirect object, a recipient, in a ditransitive clause.

(31) *Spanish* (Bárány 2018: 1)

a. Yo vevo el libro
I see the book
'I see the book.'

b. Yo vevo a la mujer
I see DOM the woman
'I see the woman.'

c. Yo le doy el libro a la mujer
I CL.DAT give the book DAT the woman
'I give the woman the book.'

It has been argued that in Spanish the dative DOM marker appears with a subset of direct objects (e.g., Aissen 2003; Leonetti 2004; López 2012; Zdrojewski & Sánchez 2014). Specifically, as shown in (31b), this marker is often said to be triggered by animacy and definiteness (Jaeggli 1982; Suñer 1988; Torrego 1998; Gutiérrez-Rexach 1999; Aissen 2003; López 2012; Irimia 2020, among many others).[28] While Spanish and NGA clearly differ with respect to animacy, they share the following properties: (i) dative differential marking, and (ii) definiteness. The next section discusses another shared property, namely Clitic Doubling.

2.2.3 Clitic Doubling

Clitic Doubling (CD) has been documented in certain DOM languages (Hill 2013; Kramer 2014; Kallulli 2016; Hill & Mardale 2019; Bárány & Kalin 2020; Irimia 2021). For

[28] There is a lot of variation across dialects of Spanish and the factors determining DOM are complex (see Jaeggli 1982; Torrego 1998, 2010; Gutiérrez-Rexach 1999; Aissen 2003; Leonetti 2004; Rodríguez-Mondoñedo 2007; Heusinger 2008; von Heusinger & Kaiser 2011; López 2012).

instance, in Modern Romanian, both CD and DOM must co-occur as demonstrated in (32). Consider the example in (32a): the pronoun *tine* 'you' in the accusative form is introduced by a DOM particle *pe* while also undergoing clitic doubling by means of the proclitic *te-* 'you,' which is also marked accusative.

(32) *Modern Romanian* (Hill & Mardale 2019: 2)

a. **Te**$_j$-am strigat **pe** tine$_j$
you.ACC=have.1= called DOM you.ACC
'I called you.'

b. *Am strigat pe tine. // *Te$_j$-am strigat *tine$_j$ //
have.1= called DOM you.ACC you.ACC =have.1= called you.ACC
*Am strigat tine.
have.1= called you.ACC

Contrasting with (32a), the ungrammatical sentences in (32b) illustrate not only that object marking is obligatory with this nominal category (i.e., personal pronouns), but also that neither CD nor DOM may fulfill the marking task on their own.

Similarly, Mardale and Karatsareas (2020) also show that DOM in Standard Romanian is related to CD in that CD is found only in contexts where DOM is also found. Further, Kallulli (2016) argues for Albanian, a Balkan Language, that dative/genitive objects and direct objects instantiated by local (i.e., 1st and 2nd person) full pronouns are consistently clitic doubled, hence supporting the view that CD behaves as a DOM strategy.

According to Kayne (1975), there seems to be a strong grammatical dependence between the CD and DOM mechanisms. Kayne (1975) pointed out that in Spanish and Romanian in order to have CD, DOM is required. This grammatical dependence between DOM and CD is known as Kayne's Generalization, which I reproduce in (33), from Jaeggli (1982):

(33) *Kayne's Generalization* (Jaeggli 1982: 20)

An object NP may be doubled by a clitic only if the NP is preceded by a preposition.

The idea behind the generalization is that the doubling clitic absorbs Case, so unless a preposition (or some other case-assigning device) could be inserted, the DP-argument

would remain caseless, and thus the Case Filter (Chomsky 1981) would cause the derivation to crash. Example (34) demonstrates this crash in a variety of Spanish where the generalization holds: CD only occurs when the direct object is preceded by the preposition *a* 'to,' thus dropping the DOM marker renders ungrammaticality.

(34) *Rioplatense Spanish* (Jaeggli 1986: 32)

Lo vi-mos *(a) Juan
ACC.3MS saw-1PL to Juan
'We saw Juan.'

The major problem is that there are cases that violate Kayne's Generalization. Modern Greek and Albanian do not obey Kayne's generalization as they allow clitic doubling without the presence of any preposition (Anagnostopoulou 2006 for Greek, Kallulli 2016 for Albanian). Anagnostopoulou (2006) presents counterexamples to Kayne's Generalization that in all Balkan languages that have CD (e.g., Bulgarian, Albanian, and Greek), CD is not dependent on the presence of a preposition. In fact, the presence of a preposition in Greek CD yields to ungrammaticality. This means not only that the generalization was a spurious one, but that any approach to CD based on the absorption of Case by clitics will fail to capture the Balkan CD data.

As far as NGA is concerned, as shown in section 2.1, when a clause contains the DOM marker, CD is obligatory. However, the reverse is not true, that is, DOM is not obligatory whenever a clause contains CD. We will see in Chapter 4 that in other configurations different from DOM, CD is obligatory but DOM is not and not even possible.

2.3 Conclusion

This chapter has discussed the DOM patterns in Arabic in relation to the cross-linguistic DOM facts and approaches. We will make reference to these properties throughout the thesis.

Chapter 3

The clausal syntax of DOM in NGA

3.1 Introduction and Background

Chapter 2 (section 2.1) reviewed the basic properties of DOM in Spoken Arabic. This chapter establishes and investigates in more detail the empirical generalization that DOM is an instantiation of dislocation (discussed briefly in chapter 2) and proposes a novel syntactic analysis for this structure.

The chapter examines the clausal syntax of the DOM construction in North Galilee Arabic (NGA) in three main parts. The first part advances the empirical generalization that DOM involves right dislocation. The second part explores the derivation of the dislocated DOM phrase and whether it reaches its position by means of syntactic movement or if it is base-generated in that position. Finally, the third part examines the structural height of the dislocated phrase in the clause structure.

The literature provides three competing analyses for the derivation of right dislocation: (i) base-generation, (ii) movement on the syntactic component, and (iii) movement on the level of phonological form (PF). On the one hand, some scholars argue that the dislocated element is externally merged in the C-domain (Cardinaletti 2002; Frascarelli 2004; De Cat 2002, 2007, among others). On the other hand, others contend that the displaced material is base-generated inside its host and somehow moved rightward (Ross 1967; Chomsky and Lasnik 1977; Fox & Nissenbaum 1999; de Vries 2002). Within the movement approach, rightward movement is treated as an operation that either targets the VP or *v*P (e.g., Bresnan 1976; Stowell 1981; Johnson 1985; Overfelt 2015) or targets the CP-layer (Vallduví 1992; Samek-Lodovici 2006, 2015). The last approach (iii) argues for locating the rightward movement in the PF component of the grammar (e.g., Truckenbrodt 1995; McCloskey 1999; Göbbel 2007).

In this chapter, I will argue for an analysis of DOM in NGA in which differentially marked objects are derived by movement and not by base-generation because DOM exhibits island sensitivity which is a hallmark of movement. I argue that the directionality of the DOM movement is rightward. This directionality comes from a combination between interactions of Right Roof Constraint (Ross 1967; Grosu 1973), islands, and the Right Edge Restriction (e.g., Hartmann 2000; Sabbagh 2007; Bachrach & Katzir 2009). I argue that the right-dislocated phrase i.e., the DOM phrase, targets the edge of *v*P. Specifically, it moves to a specifier of a functional projection XP at the edge of *v*P. I will elaborate on the identity of XP in Chapter 4.

This chapter is structured as follows. Section 3.2 advances an empirical generalization that DOM involves dislocation. In section 3.3, I argue that DOM is derived by movement and the directionality of this movement is rightward. Section 3.4 argues that DOM rightward movement is an instance of A-movement. In section 3.5, I use adverbial placement test and binding theory to determine the structural height of the moved differentially marked object, namely, whether the dislocated phrase targets the C-area or a clause-internal position. The results of these tests suggest that the differentially marked object targets the edge of *v*P. I also examine the landing position of the dislocated element by considering the behavior of the clitic, namely, whether the clitic is an independent pronoun with its own referential index, hence entering a binding relation. Alternatively, if

the clitic is part of an A-chain with the *la*-phrase, it will not get its own referential index and thus, cannot enter a binding relation. When clitics are independent pronouns, I argue that all the binding data are correctly accounted for under the clause-internal hypothesis, thereby making the analysis well-suited for the derivation of DOM in Arabic. Section 3.6 addresses the question of whether DOM differs from other similar phenomena such as Clitic Doubling (CD) and Clitic Right Dislocation (CLRD). I argue that DOM must be treated separately as it shares only some properties with the other phenomena. Section 3.7 summarizes the syntactic properties of DOM discussed in this chapter and will lay the foundation for Chapter 4.

3.2 DOM as dislocation

The goal of this section is to establish that DOM is an instance of dislocation. Section 2.1 briefly demonstrated that DOM involves dislocation in NGA. This section discusses this property in more detail. Brustad (2000) observes that the DOM marker *la-* in Syrian Arabic marks certain dislocated objects.[29] She uses the term right dislocation, as described by Lambrecht (2001) and Shaer et al. (2009), to refer to a construction in which the *la*-marked object is dislocated to the right edge of a clause which typically contains a clitic.[30] However, Brustad does not provide data showing that there is indeed a dislocation; the examples that Brustad (2000: 353-358) reports contain the differentially marked object to be immediately adjacent to the verb, thus it could be that the linear relation between the differentially marked object and the gap associated with it is string vacuous. Generally speaking, the differentially marked object is compatible with in-situ analysis. I offer evidence that DOM involves dislocation based on adverb ordering with differentially marked objects, but first I demonstrate the baseline order with non-differentially marked objects.

[29] Brustad (2000) investigates four Arabic dialects (Moroccan, Egyptian, Syrian and Kuwaiti Arabic) and observes that only Syrian Arabic uses *la-* to mark dislocated objects. The Syrian data provided in Brustad (2000) hold true of NGA.

[30] Brustad (2000: 353) calls the pronominal element a resumptive pronoun.

The examples in (35) and (36) illustrate that the non-differentially marked objects must appear in their canonical position (in-situ) of SVO word order. In this section I examine two types of adverbs: frequency adverb and locative adverb. Other sets of adverbs are discussed in section 3.5.1 for the purpose of the structural height of the differentially marked object.

Adverbs such as ‘very’ and ‘in front of’ cannot intervene between the verb and the unmarked object, as shown in (35a) and (36a), respectively. The following adverbs are chosen randomly. The grammatical placement of the two adverbials is given in (35a’) and (36a’).[31]

(35) *Unmarked objects with degree adverbials*

a. ma-g'dert-š ʔfham *ktiir aḥmad
NEG-able-NEG understand.1SG very Ahmad
'I can't completely understand Ahmad.'

a’. ma-g'dert-š ʔfham aḥmad ktiir
NEG-able-NEG understand.1SG Ahmad very
'I can't completely understand Ahmad.'

(36) *Unmarked objects with locative adverbials*

a. šoft *fi-l-madrse sara
saw.1SG inside-the-school Sara
‘I saw Sara inside the school.’

a’. šoft sara fi-l-madrse
saw.1SG Sara inside-the-school
‘I saw Sara inside the school.’

Thus, non-differentially marked objects exhibit the following order summarized in (37), providing evidence that the unmarked object is analyzed in-situ.

[31] For the sake of simplicity, I only present dislocation with transitive clauses where the target object is a DP.

(37) *Word order with non-DOM*

a. V *Adv DP-OBJ

b. V DP-OBJ Adv

In order to determine whether the differentially marked object counterparts undergo obligatory or optional dislocation, we need to look at the distribution of the differentially marked object in relation to adverb placement (frequency and locative). The orders are demonstrated in (38).

(38) *Word order with DOM*

a. V Adv DOM-OBJ

b. V DOM-OBJ *Adv

The first order shows that frequency and locative adverbs intervene between the verb and the differentially marked object, as illustrated in (39) below. It cannot be concluded that the differentially marked object undergoes obligatory dislocation; the examples in (39) are also compatible with optional dislocation.

(39) a. ma-g'dert-š ʔfham-o$_i$ ktiir la-aḥmad$_i$

NEG-able-NEG understand.1SG-3M.SG.OBJ very DOM-Ahmad

'I can't completely understand Ahmad.'

b. šoft-a$_i$ fi-l-madrse la-sara$_i$

saw-3F.SG.OBJ inside-the-school DOM-Sara

‘I saw Sara inside the school.’

While these two types of adverbials can intervene between the verb and the differentially marked object, they cannot follow the differentially marked object, as exemplified in (40). The distribution of DOM with adverbials in (40) strongly suggests that differentially marked objects in NGA undergo obligatory dislocation.

(40) a. ma-g'dert-š ʔfham-o$_i$ la-aḥmad$_i$ *ktiir
NEG-able-NEG understand.1SG-3M.SG.OBJ DOM-Ahmad very
'I can't completely understand Ahmad.'

b. šoft-a$_i$ la-sara$_i$ */?? fi-l-madrse
saw-3F.SG.OBJ DOM-Sara inside the-school
'I saw Sara inside the school.'

Recall from Chapter 2 that all DOM examples, including (39)-(40) obligatorily involve clitic doubling; namely, the right-dislocated DP must co-occur with a clitic.[32] In addition, co-referentiality must be established between the clitic and the dislocated DP; that is, the dislocated DP is obligatorily co-indexed with the clitic.

The behavior of differentially marked and non-differentially marked objects with respect to adverbial placement is summarized by the schematics in (41).

(41) a. **Non-DOM**: V *Adv DP-OBJ Adv
b. **DOM**: V Adv DP-OBJ *Adv

So far, we conclude that while non-differentially marked objects appear in their canonical position (in-situ), the differentially marked objects are obligatorily dislocated. The next section addresses the question of how we can account for dislocation in the NGA DOM construction.

[32] In NGA, CD is obligatory with dislocation (see examples in chapter 4.2). The obligatoriness or optionality of CD is subject to debate among scholars. It has been pointed out by Samek-Lodovici (2015), following Benincà et al. (1988) and Frascarelli and Hinterhölzl (2007), that CD of right-dislocated elements is not obligatory. His claim does not align with Cardinaletti (2002), who claims that CD is obligatory, as well as Cecchetto (1999: 65) and Cruschina (2010), who maintain that any case of right dislocation not involving CD is only apparent, actually involving marginalization.

3.3 DOM as rightward movement

In the previous section I established that DOM is an instance of dislocation. The pertinent question of this section is whether the dislocated phrase reaches its position by means of syntactic movement or is base-generated in that position.

The two accounts (movement and base-generation) vary in the predictions that they make. I assume that the differentially marked dislocated phrase originates in VP since it is connected to the verb by syntactic and semantic selection (so-called c- and s-selection), see section 3.5 for a detailed syntactic analysis. If the differentially marked object moves from this position, it is expected to display locality constraints which might involve islands and therefore be subject to island constraints. Alternatively, under the base-generation approach, locality constraints are not predicted. I argue that DOM is derived by movement because it displays island sensitivity. Further, I provide evidence that the directionality of DOM movement is rightward, which comes from a combination between the interactions of Right Roof Constraint (RRC), islands and the Right Edge Restriction.

Rightward movement is subject to RRC (Ross 1967; Grosu 1973) which states that rightward movement is clause bounded. Rightward movement must be applied to a single clause, as shown by the contrast between (42a) and (42b).[33]

(42) *Rightward movement is subject to RRC*

a. Sara bought t_i, yesterday, [the new car]$_i$.

b. *Sara said [CP that John likes t_i] yesterday, [the new car]$_i$.

In NGA, DOM is clause-bounded as it obeys RRC as shown in (43).[34]

[33] Subsequent research has gradually streamlined the locality conditions on rightward movement. Sabbagh (2007: 351(3)) adapted the Right Roof Constraint:

(i) Rightward movement may move and right-adjoin an element X to the cyclic node in which X is merged, but no further.

In other words, rightward movement of some object is bound to the edge of the first cyclic node that dominates that element (e.g., Akmajian 1975; Baltin 1981). For Ross (1967), the only cyclic node was CP, but for other scholars *v*P is a cyclic node (e.g., Chomsky 2000, 2001; Baltin 1981).

[34] The fact that DOM in NGA is subject to RRC is not limited to Arabic but to other languages involving dislocation. Kayne (1994), on the basis of French data, argues that right dislocation obeys RRC. Cecchetto (1999: 7) adds that RRC constitutes a discrepancy between right dislocation and other clitic construction, CLLD. A constituent can be long-distance left dislocated whereas right dislocation can displace a constituent

(43) a. *sara qal-at [CP ʔenno aḥmad yokol-a$_i$] mbirḥ [la-l-kaʕk-e]$_i$
Sara said-3F.SG that Ahmad eat-3F.SG.OBJ yesterday DOM-the-cake-F.SG
(Lit. 'Sara said that Ahmad eats yesterday the cake.')
Intended: 'Sara said that Ahmad eats the cake yesterday.'

b. *mntwaqqaʕ [CP ʔenno nlaqi-ha$_i$] bʕed asobʕ [la-l-ʔswar-a]$_i$
expect.1PL.FUT that find.1PL-3F.SG.OBJ next week DOM-the-bracelet-F.SG
(Lit. 'We expect that we find next week the bracelet.')
Intended: '(We) expect that we (will) find the bracelet next week.'

Unless it is otherwise indicated, none of the Arabic examples in this chapter include a pause, (see example (75) in section 3.6).

Given that rightward movement is clause-bounded in NGA, we can therefore only use tests where the relevant movement does not cross clause boundaries.[35] I use the Coordinate Structure Constraint (CSC) (Ross 1967). The general consensus appears to be that rightward movement obeys the CSC (see e.g., Wexler and Culicover 1980; McCawley 1982; Postal 1998; Johnson 2007; Sabbagh 2007; Citko 2017). There are two parts to CSC; one prevents movement of an entire conjunct, and the other prevents movement of any element contained within one conjunct. The examples below demonstrate that both are operative in DOM. The examples involve a coordinate structure consisting of two *v*Ps (44) or two DPs (45).[36]

(44) a. *sara [*v*P ḥ̲d̲d̲art-a$_i$ t_i o zayyant l-bit] mbirḥ
Sara prepared-3F.SG.OBJ and decorated the-house yesterday
[la-l-kaʕk-e]$_i$

only to the immediate periphery of the clause in which it originates. This discrepancy is relevant for NGA and will be elaborated in Chapter 4.

[35] In the domain of DOM, as Kalin and Weisser (2019) note, many islands are not possible to test since the alleged movement step is very short.

[36] Sabbagh (2007) argues that the ungrammaticality of CSC with Right Node Raising (RRC) cannot be explained under an ellipsis account. He noticed that there are some issues with this account. Based on the ellipsis account, we predict RNR to behave similarly to other types of ellipsis, such as VP ellipsis; however, it has been shown that the identity requirement is different in VP ellipsis. Moreover, the ungrammaticality of CSC examples cannot be accounted for by backward deletion as all coordination with RNR will need to be excluded. Thus, I opt to examine the movement account in this thesis.

DOM-the-cake-F.SG

(Lit. ‘Sara prepared it and decorated the house yesterday the cake.’)

Intended: ‘Sara prepared the cake and decorated the house yesterday.’

b. *aḥmad [*v*P ʔjjar-o$_i$ t_i o baʕ s-sayyara] mbirḥ [la-l-bit]$_i$

Ahmad rented-3M.SG.OBJ and sold the-car yesterday DOM-the-house

(Lit. ‘Ahmad rented it and sold the car yesterday the house.’)

Intended: ‘Ahmad rented the house and sold the car yesterday.’

(45) a. *sara zayyant-on [DP t_i o l-ʔoḏa] mbirḥ [la-l-kaʕk-e]$_i$

Sara decorated-3PL.OBJ and the-room yesterday DOM-the-cake-F.SG

(Lit. ‘Sara decorated them and the room yesterday the cake.’)

Intended: ‘Sara decorated the cake and the room yesterday.’

b. *aḥmad šara-hon [DP t_i o s-sayyara] mbirḥ [la-l-bit]$_i$

Ahmad bought-3PL.OBJ and the-car yesterday DOM-the-house

(Lit. ‘Ahmad bought them and the car yesterday the house.’

Intended: ‘Ahmad bought the car and the house yesterday.’

The examples in (44)-(45) demonstrate that DOM in NGA is subject to islands.[37,38]

One of the defining properties of rightward movement is the so-called Right Edge Restriction.[39] This restriction requires the displaced element to be the rightmost element in

[37] An additional test used to detect movement, suggested by Idan Landau, is DP sub-extraction. If DPs are islands, we predict that no element can be sub-extracted from DP. However, this test cannot be constructed with DOM as the examples are ill-formed. The only way we can extract the differentially marked *la*-DP is when the *la*-DP behaves as a possessor dative. Thus, the grammaticality of the following example is due to a different construction (possessor dative) than DOM.

(i) sara aklat-a [DP la-kaʕk-t-o t_i] ʕl-axer [la-kamal]$_i$

Sara ate-3F.SG.OBJ DOM-cake-CS-3M.SG.OBJ on-end DAT-Kamal

‘Sara ate Kamal’s cake completely.’

[38] See appendix E for data on coordination with DOM.

[39] Other terms have been used in the literature to refer to this restriction including Right Edge Generalization (Abels 2004), Right Edge Effect (Johnson 2007), or Right Edge Condition (Wilder 1999). I use ‘Right edge restriction’ in this thesis.

each conjunct before rightward movement can apply (Postal 1974; Wilder 1995, 1999; Hartmann 2000; Johnson 2007; Sabbagh 2007; Kluck & de Vries 2013).

(46) *Right Edge Restriction* (Sabbagh 2007: 356)
In the configuration: [A... X ...] Conj. [B ... X ...], X must be rightmost within A and B before X can undergo rightward movement.

This constraint explains the contrast in (47a)-(47b). The gap position within each conjunct in (47a) is rightmost, meaning that the Right Edge Restriction can be satisfied, and rightward movement is licensed. However, in (47b) the Right Edge Restriction is not satisfied due to the presence of 'to Mary,' thereby rendering ungrammaticality.

(47) a. Sara read ___ , and reviewed ___ , [a book].
b. *Sara read ___, and gave___ to Mary, [a book].

It has been claimed that rightward movement is able to feed the Right Edge Restriction (Wilder 1999; Sabbagh 2007; Kluck & de Vries 2013). Therefore, since DOM in NGA is derived by rightward movement, we predict that it feeds the Right Edge Restriction. This prediction is borne out as evidenced by the contrast between (48a) and (48b).

(48) a. sara katbat-o$_i$ t_i o baʕat-o$_i$ t_i haḏik lseni [la-l-ktaab]$_i$
Sara wrote-3M.SG.OBJ and sold-3M.SG.OBJ last year DOM-the-book.M.SG
Lit. 'Sara wrote it and sold it last year, the book.'
'Sara wrote and sold the book last year.'

b. * sara katbat-o$_i$ t_i o aʕṭat-o$_i$ t_i la-dalia haḏik lseni
Sara wrote-3M.SG.OBJ and gave-3M.SG.OBJ DAT(to)-Dalia last year
[la-l-ktaab]$_i$
DOM-the-book.M.SG
(Lit. 'Sara wrote it and gave it to Dalia last year, the book.')
Intended: 'Sara wrote the book and gave it to Dalia last year.'

The fact that the differentially marked object occupies the rightmost position inside the two conjuncts is reminiscent of across-the-board (ATB) movement. By way of illustration,

consider the following examples involving an extraction of the rightmost differentially marked object from both conjuncts simultaneously.

(49) a. sara [*v*P ḥḏḏart-a$_i$ t_i o zayyant-a$_i$ t_i] mbirḥ [la-l-kaʕk-e]$_i$
Sara prepared-3F.SG.OBJ and decorated-3F.SG.OBJ yesterday DOM-the-cake-F.SG
Lit. 'Sara prepared it and decorated it yesterday, the cake,'
'Sara prepared and decorated the cake yesterday.'

b. aḥmad [*v*P katab-o$_i$ t_i o ḥallo-o$_i$ t_i] mbirḥ [la-s-soʔal]$_i$
Ahmad wrote-3M.SG.OBJ and solved-3M.SG.OBJ yesterday DOM-the-question
Lit. 'Ahmad wrote it and solved it yesterday, the question.'
'Ahmad wrote and solved the question yesterday.'

The fact that ungrammatical CSC examples (44)-(45) become grammatical when movement takes place in an ATB-movement fashion from both conjuncts (*v*Ps) simultaneously and targets the rightmost element, further supports the argument that DOM as rightward movement obeys CSC.[40]

Under this proposal, the movement cannot involve a PF component. The Y-model of grammar, an architectural representation of how grammar is perceived in the generative theory (Chomsky 1995, 2013), leads us to expect that movement in the narrow syntax may in principle feed a semantic interpretation. If the movement takes place in the PF component of the grammar, it should be irrelevant to the semantic module. However, as will be shown in Chapter 5, the DOM movement feeds a semantic interpretation, thus it must be a syntactic movement.

This section demonstrated that DOM manifests locality constraints on movement (i.e., island tests), as well as the locality constraints on rightward movement (i.e., RRC and Right Edge Restriction). Before turning to the structural question of where the rightward

[40] Another defining property of rightward movement is that it resists P-stranding (see Ross 1967; Wexler and Culicover 1980; Pesetsky 1995; Baltin and Postal 1996; Sabbagh 2007), this test cannot be applied to Arabic because Arabic generally disallows P-stranding (Albukhari 2016 on Jordanian Arabic; Algryani 2017 on MSA).

(i) *min ḥak-at sara maʕ?
Who talked-3F.SG Sara with
('Who did Sara talk with?')

moved element appears in the syntactic structure, I argue that DOM rightward movement in NGA is an instance of A-movement.

3.4 DOM as A-movement

I argue that DOM is an instance of A-movement. I provide three diagnostics to determine what kind of movement relates the position of the clitic and the position of its dislocated phrase. These diagnostics include: (i) reconstruction effects, (ii) quantifier stranding, and (iii) idiomatic interpretation.

The first argument comes from reconstruction effects. While A-bar movement obligatorily reconstructs, A-movement is not required to do so (e.g., Chomsky 1993, 1995; Fox 1999; Lebeaux 2009). If DOM involves A-movement, we should expect that it can but does not have to reconstruct. This prediction will be confirmed in section 3.5.2 where I argue that DOM does not display reconstruction effects.[41]

Further, the behavior of A- and A'-movement is contrasted with respect to quantifier stranding: only the former kind of movement appears to license stranded quantifiers (Déprez 1989; McCloskey 2000; Bobaljik 2003). Quantifier float (Q-float) has been documented in Arabic (Benmamoun 1999; Al Khalaf 2019). Before turning to DOM data, I show that NGA matches the expected profile of Q-float in other constructions (e.g., Sportiche 1988). In the examples (50) and (51), both nominals in subject and object positions may float. The quantifier and its associate must agree in number, which is manifested by the plural clitic *-on*.[42]

(50) *Q-float in NGA*

a. kol ṭ-ṭollab sallam-o al-waẓife

[41] Unlike DOM, reconstruction has been argued to be a characteristic of CLLD in Arabic (Aoun and Benmamoun 1998; Aoun, Choueiri and Hornstein 2001; Aoun, Benmamoun and Choueiri 2010; Alzayid 2022). If reconstruction is a diagnostic of A'-movement, then the operation which yields CLLD should be an instantiation of A'-movement.

[42] The presence of the obligatory clitic associating with the floating quantifier has been discussed in Arabic works (e.g., Benmamoun 1999; Al Khalaf 2019). Under the analysis of the big-DP adopted in this thesis, it is unclear where the clitic on the floating Q appears in the big-DP since the other clitic is already occupied. I leave a more careful exploration of the facts for future research.

all the-students.BP submitted-3PL the-assignment
'All the students submitted the assignment.'

b. ṭ-ṭollab sallam-o koll-on al-waẓife
the-students.BP submitted-3PL all-3PL the-assignment
'The students all submitted the assignment.'

(51) a. sami qara kol al-kotob bi-l-ʕoṭle
Sami read all the-books.BP in-the-vacation
'Sami read all the books in the vacation.'

b. sami qara al-kotob koll-on bi-l-ʕoṭle
Sami read the-books.BP all-3PL in-the-vacation
'Sami read the books all in the vacation.'

In addition to the data in (50)-(51), the DOM construction also has Q-floating.

(52) *DOM allows Q-float*

a. qrit-on la-l-kotob koll-on
read.1SG-3PL.OBJ DOM-the-books.BP all-3PL.OBJ
'I read all the books.'

b. ḥayyakt-on la-l-jarasi koll-on
knitted.1SG-3PL.OBJ DOM-the-sweaters.BP all-3PL.OBJ
'I knitted all the sweaters.'

Therefore, given that A-movement licenses floating quantifiers, the empirical facts above provide further evidence in favor of DOM as A-movement.

Further evidence in support of treating the relation between the clitic and its differentially marked object as an instance of A-movement comes from idiomatic interpretation. It is generally assumed that A-movement retains the figurative meaning of idioms, whereas A'-movement does not. The following example shows that DOM retains

idiomatic reading,[43] suggesting that DOM is A-movement, as shown in (53b) and (54b). However, the idiomatic interpretation is not maintained with Clitic Left Dislocation, (53c) and (54c), which is argued to be A'-movement in Arabic (e.g., Alzayid 2022).

[Context: A friend has recently been getting angry at things happening at home. His friend suggested calming down, so he said:]

(53) a. ʔmsek ʔaʕsˁab-ak BASELINE
Hold.IMP nerves-your
Lit. 'Hold your nerves!'
Idiomatic reading: 'Calm down!'

[Context: A friend is *consistently* getting angry at things happening at home. His friend said, 'that's enough' and he suggests calming down so he said:]

b. ʔmsek-on la-ʔaʕsˁab-ak A-MOVEMENT
Hold.IMP-3PL.OBJ DOM-nerves-your
Lit. 'Hold them your nerves!'
Idiomatic reading: 'Calm down!'

c. * ʔaʕsˁab-ak ʔmsek-on A-BAR MOVEMENT
nerves-your hold.IMP-3PL.OBJ
(Lit. 'Your nerves, hold them!')
Intended idiomatic reading: 'Calm down!'

[Context: A friend is fighting with his friend. He said to him:]

(54) a. balleṭ al-baḥer BASELINE
Pave.IMP the-sea
Lit. 'Pave the sea!'

[43] Unlike DOM, Lyassi (2012) observes for Standard Arabic that idiom interpretation is not preserved in HTLD constructions.

Idiomatic reading: ‘Go I don’t care!’

[Context: A friend is repeatedly having a fight with his friend, and he reaches a point he starts not to care, so he says:]

b. balṭ-o la-l-baḥer A-MOVEMENT
Pave.IMP-3M.SG.OBJ DOM-the-sea
Lit. ‘Pave it the sea!’
Idiomatic reading: ‘Go I don’t care!’

c. * al-baḥer balṭ-o A-BAR MOVEMENT
the-sea pave.IMP-3M.SG.OBJ
Lit. ‘The sea, pave it!’
Intended idiomatic reading: ‘Go I don’t care!’

The three pieces of evidence demonstrated above point to the conclusion that DOM in NGA has the properties of A-movement.

The analysis that DOM has the properties of A-movement is supported crosslinguistically. As will be shown in section 3.6, DOM in NGA shares some properties with CD. Cross-linguistically, CD shows evidence of A-movement (see Alexiadou and Anagnostopoulou 1997, 2000; Anagnostopoulou 2014; Angelopoulos 2019 for Greek; Harizanov 2011, 2014 for Bulgarian; Kramer 2014 for Amharic, among others).

3.5 Examining the structural position of the dislocated DP

I argued in previous sections that DOM is derived by rightward movement and DOM movement patterns with A-movement. The purpose of this section is to examine the structural position of the moved differentially marked object i.e., the *la*-DP. I use two tests: adverb placement (section 3.5.1) and binding data (section 3.5.2).

After ruling out the possibility that dislocated DOM phrases in NGA are base-generated (shown in section 3.3), we are left with two potential landing sites for the *la*-phrase. First, some authors argue that dislocated phrases move to the right periphery of the

clause, specifically to C-domain (Vallduví 1992; Cardinaletti 2002; Frascarelli 2004; Samek-Lodovici 2006, 2015). In contrast to the peripheral approach, others contend that the rightward moved material targets a clause-internal position (see Bresnan 1976; Stowell 1981; Johnson 1985). Within the clause-internal analysis, the rightward constituent targets an intermediate position above *v*P but lower than T, an area which is typically referred to as the middle field,[44] as proposed by Villalba (1998, 2000) and López (2003, 2009) for Catalan.[45]

Cecchetto (1999) and Villalba (2000) argue for Italian and Catalan respectively, that the clause-internal area is associated with an information-structural property; it hosts dislocated topics that raise to the specifier of an intermediate topic projection located between the VP and TP projection. Outside of Romance languages, Halpert and Zeller (2015) propose that in Zulu, a Bantu language, the dislocated DP raises to a *v*P-external position, namely, the right-branching specifier of a functional category above *v*P.[46]

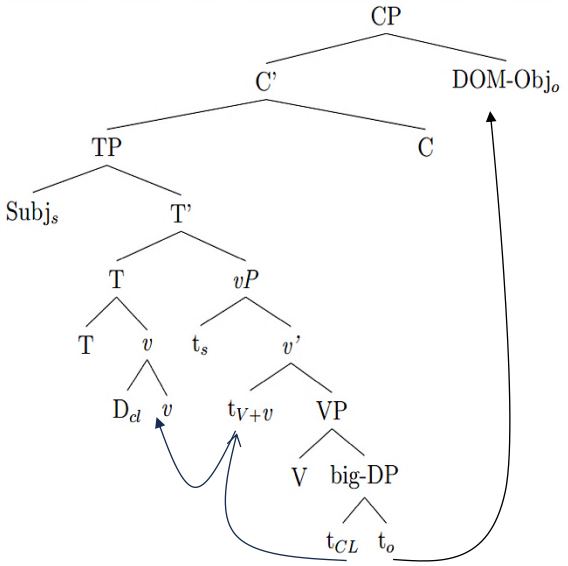
The syntactic representations for the right periphery and clause-internal approaches are shown in (55) and (56). They follow two different hypotheses: in Hypothesis A, shown in (55), the dislocated phrase moves to the right-peripheral position, whereas in Hypothesis B, shown in (56), the dislocated phrase moves to a clause-internal position at the edge of *v*P.

[44] Middle field (so called *mittelfeld*) is a descriptive term employed in V2 languages such as German and Dutch that corresponds to the area comprised between the auxiliary verb and the past participle at the right periphery of the clause.

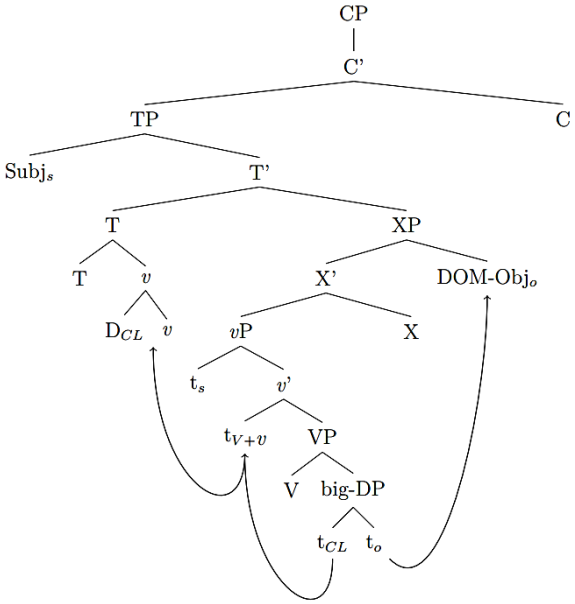
[45] Within the clause-internal analysis, there is another competing position that the right-dislocated element targets a position between *v*P and VP, as proposed by Cecchetto (1999) for Italian. The proposal argues for the first position, not the latter.

[46] In Zulu, this position is also linked to information structure; it is associated only with given constituents and not focused constituents (Buell 2005; Halpert & Zeller 2015, among others).

(55) *Hypothesis A*



(56) *Hypothesis B*



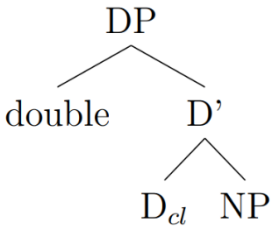
Below, I test which of these hypotheses correctly predicts the adverb placement and binding data in NGA. The differentially marked objects targets some functional projection XP at the edge of *v*P (56), I locate the 'dislocated phrase' as a specifier of XP. Before analyzing the DOM data, I first need to establish several theoretical assumptions.

The movement-based approach of DOM raises the question of the structure of the differentially marked object DP that the clitic moves out of. I assume the dislocated DP and the clitic enter the derivation together in a big-DP configuration (see (57)).[47] There are different versions of the big-DP analyses (Kayne 1972; Torrego 1998; Uriagereka 1995; Belletti 1999; Papangeli 2000; Rezac 2008a; Nevins 2011, among others). Some analyses take clitics to be the head of the big-DP (Uriagereka 1995: 81), others treat clitics as adjuncts to DP (Nevins 2011), and yet others embed clitics as specifiers within a functional projection that also hosts the DP (Arregi & Nevins 2012).[48] I assume the following structure

[47] See Paparounas and Salzmann (2023) for an argument against the big-DP analysis for CD.
[48] Another option is to view the clitic as an agreement marker (see e.g., Borer 1984; Suñer 1988; Sportiche 1996; Anderson 2005). This type of argument does not hold for Arabic; see appendix A on the distinction between CD and object agreement in Arabic.

for the big-DP where the clitic is taken to be the head of DP and the associate (the double) is generated as the specifier of DP.

(57) *The structure of the Big-DP*



I consider clitics as D heads that strand the DP in the course of the derivation by moving to a verbal functional head.[49] Since the verb moves to T in Arabic (Mohammad 1990; Ouhalla 1991; Fassi Fehri 1993; Shlonsky 1997, among others), the clitic must move to T as well (e.g., Anagnostopoulou 2003).[50]

Following Rezac (2008a) and Roberts (2010), among others, I take CD to be an instance of head movement – specifically, head movement of D. I assume that all movement depends on a prior ϕ-agreement relation (Chomsky 2000, 2001). On this view, I assume that syntactic movement involves two steps: (i) the verb enters into an Agree relation with the clitic, subsequently (ii) the clitic moves to *v* and it ends up as part of a complex verbal head including *v*, then the complex *v* undergoes head movement to T.

Finally, I assume that preverbal subjects move to Spec,TP since the unmarked word order in NGA and other Arabic dialects, as mentioned in Chapter 1, is SVO (Fassi Fehri 1993; Aoun et al. 1994; Mohammad 2000; Benmamoun 2000; Announi 2021, among others).[51]

[49] Based on the formal similarities between doubled clitics and definite determiners discussed in clitic doubling literature, doubled clitics are D's (Uriagereka 1995; Bleam 1999; Anagnostopoulou 2003, among others).

[50] Under the movement approach, clitics are D heads that move from within the DP to a verbal functional head. The identity of the verbal functional head differs depending on the analysis and language under investigation, for instance, T (Anagnostopoulou 2003) and *v* (Nevins 2010).

[51] However, in Modern Standard Arabic, the unmarked word order is, as widely assumed in the related literature, VSO word order (see Bakir 1980; El-Yasin 1985; Fassi Fehri 1993; Jarrah 2020, among many others).

3.5.1 Adverbial placement test

The adverbial placement test is one tool that I use to diagnose the structural height of the rightward differentially marked object. The adverbial placement test has been used in Johnson (1985) for determining the locus of rightward movement. I argue below that the adverbial placement data are compatible with Hypothesis B (targeting the edge of the *v*P phase) but incompatible with Hypothesis A (clause-external).

I assume that adverbs attach at different positions depending on their interpretation. I test two types of adverbs: low adverbs (degree and locative) which I assume are adjoined to *v*P and high adverbs (subject-oriented and temporal) which we assume are adjoined to TP (e.g., Thompson 1996; Cinque 1999; Jarrah 2017). In order to determine the structural height of the differentially marked object, we compare the adverbial distributional facts between differentially marked objects and their counterparts (unmarked).[52] If differentially marked objects are structurally higher than the unmarked ones, we expect different distributional results with low adverbs. The results of the adverbial placement test indicate that differentially marked objects can cross low adverbials that adjoin to *v*P but cannot cross high adverbials. Thus, differentially marked objects are compatible with Hypothesis B (the edge of *v*P) but not with Hypothesis A (CP-position).

I begin with the low adverbs (degree and locatives modifiers). I demonstrate below that the distribution of low adverbs with DOM is different than their counterparts (unmarked DPs). Consider the following data containing degree adverbs (58), and locative adverbs (59), respectively. The adverbs are in round brackets and the objects are in square brackets.

(58) *Degree adverbs with non-DOM vs. DOM*

a. baʕref t_i (*ktir) [sara]$_i$ (ktir)
know.1SG well Sara well
'I know Sara well.' Non-DOM

[52] Thanks to the CLA audience (2023) for the discussion of this diagnosis.

b. baʕref-a$_i$ t_i (ktir) [la-sara]$_i$ (*ktir)
know.1SG-3F.SG.OBJ well DOM-Sara well
'I know Sara well.' DOM

(59) *Locative adverbs with non-DOM vs. DOM*

a. sara šarat t_i (*be-ṭ-ṭariq ʕa-l-bit) [l-kaʕk-e]$_i$ (be-ṭ-ṭariq ʕa-l-bit)
Sara bought in-the-way to-the-house the-cake-F.SG in-the-way to-the-house
'Sara bought the cake on the way home.' Non-DOM

b. sara šarat-a t_i (be-ṭ-ṭariq ʕa-l-bit) [la-l-kaʕk-e]$_i$ (*/??be-ṭ-ṭariq
Sara bought-3F.SG.OBJ in-the-way to-the-house DOM-the-cake-F.SG in-the-way
ʕa-l-bit)
to-the-house
'Sara bought the cake on the way home.' DOM

While unmarked objects cannot cross over a low adverb such as *ktir* 'well' or *be-ṭ-ṭariq ʕa-l-bit* 'on the way home' which are adjoined to *v*P, differentially marked objects *must* cross over such *v*P level adverbs.

In contrast, we have the same distribution of differentially marked and unmarked objects with high modifiers e.g., subject-oriented (60) and temporal modifiers (61). The examples in (60)-(61) contain a subject-oriented adverb *be-taraddod* 'reluctantly' and temporal adverb *mbirḥ* 'yesterday,' and a configuration in which the differentially marked and unmarked object would need to move to a position higher than TP.

(60) *Subject-oriented modifiers with non-DOM vs. DOM*

a. sara šarat t_i (*bi-taraddod) [l-bit be-nosˁ al-madine] (bi-taraddod)
Sara bought in-reluctance [the-house in-middle the-city]$_i$ in-reluctance
('Sara reluctantly bought the house which is in the middle of the city.'
Non-DOM

b. sara šarat-o t_i (*bi-taraddod) [la-l-bit be-nosˁ al-madine]
Sara bought-3M.SG.OBJ in-reluctance [DOM-the-house.M.SG in-middle the-city]$_i$

(bi-taraddod)
in-reluctance
('Sara reluctantly bought the house which is in the middle of the city.')
DOM

(61) *Temporal modifiers with non-DOM vs. DOM*

a. sara qal-at [CP ʔenno aḥmad akal t_i] (*mbirḥ) [l-kaʕk-e]$_i$ (mbirḥ)
Sara said-3F.SG that Ahmad ate yesterday the-cake-F.SG yesterday
('Sara said that Ahmad ate the cake yesterday.')
Non-DOM

b. sara qal-at [CP ʔenno aḥmad akal-a$_i$ t_i] (*mbirḥ) [la-l-kaʕk-e]$_i$
Sara said-3F.SG that Ahmad ate-3F.SG.OBJ yesterday DOM-the-cake-F.SG
(mbirḥ)
yesterday
('Sara said that Ahmad ate the cake yesterday.')
DOM

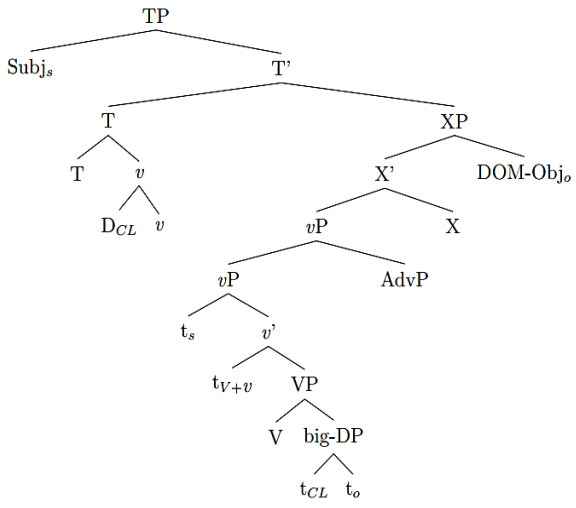
The data suggest that the ungrammaticality of the examples i.e., when the adverb is sandwiched between the verb and the object, must be attributed to the argument that the object is unable to cross high adverbs. A summary of the distribution of adverbials with both DOM and non-DOM objects is presented in Table 3.1.

	DOM	**Non-DOM**
DEGREE MODIFIERS	V Adv Obj V Obj *Adv	V *Adv Obj V Obj Adv
LOCATIVE MODIFIERS	V Adv Obj V Obj *Adv	V *Adv Obj V Obj Adv
SUBJECT-ORIENTED MODIFIERS	V *Adv Obj V Obj Adv	V *Adv Obj V Obj Adv
TEMPORAL MODIFIERS	V *Adv Obj V Obj Adv	V *Adv Obj V Obj Adv

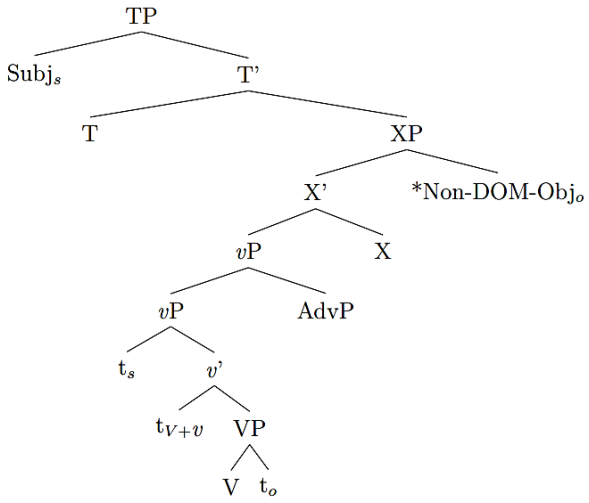
Table 3.1: The distribution of DOM and non-DOM with adverbials

Based on the adverbial placement test, I argue that DOM targets a position higher than *v*P-layer, thereby supporting Hypothesis B. The data also provide evidence that unmarked objects stay lower than differentially marked objects. The syntactic structures corresponding to DOM and non-DOM constructions with low and high adverbials are represented in (62)-(63).

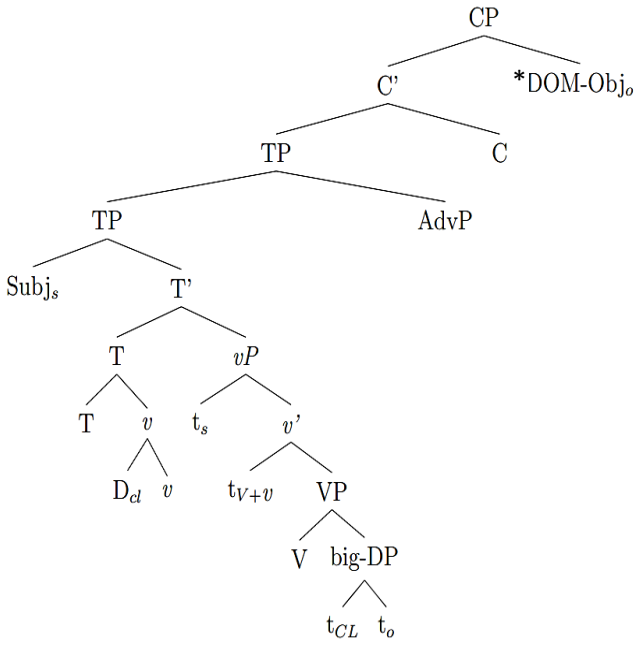
(62) a. *DOM over a low adverb*



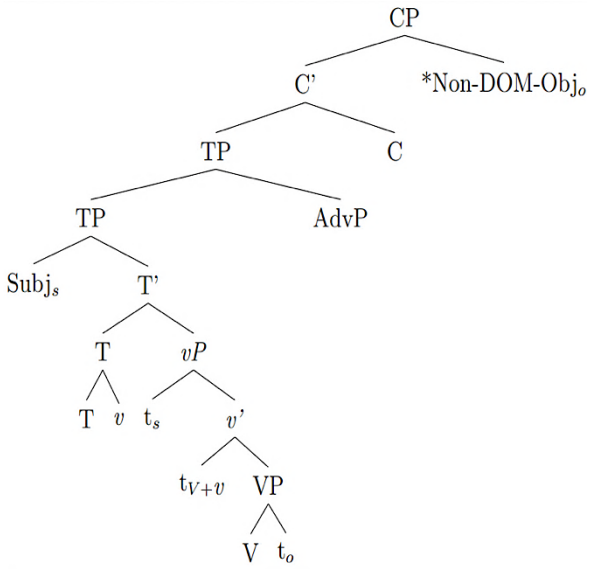
b. *Non-DOM over a low adverb*



(63) a. *DOM over a high adverb*



b. *Non-DOM over a high adverb*



3.5.2 Binding data

An additional test to diagnose the structural height of the DOM *la*-phrase comes from binding data involving dislocation. The binding data below provide evidence in favour of rightward movement targeting a clause-internal position (Hypothesis B).

Since the clitic is part of the DOM construction and enters an anaphoric relation with the *la*-phrase, I consider the syntactic behaviour of the clitic. On the one hand, it has been argued that clitics form an A-chain with their arguments and therefore are not visible for binding (See Suñer 1992; Sportiche 1996; Alexiadou and Anagnostopoulou 1997; Anagnostopoulou 2006; Preminger 2009; Harizanov 2014; Angelopoulos and Sportiche 2021, among others). On the other hand, others treat clitics as independent pronouns with their own referential index which thus enter a binding relation and are therefore visible to binding conditions (Borer 2014; Nomoto 2016, among others).

When clitics do not share an index with the *la*-phrase, I argue that the binding results are inconclusive. On the other hand, when clitics are independent pronouns (visible for binding), I argue that only Hypothesis B is supported.

Turning to the binding facts, the first binding fact involving Conditions A and B is illustrated in the following example.

(64) * sara$_i$ šaf-at-a$_i$ la-nafs-a$_i$
Sara saw-3F.SG-3F.SG.OBJ DOM-self-3F.SG.ACC
(Lit. 'Sara$_i$ saw her$_i$ herself$_i$.')
('Sara$_i$ saw herself$_i$.')

The example in (64) contains the anaphor *la-nafsa* 'DOM-herself' and a potential binder i.e., the subject 'Sara.'[53] Note that the clitic moves to a position that cannot c-command other elements, so it cannot behave as a binder (see, for example, structure (55) where the complex *v* containing the clitic is adjoined to T). If the anaphor moves to the specifier of CP, we expect a Condition A violation because the anaphor would not be c-commanded by the R-expression 'Sara.' Note that a violation of Condition B also arises because the accusative clitic *-a* 'her' is c-commanded by 'Sara.' Therefore, Hypothesis A correctly predicts the ungrammaticality of (64). Under Hypothesis B, if the anaphor moves to a clause-internal position, we do not expect a Condition A violation as the anaphor is bound by 'Sara.' However, a Condition B violation arises because the clitic is bound by 'Sara.' Thus, both Hypotheses account for the ungrammaticality of (64).

If the clitic is not visible to binding, we expect (64) to be grammatical under Hypothesis B but not under Hypothesis A. If the anaphor moves to Spec,CP, a violation of Condition A arises as the anaphor is not bound by 'Sara,' thereby yielding the ungrammaticality of (64). However, if the anaphor moves to a clause-internal position, Spec,XP, Condition A is satisfied, but nevertheless (64) is ungrammatical. Therefore, only Hypothesis A accounts for (64) when the clitic is not visible to binding, however when the clitic is visible for binding, both hypotheses correctly account for (64).

Consider another binding fact involving Condition C where the marked object contains an R-expression.

[53] *nafs-a* 'herself' is a reflexive pronoun and the example containing DOM-herself is expected to be ungrammatical according to the generalization made in Chapter 2 that DOM is illicit with pronouns.

(65) *hi$_i$ šaf-at-o$_k$ la-[ʔben sara$_i$]$_k$
She saw-3M.SG.OBJ DOM-[son Sara]
(Lit. 'She$_i$ saw him$_k$ Sara$_i$'s son$_k$.')
Intended: 'She$_i$ saw [Sara$_i$'s son]$_k$.'

The relevant elements in (65) are the subject pronoun 'She,' the clitic, 'Sara's son' and the R-expression 'Sara.' The potential binders are the pronoun 'She' and 'Sara's son.' Note that 'Sara' in a possessive expression 'Sara's son' cannot be a binder because it occupies a specifier of a possessive DP structure. As pointed out earlier, the clitic cannot be a binder because it is located in a position that cannot c-command other elements (see the structure in (55)).

If the clitic is visible to Condition B and behaves as a bindee, we do not expect (65) to be grammatical under either hypothesis. Hypothesis A predicts the ungrammaticality of (65) because 'Sara's son' in Spec,CP c-commands the clitic located within the TP, thereby triggering a Condition B violation. Note that Condition C is satisfied since 'Sara' is not c-commanded by the subject pronoun. Turning to Hypothesis B, while there is no Condition B violation because the clitic is no longer c-commanded by 'Sara's son,' there is a Condition C violation where 'She', located in Spec,TP, c-commands 'Sara' inside the right-dislocated object. We can conclude that both hypotheses correctly predict the ungrammaticality of (65).

If the clitic is not visible to Condition B as a bindee, we expect (65) to be grammatical only under Hypothesis A but not under Hypothesis B. According to Hypothesis A, Condition C is satisfied because 'Sara' is not c-commanded by the pronoun 'She.' Therefore, under Hypothesis A, (65) is predicted to be grammatical but the sentence is in fact ungrammatical. Hence, (65) cannot be predicted under Hypothesis A. Hypothesis B, on the other hand, gives rise to a violation of Condition C where 'She', located in Spec,TP, c-commands 'Sara' inside the right-dislocated DOM object. This accounts for the ungrammaticality of (65). Thus, only Hypothesis B appears to support (65).

Finally, I provide a grammatical example where only Condition B is relevant, while Condition C is irrelevant (i.e., the proper name and the pronoun are not coreferential).

Example (66) poses a challenge for movement to Hypothesis A since Condition B would not predict the grammaticality of (66).

(66) $sara_i$ šaf-at-o_k la-[ʔbn-$a_{i/j}$]$_k$
Sara saw-3F.SG-3M.SG.OBJ DOM-[son-her]
(Lit. '$Sara_i$ saw him_k [$her_{i/j}$ son]$_k$.')
'Sara saw her son.'

The pronoun 'her' is ambiguous between a pronoun that is coindexed with Sara i.e., Sara saw Sara's son, or to any other female i.e., Sara saw another woman's son. The example contains the binder 'her son.' If the clitic is visible to Condition B as a bindee, we expect example (66) to be ungrammatical under Hypothesis A but not under Hypothesis B. When 'her son' moves to Spec,CP, 'her son' c-commands the clitic thereby inducing a Condition B violation. Yet the sentence (66) turns out to be grammatical. We therefore conclude that Hypothesis A cannot account for (66).

However, if 'her son' moves to the edge of *v*P, the c-command relation can no longer be established between 'her son' and the clitic, so we do not expect a Condition B violation. This prediction is borne out, as evidenced by the grammaticality of (66). Hence, when the clitic is involved in binding, only Hypothesis B correctly predicts this fact. On the other hand, when the clitic is not involved in binding, both hypotheses correctly predict the data in (66).

Table 3.2 summarizes the results of the binding data presented above. The results are inconclusive when the clitic is not visible to binding, however, when the clitic is visible to binding (i.e., it is a bindee), all the binding data are accounted for under the clause-internal Hypothesis B.

	When clitics visible to binding			When clitics are not visible to binding		
	(64)	(65)	(66)	(64)	(65)	(66)
Hypothesis A: Right-periphery	✓	✓	✗	✓	✗	✓
Hypothesis B: Clause-internal	✓	✓	✓	✗	✓	✓

Table 3.2: Comparison between the two hypotheses

Using binding theory to evaluate movement under the two hypotheses, I have argued that when clitics are independent pronouns and subject to the binding theory conditions, the clause-internal hypothesis correctly predicts the binding data.

It is also crucial to compare the DOM data with non-DOM data because reconstruction may be a potential confound.[54] If the binding data for DOM and non-DOM play out the same way, we cannot conclude anything about the height of differentially marked objects. However, if the data play out differently, we can conclude that reconstruction is not observed. Examples (67)-(69) below show that binding data containing differentially marked objects behave differently from data containing non-differentially marked objects, thus reconstruction is not a confounding factor in determining the height of DOM.

The example in (67) shows that differentially marked and unmarked objects exhibit different binding properties. As defined in Chapter 2, the unmarked object lacks a morphological case marking (absence of accusative), whereas the marked form bears overtly dative marking.

(67) a. $sara_i$ šaf-at nafs-a_i
Sara saw-3F.SG self-3F.SG.ACC
'$Sara_i$ saw $herself_i$.'[55] UNMARKED OBJECT

[54] Thanks to Idan Landau for raising the importance of reconstruction.

[55] The sentence has two interpretations: (a) anaphoric reading (context: Sara is looking at the mirror and saw herself) and (b) idiomatic interpretation (Sara seems to be arrogant).

b. * sara$_i$ šaf-at-a$_i$ la-nafs-a$_i$

Sara saw-3F.SG-3F.SG.OBJ DOM-self-3F.SG.ACC

(Lit. ‘Sara$_i$ saw her$_i$ herself$_i$.’)

Intended: 'Sara$_i$ saw herself$_i$.' MARKED OBJECT

If DOM reconstructs, we predict that the unmarked object and the marked object exhibit the same distributional properties with respect to binding conditions. This prediction, however, is not borne out as exemplified in (67). In (67a), Condition A is satisfied as the anaphor 'herself' is c-commanded by 'Sara,' thereby giving rise to a well-formed sentence. On the other hand, (67b) includes a marked anaphor *la-nafsa* 'DOM-herself' that is co-indexed with both the clitic and 'Sara.' The ungrammaticality of (67b) cannot be accounted for by assuming that the differentially marked object reconstructs to the base position occupied by the clitic i.e., the marked anaphor is interpreted within the c-command domain of the subject.

Condition C is another piece of evidence showing that the differentially marked object does not undergo reconstruction as shown in the contrast between (68a) and (68b).

(68) a. * ho$_i$ semeʕ sˁoṭ aḥmad$_i$

He heard.3M.SG voice Ahmad

Intended: 'He$_i$ heard Ahmad$_i$’s voice.’ UNMARKED OBJECT

b. ho$_i$ semʕ-o la-sˁoṭ aḥmad$_i$

He heard.3M.SG-3M.SG.OBJ DOM-voice Ahmad

(Lit. ‘He$_i$ heard it Ahmad$_i$’s voice.’)

'He$_i$ heard Ahmad$_i$’s voice.’[56] MARKED OBJECT

The ungrammaticality of (68a) is due to the fact that ‘he’ is in Spec,TP, and c-commands the R-expression ‘Ahmad’ in its base-generated position, giving rise to a Condition C

[56] The co-indexation is possible in a scenario where friends are listening to a recording of their voices and guessing which voice belongs to who.

violation. By contrast, the grammaticality of (68b) is not expected if the differentially marked object reconstructs.

Until now we have seen that the distributional properties of binding with DOM are different than with the unmarked counterparts. The last minimal pair in (69) exhibits similar properties for both objects. I argue that although the distributional property of DOM with binding is similar to the unmarked one, we cannot conclude that the differentially marked object *must* reconstruct. In both examples of (69a) and (69b) below, Condition C is violated (i.e., the R-expression 'Sara' is c-commanded by the subject pronoun 'she'), thus resulting in ungrammaticality. The ungrammaticality of (69b) can be accounted for if we assume that the DOM-DP reconstructs. However, it does not show that the DOM-DP *must* reconstruct.

(69) a. * hi_i šaf-at ʔben $sara_i$
she saw-3F.SG son Sara
'She$_i$ saw Sara$_i$'s son.' UNMARKED OBJECT

b. *hi_i šaf-at-o_k la-[ʔben $sara_i$]$_k$
She saw-3F.SG-3M.SG.OBJ DOM-[son Sara]
(Lit. 'She saw him Sara's son.')
Intended: 'She$_i$ saw [Sara$_i$'s son]$_k$.' MARKED OBJECT

We have seen that the binding data for DOM and non-DOM phrases presented above differ as shown in (67)-(68); therefore, we can conclude that reconstruction does not take place. I claim that the lack of reconstruction is due to the interpretive effects of the differentially marked object, as will be discussed below.

3.5.2.1 Reconstruction and wide scope reading

This section addresses the question of why differentially marked objects might not reconstruct.

In Chapter 5, I argue that differentially marked objects are interpreted as aboutness topics (in the sense of Reinhart 1981). Endriss (2009) proposes that aboutness topicality

always induces wide-scope reading,[57,58] and wide-scope reading is available when the nominal does not reconstruct. Given the interpretation of the differentially marked object as aboutness topics, we do not expect them to reconstruct.

The association between topicality and wide-scope reading is also adopted by Dočekal and Kallulli (2012). They analyze clitic-doubled DPs in Albanian as aboutness topics in which the clitic-doubled DP invariably takes wide scope. They highlight that if an element is clitic doubled, then it must be interpreted as topical and consistently forces wide scope.[59]

I demonstrate below that the differentially marked object in NGA also has obligatory wide-scope reading with respect to the preverbal subject as shown in (70).[60] The narrow-scope reading is conspicuously absent.

(70) *DOM forces wide scope*

kul walad qara-hon la-th-thalath kotob
every boy read-3PL.OBJ DOM-the-three books.BP
'Every boy read the three books.'

✓ three > every: 'There are three specific books that every boy read.'

* every > three: 'For every boy, there are three (possibly different) books that he read.'

[57] Endriss (2009) argues that there are classes of topics that challenge the claim that direct aboutness topics always take wide scope. For example, topics that denote contrastivity can only receive a narrow-scope reading (Endriss 2009: 212).

[58] Alexopoulou & Kolliakou (2002) associate discourse topic with a wide-scope reading.

[59] See appendix B on the formal semantic mechanism of Endriss (2009) on deriving the meaning of topicality.

[60] The availability of a wide-scope reading with right dislocated quantifiers is similar to clitic left dislocated ones. In Macedonian, for example, clitic left dislocated quantifiers only take wide scope with respect to the subject quantifier as illustrated below (Kochovska 2010: 114,(42b)).

(i) Dve knigii , sekoj student gii pročita.
two books every student them read
* every student > two books; ✓ two books > every student

The facts are consistent with the widely held observation regarding scope facts in Greek, which can be formulated as there is only wide-scope reading when the QP is clitic left dislocated (Anagnostopoulou 1994; Iatridou 1995; Alexiadou & Anagnostopoulou 1997). Similarly, in Spanish, as argued by Arregi (2003: 41), the clitic left dislocated object has an obligatory wide scope with respect to the postverbal subject.

Compare (70) with (71). The non-differentially marked object (i.e., the unmarked counterpart) in (71) permits both wide-scope and narrow-scope readings.

(71) *Non-DOM allows both readings*

kul walad qara th-thalath kotob
every boy read the-three books.BP
'Every boy read the three books.'

✓ three > every: 'There are three specific books that every boy read.'

✓ every > three: 'For every boy, there are three (possibly different) books that he read.'

Thus, the contrast between (70) and (71) provides empirical evidence that differentially marked objects in NGA do not reconstruct as they only allow a wide-scope reading.

3.5.3 Section summary

This section examined the structural height of the moved differentially marked object. Two diagnostics were used: adverbial placement and binding data. The results of the two diagnostics suggest that DOM supports Hypothesis B, namely, DOM targets the edge of *v*P.

Having discussed the syntactic properties of DOM, I turn now to comparing DOM to other superficially similar constructions and argue that DOM must be treated as its own construction.

3.6 DOM and similar constructions

In this section I investigate whether DOM differs from other similar phenomena such as Clitic Doubling (CD) and Clitic Right Dislocation (CLRD). Given the presence of the clitic in the DOM construction and the fact that the nominal phrase is dislocated (shown in section 3.2), it is worth asking whether DOM is an instance of CD and CLRD. Based on

the observations will be presented below, I argue that DOM is in fact a separate construction, despite some surface similarities with CD and CLRD.[61]

Cross-linguistically, CLRD and CD show distinct syntactic properties (Rivas 1977; Jaeggli 1982, 1986; Borer 1984; Anagnostopoulou 1999, 2006, 2014; Harizanov 2014; Di Tullio et al. 2019, among others); thus, they are treated as two separate phenomena.[62] Anagnostopoulou (2006) discusses the cross-linguistic variation of CD and CLRD and highlights the differences between the two. Building on Anagnostopoulou's work, CD and CLRD must be distinguished in the following properties:

(72) In CLRD, the dislocated phrase is set off from the rest of the sentence with a sharp intonational break, while no such break is required before the object in CD.

(73) In the languages in which CD is subject to Kayne's generalization, CLRD is not (e.g., Rioplatense Spanish). Kayne's generalization is defined in section 2.2.3 as an object NP that may be doubled only if it is preceded by a special preposition.

The two properties are summarized in Table 3.3.

	CD	**CLRD**
Intonational property	✗	✓
Kayne's generalization	✓	✗

Table 3.3: The distinction between CD and CLRD

[61] Given that DOM obligatorily involves CD and right dislocation, it is worth asking whether DOM is a separate construction from CLLD, another clitic construction, De Cat (2002, 2007) shows that CLRD is the mirror image of CLLD, namely, CLRD displays the same properties as CLLD. See section 4.3 for the analysis that DOM deserves its own syntactic analysis as it shows several differences with CLLD.

[62] CLRD must be distinguished from a parallel construction such as Right Scrambling (Villalba 2000). Right scrambling is found in languages such as Japanese and Hindi-Urdo where the unmarked word order is verb-final, but constituents can undergo rightward scrambling (postverbally). Unlike CLRD, there is no pronominal element in the construction as shown below.

(i) Siita-ne Mohan-ko dikhaa-ii ek kitaab
Sita-ERG Mohan-DAT show-PFV a book
'Sita showed a book to Mohan.' Hindi-Urdo (Manetta 2012: 43(1))

This contrast between CD and CLRD is highlighted in the Rioplatense Spanish examples.

(74) *Rioplatense Spanish* (Anagnostopoulou 2006: 526)

a. Parece que tuvieron que llevar-la de urgencia a los Estados Unidos **, la**
Seems that had.3PL that bring-her urgently to the United States, the
hija de Coronel Martínez
daughter of Coronel Martinez
'It seems that they had to take her urgently to the United States, the daughter of Coronel Martinez.'

b. Parece que tuvieron que llevar-la **a la hija de Coronel Martinez** de
Seems that had.3PL that bring-her a the daughter of Coronel Martinez
urgencia a los Estados Unidos
urgently to the United States
'It seems that they had to take the daughter of Coronel Martinez urgently to the United States.'

The example in (74a) is an instance of dislocation as evidenced by the presence of a pause before the dislocated element (boldfaced). (74b) is a CD example as clearly shown by the clitic-doubled element preceded by the preposition *a*, which according to Kayne's generalization, is one of the defining characteristics of CD, as presented in (73).

DOM in NGA shares properties with both CD and CLRD. In particular, both CD and DOM are (i) subject to Kayne's generalization, and (ii) do not involve special prosodic phrasing. At the same time, DOM involves right dislocation as demonstrated in section 3.2.

The first characteristic that differentiates between CD and CLRD in NGA is the intonational properties of the phrase (see (72)). DOM patterns with CD, but not CLRD, as there is no prosodic boundary before the associate i.e., differentially marked object. The following examples illustrate that having a comma, indicative of the prosodic boundaries, is ungrammatical with DOM (75a), but its absence yields grammaticality (75b).

(75) a.* baʔref-a$_i$ **,** la-sara$_i$
know.1SG-3F.SG.OBJ DOM-Sara
(Lit. 'I know her, Sara.')

Intended: ‘I know Sara.’

b. baʕref-a$_i$ la-sara$_i$
know.1SG-3F.SG.OBJ DOM-Sara
Lit. ‘I know her Sara.’
‘I know Sara.’

Unlike in NGA, the right dislocated material in some Romance languages is preceded by an intonational break and optional short pause (represented by a comma). By way of illustration, consider the following examples.

(76) *Catalan* (adapted from García-García 2018 :175, footnote 10)
Le$_i$ he fet un regal, [a la meva mare]$_I$
Her-DAT have-AUX.PRS.1SG make-PTCP a present DOM the my mother
‘I made a gift to my mother.’

(77) *Sardinian* (Jones 1993: 318(18))
L’$_I$ appo vistu, [su dottore]$_i$
him have seen the doctor
‘I have seen him, the doctor.’

The second characteristic used to differentiate between CD and CLRD is whether the construction is subject to Kayne’s generalization. DOM shares this property with CD (but not CLRD). As described in section 2.2.3, in order to have a clitic-doubled element, the preposition *la-* ‘to,’ a dative marker, must co-occur with the associate (78).

(78) al-wlaad beḥbbo-ha *(la)-sara
the-kids.BP love.3PL-3F.SG.OBJ DOM-Sara
‘The kids loved her Sara.’
‘The kids love Sara.’

The observations so far suggest that DOM shares properties with CD, but not CLRD. However, once we consider the dislocation facts presented in section 3.2, we can also claim

that DOM shares properties with CLRD. Table 3.4 summarizes the properties DOM shares with CD and CLRD discussed so far.

	CD	**CLRD**	**DOM**
allow adverbs to intervene between the verb and clitic (section 3.2)	✗	✓	✓
lacking intonational break	✓	✗	✓
subject to Kayne's generalization	✓	✗	✓

Table 3.4: Properties of DOM, CD and CLRD in NGA

Thus, DOM is distinct from CD and CLRD because it only overlaps but is not identical to either CD or CLRD.

3.7 Conclusion

The goal of this chapter was to explore the clausal structure of DOM in NGA which constitutes a basis for the syntactic analysis of the internal structure of DOM in Chapter 4. I have summarized the behavior of DOM with respect to the properties below in (79).

(79) *The syntactic properties of DOM in NGA*

a. DOM involves dislocation.

b. DOM obeys RRC.

c. DOM obeys CSC.

d. DOM satisfies the Right Edge Restriction.

e. DOM is A-movement.

f. DOM affects binding.

g. The differentially marked object only takes wide scope.

I focussed on the questions: (i) how the dislocated differentially marked object is derived; (ii) what the structural height of the differentially marked object in the clause structure is, and (iii) whether DOM resembles other superficially related phenomena such as CD and CLRD.

First, I proposed that DOM is derived by movement and obeys the locality constraints of rightward movement. Second, I argued that differentially marked objects target a clause internal position, but not the right periphery. Finally, DOM must be distinguished from other superficially phenomena (CD and CLRD), and, accordingly, requires its own syntactic analysis. In Chapter 4, I continue to investigate the syntax of DOM focusing more specifically on the identity of the target XP.

Chapter 4

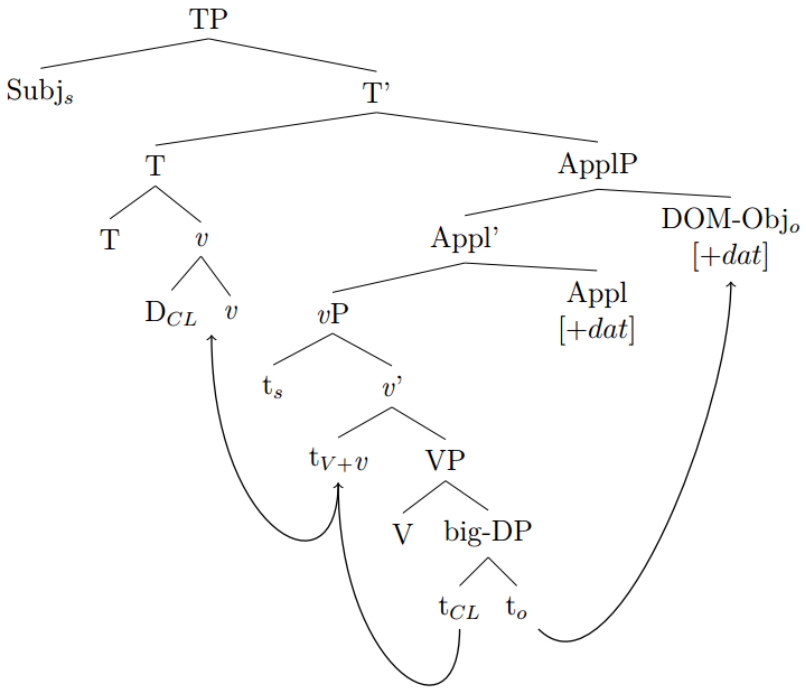
The landing position of DOM

4.1 Introduction

This chapter continues the investigation into the syntax of DOM in NGA. In Chapter 3, I argued that, based on the adverbial and binding data, the dislocated differentially marked object undergoes rightward movement to a functional projection, dubbed XP, located at the edge of *v*P. The facts strongly militate against a right periphery analysis of DOM. This chapter concerns the identity of XP as the target of the DOM movement.

I argue for the DOM syntactic structure in (80) where differentially marked objects target Spec,ApplP. Introducing applicative structure has three main empirical motivations: (i) case checking, (ii) locality conditions of the DOM movement, and (iii) aboutness topicality (will be argued in detail in Chapter 5). I will elaborate on each motivation in order.

(80) *DOM in NGA*



The chapter is organized as follows. The next three sections concentrate on the motivations behind the ApplP as an analysis for DOM in NGA.[63] In section 4.5, I lay out the predictions of the applicative analysis. Section 4.6 concludes.

4.2 Case

The first motivation for introducing the applicative structure in (80) comes from case checking.

Movement-based accounts of DOM consider raising of the object out of VP to be a necessary ingredient of DOM, as schematized in (81). This type of approach is referred to in the literature as object raising (e.g., Bhatt & Anagnostopoulou 1996; Bhatt 2007; Baker & Vinokurova 2010; López 2012; Ormazabal and Romero 2013).

[63] I will not use Pylkkänen's (2002) diagnostics for determining the attachment height of ApplP but see appendix C for the argument that the diagnostics are not reliable for DOM in NGA.

(81) [TP T ... [DOM-object ... [VP V t_{obj}]]]

There are two types of movement-based accounts (i) accounts in which the object moves to a Case position (e.g., Bhatt 2007; Rodríguez-Mondoñedo 2007; López 2012; Ormazabal and Romero 2013; Kalin 2018), and (ii) accounts in which object raising feeds case competition with the subject (e.g., Baker and Vinokurova 2010; Baker 2014; Levin and Preminger 2015). The motivation presented in this section argues in favor of the first analysis.

The first account argues that DOM arises as a result of the movement of certain objects out of VP in order to check case. Within this account, the details of the landing position of the marked object are debatable. It has been argued that differential DPs undergo syntactic movement within and subsequently higher than *v*P (e.g., Diesing 1992; Roessler 2019). López (2012) argues that for Spanish direct objects with DOM are in a higher position than zero-coded objects, namely, in a projection between *v*P and VP which he calls αP (which bundles aspectual and applicative features). When a direct object DP is in an αP, it is assigned the accusative case, and when it is incorporated, it is not assigned case at all.

It has been widely argued that the applicative head licenses the dative case to its argument (e.g., Marantz 1993; Cuervo 2003, 2010; Pylkkänen 2002, 2008; Grashchenkov & Markman 2008; Bruening 2010; Hallman 2018, 2021; Wood & Zanuttini 2018; Gogłoza 2021). As demonstrated in section 2.1 in Chapter 2, differentially marked objects in NGA are marked with dative case. In Chapter 3, I assume that the differentially marked object starts its life in a big-DP configuration where the originally base-generated position cannot be assigned dative case as there is no existing syntactic head that can assign dative to the DP. In order to satisfy the Case filter conditions (Chomsky 1981), the differentially marked object needs to move to Case position. As shown in (80), the marked object moves up to Spec,ApplP where it receives dative directly from Appl in the spec-head relation.

The second type of movement account is exemplified by Baker and Vinokurova (2010) who examine the Turkic language Sakha. They argue that the differentially marked object undergoes movement out of VP phase into the higher CP phase. Since the subject and the object are in the same spell out domain, the subject c-commands the object, and the object is therefore assigned dependent accusative case. Baker and Vinokurova (2010) argue

that dependent case provides a straightforward analysis of differential marking patterns involving movement, such as that found in Sakha. However, this type of approach is challenged by data from NGA clitic constructions. If we assume that case mismatching in DOM (the clitic is accusative and the DOM-DP is dative) is assigned by some form of dependent case, we need to explain how the same form of dependent case can also account for the absence of case in another clitic construction (Clitic Left Dislocation, CLLD). For the purpose of this thesis, I put dependent case and its challenges aside and instead follow an alternative analysis for case assignment in CLLD described below.

CD in NGA (and other closely related Arabic dialects) does not always give rise to two different morphological cases. CLLD is an example which obligatorily involves CD but it does not display two overt distinct case markings (see (82b) below).[64] Part of the motivation behind the ApplP as the target of the DOM movement stems from a comparison of DOM to CLLD. As shown in Chapter 2, the dislocated element in DOM, illustrated in (82a), does not share the same morphological case as the clitic; the marked object shows up dative while its corresponding clitic is accusative. In CLLD, the dislocated DP does not bear overt case as case is not morphologically marked in Arabic dialects including NGA. In (82b), the clitic is marked accusative, and the associate is unmarked (absence of accusative).[65]

(82) a. aḥmad šara-ha la-s-sayyara DOM
Ahmad bought-3F.SG.OBJ.ACC DOM(DAT)-the-car
Lit. 'Ahmad bought it the car.'
'Ahmad bought the car.'

b. s-sayyara, aḥmad šara-ha CLLD
the-car, Ahmad bought-3F.SG.OBJ.ACC
Lit. 'the car, Ahmad bought it.'

[64] Ouhalla (1997) refers to CLLD as instances of "left-dislocated phrases" (LD-phrases) (see also Soltan 2007).

[65] Unlike in NGA, in MSA, CLLD can exhibit case-mismatching: the left-dislocated element is marked with nominative and the clitic is marked accusative. This construction is known in the literature as Hanging Topic Left Dislocation (see Alzayid 2022 for more details on this construction).

'Ahmad bought the car.'

The question arises now is how the contrast in case between DOM and CLLD motivates our ApplP analysis for DOM.

Caha (2023) builds on Schütze's (2001) idea that default case is the absence of case. Default case forms of a language are defined as the case used to spell out nominal expressions (e.g., DPs) that are not associated with any case feature or otherwise determined by syntactic mechanisms (Schütze 2001). Caha (2023) presents a theory of default case in the spirit of Schütze, but is more restrictive in terms of which case values are allowed for default case and its syntactic distribution.[66] Caha shows that only certain configurations are default-case environments. For example, in German, left-dislocated positions hosting topics are the default case (Schütze 2001: 224).

Building on Caha's analysis, I argue that the difference in case in CLLD and DOM constructions is attributed to distinct case assignment mechanisms applied for each construction. I argue that CLLD is a default case environment (see example (82b)) while DOM is a proper syntactic case assignment. Appl is necessary to assign the syntactic case for the differentially marked DP (see the structure in (80)), but not for CLLD. Hence, the applicative analysis is tenable for deriving the case marking in NGA DOM.

4.3 The locality condition of DOM movement

The second empirical motivation for ApplP as the target position of DOM comes from the locality restriction of DOM movement. As demonstrated in section 3.3, DOM is subject to the Right Roof Constraint (RRC). In contrast to DOM, CLLD is an unbounded dependency in which the relation between the clitic and the dislocated element is not instantiated in the same clause (e.g., Cinque 1990; Iatridou 1991; Anagnostopoulou 1994; Cecchetto 1999; Gregoromichelaki 2013). In NGA, a left dislocated DP undergoes long distance left dislocation, that is, it escapes the boundary of the clause in which it originates.

[66] Thanks to Ivona Kučerová for bringing Caha's account to my attention.

(83) a. sara waʕdat-na ʔenno raḥ nzor šallalat ʔl-niagara BASELINE
Sara promised-us that will visit falls the-Niagara
'Sara promised us that we will visit Niagara Falls.'

b. šallalat ʔl-niagara, sara waʕdat-na ʔenno raḥ nzor-on CLLD
falls the-Niagara, Sara promised-us that will visit-3PL
Lit. 'Niagara Falls, Sara promised us that we will visit them.'
'Sara promised us that we will visit Niagara Falls.'

The example in (83a) is the standard transitive sentence on which we operate the relevant permutation. In (83b), the DP *šallalat ʔl-niagara* (the direct object of the embedded clause) undergoes long distance left dislocation (it surfaces to the left of the matrix clause).

It has been argued that CLLD targets the left periphery (Benmamoun 2000; Aoun, Benmamoun, and Choueiri 2010; Al-Balushi 2011, 2012; Makkawi 2021).[67] In contrast, Chapter 3 provided evidence from the adverbial placement test and binding data that DOM is incompatible with the right periphery analysis (dubbed as Hypothesis A) but compatible with the internal-clausal analysis (dubbed as Hypothesis B). The syntactic difference between DOM and CLLD lies in their locality conditions. The crucial question at this point is how the applicative analysis accounts for the ban on differentially marked object to move to a higher position.

I argue that the applicative position of the dislocated differentially marked DP at the edge of the *v*P displays freezing effects. I take this as a "criterial position," from which further movement is impossible. This phenomenon is the effect of what Rizzi (2006) calls Criterial Freezing.

(84) *Criterial Freezing* (Rizzi, 2006: 112))
A phrase meeting a criterion is frozen in place.

[67] It has been argued that CLLD is derived by a hybrid approach, namely, CLLD in Arabic can be derived by both base-generation and PF movement (Aoun and Benmamoun 1998). After Iatridou (1995), Alzayid (2022) calls this simultaneous behavior of movement and base-generation 'Cinque's paradox.

Rizzi (2003, 2006) observes that a phrase meeting a Criterion (= reaching a position dedicated to particular discourse interpretive property in the term of Chomsky 2001), is frozen in place and resists further movement to a higher position.[68]

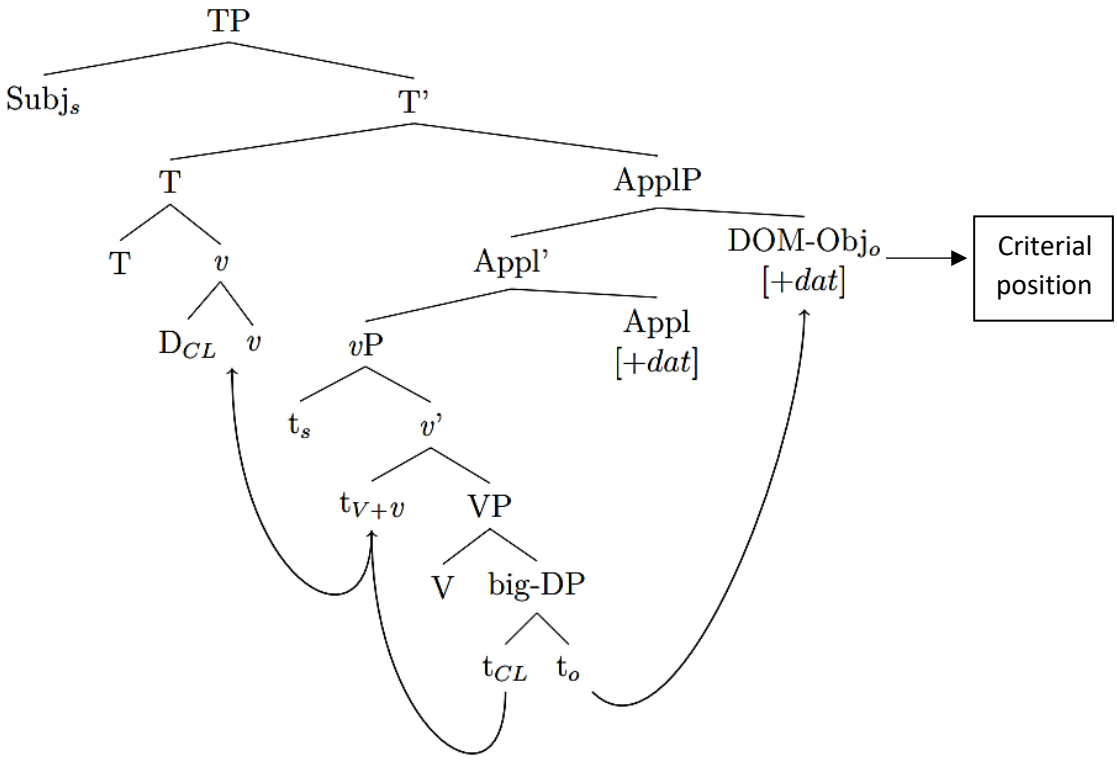
I propose that the Criterial Freezing approach accounts for the ban of DOM on further movement. The differentially marked object has already satisfied an interpretive property (details discussed in Chapter 5), i.e. a criterion, and thus it is "arrested" in its criterial position and cannot move further. The freezing effect accounts for RRC. Criterial Freezing restricts DOM movement to higher position i.e., higher than ApplP, and in this way, DOM respects the RRC locality constraint.

Freezing effects can be observed with A-movement (e.g., Müller 2013; Corver 2014; Blümel 2017), and as argued in section 3.4, DOM is an instance of A-movement. In German, Müller (2013) argues that freezing effects appear regularly with A-movement such as scrambling and topicalization. Freezing effects are also observed in rightward movement (Wexler & Culicover 1980; Corver 2014; Overfelt 2015).

The corresponding structure of DOM adopting the Criterial Freezing is represented in (85) below.

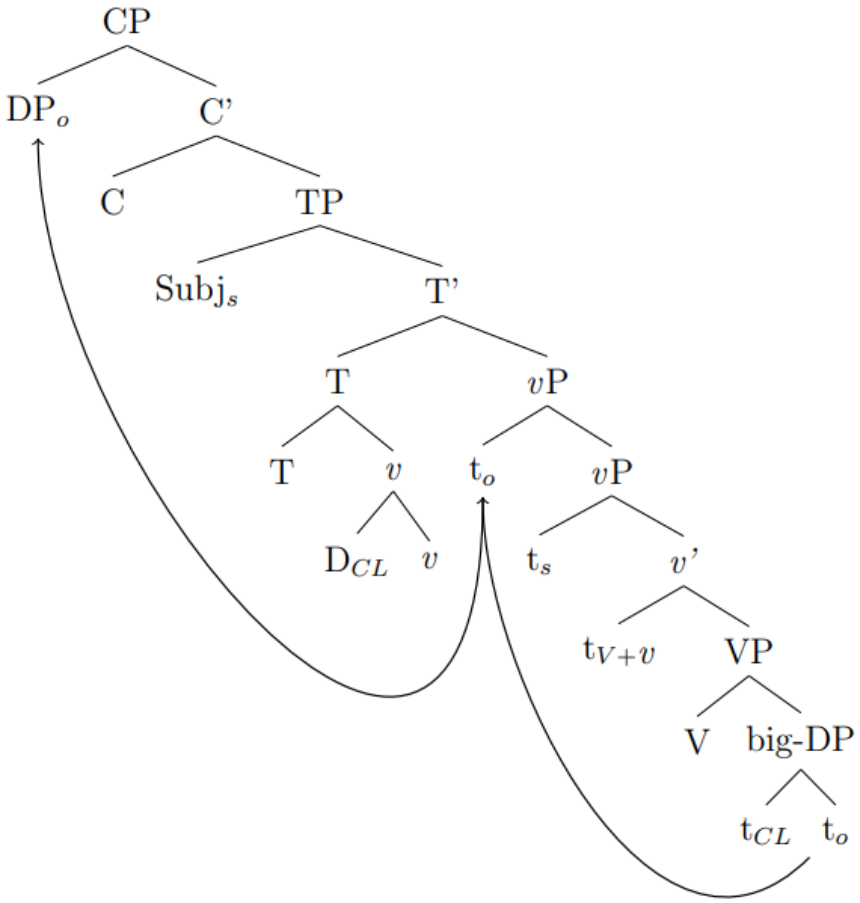
[68] Note that Criterial Freezing is different from a probe-goal theory of derivational syntax (Chomsky 2001) in that it demands a Spec-head relation between the criterial head and a phrase with a matching criterial feature. Additionally, unlike probe-goal relations, movement is not a reflex of Φ-Agree in Criterial Freezing

(85) *DOM in NGA*



In contrast to DOM, CLLD does not induce freezing effects and therefore does not obey RRC. Following Aoun and Benmamoun (1998) in Lebanese Arabic, I adopt the big-DP analysis for CLLD. It is assumed that CLLD undergoes cyclic-movement: first movement of the DP dislocated element to an intermediate position (adjunction to *v*P), and second, successive movement to the final landing position (Spec,CP). The corresponding structure for CLLD in NGA is represented in (86).

(86) *CLLD in NGA*



Thus, the comparison of DOM to CLLD motivates the applicative analysis as a target for DOM in which the ApplP, being at the edge of *v*P, induces freezing effects which prevent the differentially marked object to move further.[69]

[69] The comparison of DOM to CLLD raises the question of whether DOM and CLLD are information-structural related. It is argued that the determining factor underlying the interpretation of CLLD is contrastiveness (Alzayid 2022), whereas aboutness topicality for DOM (argued in chapter 5). Yet the dislocated element in CLLD, shown in (86), is derived by movement to the left periphery (Benmamoun 2000; Aoun et al. 2010; Al-Balushi 2011, 2012; Makkawi 2021), and the dislocated differentially marked object undergoes shorter movement to ApplP (see (85)).

4.4 Aboutness topicality effects

I have argued that the case checking and locality condition motivate DOM movement to ApplP. Another motivation for this movement is concerned with the aboutness topic status of DOM in NGA. As will be argued in Chapter 5, the determining factor underlying the interpretation of DOM in NGA is aboutness topicality (in the sense of Reinhart 1981). An aboutness topic (i.e., a sentential topic) is defined as what a sentence is about. From a semantic perspective, this means that for something to be a sentential topic, it needs to be associated with a referential address which, in and of itself, is associated with the common ground (e.g., Endriss 2009).

Aboutness topics require referential anchoring and referential anchoring requires a particular syntactic position. It has been argued that aboutness topics must move to the phase edge of *v*P (e.g., Frey 2000, 2004). I follow Johns and Kučerová's (2017) argument that the obligatory movement to the edge of the phase is a direct consequence of topics requiring association with a referential address. Thus, movement of the differentially marked object to Spec, ApplP, being at the phase edge, allows the object to be anchored to a referential address, making it the aboutness topic.

4.5 Predictions

I adopt the object raising analysis for DOM as shown in (80). I argued in Chapter 3 that based on adverbial placement and binding data, the differentially marked object moves to a clause-internal position, at the edge of *v*P, and not to the right periphery.

In order to satisfy the Case filter conditions (Chomsky 1981), the differentially marked object needs to move to Case position. Thus, attracted by the case feature (dative) of the Appl head, the marked object moves up to Spec,ApplP. Further, I assumed in Chapter 3 that the clitic moves to *v* and ends up as part of a complex verbal head including *v*, then this complex *v* undergoes head movement to T. In turn, little v assigns accusative case to an object, which generates an underlyingly accusative clitic that undergoes head movement to T.

There are two immediate predictions to this proposal. First, the structure in (80) accounts for the fact that DOM obeys RRC (shown in Chapter 3) i.e., DOM is a clause-bounded operation. DOM is compatible with an ApplP as it is situated in a position that is not as high as TP.

Crucially, the analysis also makes predictions about the type of nominals participating in DOM. As will be argued in Chapters 5 and 8, not every nominal can be differentially marked. The big-DP analysis is a necessary pre-condition for objects participating in DOM. Given that the big-DP combines together the DP and the corresponding clitic which together form a single constituent, it predicts that only nominals that are mapped onto a referential address (fall under the definition of aboutness topicality of Reinhart, 1981 and Endriss, 2009) can be linked to a pronoun. This prediction is borne out in DOM since nominals that map onto a referential address are able to combine with pronominal elements, and consequently, are differentially marked. This will be shown in detail in Chapter 8.

4.6 Conclusion

This chapter has argued for ApplP as the target position for DOM in NGA. The applicative analysis is motivated by (i) dative case marking, (ii) the locality conditions of DOM movement, and (iii) aboutness topicality effects. In the following chapter, I will expound on aboutness topicality as a motivation for the DOM movement to ApplP at the edge of *v*P.

Chapter 5

DOM and information structure in NGA

5.1 Introduction

I have argued in chapters 3 and 4 that differentially marked right-dislocated objects move to a clause-internal position, specifically, to Spec,ApplP. Similar analyses maintaining that right-dislocated phrases move to an intermediate position between TP and VP have been proposed by Belletti (2005), Villalba (1998, 2000), and López (2003, 2009). In the literature, the clause-internal position has been associated with the interpretive effects of the right-dislocated constituent, specifically topicality (Cecchetto 1999; Villalba 2000; Belletti 2001, 2005; Cruschina 2021).[70,71] The main purpose of this chapter is to argue that the defining property of DOM is aboutness topicality (in the sense of Reinhart 1981). I

[70] According to the split-CP proposal initially presented by Rizzi (1997), the discourse-oriented projections are available at the C-area which are presented as a sequence of functional elements. Other scholars propose that these discourse properties can also be available in the area referred to as the low periphery (an area between the TP and the VP).

[71] Having the clause-internal position hosting topics aligns with the view that topics do not necessarily have to appear on the left periphery (see Reinhart 1981; Lambrecht 1994; Frey 2000, 2004; Endriss 2009) contra, for example, Molnár (1993) and Jacobs (2001).

follow the intellectual lead of Dočekal and Kallulli (2012) in arguing that aboutness topicality is a requirement for DOM. Dočekal and Kallulli build on Endriss' (2009) formalization of aboutness topicality. Their core argument comes from the observation that clitic doubling is required in contexts where a topical interpretation arises.

The rest of this chapter is structured as follows: section 5.2 claims that aboutness topicality is a necessary characterization for DOM in NGA. To test whether the differentially marked object is an aboutness topic, I apply two tests: (i) the first test builds on Reinhart's (1981) "said about" test on what is considered to be an aboutness topic; and (ii) Endriss' (2009) test on topicality, which builds on the observation that only quantifiers with particular lexical semantic properties can be aboutness topics. I show that only quantifiers which can be topics can be DOM. In this section, I also discuss other notions suggested to be connected to DOM including referentiality (Lambrecht 1994), specificity (Enç 1991), and givenness (Schwarzschild 1999). I propose that none of these notions are clear-cut properties for DOM in Arabic. Section 5.3 excludes other information-structural characterizations of DOM such as types of foci. Section 5.4 discusses the association between topicality and other CD languages. Section 5.5 concludes.

5.2 The information-structural characterization of DOM

The goal of this section is to address the main question of what information-structural property underlies the interpretation of DOM. The role of information structure has been widely discussed in the DOM literature (e.g., Aissen 2003; Leonetti 2004, 2008; Klumpp 2012; Belletti 2018; Cristofaro 2019; Onea and Mardale 2020; Hill and Mardale 2021; Irimia 2022). I argue that aboutness topicality (Reinhart 1981) is the defining property for DOM in NGA. I adopt Reinhart's (1981) aboutness notion of topicality stated in (87).[72]

(87) *Aboutness topicality*: (after Reinhart, 1981)

A sentence topic is what a specific sentence is about.

[72] From the semantic perspective, this means that for an element to be an aboutness topic, it needs to be associated with a referential address which in and of itself is associated with the common ground. This view will be discussed more extensively as it is crucial for the analysis in Chapter 8.

I follow the intellectual lead of Dočekal and Kallulli (2012) in arguing that aboutness topicality is a requirement for DOM. Dočekal and Kallulli closely follow the formalization of aboutness topicality proposed in Endriss (2009) to provide a convincing case that CD is tied to aboutness topicality in Albanian.[73] Their core claim comes from their chief observation that only a specific set of quantifiers can be aboutness topics, and subsequently can be clitic doubled. Only quantifiers that are mapped onto a minimal witness set representation i.e., the maximal set of which a certain property must hold, are topical (Endriss 2009). See appendix B for the exact semantic mechanism.

5.2.1 Aboutness topicality

Our theoretical understanding of topicality is rooted in Reinhart's (1981) work. Reinhart (1981: 4) writes: "Although the linguistic role of the relation topic is widely acknowledged, there is no accepted definition for it and not even full agreement on the intuitions of what counts as topic."[74] It is beyond the scope of this thesis to go into all the notions of topicality that have been discussed throughout the literature of information structure; here I focus on aboutness topicality.

I follow Reinhart's (1981) notion of aboutness topicality (which is based on Strawson 1964) wherein a sentence topic is what the sentence is about. Reinhart's (1981) work seeks to answer what it means for an expression to be about something by employing the notion of a "context set" as originally assumed in Stalnaker (1978).[75] The context set represents the shared knowledge of a speaker and addressee in a given context. This context can be updated by adding new propositions to the context set according to the utterances created. Reinhart observes that the context set as proposed in Stalnaker (1978) is not organized in any way and suggests fixing the internal organization of the context set.

[73] The proposal of Dočekal and Kallulli has also been followed by Johns and Kučerová (2017) for the Inuit language. Johns and Kučerová argue that topichood straightforwardly derives the core morphosyntactic properties of the Inuit object case marking. Specifically, if the object is a sentential topic, then it triggers double agreement on the verb. They use the term topic as *sentential topic* in the sense of Reinhart (1981).

[74] van Bergen and de Hoop (2009 :173) wrote "there is very little consensus among linguists on any … specific condition. Multiple properties contributing to topichood have been described, but none of these properties seem either necessary or sufficient to classify something as topic."

[75] Keeping things as simple as possible, Reinhart restricts herself to DP-topics, although she notes that there may be topics from other categories such as frame setting adverbs etc.

Reinhart proposes that the organization of information is similar to a library catalogue, a metaphor that is comparable to the file change semantics of Heim (1982) and Erteschik-Shir (1997). Specifically, in this metaphor, each book entry corresponds to a proposition. The idea is that all the information of the library is structured in a way that it is stored under a particular topic or keyword. Based on this metaphor, the topic of the sentence is understood to be the address for the context update, which points to a place where the information conveyed by the sentence will be stored during the context update. This information can then be accessed via the topic entry.[76] In Reinhart's view, 'sentence topics [...] are one of the means available in the language to organize, or classify the information exchanged in linguistic communication' (Reinhart 1981: 24).

There are different ways that have been argued to test for aboutness topicality. For example, certain syntactic structures, morphological markers and intonational means have been argued to determine the topical status of a constituent (Endriss 2009). In Arabic, however, topicality cannot be tested via such means because, unlike other languages like Japanese (see Kuno 1972) and Korean (see Tomioka 2007), it lacks overt morphological marking for topics. Moreover, similar to English and German,[77] intonation also cannot fully serve as a topic marking device because the specific (rising) intonation in Arabic also marks contrastivity, which will be examined further in section 5.3.2. Thus, one has to be cautious when applying these devices as topic tests, because often the same construction marks both topicality and contrastivity. Therefore, contrastivity must be controlled for to definitively claim that topicality is involved.

To control for contrastivity, there are two reliable tests for aboutness topicality which can be used in Arabic. Building on Reinhart (1981), the first test is a simple test of what the sentence is about. The second test is based on Endriss' (2009) observation that only quantifiers with certain lexical semantic properties can behave as aboutness topics. I start by applying the first test to decide whether the differentially marked objects are interpreted as aboutness topic.

[76] Vallduví (1992) also used this metaphor. In particular, what are taken as entries in Reinhart's proposal are dubbed 'addresses' in Vallduví's proposal. In Vallduví's work, each topic is a salient address under which the remaining information is stored.

[77] See data on German and English in Endriss (2009: 54-57).

5.2.1.1 Test #1: Reinhart (1981)

To test if a given constituent can behave as an aboutness topic, topic tests such as 'what about X' and 'say something about X' (Gundel 1975; Reinhart 1981; Erteschik-Shir 1997) have been proposed in the literature. These tests are built on the individual's intuitions concerning what a specific sentence is about. Suppose that in a sentence *s* there is an element *x* which is believed to be the aboutness topic of *s*, one can check their intuitions by testing whether *s* still sounds natural continuation if it is preceded by *I will tell you something about x*. If *s* sounds unnatural, *x* cannot be the aboutness topic of *s*. However, if *s* still sounds natural, *x* is at least a likely candidate to function as the topic for *s*. Reinhart (1981) proposes the following method to test the aboutness status of a given element: if a sentence *s* with the assumed topic *x* can be safely paraphrased by *It was reported of x that s*, then the aboutness topic of *s* is *x*.

The shortcoming of these tests, as pointed out by Endriss (2009: 31), is that: (i) they "heavily rely on rather vague intuitions that are not independently verifiable," and (ii) they only apply to a small subset of cases and are not applicable for items such as indefinites and quantifiers. Roberts (2011) adds that none of these tests seem adequate or necessary to characterize topics, nor is it obvious that they in fact test for the same entity.[78]

Despite the shortcomings of such intuition-based tests, the judgments across the NGA informants consulted in this study were clear-cut and consistent. I use Reinhart's test to determine whether or not the differentially marked object can be regarded as the aboutness topic.[79] Examples (88) and (89) illustrate that differentially marked objects are interpreted as aboutness topic because the DOM sentence can be paraphrased by *they were saying about x that s*, the aboutness topic of *s* is *x*, as shown in examples (88b)-(89b).[80]

[78] In the same vein, Casielles-Suárez (2004: 24) notes "(t)hese tests have been found to be too strong in some instances, since some topicalized phrases fail to pass the tests, and also too weak in other instances, since they can identify as topics too many elements."

[79] I had to modify the original test by using an active form for the verb because the original test uses the passive form which results in ill-formedness.

[80] Other tests proposed in the literature that test for topicality do not actually test for topicality. One example is the subjecthood test, i.e., every grammatical subject must be a topic. Following Reinhart (1981) and Endriss (2009), I will not treat this test as a test for topicality because Arabic has cases where subjects do not have to be topics as they can be focus.

(88) *DOM-objects are aboutness topics*

a. sara qatlat-a la-l-ʕaqrb-e
Sara killed-3F.SG.OBJ DOM-the-spider-F.SG
'Sara killed the spider.'

b. = kano yqolo ʕan al-ʕaqrbe ʔenno sara qatlat-a
were said.3PL about the-spider that Sara killed-3F.SG.OBJ
'They have said about the spider that Sara killed it.'

(89) a. sara marqat-o la-l-amtḥan
Sara passed-3M.SG.OBJ DOM-the-exam.M.SG
'Sara passed the exam.'

b. = kano yqolo ʕan l-amtḥan ʔenno sara marqat-o
were said.3PL about the-exam.M.SG that Sara passed-3M.SG.OBJ
'They have said about the exam that Sara passed it.'

The question now is: can non-differentially marked objects also be aboutness topics? In contrast to differentially marked objects, unmarked objects are, in fact, not interpreted as aboutness topics. In (90), DOM cannot appear in an answer to a focus question; namely, DOM is prohibited in contexts where the definite object is not the topic. See section 5.3 for the discussion of types of focus (information and contrastive focus) with DOM.

(90) *DOM is banned when the DP object is not a topic*

A: šu ʕemlet sara
what did Sara
'What did Sara do?'

B: # sara qarato la-l-ktaab
Sara read-3M.SG.OBJ DOM-the-book
('Sara read the book.')

As for the unmarked objects, the examples (91) and (92) below show that sentences containing unmarked objects cannot be paraphrased by *they were saying about x that s*, the aboutness topic of *s* is *x*. [81]

(91) *Non-DOM objects are not aboutness topics*

a. sara qatlat l-ʕaqrb-e
Sara killed the-spider-F.SG
'Sara killed the spider.'

b. # kano yqolo ʕan al-ʕaqrbe ʔenno sara qatlat-a
were said.3PL about the-spider that Sara killed-3F.SG.OBJ
('They have said about the spider that Sara killed it.')

(92) a. sara marqat l-amtḥan
Sara passed the-exam.M.SG
'Sara passed the exam.'

b. # kano yqolo ʕan l-amtḥan ʔenno sara marqat-o
were said.3PL about the-exam.M.SG that Sara passed-3M.SG.OBJ
('They have said about the exam that Sara passed it.')

I conclude, therefore, that the 'aboutness' property is what underlies the interpretation of elements occurring in the DOM construction in NGA. This conclusion does not align with Brustad (2000: 354) who describes the DOM phrase in Syrian Arabic as a discourse topic but not as a sentence topic that only operates on a single sentence. Using Syrian data contexts, she claims that the DOM phrase is not a sentence topic because it does not operate on the sentence level but only on the context level where the context plays an important role in using DOM. On the contrary, the examples above indicate that DOM strictly operates at the sentential level. In Chapter 6, I discuss scenarios where DOM in NGA yields

[81] I assume that indefinite objects in Arabic cannot be aboutness topics, so there is no reason to compare indefinite nominals but the relevant comparison is between unmarked definite DP and definite differential marked objects.

pragmatic effects and argue that aboutness topicality alone cannot account for the distribution of DOM in NGA.

5.2.1.2 Test #2: Endriss (2009)

In the previous section, I demonstrated that DOM can be paraphrased as an aboutness topic based on Reinhart's test. This section provides further evidence that differentially marked objects are interpreted as topics. I follow Dočekal and Kallulli (2012), who build on insights from Endriss (2009), that only quantifiers with particular lexical semantic properties can be mapped onto aboutness topics, and consequently they can be clitic doubled.

Crucially, Endriss studies aboutness topicality in her work and provides a formal semantic definition for it. Specifically, she proposes a notion of topicality that accounts for the fact that only some but not other quantifiers can be topic. See appendix B for an explanation of the exact mechanism. For our purposes the observation that only certain quantifiers can be topics is sufficient. Endriss observes that weak quantifiers, indefinites, and the universal *all*-quantifier are topicable, whereas monotone decreasing quantifiers, non-monotone quantifiers, the universal quantifier *every*, and monotone increasing quantifiers are non-topicable. I predict that if the same set of quantifiers that are argued to be topical are the only ones that can be differentially marked, then differentially marked objects are topics. As we will see below, the set of topical quantifiers, according to Endriss, is identical to the set of quantifiers that can be differentially marked in NGA.

The starting point is to demonstrate that not all contexts allow DOM. With the exception of 'the' and 'all,' definite nouns whose heads are strong determiners (in the sense of Barwise and Cooper 1981)[82] cannot be DOM in NGA as illustrated in (93c) – (93d).

(93) a. sara šara-at-o la-l-ktaab
Sara bought-3F.SG-3M.SG.OBJ DOM-the-book.M.SG
'Sara bought the book.'

[82] Quantifiers fall into two categories: weak and strong (Milsark 1977; Barwise and Cooper 1981). Strong quantifiers are those whose heads are strong determiners such as the, all, most, every, each, etc., while weak quantifiers are those whose heads are weak determiners a, some, many, a few, ∅, two… etc.

b. sara šara-at-on la-kull al-kotob
Sara bought-3F.SG-3PL.OBJ DOM-all the-books.BP
'Sara bought all the books.'

c. * sara šara-at-on la-ʔg'lab al-kotob
Sara bought-3F.SG-3PL.OBJ DOM-most the-books.BP
('Sara bought most of the books.')

d. * sara šara-at-o la-kull ktaab
Sara bought-3F.SG-3M.SG.OBJ DOM-every book.M.SG
('Sara bought every book.')

In NGA, *kull* 'every' is ambiguous between two readings 'all' and 'every.' When 'every' is interpreted as 'all,' the following DP must be definite and plural as in (93b). On the other hand, when 'every' is interpreted as 'every', then it must be followed by an indefinite noun, and in this case, it also has to be singular as shown in (93d). For example, while *kull al-kotob* 'all the-books' allows DOM (93b), while *ktaab* 'a book' headed by 'every' in (93d) does not. One might claim that only quantifiers that select for DPs but not NPs can be overtly marked. However, a counterexample to this claim is (93c) because 'most' selects for definite nouns in NGA yet still bans DOM.

Having shown the distribution of strong quantifiers with DOM, I turn to the distribution of weak quantifiers. I demonstrate below that the data are more complicated since not all weak quantifiers behave the same with respect to DOM. While definite nouns whose heads are numerals can appear in DOM contexts in NGA (94a), weak monotone increasing (such as *at least three books* (94b)), weak monotone decreasing (such as *at most three books* (94c)) and non-monotone (94d) quantifiers cannot.[83]

[83] The behavior of partitives with DOM seems to be puzzling. On the one hand, 'some,' as a weak quantifier, allows DOM only if it is interpreted as partitive:

(i) sara šaf-at-on la-qesm men t-tollaab mbirḥ
Sara saw-3F.SG-3PL.OBJ DOM-some from the-pupils.BP yesterday
'Sara saw some of the pupils yesterday.'

(94) a. sara šara-at-on la-l-thalth/arbʕ kotob
Sara bought-3F.SG-3PL.OBJ DOM-the-three/four books.BP
'Sara bought three/four books.'[84]

b. * sara šara-at-on la-th-thallath kotob ʕl-ʔaqal
Sara bought-3F.SG-3PL.OBJ DOM-the-three books.BP on-less
('Sara bought at least three (of the) books.')

c. * sara šara-at-on la-l-thallath kotob ʕl-ʔakthar
Sara bought-3F.SG-3PL.OBJ DOM-the-three books.BP on-most
('Sara bought at most three books.')

d. * sara šara-at-on la-l-thallath kotob bi-ẓẓabṭ
Sara bought-3F.SG-3PL.OBJ DOM-the-three books.BP in-exactness
('Sara bought exactly three books.')

According to Endriss' (2009) observation relying on the lexical semantic properties of a quantifier and its ability to function as an aboutness topic, weak quantifiers, indefinites, and the universal *all*-quantifier are topicable, however, monotone decreasing quantifiers, non-monotone quantifiers, the universal quantifier *every*, and monotone increasing quantifiers are non-topicable. We found exactly the same set of quantifiers which are topicable can also be clitic doubled in Arabic. Specifically, this set consists of numerals and the universal quantifier 'all DP' which freely allow DOM, but does not include weak monotone increasing, weak monotone decreasing, and non-monotone quantifiers, hence offering evidence that differentially marked objects are topics.

The pattern in NGA is close, though not identical to the Albanian CD patterns presented in Dočekal and Kallulli (2012). The only exception is monotone increasing

On the other hand, 'most (of)' is understood as partitive without the overt partitive marking but prohibits DOM (see example (93c)). I leave the investigation of partitives for future work. For more on partitives with DOM in Hebrew, see a recent work by Hacohen et al. (2021).

[84] Both numbers expressing a low quantity such as 'three' and transdecimal numerals (numerals larger than 10) can appear with DOM. As Ouwayda (2014) observes, nouns occur in the singular when they follow transdecimal numerals; thus, the dislocated DP needs to be in singular. The clitic agrees with the DP and must be plural-marked.

quantifiers. Endriss (2009: 252) acknowledges that in some cases monotone increasing non-exhaustive quantifiers such as 'several' can be interpreted as a bare numeral weak quantifier similar to *n,* and under this reading, it can be interpretable as a sentence topic. While monotone increasing quantifiers can be clitic doubled in Albanian (Dočekal and Kallulli 2012), they cannot be clitic doubled in NGA.

So far, we have established that differentially marked objects are aboutness topics based on two tests: (i) Reinhart's test based on intuition, and (ii) Endriss' (2009) quantifier distributional test. In the next section, I discuss other notions used to describe DOM such as referentiality, specificity and givenness. I argue that none of these notions provide an accurate account for DOM in NGA.

5.2.2 Referentiality

In this section, I test whether differentially marked objects are referential following the notion of referentiality in Lambrecht (1994: 127).

(95) Under the ***referentiality notion***, a topic is the referent that the proposition is about.

Lambrecht (1994: 335) states that "A topic entity must exist in the universe of discourse independently of what is being predicated of it in each proposition, i.e., it must be a discourse referent." According to this definition, referentiality is understood as a requirement on topics. This restriction prevents non-referring expressions such as quantified NPs from appearing in topic positions. Endriss' (2009) point is that topicality is about referentiality and individuals. This raises the question of why some quantifiers can be topics since quantifiers are of type ⟨⟨e,t⟩,t⟩, which by definition are not-referential. In response, Endriss (2009) proposes that quantifiers can obtain a referential-like behaviour via the anchoring to a referential address. We have seen in section 5.2.1.2 that certain quantifiers which function as topics (Endriss 2009) are found to allow DOM. I therefore conclude that referentiality cannot be the relevant dimension for DOM.

Another reason to dismiss referentiality as a description of DOM is that referentiality cannot distinguish between marked and unmarked objects. According to

Turner (2018: 146), the definite article in Levantine Arabic dialects (spoken in Lebanon and Syria) occurs with both anaphoric and non-anaphoric definite referents. Turner (2018) proposes a model, based on Dryer's (2014) reference hierarchy, that can systematically account for variation in the morphosyntactic strategies used to mark different degrees of definiteness in Arabic dialects. Anaphoric definite nominals describe a unique entity, known to the speaker, that is assumed to be retrievable for the listener because it is already actively present in the immediate discourse. Non-anaphoric definites describe a unique entity, known to the speaker, that is assumed to be retrievable for the listener because it is unique within the shared world of the speaker and listener. In Levantine, both nominals can be marked with the definite article *al-*.[85] Thus, unmarked definite objects are not necessarily non-referential in Arabic.

5.2.3 Specificity

I argue that specificity cannot be a possible characterization for DOM in NGA. I follow the conception of specificity in the sense of Enç (1991).

(96) *Specificity*: (after Enç 1991: 24)
Specificity involves […] being a subset of or standing in some recoverable relation to a familiar object.

According to Enç, specific DPs are linked to another DP that has already been established in the context. A definite DP DP2 (which counts as specific in her approach) picks up an entity that is already known to the hearer (= that of another DP DP1), i.e. the link is established via identifying DP1 as DP2. According to Enç (1991), all strong determiners (which are also specific in Enç's terminology) pick up an already established set in their restrictor. If specificity is the right dimension for DOM in Arabic, we would expect that DOM is compatible with all strong determiners, with no exception. However, this is not

[85] For a comprehensive discussion of definiteness in Levantine dialects, see Turner (2018: 146-162).

what we find. The examples in (93) show that only certain strong determiners are possible with DOM.

Another reason to rule out specificity as a relevant criterion for DOM is that indefinite nominals can be interpreted as specific. This is demonstrated in section 2.1 and repeated below.

(97) *Indefinites nominals can be specific (Syrian Arabic)*
laazim nʕmil-l-u **ši muqaddime** la-ḥatta maa-yinsʿidim
necessary make.IMPF.1PL-DAT-3M.SG some preparation so-PURP NEG-be.shocked.IMPF.3M.SG
'We need to arrange some sort of preparation for him so he won't be shocked.'
(Brustad 2000: 27)

Therefore, I exclude the notion of specificity to be relevant for describing DOM.

5.2.4 Givenness

This section argues that givenness cannot be the right characterization for DOM either because it is possible to have definite given NPs that are unmarked in NGA.

There has been a long-standing debate about whether givenness is an obligatory property of topics. Some scholars (e.g., Hockett 1958; Kuno 1972; Gundel 1985, 1988; Hedberg 1990) defend the view that topics must denote old information. However, others hold the opposite view (Reinhart 1981; Molnár 1993; Lambrecht 1994; Frey 2000, 2004; Endriss 2009) in which the fact that most sentence topics are given is not an inherent feature of the topics themselves, but rather due to general requirements on discourse structure (Reinhart 1981). The idea is that in the discourse, we typically have two adjacent sentences that have to be linked to each other, either because they provide information about the same referent or because there is a clear relation between the two propositions expressed in the two sentences. Most discourses are connected by a referential link. This is the reason why most topics are discourse given. Based on Reinhart (1981) if a topic is discourse-given, this

means that a discourse referent for this topic already exists in the common ground, which in turn means that the information conveyed by the sentence can straightforwardly be added to the common ground, because it can be 'stored' under the 'address' of this discourse referent. In this case, the common ground can be updated with the conventional meaning.

I adopt Schwarzschild's (1999) definition of givenness, as stated in (98).

(98) *Givenness*: (after Schwarzschild 1999: 151 (25))
An utterance U counts as GIVEN iff it has a salient antecedent A and
a. if U is type e, then A and U corefer;
b. otherwise: modulo ∃–type shifting, A entails the Existential F-closure of U.

Recall that differentially marked objects must be definite NPs and definite NPs in NGA are consistently given. I demonstrate below that a definite given NP can be non-differentially marked.

In (99), a friend was curious about submitting an assignment after some time they did not talk. A friend asks her friend after some time of not talking:

(99) *Unmarked DP is given*
šu sar? sallamt-i **l- waẓif-e** ?
What happened? submitted-2F.SG the-assignment-F.SG
'What happened? Did you submit the assignment?'

According to Schwarzschild (1999), the DP object is given since it has a salient antecedent, 'the assignment' would be correctly interpreted as given.

Similarly, in the following scenario we find that the given DP which is compatible with the definition of givenness in (98) is unmarked.

(100) *Unmarked DP is given*
roḥet alyom la-ʕend **d-daktor** ʕašin axoḏ al-fḥosʿaat
went.1SG today to-loc the-doctor in order to take.1SG the-tests
'I went to the doctor in order to take the tests.'

Givenness appears to be an accurate characterization for DOM since differentially marked objects are definite and definite NPs are given in NGA. However, as shown above, we can have definite given NPs that are unmarked. We can conclude then that givenness is not a relevant dimension for DOM.

5.2.5 Section summary

I have concluded that the core property for DOM is aboutness topicality. I have dismissed other properties such as referentiality, specificity, and givenness suggested in the DOM literature. As shown earlier, referentiality and specificity can both be rejected based on the discussion with quantifiers (section 5.2.1.2), thus it was superfluous to distinguish between marked and unmarked objects within these notions. Moreover, I have dismissed the conception of givenness in the previous section since unmarked objects are given; see Table 5.1 for a summary.

	DOM	**NON-DOM**
Givenness	✓	✓
Aboutness topicality	✓	✗

Table 5.1: Interpretation distinction between marked vs. unmarked objects

Thus far, I have established that the relevant description of DOM is aboutness topicality, making it a clear-cut diagnostic for the characterization of DOM in Arabic. I turn now to discuss other information-structural properties such as types of foci.

5.3 Excluding other informational-structural properties

I reached the conclusion that differentially marked objects i.e., *la*-phrase are aboutness topics. In order to make sure that other information-structural properties do not interfere with this conclusion, I still need to examine how types of foci, information focus and contrastive-focus, interact with DOM. I argue that DOM cannot appear under either information focus and contrastive focus. Given that in Arabic contrastive focus behaves

similarly with contrastive topic with respect to prosody, I also set aside the possibility that DOM is possible under contrastive topics.

5.3.1 DOM in NGA cannot be information focus

This section rules out the possibility that information focus can appear with DOM.[86] I use Kiss' (1998) definition of information focus.

(101) *Information focus* (or presentational focus) is an element that denotes purely new and non-presupposed information. Information focus operates on an open set of entities, and it is typically tested by recourse to a question/answer form.

I demonstrate below that DOM is illicit under information focus. By way of illustration, consider the following scenario:

Two classmates were talking. Classmate A asks B:

(102) A: min ʕazam-et ʕala-ʕidak?
who invited-2M.SG on-your_birthday
'Who did you invite to your birthday?'

B: ʕazam-et sˁḥab-i men l-jiš
invited-1SG friends-my from the-army
'I invited my friends from the army.'

B': *ʕazam-at-on la-sˁḥab-i men l-jiš
invited-1SG-3PL.OBJ DOM-friends-my from the-army
('I invited my friends from the army.')

As (102B') shows, the information focus *la-sˁḥabi* 'DOM-my friends' is not allowed to appear with DOM. The phrase 'my friends' conveys non-presupposed information and is chosen from a set of alternatives involving an open set of possible referents.

[86] Brustad (2000: 354) observes that the marked material cannot be new information. She does not apply a test to check this generalization.

Another example supporting the claim that information focus constituents are barred with DOM is given in (103).

(103) *Information focus is banned with DOM in NGA*

A: maʕ min raqasˁet fi-l-ḥafle?
with who danced at-the-party?
'Who did you dance with at the party?'

B: raqasˁet maʕ aḥmad
danced.1SG with Ahmad
'I danced with Ahmad.'

B': * raqasˁet maʕ-o la-aḥmad
danced.1SG with-3M.SG.OBJ DOM-Ahmad
('I danced with Ahmad.')

In (103), *aḥmad* cannot be differentially marked since it is chosen from a set of alternatives involving an open set of possible referents who danced in the party; thus we conclude that DOM is infelicitous under information focus DOM which operates on an open set of alternatives. I turn to the following section where I argue that DOM in NGA cannot host contrastive foci either.

5.3.2 DOM in NGA cannot be contrastive focus or contrastive topic

Contrastive focus occurs when constituents are linguistically prominent (Gundel 1999), evoking alternatives within a set of entities. Kiss (1998) distinguished two types of focus: information focus and contrastive focus.[87] I assume the following definition of contrastive (i.e., identificational) focus (Kiss 1998: 245).

[87] Kiss (1998) connects information focus to the terms *presentational focus* and *broad focus* and *identificational focus* to the terms *contrastive* and *narrow focus* (cf. Jackendoff 1972; Rochemont 1986; Ladd 1980). The concept of contrastive focus has been equated with narrow focus in the literature (Kiss 1998), Ladd (1980) argues that these in fact represent two different parameters of focus. While narrow focus refers to the size of a focus constituent, focus refers to the discourse characteristics of a focus constituent. See more on the differences between the two terms.

(104) *Contrastive focus* represents a subset of the set of contextually or situationally given elements for which the predicate phrase can potentially hold; it is identified as the exhaustive subset of this set for which the predicate phrase actually holds." In other words, contrastive focus exhaustively identifies the items for which the predicate holds, excluding any other items.

Given that contrastive focus by definition involves exhaustivity (i.e., it identifies a subset), it follows that quantifiers that do not involve 'exhaustive identification' are not good candidates for contrastive focus. Based on the DOM data containing quantifiers presented in section 5.2.1.2, quantifiers such as 'all' in (105), repeated from (93b), which do not involve exhaustive identification, can be DOM.

(105) sara šara-at-on la-kull al-kotob
Sara bought-3F.SG-3PL.OBJ DOM-all the-books.BP
'Sara bought all the books.'

Consider another example illustrating that DOM is incompatible with contrastive focus adapted from Erteschik-Shir (1997: 12(5)).

(106) *Contrastive focus is incompatible with DOM*

A. min sara tzawwaj-at-(*o), (*la)-aḥmad wlla (*la)-mḥemmad?
who sara married-3F.SG-3M.SG.OBJ DOM-Ahmad or DOM-Mohammad
'Who did Sara marry, Ahmad or Mohammad?'

B. tzawwaj-at-(*o) (*la)-aḥmad
married-3F.SG-3M.SG.OBJ DOM-Ahmad
'(She) married Ahmad.' UNMARKED; CONTRASTIVE FOCUS

B.' * tzawwaj-at-o la-aḥmad
married-3F.SG-3M.SG.OBJ DOM-Ahmad
Lit. ('(She) married him Ahmad.')
Intended: ('She married Ahmad.') *DOM; CONTRASTIVE FOCUS

I conclude, therefore, that marked phrases cannot be contrastive focus in NGA.[88] Similarly, in other languages such as Manambu (spoken in Sepik region), Burmese (Sino-Tibetan, Tibeto-Burman) and Neo-Aramaic dialect of Telkepe (spoken in Iraq), DOM is illicit when direct objects are interpreted as contrastive focus (Aikhenvald 2008; Coghill 2014).

Further argument against a contrastive focus analysis of DOM is that contrastive focus is argued to carry a pitch accent in Arabic which aligns with the intonation pattern, as has been pointed out in previous work (Bolinger 1961; Lambrecht 1994; Zimmermann 2008). This behaviour is incompatible with the prosodic properties of right dislocation as being referred to DOM (see Chapter 3), which generally involves prosodic deaccentuation, i.e., the lack of pitch accent (Escandell-Vidal 2009; Feldhausen 2010; López 2016; Fernández-Sánchez 2017).

Contrastive focus behaves prosodically in a similar way with contrastive topic. It has been argued that contrastive topics receive a rising intonation (see e.g., Gundel 1985; Hedberg 1990; Umbach 2001; Yeou et al. 2007; Krifka and Musan 2012; Büring 1997, 2014). We expect that, similar to contrastive focus, DOM is illicit with contrastive topic since contrastive topics in Arabic receive a secondary pitch accent which cannot characterize the differentially right-dislocated object. The ungrammaticality of DOM with a contrastive topic is exemplified in (107).

Suppose that a family is moving out of the country, a friend asks:

(107) A: šu maʔ al-bit w s-sayyara?
What with the-house and the-car
'What about the house and the car?'

B: *ʔajjarn-a la-l-bit w beʔn-ha la-s-sayyara
rented-3M.SG.OBJ DOM-the-house and sold-3F.SG.OBJ DOM-the-car-F.SG
('We rented the house and we sold the car.')

[88] Alzayid (2022), however, argues that in MSA right dislocated phrases are interpreted as contrastive focus. He provides evidence for this generalization by showing cases which cannot be right-dislocated: indefinites and universal quantifiers. The quantifier 'all DP' and the indefinite nouns do not involve a subset of a set whose members are presupposed; thus, they cannot behave as identificational (contrastive) focus according to Kiss' (1998) definition.

The example in (107) illustrates that although *the house* and *the car* are aboutness topics, they cannot be contrasted with each other when they are differentially marked; thus, no rising accent is indicated. From this, I conclude that differentially marked DPs cannot behave as contrastive topics.[89]

The fact that DOM is incompatible with both contrastive topic and contrastive focus is also not surprising. Contrastiveness is what generally underlies the interpretation of left dislocation constructions in Arabic as shown in (108) (see Moutaouakil 1989; Ouhalla 1994; Aoun and Benmamoun 1998; Albuhayri 2019; Alzayid 2022).

(108) A: šu šereb aḥmad?
what drank Ahmad
'What did Ahmad drink?'

B: **qahwe,** šereb aḥmad, meš šay
coffee drank Ahmad not tea
'It was coffee that Ahmad drank, not tea.'

Differentially marked objects are restricted in their position, that is, they cannot occur in the left edge of the clause. Thus, given that left dislocation has interpretive import rooted in contrastiveness, it is expected that marked objects cannot be interpreted with a contrastive flavour.[90]

The list of information-structural properties of the marked object in NGA is summarized in (109).

[89] Unlike in NGA, DOM in Early Modern Romanian has a strategy to achieve a contrastive topic reading (in Lambrecht's 1994 sense) on the direct object (Hill 2013). This observation, however, is not found in Modern Romain as pointed out by Hill (2013).

[90] Further evidence for the claim that differentially marked objects cannot be contrastive focus is that contrastive focus can be iterated (Kiss 1998: 248), and differentially marked objects lack this property:

(i) * sallamt-a_k la-ʔlo$_i$, la-l-waẓife$_{1k}$, la-ʕli$_{2i}$
submitted.1SG-3F.SG.OBJ DAT-him, DAT-the-homework-F.SG, DAT-Ali
('I submitted the homework to Ali.')

(109) *The NGA differentially marked object cannot be …*

a. information focus

b. contrastive focus

c. contrastive topic

Additionally, I add the information-structural properties examined in section 5.2 in (110) to the summary.

(110) *The NGA differentially marked object also cannot be characterized as …*

a. referential

b. specific

c. given

Having established that differentially marked objects in NGA are interpreted as aboutness topics in section 5.2.1, I turn now to reviewing a relevant information-structural property of DOM across languages.

5.4 A note on givenness and clitic doubling across languages

Recall that DOM in NGA obligatorily involves CD and it has been shown in section 3.6 that DOM shares some properties with CD. In this section, I set out to demonstrate that cross-linguistically, specifically throughout the Balkan languages, givenness has been argued to be associated with CD structures (Anagnostopoulou 1994; Leafgren 1997; Rudin 1997; Kallulli 2000, 2008; Harizanov 2014; Runic 2014, among others).[91] Crucially, scholars have linked CD to givenness when they mean 'given' as an element previously mentioned in the preceding discourse. In the CD literature, some scholars equate the notion of givenness to topicality and the two are often used interchangeably, although semantically, givenness must be distinguished from topicality (Halliday 1967; Krifka 2008; Kučerová and Neeleman 2012).

[91] Some authors proposed that givenness is a defining factor in other Semitic languages (e.g., Kifle 2007; Bekins 2014; Coghill 2014) but this claim is hard to evaluate since the data reported in the literature are not sufficient.

In Bulgarian, for example, CD is felicitous with a DP that has been previously mentioned (e.g., Ivanov 2009; Runic 2014). By way of illustration, consider the following examples in (111). The lack of CD is considered semantically infelicitous.

(111) *Bulgarian* (Ivanov 2009: 19, (4))

A: Njakoj vižda**l** li e Ivan dnes?
somebody seen-PART.Q is Ivan today
'Has anybody seen Ivan today?'

B: Ivan #(go) vidjax sutrinta
Ivan him-CL.DAT see-1SG.PST in the morning
'I saw Ivan in the morning.'

Another Balkan CD language discussed in relation to givenness is Albanian. Kallulli (2000, 2008, 2016) and Dočekal and Kallulli (2012) show that when clitic-doubled objects are discourse-given as they appear in the question, they are obligatorily marked. The data in (112) demonstrate that clitic doubled objects must be discourse-given.

(112) *Albanian* (Dočekal and Kallulli 2012: 117 (13), (15))

A: Who read the book?

B: Ana *(e) lexoi librin
Anna CL.ACC.3S read book.the
'Anna read the book.'

When 'the book' is discourse-new, CD is ungrammatical (Dočekal and Kallulli 2012). While in Albanian CD must be present with discourse-given objects, it is strongly preferred in Greek (Kallulli 1999, 2016) as shown below.[92,93]

[92] Kallulli (1999, 2016) uses the term familiarity to refer to clitic doubled object as given.

[93] Crucially, in both languages, Albanian and Greek, direct object CD is incompatible with direct object DPs that are contained in focus domains (Kallulli 1999).

(113) *Greek* (Kallulli 2016: 164, (9))

A: What did Ana do with the book?

B: I Ana ?(to) dhiavaseto vivlio
the Anna CL read the book
'Ana read the book.'

The presence of CD in Greek has also been attributed to discourse givenness by Anagnostopoulou (1994, 2006): the doubled DP is linked to a previously mentioned element.[94]

All in all, CD has been widely linked to givenness,[95] but it is crucial to note that the view adopted in the literature is different from the view adopted here for givenness. As stated in (98), for an element to be given, it needs to have a contextually salient antecedent.

Future research is still required to examine whether this generalization holds true in other CD languages not discussed here.

As far as Arabic is concerned, we have examined several scenarios in section 5.2.4, where unmarked objects are interpreted as given, and accordingly concluded that givenness is not a defining property for DOM.

5.5 Conclusion

I have proposed that aboutness topicality underlies the interpretation of DOM in Arabic. I have also shown that aboutness topicality is a clear-cut diagnostic between marked and unmarked nominals. To ensure that no other information-structural properties are involved, I also tested other information-structural notions such as referentiality, specificity, and givenness which are also used to describe DOM, and demonstrated that none of these

[94] See also Kouneli and Kushnir (2021).

[95] Other CD languages have been discussed under the notion of discourse givenness. For example, in French, "The [doubled] DP must already have been mentioned in the discourse. The construction does not introduce a new entity" (Chapman 2012: 7). Similarly, Sikuku and Diercks (2022) argue that in Banu languages (Lubukusu) the object-marked doubled must be discourse-given.

notions are relevant for the characterization of DOM in NGA.

This chapter has motivated the analysis of DOM movement to the phase edge of *v*P. Aboutness topics require referential anchoring and referential anchoring requires a particular syntactic position (e.g., Frey 2000, 2004). I have argued that the obligatory movement to the edge of the phase is a direct consequence of topics requiring association with a referential address. Thus, movement of the differentially marked object to Spec, ApplP, being at the *v*P phase edge, allows the object to be anchored to a referential address, making it the aboutness topic. Since information structure is involved, the DOM movement displays freezing effects in which differentially marked objects cannot move to a higher position, as shown in chapter 4.

In the following chapters, I argue that aboutness topicality will have consequences for: (a) how the overall discourse will be modulated, and (b) what nominals can partake in DOM. The rest of the thesis discusses these two topics in order.

Chapter 6

DOM as a non-default Discourse Commitment Device

6.1 Introduction

I have argued in Chapter 5 that aboutness topicality is the defining property for DOM in NGA. By analyzing various DOM scenarios, I observe that aboutness topicality does not fully account for the distribution of DOM in NGA. I expound on cases of DOM in NGA where using DOM, as opposed to its non-differentially marked counterpart, may yield certain pragmatic interpretations ranging from the speaker's emotive content, to making a correction, to accommodation, to activating a parallel Question Under Discussion. The chapter fully adopts recent work by Kučerová and Zarka (in prep) which argues that DOM in NGA behaves as an illocutionary marker and that it grammatically marks the asserted proposition as a non-default Discourse Commitment (Gunlogson 2001; Farkas & Bruce

2010; Rett 2021). Following Farkas and Bruce (2010), I assume that non-default assertions have additional discourse pragmatic effects and override the default updates of the common ground (*cg*, a set of propositions already confirmed by the discourse participants).

Crucially, as argued in Chapter 4, DOM marking syntactically encodes aboutness topicality such that the obligatory movement to the edge of the phase is an outcome of aboutness topics requiring association with a referential address. I follow Endriss' (2009) argument that aboutness topic marking triggers anchoring of the asserted proposition to a referential address. Using the *Table* of Farkas and Bruce (2010), a discourse tracking device, we implement the illocutionary contribution of DOM as an addition of a proposition that is not a member of the projected set (a superset of the *cg*). It is suggested that the non-default Discourse Commitment function arises from the speaker using the aboutness topicality to grammatically mark that the union of the Discourse Commitment of the speaker and the Discourse Commitment of the other interlocutor(s) is incoherent. As a result of the incoherent communication, the move calls for a revision of the *cg*. This chapter also clearly connects the illocutionary effects of DOM derived in the chapter to the additional applicative syntactic structure I proposed for DOM in Chapter 4.

6.2 DOM and pragmatic effects

This chapter's core claim is that aboutness topicality alone does not fully explain the distribution and interpretive effects of DOM in NGA. As will be illustrated below, DOM systematically contributes additional illocutionary content which pertains to how the speaker is using the utterance in context.

First, DOM is frequently used to correct a propositional content, as exemplified in (114).

[Context: The two older sisters were preparing for their younger sister's birthday party. They invited some friends. They look at the homemade cake on the table and one of the friends asks:]

(114) A: min zayyan al-kaʕke?
who decorated.3M.SG the-cake
'Who decorated the cake?'

B. One sister replies:

ʔana
I.NOM
'Me'

C. Another sister heard that her sister lied:

laʔ, ʔana zayyant-a **la-l-kaʕk-e**
no, I decorated.1SG-F.SG.OBJ DOM-the-cake F.SG
'No, the cake was decorated by me.'

Speaker C knows that her sister (B) did not tell the truth, so she feels the necessity to correct previous information. When the same scenario does not use DOM, the absence of DOM in (114C) does not yield the speaker's emotive content. This will be explained later in the section.

DOM is also used to trigger an accommodation. Accommodation is usually conceived as a repair strategy: If a piece of information cannot be interpreted with respect to the current common ground, then the current common ground can be minimally changed in a way that fits the requirement of the piece of information (see, e.g., Stalnaker 1978; Lewis 1979; Heim 1982; von Fintel 2008). Consider the following scenario.

[Context: During the Covid-19 time, a friend knows that his friend is afraid of getting covid and he always wants to keep himself safe, so he declares:]

(115) A: ʔnti ʔakid meš jay ʕ-l-ḥafle
you sure NEG come to-the-party
'You are definitely not going to the party.'

B: bas maʔana ʔaxaḏt-o **la-ṭ-ṭʕiim**
but I took-3M.SG.OBJ DOM-the-vaccine
'But I took the vaccine.'

The speaker uses DOM to accommodate his friend's declaration not by changing his own beliefs about covid, but by declaring that he took the vaccine so it should not be an issue joining the party. However, when the same scenario is uttered without DOM, the speaker utters (115B) in a more natural way and their goal, unlike when DOM is used, is merely to provide new information.

Lastly, speakers use DOM to express unexpectedness towards the addressee. Consider the following scenario.

[Context: a grandfather saw his grandchildren playing in the yard and said:]

(116) ʔenzalo, kassart-o-ha la-š-šajara
go_down, broke-3PL-3F.SG.OBJ DOM-the-tree
'Get down! You broke the tree.'

The grandfather did not expect his grandchildren to climb or break the tree; that is, it is not normally an expected behavior.

When the same context contains the non-differentially marked object 'the tree,' according to the speaker's intuitions, the speaker does not really care about the tree being damaged. It would also appear that the kids usually misbehave, and it is something the grandfather deals with more frequently.

Consider another scenario involving unexpectedness towards the addressee, the scenario adapted from Levin (1987).

[Context: Mother was angry about her children misbehaving. Father exclaimed to them:]

(117) ḥaraam … ɣallab-to-ha la-ʔem-ko
pity tired-3PL-3F.SG.OBJ DOM-mother-2PL.POSS
'It's a shame! You have tired your mother!'

The goal of the DOM propositional content is to tell the children to stop misbehaving towards their mother. The unexpectedness comes from the father's expectation that his children should behave well and that they are not meeting the father's expectation. Example (117) is grammatical without DOM but signals that the speaker prefers not to utter it under the context of unexpectedness since the children most likely misbehave often, thus there is nothing mismatching the father's expectation.

Moreover, according to the speaker's native intuitions, DOM is frequently associated with emotive content. Building on Levin (1987), all the scenarios described above involve different emotions such as anger, disappointment, sadness, upset etc. For instance in (114), the speaker is upset or angry about the sister lying; in (115), the speaker is upset when they heard that their friend is assuming that they are not coming to the party, and in (116) the grandfather is surprised about his grandchildren climbing and breaking the tree, etc. DOM scenarios may involve more than one emotion.

So far, we have shown that DOM exhibits illocutionary content. This content is absent with other instances of aboutness topics, such as sentential subjects. By way of illustration, consider the following examples.

(118) a. [sewaʔ eḏa mneʕmal-o essa aw baʕdin] raḥ ykon bešeʕ
whether if do-it now or later FUT becomes ugly
'Whether we do it now or later it is going to be ugly.'

b. [ʔenno hi ʕerfet] meš ktiir fareq maʕ-i
that she knew NEG very matter with-me
'That she knew does not matter to me.'

It has been argued that sentential subjects are associated with the aboutness topic interpretation (e.g., Lacerda 2020). In the examples above, the sentential subject is interpreted as an aboutness topic (what the sentence is about), but unlike DOM, sentential subjects in NGA do not yield additional illocutionary content such as correcting previous information, triggering accommodation, or expressing unexpectedness towards the addressee. This supports the claim that aboutness topicality cannot fully account for the DOM interpretation.

6.3 Deriving the illocutionary content of DOM

The question that raises now is how to account for the illocutionary properties of DOM illustrated in the previous section. The proposal put forth by Kučerová and Zarka (in prep) detailed below accounts for the illocutionary content of DOM which pertains to how the speaker is using the DOM-utterance in context.

The proposal implements Farkas and Bruce's (2010) framework. Before delving into the proposal, let me first review Farkas and Bruce's model of a discourse structure K which includes the following components:

(119) The common ground (*cg*), the set of propositions believed by all discourse participants (for the purpose of the conversation).

(120) Sets of discourse commitments (*DC*): for each participant *x*, the set of propositions *x* publicly commits to in the conversation.

(121) The Table *T*, a discourse tracking device.

(122) The projection set (*ps*), the set of beliefs that are being considered for addition into the *cg*. A conversational move that places an item on the Table simultaneously projects a set of future common grounds relative to which the issue on the Table is decided. These projected sets are supersets of the current *cg*.

Gunlogson (2001) defined the *cg* in terms of the participants' commitment sets, specifically, the *cg* is the union of the discourse commitments of the participants in the conversation. A crucial innovation of Farkas and Bruce's (2010) approach is the separation of the *cg* and Discourse Commitment sets. "The discourse commitment set of a participant A at a time *t* in a conversation *c* contains those propositions A has publicly committed to in the course of *c* up to *t* and which have not (yet) become mutual commitments. The *cg*, on the other hand, is that set of propositions that have been agreed upon by all participants in *c* at *t*

together with the propositions that represent the shared background knowledge of the discourse participants" (p.85). This allows for participants to negotiate the *cg* independently of their own public beliefs.

Following Farkas and Bruce (2010: 90), I take non-default assertions to have additional effects and override the default updates. Default assertions, on the other hand, are performed by uttering a simple declarative.

Rett (2017)'s detailed description of different types of non-default content, focusing specifically on the difference between descriptive and illocutionary content, is represented in Figure 6.1.

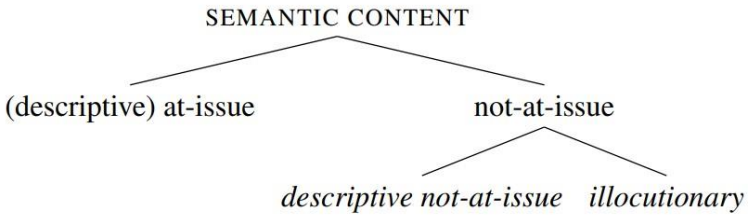


Figure 6.1. Types of semantic content

Rett argues that illocutionary encoders contribute information about how the speaker intends the utterance to be understood. Descriptive not-at-issue markers are part of the descriptive content of an utterance, but which nevertheless contribute not at-issue content. For example, appositive constructions, as in (123) contribute descriptive not-at-issue content.

(123) John, a former syntactician, lives in Canada.

The description in (123) is classified as "canonical not-at-issue content" and not as illocutionary content since the speaker does not comment on a speaker's attitude toward the content of any part of the proposition.

Teasing apart the difference between descriptive and illocutionary not-at-issue content (parallel to non-default content) is important for the context of DOM. Specifically, I argue that DOM behaves as a non-default Discourse Commitment device. DOM in NGA functions as an illocutionary marker and it grammatically marks the asserted proposition as a non-default Discourse Commitment (Gunlogson 2001; Farkas & Bruce 2010). Non-

default utterances, unlike default utterances, are marked with regard to how they affect the discourse themselves, especially in terms of the commitments they impose. Using the Table of Farkas & Bruce (2010), Kučerová and Zarka (in prep) implement the illocutionary contribution of DOM as an addition of a proposition that is not a member of the projected set. The proposition must be in the *ps* if it is coherent with the conversation state. The addition of the proposition makes the union of the speaker's Discourse Commitment and the Discourse Commitment of the other participant(s) ($DC_A \cup DC_B$ which A and B are the speaker and the participant), internally incoherent. For example, in the scenario in (115), participant A believes that participant B will not attend the party since he assumes that he did not get the Covid-19 vaccine. However, this public commitment is incoherent with participant B's belief that he indeed received the vaccine. This incoherency is a result of the commitment lists of the two participants (speaker and addressee) being mutually incoherent. Although $DC_A \cup DC_B$ is incoherent, this does not result in oddness or the conversation being in crisis, and consequently, the proposition calls for the revision of the *cg*.[96]

Given that differentially marked objects are interpreted as aboutness topics, I follow Endriss (2009) in that the aboutness topic serves as an argument in the asserted proposition, and thus aboutness topic markings trigger the anchoring of the asserted proposition to a referential address. Aboutness topicality is a necessary precondition. I argue that a proposition can only be added to the Table if it can be anchored. Such anchoring is either done via an existing *ps* (if the proposition has been projected as a member of *ps*), or the proposition must invoke its independent anchoring. In DOM contexts, the differentially marked object is the topic of the entire utterance or, more precisely, the speech act. For instance, 'the mother' in (117) constitutes the topical constituent in this way, the entire speech act would be an assertion about the mother, about whom it is asserted that the father declared that the children misbehaved toward her. In such DOM propositions, the aboutness topic DP anchors the proposition via its referential address.

[96] The current proposal explains when DOM arises and that is a result of a pragmatic nature. DOM must arise when the propositions are incoherent. When the propositions of the conversation are coherent, the use of DOM is redundant.

(124) *Deriving the illocutionary content of DOM*

a. Differentially marked objects are interpreted as aboutness topics.

b. Aboutness topic marking triggers anchoring of the asserted proposition to a referential address.

c. Once the DOM-proposition (which is not a member of the *ps*) is anchored, it can be added to the Table.

d. Once it is added, the addition of the proposition makes $DC_A \cup DC_B$ internally incoherent.

e. The proposition calls for a revision of *cg*.

Now the question is how the emotive content of the speaker is derived in certain DOM contexts. I follow Rett (2021) in that what constitutes an 'emotive attitude' is "a strict subset of epistemic attitudes which characterize the speaker's emotion towards a proposition in addition to their epistemic relation to that proposition." Rett further defines emotive markers as they encode the speaker's emotive attitude towards some proposition made salient by the utterance in which they occur. In DOM, the emotive content ties to the DOM proposition being anchored (i.e., the proposition added to the Table). The most notable consequence of this fact is that the emotive content cannot target the differentially marked entity, instead it is associated with the whole proposition.

6.4 Structural economy

The remaining question is why all aboutness topics do not yield the same illocutionary effects. Kučerová and Zarka (in prep) suggest that the obligatory illocutionary effect of DOM arises from structural economy. In chapters 3 and 4, I have argued that DOM involves an additional structure: clitic doubling (big-DP) and an applicative projection. This structure, however, is absent with non-DOM objects and with non-object aboutness topics (e.g., sentential subjects). This additional structure licenses interpretive effects (illustrated in section 6.2) that would not be available otherwise (e.g., Fox 2000; Sichel and Wiltschko 2021). The DOM scenarios shown in section 6.2, describe that the speaker uses DOM to accomplish communicative goals. According to the speaker's intuitions these goals, as

noted, are not preferred under a non-DOM context. This follows from structural economy, which dictates that an ApplP projection is more costly than a simple structure and needs to be motivated. The use of a larger structure of DOM when there is a simpler non-DOM structure available yields an additional pragmatic interpretation.

6.5 Conclusion

This chapter has argued that aboutness topicality alone does not fully explain the distribution and interpretation of DOM in NGA. It has focussed on the pragmatic effects of DOM and provided an account of how these effects arise. The next two chapters will concentrate on which types of nominals may be differentially marked and specifically how aboutness topicality affects which nominals partaking in DOM.

Chapter 7

Individuation: Countability or atomicity?

7.1 Introduction

This chapter examines (i) which types of nominals may be differentially marked, and (ii) what determines why certain types of nominals may be differentially marked in DOM but not others. As shown in Chapter 2, DOM is licit only with definite objects (e.g., Abu-Haidar 1979; Levin 1987; Aoun 1999; Brustad 2000). Beyond definiteness, previous theoretical works argue that individuation is the key factor in licensing Arabic DOM (e.g., Brustad 2000, 2008; cf. Khan 1984). I follow the view that individuation is the key for licensing DOM, but a finer understanding of individuation is needed to capture the DOM facts in NGA. Thus, in this chapter I examine individuation from the perspective that individuation has both morphosyntactic and semantic properties (e.g., Borer 2005; Deal 2017; Rothstein 2017; Grimm 2018; Grimm and Dočekal 2021). While morphosyntactically, individuation

parallels countability; semantically, individuation parallels atomicity: a denotation of an individuated noun is atomic if it has salient individuable entities (e.g., Rothstein 2010).

The chapter consists of three main parts. The first part reviews previous work by Khan (1984) and Brustad (2000) on individuation and object marking. I show that their descriptions are not quite accurate. I then review Zarka's (2021) generalization on DOM and countability in Arabic. Zarka (2021) shows that countability is a relevant dimension for describing the distribution of nominals with DOM. I advance an empirical generalization that only countable nominals can participate in DOM. I describe nouns as countable or non-countable, where these two terms are restricted to designating nominal behavior in terms of morphosyntactic characteristics and make no reference to semantic characteristics.

I argue that Zarka's (2021) generalization that countability is the right dimension for describing the distribution of nominals with DOM is inaccurate. I examine DOM facts with nominals and the findings show that DOM is illicit with countable nominals denoting a kind interpretation. Given that kinds are morpho-syntactically countable, I conclude that countability, contra Zakra (2021), cannot be the relevant characterization.

Having dismissed the possibility that individuation under the notion of countability is relevant, I examine an alternative notion of individuation: atomicity. The second part of the chapter presents a recent experimental study conducted by Zarka and Hacohen (2023). The study tests whether individuation under the notion of atomicity, a semantic property concerning the atoms e.g., salient individuable entities, best describes the distribution of nominals with DOM in NGA. I tested count, object mass, substance mass, and collective nouns. The results of the experimental study challenge the hypothesis that atomicity is the correct dimension for characterizing the DOM nominal distribution. The results are better described following Grimm's (2012) scalar view of atomicity. I take the distribution of the nominals to be based on a binary distinction, while I treat atomicity itself as a scale: only nominals that are high on the atomicity scale may be differentially marked.

This chapter is organized as follows: section 7.2 reviews the preliminary observations about the types of nominals occurring in Arabic DOM (Khan 1984; Brustad 2000; Zarka 2021). Section 7.3 discusses an experimental study on the nominal distribution of DOM (Zarka & Hacohen 2023) which examines whether atomicity is the right

dimension for describing DOM in NGA. Section 7.4 discusses the DOM results presented in the previous section with respect to Grimm's (2012) view of atomicity. Section 7.5 concludes.

7.2 Preliminary observations about DOM in Arabic

This section first reviews observations made by Khan (1984) and Brustad (2000) with respect to which types of nominals can partake in DOM. I show that their views do not entirely capture the DOM data in NGA. Then I review Zarka's (2021) observation that only countable nominals can participate in DOM.

7.2.1 Previous work on individuation and object marking in Arabic

Khan (1984) investigates the patterns of occurrences of object marking across ancient Semitic languages. In particular, he demonstrates that in several languages, definiteness is not a sufficient condition for licensing object marking, that is, object marking does not occur with all definite nominals. Moreover, in some contexts, definiteness is not even a necessary condition for object marking, in that object marking may also occur with indefinite nominals. However, this is not true for NGA since DOM must only appear with a subset of definite objects (shown in section 2.1).

In addition to the definiteness, Khan also explores what he calls 'Individuation.' Khan uses the term individuation as a cover term for notions such as *specificity, concreteness, qualification,* etc. Khan (1984) observes that individuated nominals are more likely to participate in DOM than non-individuated ones. Different definitions of individuation have been proposed in the literature. Individuation was developed to explain the qualities that account for object marking in Semitic languages. Further, Brustad (2000) modifies Khan's (1984) hierarchies of individuation by adding a new characteristic to Khan's list called *quantification*. According to Brustad (2000), quantification is related to the status of the noun as either a collective or countable. Specifically, a quantified nominal

marked with a numeral from 2-10, is individuated.[97] Thus, it is more likely that quantified nominals trigger object marking or any other syntactic marking such as agreement.

Although Brustad has shown that nominals that possess the quantification property are more likely to trigger object marking, she does not provide data showing that the reverse is true, namely that nominals that do not possess this property are less likely to be differentially marked. Importantly, Brustad's definition of quantification does not consider other classes of nominals such as singular marked, plural marked, and nouns combined with quantifiers. In Zarka (2021), I address this gap by examining *countability* more closely with respect to DOM. Based on distributional data, I advance a novel generalization that only countable nouns can be part of the DOM construction.[98] This generalization will be detailed in the next section.

7.2.2 Zarka (2021): Countability and DOM

Quantification or number have rarely been discussed as a factor that influences DOM (Woolford 1995, 1999; de Swart 2003) in comparison to other properties such as animacy and definiteness, as shown in Chapter 2, which have received much attention in the DOM literature.[99] Zarka (2021) chooses to focus on one property of individuation, called *countability*, and examines this property closely with respect to DOM. The aim of this section is to elaborate on Zarka's (2021) empirical generalization stating that the DOM construction in NGA is attested only with countable nominals. This restriction can be shown in (125), thereby barring the presence of DOM with non-countable nominals.

(125) *Generalization I: nominal type with DOM*

a. ✓DOM: countable nominal.

b. * DOM: non-countable nominal.

[97] Brustad does not mention the behavior of object marking with numerals higher than 10.

[98] Zarka (2021) did not test numerals higher than 10 with DOM. See footnote 84 that such numerals are compatible with DOM.

[99] In some languages such as in South Saami, number has been observed to correlate with DOM. Singular direct objects, regardless of their definite status, are always marked with accusative; whereas plural objects bear accusative case only when they are definite (Kroik 2016). In Palauan, an Austronesian language, DOM is affected by number, in addition to specificity and animacy (Woolford 1995).

On the other hand, we do not find such restriction for unmarked nominals as illustrated in (126).

(126) *Generalization II: nominal type with non-DOM*

✓Non-DOM: countable and non-countable nominal.

Countable nouns are licit with DOM, irrespective of whether their number morphology is suffixal (127a)-(127b), or templatic (127c). The first (suffixal) involves a concatenative process, while the templatic is a non-concatenative process.

(127) *Countable nouns are licit with DOM*

a. sara šara-at-a la-š-šant-a
Sara bought-3F.SG-F.SG.OBJ DOM-the-bag-F.SG
'Sara bought the bag.'

b. sara šara-at-on la-š-šant-ein/-āt
Sara bought-3F.SG-3PL.OBJ DOM-the-bag-DUAL/-SP_F[100]
'Sara bought the two bags/the bags.'

c. dalia šara-at-on la-l-karaasi
Dalia bought-3F.SG-3PL.OBJ DOM-the-chairs.BP
'Dalia bought the chairs.'

Their equivalent unmarked versions, as shown in (128), are also grammatical.

(128) a. sara šara-at š-šant-a
Sara bought-3F.SG the-bag-F.SG

[100] The sound feminine plural (glossed as SP_F) can occur with both human nouns with conceptual gender (a) and grammatically feminine nouns whose denotation is genderless i.e., non-human nouns (i).

(i) a. mhands-e → mhnds-āt b. ṭawl-e → ṭawl-āt
engineer-F.SG engineer-SP_F table-F.SG table-SP_F
'engineer' 'engineers' 'table' 'tables'

‘Sara bought the bag.’

b. sara šara-at š-šant-ein/-āt
Sara bought-3F.SG the-bag-DUAL/-SP$_F$
‘Sara bought the two bags/bags.’

c. dalia šara-at l-karaasi
Dalia bought-3F.SG the-chairs.BP
‘Dalia bought the chairs.’

Such countable nominals are freely combined with cardinal numerals and thus are allowed with DOM. Chierchia (2010) calls this property of a noun being directly modifiable by a numeral as ‘signature property’ of countable nouns.

(129) a. sara šara-at-on la-l-xams šant-āt
Sara bought-3F.SG-3PL.OBJ DOM-the-five bag-SP$_F$
‘Sara bought the five bags.’

b. dalia šara-at-on la-l-arbaʕ karaasi
Dalia bought-3F.SG-3PL.OBJ DOM-the-four chairs.BP
‘Dalia bought the four chairs.’

Another type of countable nouns in Arabic is the broken plurals (BPs). As observed in Ouwayda (2014), BPs allow both kind and unit readings.[101] In (130), *karaasi* 'chairs' can be interpreted as either 'kind' or 'unit.'

(130) sara šara-at arbaʕ karaasi
Sara bough-3F.SG four chairs.BP
'Sara bought four chairs.'

[101] It has been noted for Modern Standard Arabic that certain patterns of broken plurals are connected to different meanings, paucity vs. abundance (Wright 1898, 1:233–4; Fischer 2002, 53–64).

→ Sara bought exactly four individual chairs (4 total)

→ Sara bought exactly 4 kinds of chairs (e.g., if 2 of each kind, then 8 chairs total)

Interestingly, DOM causes the BP to have a unit interpretation but blocks the kind interpretation, as shown in (131).

(131) sara šara-at-on la-l-arbaʕ karaasi
Sara bought-3F.SG-3PL.OBJ DOM-the-four chairs.BP
'Sara bought the four chairs.'

→ Sara bought exactly four individual chairs (4 total)

↛ Sara bought exactly 4 kinds of chairs (e.g., if 2 of each kind, then 8 chairs total)

The BP data suggest that DOM is possible when the BP denotes a unit reading but is impossible with a kind reading even if the kinds are generally countable (see more on the discussion of nominals with kind-denoting reference in section 7.3.6 and chapter 8).

Having discussed countable nouns, I turn now to non-countable nominals. Zarka (2021) examines two classes of non-countable nominals: substance-mass and collectives. Unlike countable nominals, substance-mass nominals (glossed as SM) cannot be pluralized (132), and cannot be combined with cardinal numerals (133). Note that in certain contexts, (133) can have the reading of 'four kinds/types of asphalt.'[102]

(132) zeft-e → * zeft-āt
asphalt-F asphalt-SP_F
('asphalts')

(133) * arbaʕ zefte
four asphalt.SM
('four asphalt')

[102] It has been noted that some substance mass nouns can be pluralized and that depends on the context. The following context has the reading of 'two kinds of wine' (Rothstein 2017: 85(3b)).

(i) We had two (different) wines with dinner, a Chablis and red Beaune. We drank two glasses of each.

The following examples illustrate that mass nouns which denote substances like ‘asphalt’ and ‘blood’ are incompatible with DOM (134). Note that the kind reading of such substance mass nouns is also not possible under a DOM context.

(134) a. * aḥmad baʕ-a la-z-zeft-e
Ahmad sold.3M.SG-3F.SG.OBJ DOM-the-asphalt.SM-F
(‘Ahmad sold the asphalt.’)

(Zarka 2021: 6(12))

b. * n-narse šaf-at-o la-d-dam
the-nurse-F.SG saw-3F.SG-M.SG.OBJ DOM-the-blood.SM
(‘The nurse saw the blood.’)

However, substance-mass counterparts i.e., non-differentially marked objects, can appear in baseline contexts.

(135) a. aḥmad baʕ z-zeft-e
Ahmad sold.3M.SG the-asphalt.SM-F
‘Ahmad sold the asphalt.’

b. n-narse šaf-at d-dam
the-nurse-F.SG saw-3F.SG the-blood.SM
‘The nurse saw the blood.’

The second class of nouns that behaves like substance mass with respect to the countability property is collectives (glossed as coll). Collectives in Arabic are a class of nouns that have a singular morphological form but are neither singular nor plural in meaning.[103] Their

[103] The fact that collectives are morphologically singular can be supported through the singular agreement they trigger on the adjectives that modify them. The third masculine singular is the default realization in NGA.

(i) a. šajar yabis
trees.COLL dry-Ø

b. samak maṭbox
fish.COLL cooked-Ø

c. namel sɣir
ant.COLL small-Ø

meaning is described by traditional Arabicists as generic, group-like, or mass-like. Intuitively, collectives refer to groups of people as well as natural classes of animals, insects, vegetables, minerals, etc. rather than to the individual members of each class (Gaudefroy-Demombynes & Blachère 1952).[104,105]

In Arabic, a singular form called the SINGULATIVE (individual-denoting) can be derived from the collective form through a gender shift, as shown in (136).

(136) a. *baqar* cows.COLL.M 'herd of cows' → *baqar-a* cow-F.SG 'A cow'

b. *šajar* trees.COLL.M 'group of trees' → *šajar-a* tree-F.SG 'A tree'

c. *jaaj* chickens.COLL.M 'flock of chickens' → *jaaj-e* chicken-F.SG

A sound plural can in turn be derived from this newly formed singulative noun. Arabic grammarians refer to the singulative as *esm el waḥda* 'noun of unity' (Wright 1933: 147; Ojeda 1992). Ojeda (1992: 307) defines singulatives as "lexical items that are derived from collectives and refer either to a specific quantity of the substance or to an individual member of the collect." The feminine suffix *-a*, identified in the literature as a singulative, plays an individuative role and acts as a classifier (Cowell 1964; Wright 1967; Greenberg

Also, Abd-Rabbo (1990) notes for Classical Arabic that 'cows' *baqar* (the collective form) triggers masculine singular agreement.

(ii) haḏa baqar
this.M.SG cow.M.SG
'this group of cows'

[104] There are different definitions for collective nouns in the existing literature. For instance, Talmoudi (1980: 132) defines the collective as denoting either "a collection of things or animals regarded as a unit," or else "a mass or volume." Wright (1933: 147) describes the collective as expressing "the genus or whole," and Abdel-Massih et al. (1981, 49) define the collective as a noun that designates "a class or mass of like things without counting the units that make up the mass." For Fleisch (1961: 65), a collective noun is "the mass wherein the individuality of the 'amassed' is effaced." All of these definitions share the idea of inherent plurality of objects viewed as substances. Ojeda (1992) defines Arabic collective nouns as basic lexical items that indicate either "a substance or material in the mass" or "a collection of objects viewed as a totality without reference to the individual members" (referring to Erwin 2004: 166).

[105] This observation, however, is challenged by Dali (2020: 85) where it is argued that collectives are ambiguous between individuated and unindividuated semilattices. She tests individuation based on Bale and Barner's (2009) comparative construction test. To control for each reading of the collective, see Zarka and Hacohen (2023), presented in section 7.3.

1972; Ojeda 1992; Mathieu 2009, 2012, 2014; Zabbal 2002; Fassi Fehri 2003, 2012; Ouwayda 2014; Dali 2020, among others). Examples in (137) illustrate the derivation steps.

(137) a. *baqar-a* cow-F.SG 'A cow' → *baqar-a-āt* cow-F.SG-SP$_F$ 'individual cows'

b. *šajar-a* tree-F.SG 'A tree' → *šajar-a-āt* tree-F.SG-SP$_F$ 'individual trees'

c. *jaaj-e* chicken-F.SG 'A chicken' → *jaaj-a-āt* chicken-F.SG-SP$_F$ 'individual chickens'

In contrast to count nouns, but similar to substance mass nouns, collectives are non-countable. As we can see from the derivation steps of collectives in (137), collectives cannot be directly pluralized as two operations need to be applied. In addition, collectives cannot combine with a cardinal numeral as shown in (138) below.[106,107]

(138) a. *arbaʕ baqar
four cows.COLL
('four group of cows')

b. *xamse jaaj
five chickens.COLL
('five flocks of chickens')

The following examples show that while the singulative and sound plural are acceptable with DOM, collectives are not.

[106] According to Ouwayda (2014), collectives (*batch* nouns in Ouwayda's term) can combine with numerals if they are used as grocerese nominals. They behave as grocerese nominals when used as ordering portions of food or counting portions of grocery items. An example from Lebanese Arabic with collectives as grocerese is provided below (Ouwayda 2014: 109(186)).

(i) xamseh teffeeḥ iza betriid
Five apple.COLL if will.3MS
'Five apples please'

I add that the interpretation of (i) is 'five kilos of apples.' This reading has been called in Rothstein (2017: 236) as pseudopartive measures and it has been distinguished from attributive measures such as 'a five-kilo apples.'

[107] The same facts hold true of Tunisian Arabic (see Dali 2020).

(139) a. sara šara-at-a la-l-jaaj-e SINGULATIVE
Sara bought-3F.SG-3M.SG.OBJ DOM-the-chicken-F.SG
'Sara bought the individual chicken.'

b. sara šara-at-on la-l-jaaj-āt SOUND PLURAL
Sara bought-3F.SG-3PL.OBJ DOM-the-chicken-SP
'Sara bought the individual chickens.'

c. * sara šara-at-on la-l-jaaj COLLECTIVE
Sara bought-3F.SG-3PL.OBJ DOM-the-chicken.COLL
('Sara bought the flocks of chickens.')

However, the ungrammatically of (139c) disappears when collectives appear in non-DOM contexts.

(140) sara šara-at l-jaaj
Sara bought-3F.SG the-chicken.COLL
'Sara bought that kind/flock of chickens.'

Table 7.1 summarizes the facts presented in this section. Thus far, the data support the empirical generalization from Zarka (2021) that only countable nominals can participate in DOM.

		DOM	Non-DOM
COUNTABLE NOUNS	Singular, dual, sound plural, singulative	✓	✓
	Broken plural	✓	✓ count reading * kind reading
NON-COUNTABLE NOUNS	Substance mass	✗	✓
	Collectives	✗	✓

Table 7.1: Countability and DOM in NGA

However, Zarka's generalization cannot be accurate when we consider the facts about DOM with nominals denoting kinds (130)-(131). The data show that the interpretive ambiguity of broken plurals disappears when the DP is differentially marked: the kind interpretation is no longer available, but only the individuated reading. Kinds are morpho-syntactically countable and yet they are incompatible with DOM, therefore, I conclude that countability cannot be the right characterization of DOM in NGA.

The observation that nouns denoting kinds are illicit with DOM in NGA is also found in other languages like Romanian and Spanish. Mardale (2008) illustrates that in both Romanian and Spanish bare plurals denoting kinds cannot be differentially marked.

The next section uses an experimental study of DOM in NGA (Zarka and Hacohen 2023) to examine an alternative notion of individuation for describing which noun phrases may be differentially marked.

7.3 Experimental study on DOM: Zarka and Hacohen (2023)

The previous section concluded that Zarka's (2021) generalization that countability plays a role in DOM does not hold true in NGA. As demonstrated in the previous section, although countable nouns may be differentially marked, kinds are morpho-syntactically countable and still they are illicit with DOM, thus I concluded, contra Zarka (2021), that individuation under the notion of countability cannot be the relevant dimension for DOM.

It has been argued that individuation has both syntactic and semantic properties (e.g., Borer 2005; Zabbal 2002; Lima 2014; Deal 2017). Morphosyntactically, individuated nouns are countable. Countability is evidenced by using distributional tests such as susceptibility to pluralization and the ability of the noun to combine directly with

numerals.[108] Semantically, individuation parallels atomicity: a denotation of an individuated noun is atomic if it has individuable entities.[109]

Having established that countability is irrelevant for describing DOM, in the experimental study (Zarka and Hacohen 2023), I test the hypothesis of whether individuation under the notion of *atomicity* is the right dimension for describing DOM in NGA. We collected Arabic speakers' acceptability judgments for various types of nominals in DOM and non-DOM sentences in order to explore which type of nominals are allowed with DOM in NGA, and their degree of acceptability. We measured speakers' responses to DOM constructions while manipulating nominal type. The stimuli were presented auditorily and included count nouns, collectives, substance mass nouns, and an additional class of so-called "object mass" nouns e.g., *ʔθaθ* 'furniture.'[110] Object mass nouns pattern like substance mass nouns in Arabic with respect to their distributional properties and like count nouns with respect to their atomic denotation. Object mass nouns cannot be pluralized (141a) and cannot be combined with numerals (141b).

(141) a. ʔθaθ → * ʔθaθ-āt
furniture furniture-SP_F
(*furnitures')

b. *xamse ʔθaθ
five furniture
('five pieces of furniture')

[108] There are other distributional tests such as the choice of determiners. Determiners are sensitive to the countability distinction. This test cannot be applied in NGA as the quantifier 'many' 'ktiir' is ambiguous between 'much' and 'many.' *ktiir* can select for substance and object mass and plural nouns, as demonstrated below:

(i) ktiir waḥle
much mud

(ii) ktiir ɣasil
much laundry

(iii) ktiir sayyar-āt
many car-F.PL

[109] In the literature, there are analogous replacement notions for 'atomicity' such as Landman's (2011) notion of overlap, Chierchia's (2010) stable atomicity and Krifka's (1989) divisiveness. It is beyond the scope of this thesis to go over each definition, here I choose to focus on atomicity.

[110] Object mass nouns have been identified by this name, or others like 'fake mass nouns' (Chierchia 1998, 2010), 'naturally atomic mass nouns' (Rothstein 2010), 'aggregate' (Huddleston and Pullum 2002), 'neat mass nouns' (Landman 2011) and 'artifactual aggregate' (Grimm 2012).

Atomicity has been used to distinguish between count, object mass, and substance mass nouns (Quine 1960; Chierchia 1998, 2010; Rothstein 2010; Landman 2011; Schwarzschild 2011; Grimm 2012, among others). Unlike substance mass nouns however, the denotation of object mass nouns denotes atoms; the minimal parts of ⟦*furniture*⟧ are individual, non-overlapping pieces of furniture such as chairs (Barner and Snedeker 2005; Bale and Barner 2009; Rothstein 2010; Schwarzschild 2011; Sutton and Filip 2016; Deal 2017; Erbach 2020).[111] Table 7.2 below summarizes the atomicity distinction across nominal types. Atomicity is presented as a binary distinction i.e., ± *atomic*. I will demonstrate below how atomicity plays out with collectives.

	COUNT	**OBJECT MASS**	**SUBSTANCE MASS**	**COLLECTIVES**
ATOMICITY	+	+	-	??

Table 7.2: The atomicity distinction across types of nominals

I use the distinction presented in Table 7.2 to test whether atomicity is the key factor in determining whether DOM is licit in NGA. I predict that if count and object mass nouns receive high acceptability scores compared to other types of nominals, then the DOM results are most likely characterized by atomicity. On the other hand, if substance and object mass nouns receive low acceptability scores compared to other types of nominals, then atomicity cannot be the right characterization since object-mass nouns are atomic.

An additional way to test whether atomicity is the right dimension for characterizing the distribution of nominals with DOM is the class of collectives. Recall that collectives are non-countable and, as shown earlier, are incompatible with DOM (see section 7.2.2). However, based on the atomicity distinction, collectives can be either atomic

[111] Barner and Snedeker (2005) and Bale and Barner (2009) show, experimentally, that object mass nominals must denote atoms. They use comparative tests like 'who has more N.' Such comparatives are assessed in different ways depending on whether they feature a noun with individuable entities, like *cat* or *furniture*, or a noun without individuable entities like *water*. The results indicate that object mass nouns like *furniture* behave as plural count nominals like *cats* and require comparison in terms of cardinality. Thus, both count and object mass nouns have denotations grounded in sets of individuals (however see Rothstein 2016 and Rothstein and Pires de Oliveira (2020) for an argument against this claim).

or non-atomic. Dali (2020: 85) claims that collectives in Arabic are argued to be ambiguous between having atomic and non-atomic reference. In order to determine whether atomicity is the right dimension for DOM, I manipulated the clitic form in the DOM construction (see the stimuli in section 7.3.2). Recall that the DOM construction obligatorily involves co-referentiality between the clitic and the associated differentially marked nominal. There are two possible forms of the clitic that agree with the collective: singular and plural clitic. Each form is associated with a distinct reading. The agreeing singular clitic provides the group reading of the collective, e.g., *group of chickens* (142a), whereas the plural one gives the atomic reading in which it refers to individual entities (not a group) e.g., *individual chickens* (142b).[112] Examples of collectives with each form of clitic will also be presented in the stimuli section.

(142) a. ?? dalia baʕat-**o** la-l-jaaj
Dalia sold-3M.SG.OBJ DOM-the-chicken.COLL
'Dalia sold the group of chicken.'

b.?? dalia baʕat-**on** la-l-jaaj
Dalia sold-3PL.OBJ DOM-the-chickens.COLL
'Dalia sold the individual chickens.'

The collective with the plural clitic in (142b) cannot be counted as shown in (143), thus regardless of the associated clitic, collectives can never be counted therefore the countability dimension is irrelevant for collectives.

[112] The idea to manipulate the clitic form comes from the observation that collectives in NGA can trigger singular and plural agreement, and each agreement is associated with a different reading. As mentioned in footnote 103, the collective 'cows' *baqar* typically triggers masculine singular agreement, which is associated with the group interpretation, but in some contexts, we can see the plural agreement which is associated with individuation reading.

(i) al-baqar berʕa barra
the-cow.COLL graze.3M.SG.PRS outside
'A group of cows graze outside.

(ii) al-baqar berʕ-o barra
the-cow.COLL graze-3PL.PRS outside
'Individual cows graze outside.'

This 'hybridity' in agreement is attested in another class of nouns called *plurative* (Fassi Fehri 2016, 2020). It is a type of plural often neglected, although quite productive, which refers to a group or a collection of individuals. Similar to collectives, pluratives control feminine singular agreement and plural agreement (see Fassi Fehri 2016: 229). I have not investigated this class of nominals with DOM, so it remains an area for future research.

(143) * dalia baʕat-**on** la-l-xams jaaj
Dalia sold-3PL.OBJ DOM-the-five chickens.COLL
('Dalia sold the five individual chickens.')

The collective test predicts that if DOM sentences containing collectives with the plural clitic receive higher acceptability scores than collectives combining with the singular clitic, then atomicity is a likely relevant dimension for characterizing the distribution of nominals with DOM.

7.3.1 Participants

A sample of 52 participants were recruited. Of these, 4 participants did not complete the questionnaire, leaving a final sample size *n* = 48 (Male = 19; Female = 29). Participants were born and currently live in Peki'in (a town located in North Galilee of Israel). Most of them are relatives and friends of the author and were recruited online through social media (Facebook, Instagram, WhatsApp groups).[113] The age of the participants ranged from 18 to 73 years old (Mean = 33.57; Median = 31). Participation in the study was on a voluntary basis and no compensation was offered.

7.3.2 Stimuli

Participants heard Arabic sentences with DOM. We manipulated the nominal type which included the following four classes: count, object mass, substance mass, and collectives. Within collectives, we also manipulated the clitic form (singular, plural), thus there were five different conditions in total. For each condition, there were six items for a total of 30

[113] Studies have shown that geographical region and religion affect the variability of using certain sounds among dialects spoken in Israel (Habib 2005; Hijjo & Fannouna 2014). In order to avoid potential confounds wherein participants would rate the sentences based on phonology, we only asked individuals who are Druze religion and live in Peki'in to participate in the study. Women who are married and moved with their spouses to live outside of Peki'in were also excluded from the study as their dialect may be affected by the region.

DOM-sentences. In addition to the DOM sentences, there were 10 non-DOM distractor sentences. An example of each experimental condition is presented in Table 7.3.

Condition	**Example**
1. COUNT[114]	al-walad rama-ha la-ṭ-ṭab-e The-boy threw-3F.SG.OBJ DOM-the-ball-F.SG 'The boy threw the ball.'
2. OBJECT MASS	nadia našrat-o la-l-ɣasil Nadia hung out-3M.SG.OBJ DOM-the-laundry.OM 'Nadia hung out the laundry.'
3. SUBSTANCE MASS	*aḥmad ɣassal-o la-d-dam Ahmad washed-3M.SG.OBJ DOM-the-blood.SM ('Ahmad washed the blood.')
4. COLLECTIVE +SG. CLITIC	dalia baʕat-**o** la-s-samak Dalia sold-3M.SG.OBJ DOM-the-fish.COLL 'Dalia sold the group of fish.
5. COLLECTIVE +PL. CLITIC	tia šafat-**on** la-l-baṭ Tia saw-3PL.OBJ DOM-the-ducks.COLL 'Tia saw the individual ducks.'

Table 7.3: Example of items in each condition

As illustrated by the examples, the structure of the sentences was kept uniform across conditions, with variations restricted to changes in the lexical items (subject, verb, object). Animacy was not controlled for in the experiment since DOM is not sensitive to animacy (shown in section 2.1). The two conditions with DOM sentences containing collectives (singular and plural) did not form minimal pairs in order to avoid repetition which could lead to boredom or increased acceptability ratings.

[114] In the study, the class of count nouns includes singular marked, plural marked, and broken plural.

7.3.3 Procedure

The study took place online. Participants heard the 40 sentences one at a time and provided acceptability ratings after each sentence. All test sentences were pre-recorded by a native speaker and produced using monotone, neutral intonation to control for any prosodic effects.[115] Stimuli were then presented via Qualtrics[xm] as an audio file that could be replayed with no limits. All sentences were automatically randomized using the Qualtrics randomizer. The average length for completing the task was around 8 minutes.

Participants were presented with the following instructions, translated here from Arabic: “for each sentence you hear, you need to determine if it is a sentence you would expect in our spoken dialect”. They were asked to rate their judgment using a 6-point Likert scale of acceptability, with only the extreme options explicitly stated: 1 (*akid laʔ* 'absolutely not') and 6 (*akid ʔaa* 'absolutely yes'). In the study, we took plausibility to mean acceptability.

We decided to use a scale from 1-6 which forces a choice between agreement (4-6) and disagreement (1-3). We opted not to use a 7-point scale to avoid having a midpoint rating which is often associated with different interpretations: "unsure," "I don't know," or "I don't understand the task" (e.g., Chyung, Swanson, Roberts & Hankinson 2018). It has also been pointed out that the middle-point behaves as a "get-away" option from extreme alternatives. Using a 6-point scale forces participants to choose between agreement or disagreement to different degrees without the alternative of not expressing their opinion at all (e.g., Pimentel 2019).

A gradable paradigm was used instead of the more commonly used binary judgment task because it is better able to reflect the scalar nature of acceptability judgments. Unlike grammaticality judgments, which are typically categorical (a sentence is either grammatical or not), acceptability judgments are considered inherently gradable (Chomsky 1965; Sprouse 2007 and references therein). Moreover, since “[i]n natural conversation, there are many moves available to an interlocutor who is asked to judge the validity of a statement” (Sikos, Kim & Grodner 2019: 2), reducing speakers’ available choices to only

[115] As mentioned in section 3.6, the differentially marked object is not separated from the rest of the sentence with an intonational break.

two possible response options artificially constrain their behavior (ibid. and cf. Sorace & Keller & 2005).[116]

7.3.4 Hypothesis and predictions

In the experimental study I test whether the atomicity notion of individuation best describes the distribution of nominals with DOM in NGA. This is presented in (144) below.

(144) *Hypothesis*

> If participants rate DOM sentences with count and object mass nouns as more acceptable than substance mass nouns, atomicity is a likely dimension to characterize the distribution of nominals with DOM.

Under this hypothesis, I predict that differentially marked nominals which denote atomic reference such as count and object mass nouns will receive much higher acceptability scores than nouns whose denotation is non-atomic such as substance mass. Object mass nouns cannot be atomic. If participants rate DOM sentences with object mass as highly acceptable I conclude that atomicity is a potential notion for characterizing the DOM results.

As stated earlier, atomicity also clearly manifests in the collective conditions. While the collectives with plural clitics denote an atomic reference, collectives with singular clitics denote a non-atomic reference. Thus, if participants rate DOM sentences containing collectives with the plural clitic as more acceptable than collectives combining with the singular clitic, atomicity is, again, a likely dimension to characterize the distribution of nominals with DOM.

[116] See similar use of this paradigm in Kasher and Hacohen (2023).

7.3.5 Results and analysis

We used mixed ordinal logistic regression to analyze participants' acceptability ratings for the different nominal types. The model indicated no significant main effect for participant gender (p = 0.7), therefore we discuss the ratings from participants as a whole for the remainder of the analysis. The results of the DOM response distribution are summarized in Figure 7.1.

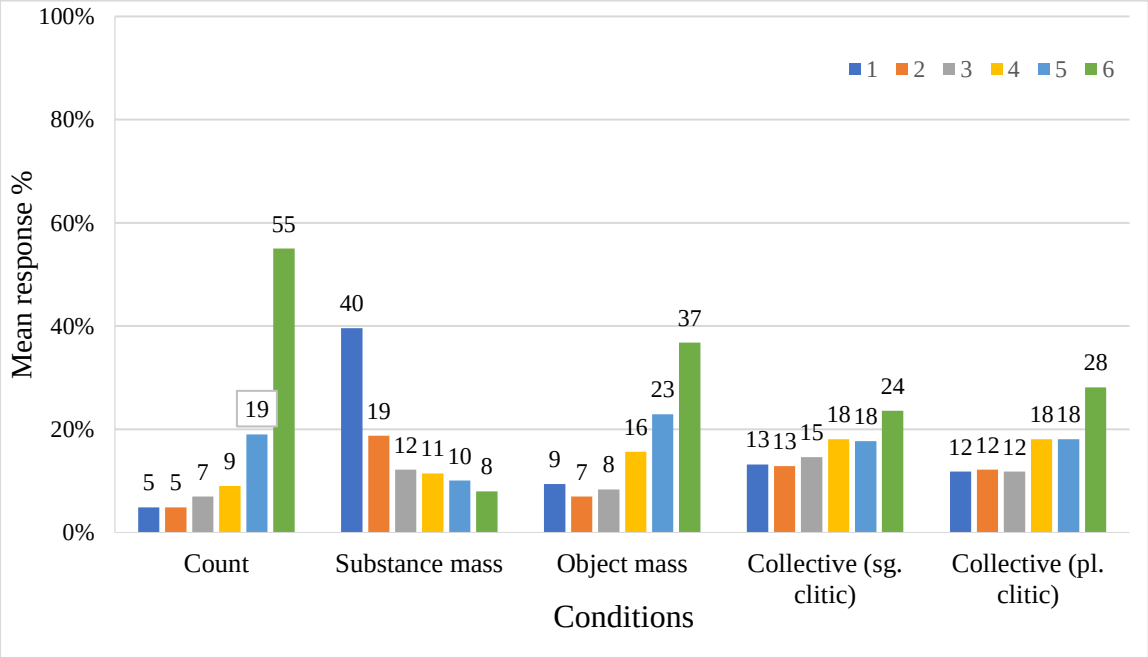


Figure 7.1. The distribution of DOM acceptability ratings across conditions

As we see in Figure 7.1, DOM sentences with count and object mass nouns received the highest acceptability scores (74% and 60% 5-6 ratings, respectively). Judgments for the two collective conditions were evenly dispersed across response options, with no apparent difference between collectives with singular clitics and collectives with plural clitics. Lastly, the substance mass condition received the lowest acceptability scores at only 18% 5-6 ratings, which had the largest proportion of 1 (completely unacceptable) rating. This is an important finding because it shows a strong preference for unacceptability as opposed to the collectives where participants' responses are equally acceptable (24% 6-ratings for the two collective conditions). Additionally, the more compelling argument for low

acceptability is that substance mass received 40% strongly unacceptable ratings (which is at least 4 times higher than any other condition). One can also argue that substance mass is as unacceptable as object mass nouns are acceptable, which constitutes a mirror image of each other.

Figure 7.2 shows the average acceptability ratings for each condition. The ordinal logistic model indicated no significant difference between the two conditions (p = 0.22). The model showed all other ratings different from each other and collectives (all p < 0.0001).

Although the distribution of ratings varied across conditions (Figure 7.1), the average acceptability rating for object mass and count nouns is quite similar. In addition, the average rating of the substance mass condition is quite similar to collectives; however combined with Figure 7.1, we notice that there is some convergence on a strong opinion that DOM sentences with substance mass nouns are not acceptable. The average rating for collectives with singular clitics and collectives with plural clitics was 3.05 and 3.12 respectively.

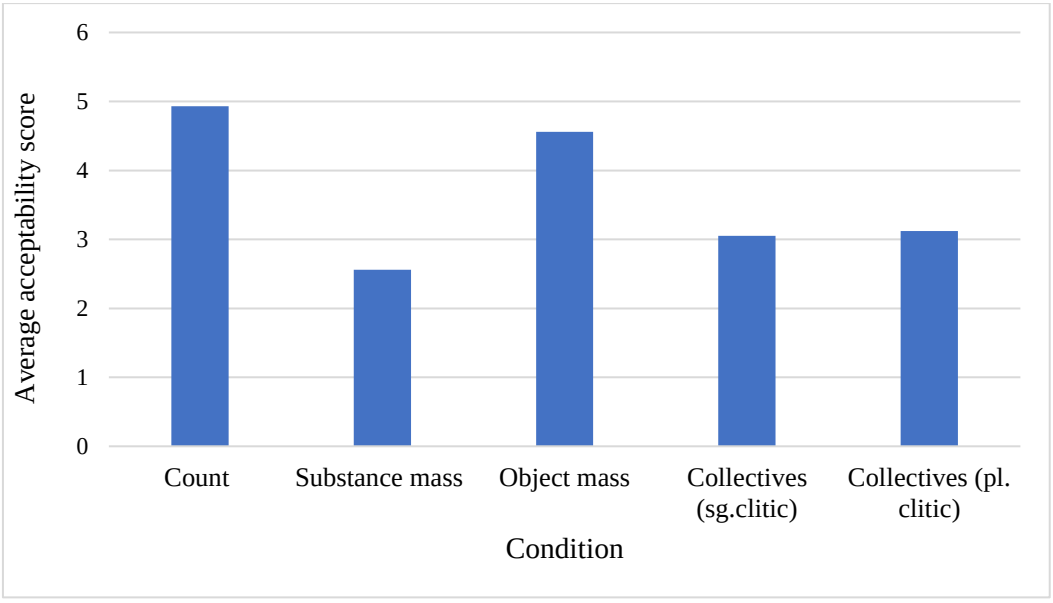


Figure 7.2. Average acceptability rating by condition

As for the non-DOM sentences, the results indicate an overall high acceptance average score (M = 5.78).

The item effect was also measured to ensure that all items are related and measured in a similar way. To check for an item effect, Cronbach alphas were calculated by condition and summarized in Table 7.4. Cronbach alpha gives an overall measure of the internal consistency, that is, how closely related a set of items are as a group. Values of 0.7 and above indicate strong similarity between items. In other words, the response values for each participant across a set of items are consistent suggesting that our measure is reliable.

Condition	**Cronbach alpha**
Count	0.84
Object mass	0.77
Substance mass	0.77
Collectives with pl. plural	0.82
Collectives with sg. Clitic	0.70

Table 7.4: Cronbach alphas per condition

As we can see in Table 7.4, all conditions obtained Cronbach alphas at 0.7 or higher, therefore the ratings for each item are reliable. Cronbach alphas were also measured when a single item is removed. If removing an item improves the value of said alpha score that means that the item should be omitted. Removing an item did not increase the value of the Cronbach alpha for any conditions, therefore we conclude that there was no significant item effect.

7.3.6 Discussion

Under the hypothesis (144), we expect that atomic nouns such as count and object mass nouns would receive high scores. On the other hand, substance mass nouns, which are non-

atomic should receive low scores. These predictions are borne out in the experimental findings ($M_{\text{count}} = 4.93$; $M_{\text{object mass}} = 4.56$; $M_{\text{substance mass}} = 2.56$).

Also under the atomicity hypothesis, we expect DOM sentences containing collectives with a plural clitic denoting an atomic reference to receive much higher acceptability scores than collectives with a singular clitic (i.e., denoting a non-atomic reference). However, this prediction is not borne out in the data. Based on the results of the experiment, the average acceptability ratings of the two conditions of collectives were similar regardless of the form of the associated clitic (singular $M = 3.05$ or plural $M = 3.12$). Hence, the results of collectives challenge the atomicity hypothesis (144). The atomicity results are summarized in Table 7.5.

		ATOMICITY	QUALITATIVE RESULTS	MEAN RATINGS
COUNT		✓	High	$M = 4.93$
SUBSTANCE MASS		✗	Low	$M = 2.56$
OBJECT MASS		✓	High	$M = 4.56$
COLLECTIVES	SG CLITIC	✗	No difference	$M = 3.05$
	PL CLITIC	✓		$M = 3.12$

Table 7.5: Atomicity and DOM across nominals in NGA

To summarize, the results of the experimental study provide evidence in favour of the atomicity hypothesis (144).[117] I further demonstrate how the distribution of DOM with

[117] I concluded in section 7.2.2 that countability is irrelevant for describing the nominal distribution of DOM in NGA based on the distributional facts of DOM with nominals denoting kind. If countability is the right dimension, we would predict count nouns to receive high acceptability scores, whereas the other nouns are non-countable and are predicted to receive low acceptability scores. However, this prediction is not borne out in the results. Object mass nouns are non-countable in Arabic yet received similar average acceptability rating ($M = 4.56$) to DOM sentences with count nouns ($M = 4.93$). Thus, contrary to Zarka's (2021) generalization,

nominals denoting kinds are predicted under the atomicity hypothesis. Recall the observation in section 7.2.2 about how the interpretive ambiguity of broken plurals disappears when the DP, specifically when the broken plural is differentially marked, the kind interpretation is no longer available. I assume, following Krifka (2003) and Dayal (2004), that kinds are represented as individual concepts of type ⟨s,e⟩ (i.e., functions that yield the totality of the manifestations of that kind in that world in any world). Thus, given how kinds are semantically represented, it is unexpected under the atomicity hypothesis that nominals denoting kinds would appear with DOM.

The atomicity distinction is limited only to a subset of nominals. In Chapter 9, I will elaborate on the atomicity distinction and argue that it does not account for all the types of nominals which appear with DOM. In particular, while certain quantifiers can appear with DOM (shown in section 5.2.1.2), quantifiers are not atomic. This issue serves as an avenue for future work.

Further, the results for the collective class pose a challenge for the atomicity view, therefore further investigations are needed to account for the non-uniform effect of atomicity across noun-types. Below, I demonstrate how the results of the experimental study align with Grimm's (2012) view of atomicity as a scalar phenomenon and are therefore expected.

7.4 Grimm's (2012) scale of individuation

In this section I discuss the results of the experimental DOM study in light of Grimm's (2012) analysis of individuation. Similar to Brustad's (2000) view of individuation as a continuum, Grimm argues that individuation, while often taken to be a binary distinction (atomic vs. non-atomic), should be viewed as a scalar phenomenon. I adapt Grimm's (2012) scale of atomicity to account for the non-uniform effect of atomicity across noun types in the experimental study in section 7.3.

Grimm argues that the mass/count distinction is best viewed as a scalar phenomenon, with languages dividing the scale of individuation in (145) in different

individuation defined by countability is not the right dimension for characterizing the distribution of nominals with DOM.

ways.[118,119] Nouns of different types are individuated to varying degrees and can be ordered accordingly along the scale of individuation:

(145) Scale of individuation (Grimm 2012: 68)

liquids/**substances** < granular aggregates < **collective** aggregates < **individuals**

Based on the individuation scale (145), nominals that denote individuals are highly individuated. Substance mass nouns are least individuated. Collective aggregates lie between substance mass and nouns denoting individuals.[120] Here, I discuss only these three classes of nouns as they are most relevant to the present study.

Working in the mereotopological theory i.e., the theory of part-whole relation, Grimm (2012) argues that the scale of individuation is derived via a *connection-relation* and he uses the notion of connectedness as a thesis concerning parts (atoms). The intuitive definition of connection is that two entities are connected if they share a common boundary. The connection relation may be strongly connected, proximate, or separated (Grimm 2012: 149). For instance, the denotation of substance mass such as *water* has *strongly connected* parts, i.e., parts that are internally connected to at least one other element in water.[121] In contrast to substance mass, the denotation of count nouns consists of *separated* individuals, e.g., a *bicycle* includes – among other parts – a frame and wheels, these parts are separated from each other. The last group of nouns are collectives which contain clustered individuals, related by *proximate connection* relations where two entities are co-located and near one another. A clustered individual under the proximately connected relation, then, will specify a group of individuals all within a particular distance of one another. For

[118] Grimm mentioned previous work by Lucy (1992) and Gentner and Boroditsky (2001) claiming that individuation has been related to scalar structure, however, the facts examined in these works were different.

[119] Grimm (2018) noted that cross-linguistically the scale of individuation has been adopted in languages including Arabic (Acquaviva 2008; Mathieu 2012), Czech, Hebrew (Doron and Muller 2013) and Swahili (Contini-Morava 2000).

[120] The class 'collective aggregate' is identical to collectives in Arabic as the nouns of the collective aggregate class specify a group of (clustered) individuals.

[121]Another way to characterize the relation between the minimal parts of the substance is 'locally strongly connected' (Grimm 2012: 142). The relation 'locally' connected is defined such that not every occurrence of water is connected, since clearly water may appear parceled out into puddles, rivers etc. However, locally, each occurrence of water is always attached to another instance of water, where the second instance may contain or be contained in the first.

example, insects or berries that do not appear in groups where each individual touches another, and are all at some distance from one another are considered proximately connected.[122] These connection distinctions are summarized in Table 7.6 where the nominal types are ordered based on the degree of the connection relation (adapted from Grimm 2012: 149).

Nominal type	**Connection relation type**
Count	Separated (none)
Collectives	Proximate
Substance	Strong

Table 7.6: Connection relation type and nominals based on Grimm (2012)

Having viewed individuation as a scalar phenomenon and its link to the degrees of connectedness, I now adapt Grimm's (2012) analysis to account for the results of the DOM experiment.

The acceptability ratings for the nominal types can be ordered along a scale. We have seen from the results that nominals that denote individuals such as count nouns are most likely to be differentially marked. Additionally, substance mass nouns such as *mud* which are less individuated are less likely to be differentially object marked. These findings align with the endpoints of Grimm's (2012) individuation scale.

Grimm (2012) does not, however, address object mass in his scale.[123] The experimental DOM results therefore extend our understanding of the behavior of object mass with respect to individuation. The acceptability ratings for object mass nouns with DOM received slightly lower ratings than count nouns. Unlike count nouns, object mass nouns are non-countable, but similar to count nouns, they denote atoms. Thus, I place object

[122] I do not expect a rating difference between DOM with collectives such as *berries* and those with collective such as *cows* since animacy does not play a role in DOM in NGA.

[123] Grimm (2012) restricts his research to what he called natural concrete entities (dog, water), to the exclusion of artifactual entities (hammer, furniture) or abstract entities (arrival, happiness). According to Grimm, the reason for this restriction is that natural concrete entities provide the firmest foundation for comparative studies. Only natural concrete entities were included in the current experimental DOM study.

mass nouns below count and above substance mass nouns, but much closer to count than to substance mass.

As for collectives, it has been argued that collectives in Arabic are ambiguous between individuated and non-individuated reference (Dali 2020). This claim is challenged by the collective DOM results where the average acceptability ratings for collectives were relatively low regardless of the form of the clitic. I argue that collectives occupy a lower position in the individuation scale closer to the lowest extreme of substance mass i.e. non-atomic nouns. Thus, both collectives and object mass nouns occupy the middle region of the atomicity hierarchy. While object mass and count nouns occupy the upper region (atomic), collectives and substance mass occupy the lower regions (non-atomic). I present my modified scale of individuation for Arabic based on the acceptability ratings from 7.3.5 in Figure 7.3. No distinction is made between the position of collectives associated with singular or plural clitics in the scale as their average rating results were statistically identical.

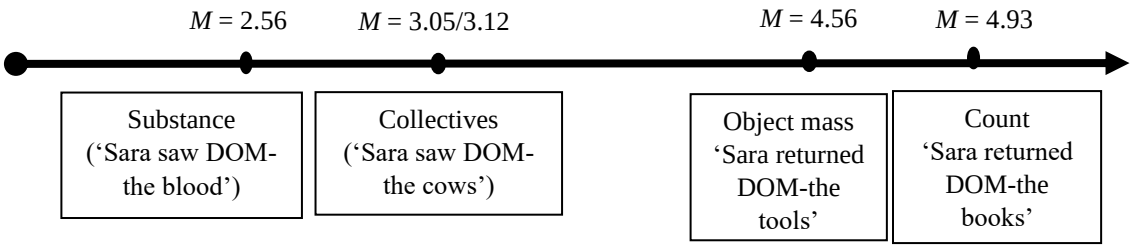


Figure 7.3. The scale of nominal types based on the acceptability ratings of DOM

To summarize this section, the results of the experimental study suggest that neither countability nor atomicity are relevant dimensions for the nominal distribution of DOM. The results are better described if we consider atomicity as a scale in Grimm's (2012) sense rather than a binary feature. The modified scale is only a descriptive way of restating the DOM experimental results in Grimm's view of atomicity.

7.5 Conclusion

I have established that contrary to Zarka's (2021) and Brustad's (2000) observations on countability presented in section 7.2, countability cannot be a plausible characterization for the nominals participating in the DOM construction. By examining the results of the experimental study, I have also established that DOM is instead attested with highly atomic nouns such as count and object mass nouns, and less likely to appear with nominals that are located low on the atomicity scale such as substance mass nouns. Crucially, the atomicity distinction is relevant only for a subset of nominals that can be differentially marked. I will discuss this limitation in Chapter 9 under open questions.

In the next chapter I lay out a semantic account in which aboutness topicality explains the atomicity restriction with DOM.

Chapter 8

Explaining the atomicity restriction

8.1 Why is there an individuation restriction on nominals?

The goal of this chapter is to answer the question of what factor(s) determine why certain types of nominals can participate in DOM but not others. In the previous chapter, I established that contrary to Zarka's (2021) and Brustad's (2000) observations on countability, countability cannot be a plausible characterization for the nominals partaking in the DOM construction. I have also established that DOM is instead attested with highly atomic nouns such as count and object mass nouns and less likely to appear with nominals that are located low in the atomicity scale such as substance mass nouns.

This chapter determines *why* highly atomic nouns nominals in particular are the most likely to be differentially marked. I explore whether the atomicity scale presented in section 7.4 can be tied to the DOM properties discussed in the previous chapters. Specifically, I ask:

(146) a. Is this restriction due to dative marking?
b. Is this restriction due to dislocation?
c. Is this restriction due to aboutness topicality?
d. Is this restriction due to clitic doubling?

In sections 8.1.1 and 8.1.2, I establish that neither dative marking (146a) nor dislocation (146b) can be responsible for the distribution restriction with DOM. Further, I discuss in section 8.2.1 how topicality explains why certain nominals denoting atomic references are found to be more highly differentially marked than other nominals. Section 8.2.2 builds on the previous section where I demonstrate that Clitic Doubling (CD) also matches the requirement on nominals established by topicality.

I propose that topicality and CD are interconnected. CD involves anaphoricity which is indicated by the presence of the pronoun. I show that in general nominals that are able to be linked to a pronoun can be differentially marked. The connection between topicality and anaphoricity plays out in the ability of the pronoun to refer to aboutness topics i.e., nominals mappable onto a referential address. In section 8.3, I examine whether the anaphoric relation established between the differentially marked object and the clitic (shown in section 8.2.2) can be extended to other environments with CD, such as in CLLD. The data reinforces the empirical generalization that pronouns in Arabic are restricted in their anaphoric capacity in that they can only refer to highly atomic nominals. Section 8.4 concludes.

8.1.1 Ruling out dative marking

Recall from Chapter 2 that differentially marked objects are marked with dative. In order to determine whether dative marking is responsible for the individuation restriction, I check whether non-individuated nouns i.e., substance-mass, can be dative-marked. The following examples in (147) exemplify that the dative marking cannot be the reason for the individuated noun restriction because substance mass nouns (which are incompatible with DOM as shown in section 7.2.2) can also appear with dative marking. The dative preposition used below is *la-* and it has the meaning of ‘for; for the purpose of.’

(147) *Substance mass can be dative marked*

a. jeb-et kyaas la-s-smide
brought.1SG bags.BP DAT-the-bulgur.SM
'I brought bags for the bulgur.'

b. aʕtin-i kyaas kbir-e la-l-baḥes
give.2SG.IMP-1SG.POSS bags.BP big-F.SG DAT-the-gravel.SM
'Give me large bags for the gravel!'

Thus, we conclude that there is no straightforward morphosyntactic restriction on non-individuated nouns.

8.1.2 Ruling out dislocation

Recall that in section 3.2 in chapter 3 I argued that differentially marked objects are derived by right dislocation. I show below in (148) that dislocation does not explain the individuation nominal restriction with DOM since nominals denoting non-atomic reference can be dislocated to the left.[124] Note that these nominals are indefinite and they are not anaphorically linked to a pronoun (see data on CLLD in (158)-(160)).

(148) *Substance mass can be dislocated*

a. smide, aštri-na l-yom
bulgur.SM bought-1PL the-day
'Bulgar, we bought today.'

b. baḥes, aštri-na mn ʕend yosef
gravel.SM bought-1PL from LOC Yousef
'Gravel , we bought from Yousef.'

Substance mass nouns can also be left dislocated in other languages like Spanish and Hebrew respectively.

[124] The indefinite left dislocated DPs in (148) are characterized as fronted focus and they are interpreted as contrastive focus (Aoun, Benmamoun, and Choueiri 2010).

(149) *Spanish*

Agua, toma por la mañana
Water, s/he-drinks in the morning
'Water, s/he drinks in the morning.' (Smeets 2019 :62(3b))

(150) *Hebrew*

mayim, šatiti ha-yom
Water, drank.1SG the-day
'Water, I drank today.' (Michal Domer, p.c.)

Thus, we can conclude from the data that the DOM individuation restriction in NGA cannot be explained through dative marking nor dislocation properties.

8.2 Proposal

I argue that aboutness topicality explains the nominal distribution presented with DOM. Aboutness topicality accounts for why nominals that are high on the atomicity scale (presented in section 7.4) are most likely to be differentially marked.

8.2.1 Aboutness topicality

In this section I argue that aboutness topicality explains the distribution of nominals with DOM and I show how the categories tested in the DOM experimental study (shown in section 7.3) play out with respect to topicality.

The analysis builds on the argument in Chapter 5 where I argued that differentially marked objects in NGA are interpreted as aboutness topics. A topic is understood to be the address for the context update, which points to a place where the information conveyed by the sentence will be stored during the context update (Reinhart 1981).[125] I follow Endriss'

[125] Reinhart's (1981) idea of topicality, implemented in Endriss (2009), is that if the referential address for topic-marked constituent α_T is already accessible in the common ground *c*, then an update with the conventional meaning α_C (α_T) is carried out.

(2009) proposal that only certain semantic types are compatible with the definition of aboutness topics. In particular, once we delve deeper into the semantics of nominals participating in DOM, we will see that mainly nominals denoting individuals (type *e*) and sets of individuals can be aboutness topics because their referential addresses already exist in the common ground (Reinhart 1981; Krifka 2008; Endriss 2009).

I examine the categories of count, object mass, substance mass, and collective nouns. I argue that nominals located high on the atomicity scale (shown in section 7.4) such as count or object mass nouns which denote individuals or sets of individuals are able to function as aboutness topics, thus they are highly marked in Arabic. In contrast to count and object mass nouns, the denotation of collectives and substance mass nouns makes them less likely to be anchored onto a referential address and ultimately behaves as an aboutness topic. Substance mass and collectives employ certain type-shifting operations in order to derive their meanings (e.g., Chierchia 1998; Dayal 2004, 2011; Rothstein 2010; Despić 2019; Erbach et al. 2019).[126] I assume that these operations make anchoring less likely to occur and therefore substance and collectives are less likely to be differentially marked.

The proposal on aboutness topicality accounts nicely for another class of objects that participates in DOM: quantifiers. As we have seen in Chapter 5, certain quantifiers are found to be licit with DOM in NGA. The same set of quantifiers can be aboutness topics (Endriss 2009). Although aboutness topicality straightforwardly accounts for the distribution of quantifiers with DOM, the atomicity distinction fails to do so since quantifiers are non-atomic. This issue is discussed in Chapter 9.

I begin with count nouns whose denotation makes them fall under the classification of aboutness topics. A denotation of a count noun is generated from atoms, or a set of atoms when the count noun is a plural (e.g., Link 1983; Chierchia 2010; Deal 2017). For something to be an aboutness topic, it needs to be associated with a referential address

---

[126] Since all differentially marked objects in NGA are definite and since Arabic has the *al-* which is the lexical version of ι, I assume it will always block ι. This has been formulated as the Blocking Principle, defined below. The application of this principle is not trivial for Arabic; that is, whenever the lexical item (in our case, for example, the definite article) is present, then that item must be used instead of the covert version.

(i) Blocking Principle (Type Shifting as Last Resort)
For any type-shifting operation ϕ and any X: $^{*}\phi(X)$ if there is a determiner D such that for any set X in its domain, $D(X) = \phi(X)$. (Dayal 2004: 216)

which in and of itself is associated with the common ground (Reinhart 1981; Molnár 1993; Krifka 2008; Endriss 2009). I follow Endriss (2009) who claims that individuals and sets can be topics because referential addresses can already exist in the common ground only for these types. I have argued that differentially marked objects are aboutness topics, as such, we expect that in order for a noun to be differentially marked, it must denote an individual, a set, or both. Given that count nouns denote individuals and sets, they are found to be highly marked.

Similar to count nouns, object mass nouns (e.g., *furniture*) denote sets of atoms i.e., individuals (e.g., McCawley 1975; Gillon 1992; Chierchia 1998; Barner and Snedeker 2005; Bale and Barner 2009; Rothstein 2010). Under the analysis of Rothstein (2017), object mass would be type ⟨e,t⟩. I also assume that these nouns denote individuals. Thus, given their denotations (sets and individuals) we expect object mass nouns to be able to establish a referential address, and in turn, are also highly marked.

On the other hand, substance mass nouns denote sets of quantities of non-individuated or atomic entities/objects. Scholars argue that substance mass nouns denote a kind reference (Krifka 2003; Despić 2019; Rothstein 2021; Köylü 2023, among others). According to Krifka (2003) and Dayal (2004), kinds are represented as individual concepts of type ⟨s,e⟩ (i.e., functions that at any world yield the totality of the manifestations of that kind in that world). Kinds are also related to properties by a nominalization operation ('down') $^{\cap}$. The $^{\cap}$ operation is "a function from properties to functions from situations to the maximal entity that satisfies that property in that situation. The function is partial in that it requires the kind term to pick out distinct maximal individuals across situations, thereby capturing the inherently intensional nature of the term" (Despić 2019: 273 based on Dayal 2004). I follow Despić (2019) who claims that $^{\cap}$ cannot be linked to a contextually anchored element. Therefore, I conclude that substance mass nouns are less likely to create a referential address, and consequently, are very unlikely to be differentially marked.

Collectives in Arabic are similar to substance mass in that they denote kinds and sets. The basic interpretation of collectives in Arabic is kind-denoting (see Ojeda 1992; Zabbal 2002; Mathieu 2014). I assume that the kind-denoting meaning does not require any type-shifting operation to derive the kind interpretation. Unlike substance mass, collectives can also denote individuals. In order to derive the individuation (atomicity) interpretation of

the collective, which I assume is not the basic or typical interpretation, a type-shifting operator i.e., the predicativizer or 'up' is required to take the basis as kinds and return individuals (see Chierchia 1998: 364; Dayal 2004: 399). Given the additional operation to derive the individuation reading of collectives, I assume collectives cannot be fully anchored onto a referential address. Thus, they are less likely to be differentially marked which is supported by the results of the experimental study (see section 7.3.5).

I assume that both substance mass and collectives are generally less likely to be anchored onto a referential address. For collectives, the situation is more complicated since a referential address can still be created in the special case where the collective denotes an atomicity interpretation.

So far, we have shown that topical elements which establish a referential address, and thus can be conceived of as individual-denoting expressions of semantic type *e* or sets of entities, are able to be differentially marked. The topicality component also explains why certain quantifiers are compatible with DOM (see data on the distribution of quantifiers presented in section 5.2.1.2 in chapter 5). Quantifiers are of type ⟨⟨e,t⟩,t⟩ which cannot be picked up by anaphoric expressions, hence no referential address is available for the context update. The question is how quantifiers can behave as an aboutness topic of the sentence. Endriss (2009) proposes that certain quantifiers can still function as aboutness topics via a mechanism that selects a suitable representative of the entire quantifier. The suitable representative stands proxy for the quantifier itself and at the same time delivers an address for the information structuring of the context update. This comes down to the creation of a suitable discourse referent.[127] Thus, certain quantifiers which are able to provide suitable addresses can serve as sentence topics. Crucially, as demonstrated in section 5.2.1.2, the class of quantifiers which can be differentially marked is the same as the class of quantifiers which are argued by Endriss (2009) to be good candidates for topics.

In summary, differentially marked objects are aboutness topics (mapped onto a referential address) in NGA and thus must be individual-denoting or denoting a set of individuals (Endriss 2009). Count and object mass nouns, classified high on the atomicity scale as discussed in section 7.4, are found to be highly marked as they are conceived as individual-denoting expressions or sets of individuals, while collectives and substance

[127] For the exact semantic mechanism of Endriss' (2009) proposal, the reader is referred to appendix b.

mass, which are classified low on the atomicity scale, require further type-shifting operations in order to derive their meanings, making them less likely to be marked. Table 8.1 below summarizes the distribution of nominals with DOM, as well as their interpretation, availability to be anchored onto a referential address, and their rank of acceptability with DOM.[128]

Noun type	**Interpretation of the noun**	**Is there mapping onto a referential address?**	**Acceptability with DOM**
COUNT	Individual, set of individuals	Yes	High (M = 4.93)
OBJECT MASS	Individual, set of individuals	Yes	High (M = 4.56)
QUANTIFIERS	Sets of sets ⟨⟨e,t⟩,t⟩	Yes (address will be created)	Acceptable
COLLECTIVES[129]	Individual, set of individuals, kind	Yes: when the collective denotes (set of) individuals. No: when the collective denotes kind reference.	Low (M = 3.05; 3.12)
SUBSTANCE MASS	Set, kind	No. Kind-denoting reference can't be anchored to a referential address (Despić 2019)	Low (M = 2.56)

Table 8.1: Aboutness topics across nominals and DOM results

[128] We did not test quantifiers in the experimental study. A follow-up study will test the acceptability judgment of types of quantifiers.

[129] For the purpose of this table which relates DOM to the ability of the noun to be an aboutness topic, collectives are treated as a whole; no distinction has been made here with respect to the associated clitic.

Having demonstrated how aboutness topicality (anchoring onto a referential address) explains the distribution of nominals with DOM, I now show that the same set of nominals mapped onto a referential address also appears in a clitic doubling configuration.

8.2.2 Clitic Doubling

In this section I demonstrate that CD also matches the requirement on nominals established by topicality. Recall that DOM in NGA obligatorily involves CD, namely, the double DP and the clitic must hold an anaphoric relation. I show that the referential relationship with the clitic is restricted to nominals denoting individuals or sets of individuals. The same set of nominals are also mapped onto a referential address i.e., aboutness topics (shown in section 8.2.1). I further demonstrate that the anaphoric relation between the clitic and the differentially marked object extends to encompass other pronominal relations.

Below, I present data showing that nouns denoting individuals or sets of individuals such as count and object mass nouns can stand in an anaphoric relation with the clitic in non-DOM contexts. Examples (151) and (152) illustrate for count nouns and (153) for object mass.

(151) *Count nominals hold an anaphoric relation with the clitic*

a. raḥ aštre haḏok ṭ-ṭawl-āt bas mlaqit-**on** šwai ɣaly-āt
will buy.1SG those the-tables-SP_F but find.1SG-3PL.OBJ little expensive-SP_F
'I will buy those tables, but I found them a little bit expensive.'

b. dahan-na kul ṭ-ṭawl-āt o ḥaṭṭi-na-**hon**[130] barra
painted-1PL all the-tables-SP_F and put-1PL-3PL.OBJ outside
'We painted all the tables and put them outside.'

Similarly, the anaphoric relation between the clitic and other plural nouns such as broken plurals can also be observed below.

[130] In NGA, the plural pronoun (referring to object) can be pronounced as *hon* or *on* which depends on its phonological environment.

(152) *Broken plurals hold an anaphoric relation with the clitic*

a. al-**karaasi** ʔelli fi-l-maxzan badna nestʕml-**on** la-l-ḥafle
the-chairs.BP that in-the-storage want.1PL use.1PL-3PL.OBJ for-the party
'The chairs that are in the storage we want to use them in the party.'

b. fii ktiir **šababik** fi-l-bit, dayman banaḏef-**on** laḥali
exist many windows.BP in-the-house always clean.1SG-3PL.OBJ myself
'There are many windows at home, I always clean them alone.'

So far, we have shown that nouns denoting individuals or sets as count nominals can appear in an anaphoric relation with the clitic in non-DOM contexts. As shown in section 7.2.2, these same types of nominals can also occur in DOM contexts.

Like other nouns denoting individuals or sets we have seen so far, object-mass nouns such as *bariid* 'mail' and *ɣasil* 'laundry' can be anaphorically linked to a pronominal element. Object mass nouns are morphologically singular in Arabic, thus, the only possible form of a clitic is singular.

(153) *Object mass nominals hold an anaphoric relation with the clitic*

a. laqit al-**bariid** o aʕṭit-**o** la-aḥmad
found.1SG the-mail.OM and gave.1SG-3M.SG.OBJ to-Ahmad
'I found the mail and give it to Ahmad.'

b. sara al-yom našr-at al-**ɣasil** o ṭaww-at-**o**
Sara the-day hung-3F.SG the-laundry.OM and folded-3F.SG-3M.SG.OBJ
be-sorʕa
in-quickness
'Today, Sara hung out the laundry and folded it quickly.'

We have seen in section 7.3.5 that collectives with DOM received low acceptability ratings regardless of whether the associated clitic is singular or plural. Recall that while the singular clitic is associated with a group reading of the collective, the plural clitic is

associated with the individuation reading. In an anaphoric context (154), the judgments containing collectives *toffaḥ* 'apples' and *ward* 'plants' are not very clear whether collectives can participate in an anaphoric relation irrespective if the pronoun is singular or plural. Based on the data below, I conclude that there is no difference between the way the singular clitic and plural clitics behave with collectives in any anaphoric context.

(154) *Collectives cannot hold an anaphoric relation with the clitic*

a. ?? šarit kthir **toffaḥ** o ḥaṭṭ-et-**o**/ ḥaṭṭet-**hon** fi-l-kyaas
bought.1SG a lot apple.COLL and put.1SG-3M.SG/3PL.OBJ in-the-bags.BP
'I bought a lot of apples and I put it/them in bags.'[131]

b. ?? jebit al-**ward** men d-dokkane bs tarakt-**o**/tarakt-**on**
brough.1SG the-flowers.COLL from the-store but left.1SG-3M.SG/PL.OBJ
barra
outside
'I brought the flowers/plants from the store and left it/them outside.'

The facts with collectives are in accordance with the findings of the DOM study indicating that collectives received relatively low acceptability ratings with DOM.

On the end of the atomicity hierarchy, it has been shown that substance mass nouns are incompatible with DOM. This fact is extended to other pronominal relations as demonstrated in the examples in (155): substance mass nouns such as *dam* 'blood or *bḥesʕ* 'gravel' cannot hold an anaphoric relationship with the clitic.

(155) *Substance mass nominals cannot hold an anaphoric relation with the clitic*

a. * lyom roḥ-et tbarraʕ-et bi-d-**dam** o laqit-**o**
today went-1SG donated-1SG in-the-blood.SM and found. 1SG-3M.SG.OBJ
ktiir aḥmar

[131] This example is judged to be grammatical with both pronouns (Hnout, Laks and Rothstein 2021: 162, (32a)), however, my native speaker consultants did not like it.

very red

(‘I went today to donate blood and I found it very red.’)

b. * z-zalame axaḏ al-**baḥesˁ** o jab-**o** ʕal-bit

the-man took.3M.SG the-gravel.SM and brought.3M.SG-3M.SG.OBJ to-home

(‘The man took the gravel sand and brought it to the home.’)

Thus, the data points to the conclusion that nouns that are high on the atomicity scale, denoting individuals or sets of individuals such as count and object mass nouns, can stand in an anaphoric relation with the clitic in non-DOM contexts. However, substance mass, located very low in the atomicity scale, cannot participate in an anaphoric relation. Based on the data, it is unclear whether collectives can participate in an anaphoric relation irrespective of whether the pronoun is singular or plural. As summarized in Table 8.2, the anaphoric relationship established between the pronominal clitic and the corresponding nominal in non-DOM context is identical to DOM-context. This generalization, as will be explained in the following section, is attributed to the anaphoric capacity of the pronominal clitic in NGA.[132]

Type of nominal	Availability with DOM	Availability with anaphoric reading (non-DOM contexts)
Sound plural (*tables*)	✓	✓
Broken plural (*chairs*)	✓	✓
Object mass (*laundry*)	✓	✓
Collectives (*apples*)	??	??
Substance mass (*blood*)	✗	✗

Table 8.2: Anaphoric relation with types of nouns

[132] Preliminary data show that Hebrew has a similar distinction in pronominal reference. Future research is needed to explore the data more closely.

I have demonstrated so far that nominals linked to a pronominal element can be differentially marked. As shown in section 8.2.1, nominals which fall under the classification of aboutness topics, i.e., those denoting individuals and sets of individuals, can be differentially marked. In the following section I will argue that topicality is semantically connected to anaphoricity. I begin with a discussion of the definition of anaphoricity below.

8.2.2.1 Anaphoricity

In section 8.2.1 I argued how aboutness topicality explains the distribution of nominals with DOM. Then in section 8.2.2, I demonstrated that CD also matches the requirement on nominals established by topicality. This section now examines the connection between topicality and anaphoricity.

The data presented in section 8.2.2 lead to the empirical generalization that pronouns in Arabic are restricted in their anaphoric capacity, specifically, they cannot refer to substance mass nouns. As shown in the previous section, the same set of nouns that cannot be anchored to a referential address (i.e., an aboutness topic), also cannot be linked to a pronoun.

I follow Heim and Kratzer's (1998) hypothesis that all pronouns are variables. A pronoun is a numerically indexed expression whose semantic value is a function of a variable assignment g (where g is a function from numbers to individuals in the domain of the relevant model). The variable assignment is generally assumed to be a function of the discourse context and can informally be thought of as a model of the speakers' intentions. This is stated in the Pronoun rule in (156).

(156) The pronoun rule [Heim & Kratzer (1998: 111)]

If X is a pronoun, and *g* is a variable assignment, and n is an index in the domain of *g*, then $[[X_n]]^g = g(n)$

The assignment function must be set up with respect to a referential address; pronouns should be able to refer only to elements that can be anchored. Anchoring something to a

referential address allows the corresponding nominal to feed an assignment function and the corresponding nominal needs to have a value in the assignment function in order to identify the exact denotation of the pronoun. The behavior of aboutness topicality and the pronouns accounts for the nominal distribution of DOM as follows:

(157) a. Differentially marked objects in NGA are interpreted as aboutness topics (Chapter 5).
b. By definition, topics are mapped onto referential address.
c. A referential address corresponds to a file card and each file card is associated with an individual (Reinhart 1981).
d. Pronouns are limited in their anaphoric capacity: they can only hold an anaphoric relation with nominals denoting individuals.

If the corresponding nominals are count and object mass, the referential address is created and accordingly the assignment function of the pronoun is applied. The logic is true for nouns that are less likely to be mapped onto a referential address and cannot be linked to a pronoun such as substance mass.[133]

Thus, the generalization that DOM is attested with highly atomic nominals is explained through aboutness topicality, i.e., mapping onto a referential address is successful with highly atomic nominals. There is a clear connection between topicality and anaphoricity. The step in (157d) assumes that the referential address that represents the individual (numeric index) is picked out by the pronoun. This connection accounts for why the same nominals that can behave as aboutness topics (be anchored onto a referential address) can also be linked to a pronoun, and why the nouns that cannot behave as aboutness topics, cannot be linked to a pronoun.

Syntactically, topicality and anaphoricity are interconnected. The big-DP analysis is a necessary pre-condition for objects participating in DOM. Given that the big-DP

[133] Based on the proposed analysis, it is unclear why DOM is illicit with pronouns (shown in section 2.1). My very preliminary semantic explanation for this restriction involves a combination of topicality and clitic doubling. When the clause has two pronouns (one is differentially marked), they must be co-indexed with an individual being interpreted as a topic that appears in the previous text. The assignment function of the pronoun is applied twice and this is unnecessary, so that's why DOM is not possible with pronouns.

combines the DP and corresponding clitic which together form a single constituent, it is predicted that only nominals that are mapped onto a referential address (fall under the definition of aboutness topicality of Reinhart, 1981 and Endriss, 2009) can be linked to a pronoun. This prediction is borne out in the NGA DOM data. As explained in this chapter, nominals that map onto a referential address such as count and object mass nouns are able to hold an anaphoric relation with pronominal elements, and consequently, are differentially marked. However, nominals that are less likely to be mapped onto a referential address, cannot be linked to a pronominal element, and thus, are less likely to be differentially marked.

Given that only aboutness topic elements such as count and object mass nouns can be linked to a pronoun, and thus are able to be marked, we expect them to hold an anaphoric relation in other environments of clitic doubling. One such environment, clitic left dislocation, will be examined in the next section.

8.3 CLLD and types of nominals

In this section I examine whether the anaphoric relation established between the differentially marked object the clitic (shown in section 8.2.2) can be extended to other environments with CD (see appendix A.2.1 for CD data in anti-pronominal contexts). The data shown below reinforces the empirical generalization made in section 8.2.2.1 that pronouns in Arabic are restricted in their anaphoric capacity, specifically, they can only refer to highly atomic nominals.

An environment that obligatorily contains CD is CLLD, which, as defined in section 4.2, contains a left dislocated nominal that is obligatorily associated with a clitic. I demonstrate below that in CLLD we find the same profile of data as in DOM with respect to the distribution of nominals.[134] Similar to the nominal distribution shown with DOM, while count and object mass nouns can be CLLDed, substance mass nouns cannot. The following examples in (158a) and (159a) illustrate that count and object mass nouns can be

[134] The observation that the same class of nouns participates in DOM (instance of clitic doubling) also appears in CLLD does not align with Anagnostopoulou's (2006) conclusion for Greek that there are semantic classes of DPs that resist clitic doubling, but at the same time can occur in CLLD.

CLLDed, and (158b) and (159b) illustrate these same nominals can also be differentially object marked.

(158) a. al-bloza-āt, sara baʕ-at-on CLLD
the-shirt-SP$_F$ Sara sold-3F.SG-3PL.OBJ
'The shirts, Sara sold them.'

b. sara baʕ-at-on la-l-bloza-āt DOM
Sara sold-3F.SG-3PL.OBJ DOM-the-shirt-SP$_F$
'Sara sold the shirts.'

(159) a. al-ɣasil, sara našr-at-o CLLD
the-laundry.OM, Sara hung-3F.S-3M.SG.OBJ
'The laundry, Sara hung out it.'

b. sara našr-at-o, la-l-ɣasil DOM
Sara hung-3F.SG-3M.SG.OBJ DOM-the-laundry.OM
'Sara hung the laundry.'

In contrast to the count and object mass nouns, substance mass nouns are illicit with CLLD as shown in (160a). The same fact applies to DOM.

(160) a. * s-smide, sara šar-at-a CLLD
the-bulgur.SM, Sara bought-3F.SG-3F.SG.OBJ
('The bulgur, Sara bought it.')

b. * sara šar-at-a la-s-smide DOM
Sara bought-3F.SG-3F.SG.OBJ DOM-the-bulgur.SM
('Sara bought the bulgur.')

So far, we have shown that while count and object mass nouns are compatible with CLLD, substance mass nouns are not.

The last nominal discussed with DOM is collectives. As shown in section 7.3.5, the findings of the DOM study indicate that collectives received relatively low acceptability ratings with DOM. In CLLD, speakers display a high degree of variability in judging

collectives (161a). Given the inter-speaker variability of judging collectives with CLLD, I leave this issue for future research. Similar judgments are given when the collective is associated with a plural clitic.

(161) a. ?? al-baqar, sara ṭaʕm-at-o CLLD
the-cows.COLL, Sara fed-3F.SG-3M.SG.OBJ
('The group/kind of cows, Sara fed it.')

b. ?? sara ṭaʕm-at-o la-l-baqar DOM
Sara fed-3F.SG-3M.SG.OBJ DOM-the-cows.COLL
('Sara fed the group/kind of cows.')

As it stands, the emerging picture points out that differentially marked objects, which are argued to be compatible with the definition of an aboutness topic such as count nouns and object mass nouns, are able to be used anaphorically in other environments involving clitic doubling. Thus, the observation of which types of nominals can partake with DOM can be extended to CLLD as well (see sections 4.2 and 4.3 on the distinction between CLLD and DOM).

8.4 Conclusion

In this chapter I argued that nominals that are high on the atomicity scale are conceived as individual-denoting expressions or sets of individuals and thus are able to be anchored onto a referential address and consequently are found to be highly marked. On the other hand, nominals low on the atomicity scale, are less likely to be anchored onto a referential address (being an aboutness topic), making them less likely to be marked. I have also established that the same set of nouns that can be anchored onto a referential address and participate in DOM can also be linked to pronouns. The set of nouns that cannot be anchored to a referential address, such as substance mass nouns, cannot be linked to pronouns. I have also shown that this anaphoric restriction is attributed to the fact that pronouns in NGA are restricted in their anaphoric capacity in that they can only refer to highly atomic nominals.

Future work will continue to explore this generalization in other anaphoric contexts in more depth.

The current analysis neatly captures the nominal properties of atomicity and anaphoricity associated with DOM. The big-DP analysis is a necessary pre-condition for objects participating in DOM. Given that the big-DP combines the DP and its corresponding clitic which together form a single constituent, it is predicted that nominals that are mapped onto a referential address (i.e., those which fall under the definition of aboutness topicality of Reinhart, 1981 and Endriss, 2009) can be linked to a pronoun.

I have shown that aboutness topicality also has consequences for how the overall discourse is modulated. As assumed in Chapter 6, aboutness topic marking triggers the anchoring of the asserted proposition to a referential address. The illocutionary effects of DOM derived in Chapter 6 arise from the speaker using the aboutness topicality to grammatically mark the union of the Discourse Commitment of the speaker and the Discourse Commitment of the other interlocutors. The applicative analysis proposed in this thesis ties nicely to the illocutionary properties of DOM. The use of a larger structure of DOM (clitic doubling and applicative projection) when there is a simpler non-DOM structure available yields an additional pragmatic interpretation.

Chapter 9

Conclusion

This thesis has offered a comprehensive analysis of the Differential Object Marking phenomenon in Arabic. While DOM is well studied in languages like Turkish, it has been less studied in Arabic dialects, and Northern Galilee Arabic (NGA) in particular.

The thesis began in chapter 2 by presenting the DOM patterns in NGA and other closely related Arabic dialects, and then discussing the Arabic facts in relation to the existing DOM literature in general. The NGA facts demonstrated that differentially marked objects are marked with dative and the DOM construction obligatorily involves clitic doubling. Further, only definite DPs and proper names can be DOM in NGA (Abu-Haidar 1979) but not pronouns, contrary to what the definiteness scale predicts (e.g. Silverstein 1976; Aissen 2003). Crucially, not all definite DPs are DOM, that is, definiteness cannot be the defining semantic property of DOM; and neither can specificity. Indefinite nominals cannot be DOM even if they are specific (Brustad 2000). DOM in NGA also does not

depend on animacy, which was previously proposed as a defining factor, for example, in Silverstein (1976), Aissen (2003), and Krause & von Heusinger (2019).

In Chapters 3 and 4 I investigated the syntax of DOM where I established an empirical generalization that DOM is an instantiation of dislocation and provided a syntactic analysis for this structure. I argued that DOM is derived by rightward A-movement and that this movement targets an ApplP at the edge of *v*P. The movement is accompanied by clitic doubling, where the DP object and the corresponding clitic start the derivation by forming a big-DP.

Arguing for the applicative structure was motivated by (i) case checking, (ii) locality conditions of DOM, and (iii) aboutness topicality effects (argued in detail in chapter 5). Another motivation for the applicative analysis comes from a comparison of DOM to Clitic Left Dislocation (CLLD), another clitic construction in NGA. This comparison deepened our understanding of the syntax of DOM. I argued that unlike CLLD which undergoes movement to the left periphery, DOM undergoes rightward movement to ApplP. I argued that differentially marked objects are unable to move to a higher position because the ApplP, being at the edge of *v*P, is a criterial position (in the sense of Rizzi 2006) where once differentially marked objects move to this position, they are rendered inaccessible to further syntactic movement. Further, I demonstrated that DOM is clause-bounded as it obeys the Right Roof Constraint (RRC). The freezing effect accounts for the RRC: Criterial Freezing restricts DOM movement to a higher position (i.e., higher than ApplP) and in this way, DOM respects the RRC locality constraint.

Having investigated the syntax of DOM, I examined the question in Chapter 5 of whether an information-structural property associated with DOM motivates the movement to ApplP. I argued that in NGA what distinguishes DOM structures from their non-DOM counterparts is aboutness topicality (Reinhart 1981). I further observed that only nominals that can be mapped onto a referential address (in the sense of Endriss, 2009) can be differentially marked. Movement of the differentially marked object to the edge of *v*P allows it to anchor to a referential address, making it an aboutness topic.

However, the syntactic analysis does not fully account for the pragmatic properties of DOM, which raises the question of *how* they arise. In Chapter 6, I showed that DOM, as opposed to its non-DOM counterpart, may yield interpretations of the speaker's emotive

content, a correction, an accommodation, or activate a parallel Question Under Discussion. Adopting Kučerová and Zarka (in prep), I argued that DOM in NGA behaves as an illocutionary marker and that it grammatically marks the asserted proposition as a non-default Discourse Commitment (Gunlogson 2001; Farkas & Bruce 2010; Rett 2021). This follows Endriss (2009) in that the aboutness topic marking triggers the anchoring of the asserted proposition to a referential address. I argued that this anchoring allows the proposition to be added to the Table (Farkas & Bruce 2010) when the asserted proposition is not a member of the projected set. Aboutness topicality is a necessary precondition such that a proposition can only be added to the Table if it can be anchored. In DOM propositions, the marked DP aboutness topic anchors the proposition via its referential address. I concluded this chapter by proposing that the obligatory illocutionary effect of DOM is a result of structural economy. Since DOM involves an additional structure combining clitic doubling and an applicative projection, which is absent with non-DOM objects, this additional structure triggers interpretive effects that would not be available otherwise (e.g., Sichel and Wiltschko 2021).

In Chapter 7, I examined *which* types of nominals may be differentially marked. I initially observed that only individuated nouns, that is, countable nouns, can be differentially marked. After reviewing the various views of individuation affecting DOM, I argued that a finer distinction of individuation is necessary to capture the DOM facts in NGA. Based on the results of Zarka and Hacohen's (2023) experimental study, I concluded that nominals high on the atomicity scale are more likely to be differentially marked. I, therefore, proposed that individuation, as a gradable treatment of atomicity (in the sense of Grimm 2012) rather than countability, is the right characterization for the distribution of nominals with DOM in NGA.

In Chapter 8, I investigated the question of *how* to account for the fact that nominals that are high on the atomicity scale are more likely to be differentially marked. I developed a semantic analysis in which aboutness topicality explains the distribution of nominals with DOM. I further demonstrated that topicality is tied to anaphoricity. I demonstrated that the same set of nominals that can be mapped onto a referential address can also hold an anaphoric relation with a clitic. The connection between topicality and anaphoricity is established through the behavior of pronouns in general. Pronouns are analyzed as variables

that are mapped onto a referential address via a variable assignment function (Heim and Kratzer 1998). Thus, pronouns are able to refer to aboutness topics being anchored onto referential address. Therefore, in NGA DOM, clitic doubling arises through the presence of the pronominal element whose anaphoric capacity is limited in NGA: it can only refer to highly atomic nominals, and its general behavior is to be mapped to a referential address.

To my knowledge, this thesis is the first work to thoroughly investigate DOM in Arabic dialects and, in particular, NGA. It provides a detailed investigation of DOM in NGA which is complemented by novel arguments about the syntax, semantics and pragmatics of DOM. Previous descriptive works have focused on each level separately without providing an analysis of the entire phenomenon. Novel to the current thesis is that the defining property for DOM in Arabic dialects is not definiteness as was previously claimed (Aoun 1999; Brustad 2000). Rather, DOM is sensitive to atomicity, and atomicity in NGA is tightly correlated with anaphoricity. Another novel contribution of the thesis is demonstrating that DOM also operates at the context level where the use of DOM, as opposed to non-DOM counterparts, yields certain pragmatic effects.

The current analysis neatly captures the nominal properties of atomicity and anaphoricity associated with DOM. The two properties are derived by the big-DP. Since the big-DP combines the DP and corresponding clitic which together form a single constituent, it is predicted that nominals which are mapped onto a referential address i.e., aboutness topics (Reinhart, 1981 and Endriss, 2009) can be linked to a pronoun. Importantly, aboutness topicality impacts how the discourse is modulated. I argue that the illocutionary properties of DOM stem from the speaker using aboutness topicality to grammatically mark that the union of the Discourse Commitment of the speaker and the Discourse Commitment of the other interlocutor(s) is incoherent.

I end the dissertation by discussing some of the open questions which serve as avenues for future research.

9.1 Open questions

I have argued in Chapter 7 that atomicity is the right characterization for describing which types of nominals may be differentially marked. The atomicity distinction fails to account

for the distribution of quantifiers with DOM. As shown in Chapter 3, certain quantifiers can be differentially marked yet quantifiers themselves cannot denote atoms. Consequently, atomicity cannot account for all the types of nominals partaking in DOM. On the other hand, we have seen that aboutness topicality straightforwardly accounts for all the types of nominals participating in DOM without exception. This raises the question of whether topicality could resolve some of the restrictions that atomicity cannot. The puzzle that quantifiers can appear with DOM and yet are themselves non-atomic is reminiscent of Endriss' problem with quantifiers being topicable, despite not being individuals. Endriss solves this puzzle by providing a formal semantic analysis in which certain quantifiers can still function as topics if they are mapped onto a minimal witness set.

Another open question for future research concerning DOM in NGA is how DOM in NGA differs from ethical datives. I have argued that DOM targets a position higher than *v*P, and it has been argued that ethical datives attach to a position higher than *v*P or even in the TP area. Future research is needed to shed light on the pragmatic and semantic properties of ethical dative and how they compare to DOM.

The proposal in this dissertation has been motivated primarily based on Arabic. It is an open question whether it extends to other languages with DOM. Similar to NGA, DOM in languages like Romanian (e.g., Hill & Tasmowski 2008) and Catalan (e.g., Escandell-Vidal 2009) has been associated with discourse pragmatic effects, but further research is required for these and other languages. The proposed analysis of DOM raises the question whether a uniform analysis of DOM is possible by unifying the distinct DOM systems as instantiations of economy-driven interpretive effects.

Appendix A

Clitic doubling vs. object agreement

In the thesis I adopt the big-DP analysis for DOM. Here I provide two pieces of evidence that the agreeing morpheme behaves as a clitic.

The literature has identified two operations, object agreement and clitic doubling, that give rise to the relation between the φ-bearing morpheme and a corresponding full nominal phrase as in (162).

(162) host+[morpheme]$_{\varphi i}$ . . . (other material) . . . [full noun phrase]$_{\varphi i}$

There has been a substantial amount of work attempting to provide diagnostics to tease apart agreement and clitic doubling (e.g., Preminger 2009; Nevins 2011; Kramer 2014). Many scholars have focused on morphosyntactic diagnostics to identify whether a given φ-bearing morpheme results from true agreement or clitic doubling (Zwicky and Pullum 1983; Woolford 2008; Nevins 2011). Recent work by Yuan (2018, 2021) on Inuktitut (one of the Inuit varieties) focuses on a semantic diagnostic for φ-agreement vs. clitic doubling. Specifically, the interpretation of the doubled objects is correlated with a semantic property and hence they are obligatorily interpreted as *strongly D-linked*. D-linking is understood as

a semantic restriction on the domain of individuals to those that are contextually salient (e.g., Pesetsky 1987).

Based on these morphological and semantic diagnostics, I analyze the behavior of the φ-bearing morpheme in the DOM construction as a clitic but not as a result of object agreement.

A.1 Morphological properties

Zwicky and Pullum (1983) argue that φ-bearing morphemes are often subject to allomorphy and exhibit morphological irregularities, while clitics are expected to be regular; furthermore, while clitics are able to attach to stems that contain affixes, affixes cannot attach outside of clitics. However, these morphological distinctions are not universal and are not clearly driven by any theoretically-grounded differences between the two (Yuan 2018). Thus, it is unclear whether there is a reliable link between affixes and agreement, or between morphophonological clitics and pronominal clitics.

As a result, Nevins (2011) provides an alternative diagnostic based on contextual morphological variance. He proposes that since agreement is a realization of phi-features on a functional head, the realization of those phi-features may vary depending on other features that the functional head itself has, for instance, the tense feature on T. Nevins argues that since clitics are Ds, they are expected to not vary with respect to tense.

In NGA, the φ-bearing morpheme does not vary with respect to tense. The verb in Arabic has two tenses, past and present, or as they are referred to respectively, perfective and imperfective aspects. The agreeing morpheme does not vary whether the verb it is attached to is past or present as in (163).

(163) a. past (perfective aspect)

akl-at-**o**

eat-3F.SG.PST-3M.SG.OBJ

b. present (imperfective aspect)

b-t-akl-**o**

PRS-3F.SG-eat-3M.SG.OBJ

Arabic φ-bearing morpheme on the object behaves differently than subject agreement. Subject agreement varies depending on tense, so the realization of phi-features does vary depending on the tense feature that the verb has.

Kramer (2014) observes that the tense-variance is a reliable diagnostic in Amharic, in which the realization of subject agreement (φ-agreement, italicized) is tense-dependent, but doubled clitics (bolded) are invariant.

(164) *Amharic* (Kramer 2014 :607(28))

Perfect:	**Imperfect:**
a. säbbär-*ä*-**ññ**	*yi*-säbr-**äññ**
break.PERF-3MS.S-1S.O	3MS.S-break.IMPF-1S.O
b. säbbär-*ä*-**h**	*yi*-säbr-i**h**
break.PERF-3MS.S-2MS.O	3MS.S-break.IMPF-2MS.O
c. säbbär-*ä*-**w**	*yi*-säbr-**äw**
break.PERF-3MS.S-3MS.O	3MS.S-break.IMPF-3MS.O

In addition to tense-(in)variance, Nevins (2011) also proposes person complementarity effect as a diagnostic for clitic doubling. Person-complementarity is essentially the Person-Case Constraint (i.e. PCC), which bans φ-feature combinations of ditransitive internal arguments in which one of the arguments is 1st/2nd person (Perlmutter 1971; Bonet 1994; Béjar and Rezac 2003, among others). PCC effects have been documented for Classical Arabic (Fassi Fehri 1988). Similarly, NGA is subject to PCC within the object marking in the ditransitive constructions. When both direct and indirect objects are 3rd person, the combination is possible as in (165), however, the clitic combination is not allowed with 3rd person indirect object and 2nd person direct object as in (165b) but is possible in the reversed order as in (165c).[135]

[135] It is not possible to test PCC in DOM as pronouns are not allowed to be differentially marked (see section 2.1)

(165) a. 3-IO+ 3-DO:

IO	**DO**
nasˁḥ-at-**o**	ʔeyya-**ha**
recommended-3F.SG-3M.SG	ACC-3F.SG

'She recommended her to him.'

b. *3-IO+ 2-DO:

IO	**DO**
*nasˁḥ-at-**o**	ʔeyya-**ak**
recommendeded-3F.SG-3M.SG	ACC-2M.SG

('She recommended you to him.')

c. 2-IO+ 3-DO: (reverse order)

IO	**DO**
nasˁḥ-at-**ak**	ʔeyya-ah
recommended-3F.SG-2M.SG	ACC-3M.SG

'She recommended him to you.'

Thus, NGA exhibits PCC with object marking, which is entirely unexpected if the object morpheme were a genuine instantiation of φ-agree.

Moreover, Nevins proposes an additional morphosyntactic diagnostic to distinguish between pronominal clitics and genuine φ-agreement called the *omnivorous number* effect. Omnivorous agreement also appears in languages with object agreement (e.g., Preminger 2011, 2014), therefore it is not specific to clitic doubling. For this reason, I opted not to discuss this diagnostic in this thesis.

Based on the morphological properties discussed above, the agreement morpheme in the DOM constructions behaves as a clitic. I turn now to examine its behavior with respect to semantic diagnostics.

A.2 Semantic properties

Many languages forbid clitic doubling for non-D-linked wh-phrases and require clitic doubling for D-linked wh-phrases (e.g., Suñer 1988; Dobrovie-Sorin 1990; Franks and Rudin 2005; Kallulli 2008; Kramer 2014; Yuan 2018, 2021). The directionality of the pattern is similar cross-linguistically; there do not appear to be any languages in which non-D-linked objects undergo clitic doubling when D-linked objects do not. Yuan (2018) provides novel evidence for the D-linking effect in clitic doubling. The evidence comes from a discussion of "anti-pronominal" contexts (cf. Postal 1994) or what refers to these contexts as Π-positions (Poole 2017). These are syntactic environments where a DP denotes a property (type ⟨e, t⟩). Π-positions provide novel support for the idea that clitics trigger a D-linking effect on their DP associates.

A.2.1 D-linking effect

The interpretation of the clitic doubled objects may be viewed as a D(iscourse)-linking effect, which is independently observable on wh-phrases in the language. Crucially, whereas the φ-bearing morpheme is illicit with non-d-linked wh-phrases (166a), it must appear with d-linked wh-phrases which quantify over a given set (166b).

(166) a. **Non-D-linked wh-phrase**

min šaf-at-**(*o)** sara mbirḥ?
Who saw-3F.SG-M.SG.OBJ Sara yesterday
'Who did Sara see yesterday?'

b. **D-linked wh-phrase**

ʔay walad šaf-at-*(**o**) sara mbirḥ?
Which boy.M.SG saw-3F.SG-M.SG.OBJ Sara yesterday
'Which boy did Sara see yesterday?'

To test if DOM is indeed associated with the D-linking effect, I use Yuan's (2018) test building on Poole's (2017) work, for identifying whether clitic doubling arises in

configurations that are anti-pronominal. Π-positions are known to be characterized as disallowing pronouns and definite DPs. In fact, Poole (2017) observes that only a subclass of DPs i.e., anaphoric definite nominals are excluded. By way of illustration, consider the contrast between (167)-(168). The following examples in English are taken from Poole (2017).

(167) **Π-positions permits non-anaphoric NPs and DPs**

a. *Naming verbs*

(i) Irene called the cat [Snowflake]. NP

(ii) Irene called the cat [that dumb nickname]. DP

b. *Predicate nominals*

(i) Erika became [a teacher]. NP

(ii) Erika became [the CEO]. DP

(168) **Π-positions ban anaphoric definite DPs**

a. *Naming verbs*

My mother liked one of the names in the baby book.

(i) ✓My grandmother had wanted to give the name to my uncle.

(ii) #My grandmother had wanted to call my uncle [the name].

b. *Predicate nominals*

Anna decided on a type of doctor to become.

(i) ✓The type made a lot of money.

(ii) #And she became [the type].

Poole (2017) shows that the same contexts are necessarily antipronominal, Π-positions are antipronominal i.e., they reject pronominals like *it*. As it is pointed by Poole, it is crucial to draw attention to the fact that antipronominality does not extend to strong pronouns like *that*.

(169) **Π-positions are antipronominal**

a. *Naming verbs*

Irene liked the name Snowflake, and she called the cat { *it /✓that }.

b. *Predicate nominals*

Erika wanted to become a teacher, and she became { *it /✓that }.

Having shown Π-positions in English, I turn to discuss the Arabic data to provide support for treating the φ-bearing morpheme as a clitic because it triggers a D-linking effect on the DP associate. A prediction that arises from this approach is that, if NGA pronominal clitics are anaphoric definites, then clitic doubled DPs should display properties anaphoric definites as well. This prediction is borne out as shown below. Specifically, we will see that clitic-doubled DPs may not be licit in Π-positions.

First, I must establish that NGA has Π-positions to begin with. This is illustrated in (170).

(170) **Π-positions in NGA**

a. *Naming verbs*

sara sammat l-besse šokolaṭa/ haḏak al-ʔesem al-msˁaṭṭel
Sara called the-cat Chocolate/ that the-name the-dumbed
'Sara named the cat Chocolate/ that dumb name.'

b. *Predicate nominals*

sara sˁarat mʕalme/ al-maleke
Sara became teacher/ the-queen
'Sara became a teacher/the queen.'

Parallel to the English data, non-anaphoric definite nominals in NGA are also allowed in Π-position.

Context: [I am scanning the names in the baby book]

(171) badde ʔsamme ʔebni haḏa al-ʔesem
want.1SG name my son this the-name
'I want to name my son this name.'

Crucially, Π-positions may not be compatible with clitic-doubled DPs.

(172) a. sara sˁarat ʔl-maleke
Sara became the-queen
'Sara became the queen.' BASELINE

b. * sara sˁarat-a la-l-maleke
Sara became-3F.SG.OBJ DOM-the-queen
('Sara became the queen.') CLITIC DOUBLING

While the same DP 'the queen' is licit with Π-position (170b), it is illicit when it is clitic doubled (172b). This contrast supports the idea that φ-bearing morphemes in DOM are clitics in which they trigger D-linking effect on their DP associates.

Recall from section 3.5.2.1 that, in addition to being understood as D-linked, differentially marked objects (clitic doubled) in NGA also obligatorily trigger a wide-scope reading and are interpreted as aboutness topic. I propose that the D-linking effect and the availability of wide-scope reading can be unified under the analysis of aboutness topicality. First, what is normally described as "wide scope" has been given a number of other explanations in the literature; namely, the availability of a wide-scope interpretation arises because these arguments are actually scope-rigid due to being interpreted as topical (Endriss 2009), which are also argued to be properties of D-linked elements (e.g. López 2000).

Appendix B

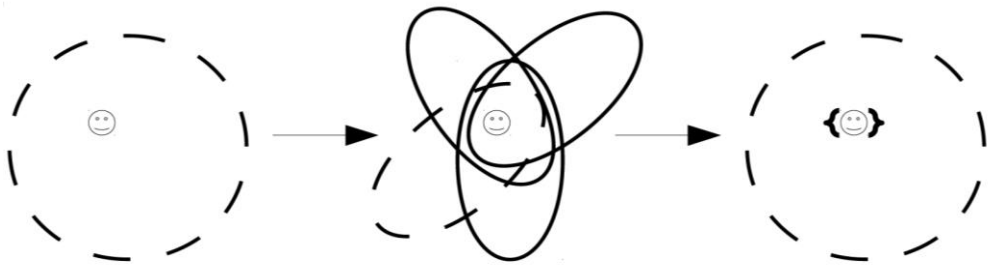
Technical mechanism for topicality- Endriss (2009)

The semantic type of generalized quantifiers i.e., ⟨⟨e,t⟩,t⟩ (Barwise and Cooper 1981) is incompatible with the semantic type of topics, *e*. These semantic types lead to a puzzle: generalized quantifiers cannot function as sentence topics, although it was shown that certain quantifiers can function as topics because they appear in a left-dislocated position which has been shown to be a topic position (e.g., Szabolcsi 1997; Frey 2004; Endriss 2009; Dočekal and Kallulli 2012).

Now the question is: how do quantifiers receive topical interpretation when their semantic types are not compatible with the semantic type of topics which denote individuals? One way of formalizing the idea of topic quantifiers is by invoking the notion of a minimal witness set for the quantifier, that is, a subset of the restrictor set that does not include irrelevant elements (Endriss 2009). The idea of the witness set (based on Barwise and Cooper 1981) is illustrated through the example 'Sara is a teacher,' the witness set for

Sara is the singleton set {s}. The sentence *Sara is a teacher* is true if and only if the witness set for *Sara* is a subset of the set of teachers. I use the figure from Dočekal and Kallulli (2012: 119(21)) to illustrate the steps of deriving the meaning 'Sara is a teacher':

(173)



Now we must determine how the denotation of generalized quantifiers compare to their witness set. I will use the generalized quantifier *at least one girl* which denotes the set of sets of which at least one girl is a member. I assume that only individuals {a,b} are girls. The denotation of this quantifier can contain 'disturbing' elements besides the relevant elements that include at least one girl. I follow the definition of a witness set as defined in Barwise and Cooper (1981):

(174) A set L is a live-on set for a generalized quantifier G if it holds that
$A \in G$ iff $(A \cap L) \in G$

Therefore, the witness sets of the quantifier *at least one girl* are the sets {a}, {b}, and {a, b}. No unwanted elements are in these sets.

Endriss (2009) proposes that only quantifiers that can be mapped on a minimal witness set can function as topics. Based on the definition of the minimal witness set (M) adopted by Szabolcsi (1997), the minimal witness set for at least one girl is therefore the sets {a}, {b} only, and not {a,b}, since {a,b} does not count as a minimal witness set. These two minimal witness sets can then function as the address where the information conveyed by the comment is stored (Endriss 2009). How is this possible? The denotation of the topic, an object of type, needs to be combined with the denotation of the comment which is a

predicate of type. This creates a problem since both the function and argument are of the same type, thus functional application rules cannot be applied here. Endriss proposes a solution in that the elements of the minimal witness set corresponding to the topicalized quantifier are distributed over the elements of the set denoted by the comment.

Furthermore, Endriss continues to develop the classification of topical quantifiers. She follows Kadmon (1985) who claims that all generalized quantifiers introduce (plural) discourse referents. The discourse referents that are created can be different depending on the type of the quantifier. Quantifiers are divided into two groups: those that allow non-exhaustive plural discourse referents, and those that do not. The following pair of examples shows the difference between *three* and *at least three* with respect to anaphoric possibilities and exhaustivity. While *at least three* allows exhaustive anaphoric referent, *three* does not allow it (see Kadmon 1985; Endriss 2009).

(175) a. Yesterday, three teachers came to the wedding. They ate a lot of chocolate.

b. Yesterday, at least three teachers came to the wedding. They ate a lot of chocolate.

The first sentence of the statement in (175a) has three teachers and it can be uttered in a context where there were more than three teachers at the wedding. In contrast, in the second sentence of (175a), the speaker does not tell us that more than three teachers ate chocolate, namely, *they* refers to the three teachers as a set, regardless of how many teachers attended the wedding. On the other hand, in (175b) the pronoun *they* refers to the totality of teachers that ate lots of chocolate. For instance, in a scenario where five teachers were at the wedding, the anaphor *they* in (175b) refers only to the set of all five teachers and cannot refer to the set of four teachers. Going back to the witness set, *they* refers back to one of the witness sets created by the quantifier. However, in (175b) the anaphor cannot refer to any witness of the quantifier *at least three teachers*, but rather refers to the maximal exhaustive intersection of the set of teachers with another set, that is, the set of individuals eating lots of chocolate.

The feature of exhaustivity is determined by the quantifier itself (*three* versus *at least three*) and is not affected by any information-structure properties (Endriss 2009). According to Endriss, the distinction between exhaustivity and non-exhaustivity is anchored in the lexical semantics of the quantifiers. Importantly, the difference between three and at least three depends on how the discourse referent is introduced since we know that all quantifiers introduce plural discourse referents. For instance, for weak quantifiers (numeral) such as three, the discourse referent X refers to a subset of the intersection of set P and Q. This gives the following semantics for a numeral determiner n:

(176) *n*: $\lambda P \lambda Q . \exists X [\, |X| = n \wedge X \subseteq P \cap Q]$

However, for *at least three,* the discourse referent refers to the maximal intersection between P and Q. For *at least n*, the semantics need to be modified in order to cover the exhaustivity property shown above:

(177) *at least n*: $\lambda P \lambda Q . \exists X [\, |X| \geq n \wedge X = P \cap Q]$

Endriss introduces this existentially bound plural variable in order to separate topical quantifiers from other quantifiers. Her main idea is to come up with a test to decide whether a quantifier can function as a topic or not. The test she formulated, called the topic condition, is applied to the semantics of the quantifier.

In order to decide if the quantifier can pass the topic condition, one needs to compare the topic interpretation (by minimal witness sets) to the standard semantics of the quantifier. More specifically, the topic interpretation needs to be equal to the standard semantics of the quantifier in order for it to be interpreted as a topic.

I illustrate Endriss' topic condition using two types of quantifiers: monotone decreasing quantifiers (at most three girls) and weak quantifiers (three girls). The topic interpretation appears on the left-hand side, whereas the normal semantics of the quantifier appears on the right-hand side. I show first that the topical interpretation for three girls is equal to the normal semantics of three girls because the lexical semantics only allows for

reference to sets of 'three girls,' which are considered as the minimal witness sets of the quantifier:

(178) $\exists P[P \subseteq [\![girl]\!] \wedge |P|=3 \wedge P \subseteq Y] = \exists X[|X| = 3 \wedge X \subseteq [\![girl]\!] \cap Y]$

On the other hand, this equivalence is not satisfied when the quantifier is monotone decreasing:

(179) $\exists P[P = \emptyset \wedge P \subseteq Y\ ]\ 0 \equiv \exists X[|X| \leq 3 \wedge X = [\![girl]\!]\ \cap Y]$

The minimal witness set for monotone decreasing quantifiers such as 'at most three girls' is the empty set. Given that the empty set is a subset of any set, the left side is therefore always true. However, the left side is not equal to the semantics of the quantifier, namely the sets of at most three girls.

Endriss expands her analysis using the same test with non-monotone and monotone-increasing quantifiers. She summarizes which quantifiers can be topics and which cannot in Table B.1.

TOPIC	NON-TOPIC
- n	- at most n
- a, some	- at least n
- every	- exactly n
- all	- few

Table B.1: Topic vs. non-topic quantifiers (Endriss 2009: 250)

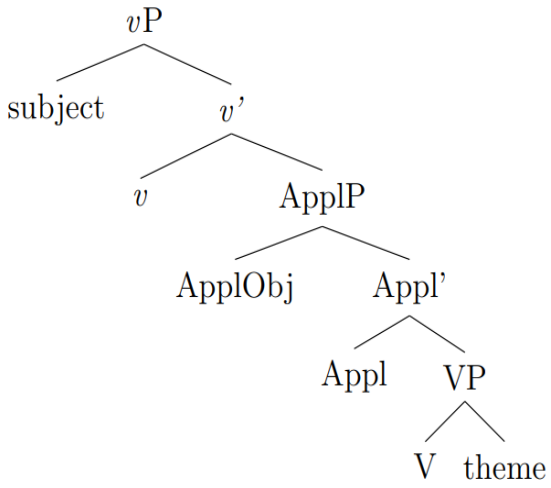
We can see from Table B.1 that weak quantifiers, indefinites, and the universal quantifiers such as 'all' and 'every' are all topics. However, monotone decreasing quantifiers, monotone increasing quantifiers and non-monotone quantifiers are not topics.

Appendix C

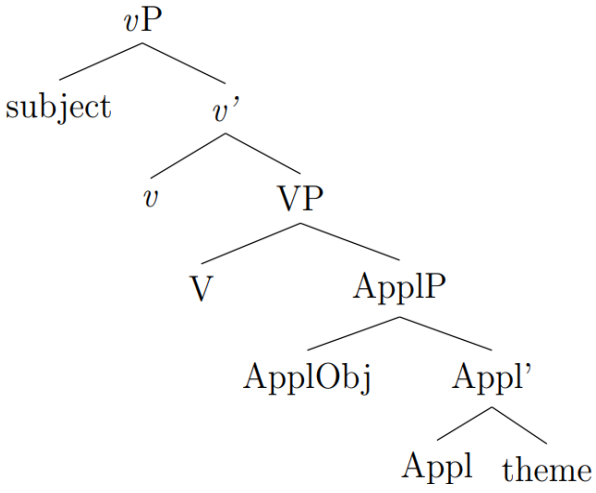
Pylkkänen (2002): high vs. low applicative

The diagnostics proposed in the literature for determining the height of the applicative projection cannot be applied to DOM in Arabic. In her influential work on the syntax of applicatives, Pylkkänen (2002, 2008) proposes two positions of applicative projections: high and low. One possibility is that Appl relates an individual to the event denoted by the VP; consequently, this Appl head is located above VP and is labeled "high" (see (180a)). The other possibility is that the Appl head relates two individuals in a possessor-possessee relationship, the benefactive and the theme. This type of Appl is located within VP and labeled "low" (see (180b)).

(180) a. *High applicative*



b. *Low applicative*



Pylkkänen (2002, 2008) provides three diagnostics that can be used to determine whether a particular applicative construction is high or low: transitivity restrictions, passivization, and depictive secondary predication. As for the first test, Pylkkänen (2002: 23) states that "only high applicative heads should be able to combine with unergatives." We predict that if DOM is a result of a high applicative, it should combine with unergative predicates.

(181) raqasˁt-lo la-aḥmad
danced.1SG-3M.SG.DAT DAT-Ahmad
'I danced for Ahmad.'

On the surface, it seems that DOM is licit with unergatives. However, once we look closely at (181), the construction cannot be characterized as DOM because the clitic-doubled associate *aḥmad* is interpreted as the benefactive, and DOM cannot carry a benefactive meaning.[136]

[136] This type of construction has been referred to in the literature as *ethical datives*. Ethical datives are attested in NGA, Syrian Arabic (Al-Zahre & Boneh 2010) and Lebanese Arabic (Haddad 2013). The ethical dative construction is different from DOM. Although they both have the pronominal element as optional, they differ with respect to the semantic effect. Affectedness seems to play a role in ethical datives, but it does not in DOM as shown below. Al-Zahre & Boneh (2010: 3) wrote that "the eventuality described in the clause affects the individual introduced by the dative in a certain way." The ethical dative example below is from Al-Zahre & Boneh (2010: 2 (5b)).

As for the passivization test, it has been argued that low applicatives only allow passivization of indirect objects (Pylkkänen 2002, 2008). This test cannot be used in Arabic as DOM cannot appear with passivization, regardless of whether the object is direct or indirect.

Third, Pylkkänen (2002, 2008) uses diagnostics such as secondary predication to diagnose low versus high applicatives. I argue that this test, under Pylkkänen's analysis, is expected to be unreliable for Arabic. Pylkkänen suggests that the depictive secondary predication test is only available for secondary depictive predicates of the English type, which Arabic lacks.

According to Pylkkänen, in depictive secondary predication constructions, an adjective attributes a property to one of the arguments of the verb and "the state described by the adjective holds during the event described by the verb" (2002: 27). In order to derive this property of depictive predicates compositionally, she proposes that depictive predicates are composed of an adjective and a depictive head that relates the state denoted by the adjective to an event (2002: 28). The resulting type ⟨e,⟨s,t⟩⟩ phrase is predicted to be able to combine with other arguments of type ⟨e,⟨s,t⟩⟩ by using predicate modification. According to Pylkkänen's analysis, these include high applied arguments and exclude low applied arguments (2002: 31). Thus, a third diagnostic is as follows: "If a language has an

---

(i) ʕali ʕam-ytfalsaf-l-a la-salma
Ali PROG-3SG.M.philosophize-to-3SG.F to-Salma
'Ali is philosophizing on Salma (this aggravates Salma in a certain way).'

It is widely assumed that the Appl head licenses the affectedness component of the meaning of datives. Gogłoza (2021) assumes for Polish that Appl hosts a [+affected] feature. Similarly, based on French data, Boneh and Nash (2011: 64) take the notion of affectedness to be "the intrinsic interpretable feature of Appl." The case for NGA is different. Affectedness does not seem to play a role in DOM in NGA because non-affected objects can be differentially marked.

(i) aḥmad šaf-o la-l-bit
Ahmad saw-3M.SG-3M.SG.OBJ DOM-the-house.M.SG
'Ahmad saw the house.' Non-affected

(ii) aḥmad dahan-o la-l-bit
Ahmad painted-3M.SG-3M.SG.OBJ DOM-the-house.M.SG
'Ahmad painted the house.' Affected

English type depictive secondary predicate, the depictive can modify an applied argument only if the applied argument is high" (Pylkkänen 2002: 31).

In English, while depictive predicates can modify direct objects and subjects, they cannot modify implied passive external arguments. In contrast, depictive predicates cannot modify indirect objects, i.e. low applicatives. Consider the following examples illustrating this difference.

(182) a. John ate the meat raw.
b. John wrote this letter drunk.
c. *This letter was written drunk.
d. *John$_i$ told Mary$_j$ the news drunk$_{i/*j}$. (Pylkkänen 2002: 26 (35))

Based on (182) and the observation that English lacks high applicatives, Pylkkänen proposes that cross-linguistically, high applicatives can be modified by secondary depictives, whereas low applicatives cannot. However, the English examples (182b) and (182d) differ from the NGA counterparts. The examples below are ungrammatical regardless of whether the object is unmarked (baseline) or marked.

(183) a. *aḥmad katab haḏa al-maktoob sakraan/taʕban
Ahmad wrote this the-letter drunk/tired
(‘Ahmad wrote this letter drunk.’) BASELINE

b. * aḥmad katab-o la-l-maktoob sakraan/taʕban
Ahmad wrote-3M.SG.OBJ DOM-the-letter drunk/tired
(‘Ahmad wrote this letter drunk.’) DOM

(184) a. * aḥmad$_i$ ḥaka la-sara$_j$ al-ʔxbaar sakran$_{*i}$/sakran-e$_{*j}$
Ahmad told DAT-Sara the-news drunk.M.SG/-F.SG
(‘Ahmad$_i$ told Sara$_j$ the news drunk$_{*i/*j}$.’) BASELINE

b. * aḥmad$_i$ ḥaka-la la-sara$_j$ al-ʔxbaar sakran$_{*i}$/sakran-e$_{*j}$

Ahmad told-3F.SG.DAT DOM-Sara the-news drunk.M.SG/-F.SG

(‘Ahmad$_i$ told Sara$_j$ the news drunk$_{*i/*j}$.’) DOM

As Pylkkänen notes, the secondary predication test is reliable only for languages that license secondary depictives of the English type. Therefore, for Pylkkänen, this test is not available for languages such as NGA which does not allow subjects to control secondary predicates, in contrast to the English example in (184b).

Thus, based on the results of the tests above, I conclude that neither of Pylkkänen’s diagnostics for determining the attachment height are reliable for DOM in NGA.

Appendix D

DOM vs. Coreferential Dative Construction

Having argued for the applicative structure as the target for DOM, I now compare the height of the ApplP proposed for DOM to another dative construction called *Coreferential Dative Construction* (CDC) (Al-Zahre & Boneh 2010, 2016; Haddad 2013, 2014). I use Al-Zahre and Boneh's (2016) diagnostics for determining the height of the applicative. I show that the results are puzzling and remain an area for future research.

CDC involves a dative element which bears agreement features that are identical to those of the subject in the clause, the Coreferential Dative, as shown in (185). This construction has been documented in Arabic dialects such as Syrian and Lebanese (Al-Zahre & Boneh 2010; Haddad 2013, 2014). The following example of CDC holds true of NGA as well.

(185) *Syrian Arabic* (Al-Zahre & Boneh 2010: 249(1))

salma raʔsˁet-l-a šway

Salma danced.3F.SG-to-3F.SG a little

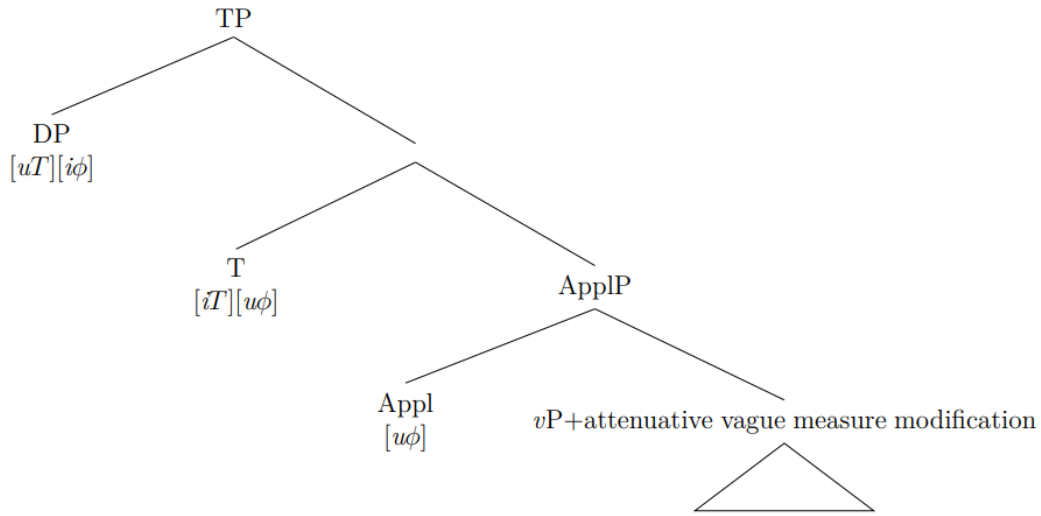
‘≈Salma (just) danced a little (it's a minor issue).’

The presence of the Coreferential Dative in (185) implies that the eventuality, *dancing*, is judged by the speaker to have little significance or weak relevance. The Coreferential Dative expresses the speaker’s attitude toward the degree of relevance of the described eventuality and towards the referent of the subject DP. In Modern Hebrew, Syrian Arabic and French, CDC are often listed with other non-selected datives such as possessive and ethical datives or interested-hearer datives (cf. Borer & Grodzinsky 1986 for Modern Hebrew; Al-Zahre 2003 for Syrian Arabic; Rooryck 2001 for French).

Al-Zahre and Boneh (2016: 22) developed an applicative analysis for CDC in Syrian Arabic represented in (186). Specifically, they argue that CDC merges in the ApplP above *v*P and below TP.[137] Their argument is supported by several distributional tests suggesting that the attachment site of ApplP needs to merge above an event denoting *v*P.

[137] Al-Zahre and Boneh extend this analysis to other non-core dative constructions such as ethical and discourse datives. ApplP in Syrian Arabic always merges in the same position, namely above vP and below TP. Syntactically, this means that Appl has to probe for a goal in the tree to get its features checked by a local DP. In CDC, since the subject is the main participant, Appl therefore can Upward Agree with the subject DP in Spec,TP, which is the closest c-commanding goal with respect to the probe, Appl. The subject DP in CDCs ends up checking the uninterpretable phi-features on both T and Appl, resulting in multiple agree.

(186) *Al-Zahre and Boneh's (2016) analysis of CDC*



Al-Zahre and Boneh propose a diagnostic analysis for the height of ApplP in CDC. In their account, as shown in (186), ApplP must attach to the eventive *v*P, modifiable by an expression of attenuative vague measure.

Al-Zahre and Boneh (2010, 2016) observe that CDC in Arabic is only found with VPs that contain an attenuative vague measure modification, such as *ši* 'some' and *šway* 'a little,' which "denote small quantities of the lower part of a scale" (Al-Zahre and Boneh 2010: 10).[138] The VP needs to contain some scale that can be attenuated. Such a scale can be for instance an incremental theme (187a), the temporal extent of the eventuality (187b), or a scale related to some aspect of manner (187c). The examples below are from Al-Zahre and Boneh (2016: 23(36)).

(187) a. salma ʔakl-t-l-a ši tffaaḥa
Salma ate-3F.SG-to-3F.SG some apple
'≈Salma ate some apple.'

[138] Al-Zahre and Boneh (2016) adopted this term from Filip (2000), who uses it to refer to the verbal prefix *po-* in Russian.

b. rakaḍt-l-i ši rbeʕ saʕa
ran-DAT-1SG some quarter hour
‘‘≈I ran around some 15 minutes or so.’

c. bitkuun salma ḥabbt-l-a (ši) waaḥed mažnuun
is Salma loved-DAT-3F.SG some one crazy.M.SG
‘‘≈Salma must have loved someone crazy.’

Based on these facts, Al-Zahre and Boneh (2016) suggest that the structure of CDC contains ApplP, which attaches to an event denoting *v*P obligatorily containing an attenuative vague measure modification. Al-Zahre and Boneh further argue that an additional indication for this attachment may be the fact that the attenuative vague measure modification can target the temporal extent of the event, and this presumably is available at the *v*P level, not the VP.

Under these constraints, Al-Zahre and Boneh predict that achievements are excluded with CDC, since the attenuative vague measure modification has no appropriate scale to operate on. The following examples from Al-Zahre and Boneh (2016: 24(34a,c)) in (188) show that achievements are incompatible with CDC.

(188) a. *weled-lo žihaad
born-DAT-3MS Jihad
b. * walladet-la laana
gave.birth-DAT-3F.SG Lana

I extend Al-Zahre and Boneh’s diagnostic for the height of ApplP to DOM. In contrast to CDC, DOM does not involve expressions of attenuative vague measure. In the following examples, the adverb in the form of vague measures like *ši* ‘some’ and *šway* ‘little’ are optional and never obligatory, however, when they appear in the DOM context they do not give rise to an inference that the described event is of weak significance, as is the case with CDC.

(189) a. # l-baladyye wassaʕat-a (ši xams mtaar) la-ṭ-ṭariiʔ
the-municipality widened-3F.SG some five meters DOM-the-road
Intended interpretation: (‘≈The municipality widened the road just some five meters (it's a minor issue)).’

b. # sara semʕet-a šway la-l-ɣonnay
Sara heard-3F.SG.OBJ little DOM-the-song
Intended interpretation: (‘≈Sara listened to the song a little (it's a minor issue)).’

The speaker in (189) only declares facts without making any implication towards the described eventuality by seeing it as having weak relevance. Thus, given that DOM, unlike CDC, does not yield the meaning of vague measure modification, we expect that DOM is compatible with achievements since they do not involve an appropriate scale on which to operate. This prediction is borne out as shown below.

(190) a. sara ẓawat-o la-ẓ-ẓay
Sara turned-3M.SG.OBJ DOM-the-light
‘Sara turned on the light.’

b. sara laqat-o la-l-jozdan
Sara found-3M.SG.OBJ DOM-the-purse
‘Sara found the purse.’

I have shown that DOM, unlike CDC, is incompatible with the reading of attenuative vague measure expression. This might imply that DOM cannot target a position higher than vP. However, this test alone cannot be reliable; I used other tests in the thesis such as binding and adverbial placement, and the results are compatible with a position higher than vP. The results of the test used here are nevertheless puzzling and warrant more research in the future.

Appendix E

Coordination data with DOM

This part of the appendix presents coordination data with DOM and offers a preliminary direction for how the movement approach of DOM argued for in this thesis derives the behavior of coordination.

Consider firstly the fact that NGA does not allow first conjunct agreement with postverbal subjects. The examples below also show that when the order of the conjuncts is flipped, first conjunct agreement is still ungrammatical.

(191) a. sara w aḥmad *wesl-et/wesl-o
Sara.F and Ahmad.M arrived-3F.SG/ arrived-3PL
'Sara and Ahmad arrived.'

b. aḥmad w sara *wesel/wesl-o
Ahmad.M and Sara.F arrived.3M.SG/ arrived-3PL
'Ahmad and Sara arrived.'

In NGA DOM, the clitic doubles the features of the entire coordination and the clitic will always be marked plural. Paparounas and Salzmann (2023) refer this this agreement as a resolved agreement.

(192) šoft-**on** la-[sara w aḥmad] fi-l-jamʕa
saw.1SG-3PL DOM-[Sara.F and Ahmad.M]PL in-the-university
'I saw Sara and Ahmad at the university.'

First conjunct agreement is not possible with DOM in NGA (see similar facts for Lebanese Arabic in Akkuş 2021). The examples in (193) demonstrate that even when the order of the conjuncts has been flipped, the results show that the clitic cannot agree with one conjunct regardless of the features.

(193) a. šoft-***a** la-[sara w aḥmad] fi-l-jamʕa
saw.1SG-3F.SG.OBJ DOM-[Sara.F and Ahmad.M]PL in-the-university
'I saw Sara and Ahmad at the university.'
b. šoft-***o** la-[aḥmad w sara] fi-l-jamʕa
saw.1SG-3M.SG.OBJ DOM -[Ahmad.M and Sara.F]PL in-the-university
'I saw Ahmad and Sara at the university.'

This raises the question of how the DOM movement analysis captures the coordination facts. Recall from chapters 3 and 4, that the clitic and differentially marked object start their life in big-DP. The clitic moves to T and the differentially marked objects move to Spec, ApplP. In order to derive the agreement with the entire coordination, the clitic in its landing position needs to probe for features of &P. This is a preliminary analysis and requires further investigation.

Bibliography

Abd-Rabbo, Mohamad. 1990. Sound plural and broken plural assignment in classical Arabic. *Perspectives on Arabic linguistics I*, 55–94.

Abdel-Massih, Ernest, Zaki N. Abdel-Malek, and El-Said M. Badawi. 1981. *A comprehensive study of egyptian arabic*. Center for Near Eastern and North African Studies, Univ. of Michigan.

Abels, Klaus. 2004. Right node raising: Ellipsis or across the board movement. In *Proceedings of NELS 34*, ed. Keir Moulton & Matthew Wolf. Amherst, MA: GLSA.

Abu-Haidar, Farida. 1979. *A Study of the Spoken Arabic of Baskinta*. Leiden and London: E, J. Brill.

Acquaviva, Paolo. 2008. *Lexical plurals: A morphosemantic approach*. Oxford: Oxford University Press.

Aikhenvald, Alexandra. 2008. *The Manambu Language of East Sepik, Papua New Guinea*. Cambridge: Oxford University Press.

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21:435–483.

Akkuş, Faruk. 2021. *On the locus of A'-agreement. Handout of a talk at* Move & Agree Forum, May 31 202, *University of British Columbia.*

Akmajian, Adrian. 1975. More evidence for an NP cycle. *Linguistic Inquiry* 6:115–129.

Al-Balushi, Rashid Ali. 2011. *Case in standard Arabic: The untraveled paths*. PhD thesis, University of Ottawa.

Al-Balushi, Rashid Ali. 2012. Why verbless sentences in Standard Arabic are verbless. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, *57*(1), 1-30.

Al Khalaf, Eman. 2019. Floating quantifiers are autonomous phrases: A movement analysis, *Glossa: a journal of general linguistics* 4(1).

Albuhayri, Salem. 2019. Information structure in standard Arabic verbal sentences. PhD thesis, University of Wisconsin-Milwaukee.

Albukhari, Juman. 2016. The syntax of elliptical constructions in Jordanian Arabic. PhD thesis, University of Wisconsin-Milwaukee.

Alexiadou, Artemis, and Elena Anagnostopoulou. 1997. Toward a uniform account of scrambling and clitic doubling. In Werner Abraham and Elly van Gelderen, eds. *German: Syntactic Problems –Problematic Syntax*. Tübingen: Niemeyer. 143-161.

Alexiadou, Artemis, and Elena Anagnostopoulou. 2000. Clitic doubling and (non-) configurationality. In *Proceedings of the North East Linguistic Society*. 18–28. Amherst, MA: GLSA.

Alexopoulou, Theodora, and Dimitra Kolliakou. 2002. On linkhood, topicalization and clitic left dislocation. *Journal of Linguistics*, *38*(2), 193-245.

Algryani, Ali. 2017. Ellipsis in Arabic fragment answers. *Order and structure in syntax II*, page 319.

Altamimi, Mohammed, Osama Alruwaili, and William J Teahan. 2018. BTAC: A twitter corpus for Arabic dialect identification. In *Proceedings of the 6th conference on computer-mediated communication (CMC) and social media corpora (CMC-corpora 2018)*.

Al-Zahre, Nisrine. 2003. *La Structure du Groupe Verbal en Arabe: Trois arguments empiriques en faveur d'une vision syntaxique de la structure argumentale*. Thèse de Doctorat. Université Paris 8, Saint-Denis.

Al-Zahre, Nisrine, and Nora Boneh. 2010. Coreferential dative constructions in Syrian Arabic and modern Hebrew. *Brill's Journal of Afroasiatic Languages and Linguistics*, *2*(1), 248-282.

Al-Zahre, Nisrine, and Nora Boneh. 2016. Pronominal non-core datives in Syrian Arabic: pragmatics, syntactic, and lexical semantic properties. *Brill's Journal of Afroasiatic Languages and Linguistics*, *8*(1), 3-36.

Alzayid, Ali. 2022. *Arabic dislocation*. Amsterdam: John Benjamins.

Anagnostopoulou, Elena. 1994. Clitic dependencies in Modern Greek. PhD dissertation, University of Salzburg.

Anagnostopoulou, Elena. 1999. Conditions on clitic doubling in Greek. *Empirical Approaches to Language Typology*, pages 761–798.

Anagnostopoulou, Elena. 2003. *The syntax of ditransitives: evidence from clitics*. Berlin: Mouton de Gruyter.

Anagnostopoulou, Elena. 2006. Clitic doubling. In *The Blackwell companion to syntax*, eds. Martin Everaert and Henk van Riemsdijk, 519–581. Malden: Blackwell Publishing. Chap. 14.

Anagnostopoulou, Elena. 2014. The Person Case Constraint (PCC). *The Blackwell Companion to Syntax, 2nd ed. Oxford: Blackwell Publishing*.

Anderson, Stephen. 2005. *Aspects of the theory of clitics*. Oxford: Oxford University Press.

Angelopoulos, Nikos. 2019. Reconstructing clitic doubling. *Glossa: A Journal of General Linguistics* 4(1)(118).

Angelopoulos, Nikos, and Dominique Sportiche. 2021. Clitic dislocations and clitics in French and Greek: from interpretation to structure. *Natural Language & Linguistic Theory*.

Announi, Inas. 2021. The Problem of Word Order and Verbal Movement in Moroccan Arabic. International Journal of Linguistics, Literature and Translation, 4(4): 34-54.

Aoun, Joseph. 1981. *The Formal Nature of Anaphoric Relations*. PhD thesis, MIT.

Aoun, Joseph. 1999. Clitic-doubled arguments. In *Beyond principles and parameters*, pages 13–42. Springer.

Aoun, Joseph, Elabbas Benmamoun and Dominique Sportiche. 1994. Agreement, Word Order, and Conjunction in Some Varieties of Arabic. *Linguistic Inquiry*, 25(2): 195-220.

Aoun, Joseph and Elabbas Benmamoun. 1998. Minimality, reconstruction, and PF movement. *Linguistic Inquiry* 29: 569-597.

Aoun, Joseph, Lina Choueiri, and Norbert Hornstein. 2001. Resumption, movement, and derivational economy. *Linguistic inquiry*, 32(3): 371–403.

Aoun, Joseph, Elabbas Benmamoun and Lina Choueiri. 2010. *The syntax of Arabic*. Cambridge syntax guides. Cambridge University Press.

Arregi, Karlos. 2003. Clitic left dislocation is contrastive topicalization. *University of Pennsylvania Working Papers in Linguistics* 9(1): 31-44.

Arregi, Karlos, and Andrew Ira Nevins. 2012. *Morphotactics: Basque auxiliaries and the structure of spellout*. Studies in Natural Language and Linguistic Theory 86, Dordrecht: Springer.

Asher, R. E., and T. C. Kumari. 1997. *Malayalam. Descriptive grammars*. London: Routledge.

Bachrach, Asaf, and Roni Katzir. 2009. Right-node raising and delayed spellout. In *Interphases: Phase-theoretic investigations of linguistic interfaces*, ed. Kleanthes K. Grohmann. Oxford University Press.

Baker, Mark. 2014. On Dependent Ergative Case (in Shipibo) and Its Derivation by Phase, *Linguistic Inquiry* 45.3: 341—379.

Baker, Mark, and Nadezhda Vinokurova. 2010. Two modalities of case assignment: case in Sakha. *Natural Language and Linguistic Theory* 28: 593–642.

Bakir, Murtadha Jawad. 1980. Aspects of clause structure in Arabic: A study in word order variation in literary Arabic. *Doctoral dissertation, University of Indiana-Bloomington.*

Bale, Alan, and David Barner. 2009. The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics 26* (3), 217–252.

Baltin, Mark. 1981. Strict bounding. In *The logical problem of language acquisition*, ed. C. L. Baker & J. McCarthy. Cambridge, Massachusetts: MIT Press.

Baltin, Mark, and Paul M. Postal. 1996. More on reanalysis hypothesis. *Linguistic Inquiry* 27: 127–145

Bárány, András. 2018. DOM and dative case. *Glossa: a journal of general linguistics* 3(1). 97.

Bárány, András. 2021. Partially ordered case hierarchies. *Glossa: a journal of general linguistics* 6: 76.

Bárány, András, and Laura Kalin. 2020. *Case, agreement, and their interactions: New perspectives on differential argument marking* (Linguistische Arbeiten 572). Berlin: De Gruyter

Barner, David, and Jesse Snedeker. 2005. Quantity judgments and individuation: evidence that mass nouns count. *Cognition* 97:41–66.

Barwise, John, and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4:159–219.

Bateson, Mary Catherine. 1967. *Arabic language handbook*. Georgetown University Press.

Bekins, Peter. 2014. *Transitivity and Object Marking in Biblical Hebrew: An Investigation of the Object Preposition ʾet*. Winona Lake: Eisenbrauns.

Béjar Susana, and Milan Rezac. 2003. Person licensing and the derivation of PCC effects. In *Romance linguistics: Theory and acquisition*, eds. Ana Teresa Perez-Leroux and Roberge Yves, 49–62. Amsterdam: John Benjamins.

Belletti, Adriana. 1999. Italian/Romance clitics: Structure and derivation. In Henk van Riemsdijk (ed.) *Clitics in the Languages of Europe*. Mouton, Berlin-New York, pp. 543-579.

Belletti, Adriana. 2001. Inversion as focalization. *Subject inversion in Romance and the theory of Universal Grammar*, 60:90.

Belletti, Adriana. 2005. Extended doubling and the VP-periphery. *Probus* 17, pp. 1-35.

Belletti, Adriana. 2018. Locality in syntax. In *Oxford research encyclopedia of linguistics*. Oxford, New York.

Benincà, Paola, Giampaolo Salvi, and Lorenza Frison. 1988. L'ordine degli elementi della frase e le costruzioni marcate. In Renzi Lorenzo (ed), *Grande grammatica italiana di consultazione* 1, pp. 129–239.

Benmamoun, Elabbas. 1999. The syntax of quantifiers and quantifier float. *Linguistic Inquiry* 30:621–642.

Benmamoun, Elabbas. 2000. *The Feature Structure of Functional Categories: A Comparative Study of Arabic Dialects*. Oxford: OUP.

Benmamoun, Elabbas. and Heidi Lorimor. 2006. Featureless expressions: When morphophonological markers are absent. *Linguistic Inquiry* 37: 1-23.

Bhatt, Rajesh. 2007. Unaccusativity and case licensing. Talk presented at McGill University, May 7.

Bhatt, Rajesh, and Elena Anagnostopoulou. 1996. Object shift and specificity: Evidence from ko-phrases in Hindi. In Lise Dobrin, Kora Singer & Lisa McNair (eds.), *Papers from the main session of CLS* 32, 11–22.

Blanc, Haim. 1953. *Studies in North Palestinian Arabic*. Jerusalem: The Israel Oriental Society, No. 4.

Bleam, Tonia. 1999. *Leísta Spanish and the syntax of clitic doubling*. Doctoral dissertation, Newark, DE: University of Delaware.

Bleam, Tonia. 2005. The role of semantic type in differential object marking. *Belgian Journal of Linguistics* 19: 3–27.

Blümel, Andreas. 2017. Exocentric root declaratives: Evidence from V2. In *Labels and roots*, ed. Leah Bauke & Andreas Blümel, 468-471. Berlin: Mouton de Gruyter.

Bobaljik, Jonathan. 2003. Floating quantifiers: Handle with care. In *The second Glot International state-of-the-article book*, eds. Lisa L.-S. Cheng and Rint Sybesma, 1.

Bolinger, Dwight. 1961. Contrastive accent and contrastive stress. Language 37(1), 83-96.

Boneh, Nora, and Lea Nash. 2011. *High and higher applicatives: The case of French non-core datives*. Somerville, MA: Cascadilla Proceedings Project. 60–68.

Bonet, Eulàlia. 1994. The Person-Case Constraint: A Morphological Approach. *MIT Working Papers in Linguistics: The Morphology–Syntax Connection*, 22: 33–52.

Borer, Hagit. 1984. *Parametric syntax: Case studies in Semitic and Romance languages*. Dordrecht: Foris.

Borer, Hagit. 2005. *Structuring sense: Volume 1: In name only*. Oxford: Oxford University Press.

Borer, Hagit. 2014. Derived nominals and the domain of content. *Lingua* 141:71–96.

Borer, Hagit, and Yosef Grodzinsky. 1986. Syntactic vs. lexical cliticization: The case of Hebrew dative clitics. In Borer, Hagit (ed.), *The syntax of pronominal clitics*, 175–217. San Francisco: Academic Press.

Bossong, Georg. 1983. Animacy and markedness in universal grammar. *Glossologia* 2:7-20.

Bossong, Georg. 1991. Differential object marking in Romance and beyond. In *New analyses in Romance linguistics*, ed. Douglas A. Kibbee and Dieter Wanner, 143–170. Amsterdam/Philadelphia: Benjamins.

Bresnan, Joan. 1976. On the form and functioning of transformations. *Linguistic Inquiry* 7:3–40.

Bruening, Benjamin. 2010. Double object constructions disguised as prepositional datives. *Linguistic Inquiry* 41(2). 287–305.

Brustad, Kristen. 2000. *The syntax of spoken Arabic*. Georgetown University Press.

Brustad, Kristen. 2008. -āt Drink Your Milks! āt as Individuation Marker in Levantine Arabic. In *Classical Arabic Humanities in Their Own Terms*, ed. B. Gruendler, 1–19.

Buell, Leston. 2005. *Issues in Zulu morphosyntax*. Los Angeles, CA: UCLA dissertation

Büring, Daniel .1997. The meaning of topic and focus—The 59th Street Bridge accent. Routledge, London.

Büring, Daniel. 2014. (Contrastive) Topic. In Caroline Féry & Shin Ishihara (eds.), *Handbook of Information Structure*. Oxford: Oxford University Press

Butt, Miriam. 1993. Object specificity and agreement in Hindi/Urdu. In Katharine Beals, Gina Cooke, David Kathman, Sotaro Kita, Karl-Erik McCullough & David Testen (eds.), *Papers from the 29th regional meeting of the Chicago Linguistic Society (CLS)*, 89–103. Chicago Linguistic Society

Cardinaletti, Anna. 2002. Against optional and null clitics. right dislocation vs. marginalization. *Studia linguistica*, 56(1):29–57.

Caro Reina, Javier. 2020. Differential Object Marking with proper names in Romance languages, in: Kempf, Luise/Nübling, Damaris/Schmuck, Mirjam (edd.), *Linguistik der Eigennamen*, Berlin/Boston, de Gruyter, 225–258.

Casielles-Suárez, Eugenia. 2004. *The syntax–information structure interface: Evidence from Spanish and English*. Oxford: Routledge.

Cecchetto, Carlo. 1999. A comparative analysis of left and right dislocation in Romance. *Studia Linguistica*, 53(1):40–67.

Chapman, Cassandra. 2012. Clitic-doubling in Laurentian french: An auditory acceptability judgment task. *Proceedings of the 2012 Annual Conference of the Canadian Linguistic Association.*

Chierchia, Gennaro. 1998. Plurality of mass nouns and the notion of "semantic parameter". In *Events and grammar*, pp. 53–103. Springer

Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174:99–149.

Chiriacescu, Sofiana. 2014. *The discourse structuring potential of indefinite noun phrases. Special markers in English, German and Romanian*. Stuttgart: University of Stuttgart.

Chomsky, Noam. 1965. Aspects of the theory of syntax. *MIT Press*.

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

Chomsky, Noam. 1993. A Minimalist Program for Linguistic Theory. In *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Ken Hale & Samuel J. Keyser, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework (mitopl 15*). Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, pages 89–155.

Chomsky, Noam. 2001. Derivation by phase. *In Michael Ken-stowicz (ed.), Ken Hale: A life in language, 1–52. Cambridge, MA: MIT Press.*, pages 1–52.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130:33 – 49.

Chomsky, Noam, and Howard Lasnik. 1977. On filters and control. *Linguistic Inquiry* 8:425– 504.

Chyung, Seung Youn (Yonnie), leva Swanson, Katherine Roberts, and Andrea Hankinson. 2018. Evidence-based survey design: The use of continuous rating scales in surveys. *Performance Improvement* 57(5). 38–48.

Cinque, Guglielmo. 1990. *Types of $\bar{A}$-Dependencies*. MIT press Cambridge, MA.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford University Press.

Citko, Barbara. 2017. Right node raising. *The Wiley Blackwell Companion to Syntax, Second Edition*, 1-33.

Coghill, Eleanor. 2014. Differential object marking in Neo-Aramaic. *Linguistics* 52:335–364.

Comrie, Bernard. 1979. Definite and animate direct objects: A natural class. *Linguistica Silesiana* 3:13–21.

Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology*. Chicago, IL: University of Chicago Press.

Contini, Riccardo. 1999. Le substrat araméen en néo-arabe libanais: Préliminaires à une enquête systématique. In Marcello Lamberti & Livia Tonelli (eds.), *Afroasiatica tergestina: Papers from the 9th Italian meeting of Afro-Asiatic (Hamito-Semitic) linguistics, Triest, April 23-24, 1998*, 101–122. Padova: Unipress.

Contini-Morava, Ellen. 2000. Noun class as number in Swahili. In Contini Morava Ellen & Yishai Tobin (eds.). *Between Grammar and Lexicon*, Current Issues in Linguistic Theory 183. Amsterdam/Philadelphia: Benjamins. 3–30.

Corver, Norbert. 2014. Freezing effects. In *The Blackwell companion to syntax,* ed. by Martin Everaert and Henk C. van Riemsdijk, 2:383–406. Malden, MA: Blackwell.

Cowell, Mark. 1964. *A reference grammar of Syrian Arabic (based on the dialect of Damascus).* Washington, D.C.: Georgetown University Press

Croft, William. 1988. Agreement versus case marking and direct objects. In Michael Barlow & Charles A. Ferguson (eds.), *Agreement in natural language: Approaches, theories, descriptions*, 159–286. Stanford, CA: CSLI Publications.

Cruschina, Silvio. 2010. Syntactic extraposition and clitic resumption in Italian. *Lingua*, 120(1):50–73.

Cruschina, Silvio. 2021. The greater the contrast, the greater the potential: On the effects of focus in syntax, *Glossa: a journal of general linguistics* 6(1): 3.

Cuervo, María Cristina. 2003. *Datives at large*. Cambridge, MA: MIT dissertation.

Cuervo, María Cristina. 2010. Against ditransitivity. *Probus* 22. 151–180.

Dali, Myriam. 2020. *Gender and number in Tunisian Arabic: A case of contextual allosemy* (Doctoral dissertation, Université d'Ottawa/University of Ottawa).

Dalrymple, Mary, and Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press.

Danon, Gabi. 2001. Syntactic definiteness in the grammar of modern hebrew. *Linguistics* 39: 1071-1116.

Danon, Gabi. 2008. Definiteness spreading in the hebrew construct state. *Lingua* 118(7): 872-906.

Dayal, Veneeta. 2004. Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy* 27(4). 393–450.

Dayal, Veneeta. 2011. Bare noun phrases. In Paul Portner, Claudia Maienborn & Klaus von Heusinger (eds.), *Semantics: An international handbook of natural language meaning*, vol. 2, 1088–1108. Berlin: De Gruyter Mouton.

Deal, Amy Rose. 2017. Countability distinctions and semantic variation. *Natural Language Semantics 25* (2), 125–171.

De Cat, Cécile. 2002. *French dislocation*. PhD thesis, University of York.

De Cat, Cécile. 2007. *French dislocation: Interpretation, syntax, acquisition*. Oxford University Press.

Déprez, Viviane. 1989. On the Typology of Syntactic Positions and the Nature of Chains. PhD thesis, MIT, Cambridge, Mass.

Despić, Miloje. 2019. On kinds and anaphoricity in languages without definite articles. *Studies in Diversity Linguistics*, (25).

De Hoop, Helen, and Peter de Swart. 2009. Cross-linguistic Variation in Differential Subject Marking. In Helen De Hoop & Peter de Swart (eds) *Differential Subject Marking*, 1-16. Dordrecht: Springer.

de Swart, Peter. 2003. *The Case Mirror*. Nijmegen, The Netherlands: Radboud University Master's thesis.

de Swart, Peter. 2007. *Cross-linguistic variation in object marking*. Nijmegen: Radboud University Nijmegen dissertation.

de Swart, Peter, and Helen de Hoop. 2007. Semantic aspects of differential object marking. *Proceedings of Sinn und Bedeutung* 11, Estela Puig-Waldmüller (ed.), 598–611. Barcelona: Universitat Pompeu Fabra.

de Vries, Mark. 2002. *The syntax of relativization*. Utrecht: LOT.

Diesing, Molly. 1992. *Indefinites*. Cambridge, Massachusetts: MIT Press.

Di Tullio, Ángela, Andrés Saab, and Pablo Zdrojewski. 2019. Clitic Doubling in a Doubling World: The Case of Argentinean Spanish Reconsidered. In *The Syntactic Variation of Spanish Dialects*, pp. 215-244. Oxford University Press.

Dobrovie-Sorin, Carmen. 1990. Clitic Doubling, Wh-Movement, and Quantification in Romanian. *Linguistic Inquiry*, 21, 351–397.

Dočekal, Mojmír, and Dalina Kallulli. 2012. More on the semantics of clitic doubling: principal filters, minimal witnesses, and other bits of truth. In *Empirical issues in syntax and semantics 9*, ed. Christopher Piñón, 113–128. CSSP CNRS.

Doron, Edit, and Ana Muller. 2013. The Cognitive Basis of the Mass-Count Distinction: evidence from bare nouns. In: Patricia Cabredo-Hofherr & Anne Zribi-Herz. (orgs.). *Crosslinguistic Studies on Noun Phrase Structure and Reference*. Leiden: Brill : 73-101.

Döhla, Hans-Jörg. 2016. The origin of differential object marking in Maltese. In Gilbert Puech & Benjamin Saade (eds.), *Shifts and patterns in Maltese*, 149–174. Berlin & Boston: De Gruyter Mouton.

Dryer, Matthew. 2014. Competing methods for uncovering linguistic diversity: The case of definite and indefinite articles (commentary on Davis, Gillon, and Matthewson). *Language* 90(4). e232–e249.

Edwards, Malcolm. 2006. Pronouns, agreement and focus in Egyptian Arabic. *SOAS Working Papers in Linguistics*14: 51-62.

El-Yasin, Mohammed. 1985. Basic Word Order in Classical Arabic and Jordanian Arabic. *Lingua* 65. 107-122.

Enç, Mürvet. 1991. The semantics of specificity. *Linguistic inquiry*, pages 1–25.

Endriss, Cornelia. 2009. Exceptional wide scope. In *Quantificational Topics*, pages 107–185. Springer.

Erbach, Kurt. 2020. Predicting object mass nouns across languages. *Proceedings of the Linguistic Society of America* 5.1 (2020): 228–241.

Erbach, Kurt, Peter Sutton, and Hana Filip. 2019. Bare nouns and the Hungarian mass/count distinction. In Silva, Alexandra & Staton, Sam & Sutton, Peter & Umbach, Carla (eds.), *Language, logic, and computation*, 86–107. Dordrecht: Springer.

Erteschik-Shir, Nomi. 1997. *The dynamics of focus structure*. Cambridge University Press.

Erwin, Wallace M. 2004. *A short reference grammar of Iraqi Arabic*. Georgetown University Press.

Escandell-Vidal, Victoria. 2009. Differential object marking and topicality: The case of Balearic. *Studies in Language* 33(4). 832–885.

Fábregas, Antonio. 2013. Differential object marking in Spanish: State of the art. *Borealis: An International Journal of Hispanic Linguistics* 2(2). 1–80.

Farkas, Donka and Kim Bruce. 2010. On Reacting to Assertions and Polar Questions. *Journal of Semantics* 27.1, p. 81 – 118.

Fassi Fehri, Abdelkader. 1988. Agreement, binding, and coherence. In Michael Barlow & Charles Ferguson (eds.), *Agreement in natural language*, 107–158. Stanford, CA: CSLI.

Fassi Fehri, Abdelkader. 1993. *Issues in the structure of Arabic clauses and words*. Studies in natural language and linguistic theory.

Fassi Fehri, Abdelkader. 2005. Verbal and Nominal Parallelisms. *Documents and Reports* 8. Rabat: Publications IERA.

Fassi Fehri, Abdelkader. 2012. *Key features and parameters in Arabic grammar*, volume 182. John Benjamins Publishing.

Fassi Fehri, Abdelkader. 2016. New roles and architecture for gender: Evidence from Arabic, Semitic, Berber, and Romance. In Michelle Sheehan & Laura R. Bailey (eds.), *Order and structure in syntax II: Subjecthood and argument structure*, 223–258. Berlin: Language Science Press.

Fassi Fehri, Abdelkader. 2020. Number and Gender Convergence: The Arabic Plurative. *Catalan Journal of Linguistics* 19. 75–126.

Feghali, Michel. 1928. *Syntaxe des parlers Arabes actuels du Liban*, Librairie Orientaliste Paul Geuthner, Paris

Feldhausen, Ingo. 2010. *Sentential form and prosodic structure of Catalan*. John Benjamins Publishing Company Amsterdam.

Fernández, Beatriz, and Milan Rezac. 2016. Differential Object Marking in Basque varieties. In Microparameters in the grammar of Basque (Language Faculty and

Beyond 13), ed. B. Fernández & J.O. de Urbina, 93–138. Amsterdam: John Benjamins.

Fernández-Sánchez, Javier. 2017. *Right dislocation as a biclausal phenomenon: evidence from Romance languages*. PhD thesis, Universitat Autonoma de Barcelona.

Filip, Hana. 2000. The Quantization Puzzle. In Pustejovsky, James & Tenny, Carol L. (eds.), *Events as Grammatical Objects, From the Combined Perspectives of Lexical Semantics, Logical Semantics and Syntax*. Stanford, CA: CSLI Publications. 39-95.

Firro, Kais. 2001. Reshaping Druze Particularism in Israel. *Journal of Palestine Studies* 30(3): 40–53.

Fischer, Wolfdietrich. 2002. *A grammar of classical Arabic*. New Haven: Yale University Press 3rd edn. Transl. Jonathan Rodgers.

Fleisch, Henri. 1961. *Traité de philologie Arabe*. Beirut: Dar El-Machreq.

Fox, Danny. 1999. Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry*, 30(2):157–196.

Fox, Danny. 2000. *Economy and Semantic Interpretation*. MIT Press.

Fox, Danny, and Jon Nissenbaum. 1999. Extraposition and Scope: A case for overt QR. In *The Proceedings of WCCFL 18*, ed. Sonya Bird, Andrew Carnie, Jason D. Haugen, & Peter Norquest, 132–144. Sommerville, MA: Cascadilla Press.

Franks, Steven, and Catherine Rudin. 2005. Bulgarian clitics as K^0 heads. In *Formal approaches to Slavic linguistics: The South Carolina meeting*, eds. Steven Franks, Frank Y. Gladney, and Mila Tasseva-Kurktchieva, 106–118. Ann Arbor: Michigan Slavic Publications.

Frascarelli, Mara. 2004. Dislocation, clitic resumption and minimality. In R. Bok-Bennema, B. Hollebrandse, B. Kampers-Manhe & P. Sleeman (Eds.), *Romance languages and linguistic theory 2002* (pp. 99–118). Amsterdam: John Benjamins.

Frascarelli, Mara, and Roland Hinterhölzl. 2007. Types of Topics in German and Italian. In K. Schwabe & S. Winkler (eds.), *On Information Structure, Meaning and Form*, Amsterdam, John Benjamins, 87-116.

Frey, Werner. 2000. Syntactic requirements on adjuncts. *ZAS Papers in Linguistics*, 17:107–134.

Frey, Werner. 2004. Notes on the syntax and the pragmatics of German. *In Horst Lohnstein & Susanne Trissler (eds.), The syntax and semantics of the left periphery*, 9:203.

García García, Marco. 2005. Differential object marking and informativeness. In Klaus von Heusinger, Georg Kaiser & Elisabeth Stark (eds.), Proceedings of the workshop 'Specificity and the evolution/emergence of nominal determination systems in Romance', 17–32. Universität Konstanz: Fachbereich Sprachwissenschaft.

García García, Marco. 2018. *Nominal and verbal parameters in the diachrony of Differential Object Marking in Spanish*, in: Seržant, Ilja A./Witzlack-Makarevich, Alena (edd.), *Diachrony of Differential Argument Marking*, Berlin, Language Science Press, 2018, 209–242.

Gaudefroy-Demombynes, Maurice, and Régis Blachère. 1952. Grammaire de l'arabe classique (morphologie et syntaxe), troisième édition revue et remaniée. Paris, G.P. Maisonneuve et Larose.

Gentner, Dedre, and Lera Boroditsky. 2001. Individuation, relational relativity and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development*. Cambridge, UK: Cambridge Univ. Press.

Gillon, Brendan. 1992. Towards a common semantics for English count and mass nouns. *Linguistics and Philosophy* 15. 597–640.

Givón, Talmy. 1978. Definiteness and referentiality. In J. Greenberg, (ed.*) Universals of human language*, Stanford, CA: Stanford University Press, pp. 291-330

Gogłoza, Aleksandra. 2021. *Polish datives - an applicative analysis*. Berlin: Humboldt Universitat zu Berlin dissertation.

Göbbel, Edward. 2007. Etraposition as PF movement. In *The thirty-fourth western conference on linguistics (WECOL 2006)*, ed. Erin Bainbridge and Brian Agbayani, 132-145. Fresno: Department of Linguistics, California State University, Fresno.

Grashchenkov, Pavel, and Vita. Markman. 2008. Non-Core Arguments in Verbal and Nominal Predication: High and Low Applicatives and Possessor Raising. *Proceedings of the 27th West Coast Conference on Formal Linguistics*, ed. Natasha Abner and Jason Bishop, 185-193. Somerville, MA: Cascadilla Proceedings Project.

Greenberg, Joseph. 1972. Numeral classifiers and substantive number: problems in the genesis of a linguistic type. *Working Papers on Language Universals* 9, 2–39.

Gregoromichelaki, Eleni. 2013. Clitic left dislocation and clitic doubling: A dynamic perspective on left-right asymmetries in Greek. In G. Webelhuth, M. Sailer & H. Walker (eds.), *Rightward movement in a comparative perspective*. Amsterdam: John Benjamins.

Grimm, Scott. 2012. Inverse number marking and individuation in Dagaare. In D. Massam (Ed.), *Count and mass across languages*, pp. 75–98. Oxford: Oxford University Press.

Grimm, Scott. 2018. Grammatical number and the scale of individuation. *Language* 94(3): 527–574.

Grimm, Scott, and Moymír Dočekal. 2021. Counting aggregates, groups and kinds: Countability from the perspective of a morphologically complex language. In Filip, Hana (ed.), *Countability in natural language*, 85–120. Cambridge: Cambridge University Press.

Grosu, Alexander. 1973. On the status of the so-called Right Roof Constraint. *Language* 49:294–311.

Gundel, Jeanette. 1975. *The role of topic and comment in linguistic theory*. New York: Garland.

Gundel, Jeanette. 1985. Shared knowledge and topicality. *Journal of Pragmatics* 9(1). 83–107.

Gundel, Jeanette. 1988. Universals of topic-comment structure. In M. Hammond, E. A. Moravcsik, and J. R. Wirth (eds.), *Studies in Syntactic Typology*, 209–42. Amsterdam: John Benjamins.

Gundel, Jeanette. 1999. Topic, focus and the grammar pragmatics interface. *Penn Working Papers in Linguistics*, 6, pp. 1885-200.

Gunlogson, Christine. 2001. *True to Form: Rising and Falling Declaratives as Questions in English*. Routledge. London.

Gutiérrez-Rexach, Javier. 1999. The structure and interpretation of Spanish neuter degree constructions. *Lingua* 109, 35-63.

Habash, Nizar. 2010. *Introduction to Arabic natural language processing*. Morgan & Claypool Publishers.

Habib, Rania. 2005. *The role of social factors, lexical borrowing and speech accomodation in the variation of [g] and [ʔ] in the Colloquial Arabic of Rural Migrant Families in Hims, Syria*. Unpublished masters thesis, Sracuse University – USA.

Hacohen, Aviya, Olga Kagan, and Dana Plaut. 2021. Differential Object Marking in Modern Hebrew: Definiteness and Partitivity. *Glossa: a journal of general linguistics* 6(1).

Haddad, Youssef. 2013. Pronouns and intersubjectivity in Lebanese Arabic gossip. *Journal of Pragmatics* 49, 57-77.

Haddad, Youssef. 2014. Attitude datives in Lebanese Arabic and the interplay of syntax and pragmatics. *Lingua* 145. 65–103.

Halabi, Rabah. 2006. *Izrakhim Shave Khuvot: Zihot Druzit Vehamedina Hayehudit*: (Citizens equal in duties: Druze identity and the Jewish state). Tel-Aviv: Hakibbutz Hameʔukhad.

Halliday, Michael AK. 1967. Notes on transitivity and theme in English. Part II. *Journal of Linguistics* 3:199-244.

Hallman, Peter. 2018. Double object constructions in Syrian Arabic. *Syntax* 21. 238–274.

Hallman, Peter, and Rashid Al-Balushi. 2022. Pronominalization and clitic doubling in Syrian and Omani Arabic. *Linguistics* 60(5):1295–1336.

Halpert, Claire, and Jochen Zeller. 2015. Right dislocation and raising-to-object in Zulu. *The Linguistic Review* 32:475–513.

Harizanov, Boris. 2011. NonInitiality within spell-out domains: Unifying the post-syntactic behavior of Bulgarian dative clitics. In *Morphology at Santa Cruz: Papers in honor of Jorge Hankamer*, eds. Nick LaCara, Anie Thompson, and Matt A. Tucker. Santa Cruz: UCSC Linguistics Research Center.

Harizanov, Boris. 2014. Clitic doubling at the syntax-morphophonology interface: A-movement and morphological merger in Bulgarian. *Natural Language & Linguistic Theory* 32:1033–1088.

Hartmann, Katharina. 2000. *Right node raising and gapping: Interface conditions on prosodic deletion*. Amsterdam: John Benjamins.

Hedberg, Nancy. 1990. *Discourse pragmatics and cleft sentences in English* (Unpublished doctoral dissertation). University of Minnesota, St. Paul.

Heim, Irene. 1982. The Semantics of Definite and Indefinite Noun Phrases. Doctoral dissertation, University of Massachusetts, Amherst.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Hill, Virginia. 2013. The direct object marker in Romanian: a historical perspective. *Australian Journal of Linguistics* 33. 140–151.

Hill, Virginia, and Alexandru Mardale. 2019. Patterns for differential object marking in the history of Romanian. *In* Journal of Historical Syntax 3(5): 1–47.

Hill, Virginia and Alexandru Mardale. 2021. *The diachrony of differential object marking in Romanian*. Oxford: Oxford University Press.

Hnout, Christine, Lior Laks, and Susan Rothstein. 2021. Mass-to-Count Shifts in the Galilee Dialect of Palestinian Arabic. *Things and Stuff: The Semantics of the Count-Mass Distinction*, 151.

Hockett, Charles F. 1958. A Course in Modern Linguistics. McMillan, New York.

Holes, Clive. 2004. *Modern Arabic: Structures, functions, and varieties*. Washington, DC: Georgetown University Press.

Huddleston, Rodney, and Geoffrey Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge UP.

Iatridou, Sabine. 1991. Topics in conditionals. Cambridge: Massachusetts Institute of Technology PhD dissertation.

Iatridou, Sabine. 1995. Clitics and island effects. *University of Pennsylvania working papers in linguistics*, 2(1):3.

Iemmolo, Giorgio. 2010. Topicality and differential object marking. Evidence from Romance and beyond. *Studies in Language* 34(2). 239–272.

Irimia, Monica Alexandrina. 2020. Types of structural objects. Some remarks on differential object marking in Romanian. In *Case, Agreement and Their Interactions. New Perspectives on Differential Argument Marking*. Edited by András Bárány and Laura Kalin. Berlin: Mouton de Gruyter, pp. 78–126.

Irimia, Monica Alexandrina. 2021. Oblique differential object marking and types of nominals. *Canadian Journal of Linguistics* 66: 486–518.

Irimia, Monica Alexandrina. 2022. DOM and nominal structure. Some notes on DOM with bare nouns. *Languages* 7: 175.

Isleem, Martin. 2016. Arabic-Hebrew code-switching: The case of the Druze community in Israel. *International Journal of Applied Linguistics* 26(2): 228–244.

Ivanov, Ivan. 2009. Topicality and clitic doubling in L2 Bulgarian: A test case for the interface hypothesis. In *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)* (pp. 17-24).

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, Massachusetts.

Jacobs, Joachim. 2001. The Dimensions of Topic-Comment. *Linguistics*, 39:641–681.

Jaeggli, Oswald. 1982. *Topics in Romance syntax* (Studies in Generative Grammar 12). Dordrecht: Foris.

Jaeggli, Oswald. 1986. Three issues in the theory of clitics: Case, doubled NPs, and extraction. In Hagit Borer (ed.), *Syntax and semantics 19: The syntax of pronominal clitics*, 15–42. Orlando, FL: Academic Press.

Jarrah, Marwan 2017. A Criterial Freezing Approach to Subject Extraction in Jordanian Arabic. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 62. 411-448.

Jarrah, Marwan. 2020. Complementizer agreement and the T0-Φ parameter in Jordanian Arabic. *Studia Linguistica*, *74*(1), 139-164.

Johns, Alana, and Ivona Kučerová. 2017. Towards an information structure analysis of ergative patterning in the Inuit language. In *The Oxford handbook of ergativity*, ed. Jessica Coon, Diane Massam, and Lisa Travis, 397–418. New York: Oxford University Press.

Johnson, Kyle. 1985. A case for movement. Doctoral Dissertation, MIT, Cambridge, MA.

Johnson, Kyle. 2007. *LCA+Alignment=RNR*. Talk presented at the Workshop on Coordination, Subordination and Ellipsis, Tubingen, June 2007.

Jones, Michael Allan. 1993. *Sardinian Syntax*. London: Taylor and Francis Routledge.

Kadmon, Nirit. 1985. The Discourse Representation of Noun Phrases with Numeral Determiners, in S. B. J. Choe and J. McDonough (eds), *Proceedings of NELS 15*,

Linguistics Deptartment, University of Massachusetts, GLSA, Amherst, pp. 183–216.

Kagan, Olga. 2020. *The semantics of case* (Key topics in semantics and pragmatics). Cambridge: Cambridge University Press.

Kalin, Laura. 2018. Licensing and differential object marking: The view from Neo-Aramaic. *Syntax* 21:112–159.

Kalin, Laura, and Philipp Weisser. 2019. Asymmetric DOM in coordination: A problem for movement-based approaches. *Linguistic Inquiry* 50:662–676.

Kallulli, Dalina. 1999. The Comparative Syntax of Albanian. On the Contribution of Syntactic Types to Propositional Interpretation. PhD dissertation, University of Durham.

Kallulli, Dalina. 2000. Direct object clitic doubling in Albanian and Greek. In *Clitic Phenomena in European Languages*, ed. by Frits Beukema and Marcel den Dikken, 209-248. Amsterdam: John Benjamins.

Kallulli, Dalina. 2008. Clitic doubling, agreement, and information structure. In *Clitic doubling in the Balkan languages*, ed. Dalina Kallulli and Liliane Tasmowski, 227–255. Amsterdam: John Benjamins.

Kallulli, Dalina. 2016. Clitic doubling as Differential Object Marking. In: *Rivista di Grammatica Generativa: Research in Generative Grammar* 38, pp. 161–171.

Kasher, Natasha, and Aviya Hacohen. 2023. A star is drawn: Testing the culmination inferences of Russian perfective accomplishments, *Glossa: a journal of general linguistics* 8(1).

Kayne, Richard. 1972. Subject Inversion in French Interrogatives, in J. Casagrande and B. Saciuk (eds.) *Generative Studies in Romance Languages*. Newbury House, Rowley, Mass., 70–126.

Kayne, Richard. 1975. *French syntax: The transformational cycle*. Cambridge, MA: MIT press.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Khan, Geoffrey. 1984. Object Markers and Agreement Pronouns in Semitic Languages. *Bulletin of the School of Oriental and African Studies, University of London* 47(3), 468-500.

Kheir, Afifa Eve. 2019. The Matrix Language Turnover Hypothesis: the case of the Druze language in Israel. *Journal of Language Contact* 12(2): 479–512.

Kheir, Afifa Eve. 2023. Codeswitching as an Index and Construct of Sociopolitical Identity. Leiden, The Netherlands: Brill.

Kifle, Nazareth Amlesom. 2007. Differential object marking and topicality in Tigrinya. In *Proceedings of the LFG07 Conference*, ed. Miriam Butt & Tracy Holloway King, 5–25. Stanford, CA: CSLI.

Kiss, Katalin. 1998. Identificational focus versus information focus. *Language*, 74(2):245–273.

Kluck, Marlies, and Mark de Vries. 2013. Cumulative rightward processes. In *Rightward movement in a comparative perspective*, ed. Gert Webelhuth, Manfred Sailer, & Heike Walker, 281–318. Amsterdam: John Benjamins.

Kochovska, Slavica. 2010. Macedonian direct objects, clitics and the left periphery. Doctoral dissertation, Rutgers University.

Kornfilt, Jaklin, and Klaus von Heusinger. 2009. Specifity and partitivity in some Altaic languages. In Ryosuke Shibagaki & Reiko Vermeulen (eds.*), Proceedings of the 5th Workshop on Formal Altaic Linguistics (WAFL 5)*, 19–40. Cambridge, MA: MIT Working Papers in Linguistics.

Kouneli, Maria, and Yuriy Kushnir. 2021. *Failed clitic doubling of plural indirect objects in Greek.* Paper presented at the 52nd annual meeting of the North East Linguistic Society. Rutgers, NJ.

Koutsoudas, Andreas. 1967. Object Particles in Lebanese. *Journal of the American Oriental Society 87*/4. 512—517.

Köylü, Yılmaz. 2023. A-feature based account of genericity and kind reference. *Studia Linguist*.

Kramer, Ruth. 2014. Clitic doubling or object agreement: The view from Amharic. *Natural Language & Linguistic Theory* 32(2). 593–634.

Krause, Elif, and Klaus von Heusinger. 2019. Gradient effects of animacy on Differential Object Marking in Turkish. *Open Linguistics* 5(1):171–190.

Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification. In R. Bartsch, J. van Benthem, and P. van Emde Boas (Eds.), *Semantics and contextual expression*, pp. 75–111. Dordrecht: Springer.

Krifka, Manfred. 2003. Bare NPs: Kind-referring, indefinites, both, or neither? *Proceedings of SALT* 13. 180–203.

Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica*, 55(3-4):243–276.

Krifka, Manfred, and Renate Musan. 2012. Information structure: Overview and linguistic issues. In M. Krifka, & R. Musan (Eds.), *The Expression of Information Structure* (The Expression of Cognitive Categories 5, pp. 1–44). Berlin/Boston: De Gruyter Mouton.

Kroik, David. 2016. Differential Object Marking in South Saami. Licentiate thesis, Umeå: Umeå University.

Kučerová, Ivona, and Ad Neeleman. 2012. Introduction. In *Contrasts and positions in information structure*, 1–23. Cambridge: Cambridge University Press.

Kucerová, Ivona, and Aya Zarka (in preparation). *DOM as a non-default Discourse Commitment Device.*

Kuno, Susumu. 1972. Functional Sentence Perspective: A Case Study from Japanese and English. *Linguistic Inquiry*, 3:269–336.

Lacerda, Renato. 2020. Middle-field syntax and information structure in Brazilian Portuguese. Doctoral dissertation, University of Connecticut.

Ladd, Robert. 1980. *The Structure of Intonational Meaning: Evidence from English*. Bloomington: Indiana University Press.

Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*, volume 71. Cambridge university press.

Lambrecht, Knud. 2001. Dislocation. *Language typology language universals: An international handbook*, 2:1050–1078.

Landman, Fred. 2011. Count nouns, mass nouns, neat nouns, mess nouns. In *The Baltic International Yearbook of Cognition, Logic and Communication, vol. 6: Formal semantics and pragmatics. Discourse, context and models*, ed. B.H. Partee, M. Glanzberg, and J. Skilters, 1–67. Manhattan, KS: New Prairie Press

Leafgren, John R. 1997. Bulgarian clitic doubling: Overt topicality. *Journal of Slavic Linguistics* 5: 117- 143.

Lebeaux, David. 2009. Where Does Binding Theory Apply? *Linguistic Inquiry Monographs*, MIT Press.

Lentin, Jérôme. 2018. The Levant. In Clive Holes (eds.), *Arabic historical dialectology: Linguistic and sociolinguistic approaches*, 170–205. Oxford: Oxford University Press.

Leonetti, Manuel. 2004. Specificity and differential object marking in Spanish. *Catalan Journal of Linguistics* 3. 75–114.

Leonetti, Manuel. 2008. Specificity in clitic doubling and in differential object marking. *Probus* 20(1). 33–66.

Levin, Aryeh. 1987. The Particle la as an Object Marker in some Arabic Dialects of the Galilee. *Zeitschrift für arabische Linguistik*, (17), 31-40.

Levin, Theodore, and Omer Preminger. 2015. Case in Sakha. *Natural Language & Linguistic Theory* 33. 231– 250.

Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8, 339- 359.

Lima, Suzi. 2014. The grammar of individuation and counting. PhD dissertation, University of Massachusetts at Amherst.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, use and interpretation of language,* ed. R Bäuerle, 302–323. Berlin: DeGruyter.

López, Luis. 2000. Ellipsis and discourse-linking. *Lingua* 110:183–213.

López, Luis. 2003. Steps for a well-adjusted dislocation. *Studia Linguistica* 57: 193–231.

López, Luis. 2009. *A derivational syntax for information structure,* volume 23. Oxford University Press.

López, Luis. 2012. *Indefinite objects: scrambling, choice functions and differential marking*. Cambridge, MA & London: MIT Press.

López, Luis. 2016. Dislocations and information structure. *In The Oxford handbook of information structure.*

Lucy, John. 1992. *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge, MA: Cambridge University Press.

Lyassi, Naima. 2012. Standard Arabic Left-dislocation without Movement: Evidence for the presence of a Hanging Topic Left-dislocation Construction. *International Journal of Linguistics*, 4(4):590.

Makkawi, Amani. 2021. *Preverbal subjects in Makkan Arabic: A feature-inheritance approach*. Manitoba: University of Manitoba Winnipeg dissertation.

Manetta, Emily. 2012. Reconsidering rightward scrambling: postverbal constituents in Hindi-Urdu. *Linguistic Inquiry* 43:43 – 74.

Manzini, M. Rita, & Ludovico Franco. 2016. Goal and DOM datives. *Natural Language and Linguistic Theory* 34(1). 197-240.

Manzini, Rita, Leonardo Savoia, and Ludovico Franco. 2020. DOM and Dative in (Italo-)Romance. In *DOM and Dative in (Italo-)Romance*, 219–68. DeGruyter. https://doi.org/10.1515/9783110666137-006.

Marantz, Alec. 1993. Implications of asymmetries in double object constructions. In *Theoretical Aspects of Bantu Grammar*, ed. Sam Mchombo, 113–150. Stanford, CA: CSLI.

Mardale, Alexandru. 2008. Microvariation within Differential Object Marking: Data from Romance, *in Revue roumaine de linguistique* No 4, LIII, Editura Academiei Române, Bucuresti, pp. 448–467.

Mardale, Alexandru, and Petros Karatsareas. 2020. Differential object marking and language contact. *Special issue of Journal of Language Contact.*

Mathieu, Eric. 2009. On the mass/count distinction in Ojibwe. Article presented at the Mass/count workshop, organized by Diane Massam, University of Toronto. Feb. 7-8.

Mathieu, Eric. 2012. Flavors of division. *Linguistic Inquiry 43*, 650–679.

Mathieu, Eric. 2014. Many a plural. In A. Aguilar-Guevara, B. L. Bruyn, and J. Zwarts (Eds.), *Weak referentiality*, pp. 157–181.

McCawley, James D. 1975. Lexicography and the count-mass distinction. In *Annual Meeting of the Berkeley Linguistics Society*, Volume 1, pp. 314–321.

McCawley, James D. 1982. Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* 13:91–106.

McCloskey, James. 1999. On the right edge in Irish. *Syntax*, 2(3):189–209.

McCloskey, James. 2000. Quantifier float and wh-movement in an Irish English. *Linguistic Inquiry* 31: 57–84.

McGregor, Ronald Stuart. 1972. *Outline of Hindi grammar*. Oxford: Clarendon Press.

Milwark, Gary. 1977. Toward an explanation of certain peculiarities of the existential construction in English. *Linguistic analysis* 3(2):1–29.

Mohammad, Mohammad. 1990. The problem of subject-verb agreement in Arabic: Toward a solution. In M. Eid (ed.). *Perspectives on Arabic Linguistics I: Papers from the First Annual Symposium on Arabic Linguistics*, Amsterdam: John Benjamins. 95-125.

Mohammad, Mohammad. 2000. *Word order, agreement, and pronominalization in Standard and Palestinian Arabic* (181). Amsterdam: John Benjamins.

Mohanan, Tara. 1990. Arguments in Hindi. Stanford, CA: Stanford University PhD dissertation.

Molnár, Valéria. 1993. Zur Pragmatik und Grammatik des TOPIK-Begriffes [On pragmatics and grammar of the term topic]. In *Wortstellung und informationsstruktur [Word order and information structure]*, ed. Marga Reis, volume 306, 155–202. Tübingen: Linguistische Arbeiten.

Moutaouakil, Ahmed. 1989. *Pragmatic Functions in a Functional Grammar of Arabic* (Reprint 2014). Berlin, Boston: De Gruyter Mouton.

Müller, Stefan. 2013. Unifying Everything: Some Remarks on Simpler Syntax, Construction Grammar, Minimalism, and HPSG. *Language* 89: 920–950.

Næss, Ashild. 2004. What markedness marks: The markedness problem with direct objects. *Lingua* 114. 1186–1212.

Nevins, Andrew. 2010. *Locality in vowel harmony*. Cambridge, Mass.: MIT Press.

Nevins, Andrew. 2011. Multiple agree with clitics: person complementarity vs. omnivorous number. *Natural Language & Linguistic Theory* 29:939–971.

Nomoto, Hiroki. 2016. Passives and clitic doubling: A view from Classical Malay. *In AFLA 23: The Proceedings of the 23rd Meeting of the Austronesian Formal Linguistics Association*, ed. Hiroki Nomoto, Takuya Miyauchi, and Asako Shiohara, 179-193. Canberra: AsiaPacific Linguistics.

Odria, Ane. 2014. Differential Object Marking and the nature of dative case in Basque varieties. *Linguistic Variation*, 289–317.

Ojeda, Almerindo E. 1992. The semantics of number in arabic. In *Semantics and Linguistic Theory*, Volume 2, pp. 303–326.

Onea, Edgar, and Alexandru Mardale. 2020. From topic to object: Grammaticalization of Differential Object Marking in Romanian. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 65. 1–43.

Ormazabal, Javier, and Juan Romero. 2007. The object agreement constraint. *Natural Language & Linguistic Theory* 25. 315–347.

Ormazabal, Javier, and Juan Romero. 2013. Differential Object Marking, case and agreement. *Borealis: An International Journal of Hispanic Linguistics* 2.2:221–239.

Ouhalla, Jamal. 1991. Functional Categories and Parametric Variation. Routledge, London.

Ouhalla, Jamal. 1994. Focus in Standard Arabic. *Linguistics in Potsdam*, 1:65–92.

Ouhalla, Jamal. 1997. Remarks on focus in Standard Arabic. *Amsterdam Studies in the Theory AND History OF Linguistic Science Series 4*, pages 9–46.

Ouwayda, Sarah. 2014. *Where Number Lies: Plural marking, numerals, and the collective-distributive distinction*. University of Southern California.

Overfelt, Jason. 2015. Rightward movement: A study in locality. Ph.D. dissertation, University of Massachusetts at Amherst.

Owens, Jonathan. 2006. Case and Proto-Arabic. *A Linguistic History of Arabic*. Oxford: Oxford University. Press.

Papangeli, Dimitra. 2000. Clitic doubling in Modern Greek: a head-complement relation. *UCL Working Papers in Linguistics* 12. 473-498.

Paparounas, Lefteris, and Martin Salzmann. 2023. First conjunct clitic doubling in Modern Greek: Evidence for Agree-based approaches to clitic doubling. *Nat Lang Linguist Theory*.

Pavel, Caha. 2023. A default theory of default case. In *GLOW 46*, Vienna.

Perlmutter, David. 1971. *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart and Winston.

Pesetsky, David. 1987. *Wh*-in-situ: Movement and unselective binding. In *The representation of (in)definiteness*, ed. by Eric Reuland and Alice G. B. ter Meulen, 98–130. Cambridge, Mass.: MIT Press.

Pesetsky, David. 1995. *Zero Syntax - Experiences and Cascades*. Cambridge, Mass: MIT Press.

Pimentel, Jonald. 2019. Some biases in Likert scaling usage and its correction. *International Journal of Sciences: Basic and Applied Research, 45*(1), 183–191.

Poole, Ethan. 2017. Movement and the semantic type of traces. Doctoral Dissertation, University of Massachusetts Amherst, Amherst, MA.

Postal, Paul. 1974. *On raising*. Cambridge, MA: MIT Press.

Postal, Paul. 1994. Contrasting extraction types. *Journal of Linguistics* 30:159–186.

Postal, Paul. 1998. *Three investigations of extraction*. Cambridge, MA: MIT Press.

Preminger, Omer. 2009. Breaking agreements: distinguishing agreement and clitic doubling by their failures. *Linguistic Inquiry* 40:619–666.

Preminger, Omer. 2011. Agreement as a fallible operation. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Preminger, Omer. 2014. *Agreement and its failures*. Cambridge, MA: MIT Press.

Pylkkänen, Liina. 2002. Introducing arguments. Doctoral dissertation, Massachusetts Institute of Technology.

Pylkkänen, Liina. 2008. *Introducing arguments* (Linguistic Inquiry Monographs 49). Cambridge, MA: MIT Press.

Quine, Willard Van Orman. 1960. Word and object: An inquiry into the linguistic mechanisms of objective reference.

Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics in pragmatics and philosophy I. *Philosophica anc Studia Philosophica Gandensia Gent, 27*(1):53–94.

Rett, Jessica. 2021. The semantics of emotive markers and other illocutionary content. *Journal of Semantics* 38. 305–340.

Rezac, Milan. 2008a. Phi-agree and theta-related case. In *Phi Theory: phi-features across interfaces and modules*, eds. Daniel Harbour, David Adger & Susana Béjar, 83–129. Oxford: Oxford University Press.

Rivas, Alberto. 1977. A theory of clitics. Doctoral dissertation, MIT, Cambridge, Mass.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, pages 281–337. Springer.

Rizzi, Luigi (ed.). 2003. *The Structure of IP and CP*. *The cartography of syntactic structures, vol. 2*. Oxford: Oxford University Press.

Rizzi, Luigi. 2006. On the Form of Chains: Criterial Positions and ECP Effects. In L. Cheng & N. Corver (eds.), *Wh-Movement: Moving on*. Cambridge, MA: MIT Press, 97-133.

Roberts, Craige. 2011. *Only*: a case study in projective meaning. In *The Baltic International Yearbook of Cognition, Logic and Communication*, vol. 6: *Formal Semantics and Pragmatics. Discourse, Context, and Models*, ed. M Glanzberg, B Partee, J Skilters, pp. 1–56. Manhattan, KS: New Prairie

Roberts, Ian. 2010. *Agreement and head movement: clitics, incorporation, and defective goals*. Linguistic Inquiry Monographs 59, Cambridge, MA: MIT Press

Rochemont, Michael. 1986. *Focus in generative grammar*. Amsterdam: John Benjamins.

Rodríguez-Mondoñedo, Miguel. 2007. The syntax of objects. Doctoral Dissertation, University of Connecticut, Storrs.

Roessler, Eva-Maria. 2019. Differential object marking and object scrambling in the Guaraní language cluster. *Lingvisticæ Investigationes*, 42(1), 31-55.

Rooryck, Johan. 2001. Evidentiality, part I. *Glot International Vol. 5, No. 4*, April 2001 Blackwell Publishers.

Ross, John R. 1967. *Constraints on variables in syntax*. Doctoral Dissertation, MIT.

Rothstein, Susan. 2010. Counting and the mass/count distinction. *Journal of semantics 27* (3), 343–397.

Rothstein, Susan. 2016. *Semantics for Counting and Measuring*. Forthcoming Cambridge University Press. Schvarcz, Brigitta. (2014) Issues in numbers, counting and measuring in Hungarian. MA thesis Bar-Ilan University.

Rothstein, Susan. 2017. *Semantics for counting and measuring*. Cambridge: Cambridge University Press.

Rothstein, Susan. 2021. Counting, plurality, and portions. In Kiss, Tibor & Pelletier, Francis Jeffry & Husić, Halima (eds.), *Things and stuff: The semantics of the count-mass distinction*, 55–80. Cambridge: Cambridge University Press.

Rothstein, Susan, and Roberta Pires de Oliveira. 2020. Comparatives in Brazilian Portuguese: Counting and measuring. In *Mass and count in linguistics, philosophy, and cognitive science*, ed. Friederike Moltmann, 141–158. Amsterdam: John Benjamins.

Rubin, Aaron. 2005. *Studies in Semitic grammaticalization* (Harvard Semitic Studies, vol. 57). Winona Lake: Eisenbrauns.

Rudin, Catherine. 1997. Agr-O and Bulgarian pronominal clitics. In *Formal approaches to Slavic linguis- tics: The Indiana meeting, 1996*, eds. Martina Lindseth and Steven Franks, 224-252. Ann Arbor: Michigan Slavic Publication.

Runic, Jelena. 2014. A new look at clitics, clitic doubling, and argument ellipsis: Evidence from Slavic. Doctoral Dissertation, University of Connecticut, Storrs, CT.

Ryding, Karin. 2005. *A reference grammar of modern standard Arabic*. Cambridge university press.

Sabbagh, Joseph. 2007. Ordering and linearizeing rightward movement. *Natural Language & Linguistic Theory* 25:349–401.

Samek-Lodovici, Vieri. 2006. When right dislocation meets the left-periphery.: A unified analysis of Italian non-final focus. *Lingua*, 116(6):836–873.

Samek-Lodovici, Vieri. 2015. *The interaction of focus, givenness, and prosody: A study of Italian clause structure*, volume 57. OUP Oxford.

Schütze, Carson. 2001. On the nature of default case. *Syntax,* 4(3):205–238.

Schwarzschild, Roger. 1999. GIVENness, AvoidF and other constraints on the placement of accent. *Natural language semantics* 7:141–177.

Schwarzschild, Roger. 2011. Stubborn distributivity, multiparticipant nouns and the count/mass distinction. *North East Linguistic Society (NELS)* 39(2).661–78.

Shaer, Benjamin, Philippa Cook, Werner Frey, and Claudia Maienborn. 2009. *Dislocated elements in discourse:Syntactic, semantic, and pragmatic perspectives*. Routledge.

Shlonsky, Ur. 1997. *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax*. Oxford University Press.

Sichel, Ivy, and Martina Wiltschko. 2021. The logic of Person markedness: Evidence from pronominal competition. *Language* 97(1). 42–71.

Sikos, Les, Minjae Kim, and Daniel J. Grodner. 2019. Social context modulates tolerance for pragmatic violations in binary but not graded judgments. *Frontiers in Psychology*.

Sikuku, Justine, and Michael Diercks. 2022. Object marking in Bukusu: At the interface of pragmatics and syntax. Ms, Moi University and Pomona College.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Robert M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian National University.

Sinnemäki, Kaius. 2014. A typological perspective on Differential Object Marking. *Linguistics*, *52*(2), 281-313.

Smeets, Liz. 2019. The acquisition of object movement in Dutch: L1 transfer and near-native grammars at the syntax–discourse interface. *Second Language Research, 35*(4), 479-504.

Soltan, Usama. 2007. *On formal feature licensing in minimalism: Aspects of Standard Arabic morphosyntax*. PhD thesis, University of Maryland.

Sorace, Antonella, and Frank Keller. 2005. Gradience in linguistic data. *Lingua, 115*, 1497–1524.

Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19:425-449

Sportiche, Dominique. 1996. Clitic constructions. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon* (Studies in Natural Language and Linguistic Theory 33), 213–276. Dordrecht: Springer.

Sprouse, Jon. 2007. A program for experimental syntax. Ph.D. thesis, University of Maryland.

Stalnaker, Robert. 1978. Assertion. In Peter Cole, editor, *Syntax and Semantics*, pages 315–332. Academic Press.

Sterian, Laura-Andreea. 2016. Obligatory and optional resumption; case studies in the syntax of Iraqi Arabic and Romanian. Doctoral thesis, University of Edinburgh.

Stowell, Tim. 1981. *Origins of Phrase Structure*. Ph.D. thesis, MIT.

Strawson, Peter. 1964. Identifying Reference and Truth Values. *Theoria*, 30. reprinted in Steinberg and Jakobovits (eds.), Semantics, 1971.

Suñer, Margarita. 1988. The role of agreement in clitic-doubled constructions. *Natural Language & Linguistic Theory*, 6(3):391–434.

Suñer, Margarita. 1992. Subject clitics in the Northern Italian vernaculars and the matching hypothesis. *Natural Language & Linguistic Theory* 10:641–672.

Sutton, Peter, and Hana Filip. 2016. Mass/count variation, a mereological, two-dimensional semantics. *The Baltic International Yearbook of Cognition, Logic and Communication* 11, 1–45.

Szabolcsi, Anna. 1997. Strategies for Scope Taking. In Anna Szabolcsi (ed.), *Ways of Scope Taking*, Kluwer, Dordrecht, 109–154.

Tal, Shira, Kenny Smith, Jennifer Culbertson, Eitan Grossman, and Inbal Arnon. 2022. The impact of information structure on the emergence of differential object marking: an experimental study. *Cognitive Science*, *46*(3), e13119.

Talmoudi, Fathi. 1980. The Arabic dialect of szisa. *Acta Universitatis Golhoburgensis: Orientalia Gothoburgensia 4*.

Thompson, Ellen. 1996. *The syntax of tense*. Doctoral Dissertation, University of Maryland at College Park.

Tomioka, Satoshi. 2007. Pragmatics of LF Intervention Effects: Japanese and Korean Wh-Interrogatives. *Journal of Pragmatics*, 39(9):1570–1590.

Torrego, Esther. 1998. *The dependencies of objects*. Cambridge, MA: MIT Press.

Torrego, Esther. 2010. Variability in the case patterns of causative formation in Romance and its implications. *Linguistic Inquiry* 41(3). 445–470.

Truckenbrodt, Hubert. 1995. *Phonological phrases: their relation to syntax, focus, and prominence*. Cambridge, Mass.: MIT dissertation.

Turner, Michael. 2018. Definiteness in the Arabic Dialects. Ph.D. dissertation, The University of Texas at Austin, Texas, USA.

Umbach, Carla. 2001. Contrast and contrastive topic. In *Proceedings of esslli 2001 workshop on information structure, discourse structure and discourse semantics: University of helsinki* (pp. 175-188).

Uriagereka, Juan. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26:79–124.

Vallduví, Enric. 1992. *The informational component*. New York: Garlnad Pub.

van Bergen, Geertje, and Peter de Swart. 2009. Definiteness and scrambling in Dutch: Where theory meets practice. In *North East Linguistic Society (NELS)*, eds. Anisa Schardl, Martin Walkow, and Muhammad Abdurrahman. Vol. 38, 89–100. Amherst: GLSA.

Van der Wal, Jenneke. 2022. *A featural typology of Bantu agreement*. Oxford University Press.

Villalba, Xavier. 1998. Right dislocation is not right dislocation. *Studies on the syntax of central Romance languages. Proceedings of the III Symposium on the Syntax of Central Romance Languages*, Olga Fullana & Francesc Roca (eds.), 227–241. Girona: Universitat de Girona.

Villalba, Xavier. 2000. *The syntax of sentence periphery*. PhD thesis, Universitat Autonoma de Barcelona.

von Fintel, Kai. 2008. What is accommodation, again? *Philosophical Perspectives, 22, Philosophy of Language.*

von Heusinger, Klaus, and Georg A. Kaiser. 2003. Animacy, Specificity, and Definiteness in Spanish. In: K. von Heusinger & G.A. Kaiser (eds.) *Proceedings of the Workshop "Semantic and Syntactic Aspects of Specificity in Romance Languages*". Arbeitspapier 113. Fachbereich Sprachwissenschaft, Universität Konstanz, 41-65.

von Heusinger, Klaus 2008. Verbal semantics and the diachronic development of Differential Object Marking in Spanish. *Probus* 20 (1), 1-31.

von Heusinger, Klaus, and Georg A. Kaiser. 2011. Affectedness and differential object marking in Spanish. *Morphology* 21(3–4). 593–617.

von Heusinger, Klaus, Georg A. Kaiser, and Alazne Arriortua. 2018. Differential Object Marking in ditransitive constructions in Basque. In *Proceedings of the VIII Nereus International Workshop "Referential properties of the Romance DP in the context of multilingualism", ed.* T. Parodi, 25–61.

Verbeke, Saartje, and Aaricia Ponnet. 2022. Animacy, Specificity, and Verb Semantics: What Drives Differential Object Marking in Hindi?. *The Heart of Change-Issues on Variation in Hindi*.

Wexler, Kenneth, and Peter Culicover. 1980. *Formal principles of language acquisition*. Cambridge, MA: The MIT Press.

Wilder, Chris. 1995. Rightward movement as leftward deletion. In *On extraction and extraposition in German*, ed. Uli Lutz & Jurgen Pafel, 273–309. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Wilder, Chris. 1999. Right node raising and the LCA. In *Proceedings of WCCFL 18*, ed. SS Bird, A. Carnie, J. D. Haugen, & P. Norquest. Tucson, Arizona: University of Arizona.

Wintner, Shuly. 2000. Definiteness in the Hebrew noun phrase. *Journal of Linguistics* 36(2): 319–363.

Wood, Jim, and Raffaella Zanuttini. 2018. Datives, Data, and Dialect Syntax in American English. *Glossa: a journal of general linguistics* 3 (1): 87.

Woolford, Ellen. 1995. Object agreement in Palauan: Specificity, humanness, economy and optimality. In *University of Massachusetts occasional papers in linguistics* 18: Papers in Optimality Theory, 655–700. GLSA, University of Massachusetts, Amherst.

Woolford, Ellen. 1999. Animacy hierarchy effects on object agreement. In *New dimensions in African linguistics and languages*, ed. Paul Kotey, 203–216. Trenton, NJ: Africa World Press.

Woolford, Ellen. 2008. Active-stative agreement in Lakota: Person and number alignment and portmanteau formation. Ms., University of Massachusetts, Amherst

Wright, William. 1898. *A grammar of the Arabic language* (3rd ed., Vol. 2). Cambridge: Cambridge University Press.

Wright, William. 1933. *A Grammar of the Arabic Language (Volume I*). Cambridge, MA: MIT Press.

Wright, William. 1967. *A grammar of the Arabic language*. Cambridge: Cambridge University Press.

Yeou, Mohamed, Mohamed Embarki, and Sallah Al-Maqtari. 2007. Contrastive focus and F0 patterns in three Arabic dialects. *Nouveaux cahiers de linguistique française, 28*, 317-326.

Yuan, Michelle. 2018. *Dimensions of ergativity in Inuit: theory and microvariation*. M.I.T. PhD Thesis.

Yuan, Michelle. 2021. Diagnosing object agreement vs. clitic doubling: An Inuit case study. *Linguistic Inquiry* 52(1): 153–179

Zabbal, Youri. 2002. *The semantics of number in the Arabic noun phrase*. MA dissertation, University of Calgary.

Zarka, Aya. 2021. Individuation and differential object marking. In *Proceedings of the Canadian Linguistic Association*.

Zarka, Aya, and Aviya Hacohen. 2023. *On the distribution of Differential Object Marking in Levantine Arabic: An experimental study*. Talk presented at Annual Symposium on Arabic Linguistics (ASAL). University of Massachusetts Amherst, March 25, 2023.

Zdrojewski, Pablo, and Liliana Sánchez. 2014. Variation in accusative clitic doubling across three Spanish dialects. *Lingua*, *151*, 162-176.

Zimmermann, Ilse. 2008. On the Syntax and Semantics of kakoj and čto za in Russian. *Journal of Slavic Linguistics, 1*(2), 289-305.

Zubizarreta, Maria Luisa, and Roumyana Pancheva. 2017. A formal characterization of person-based alignment: The case of Paraguayan Guaraní. *Natural Language & Linguistic Theory*, *35*, 1161-1204.

Zwicky, Arnold, and Geoffrey Pullum. 1983. Cliticization vs. inflection: English *n't*. *Language* 59. 502-513.