Who is What and What is Who:
The Morpho-syntax of Arabic WH

Issa M. M. Abdel Razaq

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Queen Mary, University of London
School of Languages, Linguistics & Film

November 2011
Abstract

This thesis advances a micro-parametric analysis for the variation in wh-dependencies in a number of modern Arabic dialects, especially, Iraqi, Lebanese and Jordanian. It will be shown that although these dialects have much in common, there are certain differences in the strategies used in the formation of wh-questions. At a narrower level, it will also be shown that argument wh-phrases such as ‘who’ and ‘what’ in these dialects display asymmetric behaviour in the various wh-questions. In this thesis, I argue that cross-linguistic variation can only be accounted for in terms of morpho-syntactic properties of individual wh-phrases. As far as the Arabic dialects investigated here are concerned, I propose that wh-expressions such as Iraqi meno ‘who’ and Lebanese šu ‘what’, unlike what has been assumed, are copular wh-phrases and, as such, have internally complex structures. It is this internal complexity, I argue, that directly affects their external syntax. To put the findings in perspective, this thesis examines the possibilities that Universal Grammar offers languages in terms of building wh-dependencies ranging from topicalisation and variable binding to relativisation and equation. The thesis, departs away, however, from mainstream approaches to cross-linguistics variation couched in the P&P framework (Chomsky 1981, 1986, 1995), such as LF-movement and binding, on the grounds that they are too rigid to capture the variation observed here. Instead, the thesis supports, and makes a contribution to, novel approaches to cross-linguistic variation, such as the Nanosyntax framework (Starke 2010, 2011), which take syntax to operate on (sub)-morphemic levels. Overall, the analysis has implications for the syntax of wh-constructions in general and the interaction at the morphology-syntax interface in particular.
I consider myself fortunate to have met David Adger, my supervisor, who made the writing of this thesis not only possible, but also enjoyable. Whatever words of thanking and appreciation are said, they cannot be near enough to express my gratitude to David. My sincere deepest thanks to him for his continuous support and for believing in me, even in the hardest times when confidence and enthusiasm were ebbing away. I am also indebted to David for helping me overcome all the difficulties to do with the administrative and financial sides of my PhD. I owe David a great deal for his insightful comments, one-to-one discussions, suggested readings and, above all, for his enormous and impressive knowledge in the various fields of linguistics that will be the beacon of my academic ambition in the years to come.

To my examiners, Dr. Hans van de Koot and Dr. Bernadette Plunkett, I want to express my gratitude for the insightful comments and valuable suggestions which tremendously helped me elaborate on certain aspects of my theory. I would like also to extend to my thanks for all the teaching staff in the School of Languages, Linguistics & Film at Queen Mary. Special thanks to Paul, Sue, Daniel, Luisa, and Linnaea for their vibrant contribution in the seminars and workshops held in the School. I am also grateful to all PhD students in our Linguistics Laboratory for their help in organising all background work for the Syntax-Semantics Reading Group and for being so pleasant. Thank you for such a wonderful experience. I am also thankful to the administration staff at the
School, especially Sharon who has always been a great help. I am also grateful to the staff who organised and taught for the ACTL program at the UCL which provided a wealth of knowledge in the many fields of linguistics.

To my friends in the UK, I am very grateful to you for all the help and support you have given me. A huge thank you to my friend Haythem whose advice was always a guiding light for me. Special thanks to my friends in Reading for easing the burden of being away from family: Skeik, Thanos, Mohsen, Motez, Khalid, Hakeem and Abu-Raghan. To Paulina, a very warm thank you for making my readings away from university surroundings so wonderful.

I would like to take this opportunity and thank my friends from Jordan, Palestine, Lebanon, Iraq, Syria, Libya and Egypt for their judgements and for making the task of collecting data such a valuable and pleasant experience. I would like also to extend my gratitude to Dr. Kawakib Al-Momani at Jordan University of Science and Technology (JUST) for her trust and encouragement. To all my BA and MA/MSC students at JUST, thank you for looking up to me and making me continue my journey more determined than ever before.

I am also indebted to Al-Israa University in Jordan for granting me the scholarship in the last two years of study. I would like to express my appreciation to Prof. Ghanim, Dr. Wajih and Sanaa Saadeh for their support and keeping me up-to-date with the university rules and regulations.

To my family, thank you for putting up with me, especially in the writing up period and the few weeks before submission. I feel blessed to have my mother who has been an example of sacrifice, tolerance and unconditioned love. I am deeply indebted to my father for his prayers for me, his caring and his support that made the last few months painless. A special and warm thank you to my sister Rahma who selflessly offered with pleasure the moral and financial
support in times when I needed it. I would like also to express to thanks to my sister Fatma for her love and confidence in me. She has always been my sister and my friend, and for that I am very grateful. To my little sister, Aisha, thank you for being so tolerant and so pleasant. I am also blessed to have my brothers supporting me all the way through my PhD. For my little brother Abdel-Wahhab, thank you for being there when and where I needed you. A special thank you to Ali, my younger brother for your financial support in my difficult times. Finally, I would like to extend my deepest thanks to my brothers Gassim, Abdel-Ghaffar, Ibrahim as well as the rest of the family for believing in me.
To Muhammad PBUH . . .
# Contents

Introduction 1

1 Copular Constructions in Arabic 12
   1.1 Introduction 12
   1.2 Proposal 15
   1.3 Copular Constructions 17
   1.4 Brief Review 23
   1.5 A Unified Analysis for Copular Constructions 30
      1.5.1 PRON-less Equatives 30
      1.5.2 The Structure of Simple Copular Sentences 37
      1.5.3 Equatives are Left-Dislocations 40
      1.5.4 SV in Standard Arabic as Left-Dislocation 45
      1.5.5 Left-Dislocated Equatives 50
      1.5.6 The structure of LD’ed Equatives 56
      1.5.7 Semantics Issues 60
      1.5.8 Consequences of the LD Analysis 62
   1.6 Summary and Conclusion 65

2 *Meno* is ‘Who is he’ 69
   2.1 The Core Problem 69
   2.2 Approaches to wh-in-situ 73
      2.2.1 LF wh-movement 73
4.2 Arabic (Wh-) Clefts ........................................... 192
4.3 Previous Analyses ............................................. 198
  4.3.1 The Reduced Clefts Analysis: Cheng 1991 ............. 198
  4.3.2 Shlonsky’s 2002 Equative Analysis ...................... 202
4.4 A unified analysis for Arabic (Wh-) Clefts ................. 212
  4.4.1 The cleft clause is a definite relative clause .......... 213
  4.4.2 The structure of the cleft clause ...................... 221
4.5 (Wh-) clefts are equative copular (Wh-) constructions .... 227
  4.5.1 Evidence from the syntax and semantics .............. 227
  4.5.2 The structure of (wh-) clefts ......................... 232
  4.5.3 Consequences ......................................... 238
  4.5.4 English Clefts ........................................ 245
4.6 šu-lli-questions .............................................. 248
4.7 Summary and Conclusion .................................... 254

5 Summary and Conclusion ........................................ 256

Bibliography ..................................................... 266
Introduction

The present thesis is essentially concerned with the typological variation in wh-constructions in Arabic. Many spoken Arabic dialects like Iraqi Arabic (Wahba 1991, Ouhalla 1996 and Simpson 2000), Lebanese Arabic (Aoun and Choueiri 1999, Aoun and Li 2003, and Aoun, Benmamoun and Choueiri 2010), Egyptian Arabic (Wahba 1984, Cheng 1991 and Soltan 2009) and Jordanian Arabic use more than one strategy in the formation of wh-questions. At least four wh-formation strategies avail themselves to speakers of Arabic. These are best illustrated by the following paradigm from Lebanese (Aoun, Benmamoun and Choueiri 2010: 128).

(1) a. šāft ?ayya mmasil b-l-maT’ām?
   saw.2ms which actor in-the-restaurant
   ‘Which actor did you see in the restaurant?’

   b. ?ayya mmasil šāft ∅ b-l-maT’ām?
      which actor saw.2ms in-the-restaurant
      ‘Which actor did you see in the restaurant?’

   c. ?ayya mmasil šāft-o b-l-maT’ām?
      which actor saw.2ms-him in-the-restaurant
      ‘Which actor did you see in the restaurant?’

   d. miin (huwwa) (ya)lli šāft-o b-l-maT’ām?
      who (he) that saw.2ms-him in-the-restaurant
      ‘Who is it that you saw in the restaurant?’

1Jordanian is the author’s native language.
Whereas the question in (1a) represents the wh-in-situ strategy, in which the (d-linked) wh-phrase occurs in the internal argument position, (1b) represents the movement strategy, in which the moved wh-phrase is related to a gap in the argument position. (1c) represents the resumptive strategy whereby wh-phrase appears clause initially and is related to a resumptive pronoun in the position corresponding to the wh-phrase. Finally, (4c) represents what is known as reduced cleft wh-questions (Cheng 1991 and Ouhalla 1996) or Class II interrogatives (Shlonsky 2002 and Aoun, Benmamoun and Choueiri 2010) which are restricted to argument wh-phrases only. In this type of question, the argument wh-phrase occurs in the initial position followed by an (apparently) optional copula pronoun and a relative clause headed by the complementiser (ya)lli ‘that’.

In this thesis I will be essentially dealing with syntactic variation, hence, macro-variation, and morphological variation, hence, micro-variation, phenomena in the Iraqi as well as in the Lebanese dialects. To illustrate, consider, first, the following paradigm from Iraqi.

(2) a. Mona shaafat meno?
   Mona saw.3fs whom
   ‘Who did Mona see?’

 b. Mona ishtarat šeno?
   Mona bought.3fs what
   ‘What did Mona buy?’

(3) a. meno Mona shaafat?
   Mona saw whom
   ‘Who did Mona see?’

 b. šeno Mona ishtarat?
   what Mona bought
   ‘What did Mona buy?’
(4) a. *Mona shaafat men?  
    Mona saw.3fs whom  
    ‘Who did Mona see?’  

    b. *Mona ishtarat šen?  
    Mona bought.3fs what  
    ‘What did Mona buy?’

(5) a. men Mona shaafat?  
    Mona saw whom  
    ‘Who did Mona see?’

    b. šen Mona ishtarat?  
    what Mona bought  
    ‘What did Mona buy?’

As originally reported in Wahba 1991, Iraqi is a wh-in-situ language with optional wh-fronting as can be seen in (2) and (3), respectively. However, the data in (4) and (5), for example, show that Iraqi has another set of argument wh-phrases which display a behaviour that differs from the ones in (2) and (3). More specifically, the wh-phrases men ‘who’ and šen ‘what’ are only permitted in a clause initial position as in (5), but not in situ as shown in (4).

A similar situation to that of the Iraqi wh-questions presented above can also be found in Lebanese wh-questions. In this dialect, there is a sharp contrast between argument wh-phrases in various wh-constructions, as illustrated in the following examples.

(6) a. šeft miin mbeerih?  
    saw.2sm who yesterday  
    ‘Who did you see yesterday?’
b. *şarayte  şu mbeerih  
bought.2sf what yesterday  
‘What did you buy yesterday?’

(7) a. miin şeft mbeerih?  
who  saw.2sm yesterday  
‘Who did you see yesterday?’  

b.  şu şarayte  mbeerih  
what bought.2sf yesterday  
‘What did you buy yesterday?’

(8) a. miin şoft-o b-l-maT'am?  
who  saw.2sm-him in-the-restaurant  
‘Who did you see in the restaurant?’  

b. * şu şarayt-i mn b-l-maktabe  
what bought.2sf-it from-the-bookstore  
‘What did you buy from the bookstore?’

The data from Lebanese shows that the asymmetry lies in the behaviour of individual wh-phrases. More precisely, while the wh-phrase miin ‘who’ occurs fully formed in wh-in-situ questions (6a), wh-movement questions (7a) and resumptive wh-questions (8a), the wh-phrase şu ‘what’ is only well-formed when it occurs clause initially (7b).

The asymmetric variation in Iraqi and Lebanese wh-question poses a challenge for current approaches to parametric variation encapsulated in the well known Principles and Parameters framework proposed in Chomsky (1981, 1986, 1995, 2000). The P&P framework is primarily concerned with the notion that all grammars are invariant cross-linguistically, i.e, Universal Grammar Principles. Within this framework, typological variation follows from language specific rules operating in one language but not the other. These rules are known as Parameters which vary from one language to another, such as the head-directionally
parameter, wh-parameter, null-subject parameter, to mention a few. The one parameter pertinent to this work is the wh-parameter which will be closely examined along with the question of whether the P&P framework captures the variation observed here.

In its recent reformulation, the wh-parameter is couched within the Minimalist Program (Chomsky 1995, 2000) which seeks to divide languages into either wh-movement languages or wh-in-situ languages. The crucial factor responsible for this division is the EPP. The EPP is a property of C which requires its Specifier to project the closest wh-phrase, once Agree has taken place between C and the wh-phrase, Probe and Goal in Chomsky’s terms (Chomsky 2000), respectively. That is, when the EPP is present on C, movement of the wh-phrase to Spec,C is obligatory, otherwise, i.e, when the EPP is not present on C, wh-phrases remain in situ. Thus, languages that have an EPP feature on C are wh-movement languages and those that do not have an EPP feature on C are wh-in-situ languages.

However, such a strict division between languages as to whether or not they have an EPP property on C is both conceptually and empirically problematic. One the one hand, the data presented here constitute a challenge to the MP (Chomsky 1995) in which derivational economy is the driving motivation in seeking to reduce linguistic computation to the minimum. Languages like Iraqi and Lebanese which display wh-movement alongside wh-in-situ appear to have two derivations based on one and the same numeration. If we follow the reasoning as to whether or not a language is set to have an EPP, Iraqi and Lebanese cast a serious doubt on the plausibility of this parameter. Even if we manipulate some of the operations responsible for syntactic derivation, we are still faced with issues affecting derivations, such as discourse and pragmatic factors, that, I believe, are not encoded in the syntax.
At a narrower level, such an approach to parametric variation runs into further problems posed by the distribution of wh-movement and wh-in-situ in both Languages. The data above show that the wh-phrases in the two constructions have asymmetric distribution, be it in the same dialect or across the two dialects. Further data that will be presented in the course of this study will also show that this situation is even more complicated when dealing with long wh-dependencies in both dialects. The picture presented by wh-constructions here makes it impossible to explain within the rigid P&P framework.

Recent trends in current (bio-) linguistics research put the credibility of the P&P approach to parametric variation under scrutiny. While some authors suggest certain identity criteria for parametric variation, as opposed to non-parametric variation, in an attempt to arrive at a correct representation for what Parameters are (Smith and Law 2009), others go to the extent of questioning the very existence of parameters as an approach to linguistic variation (see Boeckx 2010 and the references cited therein). In fact, Chomsky 1995 himself acknowledges the problem encountering research and the question of how to reduce language differences and typology while stating that very “little is understood to venture any strong hypotheses” (1995:6). On this background, Starke (2010, 2011) develops a new approach to parametric variation, namely Nanosyntax, which seeks to address such variation at a (sub-) morphemic level via principles of size differences of lexical items and the phrasal spell-out. While it is still in its early stages, Nanosyntax seems to be a promising approach to cross-linguistic variation departing away from what Starke calls dedicated markers such as the EPP, while at the same time preserving the tenet of the P&P framework.

The picture emerging from the controversy surrounding parametric variation and the question of whether we really have a theory of parametric variation is not a clear one. Delving into the area of parametric variation may have a be-
ginning, but it certainly has no end. Fortunately, it is not my goal in this thesis to settle this issue insomuch as to show how the external syntax of wh-phrases in certain dialects is affected by their internal syntax. Of course, I will be dealing with cases where I show that the Parameters’ approach to cross-linguistic variation is inadequate. And, certainly, the current thesis makes a contribution, based on empirical data, to the issue of how parametric variation should proceed onwards. Still, little do we understand to venture any strong hypotheses.

Let us begin the journey, the present thesis adopts an approach that takes morpho-syntactic properties of wh-phrases to be responsible for their distribution in the various wh-constructions. In particular, I defend the view that variation in Arabic wh-constructions can be straightforwardly accounted for on the basis of the micro-parametric differences between (properties of) individual wh-phrases. Based on cross-dialectal observations, I propose that argument wh-phrases have complex syntax in Arabic, which has been long concealed by standard assumptions about argument wh-phrases being a uniform class, and that the only way to unfold their complex syntax is by identifying morpho-syntactic properties of each wh-phrase. New data from a range of Arabic dialects will be presented with the aim of bringing into light existing asymmetries in the behaviour of individual argument wh-phrases. As far as Iraqi is concerned, once each wh-phrase is assigned its true syntax, it will be revealed that wh-in-situ in Iraqi is not what it appears to be. I will show that Iraqi is a wh-movement language which uses relativization as a means of d-linking which, in turn, licenses wh-in-situ. This means that Iraqi wh-in-situ is d-linked. As for the contrast between Lebanese argument wh-phrases, I will also argue that their morpho-syntactic properties are what defines their distributional patterns. The analysis advanced here has implications for the syntax of wh-constructions in general and for Arabic wh-constructions in particular.
The thesis contains four main chapters. Chapter one lays the ground for subsequent chapters by providing an analysis for copular constructions in Arabic. In particular, the chapter argues that equative constructions of the form [DP [PRON DP]] are left-dislocation structures in which the initial DP is functionally a topic, not a subject, that is base-generated in its surface position. The actual subject is PRON, more precisely, a resumptive strong pronoun that is directly merged in the subject position inside the predicational shell. The chapter begins with an investigation about properties of copular constructions in Arabic. Data and evidence from different dialects will be brought into light with the purpose of showing that among simple copular constructions, there exists one constructions that has received little attention in the literature on Arabic copular constructions, that is, PRON-less equatives of the form [DP DP]. A major consequence of this finding is that equative constructions of the form [DP [PRON DP]] are in fact dislocation structures and have the form they do due to pragmatic reasons: obviating ambiguity and identifying the predicate. A second major consequence is that the problem posed by optionality of the so-called copula pronoun in copular wh-constructions, usually represented as [WH-DP (PRON) DP], is no longer a problem. Under the analysis presented in this study, optionality represented by the structure [WH-DP (PRON) DP] is understood to be two distinct structures. That is, copular wh-constructions involving PRON are assigned the form [Wh-DP [PRON DP]] and are, therefore, syntactically and semantically different from those copular wh-constructions of the form [WH-DP DP]. The findings and conclusion drawn from this chapter will have important implications for both the analysis of copular wh-phrases in the following two chapters and the analysis of (Wh-) clefts in Arabic in the subsequent chapter.

In chapter two, I develop a morpho-syntactic analysis for Iraqi argument wh-phrases. I present the problems associated with Iraqi and Lebanese wh-in-situ constructions and broadly discuss current approaches to wh-in-situ in
the literature. In particular, I will concentrate on two types of approaches: the LF-movement approach to wh-in-situ of Huang 1982 (and Lasnik and Saito 1992) and the binding approach of Pesetsky 1987. Presenting a number of arguments, I will show that whereas the LF-approach is incompatible with the data from Arabic, the binding approach is more plausible, but only when some refinement is undertaken. The refinement in question lies in what I take to be a micro-parametric approach to the (internal) syntax of individual wh-phrases, which will be shown to be responsible for their external syntax. To be more precise, I will argue that whereas the wh-phrases men ‘who’ and šen ‘what’ are simple DPs, the wh-phrase men-o ‘who’ and šen-o ‘what’ are copular wh-phrases which involve the wh-phrase, i.e., men and šen, as a predicate whose external subject is the subject pronoun, i.e., -o ‘he/it’, that it is attached to it. Based on the analysis provided in chapter one for copular constructions, I will argue that Iraqi argument wh-phrases have the same structure as copular wh-constructions. Furthermore, I will show that Iraqi is a wh-movement language, albeit, wh-movement is triggered by the need to check the [WH] feature on the wh-phrase itself, contra Chomsky (1995, 2000) whereby movement is triggered by the operator feature on C. As a consequence, meno will have the structure of the CP ‘who is he?’ that involves internal movement of the wh-phrase men to Spec,CP. The analysis then deals with the distributional problem posed by the occurrence of meno in an argument position in Iraqi matrix wh-in-situ questions. This problem is solved on the basis of facts from free relatives which exhibit the same behaviour. It will be argued that meno is a free relative DP with a [WH] feature. A second consequence of this analysis is that Iraqi employs relativisation as a d-linking strategem for the licensing of wh-in-situ.

Chapter three extends the logic of the analysis advanced in chapter two and tackles the problem posed by the asymmetric behaviour of Lebanese argument wh-phrases. I will argue that Lebanese ‘what’ is the copular CP ‘which thing is it’. The chapter investigates properties and distribution of argument wh-phrase
in Lebanese wh-questions. It will be shown that while the wh-phrases *miin* ‘who’ and *a?yya*-NP ‘which’-NP are DPs and display properties characteristic of argument wh-phrases, the wh-expression *šu* ‘what’ does not. I, then, put forward the hypothesis that the wh-expression *šu* is a complex phrase and it is this complexity that is responsible for the type of conflict observed. Based on observations from other Arabic dialects, I will show that the wh-expression *šu* is derived from the string *?yy š(i) hu* ‘which thing it/he’. A syntactic analysis will be provided for this string in conformity with the analysis of copular constructions proposed in chapter one. The conclusion that will be drawn is that the wh-expression *šu* is indeed complex and it behaves they way it does because it is a CP. The CP analysis of the wh-expression *šu* will account for most of the facts surrounding the asymmetries reported in this study, except for one fact about the wh-expression *šu*, namely, that it occurs in one type of question, i.e., copular wh-questions, which is restricted to DPs. This issue is left for the subsequent chapter to explore.

Chapter four begins with a question carried over from the preceding chapter concerning the occurrence of the wh-expression *šu*, a CP, in wh-equatives and wh-clefts that admit only definite DPs. The first part of this chapter investigates properties of clefts and wh-clefts and shows that such properties are characteristic of equative constructions. Two types of analysis presented in the literature on Arabic wh-clefts, i.e., Cheng’s 1991 and Shlonsky’s 2002, will be discussed to see whether they can account for all the facts presented in this study. Although the two analyses have much to commend them, they fail to explain a number of crucial facts about cleft constructions in Arabic. In the second part of this chapter, I provide an alternative analysis that is based on the analysis of equatives advanced in chapter one. The proposed analysis accounts for all the facts related to clefts and wh-clefts in the various Arabic dialects, with implications for the analysis of English clefts. Lastly, the chapter returns to the question about wh-clefts in which the wh-expression *šu* occurs and whether they receive
the same treatment as wh-clefts involving Wh-DPs like the wh-phrase ‘who’. Two alternatives will be explored, both of which will be shown to constitute further evidence for the base-generation analysis of šu-questions, regardless of question type.

Finally, chapter five summarises the results of the current analysis and brings the discussion to a close with final concluding remarks and implications for further future research.
Chapter 1

Copular Constructions in Arabic

1.1 Introduction

The aim of this chapter is to provide a unified syntactic analysis for present-tense copular constructions in modern Arabic dialects. As is well known, the Arabic languages do not use the present-tense form of the copula in such constructions. The two types of copular constructions reported in the literature are illustrated in (1) and (2) below (Eid 1983, 1991, 1992; Farghal 1986; Doron 1983, 1986; Bahloul 1993; Plunkett 1993; Fassi-Fehri 1993; Ouhalla 1999; Shlonsky and Ouhalla 2002; Benmamoun 2000; Edwards 2006; Al-Horais 2006; Soltan 2007 and Aoun, Benmamoun and Choueiri 2010).

(1) a. nada hiyye l-mudiira
   Nada she the-director
   ‘Nada is the director’

   b. ’ali huwwe l-steez
   Ali he the-teacher
   ‘Ali is the teacher’
c. l-ıwhad humme l-masʔuuliin  
   the-boys they  the-responsible  
   ‘The boys are the responsible ones’

(2) a. Majdi muhandis  
   Majdi engineer  
   ‘Majdi is an engineer’

 b. l-bint helwa  
   the-girl pretty  
   ‘The girl is pretty’

c. ʕali fi-l-gam̲Comparer{
   Ali in-the-university  
   ‘Al is in the university’

The above examples show that present tense copular sentences are verbless. In particular, the first set of examples represents one type of copular construction known as the equative construction which displays a systematic pattern consisting of two definite DPs separated by a subject pronoun. The latter is taken to perform the function of a copula in this type of construction (Eid 1983, 1991, 1992; Ouhalla 1999; Shlonsky 2002, Edwards 2006 and Al-Horais 2006). The second set of examples illustrates another type of copular construction known as the simple copular construction which involves an initial definite DP followed by either an indefinite DP, an AP or a PP (Eid 1992, Benmamoun 2000, Edwards 2006, Al-Horais 2006, and Aoun, Benmamoun and Choueiri 2010).\footnote{I am assuming along the lines of Abuzy 1987 that indefinite noun phrases are Determiner Phrases. The determiner in Arabic indefinite DPs, that is equivalent to the English determiner \textit{a}, is non-overt.}

The problem presented by Arabic present-tense copular constructions above is twofold. As far as simple copular constructions are concerned, the absence of a verbal predicate poses a syntactic problem to do with the structural rep-
representation of such constructions and raises the question of whether they are full clauses like verbal predications or whether they are small clauses. Various views have been entertained in the literature. For instance, Farghal 1986 and Fassi-Fehri 1993 adopt the view that there is a copula that does not surface at the PF component. Benmamoun 2000 and Aoun, Benmamoun and Choueiri 2010, on the other hand, advance the view that simple copular constructions are full clauses, more specifically, TP projections but without a VP projection.

As far as equative constructions are concerned, the occurrence of a subject pronoun in equative constructions presents another problem which essentially centres around the nature of this pronoun. Several studies have argued for an approach that takes equatives to have the structure of full clauses in which the pronoun is said to be the phonetic realisation of subject agreement in I(NFL), much in the same way as subject-verb agreement in verbal predications (Doron 1983, Ouhalla 1999 and Shlonsky 2002). Meanwhile, Eid (1983, 1991, 1992) suggests a functional explanation in which the pronoun is taken to be a predicate head that functions as an anti-ambiguity device in forcing sentential interpretation.

Besides the problems just outlined, there are two particular questions that have not been fully explored in the literature on Arabic copular constructions which will be the main focus of this chapter. The first question is whether simple copular sentences and equative sentences represent two different types of structure or whether they have the same structure, given the verbless nature of such constructions. The second question concerns the structure of copular wh-constructions and whether they represent a different type of wh-question from canonical wh-questions in the Arabic dialects studied here.

The chapter is organised as follows. Section 2 puts forward the proposal for the syntax of Arabic copular constructions. Section 3 presents non-verbal
predications in Arabic in more detail and concentrates on equatives and the properties associated with them. Section 3 looks into previous approaches to copular constructions and shows that they are incompatible with the data presented here. In sections 4 and 5, I put forward the proposal and discuss its consequences. Section 6 is the conclusion.

1.2 Proposal

Based on data and observations from various Arabic dialects, this chapter shows that the inventory of simple copular constructions involves an equative construction without PRON. Building on Cowell 1964 and Heggie 1988, I will adopt the view that a definite DP can be predicative and argue, therefore, that these are subject-predicate constructions. The structure I propose for such constructions is the one advanced in Benmamoun 2000 and Aoun, Benmamoun and Choueiri 2010 in which the DP predicate is merged as complement of T and the DP subject is (base-generated) in Spec,TP, as shown in (3) below.

(3)

```
TP
  /\         /
DP  T'      /
  |         /
Subject T   DP/AP/PP
  |       /
[+Present, +D] Predicate
```

Under this structure, the predicate can either be adjectival (AP), prepositional (PP) or nominal (DP). A nominal predicate can be either an indefinite DP or a definite DP. Presenting data from various dialects, I will show that Arabic does allow the type of sentence in which two definite DPs occur as subject-predicate clauses, i.e. without PRON, on a par with simple copular sentences.

As far as equatives involving PRON are concerned, I will propose an LD analysis along the lines of Cowell 1964, Plunkett 1993, Adger and Ramchand 2003, Edwards 2006 and Soltan 2007. More specifically, the proposed analysis treats equative sentences of the type illustrated in (1) above as left-dislocation structures in which the first DP is base-generated in an A’-position as a discourse topic, rather than a subject. The actual subject is the pronoun in the low subject position with which the left-dislocated DP is coreferential. The structure for equatives featuring PRON is roughly represented in (4).

(4)

The analysis will show that equatives have the form they do, i.e., [DP [PRON DP]], because they are left-dislocation structures that derive from a basic subject-predicate construction. The proposed analysis will be shown to have several consequences. First, left-dislocation will equally apply to simple copular sentences, on a par with equative sentences. Second, as concerns copular wh-constructions, the proposed analysis solves the problem of optionality of PRON in such constructions. Under the analysis proposed here, optionality of PRON is only apparent in that constructions involving PRON and those without represent two different types of construction. More specifically, the analysis
correctly divides such constructions into basic subject-predicate clauses which involve no PRON and into LD’ed constructions which involve PRON, on a parallel fashion to LD’ed equative constructions. Lastly, the findings of this chapter will be the backbone of the analysis developed for wh-clauses in the subsequent chapters.

1.3 Copular Constructions

This section reviews types of copular constructions in Arabic and highlights some of the main properties associated with them. As is well known, Arabic languages allow sentences in which no verbal predicate is involved, i.e., verbless constructions or the so-called nominal sentences (Fassi-Fehri 1993, Plunkett 1993, Benmamoun 2000 and Shlonsky 2002). There are two types of verbless constructions frequently reported in the literature on Arabic copular constructions: predicational copular sentences and equative copular sentences (Eid 1983, 1991, 1992; Farghal 1986; Doron 1983, 1986; Bahloul 1993; Plunkett 1993; Fassi-Fehri 1993; Ouhalla 1999; Shlonsky and Ouhalla 2002; Benmamoun 2000; Edwards 2006; Al-Horais 2006; Soltan 2007 and Aoun, Benmamoun and Choueiri 2010). Predicational copular sentences consist of a definite DP followed by either an indefinite DP, an AP or a PP. These are illustrated in (5a), (5b) and (5c) respectively.

(5) a. Majdi muhandis
    Majdi engineer
    ‘Majdi is an engineer’

    (Jordanian)

b. l-bint elwa
    the-girl pretty
    ‘The girl is pretty’

    (Palestinian)

c. ʿali fi-l-gamʿa
    Ali in-the-university
    ‘Al is in the university’

    (Egyptian)
In spite of the absence of a copular element, the sentences above are fully well-formed and are treated in the literature as copular constructions, on a par with the English copular sentences shown in the translation. The absence of a copula, however, is confined to present-tense copular sentences; in all other tense configurations, a verbal copula is obligatory. For instance, past tense and future tense copular sentences involve an inflected form of the copular verb \( kwn \) ‘be’, as illustrated in (6a-c).

(6) a. Majdi kan muhandis
    Majdi 3ms.was engineer
    ‘Majdi was an engineer’

    b. l-bint kaan-at helwa
    the-girl was-3fs pretty
    ‘The girl was pretty’

    c. Qali ha-y-kun fi-l-gam\'ia
    Ali will-3ms-be in-the-university
    ‘Ali will be in the university’

In addition to the absence of a verbal predicate, another property that characterises Arabic copular constructions is the definiteness of the initial DP. This is due to a general restriction in Arabic that prohibits indefinite DPs from occurring in a clause initial position. An initial DP can, however, be indefinite in this type of construction, but only when it is specific. An indefinite DP is specific in Arabic if it is modified by, say, an adjective or a relative clause (see also Al-Horais 2006). The contrast between bare indefinites and specific indefinites below illustrates this point.

(7) a. *walad b-s-sayaara
    boy in-the-car
    ‘A boy is in the car’
b. walad sgheer b-s-sayaara
   boy little in-the-car
   ‘A little boy is in the car’

c. walad laabis bhuuze hamara b-s-sayaara
   boy wearing shirt red in-the-car
   ‘A boy (who is) wearing a red shirt is in the car’

The other type of non-verbal copular construction is the equative construction. Arabic equatives systematically consist of two definite DPs separated by a subject pronoun, i.e., the so-called PRON (Doron 1983, Ouhalla 1999, Shlonsky 2002 and Edwards 2006). These are illustrated below.

(8) a. nada hiyye l-mdiira
   Nada she the-director
   ‘Nada is the director’

   b. Mhemmed hu l-m'alleem.
   Mhemmed he the-teacher
   ‘Mhemmed is the teacher’

   c. 'ali huwwa l-mudarris
   Ali he the-teacher
   ‘Ali is the teacher’

   d. l-wlaad humme l-mas?uuliin
   the-boys they the-responsible
   ‘The boys are the responsible’

   e. Majdi hu(wwa) Saahib l-benaayeh
   Majdi he owner the-building
   ‘Majdi is the owner of the building’

There are two distinct properties that characterise present-tense equatives in Arabic. The first property is the definiteness of the two DPs forming the equative construction, as seen in the sentences above. The second property is the presence of what is known as the copula pronoun (Eid 1983, 1991, 1992;
Doron 1983, 1986; Ouhalla 1999; Ouhalla and Shlonsky 2002 and Edwards 2006) between the two DPs (8a-c). Copula pronouns are homophonous with third person subject pronouns and have very limited distribution in that they appear only in present-tense equative sentences. In configurations other than present tense, a verbal predicate must be used, much in the same way as simple copular sentences. To further illustrate, consider the sentences in (9a-c)4.

(9) a. ʿali kaan l-mudarris
    Ali 3ms.was the-teacher
    ‘Ali was the teacher’

b. Majdi rah y-kuun l-mudarris
    Majdi fut. 3ms-BE the-teacher
    ‘Majdi will be the teacher’

c. Nadia ha-t-kuun l-masʿuula
    Nadia fut.-3fs-BE the-responsible
    ‘Nadia will be the responsible’

Equatives are also subject to the indefiniteness restriction observed above. Unlike simple copular sentences, however, the initial DP in equatives cannot be a specific indefinite. Consider the cases below.

(10) a. *benet hiyye l-mudiira
    girl she the-director
    ‘A girl is the director’

b. *benet helwa hiyye l-mudiira
    girl pretty she the-director
    ‘A pretty girl is the director’

c. *benet qaʿde b-l-hadeeqa hiyye l-mudiira
    girl sitting in-the-garden she the-director
    ‘A girl (who is) sitting in the garden is the director’

3This is also the case in Hebrew (Doron 1983, 1986; Heggie 1988; Shlonsky 2002, among others).
4Egyptian examples are from Eid (1992) and Soltan (2009), unless stated otherwise.
In spite of the fact that the indefinite DP is modified in (10b, c), such examples are ruled out in the same way as (10a) is. Only definite DPs can occur in this type of construction.

One more issue that is usually discussed in the context of equative constructions concerns agreement of PRON. So far, the patterns of agreement that PRON displays have been shown to be consistent in that PRON shows full agreement with the two DPs forming the equative construction. There are cases, however, where PRON shows limited agreement, such as agreement in Number and/or Gender with the first DP. This might be shown in the examples below.

(11) a. ?anta huwa l-mas?ul
    you(ms) he the-responsible
    ‘You are the one responsible’

    b. *?anta ?anta l-mas?ul
    you(ms) you(ms) the-responsible
    ‘You are the one responsible’

(12) a. enti huwwa enti
    you(fs) he you(fs)
    ‘You are you’

    b. *enti enti enti
    you(fs) you(fs) you(fs)
    ‘You are you’

    c. *enti hiyya enti
    you(fs) she you(fs)
    ‘You are you’

(13) a. ehna hummu l-mas?uullin
    we they the-responsible
    ‘We are the ones responsible’
b. *ehna huwwa l-mas?u liiin
   we he the-responsible
   ‘We are the ones responsible’

On the other hand, there are cases where PRON shows agreement with one of
the two DPs, but not with the other. In (14a), for instance, PRON agrees with
the second DP, but not with the first DP. By contrast, PRON shows agreement
with the first DP rather than with the second DP in (14b)⁵.

(14)  a. ahamm ši b-kell d-daktoraa hiyye l-oTrooha
   important thing.ms in-every doctorate she the-dissertation.fs
   ‘The most important thing in every doctorate is the dissertation’

b. ašhar asor tarixi f-l-balad huwwe l-alfā
   famous palace.ms historical in-the-country he the-fortress.fs
   ‘The most famous historical monument in town is the fortress’

To recap, this section has so far presented the two frequently reported types
of copular constructions: simple copular sentences and equative copular sen-
tences. It has been shown that in both types of construction, the verbal copula
is absent in the present tense; whereas in other tenses, a verbal copula is obliga-
tory. In simple copular constructions, the initial DP may be either definite or a
specific indefinite followed by an indefinite DP, an AP or a PP. As for equative
constructions, they differ from simple copular sentences in two respects: the
definiteness of the two DPs forming the equative sentence and the presence of
a subject pronoun between the two DPs.

The types of copular constructions presented here give rise to a number of
questions. One of the questions is to do with their syntax and whether the
two types of copular construction entail different structures, and, if so, whether
such constructions are full clauses with functional projections like verbal predi-
cations; or whether they are simple subject-predicate structures. Another ques-

⁵These examples are from Syrian Arabic (Cowell 1964).
tion is posed by equatives in which third person subject pronouns apparently perform the function of a copula, hence, presenting a conflict between their (pro)nominal form and their apparent predicative function. The third question concerns the semantics of the two constructions, i.e., whether or not their interpretive properties are the same, given the absence of a verbal predicate. A further complication resides in the irregularities of agreement between PRON and either of the two DPs forming the equative construction. Some of these questions have been addressed in a number of studies in this field. The next section looks into some of the main approaches to copular constructions and see whether the data presented here can be accounted for.

1.4 Brief Review

There is extensive literature on copular constructions in general, and on Arabic copular constructions in particular (Eid 1983, 1991, 1992; Fassi-Fehri 1993; Doron 1983, 1986; Heggie 1988; Bahloul 1993; Plunkett 1993; Shlonsky 1997; Moro 1997; Ouhalla 1999; Benmamoun 2000; Greenberg 2002; Adger and Ramchand 2003; Al-Horais 2006; Edwards 2006 and Aoun, Benmamoun and Choueiri 2010, among others). A full account of analyses presented in the literature is rather beyond the scope of this section. However, I briefly present some of the dominant approaches to the syntax of copular constructions, especially, those that are directly related to the current discussion.

The question of whether or not predicational sentences and equative sentences are syntactically the same has been entertained in Heggie 1988 for English, French and Hebrew as well as in Adger and Ramchand 2003 for Scottish Gaelic. Heggie 1988 provides a unified account for predicative and equative constructions which are said to derive from a single D-structure configuration. The core argument in her analysis of copular constructions is that the copula func-
tions as a kind of a verbal operator that can create a predicate out of any phrasal category. Given that any phrasal category can be turned into a predicate by the copula, both DP and CP predicates are allowed in Heggie’s analysis. Accordingly, this would then naturally account not only for predicative and equative constructions, (15a) and (15b), but also for clefts and pseudo-clefts, (15c) and (15d, e), respectively.

(15) a. John is a teacher.
    b. John is the teacher.
    c. It is John who married Susan.
    d. What Mary hates is Bill’s tie.
    e. Bill’s tie is what Mary hates.

The cases in (15a) and (15b) are interesting in the sense that they are taken by Heggie to be syntactically the same. For Heggie, a definite DP like the teacher in (15b) can be predicative on a par with indefinite DPs like a teacher in (15a). This is attributed to the function of the copula as an operator which seems to be blind to the semantics of DPs, or any phrasal category for that matter, that it selects as a predicate. In this respect, Heggie points out that since selectional restrictions are based on phrasal category, a distinction has to be made between predicative and referential DPs in terms of category types, a costly and an unnecessary move. This aspect of Heggie’s analysis will be crucial to the discussion that follows. Capitalising on Heggie’s treatment of definite DPs, it explicitly entails that such DPs are of type \( e, t \). Heggie’s treatment radically differs from Bowers’s 1993 original work on predicational structures, where DPs are unambiguously saturated expressions, i.e., of type \( e \).

More recently, Adger and Ramchand 2003 explore the hypothesis that apparently different types of predicational structure all reduce to one underlying case. For Scottish Gaelic, Adger and Ramchand present evidence that this is indeed the case, and that the range of variation found in Scottish Gaelic copu-
lar sentences can be attributed to the semantic properties of the constituents. Syntactically, copular constructions are said to consist of a predicational core mediated by the functional projection Predicate Phrase (PredP) (see also Bowers 1993, Chomsky 2000, 2001). The head of PredP takes as its complement only lexical projections NP, VP, AP, PP. PredP is, in turn, dominated by TP. The structure is diagrammed in (16) below.

(16) TP
    /\    /
   /\   /\  
  /\ /\ /\  
 T  T'  PredP  
     /\       /\  
    /\   /\   /\  
   /\ /\ /\ /\  
  T  T  Pred'  XP
     /\     /\  
    /\   /\   /\  
   /\ /\ /\ /\  
  Subject  Pred  XP

However, Adger and Ramchand’s claim that Pred is restricted to taking lexical projections as its complement gives rise to the question of equative copular sentences, which consist of two DPs. DPs would have to be allowed to have more than one interpretation, i.e., a referential interpretation or a predicational interpretation. Crucially, an indefinite NP in Scottish Gaelic has no determiner, and is treated by Adger and Ramchand as an NP. As for definite NPs, since they are preceded by a definite determiner, they are taken to be DPs.

Adger and Ramchand present a range of data to demonstrate the restriction on DPs occurring as a complement of Pred. They develop an analysis based on Zamparelli’s (2000) claim that there are three distinct layers within the DP: the Strong DP (SDP), which includes determiners and results in a referential expression (17a); the Predicative DP (PDP), which results in an expression

---

6In Scottish Gaelic, the head of PredP may be instantiated by a null light verb or by the defective copular verb, depending on the type of copular construction.
that can occur in the same position as a predicative AP (17b); and the Kind DP (KIP), which denotes an intrinsic property (17c) (Adger and Ramchand 2003:18).

(17) a. The dog is barking
    b. Fido is a dog
    c. Fido is a friendly kind of dog

According to Adger and Ramchand, then, Scottish Gaelic DPs headed by the definite determiner can only be SDPs, and are, therefore, restricted to appearing in non-predicative positions. Bare NPs, on the other hand, are KIPs (property denoting), and can occur as the complement of Pred. The conclusion drawn from the analysis of Adger and Ramchand is that apparently different types of predicational structure all reduce to one underlying syntactic structure. As concerns the range of variation found in this language, it is explained in terms of the semantic properties of argument phrases.

Turning to Arabic copular constructions, the question of whether or not predicational constructions and equative constructions are syntactically the same has not been fully explored in the literature. One of the earliest approaches to the syntax of copular constructions in Arabic is developed in Eid (1983, 1991 and 1992). Eid points out that Arabic has no present-tense copular verb and, as such, the language uses pronouns to perform the function of a copula, i.e., the so-called the copula pronoun. Eid divides copular constructions into two types: simple copular sentences and equative copular sentences, as we have seen in the previous section. Accordingly, a copula pronoun appears in the latter, but not in the former. The occurrence of the pronoun in equative constructions is seen by Eid as a means to both eliminate potential ambiguity and warrant sentential interpretation.\(^7\)

\(^7\)I will show, however, that ambiguity arises only in certain cases, i.e., where the two DPs forming the equative sentence are both lexical. Even if the two DPs are both lexical, ambiguity does not arise if pronounced in a certain way. This issue is discussed in detail in the sections

26
Syntactically, equatives involving a copula pronoun are treated as subject-predicate clauses in which the subject occupies Spec,IP while both the copula pronoun and the second DP occupy the predicate position. The structure Eid (1992) suggests for equative constructions is shown in (18).

(18)

\[ IP \]
\[ NP \]
\[ Subject \]
\[ I \]
\[ I' \]
\[ NP \]
\[ AGR \]
\[ PRESENT \]
\[ N \]
\[ N_i \]
\[ huwwa \]

In the structure above, the copula pronoun \textit{huwwa} originates as head of an NP predicate and forms a single constituent with this NP. As far as agreement is concerned, Eid argues that the copula pronoun gets its agreement features from the second NP in Spec position. This is motivated by the assumption that the type of agreement involved in equative constructions is NP-agreement, which does not require person agreement. Consequently, in cases where the second DP is pronominal, a clash is said to occur between person features of the two pronouns leading to suspension in person agreement on the copula pronoun (\textit{cf.} examples (11)-(13)).

However, the analysis outlined above has several problems. Starting with agreement, the assumption that agreement in equatives holds between the copula pronoun and the predicate is not empirically supported. In the previous

ahead.
section, we have seen that agreement is inconsistent in that in some cases the copula pronoun agrees with the subject, but not with the predicate, while in others it shows the opposite, and vice-versa. A similar observation is also made in Edwards (2006) who draws attention to cases where copula pronouns do not show agreement with the DP predicate. Instead, copula pronouns apparently show agreement with the subject, as shown in the following sentences.

(19) a. il-muškila  hiyya  iT-Talaba
    the-problem(f.s) she  the-students(m.pl)
    ‘The problem is the students’

          b. iT-Talaba  humma  il-muškila
          the-students(m.pl) they(m.pl) the-problem(f.s)
          ‘The students are the problem’

In terms of the structure Eid assigns for equative constructions; it can be pointed out that the structure is incompatible with the standards of syntactic linearisation (Antisymmetry theory of Kayne (1994)) whereby Spec-Head-Complement is the universal order. Under Eid’s 1992 analysis, the specifier of the predicate head N, i.e., the copula pronoun, is apparently merged to the right of its head.

The third problem observed in the analysis developed in Eid (1983, 1991 and 1992) is to do with the semantics of the copular construction. More specifically, recall the indefiniteness restriction in equatives which bans the occurrence of (both bare and specific) indefinite DPs in the initial position of the clause. Under Eid’s analysis, this kind of restriction, simply, remains unexplained.

These authors differ from Eid (1983, 1991 and 1992) in that the copula pronoun is said to be the spell-out of subject agreement in INFL much in the same way as subject-verb agreement in verbal constructions. The main assumption underlying the INFL-analysis is that copula pronouns exhibit agreement patterns parallel to standard subject-verb agreement. However, such analysis is called into question not only by the agreement facts observed here, but also by other facts about negation which dictate that the copula pronoun cannot be located in INFL (See Eid 1992 and Edwards 2006).

To sum up, this section has focused on two of the main studies on copular constructions which deal with the issue of unifying apparently different types of copular constructions. Both Heggie 1988 and Adger and Ramchand 2003 provide a unified account for copular constructions in languages such as English and Scottish Gaelic, which also have crucial implications for the analysis of copular constructions in other languages. The two differ, however, in the approach they develop, while Heggie’s approach is based on the syntax of such constructions, Adger and Ramchand take the semantics of elements that make up such constructions to determine their syntax. On the other hand, the picture seems less clear for Arabic copular constructions. The majority of studies on Arabic (Eid 1983, 1991 and 1992, Ouhalla 1999 and Doron 1983 and Shlonsky 2002) have been mainly concerned with equative constructions, especially, the question concerning the role PRON plays in this type of construction. As it stands, there is a gap in the literature on this topic and the question of whether a unified analysis for these constructions is possible remains untouched. This is precisely the question that I want to address in the next section.

Like Arabic, Hebrew present tense copular sentences are formed with a pronominal copula, identical to third person pronouns, and having masculine and feminine forms, both singular and plural.
1.5 A Unified Analysis for Copular Constructions

This section provides a unified analysis for copular constructions. The analysis is essentially motivated by two novel empirical questions, the answer of which will have crucial consequences for the analysis of copular constructions in Arabic and similar languages. The first question is whether the grammar of Arabic allows equative sentences without PRON, on a par with simple copular sentences. In other words, is the presence of PRON indeed obligatory in equative sentences as is generally reported (Eid 1983, 1991, 1992, Ouhalla 1999, Shlonsky 2002 and Edwards 2006)? The second question is whether simple copular sentences can be formed in the presence of PRON on a par with equative sentences. Put differently, is the absence of PRON obligatory? The answer to both questions should be, at least syntactically, a ‘yes’. Empirically, I will show in this section, using data from various Arabic dialects, that the constructions in question are indeed available, but have received little or no attention in the literature on copular and equative constructions in Arabic. These two issues are crucial for the analysis of copular constructions, a discussion of which, I believe, is particularly illuminating.

1.5.1 PRON-less Equatives

The first question raised above is whether equative constructions involving two DPs without PRON are available in Arabic languages, and, if so, what implications this may have for the analysis of copular constructions. Such an issue can be probed further by a closer inspection of simple copular sentences, especially, the fact that there is no restriction imposed by the syntax on the category of the XP that may follow the initial definite DP. The sentences are repeated below, for convenience.
(20)  a. Majdi muhandis  
     (Jordanian) 
     Majdi engineer 
     ‘Majdi is an engineer’ 

b. l-bint helwa  
     (Palestinian) 
     the-girl pretty 
     ‘The girl is pretty’ 

c. ūlā fi-l-gamʿa  
     (Egyptian) 
     Ali in-the-university 
     ‘Al is in the university’ 

As can be seen in (20), the second XP in this type of construction can be either 
an indefinite DP (20a), an AP (20b) or a PP (20c). Such categorial variation 
raises the question of whether the syntax admits a definite DP in the XP posi-
tion in this type of construction, given that the difference between an indefinite 
DP and a definite DP resides in their semantics. In theory, it should be possible 
to have a definite DP in the XP position in sentences like (20), there being no 
restriction in the language on the occurrence of definite DPs in such construc-
tions9. In this respect, it can be noted that such a view is supported by the fact 
that the second DP in equative constructions is a definite DP10.

The assumption that syntax imposes no restriction on the XP category in 
predicational sentences gains empirical support from various Arabic dialects. As 
it turns out, there is already evidence that constructions in which the second DP 
is a definite DP are indeed available, just like simple copular sentences, albeit 
such sentences are, in certain cases, pronounced in a certain way. The first piece 
of evidence comes from the work of Cowell 1964 on Syrian Arabic. Cowell 1964 
points out that a nominal predicate in Arabic may be definite just like other 
indefinite, adjectival and prepositional predicates. Accordingly, sentences that 
involve two definite DPs are usually pronounced with a pause between the two

9See Adger and Ramchand 2003 for a different view.
10I am assuming that the nominative pronoun that separates the two DP in equatives has no effect on the type of the XP category following it.
DPs: the end of the initial DP is drawled, usually with a rising intonation. To illustrate, consider the following examples (rising intonation is marked by arrows and the pause is marked by the comma between the two DPs.).

(21) a. ?abuu-hun↑, ?adel
    father-their Adel
    ‘Their father is Adel’

b. ra?iis l-wazaara↑, ra?iis l-hukuume l-haqiqi
    head the-ministry, head the-government the-actual
    ‘The prime minister is the actual head of the government’

Moreover, in the case where the predicate is definite, the predication is said to be equational with the order of the subject and the predicate being reversible.

(22) a. ?adel ?abuu-hun
    Adel their-father
    ‘Adel is their father’

b. ra?iis l-hukuume l-haqiqi, ra?iis l-wazaara
    head the-government the-actual, head the-ministry
    ‘The actual head of the government is the prime minister’

These observations raise the issue as to whether the so-called copula pronoun, i.e., PRON, is indeed obligatory in equative constructions as is generally assumed. To address this issue, the reason why such a pronoun is said to be obligatory must be carefully examined. Early studies take the copula pronoun to function as an anti-ambiguity device (Eid 1983, 1991 and 1992). For instance, in her work about present-tense equative constructions in Egyptian Arabic, Eid 1983 posits that without the (copula) pronoun between the two DPs, constructions involving two definite DPs would be ambiguous in the sense that they might have either a sentential interpretation or a phrasal interpretation. To illustrate, consider the following example given in Eid 1983 (example (13) in Eid’s study).
According to Eid, (23) can be analysed either as an equative sentence with the DP *il-mudarris* ‘the teacher’ as subject and the AP *il-latiif* ‘the nice’ as predicate; or as a complex noun phrase whose head is the noun ‘*mudarris* ‘teacher’, that is modified by the AP *il-latiif* ‘the nice’. Empirically, the only interpretation available in this language, according to Eid, is phrasal interpretation. For sentential interpretation to obtain, a third person pronoun that agrees with the subject in number and gender must be used. Hence, the function Eid assumes for such pronouns is as an anti-ambiguity device.

Eid’s story misses two important facts, however. The first fact is to do with intonation and prosodic status of constructions that involve two lexical DPs. Ambiguity is no longer an issue when such cases are uttered in a way that conveys the intended meaning. A copula pronoun is, therefore, not obligatory per se. In other words, the use of a copula pronoun is not the only means to obtain sentential interpretation. For the latter to obtain, prosody, too, can be employed. Speakers pronounce such cases with a prosodic break between the two DPs. Now consider the following examples from Jordanian and Egyptian.

(23) *il-mudarris il-latiif*  
the-teacher the-nice  
‘The nice teacher’

(24) a. Ahmad, l-muhandis  
Ahmad, the-engineer  
‘Ahmad is the engineer’

b. saayeq t-taxi, l-mas?uul  
driver the-taxi, the-responsible  
‘The taxi driver is the responsible one’

---

11 Note that in the example Eid provides, the phrase *il-latiif* ‘the nice’ is also ambiguous between a DP and an AP. Examples where the second element is unambiguously a DP provide a better ground for testing whether two DPs can be equated.
As marked by the comma between the two DPs in (24) and (25), such sentences are pronounced with a short pause between the two DPs, with the intonation rising at the end of the first DP. If, on the other hand, such cases are pronounced as one constituent, the (a) examples would be interpreted as complex noun phrases while the (b) examples would be interpreted as appositive relative clauses.

The second fact that is not accounted for in Eid’s story is that constructions that involve two lexical DPs are only a subcase of the overall paradigm of two definite DP constructions. In cases where one of the two DP is, for instance, a (demonstrative) pronoun, ambiguity does not arise. As a matter of fact, such constructions are common across the dialects, as can be shown in the following paradigm.

(26)  

a. huwwa l-muttaham  
he the-suspect  
‘He is the suspect’  

b. huwwa z-Za’iym  
he the-leader  
‘He is the leader’  

c. haida l-fenneen  
this the-artist  
‘This is the artist’
In all the sentences above, the initial DP is pronominal while the second DP is lexical. As such, ambiguity is not an issue. The same holds also in cases where one of the two DPs is an argument wh-phrase, as will be shown later. It is worth noting that the order of two DPs is reversible in case where they are lexical, as can be seen in (24) and in (27) below. However, equatives in which the second DP is pronominal are not so common (see (28a-c) below), except in one case, namely, where the subject DP is a demonstrative pronoun as shown in (29).

(27) a. l-muhandis, Ahmad
    the-engineer, Ahmad
    ‘The engineer is Ahmad’

b. l-mas?uul, saayeq t-taxi
    the-responsible, driver the-taxi
    ‘The responsible one is the taxi driver’

(28) a. ??l-muttaham huwwa
    the-suspect he
    ‘He is the suspect’

b. ??z-za'eeem huwwa
    the-leader he
    ‘He is the leader’

c. ??l-fenneen haida
    the-artist this
    ‘This is the artist’
The data so far constitute empirical evidence that the presence of PRON in equatives is not obligatory after all. The data have also shown that equatives of the form [DP DP], which I will call PRON-less equatives, are common across the spoken dialects. This is further supported by facts from Standard Arabic (SA). Consider the following sentences (Fassi-Fehri 1993:117).

(30) a. l-junuu-du (hum) l-mas?uul-uun
the-soldiers-Nom (they) the-responsible-nom
‘The soldiers are the responsible ones’

b. ?anta (huwa) l-mas?uul-u
you (he) the-responsible
‘You are the one responsible’

According to Fassi-Fehri 1993, the occurrence of PRON between the two DPs is optional. In fact, as will be argued in the discussion below, it is not the case that PRON is optional, or obligatory for that matter, in such sentences; rather, the case is that equatives with PRON are structurally different from PRON-less equatives. The point to be made at this stage is that PRON-less equatives are indeed available in various Arabic dialects, including the standard variety.

In light of the fact that PRON-less equative constructions are indeed available in Arabic, just like simple copular sentences, two issues may be raised at this point. The first issue concerns the structure of copular sentences, includ-
ing PRON-less equatives. The second issue concerns equatives which involve PRON and the question of whether they represent a different type of construction. These two issue are discussed in turn in the following two sections.

1.5.2 The Structure of Simple Copular Sentences

As concerns the structure of copular constructions, I propose that both PRON-less equative sentences of the from \([DP \, DP]\) and simple copular sentences of the form \([DP \, DP/APP/PP]\) be analysed as subject-predicate structures\(^{12}\). Following Bahloul 1993, Benmamoun 2000, Aoun, Benmamoun and Choueiri 2010, I take copular constructions in Arabic to be TP projections\(^{13}\). They differ from verbal predications in that they involve no verbal projection, which is precisely what distinguishes the two types of predication. The structure can be roughly represented below (notice that the second DP in predicate position may be a definite or an indefinite DP).

(31)

\[
\begin{array}{c}
TP \\
\downarrow \\
DP \\
\downarrow \\
T' \\
\downarrow \\
Subject \\
\downarrow \\
T \\
\downarrow \\
DP/APP/PP \\
\downarrow \\
[+Present, \,+D] \\
Predicate
\end{array}
\]

Under this structure, the subject is directly merged in Spec,TP. Assuming along the lines of Chomsky (2001b), I take T in Arabic equatives to have an EPP property which requires a DP in its specifier position\(^{14}\). The structure in

\(^{12}\)The view that definite DPs can be, in addition of being referential, predicative is entertained in the work of Heggie 1988 as we have seen earlier. Heggie proposes that definite descriptors are predicative based on evidence from English, French and Hebrew, which I take to be the case for Arabic DPs as well. Under this analysis, then, the difference between simple copular sentences and PRON-less equatives reduces to the semantic properties of the DP predicate.

\(^{13}\)See Aoun, Benmamoun and Choueiri 2010, chapter 2, for a full discussion on the structure of verbless sentences.

\(^{14}\)The EPP feature is understood here as the requirement to be “an occurrence of something” where an occurrence of α is a sister of α (Chomsky 2001b). As for the indefiniteness restriction
(31) also incorporates the assumption that T in present-tense copular constructions has an abstract tense feature. Following Benmamoun 2000 and Aoun, Benmamoun and Choueiri 2010, I take present tense here to be specified for the feature [+D], which does not require a verb to be attracted to it since there is no dependency established between T and the verb. Evidence for the presence of an abstract tense feature on T is based on the fact that present-tense copular constructions are compatible with present tense adverbs, but not with past or future tense adverbs. This can be illustrated in the following examples.

(32) a. huwwa l-mudarris hatha l-faSl
    he the-teacher this the-term
    ‘He is the teacher this term’

    b. *huwwa l-mudarris l-faSl l-maaDi
    he the-teacher the-term the-last
    ‘He is the teacher last term’

Given the assumptions above, consider now the structure for PRON-less equative sentences like (33a) and (34a) given in (33b) and (34b) below.

(33) a. Saahib š-šareka, l-mudiir t-tanfeethi
    owner the-company, the-director the-executive
    ‘The company owner is the executive director.’

    b. TP
        DP                     T
        Saahib š-šareka       T
        The company owner
        |                     [+D]
        l-mudiir t-tanfeethi [+Present]
        the executive director

on the DP in Spec,TP, it will be discussed in the upcoming section.

15 Unlike the feature [+D], the feature [+V] requires verb movement, hence, the fact that past tense copular constructions contain the verbal copula is explained.
(34) a. l-mudiir t-tanfeethi, Saahib š-š-areka
  the-director the-executive, owner the-company
  ‘The executive director is the company owner.’

b.

\[
\begin{array}{c}
\text{TP} \\
\text{DP} & \text{T'} \\
\text{l-mudiir t-tanfeethi} & \text{T} & \text{DP} \\
\text{The executive director} & \text{[+D]} & \text{Saahib š-š-areka} \\
\text{[+Present]} & \text{the company owner}
\end{array}
\]

Likewise, the structure for simple copular sentences is the same as that for PRON-less equatives illustrated in (33) and (34). For instance, a sentence like (35a) has the representation in (35b).

(35) a. Majdi muhandis
    Majdi engineer
    ‘Majdi is an engineer’

b.

\[
\begin{array}{c}
\text{TP} \\
\text{DP} & \text{T'} \\
\text{majdi} & \text{T} & \text{DP} \\
\text{Majdi} & \text{[+Present, +D]} & \text{muhandis} \\
\text{an engineer}
\end{array}
\]

Under the present analysis, both predicational sentences of the type illustrated in (35) and equative sentences of the type illustrated in (34) are syntac-
tically the same. Both are treated as subject-predicate clauses despite the fact that in equatives the second DP is a definite DP.

To summarise, the last two sections have entertained the question of whether PRON-less equatives are possible in Arabic. I have shown that Arabic languages do contain in their grammar the type of construction which involves two definite DPs without a pronoun separating them. I have shown that sentences which involve two lexical DPs are only a subcase of the overall paradigm of two-DP constructions. As concerns the issue of ambiguity arising with lexical DPs (Eid 1983, 1991, 1992), I have shown that such cases follow a certain pattern of intonation which renders them sentential without the use of PRON. As such, the long-standing assumption that PRON is obligatory in equative sentences is not entirely true.

In terms of analysis, I have shown that PRON-less equatives and simple copular sentences are syntactically the same. Such a treatment is supported by, in addition to facts from various dialects, an early observation made in Cowell 1964 who points out that a nominal predicate can be a definite NP. Similarly, under the analysis of Heggie 1988, definite DPs can be predicative and, hence, occur in a predicate position. The structure that has been, therefore, assigned to both types of sentence is one and the same structure. This analysis gives rise to the question about equative constructions of the form [DP PRON DP], to which I turn next.

1.5.3 Equatives are Left-Dislocations

Introduction

In the previous section, I have shown that equatives consisting of two DPs without PRON are not only available in Arabic languages, but they are also common. Syntactically, these have been treated as subject-predicate structures
on a par with simple copular sentences. As such, this poses the question of
whether equatives involving PRON are related to the basic [DP DP] construction
or whether they represent a different type of construction. In this section, I put
forward the proposal that equative constructions of the form [DP PRON DP]
are derived from an original equative construction via left-dislocation. Left-
Dislocation constructions, traditionally known in the grammar of Arabic as
Topic-Comment constructions (Farghal 1986, Fassi-Fehri 1993, Plunkett 1993
and Soltan 2007), are a common phenomenon in Arabic languages and display
a similar array of properties to those associated with equative constructions.
Before articulating my proposal, I briefly discuss (properties of) topic-comment
and left-dislocated constructions in Arabic.

**Topic-Comment and CLLD**

One particular construction that is common in Arabic and exhibits properties
similar to those associated with equatives involving PRON is Left-Dislocation
(LD, henceforth). In traditional grammars of Standard Arabic, an LD structure
is characterised as a topic-comment construction (Farghal 1986, Plunkett 1993,
Fassi-Fehri 1993). The main discourse function of LD in Arabic is to topicalise
the dislocated phrase. In such constructions, the topic is a definite DP which
introduces the comment, itself a predication, and delimits its scope or applica-
tion. A common type of topic-comment construction is one which includes a
resumptive pronoun in the comment part whose antecedent is the topic, as can
be illustrated below.

(36) a. Majdi, šoft-o mbaarih
    Majdi, saw.1s-him yesterday
    ‘Majdi, I saw him yesterday’

b. hal-walad, bte¶ref-o ¿ante?
   this-boy, know.2ms-him you.2ms
   ‘This boy, do you know him?’

41
In the sentences above, the initial DP *Majdi* is functionally a topic, known in the literature on Classical Arabic as *al-mubtada* (roughly, topic), related to the pronoun appearing in the comment part of the sentence, known as *al-xabar* (roughly, comment). Topic-comment clauses like (36) are typically derived from the original subject-predicate clauses shown in (37).

(37) a. ˇ soft Majdi mbaarih saw.1s Majdi yesterday ‘I saw Majdi yesterday’

b. bte‟ref hal-walad ?ante? know.2ms this-boy you ‘Do you know this boy?’

The sentences in (36) are the result of dislocating the DPs, *Majdi* in (37a) and *hal-walad* ‘this boy’ in (37b), from the internal argument position into the left periphery. Meanwhile, the original argument position is occupied by a resumptive pronoun that is coreferential with the displaced DP. Topic-comment clauses, thus, differ from ordinary subject-predicate clauses in that the comment itself has its own subject.

Within the recent syntactic theory, topic-comment constructions are analysed as LD structures (Fassi-Fehri 1993, Plunkett 1993 and Soltan 2007). The type of LD found in the literature on Arabic is widely known as Clitic-Left-Dislocation (CLLD). CLLD displays properties similar to those associated with equatives, especially, the definiteness of the left-most DP and its relationship with a pronominal element occupying a thematic position inside the sentence (Aoun and Benmamoun 1998, Ouhallal and Shlonsky 2002 and Aoun, Benmamoun and Choueiri 2010). Consider the following paradigm:

---

16See also Chomsky 1977, Cinque 1990 and Rizzi 1997.

17LD as discussed in Cinque 1990 differs from CLLD in Arabic. The former, but not the latter, is a root clause phenomenon and only one element can be dislocated in a sentence.
(38)  a. naadia šee-f-a saami mbaarih  (Lebanese)
Nadia saw.3ms-her Sami yesterday’
‘Nadia, Sami saw her yesterday’

b. fakkart ?enno naadia šee-f-a saami mbaarih
thought.1s that Nadia saw.3ms-her Sami yesterday’
‘I thought that Nadia, Sami saw her yesterday’

(39)  a. l-ktaab gareet-oh l-yom  (Jordanian)
the-book read.1s-it the-day
‘The book, I read it today’

b. sme‘na ?enno l-ktaab gareet-oh l-yom
heard.1pl that the-book read.2ms-it the-day
‘We heard that the book, you read it today’

The (a) examples in (38) and (39) illustrate cases of CLLD in root clauses. The DPs naadia and l-ktaab ‘the book’ are placed in the left-most position and are related to a corresponding object pronoun that is attached to the verb.

The (b) examples show that the left-dislocated DPs follow the complementiser ?enno ‘that’ in embedded contexts. This indicates that the structural position of LD’ed DPs must be below C, as will be shown shortly.

Furthermore, notice that in the CLLD constructions above, (38) and (39), the DPs are definite. Indefinite DPs cannot be used in this type of construction, regardless of whether they are specific or bare indefinites. Consider the examples shown below.

(40)  a. *?a Süide ?allaf-a Šomar  (Lebanese)
poem wrote.3ms-it Omar
‘A poem, Omar wrote’

b. *?a Süide zareef-e ?allaf-a Šomar
a-poem nice wrote.3ms-it Omar
‘A nice poem, Omar wrote’
Besides the ban on indefinite expressions, there is also a ban on quantifica-
tional expressions, i.e., QPs, in CLLD in Arabic. This can be illustrated in the
following examples taken from Aoun, Benmamoun and Choueiri (2010: 97).

(42)  a. *ma hada šeefit-o zeina ūam bizaqt-bir (Lebanese)
      no one saw.3fs-him Zeina Asp. cheat.3ms
      ‘No one, Zeina saw him cheating’

      b. *maši zeebit-o lina min d-dikkeen
      nothing brought.3fs-it Lina from the-store
      ‘Nothing, Lina brought from the store’

CLLD constructions raise a number of interesting questions to do with the
position of the left-dislocated DP, and the relationship between the resumptive
pronominal inside the sentence and the left-dislocated DP, i.e., whether it is a
movement or a base-generation relationship. It is not my goal, however, to dwell
on this issue here. For the current purposes, it is sufficient to note that as far as
the position of the left-dislocated DP is concerned, various studies have argued
that it occupies an A’-position(s) below C (the reader is referred to Plunkett
1993, Aoun and Benmamoun 1998, Soltan 2007 and Aoun, Benmamoun and
Choueiri 2010). As concerns the second question and whether the relationship
between the resumptive pronominal inside the sentence and the left-dislocated
DP is a movement relationship or whether it is a base-generation relationship,
Aoun and Benmamoun 1998, for instance, conclude, based on a number of diag-
nostic tests such as reconstruction, island effects and intervention effects, that
CLLD may or may not involve movement, depending on the context. Whereas Plunkett 1993 and Soltan argue for a base-generation analysis based on the simple fact that a left-dislocated DP may occur far away from its corresponding resumptive element.

To recap, this section has presented a common type of construction in Arabic known as the topic-comment construction. The topic-comment construction involves the presence of a definite DP in the left-most position followed by a complete predication containing a resumptive pronoun whose antecedent is the topic DP. I have shown that traditional topic-comment clauses are instances of CLLD of the type discussed in Aoun and Benmamoun 1998. I have concentrated on some of the properties that CLLD has in common with equatives, especially, the definiteness of the LD’ed DP and its relation with the pronoun inside the clause. However, in all the cases so far, we have been dealing with the left-dislocation of DPs that correspond to direct objects. An interesting question that could be asked here is whether this phenomenon extends to subjects, i.e., whether subject-dislocation is feasible in Arabic. The pros and cons of this query are discussed in the next section.

1.5.4 SV in Standard Arabic as Left-Dislocation

Among the informational purposes that LD serves is focus, achieved by placing DP in the left-most position. This is the case of CLLD in Arabic dialects presented in the previous section. The initial DP would have been inside the sentence before left-dislocation, given that the word order in these dialects is SVO (Fassi-Fehri 1993, Mohammad 2000, Ouhalla and Shlonsky 2002 and Aoun, Benmamoun and Choueiri 2010). Reasoned as such, it naturally follows that an element which usually appears in the left-most position is already focused. As far as the Arabic dialects are concerned, what this means is that left-dislocation of subjects may be all over the place. Several facts point in this direction. In addition to the fact that the subject is the left-most element in the SV(O) order,
the fact that it is subject to the indefiniteness restriction, and the fact that the subject pronoun which would appear as a result of left dislocation may be a null pro, all constitute crucial evidence that SVO may be a concealed LD construction. Compare this with the dislocation of objects, the latter is not ambiguous since the object is placed in a position other than its internal position with that position being filled by an overtly realised clitic.

Putting aside the Arabic dialects and the question of whether subject LD is indeed the structure represented by the surface SV word order, I want to shift focus to another interesting, and less ambiguous, case for the current discussion, namely, subject LD in Standard Arabic. SA differs in its surface word order from the spoken dialects in that it is predominantly a VSO language and a null-subject language. This might be shown in the following examples.

(43) a. qaraʔa ?al-ʔawlaad-u ?al-dars-a
   read.3ms the-boys.Nom the-lesson-Acc
   ‘The boys read the lesson’

   b. qaraʔ-uu ?al-dars-a
      read.3mpl the-lesson-Acc
      ‘They read the lesson’

   The sentence in (43a) exemplifies the unmarked order in SA, namely, VSO. (43b) contains no overt subject, i.e., the subject is null, though agreement manifests on the verb. As a VSO language, left-dislocation of a definite subject should be possible, on a par with object left-dislocation, since both the subject and the object are not in the left-most position. Now consider the sentences in (44) given in Plunkett (1993: 241).

(44) a. al-ʔutlab-u u-hib-u-hum
    the-students.Nom Is-like-them
    ‘The students, I like them’

---

b. al-ṭullab-u ya-drus-uuna
   the-students.Nom study-3mp
   ‘The students, (they) are studying’

Plunkett 1993 analyses the SVO order in SA as a left-dislocation structure. She argues that the initial DP in both (44a) and (44b) is a topic. In (44a), the initial DP is coreferential with an overt resumptive pronoun in object position. Whereas, in (44b), the initial DP is coreferential with a small pro subject. Because Arabic is a null-subject language, the pronoun in the subject position is non-overt.

Plunkett’s LD analysis for SV in Arabic is further supported by a more recent analysis advanced in Soltan 2007. Soltan also treats all SV orders in SA as left-dislocation structures whereby the preverbal DP is said to be a discourse topic against which the event is presented. To illustrate, consider the following examples (Soltan 2007: 34-35).

(45) a. ?al-ʔawlaad-u qaraʔ-uu ?al-dars-a
       the-boys-Nom read-3mpl the-lesson-Acc
       ‘The boys read the lesson’

b. ?al-fatayaat-u qaraʔ-na ?al-dars-a
   the-girls-Nom read-3fpl the-lesson-Acc
   ‘The girls read the lesson’

According to Soltan 2007, the preverbal DP in the sentences in (45) is not a genuine subject. The subject in this type of sentence is actually a null resumptive pronoun that occupies the genuine subject position inside the sentence. To put this in concrete terms, consider the structure Soltan assigns for such sentences in (46) below.
The boys read the lesson

Under this analysis, the left-dislocated DP is base-generated in Spec,TP, while the actual subject is a null pro merged in Spec,v*P. This is bolstered by cases where null pro surfaces obligatorily in SA, namely, where the left-dislocated DP is separated by an island from its corresponding pronoun in subject position. To illustrate consider the sentences in (47).

(47) a. haDara Zayd-un wa ʕaliyy-un  
came.3ms Zayd-nom and Ali-Nom  
‘Zayd and Ali came.’

b. Zayd-un haDara huwwa wa ʕaliyy-un  
Zayd-Nom came.3ms he and Ali-Nom  
‘Zayd, he and Ali came.

c. *Zayd-un haDara wa ʕaliyy-un  
Zayd-Nom came.3ms and Ali-Nom  
‘Zayd, he and Ali came.

48
In (47a), the postverbal DP subject occurs in a coordinate structure island. The left-dislocation of this DP entails that it has to be overtly resumed by the pronoun *huwwa* within the coordinate phrase as shown in (47b). Otherwise, the sentence would be ruled out, as shown in (47c).

**Strong Pronouns as Resumptives**

The view that third person strong pronouns can function as resumptives is advanced in Aoun, Choueiri and Hornstein 2001 (ACH henceforth). According to ACH, a strong pronoun can be used as a resumptive pronoun in all contexts where the antecedent to which the pronoun is related is non-quantificational. The following are instances of subject left-dislocation involving overt resumption:

(48) a. ha-l-muttahame ḥraforto ?onno hiyye nhabsit this-the-suspect.sf know.2pl that she imprisoned.sf
    (Lebanese)

    ‘This suspect, you know that she was imprisoned.’

b. ha-l-muttahame ṣafto l-mahamme yalli byaṭrif ?onno this-the-suspect.sf saw.2pl the-attorney that know.ms that
    hiyye harabit she ran.away.fs
    (Lebanese)

    ‘This suspect, you saw the attorney that knows that she ran away.’

In the cases above, the left-dislocated DP is related to a strong resumptive pronoun inside an embedded clause in (48a) and inside an island in (48b).

Similar cases are also found in Jordanian.

19 One of the assumptions made in Soltan 2007 is that subject pronouns in null subject languages start the derivation as the null element *pro*. To maintain this assumption, Soltan argues that the reason why the resumptive obligatorily surface in conjoined subjects may be attributed to an interface condition to do with phonological parallelism of coordinate structure, i.e., that both conjuncts must have phonetic content. Lexicalisation of a *pro* conjunct is, thus, said to follow from this condition on the interface.

20 For ACH 2001, resumptive pronouns are characterised as DPs that are locally A’-bound (see Shlonsky 1992, Shlonsky 1997 and Ouhalla 2001). ACH also point out that the HSR (which prohibits (strong) pronouns from being related to an antecedent occurring in within the same CP) is not operative in LA since strong subject pronouns do appear in the highest subject position.
To sum up so far, I have explored two types of LD constructions in Arabic which can be subsumed under traditional topic-comment clauses. The first type is the well known CLLD in the spoken dialects which applies to constructions from which an object DP is left-dislocated. The other type is subject dislocation in SA that applies to constructions from which a subject is left-dislocated. In both types of construction, a pronoun in an argument position is said to be in a coreference relationship with the LD’ed DP. Whether or not the pronoun is overtly realised has been shown to be subject to language-specific rules. For instance, whereas Arabic object pronouns must be realised overtly in CLLD contexts, this is not so in the subject-dislocation contexts.

Having established both the fact that LD constructions are common and the fact that subjects can be dislocated in Arabic, the question that I want to explore next is whether a left-dislocation analysis captures the facts from equative constructions observed earlier. In particular, given that the set of properties associated with each type of construction are similar, I will put forward the hypothesis that equatives involving PRON are best treated as instances of subject-dislocation on a par with subject-dislocation in verbal predications.

1.5.5 Left-Dislocated Equatives

The work on Arabic left-dislocation has been essentially concerned with left dislocation in verbal predications. The one particular type of left-dislocation that is most common is CLLD (see Aoun and Benmamoun 1998) whereby the dislocated DP corresponds to a pronominal element in object position. We have also
seen cases where a subject is left-dislocated in SA, which is, formally, a VSO language. The left-dislocated subject is said to be related to a strong pronoun in the subject position inside the clause. Overtness of the resumptive pronoun in the subject position, however, is overridden by the null-subject parameter that is operative in Arabic.

As concerns non-verbal predications, particularly, equative constructions involving PRON, and the claim that they are instances of subject dislocation, the facts obtained thus far strongly suggest that they be treated much in the same way as CLLD and subject left-dislocation in SA. First, the set of properties associated with equatives is, more or less, the same as those that generally characterise topic-comment constructions. The fact that only definite arguments can occur in the initial position of the clause and the fact that equative constructions include, in addition to the initial DP, a predication that consists of a subject and predicate are characteristic of left-dislocation structures. Like topic-comment clauses, equatives of the form DP PRON DP derive from basic subject-predicate equatives of the form DP DP. To put this in concrete terms, consider the sentences in (50) which are derived from the original two DP equative constructions in (51).

(50) a. Saahib š-šareka hu l-mudiir t-tanfeethi (Jordanian)
owner the-company he the-director the-executive
‘The company owner, he is the executive director.’

b. l-mudiir t-tanfeethi hu Saahib š-š-areka
the-director the-executive he owner the-company
‘The executive director, he is the company owner.’

(51) a. Saahib š-šareka, l-mudiir t-tanfeethi
owner the-company, the-director the-executive
‘The company owner is the executive director.’

21Lexicalisation of pro subject would then follow from an interface condition that requires the pronoun to have phonetic content be it for emphasis, disambiguation or contrastive focus.
b. l-mudīr t-tanfeethī, Saahib ְֶָָ‍—areka
the-director the-executive, owner the-company
‘The executive director is the company owner.’

As can be seen from these examples, the initial DP is a definite noun phrase followed by a string constituting a complete predication which contains a resumptive pronoun that is coreferential with the left-dislocated DP. The initial DP is interpreted as coreferential with this pronoun in subject position in the same way a left-dislocated DP is interpreted as coreferential with a resumptive pronoun within the thematic DP.

The LD analysis of equatives involving PRON gains extra support on the basis of observations from similar dialects, such as Syrian Arabic. Cowell 1964 points out that equatives involving PRON of the type illustrated in (52) are derived from PRON-less equatives of the type illustrated in (53).

(52) a. hal-bent, hiyye l-ʔahla
that-girl, she the-prettiest
‘That girl, she is the prettiest’

b. l-bent, hiyye l-ʔahsan
the-girl, she the-best
‘The girl, she is the best’

(53) a. hal-bent, l-ʔahla
that-girl, the-prettiest
‘That girl is the prettiest’

b. l-bent, l-ʔahsan
the-girl, the-best
‘The girl is the best’

Cowell treats (52) as topical clauses that are related to the ones in (53) by what he calls extraposition. That is, the DPs hal-bent ‘that girl’ and l-bent

Extraposition involves movement of the DP leftwards. However, as will be shown shortly,
‘the girl’, in (53a) and (53b), respectively, are made topics and replaced by a corresponding pronoun in (52). Cowell notes that extraposition in this case of equational predication, is commonly used not so much to emphasise the extraposed subject, but simply to identify the predicate as such. For instance, the predication in (53b) might in some circumstances be confused with the noun phrases l-bent l-ʔas ‘the best girl’, therefore, the predication tends to be replaced by a topical sentence when no special emphasis is intended as in (52b).

Further evidence that equatives are dislocated structures can be obtained from contexts involving quantificational elements. For instance, one of the classic tests for CLLD is whether a QP can occur in this type of construction. The data show that QPs are inadmissible in equatives. Consider the following cases.

\[(54)\]
\[
a. \quad \text{*wala hada huwwa l-mudarris} \quad \quad \text{(Jordanian)}
\]
\[
\text{no one he the-teacher}
\]
\['No one is the teacher’\]
\[
b. \quad \text{*kull wahad huwwa l-mudarris}
\]
\[
\text{every one he the-teacher}
\]
\['Every one is the teacher’\]

The ungrammaticality of the examples above is due to the fact that neither can the QPs wala hada ‘no one’ in (54a) and kull wahad ‘everyone’ in (54b) be related to the subject pronoun huwwa ‘he’ inside the clause nor can this pronoun be related to an indefinite. This is because the QP is non-referential and cannot be an antecedent for the subject pronoun, which, in turn, needs a strongly referential antecedent (Fassi-Fehri 1993 and Plunkett 1993). This is exactly what we have seen in the CLLD cases earlier, which I repeat below for ease of reference.

\[(55)\]
\[
a. \quad \text{*ma hada ʔeefit-o zeina ʔam biza?bir} \quad \quad \text{(Lebanese)}
\]
\[
\text{no one saw.3fs-him Zeina Asp. cheat.3ms}
\]
\['No one, Zeina saw him cheating’\]

equatives involving PRON are base-generated constructions.
b. *maʃi ʒeebi₃-o lina min d-dikkeen
nothing brought.3fs-it Lina from-the-store
‘Nothing, Lina brought from the store’

The facts from quantification above constitute empirical evidence that equatives involving PRON are left-dislocated structures. It is clear that the left-most position can only be occupied by a definite DP in order to provide a reference for the pronoun that occurs in the subject position inside the clause.

Although a QP cannot occur in left-dislocation structures, it can occur in basic equatives which have been analysed here as subject-predicate structures. Consider the following sentences.

(56)  
\begin{align*}
\text{a. } & \text{wala hada l-mudarris} \\
& \text{no one the-teacher} \\
& \text{‘No one is the teacher’}
\end{align*}

\begin{align*}
\text{b. } & \text{kull wahad l-mudarris} \\
& \text{every one the-teacher} \\
& \text{‘Everyone is the teacher’}
\end{align*}

Recall that an indefinite expression cannot occur in Spec,TP in such constructions. The examples above appear to contradict this view. My understanding is that cases which involve QPs like \textit{wala hada} ‘no one’ and \textit{kull wahad} ‘every one’ are instances of focus, hence, occupy a Focus position in the structure, where no restriction is imposed on the definiteness of the element that occupies its Spec position.

Before bringing this section to a close, there remains one issue to be addressed. So far, I have presented several arguments to support the LD analysis of equatives that are present-tense, but what about equatives other than present-tense, in which a verbal copula is used? The LD analysis advanced here
predicts that in past and future tense LD variant of equative copular sentences, both PRON and the verbal copula should be able to surface. This prediction is indeed borne out, as can be seen in the following sentences.

(57) a. ʿali huwwa kaan l-mudarris
    Ali  he   3ms.was  the-teacher
    ‘Ali, he was the teacher’

b.  Majdi huwwa rah ykuun l-mudarris
    Majdi he fut. 3ms.BE the-teacher
    ‘Majdi, he will be the teacher’

c.  Nadia hiyya ha-tkuun l-mas?uula
    Nadia she fut.-3fs.BE the-responsible
    ‘Nadia, she will be the one responsible’

Under the current analysis, the DPs ʿali, Majdi and Nadia, (57a), (57b) and (57c), respectively, are left-dislocated topics. As such, these DPs represent the topic part of a topic-comment construction and are related to a pronoun that functions as subject of the comment part in this type of construction. Notice also that the verbal copula in the sentences above shows full agreement with the subject pronoun, which also displays full agreement with the antecedent topic.

To recap, the aim of this section was to establish a link between the LD constructions in verbal predication discussed in the previous section and equatives involving PRON which represent the other type of predication common in Arabic, i.e., verbless predication. The facts from both types of constructions suggest that they belong to the general topic-comment construction. In particular, I have explored the hypothesis that equatives involving PRON are instances of subject LD which derive from simple subject-predicate clauses. Properties such as the definiteness of the initial DP, the restriction on QPs and the presence of a subject pronoun in a separate predication string constitute strong evidence
that equatives are a species of the LD’ed construction in Arabic\textsuperscript{23}. Next, I turn to the question about the structure of LD’ed equatives.

1.5.6 The structure of LD’ed Equatives

An appropriate structure that correctly represents Left-Dislocated equatives in Arabic must account for the following facts: the fact that although such constructions are verbless, they are tensed, i.e., they are present-tense configurations; the fact that there is an internal subject-predicate relationship; and the fact that the subject of this predication is coreferential with an antecedent in the left-periphery. These facts are best captured by PredP framework originally suggested in Bowers 1993, 2001 and adopted in Adger and Ramchand 2003 and Ouhalla 2005\textsuperscript{24}. PredP is a predicational core headed by Pred which mediates subject-predicate relationships. As such, the structure I propose for LD’ed equatives in Arabic is roughly represented in (58) below.

\textsuperscript{23}The left-dislocation analysis for equative constructions is argued for independently in Edwards 2006 for Egyptian Arabic. Edwards treats equatives involving PRON as cases of left-dislocation and assigns a CP structure in which the initial DP is located in Spec,CP. The resumptive subject pronoun, i.e., PRON, is said to be initially base-generated in Spec,*vP, then moves lower to the position of the (functional) head v*\textsuperscript{2}. Such a step is based on an assumption originally made in Simpson and Wu 2003 according to which the subject pronoun undergoes re-analysis from a Spec-element into a head-element. Such analysis is argued to account for the properties that PRON displays in equative constructions in Egyptian. However, though Edward’s analysis has the advantage of treating equatives as left-dislocation structures, it suffers a from number of problems, of which I highlight three. First, the assumption that left-dislocated DP occupies Spec,CP cannot be maintained. As we have seen earlier, left-dislocated DPs follow complementisers in Arabic; they cannot occur before C. Secondly, the assumption that the subject pronoun originates in Spec,*vP is not empirically supported since such structures are, simply, verbless constructions. The last problem is that Tense, hence a TP projection, is not present in the structure. Both the fact that equative constructions are present-tense constructions and the fact that T has an EPP property that requires a definite DP in its Spec cannot be explained.

Cowell 1964, on the other hand, treats equatives involving PRON as instances of extraposition, whereby the initial DP undergoes movement to its surface position, rather than being base-generated in that position. An analysis based on extraposition faces the question of movement that will be dealt with in the subsequent sections. It remains to be said that Cowell’s work on Syrian is not an analytical work in as much as it is a reference for the grammar of this language. Meanwhile, Phunkett’s 1993 and Soltan’s 2007 analyses are essentially concerned with left-dislocation in SV constructions in SA, though, with implications for the analysis of equative constructions involving PRON.

\textsuperscript{24}Unlike Bowers 1993 and Adger and Ramchand 2003, I follow Heggie 1988 in taking a definite DP to be predicative, i.e., type (e, t), in Arabic.
The structure in (58) incorporates a number of assumptions with regards to the position of the LD’ed DP, the subject pronoun, and the nature of the relationship between these two elements. These are discussed in turn below.

**Position of the LD’ed DP**

Following Plunkett 1993, Aoun and Benmamoun 1998, Soltan 2007 and Aoun, Benmamoun and Choueiri 2010, I take the LD’ed DP to be base-generated in Spec,TP. Evidence that the LD’ed DP is in Spec,TP comes from contexts where an LD’ed DP follows the complementiser in Arabic, as we have seen in CLLD cases earlier. The same applies to LD’ed equatives in embedded contexts where the LD’ed DP also appears after the complementiser. This is shown in the following sentences.

(59) a. bayref ?anno majdi huu(luwwa) l-muhandis (Jordanian)
    know.1s that Majdi he the-engineer
    ‘I know that Majdi, he is the engineer’

b. *bayref majdi ?anno huu(luwwa) l-muhandis
    know.1s Majdi that he the-engineer
    ‘I know that Majdi, he is the engineer’
The cases above show that the LD’ed DP must be in a position below C in the structure. Under the proposed structure in (58), this position is Spec,TP. Notice that (59b) and (60b) are ruled out because the LD’ed DP occurs before the complementiser *?anno ‘that’.

The second assumption incorporated within the current analysis is that Spec,TP is an A’-position. This is because left-dislocation structures are generally A’-dependencies (Cinque 1990, Rizzi 1997, Aoun and Benmamoun 1998, Shlonsky and Ouhalla 2002 and Aoun, Benmamoun and Choueiri 2010). Evidence that Spec,TP is an A’-position in LD’ed structures in Arabic comes from extraction facts in SA. Consider, for instance, extraction across an LD’ed DP vs. extraction across a postverbal subject as shown in the examples below.

The ungrammaticality of (61a) follows from the assumption that the LD’ed DP, i.e., *Zayd-un, is in an A’-position that, therefore, intercepts movement of the wh-phrase to Spec,CP. By contrast, extraction across a postverbal subject is possible since the postverbal subject is not in an A’-position that therefore blocks movement of the wh-phrase across to the front of the clause.

\(^{25}\)See Soltan’s 2007 analysis of SA for a full discussion.
Given the assumption that the LD’ed DP is base-generated in its surface position, namely, Spec,TP, it follows that the resumptive pronoun is also base-generated in its surface position, i.e., Spec,PredP\(^{26}\). Syntactically, the LD’ed DP in Spec,TP binds the subject pronoun in PredP as shown in the structure (58). This, in turn, warrants the coreference relationship between the two.

Before moving into the semantics of the LD’ed DP, two issues deserve discussion, one concerns the base-generation analysis, as opposed to a movement analysis, and the other concerns agreement. Staring with the first issue, the base-generation analysis adopted here is more desirable than a movement analysis, for several reasons. First, for a movement analysis, we have to stipulate that indefinite NPs cannot undergo movement, while definite DPs can. In other words, movement has to distinguish between definite and indefinite DPs, so that it applies to one to the exclusion of the other; a costly move. Secondly, in a movement configuration, it has to be the case that the resumptive pronoun is the spell-out of the trace (or copy) left by the LD’ed DP. This is a problematic issue for both the view that resumption signals absence of movement and the view that traces or copies left by movement and resumptive pronouns are different\(^{27}\). Lastly, as observed in Plunkett 1993, Aoun and Benmamoun 1998, and Aoun, Benmamoun and Choueiri 2010, left-dislocated DPs can occur far away from resumptive pronouns in CLLD constructions, hence, a movement analysis is unlikely.

The second issue that needs to be addressed is agreement in equatives involving PRON. The task of identifying the type of agreement in this type of construction is not an easy one. One may ask a number of questions in this regard: is agreement between the subject pronoun and its predicate or is it between the subject pronoun and its antecedent? Why is agreement limited in some cases, but not in others? Is it grammatical agreement like agreement in verbal predication? Or, is agreement not agreement after all?

\(^{26}\)This also allows the subject pronoun to receive its \(\theta\)-role as well as Nominative case.

\(^{27}\)See Shlonsky 1992, 2002; ACH 2001; Aoun and Li 2003; and Soltan 2007.
A number of views have been entertained in the literature which I have presented earlier. However, none of these views provides convincing answers to the questions raised above. I believe this is best explained by the facts themselves which show that this type of agreement is inconsistent when a resumptive subject pronoun stands between an antecedent and a predicate that differ in number and gender.

To sum up, this section has been dealing with various issues concerning the structure of LD’ed equatives. I have proposed that LD’ed equatives are TP projections in which the LD’ed DP is base-generated in Spec,TP, an A’-position. As concerns PRON, I have shown that it is also base-generated Spec,PredP as subject of the lower predication whose complement is the second DP. Both the LD’ed DP and PRON were said to be in a coreference relationship warranted by syntactic binding.

1.5.7 Semantics Issues

The Indefiniteness Restriction and the EPP

One other issue that remains to be addressed concerns the indefiniteness restriction that prohibits indefinite DPs from occurring in the clause initial position in left-dislocated structures. I would like to argue, building on Chomsky (2001a), that the semantics of left-dislocation in Arabic can be accounted for in the syntax. More specifically, the fact that only definite DP can occur in Spec, TP can be explained by positing a discourse feature on T in Arabic equatives.

Chomsky (2001a) suggests that phase edges are where some of the surface semantics effects, such as topic and focus, arise. This seems to be empirically
supported by the fact that left dislocation of DPs in languages such as English and Italian (Cinque 1990 and Rizzi 1997) is associated with the clause periphery, i.e., the edge of the CP phase. Recall, however, that LD’ed DPs in Arabic are in Spec,TP. If we extend Chomsky’s notion so that semantic effects arise in the peripheral position of functional heads like T, the dislocation facts from Arabic follow. The fact that only definite DPs can occur in left-dislocation structures is captured on the grounds that T has a discourse or topic feature which can only be satisfied by merging a definite DP in its Spec position. Given such a view, the indefiniteness restriction follows from the semantic effects of the position that the left-dislocated DP occupies. Assuming that T can have a discourse feature and that LD-ed DPs are merged in the Spec of T in LD’ed equatives, it follows that a structure with an indefinite DP in Spec,T position conflicts with the inherent properties of this position, hence uninterpretable at the semantic interface29.

**Interpretation**

The analysis advanced here gives rise to a question about interpretive differences between simple predicational sentences of the form [XP XP] and left-dislocated cases of the form [XP PRON XP]. Recall that structurally, the latter involves PredP in which thematic roles are licensed, but not the former. However, the difference does not seem to be of predication, it is rather to do with chain formation and how a chain is interpreted in this structure. More specifically, in constructions that have the format [XP XP] there is a single member chain. Whereas in constructions that have the format [XP PRON XP], the interpretation chain is a two-member chain that connects the LD’ed DP with the subject pronoun, i.e., PRON.

29It is noteworthy that Arabic allows more than one left-dislocated DP in verbal predications. It follows that there is more than one position LD’ed DPs may occupy. Two approaches come to mind, one approach is outlined in Rizzi’s 1997 fine structure of the left periphery and the other is outlined in Uriagereka 2006 whereby the clause periphery is a zone in which certain semantic effects arise. Since this phenomenon is found only in verbal predications, I will not be discussing it here. The reader is referred to Plunkett 1993 and Aoun and Benmamoun 1998.
1.5.8 Consequences of the LD Analysis

The LD-analysis of equatives advanced here has several consequences. The first consequence of treating equatives as LD structures is that the problems associated with the so-called *copula pronoun* can be naturally explained on the basis that it is, literally, a subject pronoun and functions as such. Secondly, like basic equatives, simple copular sentences may occur with PRON. In particular, in cases where the subject is a definite DP, it can be left-dislocated. The data in the examples below illustrates this point.

(62)  
\[
\begin{align*}
\text{(a) } & \quad \text{l-bent helwa.} & \quad \text{(Palestinian)} \\
& \quad \text{the-girl pretty} \\
& \quad \text{‘The girl is pretty’}
\end{align*}
\]

\[
\begin{align*}
\text{(b) } & \quad \text{l-bent hi(ye) helwa} \\
& \quad \text{the-girl she pretty} \\
& \quad \text{‘The girl, she is pretty’}
\end{align*}
\]

(63)  
\[
\begin{align*}
\text{(a) } & \quad \text{Majdi mas\(\text{?}"}\ uul} & \quad \text{(Jordanian)} \\
& \quad \text{Majdi responsible} \\
& \quad \text{‘Majdi is responsible’}
\end{align*}
\]

\[
\begin{align*}
\text{(b) } & \quad \text{Majdi hu(wa) mas\(\text{?}"}\ uul} \\
& \quad \text{Majdi he responsible} \\
& \quad \text{‘Majdi, he is responsible’}
\end{align*}
\]

While the (a) examples are simple subject-predicate sentences, as we have seen earlier, the (b) examples are instances of left-dislocation. Now consider the structure in (64) below for a sentence like (62b)\(^{30}\).

\(^{30}\)It it noteworthy that although LD’ed simple copular sentences are available in Arabic languages, they are not so common. This is, perhaps, because they are less susceptible to ambiguity compared to equatives involving two lexical DPs. In terms of use, thus, simple copular sentences are the opposite of equatives. Equatives involving two lexical DPs are more likely to be ambiguous than simple copular sentences, hence the use of dislocation is more common.
The third consequence of this analysis, which will be crucial for the analysis of wh-dependencies in subsequent chapters, concerns the putative optionality of PRON in copular wh-constructions such as those shown below:\(^{31}\).

(65)  

\begin{align*}
\text{a. miin (huwwa) l-muhandis} & \quad \text{who (he) the-engineer} \\
\text{‘Who is the engineer?’} &
\end{align*}

\begin{align*}
\text{b. miin (huwwa) yalli šuft-o b-l-maT\text{ʕam}} & \quad \text{who (he) that saw.2pl-him in-the-restaurant} \\
\text{‘Who is is that you saw in the restaurant?’} &
\end{align*}

Under the current analysis, optionality of PRON is only apparent. In other words, it is not the case that one structure may or may not involve PRON. Rather, a sentence with PRON differs structurally from a sentence without in that the former is a dislocation structure whereas the latter is a subject-predicate structure.

\(^{31}\)Such a view is advocated in Shlonsky 2002 and Aoun, Benmamoun and Choueiri 2010.
The fourth consequence of the current analysis, especially the view that T has a feature with semantic content, in addition to its syntactic role, is that any of the two DPs in an equative construction, can be dislocated since both DPs are definite. This straightforwardly accounts for one property of equative constructions, i.e., reversibility. That is, the DPs of Arabic equatives may undergo inversion with no grammatical consequences or alteration to the propositional content of the sentence (Ouhalla 1999 and Edwards 2006). Consider the following pair of sentences and their corresponding structures.

(66)  
\[
\begin{align*}
\text{a. } & \quad \text{\'Ali hu(wwa) l-mudarris} \\
& \quad \text{Ali he the-teacher} \\
& \quad \text{‘Ali is the teacher’}
\end{align*}
\]

\[
\begin{align*}
& \quad \text{b.}
& \quad \text{TP}
& \quad \text{DP} \\
& \quad \quad \quad \text{T’}
& \quad \quad \quad \quad \text{PredP}
& \quad \quad \quad \quad \quad \text{Pred’}
& \quad \quad \quad \quad \quad \quad \text{DP l-mudarris}
& \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{the-teacher}
\end{align*}
\]

(67)  
\[
\begin{align*}
\text{a. } & \quad \text{l-mudarris hu(wwa) \text{\'Ali}} \\
& \quad \text{the-teacher he Ali} \\
& \quad \text{‘The teacher is Ali’}
\end{align*}
\]
The difference between the sentences in (66) and (67) can be said to differ in what they are about, with that difference being essentially context dependant. In (66), the DP ’ilali ‘Ali’ is focused to obtain the interpretation that ’ilali ‘Ali’, not someone else, is the teacher. In (67), it is the DP ’il-mudarris ‘the teacher’ that is focused so as to obtain the interpretation where the teacher, not, say, the engineer, is ’ilali ‘Ali’.

The current analysis, thus, lends itself to a less rigid approach to the order and position of DPs in equative constructions. Namely, an approach that takes the positions for both the subject and the predicate not to be structurally fixed, but are, rather, contextually determined.

1.6 Summary and Conclusion

This chapter has provided a unified syntactic analysis for Arabic copular constructions including equative constructions. More specifically, I have shown that copular constructions have one basic schematic structure, that is DP XP. The latter includes simple predicational sentences where XP may either be an indef-
inite DP, an AP or a PP. It also includes equative copular sentences where XP
is a definite DP. I have argued that the two types of copular construction are
subject-predicate clauses and have one and the same structure roughly repre-
sented in (68) below.

(68)

\[
\begin{array}{c}
TP \\
\downarrow \\
DP \\
\downarrow \\
T' \\
\downarrow \\
Subject \\
\downarrow \\
T \\
\downarrow \\
XP \\
\downarrow \\
[+Present, +D] \\
Predicate
\end{array}
\]

The [DP XP] construction was also argued to be the basic construction from
which the so-called *equative sentences* of the form [DP PRON DP] (Eid 1983,
1991, 1992; Farghal 1986; Ouhalla 1999; Edwards 2006) are derived. That is,
equative sentences of the form [DP PRON DP] have been analysed here as left-
dislocated structures that derive from the basic subject-predicate construction
shown in (68). Under this analysis, the first DP is treated as a left-dislocated
topic that is coreferential with PRON. Meanwhile, PRON is treated as subject
of the lower predication which includes the second DP as predicate. The struc-
ture assigned to the LD’ed copular construction including LD’ed equatives is
the one diagrammed in (69).
The analysis of copular constructions presented in this chapter started with two empirical questions: the question of whether PRON-less equatives are available in the grammar of Arabic like simple copular constructions, and the question of whether simple copular sentences may admit PRON on a par with equatives involving PRON. Concerning the first question, the facts reported from several Arabic dialects show that PRON-less equatives are common practice across the dialects. This finding gave rise to two other questions regarding the structure of PRON-less equatives and the nature of equatives involving PRON. I have argued that PRON-less equatives are syntactically the same as predicational sentences in that they are both subject-predicate clauses, thereby, reducing the difference between the two types of copular construction to the semantics of the DP predicate.

As far as equatives involving PRON are concerned, I have presented a number of arguments to support an analysis whereby they are treated as instances of left-dislocation structures. I have shown that the properties associated with equatives involving PRON are the same as those displayed in CLLD contexts in the dialects and SV constructions in Standard Arabic, all of which have been subsumed under the general topic-comment construction. To account for their
properties, I have argued for a base-generation analysis for equatives involving 
PRON, in which both the LD’ed DP and the subject pronoun it corefers with, 
are directly merged in their surface positions, that is, Spec,TP and Spec,PredP, 
respectively. The structure I proposed is based on the PredP framework (Bow-
er 1993), whereby the head of PredP, i.e., Pred, is said to mediate the internal 
predication relationship between the DP predicate and its external subject, i.e., 
the subject pronoun, located in Spec,PredP.

The LD analysis of equatives was shown to have a number of consequences. 
The first consequence came as an answer to the second empirical question raised 
at the outset. More specifically, I have shown that LD is also possible in sim-
ple predicational sentences in cases where the subject is a definite DP. More 
importantly, the LD analysis offered a natural solution to the problems asso-
ciated with PRON that have been a thorny issue for various studies. PRON, 
under the current analysis, is simply a subject pronoun and functions as one. 
Consequently, the analysis deals with the problem of optionality of PRON in 
copular wh-constructions by showing that cases with PRON and those without 
represent two different constructions, hence, optionality is only apparent.

Overall, the analysis advanced in this chapter is merited by its unified ap-
proach to the various types of copular constructions. Unified as such, the 
analysis provided in this chapter will have valuable input for the analysis of 
wh-constructions in Arabic that embarks in the following chapter.
Chapter 2

\textit{Meno} is ‘Who is he’

2.1 The Core Problem

The core problem this chapter is concerned with is the non-uniform behaviour of argument \textit{wh}-phrases in \textit{wh-in-situ} constructions in modern Arabic dialects, especially Lebanese Arabic and Iraqi Arabic. In both languages, \textit{wh-in-situ} has been reported to be one of at least two strategies used in the formation of \textit{wh}-questions (Aoun and Choueiri 1999, Aoun and Li 2003 and Aoun, Benmamoun and Choueiri 2010 for Lebanese; and Wahba 1991, Ouhalla 1996 and Simpson 2000 for Iraqi). However, patterns of \textit{wh-in-situ} show certain variation among argument \textit{wh}-phrases across the two dialects and, in some cases, within the same dialect, as is the case in Lebanese. This is shown in the following examples\(^1\).

\begin{enumerate}
\setcounter{enumi}{0}
\item \begin{tabular}{ll}
\textbf{a.} & Mona \\
\textit{shaafat} & \textit{meno}\?
\end{tabular} \\
& Mona saw whom \textit{whom} \textquoteleft Who did Mona see?\rightquote
\begin{tabular}{ll}
\item \textit{Iraqi} \\
\textit{wh-in-situ} \textit{construction}
\end{tabular}

\item \begin{tabular}{ll}
\textbf{b.} & \textit{šeft} \\
& \textit{miin} \textit{mbeerih}\?
\end{tabular} \\
& saw.2sm who yesterday \textit{whom} \textit{who} \textquoteleft Who did you see yesterday?\rightquote
\begin{tabular}{ll}
\item \textit{Lebanese} \\
\textit{wh-in-situ} \textit{construction}
\end{tabular}
\end{enumerate}

\(^1\)Iraqi examples are taken from Wahba 1991, Ouhalla 1996; Lebanese examples are from Aoun and Choueiri 1999, Aoun and Li 2003 and Aoun, Benmamoun and Choueiri 2010, unless otherwise stated.
The data in sentences (1) and (2) present a problem for analyses that classify languages parametrically as to whether they are wh-movement languages or wh-in-situ languages (Huang 1982; Pesetsky 1987; Lasnik and Saito 1992; Cheng 1997, 2003a, 2003b, 2009; Bayer 2005; among others). Although both Iraqi and Lebanese are wh-in-situ languages, they appear to belong to different types of wh-in-situ. Whereas Lebanese wh-in-situ is of the d-linked type (Aoun and Choueiri 1999 and Aoun and Li 2003), Iraqi wh-in-situ is not (Wahba 1991, Ouhalla 1996 and Simpson 2000).

In addition to typological differences, there are asymmetries in the behaviour of in-situ wh-phrases in the two languages. For instance, both Iraqi and Lebanese wh-in-situ allow the occurrence of the wh-phrase *meno*/miin ‘who’ in argument position, as in (1a) and (1b), while there is a sharp contrast between the two languages with regards to the occurrence of the argument wh-phrase ‘what’ in situ. The Iraqi wh-phrase *sheno* ‘what’ may stay in-situ and the sentence is well-formed (2a), whereas its Lebanese counterpart, i.e., the wh-phrase *su* ‘what’ cannot, as shown by the ungrammaticality in (2b). From a theoretical perspective this is interesting because it appears that parametric variation is not sufficient if it were to capture all the differences between (types of) wh-in-situ in these languages. Micro-parametric variation is at play, too, as will be shown in this study.
The other problem imposed by the data above lies in the contrast between the two Lebanese argument wh-phrases ˇs u ‘what’ and miin ‘who’. As can be seen in (1b), the argument wh-phrase miin ‘who’ is allowed to occur in-situ while the argument wh-phrase ˇs u ‘what’, is not, hence the contrast in grammaticality between (1b) and (2b). This contrast is also interesting insofar as an analysis of the behaviour of the wh-phrase ˇs u ‘what’ gives us an insight into the mechanism that (dis)allows wh-in-situ in Lebanese Arabic and similar dialects.

Broadly speaking, there are two main approaches in the literature about wh-in-situ: a LF-movement approach and a non-movement/binding approach. Under the LF-movement approach, developed in Huang (1982) for wh-in-situ in Chinese and in Lasnik and Saito (1992) for wh-in-situ in Japanese, in-situ wh-phrases are said to undergo movement at LF into a scope position within the sentence. This approach has essentially been motivated by parallel behaviour of in-situ wh-phrases and their moved counterparts in wh-movement languages. Both in-situ and moved wh-phrases are said to attain scope by movement, the difference between the two reduces to the level at which wh-movement applies, i.e., whether it applies at LF or syntax, and to conditions that may (not) apply at each level.

However, the adequacy of LF as the single approach to wh-in-situ has been called into question by typological differences between types of wh-in-situ as well as asymmetries between in-situ and moved wh-phrases, especially, asymmetries in locality effects. As a consequence, a non-movement approach to wh-in-situ has been advanced in the works of Pesetsky (1987) and Aoun and Li (1993b), which are essentially based on Baker’s (1970) unselective binding. Unselective binding assumes the presence of one super operator in a scope position that binds multiple occurrences of (wh-) phrases with matching (wh-)features. Both the analysis of Pesetsky (1987) and that of Aoun and Li (1993) have in common the assumption that in-situ wh-phrases need not undergo movement at LF.
They differ, however, in that Pesetsky’s binding is restricted to one type of wh-phrase, i.e., d(iscourse)-linked wh-phrases such as *which*-NP phrases; whereas Aoun and Li’s analysis extends to all types of wh-phrase.

As it stands, there is still much debate as to whether wh-in-situ undergoes movement at LF, and, if so, whether this movement is subject to locality restrictions like its counterpart in overt syntax, or as to whether wh-in-situ is interpreted in terms of a bindee-binder relationship between an in-situ wh-phrase and a wh-operator in a scope position. Such controversy carries over to wh-in-situ in Arabic. Iraqi and Lebanese also present us with a non-uniform picture for wh-in-situ with (apparent) variation across the dialects and within. Altogether, such variation poses serious challenges for accounts that solely rely upon parametric differences between languages for explaining different patterns of wh-in-situ, thus, leaving characteristics of individual wh-phrases out of the picture.

In this chapter and the subsequent chapter, I defend the view that variation in Arabic wh-in-situ can be straightforwardly accounted for on the basis of micro-parametric differences between (properties of) individual wh-phrases. Based on cross-dialectal observations, I propose that argument wh-phrases have complex syntax in Arabic, which has been long concealed by standard assumptions about argument wh-phrases being a uniform class, and that the only way to unfold their complex syntax is by identifying morpho-syntactic properties of each wh-phrase. New data from a range of Arabic dialects will be presented with the aim of bringing into light existing asymmetries in the behaviour of individual argument wh-phrases. Once each wh-phrase is assigned its true syntax, it will be revealed that wh-in-situ in Iraqi is not what it appears to be. I will show that Iraqi is a wh-movement language which uses relativization as a means of d-linking which, in turn, licenses wh-in-situ. This means that Iraqi wh-in-situ is d-linked after all, like Lebanese. As for the contrast between Lebanese argu-
ment wh-phrases, it will be dealt with in the next chapter which also advances a morpho-syntactic approach to account for the asymmetry in their distribution. The analysis advanced here has implications for analyses of wh-in-situ in general and for Arabic wh-in-situ in particular.

This chapter is organised as follows. Section (2) introduces two of main approaches to wh-in-situ, the LF-movement approach and the binding approach, and provides arguments showing inadequacy of these approaches on the basis of the data presented here. Section (3) discusses the analysis outlined in Ouhalla (1996) for Iraqi wh-in-situ which takes properties of wh-phrases to be crucial for their distribution. I will show, however, that although Ouhalla’s analysis is on the right track, the assumptions that underly it are simply wrong on the basis of a wide range of data from several Arabic dialects. In sections (4) and (5), I put forward an alternative proposal and show that the variation in wh-in-situ between Iraqi and Lebanese is only apparent. Section (6) is the conclusion.

2.2 Approaches to wh-in-situ

This section provides a brief overview of two of the mainstream approaches to wh-in-situ: LF-movement and binding. In particular, I discuss the main arguments for and against the two approaches covering a range of wh-in-situ facts in different languages, including English and Chinese as well as Arabic dialects.

2.2.1 LF wh-movement

parallel behaviour of in-situ wh-phrases and moved wh-phrases. In English multiple wh-questions, for instance, LF-movement of the in-situ argument wh-phrase what in (3a) across island boundaries is said to have the same effect as its overtly moved counterpart in (3b). These are shown below.

(3) a. ??Who wonders whether John bought what?
   b. ??What do you wonder whether Jon bought?

In addition to argument wh-phrases, the parallelism between wh-in-situ and wh-movement extends also to adjunct wh-phrases. Under the LF-movement analysis, structures that contain an in-situ adjunct wh-phrase inside an island are ruled out for the same reason as their moved counterparts. This is exemplified in (4a) and (4b).

(4) a. *Who wonders whether John left why?
   b. *Why do you wonder whether John left?

Like the argument wh-phrase what in (3a), the adjunct wh-phrase why in (4a) undergoes movement at LF. LF-movement of why in (4a) is seen as a parallel to its overtly moved counterpart in (4b). Such observations have led many to maintain that in-situ wh-phrases do undergo movement at LF. The difference between movement of argument wh-phrases in sentences like (3) and movement of adjunct wh-phrases in sentences like (4) has been attributed to the nature of conditions that restrict movement of argument wh-phrases and adjunct wh-phrases (Huang 1982). That is, movement of the adjunct wh-phrase in (4) induces an ECP violation. ECP violations are known to be of a stronger nature than the Subjacency violation induced by movement of the argument wh-phrase what in (3).

---

2Examples are from Lasnik and Saito 1992.
3ibid.
4ECP refers to the Empty category Principle (Chomsky 1981) which requires empty categories to be properly governed via either lexical government, antecedent government, or both.
5Subjacency (Chomsky 1973, 1986) requires that no more than one cyclic node be crossed at a time.
Beside parallelism, there exist certain asymmetries between wh-in-situ and wh-movement. In the context of islands, for instance, LF-movement has been taken to be less restricted than syntactic movement. This is the view held in Huang (1982) and Lasnik and Saito (1992) for Chinese and Japanese wh-in-situ, respectively. To illustrate, consider the following examples from Chinese.

(5) a. Ni zui xihuan piping shei de shu? (Chinese)
you most like criticise who DE book
‘??Who do you like books that criticise?’

b. Ta xiang-zhidao shei maile shenme?
he wonder who bought what
‘What (x) he wonders who bough x?’

The sentences in (5) can be interpreted as direct questions despite the fact that the argument wh-phrases are embedded inside islands, i.e., the complex NP island in (5a) and the wh-island in (5b). Here, the absence of island effects has been taken as evidence for the claim that LF wh-movement is not subject to locality restrictions as syntactic movement. Similar observations are also found in English where LF movement of the argument wh-phrase who in (6a) is not as restricted as its moved counterpart in (6b).

(6) a. Who likes books that criticise who?

b. ??Who do you like books that criticise?

To recap, the LF-movement analysis just outlined assumes that in-situ wh-phrases undergo movement at LF like their overtly moved counterparts, as we have seen in the case of wh-movement of argument and adjunct wh-phrases in English. Where LF wh-movement is long-distance, i.e., across islands, it is taken to be less constrained than overt wh-movement, as we have seen in the case of Chinese and English. As concerns Arabic, the question is whether the LF-analysis presented in this section can account for the facts from wh-in-situ in Iraqi and Lebanese. This question is addressed in the next section.

6Examples are from Ouhalla 1996 and Bayer 2005.
LF wh-movement and wh-in-situ in Arabic

In this section, I examine whether a LF-movement analysis captures the facts from Iraqi and Lebanese wh-in-situ. I provide three arguments from both Iraqi and Lebanese which show that LF approach to wh-in-situ is incompatible. First, recall that the argument wh-phrase ‘who’ occurs in situ fully well-formed in both Iraqi and Lebanese (1a-b), whereas only Iraqi allows the occurrence of the wh-phrase ‘what’ in the base argument position (2a), but not Lebanese (2b). The relevant cases are repeated below.

(7) a. Mona shaafat *meno? (Iraqi)
   Mona saw whom
   ‘Who did Mona see?’

b. ˇ seft miin mbeerih? (Lebanese)
   saw.2sm who yesterday
   ‘Who did you see yesterday?’

(8) a. Mona ishtarat *eno? (Iraqi)
   Mona bought what
   ‘What did Mona buy?’

b. *ˇ starayte ′su mbeerih (Lebanese)
   bought.2sf what yesterday
   ‘What did you buy yesterday’

The first argument concerns the two way contrast between Iraqi wh-in-situ (8a) and Lebanese wh-in-situ (8b) and between Lebanese in-situ argument wh-phrases in (7b) and (8b). Such contrasts are impossible to explain in terms related to LF-movement. In other words, assuming that the wh-phrases in (7) and (8) undergo movement at LF, we still have no way of explaining why the LA sentence in (8b) is not well-formed like its Iraqi counterpart in (8a). Likewise, by assuming that wh-in-situ undergoes LF-movement, the contrast in the behaviour of the Lebanese argument wh-phrase miin ‘who’ in (7b) and the ar-
The case for LF-movement of wh-in-situ is further weakened by considerations of long wh-dependencies in both Lebanese and Iraqi. Facts from the two dialects show that a LF-analysis is inadequate. That is, whereas Lebanese wh-phrases may occur in-situ inside islands and have wide scope over the entire sentence, Iraqi argument wh-phrases cannot. The contrast between the Lebanese sentences in (9) and the Iraqi cases in (10) illustrates the case at hand.

(9)  a. badkun ta’yrfo [ʔeza Zeena šeefit miin want.2p to-know.2p whether zeena saw.3sf who b-l-maTṬam] in-the-restaurant ‘You want to know whether Zeena saw who in the restaurant?’

b. bta’yrfo [l-mara [yalli šeefit miin b-l-maTṬam] know.2pl the-woman that saw-3sf whom in-the-restaurant ‘You know the woman that saw who in the restaurant?’

(10) a. *Mona nasat [li-meno tinti šeno]? (Iraqi)
Mona forgot to-whom to-give what
‘?What did Mona forget to whom to give?’
‘Mona forget what she should give to whom?’

b. *Ṣurfut Mona [l-bint [illi ishtarat šeno]]?
knew Mona the-girl who bought what
‘??What did Mona know the girl who bought?’

The data from Lebanese show that LF-movement is not subject to island restrictions while Iraqi shows the opposite. Although the sentences in (9) contain the argument wh-phrase miin ‘who’ occurring inside islands, i.e., a wh-island (9a) and a complex NP island (9b), they are interpreted as direct questions in Lebanese. By contrast, the Iraqi sentences in (10) cannot be interpreted as
direct questions although the argument wh-phrase šeno ‘what’ occurs inside a wh-island (10a) and a complex NP island (10b) much in the same way as its Lebanese counterpart in (9). What these examples show is that LF-movement is constrained in one language, i.e., Iraqi, but not in the other, i.e., Lebanese. Under a (LF-) theory of movement, it is not clear how to explain inconsistency in the patterns of Lebanese wh-in-situ (9a-b) and Iraqi wh-in-situ (10a-b).

The third argument against the LF-approach to wh-in-situ comes from contrasts between wh-in-situ and wh-movement in Iraqi as well as in Lebanese. Both languages employ wh-movement along with wh-in-situ in the formation of wh-questions. However, there are asymmetries between in-situ argument wh-phrases and their moved counterparts within each dialect as well as across the two dialects. First, consider the Iraqi cases below.

(11)  
(a) *Mona tsawwarit Ali ishtara šeno?  
Mona thought ali bought what  
‘What did Mona think Ali bought’

(b) šeno tsawwarit Mona Ali ishtara?  
Mona thought ali bought what  
‘What did Mona think Ali bought’

Examples like (11) show that syntactic movement and LF movement of the argument wh-phrase šeno ‘what’ out of an embedded tensed clause have exactly the opposite properties. The sharp contrast between (11a) and (11b) cannot be accounted if one assumes that the in-situ wh-phrase šeno in (11a) undergoes movement to Comp at LF like its counterpart in (11b), there being no barriers to movement, as evidenced by full grammaticality of (11b).

Lebanese, also, shows a similar type of contrast between wh-in-situ and wh-movement, albeit in the opposite direction to that in Iraqi. Namely, wh-in-situ inside islands, like the wh-island in (12a) and the complex NP island in (12b),
is well-formed, whereas overt wh-movement across these islands is not (13a-b).

These are shown below.

(12) a. badkun taɣrfo [ʔeza Zeena šecfit miin
       want.2p to-know.2p whether zeena saw.3sf who
       b-l-maTŶam] (Lebanese)
       in-the-restaurant
       ‘Who do you want to know whether Zeena saw in the restaurant?’

       b. btaɣrfo [l-mara [yalli šecfit miin b-l-maTŶam]
       know.2p the-woman that saw-3sf whom in-the-restaurant
       ‘Who you know the woman that saw in the restaurant?’

(13) a. *miin badkun taɣrfo [ʔeza Zeena šecfit
       who want.2p to-know.2p whether zeena saw.3sf
       b-l-maTŶam] in-the-restaurant
       ‘Who do you want to know whether Zeena saw in the restaurant?’

       b. *miin btaɣrfo [l-mara [yalli šecfit b-l-maTŶam]
       who know.2p the-woman that saw-3sf in-the-restaurant
       ‘Who do you know the woman that saw in the restaurant?’

Clearly, the data from Iraqi and Lebanese posit problems for the LF-movement analysis of wh-in-situ. Both languages have wh-in-situ that appears to belong to different types. One type of wh-in-situ, i.e., the Iraqi type, shows locality restrictions, whereas its movement counterparts does not. The other type of wh-in-situ, i.e., the Lebanese type, does not exhibit locality restrictions, whereas wh-movement does. This conflict cannot be simply attributed to differences between levels of movement and is certainly beyond the reach of any LF approach to wh-in-situ. Overall, the facts from Arabic wh-in-situ and the arguments presented so far call into question the adequacy of movement approaches, such as LF-movement, to wh-in-situ in general, and to Arabic wh-in-situ in particular. The next section discusses a more promising approach to wh-in-situ that does not involve movement, namely, unselective binding.
2.2.2 Unselective Binding

The view that properties of wh-phrases and differences between types of in-situ wh-phrases are crucial in their distribution is advanced in Pesetsky (1987). Pesetsky makes a clear distinction between wh-phrases based on their discourse properties. Accordingly, there are two types of in-situ wh-phrase: d(iscourse)-linked wh-phrases like which-NP phrases, and non-d-linked wh-phrases, like who and what; both require different treatment\(^7\). In particular, Pesetsky suggests that d-linked wh-phrases be treated on a par with Heim’s (1982) indefinites in that they have no quantificational force of their own and may, therefore, be unselectively bound in-situ by an operator in a higher scope position (Baker’s 1970 unselective binding). On the other hand, Pesetsky treats non-d-linked wh-phrases like real quantifiers that therefore must undergo movement to Comp at LF for their interpretation. To illustrate, consider the contrast between (14) with non-d-linked who in situ and (15) which involves the d-linked wh-phrase which man.

(14)   a. *What did who read?
       b. *[\text{CP} [\text{who}_j [\text{what}_i], [\text{IP} e_j [\text{IP} \text{read} e_i]]]

(15)   a. Which book did which man read?
       b. [ \text{Spec} [c \text{ Q}_i, j] \text{ which book}_i \text{ did} [\text{which man}_j \text{ read} e_j]]

Based on the assumption that non-d-linked wh-phrases undergo movement at LF, movement of who in (14) induces an ECP violation (Chomsky 1981)\(^8\). LF-movement of the in-situ wh-phrase who in (14a) leaves its trace ungoverned at LF (14b). More precisely, antecedent government by the root Comp is not possible since Comp acquires the index of the wh-phrase what which has moved

---

\(^7\)What distinguishes d-linked wh-phrases such as which-NPs from other wh-phrases is that they, inherently, make reference to a set already established in discourse. Bare wh-phrases like who and what can also have a d-linked reading if they are understood to refer to previous discourse entities (see Pesetsky 1987 for a detailed discussion).

\(^8\)Pesetsky differs from Huang 1982 and Lasnik and Saito 1992 in that wh-movement is uniform in both LF and syntax.
first, i.e., in overt syntax. Lexical government is not possible either, as there is no lexical element to govern the trace of who. By contrast, (15) is grammatical since d-linked wh-phrases such as which man are not quantifiers. Like indefinites, they are not required to move at LF for their interpretation. Instead, unselective binding is said to be the mechanism that takes care of interpretation of d-linked wh-phrases. The [+Q] operator in the matrix Comp in (15b) is coindexed with and, hence, binds both the subject wh-phrase as well as the object wh-phrase.\(^9\)

Since binding does not involve movement, it follows that locality restrictions do not apply to d-linked wh-phrases occurring inside islands. Rather, d-linked wh-phrases may be bound in-situ by the higher Comp and have wide scope over the islands that contain them. Consider the contrast in the interpretation of English wh-questions like (16a) which involves a d-linked wh-phrase inside a wh-island with its non-d-linked counterpart in (16b) (Simpson 2000).

(16)  
a. Who remember where we bought which book?  
b. Who remembers where we bought what.

Under Pesetsky’s analysis, the wh-phrase which book in (16a) has matrix scope. The argument wh-phrase occurring inside the wh-island is d-linked, hence no movement at LF is required. Instead, the d-linked wh-phrase which book is bound in situ by the [+Q] Comp of the root clause and, hence, has wide scope over the island containing it. By contrast, since the wh-phrase in (16b) is non-d-linked, it must undergo movement at LF. LF-movement of what in (16b) is restricted to the lower [+Q] Comp, otherwise movement to the root Comp would yield a subjacency effect, hence the only admissible reading for (16b) is one where the wh-phrase has narrow scope.

\(^9\)As for overt movement of the wh-phrase which book in (15), Pesetsky attributes it to some morphological requirement on Comp that there be a lexical wh-element, rather than to quantifier/operator nature of the wh-elements.
To recap, this section has introduced Pesetsky’s analysis of wh-in-situ. It was shown that a Pesetsky-style approach to wh-in-situ entails two types of analysis: one that involves movement, whereby in-situ wh-phrases undergo movement at LF if they are non-d-linked, and another that involves binding, whereby wh-phrases may be bound in situ if they are d-linked. The reason why both analyses are needed is justified by empirical differences between wh-phrases in terms related to their discourse properties. In the remainder of this section, I explore the question of whether a Pesetsky-style approach can be extended to wh-in-situ in Arabic.

Unselective Binding and Wh-in-situ in Arabic

One of the advantages of Pesetsky’s analysis (1987) is the distinction it makes between types of (in-situ) wh-phrase, i.e., d-linked versus non-d-linked wh-phrases, and that this distinction entails different treatment. Whereas d-linked wh-in-situ receives a ‘Baker-style’ (non-movement) interpretation, non-d-linked wh-in-situ receives a ‘Chomsky-style’ LF-movement interpretation. This distinction seems, at first, to extend to Lebanese wh-in-situ which belongs to the d-linked type and to Iraqi wh-in-situ which belongs to the non-d-linked type. However, as the discussion unfolds, a Pesetsky-style analysis, though desirable, still needs some refinement to be able to account for all the facts here.

Starting with Lebanese wh-in-situ, the fact that it is d-linked means that in-situ wh-phrases should be interpreted in situ via binding in the sense of Baker (1970). Indeed, this is the analysis outlined in Aoun and Choueiri (1999) for wh-in-situ in Lebanese. To illustrate, consider the following example in (17a) with its representation in (17b).

(17) a. ˇ seft miin/?ayya mmasil mbeerih?
   saw.2sm who/which actor yesterday
   ‘Who/Which actor did you see yesterday?’

   b. [CP Spec C [+Q]1 [TP ... miin1/ ?ayya mmasil]]
Both argument wh-phrases *miin*/*paya mmasil* ‘who’/‘which actor’ appearing in the argument position in (17a) are d-linked, understood to refer to members of a set known to speaker and hearer. Following Pesetsky (1987), Aoun and Choueiri (1999) argue that the argument wh-phrases in (17) are bound in situ by the [+Q] operator in the Comp of the wh-interrogative which bears the same index as the wh-phrase, as shown in (17b).

Binding by an operator also accounts for wh-in-situ facts in Lebanese where argument wh-phrases occur inside islands and have scope higher than the island containing them. The relevant cases are repeated in (18a-b) for convenience.

(18) a. *badkun ṯa’rfo* [Peza Zeena ̱eefit *miin* want.2p to-know.2p whether zeena saw.3sf who in-the-restaurant](Lebanese) ‘Who do you want to know whether Zeena saw in the restaurant?’

b. *bta’rfo* [l-mara ̱eefit *miin* b-l-maṮam] know.2p the-woman that saw.3sf whom in-the-restaurant ‘Who do you know the woman that saw in the restaurant?’

The fact that the sentences in (18) are interpreted as direct questions follows from the binding relation established between the matrix [+Q] Comp and the d-linked wh-phrase in situ, in spite of the presence of a wh-island in (18a) and a relative clause island in (18b) separating the in-situ wh-phrase from its binder. This relation, according to Aoun and Choueiri (1999), cannot be established via movement since absence of island effects in sentences like (18) signals absence of movement altogether. The only way to capture the broad scope readings in cases like (18) is binding by the matrix operator of the wh-question.

However, Pesetsky’s analysis runs into several problems upon consideration of additional facts from wh-constructions in Lebanese and Iraqi. For instance,
Pesetsky’s analysis fails to capture the contrast in the distribution of Lebanese argument wh-phrases. Recall that only the argument wh-phrase *miin* ‘who’ may occur in situ, but not the argument wh-phrase *śu* ‘what’. The latter can only occur in the left periphery. Consider the cases in (19) and (20) from Lebanese.

(19) a. *śeft miin mbeerih*?
    saw.2sm who yesterday
    ‘Who did you see yesterday?’

b. *miin śeft mbeerih*?
    who saw.2sm yesterday
    ‘Who did you see yesterday?’

(20) a. *śtarayte śu mbeerih*
    bought.2sf what yesterday
    ‘What did you buy yesterday?’

b. *śu śtarayte mbeerih*
    what bought.2sf yesterday
    ‘What did you buy yesterday?’

There are actually two problems presented by the examples above. On the one hand, there is the contrast between the argument wh-phrase *miin* ‘who’ and the wh-phrase *śu* ‘what’. Whereas the wh-phrase *miin* ‘who’ can be d-linked and may, therefore, be bound in situ (19a), the wh-phrase *śu* ‘what’ cannot (20a).

On the other hand, there is the contrast in the behaviour of the wh-phrase *śu* ‘what’. Only when wh-phrase *śu* ‘what’ occurs in a clause-initial position is the sentence well-formed (20b). The fact that the argument wh-phrase *śu* ‘what’ cannot be d-linked in Lebanese and therefore cannot be bound in situ, unlike the argument wh-phrase *miin* ‘who’, does not receive an explanation under a Pesetsky-style analysis.

Furthermore, the contrast between Lebanese wh-in-situ and Iraqi wh-in-situ posits additional problems for the binding analysis of Pesetsky (1987). For ex-
ample, Iraqi wh-in-situ should receive a LF-treatment under Pesetsky’s analysis since Iraqi argument wh-phrases are of the non-d-linked type. As such, they are expected to behave like real quantifiers in that they must undergo movement into Comp at LF for their interpretation. However, as we have seen in the previous section, a LF-movement approach cannot capture the asymmetries between wh-in-situ constructions and wh-movement constructions. These are shown below.

(21)  
  a. *Mona tsawwarit Ali ishtara šeno?  
       Mona thought ali bought what  
       ‘What did Mona think Ali bought’

  b. šeno tsawwarit Mona Ali ishtara?  
       Mona thought ali bought what  
       ‘What did Mona think Ali bought’

The sentence in (21b) shows that overt movement of the wh-phrase šeno ‘what’ out of the embedded tensed clause is fully grammatical. There is no reason why the sentence in (21a), under the assumption that the in-situ wh-phrase šeno undergoes movement at LF, should not have the same status as its counterpart in (21b). This contrast is puzzling in that LF wh-movement (21a) seems to be more constrained than overt wh-movement, contra standard assumptions.

A similar observation is found in Ouhalla’s analysis (1996) for Iraqi wh-in-situ. Ouhalla observes that there is a conflict between Iraqi in-situ wh-phrases and moved wh-phrases in island contexts. To illustrate, consider the following examples.

(22)  
  a. *Nasat Mona [li-meno tinki šeno]?  
       forgot Mona to-whom to-give what  
       ‘What(x) Mona forgot to whom to give x’

  b. *Qurfut Mona [l-bint [illi ishtarat šeno]]?  
       knew Mona the-girl who bought what  
       ‘What(x) Mona knew the girl who bought x’
(23)  a. ??ˇseno nasat Mona [li-meno tinti ]?
    what forgot Mona to-whom to-give

b. ??ˇseno ˇyurfut Mona [l-bint [illi ishtarat ] ]?
    what knew Mona the-girl who bought

The sentences in (22) in which the argument wh-phrase ˇseno ‘what’ occurs in situ inside islands i.e., a wh-island (22a) and a complex NP-island (22b), display a stronger type of ungrammaticality than those in (23). Although (23a) and (23b) involve overt wh-movement across the same islands as those in (22), they show a milder Subjacency-type violation, as indicated by (??). According to Ouhalla (1996), the strong ungrammaticality in (22) cannot be attributed to LF wh-movement. That is, if LF movement were to apply in (22), one expects, at most, a mild violation similar to the one in (23). Altogether, the examples in (21)-(23) present a rather odd situation where it is LF-movement that is more constrained than overt movement.

To recap, the sections above have been concerned with the differences between (types of) wh-in-situ in Iraqi and Lebanese. These differences can be characterised as follows: on a parametric level, certain variations exist between Iraqi wh-in-situ and Lebanese wh-in-situ much in the same way as the difference between d-linked and non-d-linked wh-in-situ. On a micro-parametric level, asymmetries are between properties of argument wh-phrases within each dialect. Two of the main approaches to wh-in-situ have been discussed, i.e., an LF-movement approach and a binding approach. I have shown that the LF-analysis cannot capture the patterns of variation observed in and between the dialects. Meanwhile, (unselective) binding has been shown to provide an explanation for some of the facts from Lebanese wh-in-situ, though it appears to be too broad to account for the type of variation in Iraqi and Lebanese.
2.3 Morpho-syntactic Analysis of Iraqi Argument Wh-phrases

2.3.1 Introduction

In the previous sections, we have seen that the LF-movement analysis fails to explain the range of facts from Iraqi and Lebanese wh-constructions. The differences between Iraqi wh-in-situ and Lebanese wh-in-situ, the asymmetrical behaviour of in-situ argument wh-phrases in both languages as well as the asymmetries between in-situ wh-phrases and their moved counterparts make it impossible for a theory of movement, i.e., LF-movement, to accommodate such variation. Binding by an operator seems more plausible in that it captures some of the facts about Lebanese wh-in-situ. It was concluded that binding on its own cannot resolve all the problems related to wh-in-situ variation in Arabic.

The facts from both Iraqi and Lebanese lend themselves to the type of approach needed to account for the asymmetries in the behaviour of argument wh-phrases. Namely, an approach that takes properties of individual wh-phrases to play a crucial role in the micro and macro parametric setting of wh-licensing in each of the dialects studied here. Before articulating the proposal, it is worth discussing the approach developed in Ouhalla (1996) for Iraqi wh-in-situ that is essentially based on morpho-syntactic properties of argument wh-phrases. Although Ouhalla’s analysis appears to be on the right track, the assumptions leading to the analysis are, in fact, problematic. In what follows I present the main claim in Ouhalla’s analysis and show how it cannot be empirically supported.

2.3.2 Ouhalla’s 1996 Analysis

To account for the asymmetries in Iraqi wh-constructions, Ouhalla (1996) claims that Iraqi argument wh-phrases exhibit a parallel behaviour to compound reflex-
ive anaphors in that they both require a local antecedent\textsuperscript{10}. To illustrate, compare the Chinese sentence in (24) which contains the compound anaphor \textit{ta-ziji} ‘himself’ and the Iraqi sentence in (25) which contains the compound wh-phrase \textit{meno} ‘who’, both occurring in object position in the embedded clause.

(24) Zhangsan, shuo Lisi, chang piping \textit{ta-ziji,\!/j}
Zhangsan say Lisi often criticise himself
‘Zhangsan said Lisi often criticised himself’

(25) *Mona tsawaarat Ali ishtara \textit{šeno}?  
Mona thought Ali bought what
‘What did Mona think Ali bought?’

In the sentences above, both the anaphoric-expression \textit{ta-ziji} ‘himself’ and the wh-expression \textit{šeno} ‘who’ are said to be bound in the syntax. Based on this treatment, Ouhalla takes the difference between the two types of anaphor to be residing in the nature of the antecedent they require. Compound reflexive anaphors require an A-antecedent, which is due to the anaphoric feature [+refl(exive)]; while Iraqi argument wh-phrases require an A’-antecedent, due to the anaphoric feature [+wh]. In the latter case, it is a Comp marked with a [+wh] feature that binds the wh-expression in situ. According to Ouhalla (1996), the Chinese compound anaphor \textit{ta-ziji} ‘himself’ in (24) can have as an antecedent only the subject of the lower clause. The anaphor cannot be bound, however, by the distant antecedent since this would give rise to a binding theory violation resulting from the presence of a potential local antecedent. Likewise, the ungrammaticality of the Iraqi sentence in (25) is due to a binding theory violation. That is, the wh-phrase \textit{meno} ‘who’ cannot have as an antecedent the Comp of the matrix clause in the presence of an embedded Comp as a potential antecedent.

\textsuperscript{10}Ouhalla adopts Aoun’s (1985, 1986) Generalised Binding, in particular, Condition A which requires that anaphors be bound in a given domain. The notion \textit{given domain} is defined according to whether the (wh-)anaphor is local or long-distance, which in turn depends on whether the anaphor is compound, i.e., morphologically complex, or bare, i.e., morphologically simplex. Whereas compound anaphors are bound locally, bare anaphors are bound long-distance.
The core assumption that underlies the claim put forward in Ouhalla (1996), i.e., that Iraqi argument wh-phrases are compound anaphoric expressions, is that argument wh-phrases in Iraqi, such as \textit{meno} ‘who’ and \textit{ˇseno} ‘what’, consist of two elements: a wh-element and a pronominal element. This is illustrated in (26) below.

(26) a. men-\textbf{o}  
who-him  
b. ˇsen-\textbf{o}  
what-it

While the wh-element encodes the [+wh] feature and is realised by the wh-morphemes \textit{men} ‘who’ in (26a) and \textit{ˇsen} ‘what’ in (26b), the pronominal element encodes \(\phi\)-features and is realised by the clitic -\textbf{o} ‘him/it’. According to Ouhalla (1996), the pronominal element attached to the wh-morphemes \textit{men} and \textit{ˇsen} in (26a-b) functions independently in the language as an object clitic pronoun, as shown in (27a) and (27b), respectively (Ouhalla’s (11a-b)).

(27) a. ˇSuft-\textbf{o}.  
saw.I-him  
‘I saw him’  
b. Iˇstarit-\textbf{o}.  
bought.I-it  
‘I bought it’

Ouhalla assumes that the pronoun attached to the wh-elements in (26) is the same as that attached to the verbs in sentences like (27a) and (27b). In wh-in-situ contexts, Ouhalla takes the object clitic -\textbf{o} to be a variable that needs to be bound an by an antecedent while the wh-phrase identifies the antecedent as a Comp with a [WH] feature.
Problems with Ouhalla’s analysis

Although Ouhalla’s analysis seems attractive, it is not unproblematic. In particular, the assumption that the pronominal element is an object clitic amounts to treating Iraqi wh-phrases as clitic hosts just like verbs, as is the case in (27) above. There are two problems with this assumption: one problem concerns the treatment of the pronominal element as a clitic while the other problem concerns treatment of the wh-phrase as a potential host for clitics. Upon closer examination, the pronoun appearing with the wh-elements in (26) cannot be an object pronoun in the same way as the object clitic pronoun in (27). Likewise, the wh-element cannot be a host for clitics like the verbs in (27). As such, Ouhalla’s analysis of the morpho-syntax of Iraqi argument wh-phrases cannot be maintained. Below, I provide a number of empirical arguments against the status of the wh-elements *men* and *šen* as clitic hosts and the status of the pronominal element as a normal clitic. Meanwhile, I maintain the view that Iraqi argument wh-expressions do contain a wh-element and a pronominal element, albeit this pronominal element is a subject clitic. Such a view, in turn, will have crucial consequences for the alternative analysis proposed afterwards.

The first argument against Ouhalla’s morpho-syntactic analysis of Iraqi wh-phrases concerns the status of the wh-element *men* ‘who’ as an $X^0$ category. In particular, I show that the wh-element *men* ‘who’ does not display properties that characterise it as a potential host for the clitic -ő ‘him/it’, as argued for in Ouhalla (1996). The argument is based on a general observation about object pronouns in Arabic. Arabic object pronouns are weak and have the distinctive property of appearing as clitics attached to lexical heads that govern them (Fassi-Fehri 1993, Shlonsky 1997, Aoun, Choueiri and Hornstein 2001, Ouhalla and Shlonsky 2002 and Aoun and Li 2003, among others)\(^\text{\footnote{Weak pronouns are weak in the sense that they, unlike strong pronouns, cannot be contrastively stressed, coordinated or modified (Cardinalletti and Starke 1994, and Shlonsky 1997).}}\). Although clitics in Arabic do not select a particular category to attach to, there are restrictions on
their potential hosts. In particular, a host must be a lexical head, i.e., $X^0$, and have the ability to assign Case. Consider the following paradigm from Standard Arabic, Iraqi and Lebanese (Aoun and Li 2003).

(28)

a. raʔayto-hu
   saw.1s-him
   ‘I saw him.’

b. tahaddaʔna maʔa-hu
   spoke.1p with-him
   ‘We spoke with him.’

c. qabalna ʔumma-hu
   met.1p mother-his
   ‘We met his mother’

(29)

a. l-mudiira Tordot-uḥ
   the-principal.fs expelled.3fs-him
   ‘The principal expelled him.’

b. thaʔee-na weyyaa-h
   spoke.1p with-him
   ‘We spoke with him.’

c. šofna ʔumm-a
   saw.1p mother-his
   ‘We saw his mother’

(30)

a. l-mudiira ha-teshhat-u
   the-principal.fs will-3fs.expel-him
   ‘The principal will expel him.’

b. hekina maʔ-o
   spoke.1p with-him
   ‘We spoke with him’

c. ʃefna ʔemm-o
   saw.1p mother-his
   ‘We saw his mother’
The sentences in (28-30) above show that not only can clitics attach to verbs, as in the (a) examples, and prepositions, as in (b) examples, but they can also attach to nouns as shown in the (c) examples. As might have been noticed, all these categories have in common their status as $X^0$. Moreover, the clitics in the sentences above occur in Case-marked positions and are, therefore, assigned Case. That is, the clitic in the (a) examples is assigned accusative Case by the verbal head; oblique Case by the prepositional head in the (b) examples and genitive Case by the (possessive) nominal head in the (c) examples\(^\text{12}\). The paradigm in (28)-(30) gives rise to the question of whether the wh-element *men* is compatible with those categories that host clitics. If it were, it would behave like the N-head in the (c) cases.

The genitive Case assignment by nouns shown in the (a) examples in the paradigm above is especially interesting in that it shows that nouns in Arabic, among other properties, have the property of assigning Case, unlike nouns in other languages such as English. This point deserves further discussion. Arabic nouns assign genitive Case to the NP following them in the so-called *Construct State* in Semitic languages (Mohammad 1988, Ritter 1991, Fassi-Fehri 1993, Benmamoun 1998 and Ouhalla and Shlonsky 2002) illustrated in the following examples from Standard Arabic (SA) and Iraqi\(^\text{13}\).

(31) a. kitaab-u T-Taalib-at-i
    book-nom the-student-fs-gen
    ‘The student’s book’

    b. ?umm l-walad
    mother-nom the-boy-gen
    ‘The boy’s mother’

As the examples above show, Construct State (genitive) constructions involve, at least, two nouns whereby the first noun (possessee) assigns genitive Case to

\(^{12}\)Notice also that the clitics in (28)-(30) receive theta roles from the heads that govern them, i.e., Theme, Experiencer and Possessor, respectively

\(^{13}\)Examples from Standard Arabic (SA) are given here to, essentially, show overt Case marking since Standard Arabic is the only variety with overt Case marking.
the second noun (possessor). There are two peculiarities exhibited by Construct States like (31) which are related to the current discussion: genitive Case assignment by a noun and the possessee-possessor word order. To deal with these issues, I follow Mohammad (1988), Ritter (1991) and Benmamoun (1998), and take the first noun to originate as head of a lexical NP projection while the second noun to originate as the specifier of that NP. To illustrate, consider the structure in (32) below.

(32)

According to Benmamoun 1998, the head noun ʔumm ‘mother’ in (32) assigns genitive Case to the NP ʔal-ad ‘the boy’ in its specifier, i.e., in a Spec-head configuration. As for the possessee-possessor word order, it is derived by movement of the possessee (the head noun) from the head N position to the head of the DP (N-to-D raising), thereby leaving the possessor (genitive NP) within the lexical projection. This is shown in the structure in (33).
The generalisation that can be made from the data above is that for any nominal category to qualify as a potential host for clitics it must be both an $X^0$ category and a Case-assigner. Given that nouns are $X^0$ categories which assign genitive Case to NPs following them in Arabic, and, hence, have the ability to host clitics, the assumption that Iraqi argument wh-phrases occur with an object clitic receives no empirical support. In other words, for a clitic like *o ‘him/it’ to attach to the wh-element *men ‘who’, for instance, the latter must be an $X^0$, i.e., an N category. However, the wh-word *men ‘who’ does not behave as though it is an N head. First, consider the contrast in the following examples.

(34) a. ktaab l-benet
     book the-girl-gen
     ‘The girls’s book

b. ktaab-ha
   book-her
   ‘Her book’

(35) a. *men l-benet
     who the-girl-gen

b. *men-ha
   who-her

The contrast between the cases in (34) and (35) shows that the wh-word *men ‘who’ cannot appear in the same distributional contexts as nominal heads.
Unlike the nominal head kitaab ‘book’ of the Construct State in (34a), men ‘who’ cannot assign genitive Case to the NP following it in (35a), hence the ungrammaticality. Furthermore, men ‘who’ cannot occur with clitics usually associated with nominal heads, hence the contrast in grammaticality between (34b) and (35b).

The argument that the wh-word men ‘who’ is not, and, in fact, cannot be, an X₀ is further supported by considerations of other wh-words in Arabic. That is, there exist certain wh-words that display properties that are characteristic of (nominal) heads in Arabic. For instance, the wh-word ʔayya ‘which’, found across the spectrum of Arabic dialects (Aoun and Li 2003, Holes 2004 and Aoun, Benmamoun and Choueiri 2010), invariably assigns genitive Case to the NP following it and has the ability to host clitics. Consider the following sentences from Standard Arabic (SA), Iraqi and Lebanese, respectively.

(36) ʔayy-u kitaab-in saqaTa ʔalaʔi-arD-i?

which-nom book-gen fell.3sf on the-ground-gen?

‘Which book fell on the ground?’

(37) ʔayy dawlah tetbaʔoon?

which country-gen 2p.follow

‘Which country do you follow?’

(38) ʔayya-hun seefer?

which-them travel.3ms

‘Which one of them has travelled?’

In the sentences (36-38) above, ʔayya-NP ‘which’-NP phrases exhibit identical behaviour to Construct State phrases. More precisely, like the first member in a Construct State, the wh-word ʔayya(a) ‘which’ assigns genitive Case to the following NP in the sentences above, and hosts a clitic bearing φ-features as shown in (38).
In terms of analysis, there is a clear sense that *payya* ‘which’-NP phrases should be treated on a par with Construct States. In particular, the behaviour of the wh-word *payya* ‘which’ parallels that of the nominal head in Construct State phrases. Such a view is also confirmed by a similar observation made in Benmamoun (1998) for the quantifier (Q) *kull* ‘all’ in Q-NP constructions in Arabic14. Benmamoun argues that the quantifier *kull* ‘all’ displays all properties associated with nominal heads, especially, its ability to both assign genitive Case to the following NP (39a-b) and host clitics (39c), (Benmamoun 1998).

\[(39)\]

\[a. \quad \text{kull-u T-Tullab-i} \quad \text{zaa?aa-uu}\]
\[\text{all-nom the-students-gen} \quad \text{came-3ms}\]
\[\text{‘All the students came’}\]

\[b. \quad \text{ra?ayt-u kull-a T-Tullab-i}\]
\[\text{saw.1s all-acc the-students-gen}\]
\[\text{‘I saw all the students’}\]

\[c. \quad \text{kull-u-hum/hunna}\]
\[\text{all-nom-them(m)/them(f)}\]
\[\text{‘All of them (m)/them(f)’}\]

The patterns shown by the quantifier *kull* ‘all’ in (39) are exactly the same as those associated with the wh-word *payya* ‘which’ in (36-38). To this effect, the analysis outlined in Benmamoun (1998) for the quantifier *kull* ‘all’, which is, essentially, based on the parallelism with the Construct State, can also be extended to the wh-word *payya* ‘which’. That is, like the quantifier *kull* ‘all’, I take the wh-word *payya* to originate as head of a QP projection contained in a DP structure like that of the Construct State. As for the second NP, it originates as the Specifier of the QP projection. Thus, the structure of *payya*-NP ‘which’-NP phrases might be represented in (40).

---

14See also Shlonsky (1997) for a similar view on the quantifier *kol* ‘all’ in Hebrew.
The NP in specifier position of the head ʔayya ‘which’ in (40) is assigned genitive Case via Spec-head configuration, as happens for the NP specifier of the quantifier kull ‘all’. Furthermore, to derive the ʔayya ‘which’-NP word order, the wh-word ʔayya ‘which’ undergoes head-movement to the D position of the DP, checks and deletes the [WH] feature in D. Consequently, the genitive NP is left in situ in the Spec position within the lexical projection. The structure for ʔayya-NP ‘which’-NP phrases looks like that in (41) below, after movement of the wh-element ʔayya ‘which’ has taken place.

Turning back to the wh-word men ‘who’ in Iraqi, it does not appear to display any of the properties associated with nominal heads. This contrasts sharply with the wh-word ʔayya ‘which’ which displays properties character-
istic of nominal heads, i.e., genitive Case assignment and clitic hosting. The contrast in the behaviour of the wh-word *ayya ‘which’ and the wh-word *men ‘who’ in Arabic constitutes strong evidence that whereas the former is an X\(^0\), the latter simply cannot be.

Another observation about the categorial distinction between the wh-word *ayya ‘which’ and the wh-word *men ‘who’ comes from the fact that the wh-word *ayya ‘which’ cannot occur on its own, unlike the wh-word *men ‘who’. Examine the following cases.

\[(42)\]

\[\begin{align*}
\text{a. } & \text{*ayya dawla}\text{h tetb\'a\text{\textacute{y}oon}?} \\
& \text{which country 2p.follow} \\
& \text{‘Which country do you follow?}\end{align*}\]

\[\begin{align*}
\text{b. } & \text{*ayya tetb\'a\text{\textacute{y}oon}?} \\
& \text{which 2p.follow} \\
& \text{‘Which do you follow’?}\end{align*}\]

\[\begin{align*}
\text{c. } & \text{men tetb\'a\text{\textacute{y}oon}} \\
& \text{who 2p.follow} \\
& \text{‘Who do you follow?’}\end{align*}\]

There are two remarks that are worth pointing out on the basis of data like (42a-c). First, the fact that the wh-word *men occurs on its own (42c) indicates that it is not an X\(^0\) category like *ayya ‘which’, but rather an XP. Secondly, notice that the wh-word *men ‘who’ in (42c) occurs in the position of the whole DP *ayya-dawleh ‘which country’ in (42a). This, in turn, shows that the wh-word *men is not just a morpheme encoding the [WH] feature, but rather, a full-fledged Wh-DP.

The status of *men ‘who’ as a DP is confirmed in contexts of the Construct State, which we have examined earlier. Recall that *men ‘who’ cannot occur in the position of the first member of the Construct State in Arabic since this
position is restricted to nominal heads only. Interestingly, however, it can occur in the position of the second member, i.e., the Spec position of the (definite) genitive NP. Consider the sentence in (43a) and its partial representation in (43b).

(43)  

a.  

\[ \text{ktaab \textbf{men} treed?} \]  

book who 2ms.want  

\text{‘Whose book do you want?’} 

b.  

\[
\begin{align*}
\text{DP} \\
\quad \text{Spec} \\
\quad \text{D'} \\
\quad \text{D} \\
\quad \text{NP} \\
\quad | \\
\quad \text{ktaab} \\
\quad | \\
\quad \text{Spec} \\
\quad \text{N} \\
\quad | \\
\quad \text{book} \\
\quad | \\
\quad \text{\textbf{men}} \langle \text{ktaab} \rangle \\
\quad | \\
\quad \text{who}
\end{align*}
\]

In (43b), the wh-word \textit{men} occurs in a position restricted to DPs that are assigned (genitive) Case by their associated heads, as we have seen in the cases of Construct State, \textit{?ayga}-NPs as well as Q-NPs.

The discussion has been thus far concerned with the status of the wh-word \textit{men} ‘who’ which constitutes part of the structure of the wh-expression \textit{men-o} ‘who’ in Iraqi. I have shown that the wh-word \textit{men} ‘who’ does not display any of the properties that identify it as a potential host for clitics in the same way as other heads in the language such as N, V and P. The data presented in this section provides sufficient evidence that this wh-element is a wh-phrase like its counterpart \textit{munn} ‘who’ in Standard Arabic and \textit{miin} ‘who’ in Lebanese.
Concerning clitics, they, too, cannot occur with wh-words like <i>men</i> ‘who’ and <i>šen</i> ‘what’, as argued for by Ouhalla (1996). If the assumption that the pronominal element -<i>o</i> in <i>men-o</i> ‘who’ and <i>šen-o</i> ‘what’ is an object clitic pronoun were correct, there is no reason why other pronominal forms could not attach onto the wh-word, as is the case with other heads. This is not attested in Iraqi or in any other Arabic dialect. Consider the cases in (44) where there is no restriction on the form a clitic may have when it occurs with a verbal head, as opposed to cases where only one form of clitic is allowed with the wh-word <i>men</i> shown in (45).

(44)  

a. šuft-o/-ha/-hum  
saw.I-him/-her/-them  
‘I saw him/her/them’

b. šuft-ak/-ič/-kum  
saw.I-you(m)/-you(f)/-you(pl.)  
‘I saw you(m)/you(f)/you(pl)’

(45)  

a. men-o/-*ha/-*hum  
who-him/-her/-them

b. men-*ak/-*ič/-*kum  
who-you(m)/-you(f)/-you(pl.)

The data above show that the pronominal element occurring with the wh-word <i>men-o</i> ‘who’ cannot be be part of the paradigm of object clitics in (44). Notice, for instance, that third person and second person pronominal forms in (44a) and (44b), respectively, trivially attach to the verb, unlike the ones in (45a) and (45b)\(^{15}\). The reason why the pronominal element -<i>o</i> is odd and cannot be compared to the paradigm of clitics in (44) is because it is, simply, not an object clitic. As might be recalled, Ouhalla’s analysis predicts that cases in (45) should be as well-formed as those in (44). The prediction is not borne out and the fact that only one form of a pronoun may occur with the wh-element

\(^{15}\)Notice also that this is true for the other argument wh-phrase <i>šen</i> ‘what’ in that we do not find wh-phrases like <i>šen-ha</i> ‘what-her’ or <i>šen-hum(ma)</i> ‘what-them’.
men in (45) is left unexplained.

The second argument against treating the pronominal element in men-o ‘who’ as an object clitic concerns the prosodic status of pronouns. Since clitics are weak pronouns, they cannot be contrastively stressed nor can they be focussed (Cardinalletti and Starke 1994, and Shlonsky 1997). In Arabic, when a clitic attaches to its governing head, say a verbal head, the stress remains on the head with no change in the stress pattern, as shown below (I am notating stress with small caps).

(46) a. suft-o
     saw.I-him
     ‘I saw him’

b. *suft-O
     saw.I-him

The case for the pronominal element -o attaching to the Iraqi wh-phrase men ‘who’ is different. More precisely, attaching the pronominal element to the expression men-o ‘who’ induces a shift of stress from the former onto the latter. This, in turn, indicates that this pronominal element cannot be a clitic in the same way as object clitics in the language. Consider the contrast between the two patterns in (46) with those in (47) below.

(47) a. *MEN-o
b. men-o

To sum up, I have shown that Ouhalla’s (1996) morpho-syntactic analysis of Iraqi argument wh-phrases men-o ‘who’ and šeno ‘what’, is problematic. The assumption that they contain an object clitic pronoun receives no independent support in Iraqi or any other dialect. I have also shown that the wh-word men ‘who’ does not display properties that characterise it as N, especially, the fact
that it cannot assign genitive Case to NPs following it and the fact that it
cannot occur with normal clitics. Evidence from contexts of (nominal heads
in) Construct State phrases, ḫayya-NP ‘which’-NP phrases, and Q-NP phrases
show that the bare wh-phrase men is a DP and that the clitic annexed to it is
not an object clitic after all, as originally assumed in Ouhalla (1996).

As it stands, it appears that the pronominal element -o is the only form found
occurring with the wh-phrase men. The discussion so far indicates that the wh-
expression men-o ‘who’ is not a pure DP like men ‘who’, but, rather, involves
an extra element and that element is not an object clitic. The true identity of
this pronominal element and its impact on the structure and interpretation of
the wh-expression men-o ‘who’ are examined next.

2.4 The syntax of men-o

Based on the arguments provided in the previous sections, the assumption that
the pronominal element is an object clitic pronoun hosted by the wh-word men
‘who’ in Iraqi is simply incorrect. The fact that men ‘who’ does not behave
like a head and the fact that the pronominal element suffixed to men ‘who’ is
not an object pronoun cast serious doubt on the plausibility of Ouhalla’s (1996)
morpho-syntactic analysis for Iraqi wh-argument expressions. This section ar-

gues that the pronominal element -o in men-o is a reduced version of the third
person subject pronoun huwa ‘he’. Given that Iraqi men is a DP, the fact that
the pronominal element is a subject pronoun has crucial consequences on the
syntax of Iraqi wh-questions.

The following subsections are organised as follows. In subsection (4.1), I
provide evidence that the pronominal element is a subject pronoun. Subsec-

---

16 Additionally, the analysis outlined in Ouhalla (1996) for Iraqi argument wh-phrases men-o
‘who’ and šenb ‘what’ fails to capture the contrast observed at the outset, between Iraqi and
Lebanese wh-phrases, i.e., šu and šeno ‘what’, and between Lebanese argument wh-phrases
mim ‘who’ and šu ‘what’, respectively.
tion (4.2) looks into copular wh-questions in Arabic and argues that the Iraqi wh-expression *men-o* is a copular wh-clause. In subsection (4.3), I provide evidence from other dialects which show that copular wh-phrases are a common phenomenon in Arabic.

### 2.4.1 The pronominal element in *men-o* is a subject pronoun

The assumption that the pronominal element in *men-o* is a subject pronoun is based on an observation about the forms of third person subject pronouns in spoken Arabic dialects. Specifically, the form a subject pronoun may have appears to be subject to phonological variation within each dialect depending on the order and position it occupies in relation to other constituents. In cases where the pronoun is always preceded by the same lexical item such as the wh-phrase *miin*, many of these dialects tend to use more than one version of the same pronoun. For instance, the Iraqi subject pronoun may appear in the contracted form *o*, the reduced form *hu* or the full form *huwa*, while the Jordanian subject pronoun may appear as either *hu* or *huwa*. By contrast, there is no such variation in Standard Arabic; the only form used in this language is the classic form *huwa* ‘he’

\[17\] This can be attributed to the fact that SA is not a spoken dialect.
The table in (48) shows the range of variation in the phonological form of third person subject pronouns in Arabic. Although a full discussion of the factors that affect how these forms are produced is beyond the scope of this section, it suffices for the current purposes to mention a few points. First, note that (the lack of) variation is a way of how spoken dialects like Iraqi and Jordanian are developing away from the standard variety. In spite of this development, the fact remains that the syntax of these elements remains intact. Secondly, while the forms *huwa* ‘he’, *hu* ‘he’, *hiyya* ‘she’ and *hi* ‘she’ are free-standing pronouns, the forms -u/-o ‘he’ and -i ‘she’ are not. In the latter case, the subject pronoun is treated as phonologically deficient that must, therefore, be attached to another lexical item\(^{18}\).

A number of empirical arguments can be invoked to support the view that the pronominal element is a subject pronoun. The first argument comes from constructions which involve the wh-phrase *who* followed by a subject pronoun, such as copular wh-constructions and the so called reduced-cleft wh-questions (Cheng 1991, Ouhalla 1996, and Shlonsky 2002). First, consider examples of

\(^{18}\)This is exactly where Ouhalla’s assumption concerning the pronominal element being an object pronoun went wrong.
reduced cleft wh-questions from SA and Iraqi.

(49) a. mann huwa al-ladhii ra?ayta-hu? (SA)
   who he the-one saw.2ms-him
   ‘Who is the one that you saw?’

   b. men-o illi šuft-o? (Iraqi)
   who-he that saw.2ms-him
   ‘Who is it that you saw?’

In this type of wh-question, the wh-phrase is followed by a subject pronoun and a clause headed by the complementiser illi ‘that’. In (49a), the SA subject pronoun is a freestanding pronoun. Whereas in (49b), the Iraqi pronoun is a contracted version of the one in (49a) and is attached to the wh-phrase. Notice that despite the fact that the Iraqi subject pronoun is a reduced version of the actual form, it still functions as a subject pronoun as its counterpart in (49a). This fact can be seen more clearly in copular wh-questions of the type illustrated below.

(50) a. men-o l-muddaris? (Iraqi)
   who-he the-teacher
   ‘Who is the teacher?’

   b. *men-o hu(wwa) l-muddaris?
   who-he be the-teacher
   ‘Who is the teacher?’

(51) a. man hu l-mudarris? (Jordanian)
   who he the-teacher
   ‘Who is the teacher?’
b. *man hu hu l-mudarris?
   who he he the-teacher
   ‘Who is the teacher?’

(52)  a. mann huwa al-mo'allim(-u) (SA)
   who he the-teacher
   ‘Who is the teacher?’

b. *mann huwa huwa al-mo'allim(-u)
   who he he the-teacher
   ‘Who is the teacher?’

The cases above exemplify what I have called LD’ed equatives which systematically display the format [(Wh-) DP [PRON DP]]\(^{22}\). Recall that in such constructions the initial DP is associated with a subject pronoun inside the clause. This is exactly what the (a) examples above show. In Iraqi, the subject pronoun is cliticised to the wh-phrase, whereas in Jordanian and SA this pronoun is a freestanding pronoun. Notice that the presence of another subject pronoun in the Jordanian example (51b) and the SA example (52b) is ruled out on the grounds that the subject position can only be filled with a subject pronoun, thereby, saturating the argument structure. Interestingly, the Iraqi example (49b) is also ruled out on the same grounds since the clitic is a subject pronoun occupying the only subject position in the structure.

Further evidence that the pronominal element in \textit{men-o} is a subject pronoun comes from agreement patterns that are generally associated with subject pronouns. In particular, the pronominal element of \textit{men-o} shows gender agreement patterns that are characteristic of subject pronouns. To illustrate, consider the following examples.

\(^{22}\)LD’ed equatives are discussed in chapter 1.
The pronoun attached to the wh-phrase in (53) shows agreement with the initial DP antecedent. In (53a), the pronoun shows masculine agreement with the DP Ệaboo-k ‘your father’; while in (53b), it shows feminine agreement with the DP Ệumma-k ‘your mother’. This type of agreement is the same as that of (the reduced) subject pronouns shown in (48) above. If, on the other hand, this pronominal were an object pronoun, we would expect object agreement patterns to obtain, as we have seen earlier on. To clarify further, consider the contrast in grammaticality between the sentences in (54) and (55).

(53)  

a. Ệaboo-k men-o?  
father-your who-he  
‘Your father, who is he?’

b. Ệumma-k men-i?  
mother-your who-she  
‘Your mother, who is she?’

The pronoun attached to the wh-phrase in (53) shows agreement with the initial DP antecedent. In (53a), the pronoun shows masculine agreement with the DP Ệaboo-k ‘your father’; while in (53b), it shows feminine agreement with the DP Ệumma-k ‘your mother’. This type of agreement is the same as that of (the reduced) subject pronouns shown in (48) above. If, on the other hand, this pronominal were an object pronoun, we would expect object agreement patterns to obtain, as we have seen earlier on. To clarify further, consider the contrast in grammaticality between the sentences in (54) and (55).

(54)  

a. ụfọt-o  
saw.1s-him  
‘I saw him’

b. ụfọt-ha  
saw.1s-her  
‘I saw here’

(55)  

a. *Ệaboo-k men-o?  
father-your who-he  
‘Your father, who is he?’

b. *Ệumma-k men-ha?  
mother-your who-she  
‘Your mother, who is she?’

The last piece of evidence supporting the assumption that the pronominal element in men-o is a subject pronoun and functions as such concerns the prosodic status of the whole compound. Putting aside the Iraqi for the moment, one interesting case is Jordanian which has two forms for the wh-phrase ‘who’:
miin and men. The first form, i.e., miin, is used in wh-constructions that do not involve a subject pronoun following it. However, when this wh-phrase is followed by the subject pronoun, two things happen at the same time: reduction in the form of the wh-phrase and stress shift from the wh-phrase to the subject pronoun, i.e., miin → men hu. The same scenario applies to the Iraqi wh-phrase when it occurs with the subject pronoun except that Iraqi, unlike jordanian, has one form for the wh-phrase who, i.e., men. Consequently, contrasts in cases like (56), repeated below, are accounted for.

(56) a. *MEN-o  
    b. men-o

Before closing this section, one particular issue implied by the current analysis that deserves attention concerns the Iraqi wh-form men ‘who’. Like other dialects, Iraqi men ‘who’ may occur independently of the subject pronoun. The data show that in wh-questions where men occurs, it always appears clause initially. Questions where the wh-phrase men occurs in situ are not fully well-formed. Consider the following examples.

(57) a. men Su’aad qabalat?  
    who Suad met.3fs  
    ‘Who did Suad meet?’

    b. *Su’aad qabalat men?  
       Suad met.3fs who  
       ‘Who did Suad meet?’

(58) a. men te’taqdoon ?enno Su’aad qabalat?  
    who think.2p that Suad met.3fs  
    ‘Who do you think that Suad met?’

    b. *te’taqdoon ?enno Su’aad qabalat men?  
       think.2p that Suad met.3fs who  
       ‘Who do you think that Suad met?’
The sentences above are interesting for two reasons. First, they show that men ‘who’ is a DP in its own right just like Lebanese miin ‘who’ and SA mann ‘who’. Secondly, the fact that men ‘who’ appears only clause initially entails that it always undergoes obligatory movement. Movement of men, however, is triggered by the need to check its uninterpretable [WH] feature with the interpretable [WH] feature on C, not because of an EPP feature on C. If it were, all wh-phrases must raise overtly to Spec,C in Iraqi in order to satisfy this feature. This is not the case however, since Iraqi meno and šeno occur in situ fully well-formed.

To sum up, this section has argued that the pronominal element associated with the Iraqi argument wh-phrase men-o is a subject pronoun. To this effect, I have presented evidence from several contexts, such as reduced cleft wh-questions, copular wh-constructions along with the agreement patterns associated with subject pronouns and their prosodical status. If this analysis is correct, it will have crucial implications for the syntax of Iraqi argument wh-phrases and the wh-questions in which they occur, to which I turn next.

2.4.2 meno is a copular wh-clause

In the previous section, I have shown that the pronominal element attached to Iraqi wh-expressions is a subject pronoun. This finding has crucial implications for the analysis of the internal structure of Iraqi wh-phrases and their distribution in wh-constructions. This section argues that the wh-expression men-o is a copular wh-clause with the schematic structure [WH-DP DP] just like subject-predicate constructions of the form [DP DP] discussed in the preceding chapter. As might be recalled, I have shown that certain types of predication in Arabic may be expressed without the presence of a verbal predicate, which I called verbless constructions23. Consider the following paradigms from Iraqi, Jordanian

---

23 Others refer to verbless constructions as nominal sentences (Fassi-Fehri 1993, Shlonsky 1997 and Ouhalla and Shlonsky 2002).
nian and Lebanese.

(59) Predicational [DP XP]

a. hu(wwe) mudarris
   he      teacher
   ‘He is a teacher’

b. hi(yye) mudiira
   she    director
   ‘She is a director’

c. haida fenneen
   this   artist
   ‘This is an artist’

(60) Equative [DP DP]

a. hu(wwa) l-mudarris
   he      the-teacher
   ‘He is the teacher’

b. hi(yye) l-mudiira
   she    the-director
   ‘She is the director’

c. haida l-fenneen
   this   the-artist
   ‘This is the artist’

(61) LD’ed Equative [DP [PRON DP]]

a. Ḥali hu(wwa) l-mudarris
   Ali  he      the-teacher
   ‘Ali is the teacher’

b. Majdi hu(wa) l-muhandis
   Majdi he    the-director
   ‘Majdi is the director’
c. Mhemmed hu l-m'alleh.
   Mhemmed he the-teacher
   ‘Mhemmed is the teacher’

The class of verbless constructions have been divided into simple subject-
predicate clauses and LD’ed equatives. As for subject-predicate clauses, these
have been shown to be either predicational with the schematic structure [DP
XP] as shown in (59) above, or equational which involve two definite DPs as
shown in (60). LD’ed equatives, on the other hand, have been shown to de-
rive from two-DP equatives and have the uniform structure [DP [PRON DP]]
as illustrated in (61).

The one construction that is directly related to the current discussion is the
one involving two definite DPs shown in (60) since it is the type of construction
that the wh-phrase men-o belongs to. Two DP constructions have been analysed
as subject-predicate clauses in which the initial DP is base-generated in Spec,TP
while the second DP is merged as complement of T as shown in (62) below.

(62)

```
TP
  \-----\     \-----\
  |      |     |      |
  DP     T'    DP
  \     |     |
   hu(wwa) T
   He  [+Present, +D]

l-mudarris
the teacher
```

More importantly, the wh-counterparts of sentences like (60) are also verbless
and differ only in that the first DP is a Wh-DP. These have the schematic

---

24 Recall that an XP may be either an indefinite DP, an AP or a PP.
25 Recall that under the current analysis DPs are treated as predicative like Heggie 1988,
   and unlike Bowers 1993. See Chapter 1 for details.
Consider the following examples.

(63)  

a. men hu(wwa)?  

   who he  

   ‘Who is he?’  

b. miin hi(yye)?  

   who she  

   ‘Who is she?’  

c. miin haida?  

   who this  

   ‘Who is this?’  

Notice that the sentences in (60) represent answers for the wh-questions in (63) in that the wh-phrases in (63a-c) represent the predicates in (60a-c). This is taken to indicate that the wh-phrase originates, first, in the predicate position and, then, undergoes movement across the subject to a clause initial position.

Assuming that movement of the wh-phrase is wh-movement to Spec,CP, the structure for a sentence like (63a) is shown in (64).

In (64), the wh-phrase men is shown to have undergone movement from its original predicate position to the Spec position of the copular clause. The
assumption that the wh-phrase men undergoes movement is in line with the observation made earlier about Iraqi wh-questions in which the wh-phrase men occurs. These facts constitute strong evidence that there is wh-movement in Iraqi. Recall, however, that wh-movement in Iraqi differs from classic wh-movement in that it is triggered by the need to check the [WH] feature on the wh-phrase, not because of an operator feature on C (see section 4.1).

The analysis outlined above can be straightforwardly extended to the Iraqi wh-expression men-o. Recall that although the Iraqi third person subject pronoun, i.e., -u/-o ‘he’, can cliticise onto the wh-phrase men ‘who’, its syntactic status as subject remains intact. In other words, the compound that is made up of the wh-phrase men ‘who’ and the subject pronoun -u/-o ‘he’ is a copular wh-clause of the form [Wh-DP DP]. Hence, it receives the same treatment as the questions in (63) and their corresponding structure shown in (64). The structure represented in (65) for men-o minimally differs from that in (64); the difference being essentially phonological.\(^\text{26}\)

(65) The structure of men-o

\[
\begin{array}{c}
\text{CP} \\
\text{DP} \\
\text{men} \\
\text{Who} \\
\text{C} \\
\text{TP} \\
\text{[+WH]} \\
\text{DP} \\
\text{T} \\
\text{[+Present, +D]} \\
\text{he} \\
\end{array}
\]

---

\(^{26}\)Notice that the analysis also allows for a derivation in which the wh-phrase is base-generated as a clausal subject and undergoes movement to Spec,CP.
Granting such analysis, we now have a clear picture for the internal structure of the Iraqi wh-phrase \textit{men-o}. Contra Wahba 1991, Ouhalla 1996 and Simpson 2000 who assume that \textit{men-o} is a simple wh-phrase, the analysis here shows that it is a copular wh-phrase. The same can be said to hold of the wh-expression \textit{\textacute{s}en-o} ‘what’ involving the Wh-DP \textit{\textacute{s}en} and the subject pronoun -\textit{o} ‘it’, though a more detailed decomposition of its internal structure will be provided in the following chapter.

Two types of evidence can be obtained in support for the copular analysis of the Iraqi wh-phrase \textit{men-o}. First, the wh-phrase \textit{men-o} is used independently in the language as a full copular wh-question.

\begin{exe}
  \item \textit{a.} men-o?
  \begin{exe}
    \item who-he
    \begin{exe}
      \item (Iraqi)
      \item ‘Who is he?’
    \end{exe}
  \end{exe}

  \item \textit{b.} men-i?
  \begin{exe}
    \item who-she
    \begin{exe}
      \item (Iraqi)
      \item ‘Who is she?’
    \end{exe}
  \end{exe}
\end{exe}

The second piece of evidence comes from consideration of wh-expressions in other dialects. That is to say, the claim that the wh-expression \textit{men-o} is a copular wh-clause is strongly supported by the Moroccan wh-expression \textit{\textacute{s}-kuun} ‘who’, albeit this expression contains a wh-element and a verbal copula carrying default subject agreement, i.e., third person masculine. Consider the following question from Moroccan.

\begin{exe}
  \item \textit{\textacute{s}-kuun m\textacute{s}a?} (Morrocan)
  \begin{exe}
    \item who \quad left.3ms
    \begin{exe}
      \item ‘Who left?’
    \end{exe}
  \end{exe}
\end{exe}

According to Ouhalla (2002), the wh-expression \textit{\textacute{s}-kuun} consists of the wh-element \textit{\textacute{s}}, and inflection (Ouhalla 2002). This inflection is in fact the present
tense form of the verbal copula *kun* ‘be’ in Arabic. Putting the question of derivation aside for the moment, the point to be made is that copular wh-phrases do exist, and are, as a matter of fact, common, across the dialects.

To summarise, I have provided an analysis showing that *men-o* is a clausal wh-phrase, i.e., a CP, in which the pronominal element is a subject and the wh-phrase is a predicate. The CP analysis of *meno* gives rise to an interesting question, namely, how does *men-o* appear in a DP position in Iraqi wh-questions? This question is dealt with in the next section.

### 2.5 Derivation of *men-o* wh-questions

This section entertains the question of why, or rather, how a CP occurs in a DP position, and behaves like one, in Iraqi wh-questions. I am going to argue that the wh-expression *men-o* is a free relative projecting as DP. I base my argument on Citko’s (2008) analysis of free relatives whereby movement of the wh-phrase projects the label of the wh-phrase. In the case of the Iraqi wh-expression *men-o*, I take internal movement of the wh-phrase *men* to result in projecting the label of the wh-phrase, i.e., a DP. Respectively, the structures for free relative constructions and the construction of the wh-expression *meno* will look like those in (68) and (69).

(a) Bill reads *what(ever) Tom reads*.

(b) The derivation of *s-kuun* will be dealt with in the next chapter.
Raising the wh-phrase *men* in (69b) essentially turns it into a free relative construction in a similar way to that in (68b). The difference between *men-o* and normal free relatives reduces to the [WH] feature on the DP *men-o*. Thus, the wh-expression occurring in argument position in Iraqi wh-questions is a free relative bound in situ by the matrix operator of the interrogative clause. Relativisation in this particular instance is taken to effectively be d-linking, which, in turn, licenses wh-in-situ (Pesetsky 1987). The analysis has important implications for the analysis of Arabic wh-in-situ, especially the view that languages like Iraqi use relativisation as a strategy in formulating wh-in-situ constructions.

2.5.1 External syntax of the wh-expression *men-o*

The conclusion that the wh-expression *men-o* is a CP is problematic for Iraqi wh-questions where *men-o* has the distribution of DPs. For instance, in matrix wh-questions like (70) below, *men-o* sits in an argument position.

(70) Mona shaafat *men-o*?  
Mona saw.3fs whom  
‘Who did Mona see?’
Under the analysis proposed here, the wh-expression *meno* is treated as a CP. However, wh-questions like (70) show that the wh-expression *meno* occurs in a position designated for DPs, that is, the object argument position of the verb *shaafat* ‘saw’. To the extent that the wh-expression *meno* is indeed a CP, there appears to be a mismatch between its syntax and semantics, which must be reconciled.

The apparent conflict between the syntax and semantics of the wh-expression *meno* is not unique. There exist similar constructions in many languages where these constructions are CPs and have the distribution of DPs. This can be illustrated in the following examples from English.

(71)  
   a. I know **who he is**.  
   b. I saw **who he was**.

In (71), the embedded clauses *who he is* and *who he was* are clearly CPs which involve internal movement of the wh-phrase *who* from its predicate position to Spec,CP of the embedded clause. Consider, for example, the structure for the clause *who he was* which might be represented in (72).

(72)  
$$
\text{CP} \\
\text{who} \quad \text{C'} \\
\text{C} \quad \text{TP} \\
\text{DP} \quad \text{T'} \\
\text{I} \quad \text{I} \\
\text{he} \quad \text{T} \quad \text{DP} \\
\text{I} \quad \text{I} \\
\text{was} \quad \text{(who)}
$$
Movement of the wh-phrase *who* in the structure above results in a free relative construction. It looks like it is a CP, but actually, it has a DP reading. This is the kind of conflict we see in *meno* questions like (70) above. In structural terms, the clause *who he is* is similar to the wh-expression *men-o*. Under the current analysis, wh-expression *men-o* is assigned the structure in (73). Compare (72) and (73).

(73) CP
     /\                   /\                     /\
    DP     C'            TP                     TD
     |                   |                     |                \meno
   men    C             T  \meno
   who     DP           T'
           |               |
          hu               he
                     \meno
                        who

Structurally, what the embedded clauses from the English cases above and the Iraqi wh-expression *meno* have in common is that they are essentially free relative constructions which involve internal movement of a wh-phrase. In terms of interpretation, it appears that the interpretation of English free relatives is more restricted than that of the Iraqi wh-words under consideration. Nevertheless, the problem both English free relatives and Iraqi wh-words pose is, more or less, the same, namely, the CP/DP conflict characteristic of free relatives.

\[28\] The crucial difference between free relative and *meno* is the [WH] feature on *meno*, as will be discussed shortly.
2.5.2 The DP projection of free relatives

A Brief Overview.

Free relatives are the classic case for semantics-type clash: they are CPs with the semantics of nominal expressions. To solve the conflict, a number of proposals are entertained in the literature on free relatives which share the assumption that free relatives have a DP structure. For instance, the Head Account (Bresnan and Grimshaw 1978) and the Comp Account (Groos and van Riemsdijk 1981) assign the free relative in (74a) the structures in (74b) and (74c), respectively.

(74) a. John plays whatever he likes.

b. Head Account:

```
          DP
         /   \
       DP    CP
      /     |
 Whatever DP  TP
    /     |
   0     he likes (whatever)
```

c. Comp Account

```
          DP
         /   \
       DP    CP
      /     |
   0     DP   TP
    /     |
 whatever he likes (whatever)
```

Although the wh-phrase and the empty head occupy different positions under each of the two accounts in (74b) and (74c), the bottom line is that free relatives are DP projections. Both the Head account and the Comp account...
have their own problems, which is not a matter of concern here\textsuperscript{29}. For the purposes of this study, the analysis for free relatives I am going to follow is the one advanced in Citko 2008. Presenting her main arguments, I will show how Citko’s analysis incorporates some of the main assumptions behind the current discussion.

\textit{Citko’s analysis (2008)}

In her study, Citko (2008) examines projection possibilities created by the operation Merge, the basic structure building operation of the minimalist program (Chomsky 1995, 2001, 2001a, 2001b and 2005). There are two instances of Merge: External Merge and Internal Merge (i.e., Move). External Merge takes two objects $\alpha$ and $\beta$ and groups them into a new syntactic object K (75a). This is the case for initial Merge of the wh-phrase men with T, as illustrated in (75b)\textsuperscript{30}.

\begin{equation}
(75)\quad a.\quad K
\end{equation}

\[\quad \alpha \quad \beta\]

\begin{equation}
(75)\quad b.\quad T'
\end{equation}

\[\quad T \quad \text{men} \quad \text{who}\]

The other instance of Merge is Internal Merge, which is essentially Move, where $\beta$, a subpart of $\alpha$, is re-Merged as a specifier of $\alpha$ thereby creating a new syntactic object K. This is again the stage at which the wh-phrase men undergoes movement to Spec,CP, where K, under standard assumptions, is labelled as CP (76).

\textsuperscript{29}For a detailed discussion see Larson 1998, Donati 2006 and Citko 2008).

\textsuperscript{30}Recall that T in Arabic copular constructions has an abstract T(ense) feature and must, therefore, be present in the structure.
Until recently, it has always been assumed that Internal Merge of the wh-phrase to Spec,CP in wh-clauses results in a CP projection, as is the case in (76). However, Chomsky (2005) suggests that the labelling algorithm applies freely when the operation Merge is at work. Building on Chomsky (2005), Citko argues that merging two objects $\alpha$, a Probe, and $\beta$, a Goal, results in four projection possibilities: project $\alpha$, project $\beta$, project both $\alpha$ and $\beta$, and project none. According to Citko, not only are these projections logically possible, but they are also attested in different languages (see Citko (2008) for a full discussion on the four projections).

Of the four possible projections which might result from syntactic Merge, there is one possibility that is directly related to the present discussion, namely, Project $\beta$. Project $\beta$ results from Internal Merge, i.e., movement of $\beta$, as shown in (77), and is argued here to be the case for Internal Merge of the Iraqi argumental wh-DP men.

Returning to the classic CP/DP conflict that free relatives display, it turns out to provide evidence in favour of Project $\beta$. Under the assumption that
movement of the wh-phrase is an instance of Project $\beta$, the resulting structure projects the label of the wh-phrase. Consequently, the DP-status of free relatives is straightforwardly accounted for, along with other properties associated with free relatives in general, which remain problematic for other accounts, i.e., the Comp Account and Head Account (Citko 2008). Having established this background, now consider the structure in (78b) assigned by Citko (2008) for sentences like (78a) which involves a free relative occurring in argument position.

(78)  

(a) Bill reads what(ever) Tom reads.

(b) 

\[
\text{DP} \quad \text{DP} \quad \text{CP} \\
\text{I} \quad \text{whatever} \quad C \quad \text{TP} \\
\text{Tom reads (whatever)}
\]

2.5.3 A DP analysis for Arabic free relatives

The proposal that \textit{men-o} is a free relative in situ is bolstered by facts of free relatives in Arabic where relativization is used as a means of linking referents to discourse. Consider the following sentences which involve free relatives occurring in the same distributional contexts as DPs\textsuperscript{31}.

(79)  

(a) wasala lladhi hadda$\beta$-tani 'an-hu (SA)  
arrived-he RM.the-he talked-you-to-me about-him  
‘The one you talked to me about has arrived’

(b) l-jmud darabu ?illi habasu-u. (Palestinian)  
the-army hit that (they)-arrested-him  
‘The army beat up who they arrested.’

\textsuperscript{31}SA examples are from Farghol 1986 and Ouhalla 1999; Moroccan examples are from Ouhalla 1999; Palestinian examples are from Shlonsky 2002; and Jordanian examples are my own.
The data above show that there are two types of free relative construction in Arabic: *?illi* free relatives and *ma* free relatives. *?illi* free relatives (79a-c) involve no wh-elements, hence no wh-movement. Whereas, *ma* free relatives (80a-c) involve a fronted wh-element much in the same way as the English cases discussed above.

Concentrating on *ma* free relatives of the type shown in (80), the controversy surrounding their analysis, i.e., whether the wh-element moves to Spec,CP or whether it is base-generated as head of the free relative construction, mirrors that just outlined for their English counterparts. For instance, Farghal (1986) adopts the Head Account of Bresnan and Grimshaw (1978), whereby free relatives are derived by base-generating the wh-element as head of the free relative (see the structure in (74b)). Whereas, Shlonsky (2002) argues for the Comp Account of Groos and van Riemsdijk (1981) where *ma* free relatives are derived by movement of the wh-expression from the base argument position to Spec,CP (see the structure in (74c)).
I will extend Citko’s analysis for Arabic and assume that free relatives are derived by movement of the wh-phrase and project the label of the wh-phrase. Evidence that *ma free relatives involve wh-movement comes from Subjacency effects, as shown in (81) from Palestinian (Shlonsky 2002)32.

\[(81) *ma-\text{\`imilt-š }\text{šu}u \text{ ma } \text{\`irift } l-mara \text{ \`illi } \text{\`imlat Neg-(I)did-Neg what that (I)knew the-woman that did}
\text{‘I didn’t do what I knew the woman who did’}\]

Examples like (81), where movement of the wh-expression across the boundary of a relative clause, i.e., Complex NP island, is illicit, are taken to indicate that free relatives are formed by wh-movement. This is in line with the fact that Palestinian is a wh-movement language, as shown in (82).

\[(82) \begin{align*}
a. & \text{miin l-\text{\`asad } akal mbaarih?} & \text{(Palestinian)} \\
& \text{who the-lion ate yesterday} \\
& \text{‘Who did the lion eat yesterday?’} \\
b. & \text{\`suu } \text{\`inti katabti mbaarih?} \\
& \text{what you(F) wrote yesterday} \\
& \text{‘What did you write yesterday?’}
\end{align*}\]

Given these observations, Palestinian free relatives, therefore, can be given a structure like that assigned by Citko 2008 for English, i.e., a structure where movement of the wh-phrase projects the label for the wh-phrase as in the structure below.

\[(83) \begin{array}{c}
\text{DP} \\
\downarrow \\
\text{DP} \quad \text{CP} \\
\downarrow \\
\text{šu}u \quad \text{C} \quad \text{TP} \\
\downarrow \\
\text{ma} \quad \text{\`inti } \text{\`imilt } (\text{šu}u)
\end{array}\]

32I am leaving ?illi free relatives out of the discussion since they do not involve a wh-element, only the complementiser ?illi.
If the analysis just outlined is correct, the CP/DP status of the wh-expression men-o can be straightforwardly accounted for. The fact that free relatives are common in the dialects and the fact that they occur in argument position give substance to the claim that the wh-expression men-o is an instance of a free relative. Concerning the structure of the wh-expression men-o, the apparent conflict in the CP/DP status of the wh-expression men-o can be neatly explained. The present analysis allows for a derivation whereby the wh-phrase men, a DP, undergoes Internal Merge with Spec,C and projects as the label of the newly formed constituent, as illustrated in (84).

(84)\[\text{DP}\]
\[\text{DP}\]
\[\text{CP}\]
\[\text{L}\]
\[\text{men}\]
\[\text{C}\]
\[\text{TP}\]
\[\text{Who}\]
\[\text{-o }\langle\text{men}\rangle\]
\[\text{he }\langle\text{who}\rangle\]

There is, however, one crucial difference between the interpretation of free relatives and men-o. Recall that movement of the wh-phrase men is driven by the need to check its [WH] feature, which means that the DP must receive a wh-reading. This is precisely what distinguishes free relatives from in-situ men-o in Iraqi wh-questions. Namely, men-o is a [+WH] free relative DP. As such, men-o can occur in-situ in the presence of an operator in scope position, which binds it in that position.

2.5.4 Iraqi facts explained

We are now in a position to straightforwardly account for the facts from Iraqi wh-questions involving in-situ men-o. The fact that it occurs in-situ follows from
its status as a DP with a [WH] feature. Consequently, *meno* is bound in situ by the matrix Question operator in Iraqi wh-questions. A full representation for the wh-question in (85) is given (86).

(85)    Mona shafat *meno?*  
Mona saw whom  
‘Who did Mona see?’

(86)

In the structure (86), Internal Merge re-merges the wh-phrase *men* with C and projects the label of the wh-phrase thereby forming a DP projection. Under the present analysis, movement of *men* turns the structure into a free relative con-
struction. Interpretively, there are two issues that may bear mentioning here. First, there is a fundamental difference between normal free relatives and the in-situ wh-clause *meno* in that the latter is truly [WH]. As such, *meno* needs a binder and that would be the [+WH] matrix Operator. Consequently, binding of the whole Wh-DP *meno* would give the question above the following interpretation: *for which person x, Mona saw x?* Notice that this interpretation is similar to the interpretation of d-linked wh-questions in which the set of referents that the wh-phrases asks for is already established in discourse. This is exactly what relativization does in this type of questions in Iraqi. The second issue concerns the interpretation of *meno*-questions, as opposed to the interpretation of English free relatives. Recall that under the present analysis *meno* is analysed structurally as a free relative on a par with English free relatives. In terms of interpretation, however, it can be pointed out that there are limitations on the interpretation of English free relatives but with no exact parallel to the Iraqi wh-phrases.

The analysis just outlined neatly explains the contrast between the ungrammatical cases that involve in-situ *men*, repeated below, and those that involve in-situ *men-o* such as (85). Recall that the wh-phrase *men* obligatorily moves in this language. In cases like (86), the wh-phrase *men* moves within its own structure thereby satisfying the need to check its [WH] feature. By contrast, the [WH] feature of *men* fails to be checked in the syntax in (87). In addition, the wh-phrase *men* is not d-linked and, hence, unlike *meno*, cannot be bound in situ.

(87)  
\begin{enumerate}
  \item *Su'aad qaabalat men?*
    Suad met.3fs who
    ‘Who did Suad meet?’
  
  \item *te'ttaqdoon yenn Su'aad qabalat men?*
    think.2p that Suad met.3fs who
    ‘Who do you think that Suad met?’
\end{enumerate}
As concerns the ill-formed cases which involve the occurrence of the wh-phrase meno inside an embedded tensed clause or inside an island, it looks like binding by the Q-operator of the in-situ meno is different from standard wh-in-situ cases such as Chinese (Huang 1982 and Aoun and Li 1993b) in that such a dependency cannot be long-distance. This, in turn, can be attributed to the morphological facts surrounding Iraqi wh-phrases. Arabic wh-phrases, including Iraqi wh-phrases can only have an interrogative reading (see Ouhalla 1996 and Aoun and Li 2003). By contrast, Chinese wh-phrases are polarity items which behave like free variables in need of binding by a base-generated operator that is not part of the morphological makeup of the wh-phrase (Aoun and Li 1993b). Hence, the fact that Chinese in-situ wh-phrases are variables makes direct binding by a null operator possible across clausal boundaries. On the other hand, because Iraqi wh-phrases are morphologically composed of both the operator and the variable, it could be the case that movement of the operator across clause boundaries results in ill-formedness.

Finally, recall that Iraqi wh-questions may also involve the occurrence of the wh-phrase meno clause initially. Under the current analysis, meno is said to contain a subject pronoun forming a copular construction of type found in non-verbal predications in Arabic. Anticipating the results of the analysis of such constructions in chapter four ahead, my conjecture concerning questions that involve meno in the initial position of the clause is that such cases are instances of focus in which the whole copular clause is base-generated in a focus position in the left periphery.

33See Simpson 2000 for a different analysis.  
34For a different view about islands and the distinction between types of islands see van de Koot and Mathieu 2003 who argue that weak island and strong islands represent two different phenomena. More recently, Adger and Ramchand 2005 and Boeckx and Hornstein 2008 argue that islands can no longer be used as diagnostics for movement.  
35The Iraqi situation is reminiscent of French wh-in-situ which has been reported to have a very restricted distribution in that it is limited to matrix clauses (see, for instance, Mathieu 1999, Starke 2001 and Cheng 2003).
To sum up, I have shown in the sections above that Iraqi *meno* has the structure of a free relative with the difference reducing to the [Wh] feature that the wh-phrase *meno* has. This was taken to indicate that Iraqi uses relativisation to construe wh-in-situ questions. More precisely, relativisation is seen here as a means of d-linking which is responsible for the licensing of wh-in-situ via binding. The binding of Iraqi wh-in-situ, is, however, different from unselective binding of Pesetsky (1987). The former, but not the latter, is sensitive to clause boundaries.

2.6 Conclusion

This chapter has provided a morpho-syntactic analysis for the Iraqi wh-expression *meno* ‘who’. I began by presenting the core problem associated with wh-in-situ constructions in both Iraqi and Lebanese. In particular, I have shown that the distribution of wh-phrases in the two dialects is asymmetric to the extent that it cannot be addressed in accordance with the wh-parameter. That is, the assumption that languages follow a strict typological division in that they are either plus or minus wh-movement languages was called into question by the data presented here. To this effect, I have discussed two types of approach to wh-in-situ, namely the LF-movement approach (Huang 1982) and the binding approach (Pesetsky 1987), and have shown indeed that a parametric approach is too broad to capture the facts from Arabic.

As a point of departure, I have suggested an alternative approach that takes, seriously, the morpho-syntax of individual wh-phrases to determine their distribution. Before putting forward my proposal, I looked into a particular approach in the literature that deals with Iraqi wh-constructions on the basis of the morphological properties of argument wh-phrases, namely that of Ouhalla 1996. Upon closer inspection, however, it was revealed that the analysis of Ouhalla suffered from a major drawback at the heart of the analysis. An alternative
analysis was then suggested.

The proposed analysis started from the observation that the pronominal element attached to the wh-phrase *men* ‘who’ cannot be a normal object clitic, as originally suggested in Ouhalla (1996). I have argued that this pronominal is in fact a reduced version of the subject pronoun *hu(wwa)* ‘he/it’. Evidence from equative wh-constructions led to the conclusion that the wh-expression *men-o* is underlyingly propositional and has a CP structure. The CP analysis proposed for *men-o* is supported by a similar phenomenon from the Moroccan wh-expression *škuŋ* which constitutes a wh-element and the default form of the verbal copula. A question was then raised with regards to the occurrence of the wh-expression *men-o* in DP positions in Iraqi wh-questions. Based on observations about free relatives in Arabic, I have argued that *meno* is a free relative DP with a [WH] feature and that Iraqi uses relativization as a stratagem for licensing wh-in-situ, hence, the occurrence of *meno* in a DP position in Iraqi wh-questions.

The analysis advanced here for the Iraqi wh-expression *men-o* provides us with a unique, though complex, picture for the type of wh-in-situ involved in this language. The analysis has the advantage over other analyses in that it starts off with a micro-parametric approach to elements that make up the structure of Arabic wh-expressions. Arabic wh-expressions appear to have hybrid and intricate syntax, unlike what is standardly assumed, and only when studying properties of each wh-expression, can we explain parametric settings for each language. The analysis proposed here will also be extended in the following chapter to the other Iraqi wh-expression *šên-o* and the Lebanese wh-expression *šu* which will be shown to be syntactically more complex than the wh-phrase *meno*. Overall, the analysis will provide a uniform solution for the core problem observed at the outset as well as account for variation facts from the Arabic dialects.
Chapter 3

What is ‘Which thing is it?’

3.1 Introduction

The core problem this chapter is concerned with is the asymmetric variation between argument wh-phrases in Lebanese introduced in Aoun and Choueiri 1999 and Aoun, Benmamoun and Choueiri 2010. In particular, bare argument wh-phrases exhibit contrastive behaviour in wh-in-situ questions. Meanwhile, there is no such contrast in wh-movement questions. This is shown in (1) and (2) below.

(1) a. šeft **miin** mbeeriḥ?
saw.2sm who yesterday  
‘Who did you see yesterday?’

b. *štarayte **šu** mbeeriḥ  
bought.2sf what yesterday  
‘What did you buy yesterday?’

(2) a. **miin** šeft mbeeriḥ?  
who saw.2sm yesterday  
‘Who did you see yesterday?’

b. **šu** štarayte mbeeriḥ  
what bought.2sf yesterday  
‘What did you buy yesterday?’
The problem posed by the data above lies in the contrast between the argument wh-phrases *miin* ‘who’ and *šu* ‘what’. More specifically, whereas both argument wh-phrases can appear clause initially, (2a) and (2b), only the argument wh-phrase *miin* ‘who’ occurs in-situ fully well-formed as in (1a), but not the argument wh-phrase *šu* ‘what’ (1b). The contrast illustrated by these examples is unusual given standard assumptions about the uniform behaviour of the class of argument wh-phrases (Huang 1982; Pesetsky 1987; Rizzi 1990; Wahba 1991; Chomsky 1977, 1995; Ouhalla 1996; Shlonsky 2002; among others).

To solve the problem posed by Lebanese argument wh-phrases above, Aoun, Benmamoun and Choueiri 2010 propose an account based on Pesetsky’s 1987 theory of d-linking. They argue that the source of the asymmetry between the argument wh-phrases above is directly related to the (in)ability of a wh-phrase to be d-linked. That is, whereas the argument wh-phrase *miin* ‘who’ can be d-linked and may, therefore, stay in situ, the argument wh-phrase *šu* ‘what’ cannot. They further claim that the inability of the wh-phrase *šu* ‘what’ to be d-linked requires it to always undergo overt movement to the front of the wh-interrogative.

This chapter argues, however, that d-linking is irrelevant to the contrast in (1) and (2) and proposes, instead, a syntactic analysis that will be based on micro-differences between wh-phrases. The analysis proposed, while successfully avoiding the problems encountered in Aoun, Benmamoun and Choueiri’s 2010 analysis, explains all the facts surrounding the asymmetric variation in Lebanese wh-questions.

The sections of this chapter are organised as follows: after spelling-out the proposal in section (2), section (3) presents Lebanese wh-in-situ and concentrates on the asymmetry between argument wh-phrases. Section (3.1) investigates properties of argument wh-phrases in various other contexts, including
d-linked wh-questions, resumptive wh-questions and multiple wh-questions, and reveals that the source of the asymmetries resides mainly in the behaviour of the wh-phrase šu ‘what’. Section (3.2) looks into Aoun, Benmamoun and Choueiri’s 2010 analysis and argues that their d-linking approach is irrelevant to the asymmetries between Lebanese argument wh-phrases. Section (4) develops a syntactic analysis for the wh-phrase šu ‘what’ based on data and observations from other Arabic dialects. Finally, section (5) concludes with the main findings of this chapter.

3.2 Proposal

I propose the following structures for Lebanese argument wh-phrases miin ‘who’ (3) and šu ‘what’ (4).

(3) miin ‘who’

```
miin 'who'
   \-----
     wh
     DP
```

(4) šu ‘what’

```
šu 'what'
    \-----
      CP
      \-----
        DP
        \-----
          C'
          \-----
            C
            \-----
              T
              \-----
                TP
                \-----
                  DP
                  \-----
                    T'
                    \-----
                      DP
                      \-----
                        hu
                        \-----
                          T
                          \-----
                            ñey-š
                            \-----
                              which thing
                              \-----
                                [+WH]
                                \-----
                                  TP
```

133
In the first place, I will present data from Lebanese wh-in-situ questions and follow Aoun and Choueiri 1999 and Aoun, Benmamoun and Choueiri 2010 in assuming that the problem lies in the behaviour of the wh-expression šu ‘what’. This will be followed by an investigation of various contexts and identify other properties of argument wh-phrases in Lebanese.

Secondly, I will argue against the claim that the reason why the wh-expression šu ‘what’ cannot occur in situ is a consequence of its inability to be d-linked (Aoun and Choueiri 1999 and Aoun, Benmamoun and Choueiri 2010) and show that such a claim is problematic. On an empirical level, it will be shown that the contrast in single wh-in-situ questions is a subcase of the overall asymmetry between Lebanese wh-arguments, i.e., šu ‘what’ vs. miin ‘what’, found in other wh-constructions, including resumptive wh-questions and multiple wh-questions. For instance, the wh-phrase šu ‘what’ cannot also appear in-situ in multiple wh-questions although d-linking is not relevant to the licensing of in-situ wh-elements. The data further show that where d-linking generally licenses wh-in-situ in contexts that otherwise would be ungrammatical, i.e., superiority contexts, it does not save Lebanese wh-in-situ.

Finally, I develop a syntactic approach which will be essentially based on the structural properties of argument wh-phrases. Appealing to novel empirical evidence from other dialects, I will argue that whereas the argument wh-phrase miin ‘who’ has the simple DP structure in (3), the wh-expression šu ‘what’ is a copular clause, derived from the string ?eyš hu ‘which thing it’, and has the CP structure in (4), which involves internal movement of the wh-element from the predicate position to Spec CP. Given the CP analysis of šu, I turn, afterwards, to addressing the question of the syntax of šu-questions and argue, along the lines of Koster’s 1978 and Alrenga’s 2005 analysis of clausal subjects, for a base-generation analysis, contra Aoun and Choueiri 1999 and Aoun, Benmamoun and Choueiri 2010, who argue for a movement analysis. The structure that will be
assigned for šu-questions like (5a) is the one shown in (5b) below.

(5) a. šu štarrayte mbeerih
    what bought.2sf yesterday
    ‘What did you buy yesterday?’

b.

The proposed analysis will have a number of consequences. First, the reason why šu is always clause-initial is because it is only generable in the specifier position of a [+WH] C, i.e., it is directly merged in Spec,CP. From this position, I will further argue that šu binds a null DP in argument position. This means that šu is comparable to a wh-operator, which makes šu-questions a species of null operator constructions discussed in Chomsky 1977; albeit, this operator is an overt complex wh-operator. The analysis advanced in this chapter is, there-
fore, purely syntactic and has nothing to do with d-linking as proposed in Aoun, Benmamoun and Choueiri 2010. As a result, all the problems associated with the d-linking approach are successfully avoided. A second consequence is that šu can never show up in DP distributional contexts. This would, then, explain the asymmetries found in Lebanese wh-in-situ questions, d-linked wh-questions and resumptive wh-questions. Another related consequence is that šu is well-formed in-situ in contexts where a verb selects for a CP.

Before concluding the chapter, I discuss one remaining issue to do with the syntax-phonology interface. Specifically, I raise the question as to how the string ʿeyš hu ‘which thing it’ is spelled out as šu, given that it is a left-branching specifier embedded in a configuration as complex as (5b). I adopt both Starke’s recent Phrasal Spell-Out approach that is, in turn, based on the Nanosyntax framework (Starke 2011) as well as Uriagereka’s 1999 model of Multiple Spell-Out. The core assumption leading this part of the analysis is that syntactic derivation spells out one of the complex structures before it merges with the other. Thus, before the string ʿeyš hu merges with the rest of the tree, it undergoes early spell-out to the phonological interface and re-enters the derivation as a frozen lexical compound. As for the ordering of the terminals within the spelled out structure, I follow Uriagereka’s suggestion that it may be fixed by the Linear Correspondence Axiom of Kayne 1994. The resulting structure is roughly represented in (6) below.
3.3 Lebanese wh-in-situ

The aim of this section is to re-examine Lebanese wh-in-situ and see whether an approach that is based on d-linking, such as that outlined in Aoun and Choueiri 1999 and Aoun, Benmamoun and Choueiri 2010, can adequately capture the contrast between bare argument wh-phrases. Using data from Lebanese and similar dialects, I will show that the asymmetric behaviour of Lebanese argument wh-phrases extends to contexts other than single wh-in-situ constructions. I will argue that the asymmetries observed here must be accounted for not in terms of d-linking, but, rather, in terms of structural properties of wh-phrases. To this effect, the arguments put forward in Aoun and Choueiri 1999 and Aoun, Benmamoun and Choueiri 2010 are rendered irrelevant.
To begin, it is noteworthy that Lebanese is one of four Arabic dialects that make up the Levantine family, i.e., Palestinian, Syrian and Jordanian (see Shlonsky and Ouhalla 2002 and Aoun, Benmamoun and Choueiri 2010). Whereas wh-movement is a common strategy for question formation between these languages, wh-in-situ is reported only in Lebanese (Aoun and Choueiri 1999, Aoun and Li 2003, Aoun Benmamoun and Choueiri 2010, among others). Consider the following sentences.

(7)  
\begin{align*} 
\text{a.} & \quad \text{šeft } \text{?ayya mmasil mbeeri}\text{?} \quad \text{(Lebanese)} \\
& \quad \text{saw.2sm which actor yesterday} \\
& \quad \text{‘Which actor did you see yesterday?’} \\
\text{b.} & \quad \text{nada } \text{?aali} \text{t } \text{?anno zeena šeefit } \text{?ayya mmasil b-l-maT}\text{?am}? \\
& \quad \text{Nada said.3sf that Zeina saw.3sf which actor in-the-restaurant} \\
& \quad \text{‘Which actor did Nada say that Zeina saw in the restaurant?’} \\
\text{c.} & \quad \text{btayrfo l-mara yalli šeefit } \text{?ayya mmasil b-l-maT}\text{?am} \\
& \quad \text{know.2p the-woman that saw-3sf which actor in-the-restaurant} \\
& \quad \text{‘Which actor do you know the woman that saw in the restaurant?’} 
\end{align*}

(8)  
\begin{align*} 
\text{a.} & \quad \text{šeft } \text{miin mbeeri}\text{?} \\
& \quad \text{saw.2sm who yesterday} \\
& \quad \text{‘Who did you see yesterday?’} \\
\text{b.} & \quad \text{nada } \text{?aali} \text{t } \text{?anno zeena šeefit } \text{miin b-l-maT}\text{?am}? \\
& \quad \text{Nada said.3sf that Zeina saw.3sf who in-the-restaurant} \\
& \quad \text{‘Who did Nada say that Zeina saw in the restaurant?’} \\
\text{c.} & \quad \text{btayrfo l-mara yalli šeefit miin b-l-maT}\text{?am} \\
& \quad \text{know.2p the-woman that saw-3sf who in-the-restaurant} \\
& \quad \text{‘Who do you know the woman that saw in the restaurant?’} 
\end{align*}

The sentences in (7) and (8) show that argument wh-phrases like \text{miin} ‘who’ and \text{?ayya mmasil} ‘which actor’ may occur fully well-formed in-situ in simple wh-questions (8a and 9a), in embedded wh-questions (7b and 8b) and in wh-
questions where the wh-phrase occurs inside an island, i.e., a relative clause island, (7c and 8c). According to Aoun and Choueiri 1999, Aoun and Li 2003 and Aoun, Benmamoun and Choueiri 2010, Lebanese wh-in-situ is d-linked, i.e., a wh-phrase in situ must be d-linked. In terms of interpretation, Aoun et al. extend Pesetsky’s 1987 treatment of d-linked wh-in-situ in multiple wh-questions to Lebanese single wh-in-situ questions. According to Pesetsky, there are two types of wh-phrase: d-linked and non-d-linked. Only d-linked wh-phrases may stay and get licensed in situ via unselective binding; whereas, non-d-linked must undergo wh-movement\(^1\). It follows that d-linked wh-phrases, such as ʔayya mmasil and miin in (7-8), receive interpretation via binding by the root [+Q] Comp. In particular, the fact that the wh-phrases in (7c) and (8c) have wide scope over the entire sentence follows from the binding relationship between the wh-phrase and its binder, even though the wh-phrases occur inside islands (Aoun and Choueiri 1999, Aoun and Li 2003 and Aoun, Benmamoun and Choueiri 2010)\(^2\).

Lebanese wh-in-situ is not uniform, however. Putting aside, for the moment, lexically d-linked ʔayya-NP ‘which’-NP phrases, there is a sharp contrast between the argument wh-phrase miin ‘who’, shown in (7) above, and the wh-phrase ʔu ‘what’ shown in (9) below.

(9) a. *ʔstriito ʔu mn-l-mahall?
   bought.2p what from-the-store
   ‘What did you buy from the store?’

   (Lebanese)

b. *byiftikro anno ʔstriito ʔu mn-l-mahall?
   think.3p that bought.2p what from-the-store
   ‘What do they think that you bought from the store?’

\(^1\)Pesetsky treats non-d-linked wh-phrases as real quantifiers that undergo obligatory movement to attain scope while d-linked wh-phrases are treated as indefinites, in the sense of Heim 1982, that are unselectively bound in situ by an appropriate Comp.

\(^2\)Note that a wh-phrase may be either morphologically d-linked like ʔayya mmasil ‘which actor’ in (7a-c) above, or contextually d-linked, i.e., without a visible morpho-syntactic reflex, like bare wh-phrases like miin ‘who’ in (8a-c), understood to be linked to previous discourse.
Unlike \textit{miin} ‘who’ in sentences (8a-c), the wh-expression \textit{šu} ‘what’ cannot occur in situ, be it in simple wh-questions (9a), in embedded wh-questions (9b) or in wh-questions with islands containing the wh-phrase (9c). Such contrast is not expected given standard assumptions that argument wh-phrases \textit{who} and \textit{what} behave uniformly (Huang 1982, Chomsky 1986, Rizzi 1990, Cinque 1990 and Ouhalla 1996).

The non-uniform behaviour of Lebanese \textit{miin} ‘who’ and \textit{šu} ‘what’ raises several questions for which this study attempts to provide a solution. The first question is whether this contrast is found only in wh-in-situ questions, and if so, why? If the contrast is found in other wh-contexts, how can it be explained? Can it be accounted for within existing approaches? And finally, whether such contrast is limited to Lebanese or whether it exists in other languages? As far as Lebanese is concerned, an attempt has been made in Aoun and Choueiri 1999 and Aoun, Benmamoun and Choueiri 2010 to solve the contrast observed here based on Pesetsky’s 1987 theory of d-linking. Before discussing this account, however, the next section addresses the question of whether the asymmetry between Lebanese argument wh-phrases is limited to single wh-in-situ constructions or whether it extends to other wh-constructions.

3.3.1 Surface distribution of the wh-expression \textit{šu}

This section provides a descriptive account for the behaviour of Lebanese argument wh-phrases in constructions other than single wh-in-situ questions. The data presented in this section show that the wh-expression \textit{šu} ‘what’ has a mixed distributional pattern in the sense that, in some contexts, it has a different distribution from that of other argument wh-phrases; while, in other contexts, it
appears in distributional positions where an argument wh-phrase is expected to appear. Overall, the wh-expression šu displays a behaviour that distinguishes it from argument wh-phrases which have a uniform distribution in the different types of wh-constructions.

We have seen in the previous section that, unlike the argument wh-phrase miin ‘who’, the wh-expression šu ‘what’ cannot occur in situ in argument position in single wh-in-situ questions. In addition, the wh-expression šu cannot also occur in the base position in multiple wh-questions. This can be illustrated in the following examples.

(10) a. miin zaar miin? who visited.3ms who ‘Who visited whom?’

b. *miin šeef šu who saw.3ms what ‘Who saw what?’

c. šu daaya? miin? what upset.3ms who ‘What upset who?’

As can be seen from (10), unlike the wh-argument miin ‘who’ in (10a), the wh-expression šu ‘what’ cannot appear in situ in Lebanese multiple wh-questions (10b). Meanwhile, (10c) shows that it is possible to have šu in a multiple wh-question in general. As such, (10b) is not ruled out because šu cannot occur in multiple wh-questions; rather, it is ruled out because šu occurs in the base argument position.

On the other hand, the wh-expression šu ‘what’ may appear in positions

---

3This is assuming that we have pair-list interpretation where the šu is taking wh-scope. It will be shown later on that the structure is acceptable on the interpretation where šu is a simple echo.
where a wh-element would appear, i.e., in a clause initial position. We have already seen this in multiple wh-questions in (10c) above. As for single wh-questions, consider the sentences (11) and (12) below.

(11) a. šū štarayte mbeeriḥ  (Lebanese)
    what bought.3sf yesterday
    ‘What did you buy yesterday?’

    b. šū fakkarto ?ānno zeena štarit mn l-maktabe
    what thought.2p that Zeina bought.3sf from the-bookstore
    ‘What did you think that Zeina bought from the bookstore?’

(12) a. saʔalto šū zeena štarit mn l-maktabe?
    ask.2p what Zeina bought.3fs from the-bookstore
    ‘Did you ask what Zeina bought from the bookstore?’

    b. badkun taʔrfo šū zeena štarit mn l-maktabe?
    want-2p know.2p what Zeina bought.3fs from the-bookstore
    ‘Do you want to know what Zeina bought from the bookstore?’

The sentences in (11) show that the wh-expression šū ‘what’ occurs in the initial position of matrix clauses and have wide scope over the entire sentence⁴. The distribution of the wh-expression šū ‘what’ is not restricted to the initial position of matrix clauses, the sentences in (12) show that it may also occur at the front of an embedded clause and have scope over that clause. Embedded šū cannot, however, co-occur with complementisers such as ʔeza ‘whether’, much in the same way as its English counterpart. To illustrate, consider the following cases.

(13) a. *John asked what whether Mary bought.

    b. *They don’t know who(m) whether she saw.

⁴According to Aoun and Choueiri 1999, these questions are formed by wh-movement. However, I will argue later on that wh-questions involving the wh-expression šū are actually base-generated questions.
The examples in (13) are classic cases of the Doubly-Filled Comp Filter (Chomsky and Lasnik 1977) that is operative in English. Likewise, the Doubly-Filled Comp Filter seems to be operative in Arabic (14a-b) on the basis that the wh-expression ˇs u is in Spec,CP that therefore cannot co-occur with elements occupying C.

Other contexts where a distributional contrast is found between Lebanese wh-phrases is resumptive wh-questions and d-linked wh-questions. Like many spoken Arabic dialects, Lebanese wh-questions can also be formulated using a resumptive strategy whereby the initial wh-phrase is connected to a resumptive pronoun inside the clause. However, not all Lebanese argument wh-phrases can be related to a resumptive element in this type of wh-questions. That is, only the wh-phrase miin ‘who’ and ˇaya-NP ‘which’-NP phrases may occur in this type of question, but not the wh-expression ˇs u ‘what’. Examine the following sentence pairs.

(15) a. ˇaya mmasil šoxt-o b-l-maT Yam (Lebanese) which actor saw.2sm-him in-the-restaurant ‘Which actor did you see in the restaurant?’

b. miin šoxt-o b-l-maTYam who saw.2sm-him in-the-restaurant ‘Who did you see in the restaurant?’

(16) a. ˇaya ktaab štarayt-i which book bought.2sf-it ‘Which book did you buy?’
b. *šu štarayt-i
   what bought.2sf-it
   ‘What did you buy?’

Since resumptive clitics correspond to definite DPs in Arabic (Ouhalla 2001, Aoun, Choueiri and Hornstein 2001 and Ouhalla and Shlonsky 2002) the wh-expression ša ‘what’ is excluded from resumptive wh-questions (16b) on the grounds that it is unable to provide a reference for the resumed element, hence the ungrammaticality. By contrast, the argument Wh-DPs miin ‘who’ and ?ayya-NP ‘which’-NP, shown in (15a-b) and (16a), are able to relate to a resumptive clitic inside the clause. Similarly, both argument Wh-DPs miin ‘who’ and ?ayya-NP ‘which’-NP may be linked to previous discourse, but not the wh-expression ša ‘what’. The following cases illustrate this point.

(17) l-mudiir w l-ʔasteez mawʔudiin b-l-maktab
      the-principal and the-teacher present.p in-the-office
   ‘The principal and the teacher are in the office’

   a. ʔayya waahad baddak tšuuf
        which one.sm want.2sm see.2sm
        ‘Which one do you want to see?’

   b. miin baddak tšuuf?
        who want.2sm see.2sm
        ‘Who do you want to see?’

(18) Fii ʔande bluuze hamra w bluuze sawda
     in-it at-me shirt.fs red.fs and shirt.fs black.fs
     ‘I have a red shirt and a black shirt’

   a. ʔayya wahdə baddak tžarrib
        which one.fs want.2sm try.2sm
        ‘Which one do you want to try?’

\[5\text{Notice that these are wh-movement questions although the wh-phrases are d-linked and expected to be in-situ. This strongly suggests that Lebanese is essentially a wh-movement language and that wh-in-situ is very restricted.}\]
b. *šu baddak tərrib
   what want.2sm try.2sm
   ‘What do you want to try?’

So far, the properties that the wh-expression šu ‘what’ displays, especially its inability to be d-linked or resumed, strongly indicate that it is not a DP, unlike, for instance, the bare argument wh-phrase miin. Interestingly, however, the wh-expression šu ‘what’ frequently occurs in positions that are restricted to definite DPs only, i.e., the initial DP position in copular wh-constructions and wh-cleft questions (Cheng 1991, Ouhalla 1999 and Shlonsky 2002). Consider the following examples from Aoun and Benmamoun and Choueiri 2010 (p.151).

(19) a. šu l-ʔakl l-yom?
   what the-food today
   ‘What is on the menu today?’

  b. šu huwwa l-ʔakl l-yom?
   what it the-food today
   ‘What is on the menu today?’

(20) a. šu illi rah teeklu-u l-yom?
   what that fut. eat.2p-it today
   ‘What is that you are going to eat today?’

  b. šu huwwa illi rah teeklu-u l-yom?
   what it that fut. eat.2p-it today
   ‘What is that you are going to eat today?’

(21) a. *šu ʔakl l-yom?
   what food today
   ‘What is on the menu today?’

  b. *šu huwwa ʔakl l-yom?
   what it food today
   ‘What is on the menu today?’

In such constructions, the wh-expression šu ‘what’ appears in a clause-initial

6 Wh-clefts will be discussed in detail in the following chapter
position where only definite DPs can occur\(^7\). This property of the wh-expression ˇs u ‘what’, which seems to suggest that it is a DP, is in sharp contrast with its d-linking and resumptive properties.

To recap, the last two sections have been concerned with the behaviour of argument wh-phrases in Lebanese and have identified as the source of the problem the peculiar behaviour of the wh-expression ˇs u ‘what’. I have shown that the wh-expression ˇs u ‘what’ displays an irregular distributional pattern: it cannot occur in situ in argument position in single wh-questions, nor can it occur in situ in multiple wh-questions; it can only occur clause initially; it cannot be related to a resumptive pronoun, nor can it be d-linked; meanwhile, it can occur in copular constructions and wh-cleft questions in positions restricted to definite DPs. The data thus far present us with two problems: one problem lies in the asymmetry between the argument wh-phrases miin ‘who’ and ˇs u ‘what’, while the other problem lies in the conflict between argument vs. non-argument properties that the wh-expression ˇs u ‘what’ exhibits. In the section that follows, I look into an account for these problems which was originally put forward in Aoun and Choueiri 1999 and developed in the subsequent work of Aoun, Benmamoun and Choueiri 2010.

### 3.3.2 Aoun, Benmamoun and Choueiri’s 2010 analysis

To solve the problem posed by the asymmetries between Lebanese argument wh-phrases, Aoun and Choueiri 1999 and Aoun, Benmamoun and Choueiri 2010 propose an account that is essentially based on Pesetsky’s 1987 notion of d-linking. They argue that whereas the argument wh-phrase miin ‘who’ can be d-linked and may, therefore, occur in situ, the argument wh-phrase ˇs u ‘what’ cannot. Consequently, the wh-expression ˇs u ‘what’ is said to always undergo obligatory movement to a clause initial position. This section shows that the analysis proposed in Aoun and Choueiri 1999 and Aoun, Benmamoun and Choueiri 2010 provides a satisfactory account of the distributional properties of the wh-expression ˇs u ‘what’ in Lebanese.

\(^7\)See Chapter 1 for a discussion on copular constructions.
Choueiri 2010 cannot be maintained and argues that a d-linking approach for the asymmetries observed earlier is irrelevant.

The fact that the wh-expression ʂu ‘what’ cannot be d-linked, unlike miin and ṭayya-NP, has led Aoun, Benmamoun and Choueiri 2010 to posit that ʂu ‘what’ is a non-referential Wh-NP. Based on this fact, Aoun, Benmamoun and Choueiri suggest that wh-phrases are decomposed into two parts: a wh-element which bears the question feature, and a noun phrase, which can be either a full DP as is the case for the wh-words miin ‘who’ and ṭayya(a) ‘which’, or an NP as is the case for the wh-expression ʂu ‘what’. Under this proposal, the wh-words miin/man ‘who’ and ṭayya(a) ‘which’ have the representation shown in (22) whereas the wh-expression ʂu ‘what’ has the representation shown in (23).

(22) miin/man/ṭayya(a)

\[
\begin{array}{c}
\text{wh} \\
\text{DP}
\end{array}
\]

(23) ʂu

\[
\begin{array}{c}
\text{wh} \\
\text{NP}
\end{array}
\]

As a non-referential Wh-NP, the wh-expression ʂu ‘what’ is said to always undergo overt movement to Spec,CP of the wh-interrogative. For Aoun, Benmamoun and Choueiri, then, ʂu-questions are wh-movement questions and have the same derivation as miin-questions. Consider, for instance, the derivation in (24) below. Following Chomsky (1995, 2000, 2001a and 2001b), C would have, in addition to an EPP feature, an interpretable [WH] feature and an un-

---

8Aoun, Benmamoun and Choueiri 2010 adopt Cinque’s 1990 notion of referentiality which, in turn, is based on Chomsky’s 1986 and Rizzi’s 1990 approach to the argument-adjunct asymmetry. Under the current analysis, the referentiality account is deemed irrelevant since it deals with long argument vs. adjunct wh-extraction. However, the reader is referred to those works for a full discussion on the referential status of wh-phrases.
interpretable [Q] feature. It acts as a probe searching for a goal with matching features. The goal is the wh-phrase with an uninterpretable [WH] feature and an interpretable [Q] feature. Upon matching in their feature set, Agree between C and the wh-phrase takes place. Movement of the wh-phrase would be seen as a consequence of the [EPP] property on C which allows the projection of a specifier and triggers movement of the wh-phrase into the position of that specifier, i.e., Spec,C. The structure for (24a) may roughly be represented in (24b).

(24)  a. miin šefte mbeerih
   who saw.3sf yesterday
   ‘Who did you see yesterday?’

b.  

\[
\begin{array}{c}
\text{CP} \\
| \\
\text{DP} \\
| \\
\text{miin} \\
\text{Who} \\
\text{[Q, uWH]} \\
\text{[sQ, WH, EPP]} \\
\text{C} \\
| \\
\text{TP} \\
| \\
\text{C'} \\
| \\
\text{T'} \\
| \\
\text{T} \\
| \\
\text{pro} \\
| \\
\text{šefte} \\
\text{saw.3fs} \\
| \\
\text{v*p} \\
| \\
\text{*v+V} \\
| \\
\text{DP} \\
| \\
\text{*v} \\
| \\
\text{VP} \\
| \\
\text{*v} \\
| \\
\text{V} \\
| \\
\text{DP} \\
| \\
\text{miin} \\
\end{array}
\]
(24b) is a classic derivation for wh-movement questions that applies to *miin*-questions only. The assumption that *šu* must always undergo movement and that *šu*-questions are derived in the same manner as *miin*-questions cannot be maintained, however. Below, I highlight the problems that encounter Aoun, Benmamoun and Choueiri’s analysis and show why *šu*-questions cannot be treated in the same way as *miin*-questions.

**Problems with Aoun, Benmamoun and Choueiri’s analysis**

The analysis outlined in Aoun and Choueiri 1999 Aoun and Benmamoun and Choueiri 2010 is inadequate. Although the general notion of referentiality, which subsumes d-linking, characterises the difference in the interpretive properties between Lebanese wh-phrases, it still faces the challenge of explaining the overall asymmetry in the various wh-constructions, including single wh-in-situ questions. This section presents a number of arguments which strongly indicate that d-linking is not at play in Lebanese wh-asymmetries.

The first argument against the claim that *šu* ‘what’ cannot occur in-situ because it cannot be d-linked comes from wh-in-situ contexts other than single wh-questions, such as multiple wh-questions. As is well known, in this type of question, the wh-phrase that remains in situ need not be d-linked for it to be licensed in that position. Consider, for instance, the following English sentence with its representation.

(25)  
\[
\begin{align*}
\text{(a) Who bought what?} \\
\text{(b) } [\text{CP Who C } [TP (Who) bought what ]] \\
\end{align*}
\]

In questions like (25), the higher (subject) wh-word, i.e., the wh-phrase *who*, undergoes movement to Spec,CP, while the lower (object) wh-word, namely, the wh-phrase *what*, stays in-situ, and the sentence is well-formed. As for the licensing of the in-situ wh-phrase in cases like (25), it can be said that the wh-phrase *what* is allowed to stay in-situ because of the moved wh-phrase. Within the
Minimalist framework (Chomsky 1995, 2000), in-situ *what* in (25) need not undergo wh-movement since the strong Q feature in the matrix C° is already checked by the wh-phrase *who*. The wh-feature of *what* will not be attracted to C° as a consequence. Instead, it may be assigned an interpretation via unselective binding (Baker 1970, Pesetsky 1987 and Chomsky 1995) or choice function (Reinhart 1998). The point, then, is that in multiple wh-questions, a wh-phrase may stay and get bound in-situ regardless of whether it is d-linked or not.

Turning to Lebanese multiple wh-questions, one predicts such questions are well-formed in Lebanese since the in-situ wh-phrase need not be d-linked for it to be licensed in that position, as we have just seen in English. This prediction is not borne out, however. The asymmetry between the wh-phrases *miin* ‘who’ and *šu* ‘what’ still exists. This can be illustrated in the following examples.

\[(26)\]  
\[\begin{align*}
\text{a. } & \text{miin zaar } \text{miin?} \\
& \text{who visited.3ms who} \\
& \text{‘Who visited whom?’} \\
\text{(Lebanese)} \\
\text{b. } & \text{*miin šeef } \text{šu} \\
& \text{who saw.3ms what} \\
& \text{‘Who saw what?’} \\
\text{c. } & \text{šu daaya’ miin?} \\
& \text{what upset.3ms who} \\
& \text{‘What upset who?’}
\end{align*}\]

Unlike the wh-argument *miin* ‘who’ in (26a), the wh-expression *šu* ‘what’ cannot occur in situ in Lebanese multiple wh-questions as in (26b)\(^9\). Meanwhile, (26c) shows that it is possible to have *šu* in a multiple wh-question in general. Thus, (26b) is ruled out because *šu* occurs in the base argument position, not because it cannot occur in multiple wh-questions.

\(^9\)This is assuming that we have pair-list interpretation where the *šu* is taking wh-scope. It will be shown later on that the structure is acceptable on the interpretation where *šu* occurs in simple echo questions.
Under the analysis proposed in Aoun and Choueiri 1999 and Aoun, Bennamoun and Choueiri 2010, the fact that the wh-expression .unbind ‘what’ cannot occur in situ in multiple wh-questions cannot be explained. That is, there is no reason why (26b) should not be well-formed since, generally, there is no d-linking requirement on the in-situ wh-phrase for it to be licensed in such questions. As a matter of fact, d-linking is not relevant to the licensing of in-situ wh-phrases in multiple wh-questions and the issue of whether a wh-phrase can or cannot be d-linked should not affect its ability to occur in situ. Clearly, the examples above display the same type of contrast observed earlier in single wh-in-situ questions. By and large, the cases examined so far strongly indicate that it is a distinct property of unbind ‘what’ that prevents it from occurring in situ, be it in single wh-questions or in multiple wh-questions, and that this property has nothing to do with d-linking.\(^{10}\)

The peculiarity of the wh-expression unbind ‘what’ can be seen more clearly by comparison with the ‘what’-forms in other dialects of Arabic. For instance, there is no such contrast between argument wh-phrases in Standard Arabic (SA) or in Iraqi Arabic where the ‘what’-forms do appear in situ in multiple wh-questions. Consider the following examples.

(27) a. mann raʔat man? who saw.3ms who ‘Who saw whom?’

b. mann raʔat maadha? who saw.3ms what ‘Who saw what?’

(28) a. ʃeno natat Mona li-meno? what gave Mona to-whom ‘What did Mona give to whom?’

\(^{10}\)A similar analysis is advanced in Simpson (2000) whereby d-linking is rejected as an approach for the licensing of wh-in-situ.
Unlike the wh-expression šu ‘what’, the SA form maadba ‘what’ as well as the Iraqi form šeno ‘what’ pattern with the argument wh-phrase ‘who’ in that both occur in situ in multiple wh-questions. These cases confirm the view that the wh-expression šu ‘what’ is distinct since it differs from its counterparts in other Arabic dialects as well as from other argument wh-phrases in general. This might be taken to suggest that d-linking must be abandoned and that we should look for another explanation for the peculiar behaviour of the wh-expression šu ‘what’ in the syntax.

The second related argument in favour of abandoning the d-linking approach to Lebanese wh-argument asymmetries comes from Superiority contexts. In English multiple wh-questions, for instance, superiority effects arise when a lower wh-phrase moves across a higher wh-phrase. The contrast between (29a) and (29b) illustrates this point.

(29)  
   a. Who saw what?  
   b. *What did who see

The sentence in (29a) does not involve movement of the lower wh-phrase what across the higher wh-phrase who, and is, therefore, well-formed. Such movement induces a superiority violation, hence, the ungrammaticality in (29b). However, according to Pesetsky 1987, violations imposed by the Superiority Condition can be avoided when d-linking is at play. In other words, d-linked wh-phrases may stay and be bound in situ, without the need to undergo movement. It

---

11The Superiority Condition is best understood in terms of Shortest Move or Attract Closest in Chomsky 1995 (chapter 4:296) as stated in (i):

(i) a can raise to target K only if there is no legitimate operation Move β targeting K, where β is closer to K.

152
follows, no superiority effects arise with d-linked wh-phrases. Consider the following examples from Pesetsky (1987: 309).

(30)  
\begin{enumerate}
  \item a. I know what just about everybody was asked to do, but what did who (actually) do?
  \item b. I know that we need to install transistor A, transistor B, and transistor C, and I know that these three holes are for transistors, but I'll be damned if I can figure out from the instructions where what goes!
\end{enumerate}

In the above examples, though there is no overt mark of d-linking, the context limits reference of wh-phrases by what is already known by the speaker (and hearer). Since the wh-phrases are d-linked, superiority effects do not arise. If, on the other hand, the wh-phrases were non-d-linked, such questions would constitute a superiority violation induced by movement of the lower wh-phrase across the higher wh-phrase, as can be illustrated below.

(31)  
\begin{enumerate}
  \item a. *Guess what did who do \textit{(what)}
  \item b. *Guess where what goes \textit{(where)}
\end{enumerate}

The difference between the wh-phrases in (30a-b) and (31a-b) resides in their discourse properties, i.e., whereas in (30) the wh-phrases are d-linked, they are not so in (31). Hence, (31a-b) are ruled out on the grounds that non-d-linked wh-phrases are operators that therefore must move. Consequently, movement of \textit{what} across \textit{who} in (31a) and movement of \textit{where} across \textit{what} in (31b) yield Superiority violations.

As for Lebanese, the Superiority scenario is rather different from the one outlined above. Interestingly, d-linking plays no role in evading superiority
effects in Lebanese. In other words, superiority effects still arise even when the wh-phrases are d-linked. For instance, in the context where there is a group of boys, each invited by someone, only ‘which’-NP phrases may be used, but not the bare wh-phrase *miin ‘who’. Consider the contrast in the following examples given by Aoun and Li 2003, who situate these examples in a context so that the wh-phrases have a d-linked reading (Aoun and Li 2003: 46).

(32)  
   a. *miin ̣talabto ̣min ̣miin-un ̣yiṣem?
       who asked.2p from who of-them 3ms.invite
       ‘Who did you ask whom to invite?’

   b. ̣ayya ̣walad ̣talabto ̣min ̣ayya ̣bun ̣yiṣem?
       which boy ̣asked.2p from which girl ̣3fs.invite
       ‘Which boy did you ask which girl to invite?’

Although the in-situ wh-elements in (32) are d-linked, there is a sharp contrast between (32a) and (32b). Whereas (32b), which involves the compound wh-phrase *ayya NP ‘which’-NP, is well-formed, (32a), which involves the bare wh-phrase miin ‘who’, is not. These examples reveal two crucial facts about Lebanese wh-asymmetries. One fact is that the syntax of wh-questions must be closely tied to micro-parametric differences operating at a word-level such that the difference between Lebanese cases in (32) above is accounted for. More specifically, the fact that morphologically d-linked wh-phrases can evade superiority effects while contextually d-linked wh-phrases cannot must be explained in terms related to the internal structure of wh-words. The second fact is that, at least in Lebanese, d-linking does not have the same effects on the licensing of wh-in-situ as it does in other languages like English.12 Overall, the point to be made here is that the contrast between Lebanese argument wh-phrases, whether it is between miin ‘who’ and *ayya-NP ‘which’-NP phrases, as is the case at hand, or whether the contrast is between miin ‘who’ and *su ‘what’, as we have seen earlier, is beyond d-linking.

---

12See Pesetsky 1987 for a discussion on how d-linking licenses wh-in-situ in other languages such as Japanese and Polish.
From a theoretical perspective, notice that the fact that an argument wh-
phrase like Lebanese śu ‘what’ cannot be d-linked is surprising. According to
Pesetsky (1987), all wh-phrases can be d-linked. As we have seen earlier, bare
wh-phrases like who and what can be linked to discourse once reference is es-
tablished (see (30a-b)).

D-linking aside, one other argument against the analysis of Aoun, Ben-
mamoun and Choueiri concerns the structures they propose for wh-words. In
particular, the assumption that śu ‘what’ has the (bare) NP-structure in (23),
which, in turn, is based on the fact that śu ‘what’ cannot be d-linked, is implau-
sible. Evidence against treating the wh-expression śu ‘what’ as an NP comes
from copular constructions including LD’ed equatives in which the element that
occurs clause-initially must be a definite DP. Consider the following examples.

(33) a. l-mīlālem huwwa l-maš?ol
    the-teacher he the-responsible
    ‘The teacher is the one responsible’

b. l-mīlālem huwwa maš?ol
    the-teacher he responsible
    ‘The teacher is responsible’

c. *mīlālem huwwa l-maš?ol
    teacher he the-responsible

If the wh-expression śu ‘what’ were indeed an NP, it is predicted not to be able
to occur in such constructions. However, as the following examples show, this
prediction is not borne out.

(34) a. śu huwwa l-ʔakl l-yom?
    what it the-food today
    ‘What is on the menu today?’
3.4 The syntax of ṣu

This section develops a syntactic approach that successfully accounts for the asymmetries between Lebanese argument wh-phrases as well as the conflict observed in the behaviour of the wh-expression ṣu ‘what’. The complexity in the behaviour and distribution of this wh-expression leads the current discussion to the hypothesis that this wh-expression is not what it looks to be, i.e., a bare wh-word, but rather a complex wh-expression. Pursuing this hypothesis, I examine, first, the various ‘what’ forms in other Arabic dialects given the similarities in...
their grammar and lexicon. Specifically, I will be mainly concerned with the morphological composition of these forms and see whether a connection can be established with the wh-expression šu ‘what’. This will be followed by a study of the ‘what’-forms in Jordanian Arabic which is similar to Lebanese but differs in that it has an extra ‘what’ form, i.e. ġeyš, the analysis of which will be crucial in unfolding the internal structure of the wh-expression šu ‘what’. Finally, I put forward the proposal that the wh-expression šu ‘what’ is an overt operator with an internally complex CP structure. It will be argued that the reason why šu ‘what’ is always clause initial follows straightforwardly if we assume that it is only generable in the specifier position of the interrogative clause, in a way similar to clausal subjects discussed in Koster 1978 and Alrenga 2005.

3.4.1 Preliminaries

The mystery surrounding the wh-expression šu ‘what’ poses the crucial question about its argumental status and whether it is indeed what it is assumed to be, i.e., the equivalent of English what. As we have seen in the previous sections, the surface distribution of the wh-expression šu ‘what’ coupled with the fact that it cannot bind resumptive pronouns, unlike the wh-phrase miin, suggest that šu ‘what’ cannot be treated as a DP. Given that d-linking is irrelevant to the asymmetries observed in this study, a more plausible approach to such asymmetries might be based on morpho-syntactic properties of argument wh-phrases. Before spelling out the proposal for the syntax of the wh-expression šu ‘what’ and the questions in which it occurs, a few observations about the wh-forms in Arabic dialects must be taken into account.

The first observation concerns the nature of the wh-expression šu ‘what’ and whether it has meaning as a lexical item. Upon comparison with other wh-forms, be it in Lebanese or in other Arabic dialects, the wh-word šu does not appear to have meaning, i.e., it is not a lexical item, unlike other wh-words such as ġayy(a) ‘which’, miin ‘who’ and ween ‘where’ or their counterparts in other
dialects. This observation is based on the fact that apart from the ‘what’-forms, all other dialectal forms have identical counterparts in the standard variety of the language. Compare, for instance, the ‘who’ forms with those of ‘what’ in the following table.

<table>
<thead>
<tr>
<th></th>
<th>SA</th>
<th>Iraqi</th>
<th>Lebanese</th>
<th>Jordanian</th>
<th>Moroccan</th>
</tr>
</thead>
<tbody>
<tr>
<td>What</td>
<td>maa(dhaa)</td>
<td>šeno</td>
<td>ša</td>
<td>ša/ʔeyš</td>
<td>šeno/škun</td>
</tr>
<tr>
<td>Who</td>
<td>mann</td>
<td>meno</td>
<td>miin</td>
<td>miin</td>
<td>meno/škun</td>
</tr>
</tbody>
</table>

An obvious, and crucial, fact from the data in (35) is that SA is the only language that has the interrogative maa(dhaa) ‘what’ form which does not involve the prefix š-. This wh-form, i.e., maa(dhaa) ‘what’, is compound and consists of two parts: the interrogative particle maa ‘what’, which can be used on its own in SA as in (36a), and the proximate marker dhaa ‘this’ (Farghal 1986, Fassi-Fehri 1993 and Aoun and Li 2003). The following examples illustrate the use of this form in SA.

(36)  
a. maa huwa l-soʔaal?
    what it the-question
    ‘What is the question?’

b. maadhaaʔakalta?
    what 2sm.ate
    ‘What did you eat?’

Concerning the dialectal ‘what’ forms, shown in (35), we can see that the SA wh-form maadhaa ‘what’ is absent in those dialects and an š-form is used instead. Nevertheless, one can still find traces of interrogative maa ‘what’ in questions like (37), which are common across the dialects.

13Note that this might be also the case for English what which can be said to constitute the wh-element wh- and the demonstrative -at
Examples like (37) are, thus, taken as evidence that interrogative maa ‘what’ still exists in the dialects, but, for some reason, it has almost disappeared from the varieties. A similar observation is made in Holes 2004 who points out that the generalisation of the homophonous negative particle maa ‘not’, following the simplification of the classical and standard system of negation, led to the loss of interrogative maa ‘what’ forms almost everywhere. This is a reasonable explanation since both negative and interrogative maa are placed immediately before the verb which would routinely yield ambiguity in interrogative and negative sentence structures\textsuperscript{14}.

The second observation, then, is the common use of a š-form in the dialects instead of interrogative maa ‘what’. Pursuing this line of reasoning, a variety of forms, which minimally differ from each other (see table (35)), seem to have been developed in the dialects in order to replace the interrogative maa ‘what’. As such, the question this reasoning gives rise to is how or what mechanism is used to make up the ‘what’ forms in these dialects. An answer to this question can be found in Chomsky 1995 who posits that wh-phrases might be composed of an indefinite quantifier, a wh-feature and the restriction on the quantifier. A wh-phrase like who would be composed of [some x, wh-, x a person]. A similar view is also held in Aoun and Li 2003 for Arabic wh-phrases who suggest that a question word is a composition of three parts: Question, Quantification and Restriction. While this mechanism can be said to apply straightforwardly to wh-phrase like English who and what and the Arabic forms miin and maa(dhaa) ‘what’, the case for šu ‘what’ remains unclear, although one would expect it to

\textsuperscript{14}Ambiguity would become even stronger if we also consider the frequent use of maa-free relatives in the dialects where the relativiser maa is placed immediately after the wh-phrase. Intuitively, then, the loss of interrogative maa from the dialects became necessary.

159
contain (some of) the parts of a question word like *miin*. To give substance to this discussion, the structural composition of *šu* ‘what’ can be concretely derived on the basis of new data from a similar dialect to Lebanese, i.e., Jordanian, which uses two forms for the wh-phrase ‘what’. The properties and distribution of these forms are the subject of the following section.

### 3.4.2 Jordanian *ʔeyš* & *šu*

The purpose of this section is to uncover the morphological composition of the wh-expression *šu* ‘what’, which will be incorporated in the syntactic analysis of the questions in which it occurs. I start by examining ‘what’-questions in Jordanian Arabic and establish the morphological connection between the two forms, i.e., *ʔeyš* & *šu*, used in this dialect.

Jordanian is closely related to Lebanese except that Jordanian is a wh-movement language in which wh-phrases move in overt syntax to the front of the wh-clause\(^{15}\). This can be illustrated in the following examples.

\(\begin{align*}
\text{(38)} \quad & \text{a. } \textit{miin šoft}\ b-l-ḥafleh? \\
& \text{who saw.2s in-the-party} \\
& \text{‘Who did you see at the party?’}
\end{align*}\)

\(\begin{align*}
\text{b. } \textit{miin fakkatro } \textit{ʔenno ſomar zaar}? \\
& \text{who thought.2pl that Omar visited.3ms} \\
& \text{‘Who did you think that Omar visited?’}
\end{align*}\)

\(\begin{align*}
\text{(39)} \quad & \text{a. } \textit{ʔeyš } \textit{ʔakalti l-vous}\text{?} \\
& \text{what eat.2f the-day} \\
& \text{‘What did you eat today?’}
\end{align*}\)

\(\begin{align*}
\text{b. } \textit{ʔeyš naSahto } \textit{Sanyta todrus}? \\
& \text{wha advised.2pl Sanyta study.3fs} \\
& \text{‘What did you advise Sanyta to study?’}
\end{align*}\)

\(^{15}\)Apart from multiple wh-questions, wh-in-situ is marginally available in this languages in contexts where some background information is given.
Jordanian also differs from Lebanese in that it uses two forms in ‘what’-questions: one form is identical to the one in Lebanese, i.e., the wh-expression ūs ‘what’, and the other one is the wh-expression peyš ‘what’ exemplified in (39) above. Although both forms are used alternatively in a wide range of wh-constructions by speakers of this dialect, closer inspection reveals crucial differences between the two. Putting aside the wh-expression ūs for the moment, the wh-form peyš is bi-morphemic and derives from the two-word phrase pay-ši, namely, the wh-element pay(a) ‘which’ and the indefinite (NP) š(i) ‘thing’\textsuperscript{16}. Both of these elements may be used either separately or in conjunction with each other commonly found in contexts of quantification (40a), negation (40b) and interrogative contexts (40c).

\begin{equation}
\begin{align*}
\text{(40)} & \quad \text{a.} \quad \text{biddak pay š(i) mn l-sooq?} \quad \text{(Jordanian)} \\
& \quad \text{want.2s which thing from the-market} \\
& \quad \text{‘Do you want anything from the market?’} \\
\text{b.} \quad \text{wala š(i) helo hoon} \\
& \quad \text{Not thing beautiful here} \\
& \quad \text{‘Nothing is beautiful here’} \\
\text{c.} \quad \text{pay š(i) ajabak?} \\
& \quad \text{which thing like.2s} \\
& \quad \text{‘Which thing did you like?’}
\end{align*}
\end{equation}

Structurally, the two-word phrase pay(a) š(i) ‘which thing’ is a DP much in the same way as its English counterpart which thing (see Chomsky 1995 and Aoun and Li 2003)\textsuperscript{17}. Inside the DP structure, the wh-element pay ‘which’ occupies the D position and takes as complement the NP š(i) ‘thing’\textsuperscript{18}. This structure is shown in (41) below.

\textsuperscript{16}These elements are also used in Lebanese as shown in various parts of this study.

\textsuperscript{17}It is noteworthy here that the two-word phrase pay(a) š(i) undergoes phrasal spellout as one chunk which yields the form peyš, this will be discussed in detail in the sections ahead.

\textsuperscript{18}A similar analysis is suggested for Lebanese pay(a) ‘which’, i.e., as the D head in a DP structure, in Aoun and Li 2003.
As a DP, the wh-form "which thing" displays properties that are characteristic of argument wh-phrases. For instance, unlike the wh-expression "šu", the wh-form "eyš ‘which thing’ may occur in situ in canonical argument position in questions that follow some background information such as those in (42).

(42) a. Talabto eyš ŋa-ήrada? (Jordanian)
  ordered.2p what for-the-lunch
  ‘What did you order for lunch?’

b. l-ήlæaj eyš?
  the-treatment what
  ‘What is the treatment?’

As far as d-linked wh-questions and resumptive wh-questions are concerned, the wh-form "eyš ‘which thing’ may also occur in such contexts. It can also be noted here that Jordanian is like Lebanese in that the wh-expression "šu cannot occur in these types of questions. First, consider d-linked wh-questions illustrated below.

(43) ŋaunde kull l-ʔawaafi l-jadeede (Jordanian)
  have.1s all the-clothes the-new
  ‘I’ve got all the new clothes’

a. ʔeyš beddik tʒarribi b-l-ʔawwal
  what want.2fs try.2fs in-the-first
  ‘What do you want to try first?’
Unlike the wh-expression *šu in (43b), the wh-phrase ?eys in (43a) is fully well-formed where it used to question members of a set given in previous discourse. Similarly, in wh-questions that involve resumption, the wh-form ?eys may also occur clause initially and be related to a resumptive element inside the wh-clause. Consider the following examples.

(44) a. ?eys ṭaarrībt-ī b-l-ʔawwal?
what try.2fs-it in-the-first
‘What did you try first?’

b.  *šu ṭaarrībt-ī b-l-ʔawwal
what try.2fs-it in-the-first
‘What did you try first?’

Finally, the wh-form ?eys ‘which thing’ appears in copular wh-questions. Recall that such constructions allow only definite DPs in the initial position. Interestingly, the wh-expression *šu ‘what’ may also occur in this type of construction, as we have seen in the Lebanese cases earlier19. This can be illustrated in (45) and (46) below.

(45) a. ?eys ḥu l-soʔaal?
what it the-question
‘What is the question?’

b. *šu ḥu l-soʔaal?
what it the-question
‘What is the question?’

19 A more detailed account for these constructions is presented in the following chapter. See also Shlonsky 2002 for a discussion of such questions in Palestinian.
(46) a. ṭeys hu illi baddak tehkek-h?
what it that want.2ms say-it
‘What is it that you want to say?’

b. šu hu illi baddak tehkek-h?
what it that want.2ms say-it
‘What is it that you want to say?’

Sentences like (45) were analysed in Chapter one as left-dislocated equatives which are restricted to definite DPs only\(^{20}\). The latter is said to be in a coreference relationship with the subject pronoun inside the clause. Likewise, the sentences in (46) admit both the wh-form ṭeys (46a) and the wh-form šu (46b). As observed in Shlonsky 2002 and Aoun, Benmamoun and Choueiri 2010, only definite arguments can occur in this type of wh-questions.

To recap, this section has been mainly concerned with the morpho-syntactic properties of the Jordanian wh-form ṭeys. Unlike the wh-form šu, the wh-form ṭeys displays properties characteristic of argument wh-phrases. It has also been shown that this form, i.e., ṭeys, is a compound wh-phrase that has the structure of a DP. Apart from these differences, Jordanian uses the two wh-forms interchangeably in the rest of wh-constructions, including canonical wh-questions as well as copular wh-constructions. The facts indicate that there is a resemblance between these two forms, yet there is a crucial difference between the two.

\(\text{šu is ṭeys hu}\)

Given this background, my conjecture is that the wh-form šu ‘what’ is made up of the wh-form ṭeys and some extra material. Given the DP analysis of ṭeys, what remains to be found is the nature of the extra material, how it affects the distribution of this compound and whether this distribution can be reconciled with the distribution of šu observed here.

\(^{20}\text{See Chapter one for a thorough discussion of copular constructions.}\)
The assumption that the wh-form șu is a compound wh-expression is not implausible and appears to be the only promising solution for all the problems associated with the behaviour of this wh-expression. To substantiate this view, I would like to touch on two studies on dialectal Arabic that make similar observations. The first study is Obler’s 1976 who suggests that both the forms ʔeyš and șu are, in fact, developments from the classical phrase ʔayy-u șay-in ‘which thing’ bolstered by the pronominal huwa ‘he/it’ at the end (see 47a below). The second and more recent study is Holes’s 2004 which investigates how the dialects developed away from the classical and standard varieties. According to Holes, the wh-expression șu is derived from the post-classical phrase ʔayy šay hu ‘which thing it’ (see 47b below).

(47) a. ʔayy-u șay-in huwa (SA/CA)
    which-NOM thing-GEN it
    ‘What is it?’

b. ʔayy šay hu (Dialects)
    which thing it
    ‘What is it?’

The cases in (47) invite two further observations. First, there are clear differences in the morphology and, hence, the phonology of the standard variety and the dialectal varieties. Morphological case is overt in the former, but not in the latter. This is also accompanied by changes in the phonological component which can be seen clearly by comparing the two sentences. On the other, the syntax of (47a) and (47b) is the same. The second observation is that such composition satisfies the requirement for wh-word formation proposed in Chomsky 1995 and Aoun and Li 2003 (cf. section 3.4.1). Based on these observations made above, my contention is that the wh-phrase șu consists of the Wh-DP ʔeyš ‘which thing’ and the pronominal hu ‘he/it’.
Proceeding with the assumption that šu ‘what’ is morphologically composed of the three-word string ḫayy šay hu ‘which thing it’, two issues arise: the first issue concerns the syntax of this string while the second issue concerns the syntax of wh-questions in which it occurs. These issues are discussed in turn in the following two subsections.

3.4.3 The Syntax of ḫayy šay hu

As far as the syntax of the string ḫayy šay hu ‘which thing it’ is concerned, I propose that it is an equative construction. More precisely, the string ḫayy šay hu is a present-tense equative copular wh-construction. Recall that equatives in Arabic are verbless constructions and may involve two DPs without the so-called copula pronoun, i.e., PRON. To illustrate the case further, consider the following copular wh-questions in (48) and their corresponding answers in (49).

\[(48)\]
\[
a. \text{miin hu?} \\
   \text{who the-teacher} \\
   \text{‘Who is he?'} \\

b. \text{miin haida?} \\
   \text{who this} \\
   \text{‘Who is this?’} \\

c. \text{mann humm?} \\
   \text{who they} \\
   \text{‘Who are they?’} \\
\]

\[(49)\]
\[
a. \text{hu l-mudarris} \\
   \text{he the-teacher} \\
   \text{‘He is the teacher’} \\

b. \text{haida rfee?-i} \\
   \text{this friend-mine} \\
   \text{‘This is my friend’} \\
\]
In terms of structure, since the two DPs in this type of construction are reversible, depending on which DP receives focus, I will proceed with the assumption that the wh-phrase originates in the base predicate position\textsuperscript{21}. The initial structure for (48a), for instance, is given in (50).

\begin{equation}
(50) \quad \text{TP} \quad \text{DP} \quad T' \quad | \quad | \quad hu \quad T \quad DP \quad | \quad | \quad He \quad | \quad miin \quad [+\text{Present, } +D] \quad \text{who}
\end{equation}

Because Jordanian is a wh-movement language, the Wh-DP \textit{miin} ‘who’ undergoes movement from the predicate position to the specifier position of a [+WH] C. The structural representation for such constructions is illustrated in (51)\textsuperscript{22}.

\textsuperscript{21}It is worth noting that in assuming that the wh-phrase in equative copular constructions originates in predicate position accounts for wh-in-situ in languages like Egyptian in which sentences like \textit{huwwa miin} ‘Who is he’ are fully well-formed, and are, in fact, the unmarked type of copular wh-questions in this language.

\textsuperscript{22}Notice that the analysis also allows for a derivation in which the wh-phrase originates as a clausal subject, i.e., in Spec,TP.

By analogy, the derivation of the string ʔayy ʔay hu ‘which thing it’ is the same as the derivation provided for the questions above. That is, the string ʔayy ʔay hu ‘which thing it’ can be said to have the same structure prior to movement of the wh-phrase ʔayy ʔay ‘which thing’. The Wh-DP ʔayy ʔay ‘which thing’ is base-generated in predicate position as shown in the structure in (52b).

(52) a. hu ʔey-š
it which-thing
‘Which thing/what is it?’

again, this structure accounts for Egyptian wh-questions such as ʔeww ʔeek? ‘What is it?’ as noted in footnote 21.
Since Jordanian is a wh-movement language, the wh-phrase ʔeys ‘which thing’ raises from the predicate position to Spec,CP of the copular clause in which it occurs. The overall structure is given in (53) below.

(53) a. ʔeys hu
    which-thing it
    ‘What is it?’

b. CP
    DP C’
    |  
    ʔeys
    What
    [+WH]
    DP T’
    |  
    hu T DP
    it
    [+Present, +D] ʔeys

To summarise, the discussion thus far has been concerned with morphosyntactic properties of the wh-expression ʃu. Having shown irregularities in the distribution of this wh-expression, and rejecting the assumption that it is a simple Wh-NP, I looked into other ‘what’ forms in several Arabic dialects and
observed that, unlike ‘who’, the ‘what’ forms are synthetic variants of the string ʔayy-u ʔay-in (huwa), used to replace the classical form maadha ‘what’. This string consists of a Wh-DP, itself made up of a wh-element, and an indefinite NP, and a DP pronoun. Syntactic derivation was then provided showing that this string is, in fact, a copular wh-clause.

If this analysis is correct, Lebanese šu receives the same treatment as Jordanian šu. Thus, contra Aoun and Li 2003 and Aoun, Benmamoun and Choueiri 2010, the wh-expression šu is not a bare wh-element, but, rather, an internally complex wh-expression with a CP structure, i.e, a Wh-CP.

One advantage of the present analysis is that we can now decompose the wh-element šen. In the previous chapter, the wh-expression šen-o ‘what’ was said to consist of the the wh-DP šen and the reduced pronoun o ‘it’. Recall that there is a group of Arabic dialects which use a replacement for the wh-phrase what. I have shown, for instance, that Lebanese and Jordanian šu are derived from the string ʔayy ʔay-hu, which is, in turn, a simplified version of the classical form ʔayy-u ʔay-in huwa. Unlike these dialects, however, Iraqi retained the nunation that marks the end of the Wh-DP ʔayy-u ʔay-in.

3.4.4 Evidence from other dialects

Evidence that sentential wh-phrases do exist, and are, as a matter of fact, a common phenomenon in Arabic, comes from other dialects such as Iraqi and Moroccan. In the previous chapter, I have shown that Iraqi men-o is a copular wh-phrase that is made up of the Wh-DP men/miin ‘who’ and a reduced form of the subject pronoun hu(wwa) ‘he’. There is a crucial difference between Iraqi men-o and Lebanese šu, however. Whereas the latter is a CP, the former is a DP. Such a difference can be attributed to the labelling algorithm and the way

---

24 I assume that the same holds for Palestinian šu.
25 The analysis also extends to the Egyptian ‘what’ form ʔeek. This form can be said to consist of the wh-DP ʔeek-š, except that š-suffix is dropped altogether in this language.
how wh-movement projects in each dialect. Recall that in Iraqi, movement of
the wh-phrase *men* ‘who’ is triggered by the need to check its [WH] feature, not
because of any operator feature on C\(^{26}\). As such, movement of *men* to Spec,C
projects the label of the wh-phrase, namely, a DP. Meanwhile, wh-movement
in Lebanese is a standard wh-movement (see section 3.2). That is, movement
of the wh-phrase *†ey-š* in Lebanese is triggered by the need to check the EPP
feature on C. As with classic cases of wh-movement, the projection that results
from this type of wh-movement is a CP.

Further evidence for the existence of clausal wh-phrases comes from Moroc-
can which uses phrases such as *†s-kuun* ‘who/what’. In the preceding chapter,
Moroccan *†s-kuun* ‘who/what’ was also said to consist of a wh-element and a ver-
bal copula. According to Ouhalla 2002, the wh-expression *†s-kun* consists of the
wh-element *š*, and inflection. This inflection is in fact the present tense form of
the verbal copula *kun* ‘be’ in Arabic with default agreement features, i.e., third
person masculine singular. Adding to Ouhalla, the wh-element *š* is actually a
reduced form of the Wh-DP *†eyš* analysed as such in this study. Since Moroc-
can is a wh-movement language (Ouhalla and Shlonsky 2002), the derivation
and the structure that may, therefore, be assigned for the wh-expression *š-kuun*
looks like that in (54b).

(54)  a.  \[ \text{pro} \ y-kun \ †ey-š \]
      \[ \text{pro} \ 3ms-BE \ which-thing \]

\(^{26}\)See chapter 2, section 4.1
Although Moroccan presents a different case, in that the string in (54c) includes a present-tense form of the copular verb in Arabic, the point to be made here is that Moroccan ʃ-kuun ‘who/what’ patterns with the Lebanese ʃu and Iraqi ʃen-o in that all of these expressions are copular wh-clauses.

Next, I turn to the question of how the CP-analysis of the wh-expression ʃu reflects on the syntax of wh-questions in which it occurs and how this analysis can account for the facts from Lebanese.

3.4.5 The Syntax of ʃu-questions

We turn now to the second issue raised earlier with regards to wh-questions in which the wh-expression ʃu ‘what’ occurs, the question one might ask is how such questions are derived in light of the analysis proposed here. Given that the wh-expression ʃu ‘what’ is clausal along with the fact that it has very limited
distribution, i.e., appearing only in a clause initial position, I propose a base-generation analysis for šu-questions along the lines of Koster’s 1978 analysis of sentential subjects and objects. The derivation of šu-questions would proceed as follows: the CP šu is directly merged in the specifier position of a [+WH] C and binds a phonetically null pronominal in the argument position. The structure I have in mind for wh-questions that involve the wh-expression šu ‘what’ like (55a) might be roughly represented as in (55b).

(55) a. šu šтарайте mbeerih
    what bought.2sf yesterday
    ‘What did you buy yesterday?’

b. 

Under this analysis, šu ‘what’ is a CP that is base-generated in Spec,CP of the matrix wh-interrogative in which it occurs. Such analysis is reminiscent of
that originally proposed in Koster 1978 and elaborated in the work of Alrenga 2005 for clausal subjects and objects in English and Dutch. Koster suggests that subject and object sentences do not exist. Instead, there are satellite sentences which bind a null DP inside the main clause (the Satellite Hypothesis)\textsuperscript{27}. Consider the following examples.

(56) a. Dat hij Komt, dat/∅ is duidelijk.
   that he comes that is clear
   ‘That he will come is clear.’

   b. That he will come ∅ is clear.

(57) a. Dat hij Komt, dat/∅ betreur ik t.
   that he comes that regret I t
   ‘That he will come I regret.’

   b. That he will come ∅ I regret t.

For the Dutch examples (56a) and (57a), Koster points out that the pronoun which corresponds to the subject or object is optional. The fact that the pronoun can be dropped led Koster to treat such sentences as sentential subject constructions. For the English cases (56b and 57b), Koster assumes that there is a Wh-DP that moves from the argument position to Spec,CP and gets obligatorily deleted. In this analysis, CP subjects and objects are base-generated in the specifier position of a CP much in the same way as topic phrases discussed in Chomsky 1977. Under Koster’s proposal, a sentence like (56b), repeated in (58a), would have the structure roughly represented in (58b).

(58) a. That he will come is clear.

\textsuperscript{27}see Koster 1978.
The CP specifier in (58b) is said to bind the null DP in thematic position. This DP, then, undergoes A’-movement to Spec,CP and eventually gets deleted. Koster maintains that the same is true of clausal objects like (57b) which has a similar derivation to (58) except that the null DP is in the complement position. This way, sentences which involve clausal subjects and objects are derived in a base-generation analysis doing away with the unlikely CP movement.

The whole point that Koster and Alrenga make is that clausal subjects and objects have very restricted distribution which cannot be accounted for transformationally. To support this view, Alrenga discusses cases of CP-topicalisation which cannot be derived under a movement analysis. Consider the following examples (Alrenga 2005).

(59)  
  a. That languages are learnable, his theory fails to capture.  
  b. *The grammar captures that the rule is obligatory.  
  c. The grammar captures the fact that the rule is obligatory.
The asymmetries above do not readily follow from an analysis of topicalisation in which the CP topic itself undergoes $\lambda'$-movement from the gap position. Otherwise, the inability of verbs like *capture* in (59b) to occur with an in-situ CP complement is left unexplained. Notice also that the sentence in (59c) shows that the gap position in CP-topicalisation constructions must be a position in which a DP argument must be present. This line of analysis continues to gain widespread acceptance since it captures a number of phenomena where CP and DP asymmetries are observed (see Alrenga for a discussion of the full range of facts)\textsuperscript{28}.

The base-generation analysis of clausal elements fits the facts observed in Lebanese, especially the view that such elements are base generated in Spec,CP and are licensed by a special kind of C. Presumably there is a [WH] version of this C, which I take to be the case for $\ddot{s}u$-questions. Within the current analysis, this [+WH] C would license the CP $\ddot{s}u$ ‘what’ in its specifier. Notice, however, that whereas Koster’s and Alrenga’s clausal subjects and objects are indistinguishable from base-generated topics, the current analysis does away with the extra CP layer which hosts them. More precisely, clausal $\ddot{s}u$ is directly merged in Spec,CP of the wh-question, just like cases where null [WH] operators are said to originate in Spec,CP of an interrogative clause binding wh-elements inside the clause. Altogether, the structure for $\ddot{s}u$-questions (see (55) above) is comparable to null operator constructions discussed in Chomsky 1977, except that under the present analysis, the WH-Operator is not null, but, rather, an overt complex operator.

A number of consequences follow from the analysis just outlined. First, as a CP, $\ddot{s}u$ can only be generable clause initially, i.e, in Spec,C of the matrix CP. More precisely, $\ddot{s}u$ cannot appear in the same distributional contexts where a

\textsuperscript{28}Koster’s hypothesis, however, i.e., that sentential subjects do not occupy canonical subject position, has been challenged in Davies and Dubinsky 1999, 2000 and Haegeman 2010 who argue that such subjects may, at least for some speakers, occupy the subject position. See these studies for a full discussion.
DP is selected for. This, in turn, explains why šu never occurs in situ in a DP position as we have seen in earlier sections. Secondly, cases where a verb selects for a CP complement are predicted to be well-formed, as will be shown shortly. A third consequence is that all the problems related to d-linking are deemed irrelevant under this analysis. Before implementing this analysis for Lebanese, I address two remaining issues regarding movement diagnostics and the PF form of the string ʔayy ʔay ʔu ‘which thing it/he’ in the next two subsections.

Movement Diagnostics and Binding

The current analysis gives rise a number of issues to do with the binding relationship between the Wh-operator šu and the null pronominal DP. Given that there is a pro in the base-position that is bound by the overt Wh-operator šu, it is predicted that the traditional movement diagnostics, such as reconstruction, islandhood and intervention effects, should be absent. This section throws some light on this prediction.

Starting with reconstruction, as is well known, reconstruction corresponds to the interaction between movement and interpretation (Chomsky 1995, Aoun and Benmamoun 1998, Aoun and Li 2003, Malkawi and Guilliot 2007, Aoun, Benmamoun and Choueiri 2010). Given the copy theory of movement (Chomsky 1995), reconstruction effects arise if the base copy of the moved element feeds interpretation. To illustrate, consider the following example.

(60) a. Which picture of himself does everyone like?
    b. Which picture of himself does everyone like (Which picture of himself)

In (60), the reflexive pronoun himself is licensed under the interpretation of the copy of the wh-phrase which picture of himself in the base argument position, as shown in (60b). Otherwise, the reflexive pronoun would have no antecedent and the sentence would be ruled out.
Concerning reconstruction in Lebanese šu-questions and the question of whether its effects are detectable on a par with the English case above, the answer is negative, however. This is attributed to the fact that the wh-expression šu is only a single lexical item, which makes it impossible to test in contexts where reconstruction might be involved.

Turning to intervention effects such as the Highest Subject Restriction (McCloskey 1990) which normally bars binding of a pro in the highest subject position within a CP, the question is whether the HSR effects are observed in šu-questions. As might be recalled from the previous chapters, the HSR is not operative in Lebanese (Aoun, Hornstein and Choueiri 2001) and Jordanian and cannot therefore be used to diagnose the relationship between the wh-expression šu and the null pro that it binds. Contra standard assumptions, subject pronouns in Jordanian, for instance, may appear as resumptives in the highest subject position in various constructions such as clitic left-dislocation and relative clauses.

(61) a. l-walad haad smešna ?onno hu katab ?eSSa (Jordanian)
    the-boy this heard.1pl that he wrote.ms story
    ‘This boy, we heard that he wrote a story.’

   b. haad l-walad illi hu katab ?eSSa
      this the-boy that he wrote.ms story
      ‘This is the boy who wrote a story.’

As concerns islands, they have been traditionally used to test whether a given structure or a relationship is derived by movement. However, recent work on locality effects, especially, island effects, has revealed that actually they are not diagnostics for movement, rather, they are diagnostics for a dependancy relation, namely, Agree (Adger and Ramchand 2003 and Boeckx and Hornstein

29Similarly, Aoun, Choueiri and Hornstein 2001 point out that the HSR is not operative in Lebanese since strong subject pronouns do appear in the highest subject position of an embedded clause.
2008). For instance, Boeckx and Hornstein 2008 argue along the lines of Adger and Ramchand (2003:2) who observe that “given recent approaches to syntax, locality effects can no longer be assumed to be diagnostic of movement . . . . This is because any locality effect can be construed as deriving from constraints on the Agree operation (which is a necessary precursor to movement), rather than on the Movement operation itself.”

The non-availability of diagnostic tests for the relationship between the wh-CP șu and its pro should not be taken, however, to invalidate the analysis presented here. Whereas some of the restrictions mentioned here, such as the HSR, need to be re-instated on the basis of the data presented in this study, others have already been rendered inapplicable, i.e. island effects.

The second issue that is also raised by the current analysis is whether the binding mechanism adopted here allows binding between șu and its pro across islands like unselective binding of Pesetsky. Recall that under Pesetsky’s unselective binding, the lowest wh-element in a sentence like (62) can stay in situ if it is d-linked and be bound by the highest wh-element.

(62) Which man did you hear the rumour that which woman kissed.

However, binding within the current analysis is a slightly different phenomenon in that islands form syntactic barriers to the binding relationship between the Wh-operator in Spec,CP and the null DP in argument position. Assuming along the lines of Adger and Ramchand 2005, the fact that șu cannot bind its pro across islands can be attributed to syntactic binding, rather than pure semantic binding. In other words, the binding mechanism adopted here differs from the traditional mechanism of unselective binding advanced in Pesetsky 1987. While the latter allows binding across island boundaries, the former does not.
The PF form of šu

The last issue that is still at stake concerns the PF form of the string ḫay ṣay hu and how it is spelled out as šu, given that šu-questions contain a complex configuration with a left-branching specifier, i.e., the CP šu. The structure is repeated in (63) for convenience.

(63)

To settle this issue, I follow Starke’s 2011 Phrasal Spell-Out (PSO) approach and Uriagereka’s 1999 syntactic model of Multiple Spell-Out (MSO) and take šu as a complex case of contraction that can be accounted for if left-branching structures form a self-contained spell-out domain. The PSO is a lexical insertion procedure whereby non-terminal nodes in a syntactic tree can be targeted by early Spell-Out (see also Fábregas 2011). Under this approach, the entire sequence ḫay ṣay hu would be spelled out as one chunk, i.e., as a single lexical

180
In terms of the relevant ordering of the terminals within the spelled-out structure, Uriagereka suggests also that it may be fixed by Kayne’s 1994 Linear Correspondence Axiom which states that if $\alpha$ asymmetrically c-commands $\beta$ in the syntactic structure, $\alpha$ precedes $\beta$ in the linear ordering\textsuperscript{30}. Once the ordering is fixed, the string ?eyš hu undergoes early spell-out to the phonological interface and re-enters the derivation as “a frozen giant lexical compound” (Uriagereka 1999). The compound is later plugged into where it fits in within the entire derivational process\textsuperscript{31}. Now re-consider the domain for Spell-Out presented in the structure of this string in (64).

\textit{(64)}

\begin{itemize}
  \item \textit{CP}
  \item \textit{CP',}
  \item \textit{DP,}
  \item \textit{C',}
  \item \textit{C,}
  \item \textit{TP,}
  \item \textit{[+WH,}
  \item \textit{DP,}
  \item \textit{T',}
  \item \textit{T,}
  \item \textit{pro,}
  \item \textit{v*P,}
  \item \textit{v*}
  \item \textit{VP,}
  \item \textit{v*}
  \item \textit{VP,}
  \item \textit{DP,}
  \item \textit{pro,}
\end{itemize}

\textsuperscript{30}See also Uriagereka (1999: 252 for a discussion of the two steps that make up LCA, only one of which is kept in Uriagereka’s system, that is, the Base Step.

\textsuperscript{31}The rest of the tree is derived via Chomsky’s Phase Theory, which I leave aside for the moment since the main focus of this section is on the PF form šu.
There are two phonological operations that may be pointed out at this stage. First, there is a general loss of the glottal stop in the dialects which results from its combination with the diphthong in the NP \( \tilde{s}a\) \(\tilde{a}y\). Secondly, since the new string occurs always with the pronoun \(hu\), this leads to the loss of \([h]\) phoneme\(^{33}\).

The structural representation for (64) might be given in (65) below.

\[
(65)
\]

![Diagram](image)

Such a phenomenon is found in similar combinations such as that of Iraqi Arabic wh-expression \(men-o\) ‘who’ which derives from the string \(men hu\) ‘who he/it’. Negation contexts also provide further evidence for the loss of \([h]\) phoneme of this particular form of pronoun. Consider the following examples from Jor-

---

\(^{32}\)Note that the formation of ‘what’ forms which are derived from the string \(?ay\) \(\tilde{a}y\) \(hu\) ‘which thing it/he’ are also subject to phonological variations in each of the dialects.

\(^{33}\)The same view is held in Cowell (1964) who points out that \(\tilde{s}u(u)\) that is limited to Greater Syria (that is Syria, Lebanon, Jordan and Palestine) and that it can lose stress thereby reducing to the consonant \(\tilde{s}\). The latter being the strongest consonant, while others are defective, in a tri-consonantal system. Such operations are also common due to frequency in use (Fassi-Fehri 1993).
The sentences above contain two DPs separated by what looks like a negative particle. However, there is no such construction in Arabic as DP NEG DP. Under the current proposal, (66) receives a straightforward account if we assume that the negative particle *muš* is a compound expression that consists of the negative particle *maa*, the subject pronoun *hu* and the indefinite (NP) *š(i)* ‘thing’. In other words, the negative particle *muš* is a reduced compound of the string formed by *maa* + *hu* + *š(i)*. An additional observation suggesting that this is indeed the case comes from emphatic interpretations for sentences like (66) whereby the three elements, assumed here to make up the negative compound, are pronounced separately, as illustrated in (67) below.

\[(67)\]  
\[a. \text{Ali ma-hu-(š) l-muhandis} \quad \text{(Jordanian)}\]
\[\text{Ali Neg-he-Neg the-engineer} \]
\[\text{‘Ali is not the engineer?’}\]

\[b. \text{Magdi muš l-haraami} \quad \text{(Egyptian)}\]
\[\text{Magdi Neg the-thief} \]
\[\text{‘Magdi is not the thief’}\]

It is noteworthy that third person strong pronouns consist of two phonemes, which are structurally represented in (i) below (see also Aoun, Choueiri and Hornstein 2001 and Ouhalla 2001). The phoneme /h/ is phonologically deficient when it is intervocalic. As such, once it is preceded by a lexical item ending with a consonant, it is dropped. What remains after this dropping, is the element that bears φ-features, (pronounced as) u(wwa). However, when it is in a syllable-initial position, /h/ cannot be dropped. This means that, we do not get cases where the pronoun *hu* is amalgamated into the wh-word *pay-š* in the sequence *hu pay-š*. The phenomenon of h-dropping is also found in other languages such English, especially, the Cockney variety. For a discussion see Al-Tamimi 2002.

\[\text{(i)}\]
\[\text{DP}\]
\[\text{Spec} \quad D\]
\[\text{h} \quad \text{φ} \quad -u(\text{wwa})\]

\[\text{34It is noteworthy that third person strong pronouns consist of two phonemes, which are}\]
\[\text{structurally represented in (i) below (see also Aoun, Choueiri and Hornstein 2001 and Ouhalla}\]
\[\text{2001). The phoneme /h/ is phonologically deficient when it is intervocalic. As such, once it}\]
\[\text{is preceded by a lexical item ending with a consonant, it is dropped. What remains after}\]
\[\text{this dropping, is the element that bears φ-features, (pronounced as) u(wwa). However, when}\]
\[\text{it is in a syllable-initial position, /h/ cannot be dropped. This means that, we do not get}\]
\[\text{cases where the pronoun *hu* is amalgamated into the wh-word *pay-š* in the sequence *hu pay-š*.}\]
\[\text{The phenomenon of h-dropping is also found in other languages such English, especially, the}\]
\[\text{Cockney variety. For a discussion see Al-Tamimi 2002.}\]
\[\text{35see also Eid 1981 for a similar analysis of Egyptian.}\]
b. Magdi ma-hwaa-š l-haraami
   Magdi neg-he-neg the-thief
   ‘Magdi is not the thief’

3.4.6 Lebanese wh-in-situ revisited: predictions

To the extent that the analysis above is correct and given the fact that Jordanian and Lebanese are closely related, the conclusion that the wh-expression šu is a complex Wh-CP (overt) operator appears to be also correct for Lebanese. A second reason to believe that this analysis is on the right track is that it captures all the relevant facts from Lebanese regarding asymmetries in the distribution of wh-phrases observed earlier. The analysis predicts that if it is possible to have šu in situ at any point, the only places where it is possible to have it are positions where a CP is selected for.

As far as single wh-in-situ questions are concerned, it is impossible for the wh-expression šu, formally a CP, to appear in the same distributional contexts as DPs, hence, its ill-formedness in argument position in (68b) below.

(68) a. šeft miin mbeerih?
   saw.2sm who yesterday
   ‘Who did you see yesterday?’

b. *štarayte šu mbeerih
   bought.2sf what yesterday
   ‘What did you buy yesterday?’

Under the analysis proposed here, the contrast between (68a) and (68b) is explained syntactically. That is, the possibility of the wh-phrase miin ‘who’ in situ is based on its structural status as a DP while the impossibility of in-situ šu is a direct consequence of its status as a CP that cannot occupy a DP position. Likewise, in multiple wh-questions, we find the following pattern.
The fact that multiple wh-questions do not allow the wh-expression šu in situ is also captured on the grounds that it is a CP that cannot occur in a DP position. On the other hand, šu can occur in situ in multiple wh-questions but only when its selected for by an appropriate verb. This is shown in the following example.

Similarly, in single wh-questions, šu is expected to occur in situ as complement for verbs that select for CPs. This prediction is borne out. Consider the following examples from Lebanese.36

Notice that the cases above are impossible to capture in terms related to d-linking. This is a prediction that is not made by Aoun, Benmamoun and

---

36There might be a problem with echo questions where šu appears in DP positions such as šurayš šu? ‘You bough WHAT?’. I will assume, however, that echo-questions have a particular kind of syntax which requires the in-situ elements to be turned into free relative DPs.
Choueiri’s 2010 analysis. In terms of its referential properties, especially resumption and d-linking, it is difficult to see how the CP šu can either have an antecedent in discourse or be able to share reference with a resumptive pronoun. More precisely, d-linking and resumption are properties that characterise DPs, and the issue of whether or not šu can be d-linked is out of question due to the fact that it is a CP and has very restricted distribution. This way, the current analysis offers a unified syntactic approach that does away with the problems that encounter the d-linking approach.

3.5 Conclusion

I have argued in this chapter for a base-generation analysis of Lebanese (as well as Jordanian) šu ‘what’-questions. The chapter started by highlighting the problem between Lebanese argument wh-phrases in Lebanese single wh-in-situ questions as originally observed in Aoun and Choueiri 1999 and Aoun, Benmamoun and Choueiri 2010. I have, then, looked into properties and distribution of argument wh-phrases in other wh-constructions including wh-movement questions, resumptive wh-questions and d-linked wh-questions. The source of the problem was shown to lie in the behaviour of the wh-expression šu ‘what’ that displays properties that distinguish it from other argument wh-phrases like miin/mann ‘who’ and ?ay(ya)-NP ‘which’-NP phrases. I, then, presented Aoun, Benmamoun and Choueiri’s 2010 solution which is based on the assumption that the wh-in-situ asymmetry is a consequence of the (in)ability of a given wh-phrase to be d-linked. However, this assumption was deemed irrelevant on the basis of new data showing that the asymmetry between Lebanese argument wh-phrases extends to wh-in-situ contexts other than single wh-in-situ questions, namely, multiple wh-questions. It was concluded at this stage that a d-linking approach cannot account for the overall asymmetry observed and a syntactic approach was suggested, instead.
The proposal I put forward began with the hypothesis that the wh-expression šu is not what it looks like, i.e., a single wh-element, but rather a complex wh-expression. Defending the hypothesis, I reported crucial facts from other dialects to uncover the morphological composition of the wh-expression šu. Upon closer examination of the various wh-forms in other dialects, especially, the ‘what’ forms, the ‘who’ forms and the ‘which’-NP forms, I have argued that the wh-expression šu is a copular wh-clause deriving form the string ?ayy šay hu ‘which thing he/it’ in a similar fashion to Iraqi meno and šeno as well as Moroccan škuun. The table below shows the equivalent of each of these wh-forms in all the dialects examined here.

<table>
<thead>
<tr>
<th>(72)</th>
<th>SA</th>
<th>Iraqi</th>
<th>Jordanian</th>
<th>Lebanese</th>
<th>Moroccan</th>
</tr>
</thead>
<tbody>
<tr>
<td>mann</td>
<td>men</td>
<td>miin</td>
<td>miin</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>‘who’</td>
<td>‘who’</td>
<td>‘who’</td>
<td>‘who’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mann huwa</td>
<td>men-o</td>
<td>man-hu</td>
<td>-</td>
<td>š-kuun</td>
<td></td>
</tr>
<tr>
<td>‘who he’</td>
<td>‘who-he’</td>
<td>‘who-he’</td>
<td>-</td>
<td>‘which thing-be’</td>
<td></td>
</tr>
<tr>
<td>maadba</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>‘what’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>?ayy-u šay-in</td>
<td>šen</td>
<td>?ay-š</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘which thing’</td>
<td>‘which thing’</td>
<td>‘which thing’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>?ayy-u šay-in huwa</td>
<td>šen-o</td>
<td>šu</td>
<td>šu</td>
<td>š-kuun</td>
<td></td>
</tr>
<tr>
<td>‘which thing it’</td>
<td>‘which thing it’</td>
<td>‘which thing it’</td>
<td>‘which thing it’</td>
<td>‘which thing-be’</td>
<td></td>
</tr>
</tbody>
</table>

Syntactically, I have shown that the string ?ayy šay hu ‘which thing he/it’ is a CP involving internal movement of the Wh-DP ?ay-š ‘which thing’ from the predicate (lower DP) position to Spec,CP of the copular clause. In light of the CP-analysis of the wh-expression šu, I turned into addressing the question about the syntax of wh-questions in which it occurs and how it differs from canonical wh-movement questions. I have argued for a base-generation account...
along the lines of Koster’s 1978 and Alrenga’s 2005 analysis of sentential subjects and objects. I have proposed that the CP šu is generated in Spec,CP of the main wh-clause where it binds a null DP in the argument position inside the clause. The analysis proposed for šu-questions was said to fit in with null operator constructions discussed in Chomsky 1977, except that under the present analysis, the WH-Operator is not null, but is, rather, an overt complex operator.

The last issue raised was the PF form of šu which was resolved on the basis of Starke’s 2011 Phrasal Spell-Out approach and the MSO model of Uriagereka 1999. Following these accounts, I have taken šu, being a left-branching complex specifier, to be spelled-out as one chunk and sent out to the phonological component that gets later merged in Spec,CP of the main clause.

A number of consequences followed from the CP analysis. First, the fact that šu appeared to be found only in a clause initial position was explained on the grounds that it is a CP and can never occur in a DP position. I have also shown that šu can occur as a complement in-situ for verbs that select for CP. A second consequence was that all the problems associated with the d-linking solution were neatly dealt with from a purely syntactic point of view.

Above all, the analysis followed a micro-parametric approach based on morphological properties of wh-phrases showing how this affects their external syntax. The analysis also gave us insight into the mechanism that (dis)allows wh-in-situ in Lebanese and Jordanian and, hence, solved the asymmetries observed in all wh-contexts not only in Lebanese, but also in Jordanian, with implications for other similar dialects.

Finally, the chapter closed with an unresolved question about the conflict observed between CP/DP properties of the wh-expression šu. This will be dealt with in the following chapter.
Chapter 4

Wh-cleft questions

4.1 Introduction

One issue that carries over from the previous chapter is the conflict between the non-argument properties vs. the argument properties of the wh-expression šu. Concerning the non-argument properties, it was shown, for instance, that the wh-expression šu cannot occur in situ in argument position, be it in single wh-in-situ questions (1) or in multiple wh-questions (2); nor can it be d-linked (3) or be related to a resumptive pronoun (4), unlike argumental Wh-DPs such as miin ‘who’ and ?ayya-NP ‘which’-NP phrases.

(1) a. šeft miin mbeeri? (Lebanese)
    saw.2sm who yesterday
    ‘Who did you see yesterday?’

        b. *starayte šu mbeeri
            bought.2sf what yesterday
            ‘What did you buy yesterday?’

(2) a. miin zaar miin?
    who visited.3ms who
    ‘Who visited whom?’
b. *miin ˇseef ˇsu
   who saw.3ms what
   ‘Who saw what?’

(3) Fii ˇfonde bluuze hamra w bluuze sawda
   in-it at-me shirt.fs red.fs and shirt.fs black.fs
   ‘I have a red shirt and a black shirt’

a. ˇlayya wahdo baddak tˇzarrib
   which one.fs want.2sm try.2sm
   ‘Which one do you want to try on?’

b. *ˇsu baddak tˇzarrib
   what want.2sm try.2sm
   ‘What do you want to try on?’

(4) a. miin ˇsaft-o b-l-maˇTjam
   who saw.2sm-him in-the-restaurant
   ‘Who did you see in the restaurant?’

b. *ˇsu ˇstarayt-i mbeerih
   what bought.2sf-it yesterday
   ‘What did you buy yesterday?’

To account for the contrasts above, I have proposed that whereas argument
wh-phrases like miin and ˇlayya-NP ‘which’-NP are DPs, the wh-expression ˇsu
is a CP, i.e., a wh-clause. The non-argument behaviour of the wh-expression ˇsu
was thus explained on the basis of its status as CP that can only be generated
clause initially. To this effect, I have argued that ˇsu-questions have different
derivation from wh-questions that involve argument wh-phrases like miin ‘who’
and ˇlayya-NP ‘which’-NP. Whereas the latter are wh-movement questions, the
former are base-generated wh-questions. The conclusion was that ˇsu-questions
differ syntactically from canonical wh-movement questions.
On the other hand, one property that is typically associated with argumental Wh-DPs and one that remained unexplained is the ability of šu to occur in positions that are known to be restricted to DPs. In particular, the wh-expression šu was shown to be well-formed when it occurs in the left-most position of equative copular wh-constructions and the so-called reduced cleft wh-questions (Cheng 1991, Ouhalla 1996 and and Shlonsky 2002)\(^1\). The relevant sentences are repeated below (the Lebanese examples are Aoun, Benmamoun and Choueiri’s (66), p151).

(5) a. šu (huwwe) l-ʔakl l-yom? what (it) the-food the-day
   ‘What is on the menu today?’

   b. šu (huwwe) lii rah teeklu-u l-yom?
      what (it) that Fut. eat.2p-it the-day
   ‘What is it that you are going to eat today?’

(6) a. šu (hu) l-soʔaal?
      what (it) the-question
   ‘What is the question?’

   b. šu (hu) lii baddak tesʔalu-h?
      what (it) that want.2ms ask-it
   ‘What is it that you want to ask?’

The fact that the wh-expression šu occurs clause-initially in such constructions followed by a resumptive subject pronoun is puzzling given that it is not a DP and cannot be related to a resumptive pronoun. This property of šu is in conflict with its non-argument properties and seems to posit a challenge to the CP-analysis proposed for wh-questions.

This chapter argues that the conflict just observed can, in fact, be accounted for only if we maintain the CP analysis of šu. I will, first, identify properties

\(^1\)see Chapter 1 for a discussion of copular constructions in Arabic.
of (wh-) clefts in several Arabic dialects and show that such properties are similar to those displayed in copular constructions. I, then, propose an analysis that assimilates wh-clefts to copular wh-constructions in Arabic. I will also show that the proposed analysis extends to other Arabic dialects. As far as Lebanese wh-clefts are concerned, I will argue that they are cases of d-linked wh-in-situ by virtue of the fact that they involve presupposition: a case of d-linking. Based on the fact that šu cannot occur in situ, wh-clefts that involve this wh-expression (5b) and (6b) are treated on a par with verbal wh-questions, i.e., as base-generated questions in which šu is merged as a complex overt operator in Spec,CP binding a null DP in the argument position. Once again, the CP-analysis gains extra support in uniformly accounting for the asymmetries observed in Lebanese.

The sections of this chapter are organised as follows: section 2 presents (wh-) clefts in Arabic and the properties associated with such constructions. Section 3 discusses two types of analysis for Arabic wh-clefts, namely, Cheng’s 1991 analysis for Egyptian wh-clefts and Shlonsky’s 2002 analysis for Palestinian wh-clefts. Although both analyses have much to commend them, they will be shown to be empirically problematic. In section 4, I put forward an alternative analysis for (wh-) clefts in Arabic. Section 5 returns to šu-wh-clefts and tackles the question of how such questions are derived in light of the findings of this chapter. Section 6 is the conclusion.

4.2 Arabic (Wh-)Clefts

Arabic dialects have an additional strategy for the formation of wh-questions which has several properties that distinguish it from canonical wh-questions. This strategy involves the occurrence of a definite argument wh-phrase, i.e., a Wh-DP, in the left-most position followed by (what has been taken as an optional copula) pronoun and a clause marked by the complementiser ʾillī ‘that’ of
the type found in relative clauses. To illustrate, consider the following paradigm from a number of Arabic dialects.

(7) a. miin (huwwe) illi šoft-o b-l-maTyum? (Lebanese)
who (he) that saw.2sm-him in-the-restaurant
‘Who is it that you saw in the restaurant?'

b. miin (hu) illi hall l-muškele? (Palestinian)
who (he) that solved.3ms the-problem
‘Who is it that solved the problem?’

c. miin (huwwa) illi Darab boš (Egyptian)
who (he) that hit.3ms Bush
‘Who is it that hit Bush?’

d. men(-o) illi Mona šafat-uh (Iraqi)
who(-he) that Mona saw-him
‘Who is it that Mona saw?’

e. škun (huwwa) illi mša? (Moroccan)
who (he) that left.3ms/left.3fs/left.3mpl
‘Who is it that left?’

f. miin(-hu) illi katab haay l-riwaayeh? (Jordanian)
who(he) that wrote this the-novel
‘Who is it that wrote this novel’

Although the questions above come from different dialects, they exhibit the same schemata that consists of: a Wh-DP, a subject pronoun (PRON), and a relative clause (RC). This pattern is found in Arabic clefts which involves an initial definite DP, PRON, and a *illi-RC*. Consider the following paradigm.

(8) a. l-walad huwwe illi xazza? l-kteeb (Lebanese)
the-boy he tore.3ms the-book
‘It is the child that tore the book’

2 Notice that PRON in the Iraqi wh-phrase *men-o* is reduced, and, therefore, attached, to the wh-phrase *miin*. The latter is also reduced into *men* as a result. See Chapter 1 for a detailed discussion.
Apart from the difference in the nature of the initial DP, there seems to be another difference to do with PRON between the cleft sentences in (8a-d) and their wh-counterparts in (7a-f). Though such constructions have received little attention, the predominant view in the literature is that PRON is obligatory in the cleft sentences in Arabic while it is optional in wh-clefts (see Ouhalla 1999, Shlonsky 2002 and Aoun, Benmamoun and Choueiri 2010).

On the other hand, the pattern exhibited by Arabic clefts is similar to the one found in English clefts despite the fact that Arabic, unlike English, uses no copula in such constructions. Compare, for instance, the cleft sentences in (8) above with the ones in English shown in the following examples.

(9) a. It was John that Mary saw.

b. It was the book that Mary bought.

As can be seen in (9a-b), English clefts consist of, in addition to the DP and the relative clause, the pronominal subject it that is equivalent to PRON in Arabic clefts.
In terms of interpretive properties, both Arabic clefts and English involve presupposition. For instance, the sentence in (8d) involves exhaustive presupposition whereby the only contextually relevant person that wrote the novel is Sama’s friend, and no one else. Likewise, Reeve 2010 shows that in English, a cleft sentence like (10a) below involves existential presupposition. That is, (10a) obligatorily presupposes that there is someone that Mary hit. By contrast, presupposition in (10b) is not obligatory.

(10) a. It was John that Mary hit.
    b. Mary hit John

Presupposition is also a property of wh-clefts in Arabic of the type presented here. For instance, Shlonsky 2002 observes that a sentence like (7b) in Palestinian, (repeated in (11a) below), presupposes that a person actually solved the problem and asks for the identity of that person. By contrast, canonical wh-questions involve no such presupposition. The question in (11b), for example, simply asks who solved the problem and does not presuppose that anyone actually solved the problem.

(11) a. miin (hu) illi hall l-muşkele? (Palestinian)
    who (he) that solved.3ms the-problem
    ‘Who is it that solved the problem?’

    b. miin hall l-muşkele?
    who solved.3ms the-problem
    ‘Who solved the problem?’

In spite of these similarities, Arabic clefts differ from English clefts in a number of ways. A major distinction can be drawn between argument and non-argument phrases. Unlike English clefts (12a and b), Arabic clefts are restricted to definite DPs only (see also Ouhalla 1999). That is, both bare indefinite noun phrases, (13a) and (14a), and specific indefinite noun phrases, (13b) and (14b),

3See E Kiss 1998 on both types of presupposition.
cannot undergo clefting. The contrast is shown below.

(12)  a. It was an engineer that Nada married.
     b. It was a book that Nada bought.

(13)  a. *muhandis huu illi Nada tjawwazat-uh an-engineer he that Nada married.3fs-him ‘It was an engineer that Nada married’
     b. *muhandis b-l-ʔtissaalaat huu illi Nada tjawwazat-uh engineer in-the-communication he that Nada married.3fs-him ‘It was a communication engineer that Nada married’

(14)  a. *ktaab huu illi Nada štarat-uh a-book it that Nada bought.3fs-it ‘It was a book that Nada bought’
     b. *ktaab nahw ʔarabi huu illi Nada štarat-uh book syntax Arabic it that Nada bought.3fs-it ‘It was a book on Arabic syntax that Nada bought’

Another distinction between Arabic and English concerns the possibility of clefting PPs and adverbials phrase (AdvPs). Whereas English allows the clefting of PPs and AdvPs, Arabic does not. The contrast between (15) and (16) illustrates this point.

(15)  a. It is for this reason that Bill left.
     b. It was very angrily that John left the room.

(16)  a. *b-s-ʔaariʕ had illi Raami šaaf beyt on-the-street this that Raami viewed.3ms house ‘It is on this street that Raami viewed a house.’

In addition, notice that Arabic clefts differ from English clefts in that they involve the use of resumptive clitics inside the cleft clause where the relativised DP corresponds to an object position inside *illi-RC, as will be discussed later on.
b. *b-sorYa illi Nada raahat yala amman
  in-a-hurry that Nada went.3fs to Amman
  ‘It was in a hurry that Nada went to Amman.’

This is also true of the wh-counterparts of cleft sentences where only argumental Wh-DPs are clefted, PPs and adverbial wh-phrases cannot\(^5\). Consider the following examples from Palestinian (Shlonsky 2002).

(17) a. *la-miin illi ?enti ba`ati maktuub
    to-who that you(F) sent a-letter
    ‘To whom did you sent a letter?’

b. *winta illi katabti l-maktuub?
    when that wrote.3fs the-letter
    ‘When did you write the letter?’

c. *keef illi fahasti s-sayyara?
    how that examined.3fs the-car
    ‘How did you examine the car?’

To recap, this section showed how a productive type of wh-questions in Arabic differs, syntactically and semantically, from canonical wh-questions. In addition to similarities in their interpretive properties, the fact that they consist of a fronted (Wh-) DP followed by a subject pronoun and a relative clause suggests a resemblance with English cleft constructions. Nevertheless, the data from Arabic show that some differences exist between Arabic and English, especially, the ban on indefinite DPs, PPs and adverbials as well as the optionality of the pronoun in Arabic wh-clefts vs. its obligatoriness in other cleft sentences.

In the next section, I discuss Cheng’s 1991 analysis for Egyptian and Shlonsky’s 2002 for Palestinian which attempt to account for some of the properties that have been observed here.

\(^5\)When these items are fronted, some other syntactic operation is taking place, such as topicalisation (see Cheng 1991). Alternatively, such elements may be focused in-situ (see Ouhalla 1999).
4.3 Previous Analyses

This section presents two types of analysis for Arabic wh-clefts. One type of analysis is outlined in Cheng 1991 in which Egyptian wh-clefts are treated on a par with Irish reduced clefts discussed in McCloskey 1979. The other type of analysis is the equative analysis advanced in Shlonsky 2002 for Palestinian and adopted for Lebanese in Aoun, Benmamoun and Choueiri 2010. Although both types of analysis have their own merits, they will be shown to be empirically problematic.

4.3.1 The Reduced Clefts Analysis: Cheng 1991

Amongst early analyses in which the term ‘wh-clefts’ began to appear in Arabic is that of Cheng 1991. In particular, Cheng’s analysis assimilates Egyptian wh-fronting questions to cleft constructions. She proposes that Egyptian wh-fronting is an instance of the reduced cleft construction in the sense of McCloskey 1979. Although Cheng’s analysis is not meant to address all the issues related to Arabic wh-clefts, it has some interesting aspects that deserve some discussion. That said, this section presents the main arguments in the analysis, especially, the attempt to assimilate Egyptian wh-fronting questions to Irish reduced clefts.

One of the key issues Cheng’s 1991 study deals with is the apparent optionality between wh-fronting questions and wh-in-situ questions. Cheng observes that a number of wh-in-situ languages use, in addition to a wh-in-situ strategy, an optional wh-fronting strategy as a second alternative. Amongst these languages is Egyptian Arabic. Consider the following examples from Egyptian.

(18) a. miin illi Mona šaafit-uh ʔmbaarih? (Egyptian)
who that Mona saw.3fs-him yesterday
‘Who is it that Mona saw yesterday’

Cheng 1991 also discusses other languages similar to Egyptian like Bahasa Indonesia and Palauan. In this chapter, I will be concerned only with Egyptian.
b. Mona šaafit miin ?mbaarih?
Mona saw.3fs who yesterday
‘Who did Mona see yesterday?’

The optionality between wh-fronting questions like (18a) and wh-in-situ questions like (18b) appears to constitute a challenge for Cheng’s Clause Typing Hypothesis under which any given language is predicted not to have an option as to how it types a clause. However, Cheng proposes that this optionality is only apparent in that wh-in-situ and wh-fronting are two different types of question. She observes that the properties associated with wh-fronting in Egyptian are similar to those of cleft constructions. More precisely, the fact that such questions can occur only with a (Wh-)DP in the left periphery and the fact that this DP is always followed by an illi-clause suggest a resemblance with cleft sentences.

Based on McCloskey’s 1979 analysis for Irish clefts, Cheng posits that the syntactic structure of Egyptian wh-fronting questions is a reduced version of the normal cleft. In this respect, Cheng observes that McCloskey employs the notion of a reduced cleft for a certain type of cleft found in Irish in which an indefinite argument phrase undergoes clefting. It is reduced in the sense that the copula, which is optional in clefts with a definite argument phrase, cannot appear in this case. To illustrate, consider the following cases from Irish normal clefts (19) and reduced clefts (20).

(19) a. Is é Seán Bán a\(^b\) d’inis an scéal dom
Cop Agr Seán Bán Comp told the story to-me
‘It was Seán Bán who told the story to me’

---

7The Clause Typing Hypothesis states that ‘Every Clause needs to be typed. In the case of typing a wh-question, either a wh-particle in C\(^0\) is used or else fronting of a wh-word to the Spec of C\(^0\) is used, thereby typing a clause through C\(^0\) by Spec-head agreement’ (Cheng 1991: 29).

8These examples are provided in Cheng (1991: 60) which are, in turn, taken from McCloskey (1979: 90-91).
Cheng extends the reduced cleft analysis to Egyptian wh-fronting questions. She takes wh-questions like (21) below to be similar to the Irish cases illustrated in (20). That is, the questions in (21) are said to consist of a clefted phrase, i.e., the wh-phrase miin ‘who’ in (21a) and the wh-phrase ëeh ‘what’ in (21b), followed by the cleft clause.

(21) a. miin illi Mona šaafit-uh? (Egyptian)
    who that Mona saw.3fs-him
    ‘Who is it that Mona saw’

b. ëeh illi Mona ?arit-uh
    what that Mona read-it
    ‘What did Mona read?’

The apparent similarities between the Irish cases and the Egyptian ones led Cheng to treat the latter as instances of reduced clefts of the type exemplified in (20). Granting this analogy, Cheng assigns the structure in (22) for a cleft sentence like (21a).
Under this structure, the cleft clause, i.e., the CP headed by complementiser *illi* ‘that’, is acting as a predicate and the clefted wh-DP *miin* ‘who’ is functionally acting as a syntactic subject. The structure in (22) also incorporates the additional assumption that Spec,*illi* is occupied by a null operator that binds the resumptive pronoun inside clause.

The cleft analysis presented above, however, gives rise to a number of problems. I will briefly outline some of these problems, while a more detailed discussion is taken up in the following section. The first problem concerns the structure for wh-clefts above in that it cannot be said to correctly represent the structure of clefts in Arabic. The reason being is that Arabic clefts project Tense, and are, therefore, TP projections. As will be shown in later sections, the structure in (22) as a whole is, as a matter of fact, more appropriate as a structure for the relative clause that forms part of a wh-cleft question. In addition, there is no functional head in this structure to mediate the (syntactic) subject-predicate relationship between the wh-subject and the CP predicate.

Secondly, Cheng’s attempt to reduce Egyptian wh-clefts to Irish is not insightful since Irish indefinites are cleftable while they are not so in Egyptian.

---

201
Similarly, the analysis is also challenged by the other restriction which prohibits the clefting of PPs and AdvPs in Arabic (wh-) cLEFTs. Such restriction remains as an unresolved issue in Cheng’s work.

Nevertheless, the fact remains that Cheng’s analysis is not meant to tackle all the issues related to wh-cLEFTs in as much as it is an analysis addressing optionality in wh-in-situ languages. The next section presents another type of analysis, i.e., the equative analysis, to see whether it provides a better explanation for the facts from wh-cLEFTs observed at the outset. The analysis is advanced in Shlonsky 2002 for Palestinian and is extended to Lebanese in Aoun, Benmamoun and Choueiri 2010.

4.3.2 Shlonsky’s 2002 Equative Analysis

The last two sections have shown that the type of Arabic sentence with a [DP PRON RC] format displays an interesting array of properties while suggesting a resemblance with cleft constructions in other languages such as English and Irish. To this effect, Cheng’s 1991 study remains as an attempt to extend the reduced cleft analysis (McCloskey 1979) to Egyptian, though one that has left many questions unsolved, especially, the exact nature of the relationship between the wh-expression and relative clause, the status of the relative clause and the restriction on indefinites and non-argument phrases. Some of these issues are addressed in Shlonsky’s 2002 analysis for Palestinian. This section presents Shlonsky’s analysis and shows that although it seems attractive, it still requires some refinement.

Shlonsky 2002 proposes that wh-cLEFTs in Palestinian (Shlonsky’s class II interrogatives) are equative/identificational sentences in which the (Wh-) DP is base-generated as a clausal subject and the clause headed by illi ‘that’ as a predicate. Shlonsky’s analysis is based on an analogy drawn between the two
types of construction. To illustrate consider the following cases.

(23)  
\begin{align*}
\text{a. Faatme} & \text{illi ra}'ist l-baladiye.} \\
& \text{(Palestinian)} \\
& \text{Faatma that head the-municipality} \\
& \text{‘Fatme is (the one who is) the mayor’}
\end{align*}

\begin{align*}
\text{b. miin illi l-?asad } \text{?akal-ha mbaarih?} \\
& \text{who that the-lion ate-her yesterday} \\
& \text{‘Who did the lion eat yesterday?’}
\end{align*}

According to Shlonsky, (23a) and (23b) above have the format of equative sentences in that they involve two DPs: a DP subject, namely, \textit{Faatme} in (23a) and the wh-argument \textit{miin} ‘who’ in (23b), and a DP predicate, namely, the clause marked by \textit{illi}. For Shlonsky, the difference between wh-clefts like (23b) and sentences like (23a) is that the subject in the latter is not a wh-expression. Based on this analogy, the structure Shlonsky assigns for Palestinian wh-clefts such as (23b), prior to wh-movement, is shown in (24) below (Shlonsky’s 33).

\begin{itemize}
\item All the Palestinian examples are taken from Shlonsky 2002, without any modification on the English translation, unless otherwise stated.
\end{itemize}
The structure above is built on a number of assumptions. The first assumption is that there are two predication relations mediating the relationship between the wh-expression and the resumptive element inside the lower clause. One relation is between the two DPs forming the equative sentence: the lower DP predicate that is complement of I and the subject wh-phrase that is base-generated in Spec,IP. The other predication relation is between the null pronominal head of the free relative acting as a subject and the relative CP headed by illi acting as a predicate. The second relation is based on the assumption that illi is a [+Predicational] complementiser (see Rizzi 1990) that, therefore, transmits its feature to the clause it heads. This entails, according to Shlonsky, that illi-CP is a predicate and whose subject is the (null) head of the relative clause.

The second assumption that the structure in (24) incorporates is that there is a null operator in the specifier position of illi-CP which binds the resumptive pronoun inside the clause. In this respect, Shlonsky argues that Arabic illi carries φ-features that determine its specifier as an A-position. This explains the fact that wh-clefts do not admit PPs and adverbial phrases. Under Shlonsky’s analysis, this restriction may be understood as follows: only A-elements that are able to enter into an agreement relationship may occupy Spec,illi; since PPs and adverbial phrases do not bear φ-features, they cannot appear in Spec,illi. Consequently, PPs and adverbials cannot occur in this type of question\(^{11}\).

Finally, because Palestinian is a wh-movement language, the wh-phrase is said to undergo movement from Spec,IP to Spec,CP, yielding the structure in (25).

\(^{11}\)Notice that this view does not eliminate the occurrence of indefinite DPs. This issue will be discussed shortly.
The analysis just outlined is adopted in Aoun, Benmamoun and Choueiri 2010 for Lebanese wh-clefts. Based on the structure in (25), Aoun, Benmamoun and Choueiri further point out that wh-clefts involve a three-way agreement relationship between the null operator in Spec,CP and the resumptive element inside IP and the null pronominal pro in Spec,DP. The latter is said to be overtly realised as a strong pronoun in some cases as shown in (26a-b) below (these examples are taken from Aoun, Benmamoun and Choueiri 2010: 152).

(26) a. miin hi illi l-ʔasad ʔakal-ha mbaarih?
    who he that the-lion ate-her yesterday
    ‘Who did the lion eat yesterday?’

    (Palestinian)

b. miin hiyye illi l-ʔasad ʔaDD-a mbeerih?
    who she that the-lion ate-her yesterday
    ‘Who did the lion eat yesterday?’

    (Lebanese)
Although Shlonsky’s analysis has much to commend it, it has a number of problems. In the remainder of this section, I will highlight four major problems to do with the structure of wh-clefts, the predicational status of the DPs in such constructions, the indefiniteness restriction, and the use of subject pronouns in such constructions.

**Problems with Shlonsky’s analysis**

One of the major problems in Shlonsky’s analysis concerns the structure assigned for wh-clefts. The structure is repeated in (27), before movement of the wh-phrase to Spec,CP.

(27)

```
(27)
  IP
   / \ 
  DP  I' 
   / \ 
  I   DP
   / 
 WH-expression I
   / 
 DP
   / 
 Pro
   / 
 Op
   / 
 illi
```

As can be seen, the structure above involves two main DPs that are said to form the equative/identificational sentence, i.e., the lower DP predicate complement of I and the subject wh-phrase that is base-generated in Spec,IP. In addition to this predication, there is a second subject-predicate relationship between the head of the relative clause and *illi*-CP. Syntactically, there is a contradiction between these two relations and how they are represented in Shlonsky’s structure.
More specifically, notice that the higher relation is mediated through the functional head I, and involves a subject that is external to the lower DP predicate. Meanwhile, the lower predication is neither mediated through any functional heads, nor does it involve an external subject.

Furthermore, Shlonsky claims that the structure for wh-clefts in (27) suggests a resemblance with nominal sentences (i.e., sentences with non-verbal predicates). The following examples in (28a-c) are given as an illustration (Shlonsky’s 34a-c).

(28)  

a. Faatme raʾīṣ l-baladiye.  
Faatma head the-municipality
‘Fatme is the mayor’

b. Mhemmed Tawiil  
Mhemmed tall
‘Mhemmed is tall’

c. ḫuxu-y bi-t-taʾt  
brother-my in-the-bed
‘My brother is in bed’

There are two remarks that can be made with regards to the above examples. First, recall that there are two types of copular construction: the simple subject-predicate type which has the schematic structure [(Wh-) DP XP] as exemplified by (28a-c) and LD’ed equatives which have the format [DP PRON DP], as shown in (29) below. Recall also that each of these types of copular construction has been assigned a different structure\(^\text{12}\). For instance, the structure that has been assigned for LD’ed equatives is the one in (30), while the structure for simple subject-predicate clauses is shown in (31).

\(^\text{12}\)See Chapter 1 (section 1.5) for a full discussion on (types of) copular constructions in Arabic.
(29) a. ʕali hu l-mudarris  
Ali he the-teacher  
‘Ali is the teacher’

b. Ayman huu Saahib l-benaayeh  
Ayman he owner the-building  
‘Ayman is the owner of the building’

(30)

(31)

Given this analysis along with the structural difference between the two types of construction, Shlonsky’s attempt to assimilate wh-clefts to nominal sentences is at best ambiguous. The second point to be made in this respect is that cases which involve two lexical DPs such as (28a) in a subject-predicate relationship must be marked prosodically, otherwise sentential interpretation
does not obtain. Shlonsky’s analysis, however, fails to capture this fact about copular constructions in Arabic.

The third problem in Shlonsky’s analysis concerns the indefiniteness restriction. Shlonsky provides an explanation for why PPs and AdvPs cannot occur in this type of construction on the basis that Spec,illi is an A-position, hence, restricted to argumental DPs, i.e., elements that bear φ-features. This reasoning, however, does not explain the restriction on indefinite argumental DPs since these, too, are A-elements that bear φ-features. Early on, I have proposed along the lines of Chomsky 2001a, Uraigereka 2006 and Soltan 2007 that this restriction follows from the semantic effects of the EPP position that the left-dislocated DP occupies\(^{13}\).

The other major problem in Shlonsky analysis of wh-clefts and equative sentences concerns the nature and distribution of PRON in such constructions. Shlonsky 2002 points out that while the occurrence of *the pronominal copula* is obligatory in equative sentences in Arabic (32) as well as in Hebrew (33), it is optional in wh-clefts (34a, b)\(^{14}\).

(32) Mhemmed *(hu) Abu-Tariq
    Mhemmed he Abu-Tariq
    ‘Mhemmed is Abu-Tariq’

(33) Daniela hi ha-madrixa šel-i.
    Daniela she the-advisor of-me
    ‘Daniela is my advisor.’

(34) a. miin (hi) illi l-ʔasad ʔakal-ha mbaarih?
    who (she) that the-lion ate-her yesterday
    ‘Who did the lion eat yesterday?’

\(^{13}\)See Chapter 1, section 1.5.7
\(^{14}\)These examples are from Shlonsky (2002: 153-154).
b. miin (hi) illi ?aklat l-?asad mbaarih? (Palestinian)
    who (she) that ate the-lion yesterday
    ‘Who ate the lion yesterday?’

The obligatoriness of PRON in identificational sentences vs. its optionality in wh-clefts remain as a problematic issue for Shlonsky, who suggest that both facts fall under the same generalisation. However, as I have shown in the preceding chapters, optionality is only apparent. Namely, it is not the case that PRON is optional in cases like (34); rather, sentences involving PRON are structurally different from those without\(^{15}\).

Moreover, where PRON is obligatory, i.e., in equative constructions, Shlonsky assumes along the lines of Doron’s 1983 analysis of the pronominal copula in Hebrew that it is the the spell-out of subject agreement in I(nfl). There is abundant evidence, however, that PRON cannot be the phonetic realisation of subject-verb agreement (see also Eid 1992 and Edwards 2006). Furthermore, under the analysis proposed here, equatives involving PRON are analysed as left-dislocation structures in which PRON is treated as a resumptive subject pronoun whose antecedent is the left-dislocated DP.

Shlonsky’s analysis for Palestinian wh-clefts is also extended to Lebanese in Aoun, Benmamoun and Choueiri 2010. Putting aside the problems in Shlonsky’s analysis just outlined, Aoun, Benmamoun and Choueiri posit that the null pronominal in Spec,DP can sometimes be overtly realised as a strong pronoun. Examine the relevant cases repeated below for convenience (given in Aoun, Benmamoun and Choueiri 2010: 152).

(35) miin hi illi l-?asad ?akal-ha mbaarih?
    who he that the-lion ate-her yesterday
    ‘Who did the lion eat yesterday?’

\(^{15}\)This issue will be taken up further in the section that follows.
In the sentences above, it is not clear what the nature of this strong pronoun is. For the sake of argument, assuming with Aoun, Benmamoun and Choueiri that the pronoun is in Spec,DP, it must be the head of the relative clause. The structure for (35), for instance, may roughly be represented in (37).

If this is the case, the structure above faces an empirical problem. That is, the fact that an identical strong pronoun appears in wh-clefts cannot be captured in the analysis of Shlonsky and Aoun, Benmamoun and Choueiri. Consider the following examples from Jordanian.
The examples above show that there are two occurrences of the strong pronoun. The first of these forms is the resumptive subject pronoun that occurs in left-dislocation structures, as we have seen earlier. The second form is the head of the relative clause in Spec,DP that Aoun, Benmamoun and Choueiri 2010 posit.

To recap, this section has presented Shlonsky’s analysis for Palestinian wh-clefts. The core argument in Shlonsky’s analysis is that wh-clefts are equative or identificational sentences in which the wh-subject is base-generated in Spec,IP followed by a free relative DP predicate. However, Shlonsky’s analysis has been shown to be problematic in its approach to syntactic predication, optionality of PRON, the restriction on indefinites and, above all, the structure assigned for wh-clefts. In the following section, I provide an in depth analysis for (wh-) clefts in Arabic that captures the facts observed so far while doing away with the problems outlined here.

4.4 A unified analysis for Arabic (Wh-) Clefts

The aim of this section is to provide an alternative and more elaborate analysis for (wh-) clefts in Arabic. The analysis advances the view that Arabic clefts are a species of the equative copular construction. Recall that there are two types of equatives: basic subject-predicate equatives, i.e., PRON-less equatives, and LD’ed equatives of the form [DP [PRON DP]]. In a similar fashion, the present analysis treats clefts involving PRON as instances of left-dislocation much in the same way as LD’ed equatives. Whereas, clefts without PRON are treated
as instances PRON-less equatives of the form [DP DP].

The subsections below are organised as follows. Subsection (4.1) starts with the cleft clause and shows that the properties associated with it are typical of definite relative clauses. In subsection (4.2), I provide a derivation for the cleft clause based on the DP-analysis of definite relatives in Arabic (Ouhalla 1999, 2004; Shlonsky 2002; Aoun and Li 2003 and Al-Momani 2010). This will be followed by several syntactic and semantic arguments as support for the view that wh-clefts are equative copular constructions (subsection (4.3)). The structure for (wh-) clefts will be presented afterwards along with a discussion on the consequences and implications of the proposed analysis.

4.4.1 The cleft clause is a definite relative clause

This section advances the view that the cleft clause appearing in (wh-) clefts is a definite relative clause. In particular, I will show that the cleft clause is a headless relative noun phrase, i.e., a free relative DP. I present, first, types of relative clauses in Arabic and highlight some of the main properties associated with them. Then, I show that such properties are also characteristic of the cleft clause in cleft sentences and their wh-counterparts, especially, the presence vs. the absence of both the complementiser illi ‘that’ and the resumptive pronoun.

There are two types of relative clause in Arabic: definite relatives and indefinite relatives (Farghal 1986, Shlonsky 1992, Aoun and Li 2003, Al-Momani 2010 and Aoun, Bennamoun and Choueiri 2010). Definite relatives modify a definite noun and require the presence of the relative marker illi ‘that’. Indefinite relatives, on the other hand, modify an indefinite noun without the relativiser.

---

16 The analysis in this section is based on data from Jordanian Arabic which is similar, if not identical, to other Arabic dialects (see Aoun and Li 2003 for Lebanese, Shlonsky 2002 for Palestinian, Ouhalla 1999 for Moroccan. Examples (39)-(42) are taken from Al-Momani’s 2010 study of relative clauses in Jordanian Arabic.

17 Recall that Arabic uses another type of free relative clause which involves the relativiser maa ‘ever’ that always co-occurs with a wh-expression preceding it (see Chapter 1 for an analysis of this type of free relative). I will not have much to say about this type of relative clause here since it does not relate to the current discussion.
Consider the examples in (39) and (40), in which the relativised DP corresponds to the subject position inside the relative clause, and the ones in (40) and (41), in which the relativised DP corresponds to the object position inside the clause.

(39) a. ŝuft l-walad illi gara l-ktab (Jordanian)
saw.1ms the-boy that read.3ms the-book
‘I saw the boy that read the book.’

b. *šuft l-walad gara l-ktab
saw.1ms the-boy read.3ms the-book
‘I saw the boy that read the book.’

(40) a. ŝuft walad gara l-ktab
saw.1ms a-boy read.3ms the-book
‘I saw a boy that read the book.’

b. *šuft walad illi gara l-ktab
saw.1ms a-boy that read.3ms the-book
‘I saw a boy that read the book.’

(41) a. gareit l-ktab illi štara-h T-Talib (Jordanian)
read.1s the-book that bought.3ms-it the-student
‘I read the book that the student bought’

b. *gareit l-ktab štara-h T-Talib
read.1s the-book bought.3ms-it the-student
‘I read the book that the student bought’

(42) a. gareit ktab štara-h T-Talib
read.1s a-book bought.3ms-it the-student
‘I read a book that the student bought’

b. *gareit ktab illi štara-h T-Talib
read.1s book that bought.3ms-it the-student
‘I read a book that the student bought’
In both (39a) and (41a), the definite DPs *l-walad* ‘the boy’ and *l-ktaab* ‘the book’ are followed by a definite relative headed by relativiser *illi* and the sentences are well-formed. By contrast, absence of the relativiser *illi* from the clause modifying the definite DPs in (39b) and (41b) renders such cases ungrammatical. On the other hand, the indefinite DPs *walad* ‘a boy’ in (40a) and *ktaab* ‘a book’ in (42a) are modified by an indefinite relative which does not require the presence of the relativiser *illi*. As a matter of fact, the presence of relativiser *illi* with indefinite DPs yields ungrammaticality as shown in (40b) and (42b). The point made from these examples is that the presence of the relativiser *illi* is characteristic of definite relative clauses in Arabic. Given that *illi* appears only in definite relatives and, hence, in the cleft clause, the discussion will be limited to this class of relatives.

Another property characteristic of definite relatives is the obligatory presence of a resumptive pronoun in the relativised object position in object relatives, irrespective of whether the DP is a definite DP as in (43a-b) or an indefinite DP as in (44a-b)\(^{18}\). By contrast, a resumptive pronoun must be null in the relativised subject position when the relativised DP is an indefinite subject (45a-b). Meanwhile, a resumptive pronoun is optional in cases where the relativised DP subject is definite (46-a-b). Consider these cases as shown below.

\[(43)\]  
\[\begin{array}{ll}
a. & \text{gareit l-ktab illi stara-h T-Talib (Jordanian)} \\
& \text{read.Is the-book that bought.3ms-it the-student} \\
& \text{‘I read the book that the student bought’}
\end{array}\]

\[\begin{array}{ll}
b. & \text{*gareit l-ktab illi stara T-Talib} \\
& \text{read.Is the-book that bought.3ms the-student} \\
& \text{‘I read the book that the student bought’}
\end{array}\]

\(^{18}\)That is, RPs are obligatory in the dialects if the relativised DP is a direct object, an indirect object, as well as in prepositional relatives and genitive relatives. See Al-Momani 2010 for a more detailed account.
Examples (43b) and (44b) are ruled out on the grounds that a resumptive pronoun is absent in the VP-internal position of the relative clause. As regards subject relatives, there are two observations that may bear mentioning in (45) and (46). First, (45b) appears to be a subcase of the Highest Subject Restriction (McCloskey 1990) that bans the presence of a resumptive pronoun in the highest subject position inside the relative clause in cases where the relativised DP is indefinite. Secondly, in cases where the relativised DP is definite, a resumptive pronoun is optional, however, as shown in (46b). Thus, the data from Jordanian show that HSR is operative only when the antecedent of the pronominal element is indefinite.

A similar observation is made in Aoun, Choueiri and Hornstein 2001 for Lebanese.
Definite relatives can be further divided into two types: restrictive relatives with an overt lexical head and free relatives with a null head. Consider the following examples.

(47) a. l-muddarise najjahat T-Tollaab illi šafat-hum the-teacher(fs) passed.3fs the-students that saw.3fs-them
    b-l-mobaraah (Jordanian)
in-the-match
    ‘The teacher passed the students that she saw in the match’

b. l-muddarise najjahat pro illi šafat-hum b-l-mobaraah
    the-teacher(fs) passed.3fs pro that saw.3fs-them in-the-match
    ‘The teacher passed the ones that she saw in the match’

(48) a. l-muddarise najjahat T-Tollaab illi (hummu) šaafø
    the-teacher(fs) passed.3fs the-students that (they) saw.3mpl
    b-l-mobaraah the-match
    ‘The teacher passed the students that saw the match’

b. l-muddarise najjahat pro illi (hummu) šaafø
    the-teacher(fs) passed.3fs pro that (they) saw.3mpl
    l-mobaraah the-match
    ‘The teacher passed the ones that saw the match’

In (47a), the head of the relative clause is an overt DP, i.e., T-Tollaab ‘the students’, meanwhile the head of the relative clause in (47b) is null pro. Notice, however, that the features on the resumptive clitic attached to the verb match those on the head of the relative clause in case of object relatives, whether the head is a lexical DP (47a) or whether it is null (47b). In the subject relatives in (48), the verb bears agreement features that match those of the subject of the

contrast, Shlonsky argues that the HSR is operative in all cases in Palestinian. Upon re-examination, however, Palestinian, like Jordanian, Lebanese and SA, prohibits resumptive pronouns in the highest subject position only when the antecedent is indefinite.

217
relative clause, whether the subject is overt as in (48a) or whether it is covert as in (48b).

So far, we have been dealing with types of relative clauses and have seen that ʾilla clauses are definite relatives. In addition to the presence of ʾilla, definite relatives have been shown to involve an obligatorily overt resumptive pronoun in the relativised object position. Meanwhile, the resumptive pronoun is optional in the relativised subject position, regardless of whether the head of the relative clause is a lexical DP, i.e., a restrictive relative, or a null DP, i.e., a free relative.

Turning to Arabic clefts, the consistent presence of ʾilla in the cleft clause coupled with the resumption facts indicate that it is a definite relative clause. Although the two types of definite relatives, i.e., restrictive relatives with an overt lexical head, and free relatives with a null DP head, occur in both types of cleft construction, it is the latter that is more commonly used across the dialects and which I will be concerned with henceforth. The following paradigms pattern with those of definite relatives just observed. First, consider the sentences below where the cleft clause is an object relative.

(49) Clefts

a. Nada ʾilla l-bent ʾilla Majdi tjwaaz-ʾ(ḥa) (Jordanian)
   Nada she the-girl that Majdi married-her
   ‘It is Nada that Majdi married’
   ‘Nada is the girl that Majdi married’

b. Nada ʾilla illi Majdi tjwaaz-ʾ(ḥa)
   Nada she that Majdi married-her
   ‘It is Nada that Majdi married’
   ‘Nada is the one that Majdi married’
(50) Wh-Clefts

a. miin hii l-bent illi Majdi tjwaaz-*(ha)
who she the-girl that Majdi married-her
‘Who is it that Majdi married’

b. miin hii illi Majdi tjwaaz-*(ha)
who she that Majdi married-her
‘Who is the girl that Majdi married’

In (49a) and (50a), the cleft clause is identical to the restrictive relative in (43a) and (47a), in which an overt DP is followed by a relative clause modifying the DP l-bent ‘the-girl’. Because the cleft clause is an object relative, the presence of a resumptive pronoun inside the VP is obligatory, otherwise the sentences above are ruled out. Unlike the (a) examples, the head of the cleft clause in (49b) and (50b) is non-overt, it is, rather, a null DP, as in (47b). The only difference between the cleft clause in the (a) examples and the cleft clause in the (b) examples is that the head of the cleft clause is overt in the latter, but not in the former, much in the same way as in restrictive relatives and free relatives in (47a) and (47b), respectively. Apart from this difference, the cleft clause is the same, it involves the same complementiser, i.e., illi, and a resumptive pronoun in object position.

Secondly, where the relativised DP corresponds to a subject position inside the cleft clause, a resumptive pronoun is optional, as we have seen in the subject relatives cases in (46) and (48). To illustrate, consider the following examples.

(51) Clefts

a. Nada hii l-bint illi (hii) tjwaazat Majdi
Nada she the-girl that (she) married Majdi
‘It is Nada that married Majdi’

‘Nada is the girl that married Majdi’
b. Nada hii illi (hii) tjwaazat Majdi
Nada she that (she) married Majdi
‘It is Nada that married Majdi’

‘Nada is the one who married Majdi’

\[(52)\] Wh-Clefts

a. miin hii l-bint illi (hii) tjwaazat Majdi
who she the-girl that (she) married Majdi
‘Who is the girl that married Majdi’

‘Who is the girl that married Majdi’

b. miin hii illi (hii) tjwaazat Majdi
who she that (she) married Majdi
‘Who is it that married Majdi’

‘Who is one that married Majdi’

Like definite relatives, the cleft clause in (51a) and (52a) modifies an overt DP head and involves an optional resumptive pronoun in the (highest) subject position. On the other hand, the cleft clause in (51b) and (52b) modifies a null head and also involves a non-overt resumptive pronoun in the subject position\(^{20}\).

To sum up, this section has presented two types of relative clause in Arabic, namely, definite relatives and indefinite relatives. It was shown that the properties associated with definite relatives are identical to properties of cleft clauses, especially, the presence of the complementiser *illi*, the obligatory presence of a resumptive pronoun in object relatives vs. its optional presence in definite subject relatives and (non-/-) overtness of the antecedent DP. The parallelism between clefts clauses and definite relatives presented in this section constitutes

\(^{20}\)Though the resumptive pronoun is optional in the subject position inside the relative clause, it is rarely overt due to the presence of another identical form in the structure. There is nothing in the syntax, however, that prohibits the realisation of the subject pronoun in as much as it is a stylistic issue which disfavours the occurrence of two identical forms adjacently. As concerns the HSR of McCloskey (1990), it is only operative where the antecedent subject is indefinite, as I have shown earlier.
strong evidence that the cleft clause is indeed a definite relative clause. Given this parallelism, I turn, next, to the structure of the cleft clause.

### 4.4.2 The structure of the cleft clause

In the previous section, I presented evidence that cleft clauses are free relative clauses. I have also shown that free relatives differ from restrictive relatives in Arabic in that their head is not lexical, but, rather, a null *pro* (see also Groos and van Riemsdijk 1981, Grosu 1989 and Suñer 1984). The aim of this section is to provide a structure for cleft clauses. I assume, along with most studies on Arabic relative clauses (Ouhalla 1999, 2004; Shlonsky 2002; Aoun and Li 2003; Aoun, Benmamoun and Choueiri 2010 and Al-Momani 2010), that free relatives are contained in a DP structure. The structure for object relatives can be roughly represented in (53) and the one for subject relatives in (54)\(^{21}\).

\[\text{(53)}\]

\[
\begin{array}{c}
\text{DP} \\
\downarrow \\
\text{DP} \\
\downarrow \\
\text{pro} \\
\downarrow \\
\text{DP} \\
\downarrow \\
\text{C} \\
\downarrow \\
\text{TP} \\
\downarrow \\
\text{i}\text{lli} \\
\downarrow \\
\text{Majdi } tjawwaz-ha, \\
\text{that} \\
\downarrow \\
\text{Majdi married-RP}
\end{array}
\]

\(^{21}\)Another plausible approach for the derivation of relatives is advanced in Aoun and Li 2003 based on Kayne 1994 and Bianchi 1999.
The structures for object relatives and subject relatives incorporate the following assumptions: *illi* is the phonetic realisation of a definiteness feature on C (Al-Momani 2010). Secondly, Spec,*illi* is filled by a null operator (Shlonsky 2002, Aoun and Li 2003 and Al-Momani 2010) which binds the resumptive pronoun inside the clause. The third assumption concerns the base-generation analysis of resumptive pronouns (Shlonsky 2002, Soltan 2007 and Al-Momani 2010).

In the case of object relatives, an overt resumptive pronoun is base-generated in VP-internal position. In the case of subject relatives, the resumptive pronoun is not spelled out unless required by interface conditions, such as emphasis and contrastive focus (Soltan 2007). These assumptions are dealt with in turn below.

A crucial property of *illi* is that its presence or absence is dependent on whether the relative clause modifies a definite or an indefinite DP. This particular property of *illi* recalls the presence of the definite determiner (*a*)-'the’ in definite noun phrases and its absence in indefinite noun phrases, as can be illustrated below.

(55)  a. l-baab l-?asfar  
      the-door the-yellow  
      ‘The yellow door.’
b. baab ʿasfar
   door yellow
   ‘A yellow door’

This property of illī coupled with the morphological similarity between illī and the definite determiner l- ‘the’ in Arabic (Ouhalla 1999 and Aoun, Bennamoun and Choueiri 2010) suggest that this complementiser is the phonetic realisation of a number of features amongst which is a definiteness feature. Indeed, Al-Momani 2010 treats illī as the spell-out of a definiteness feature on C of definite relatives. More precisely, C is said to carry an uninterpretable [u,DEF]. This feature is checked (and deleted) by a null relative operator with the interpretable feature [+DEF] that is merged in Spec,C. Consequently, C is spelled out as illī. By contrast, the definiteness feature is not needed in indefinite relatives, hence, the complementiser is not spelled out, i.e., it is null. To illustrate, consider the structure for object relatives roughly represented in (56) below.

(56) DP
    / \  
   /   \  
  DP  CP
     /   \  
    /     \  
   [Def] Op, C
     /   \  
    /     \  
   [Def] C      TP
     /     \  
    /      \  
   [φ]  (ya)lli  Majdi tjawwaz-ha,
         [u,Def]  Majdi married RP,
         [φ]

Notice that illī bears φ-features in (56) that match those on the relative operator and the RP. This is in line with Aoun and Li 2003 who posit that the complementiser illī in definite relatives must agree with the head in φ-features
and with Shlonsky 2002 where illsi is said to determine its specifier as an A-
specifier.

Moreover, the assumption that a null operator is merged in Spec, illsi is bol-
stered by the presence of the resumptive pronoun with which it is co-indexed
as a consequence to the binding relationship between the two. Recall that a
resumptive pronoun (RP in (56)) is obligatory in the relativised object posi-
tion. This strongly suggests that movement is not available from that position,

hence, an RP is base-generated in the object position. This RP has to be
co-indexed with an operator. It is the null operator in Spec, illsi that binds the
RP. Following Chomsky (1995: 227-228, 348), co-indexation takes place in the
numeration set thereby establishing the matching relation between the RP and
the null operator. The presence of the operator in Spec, illsi is, thus, crucial to
the interpretation in that it mediates the relation between the antecedent and
the RP, i.e., it binds the relevant category within TP and links the relative CP to
the preceding DP (Al-Momani 2010 and Aoun, Benmamoun and Choueiri 2010).

On the other hand, a resumptive pronoun is optional in the relativised sub-
ject position. This may be taken to indicate that movement is available. Ac-
cording to Shlonsky 2002 and Aoun, Benmamoun and Choueiri 2010, movement
to Spec, illsi is A-movement. Only the highest subject, i.e., the subject closest
to illsi, can move to this position, otherwise, movement from any other posi-
tion across the highest subject position will be blocked, a case of minimality.

Because movement is available, the spelling out of subject pronoun is not oblig-
atory, except for emphatic interpretations.

22 According to Al-Momani, the obligatoriness of the RP in the case of direct object rela-
tivization might be due to two reasons: the fact that (Accusative) Case in JA is left unmarked
and the absence of agreement between the noun head and the gap left because of the absence
of the RP.

23 The violation can be ascribed to Rizzi’s 1990 Relativized Minimality or Chomsky’s 1995
Shortest Move.

24 Recall that Arabic subject pronouns are usually null though they get lexicalised for em-
phatic and contrastive purposes. See also Eid 1992, Plunkett 1993, Fassi-Fehri 1993, Ouhalla
The DP-analysis of *illi* relative receives empirical support in Arabic\(^{25}\). Evidence to this effect comes from the fact that *illi*-clauses appear in the same distributional contexts as DPs in Arabic\(^{26}\). For instance, they can appear in canonical subject positions as well as in object positions in verbal predications, (57a) and (57b), respectively; meanwhile they may occur as subjects of simple copular sentences of the type illustrated in (58a-c) below.

\[(57)\]
\[
\begin{align*}
\text{(57a) } & \text{illi that talk-\text{-}\text{-}lak \text{\text{-}ann-o rawwah} } \text{Jordanian} \\
& \text{that talk-I-you about-him left.3ms} \\
& \text{‘The one I talked to you about left’} \\
\text{(57b) } & \text{šofna illi kasar l-šobbak} \\
& \text{saw.1pl that broke.3ms the-window} \\
& \text{‘We saw the one who broke the window’}
\end{align*}
\]

\[(58)\]
\[
\begin{align*}
\text{(58a) } & \text{illi that tjawwaz-ha Majdi muhandisah.} \\
& \text{married-her Majdi an-engineer.fs} \\
& \text{‘The one who Majdi married is an engineer’} \\
\text{(58b) } & \text{illi that tjawwaz-ha Majdi Taweele} \\
& \text{married-her Majdi tall.fs} \\
& \text{‘The one who Majdi married is tall’} \\
\text{(58c) } & \text{illi that tjawwaz-ha Majdi b-l-maktab} \\
& \text{married-her Majdi in-the-office} \\
& \text{‘The one who Majdi married is in the office’}
\end{align*}
\]

Furthermore, the cleft clause may be equated with another DP on a par with two-DP equatives, thereby forming subject-predicate sentences. Crucially, when the cleft clause co-occurs with another lexical DP, a prosodic break between the two is necessary for sentential interpretation to obtain. To illustrate, consider the following examples.

\(^{25}\)See Ouhalla’s 2004 analysis of Semitic relatives and the distinction made between DP-relatives, including *illi* relatives, and CP-relatives.

\(^{26}\)See also Chapter 1.
A similar observation is made in Cowell 1964 for Syrian Arabic. Cowell analyses the sentences in (61) as subject-predicate constructions on a par with two-DP equatives. He further suggests that the order of the two DPs can be reversed as long as the intonation is marked appropriately.

Recall that in the case where one of the two DPs forming the equative sentence is pronominal, the sentence would not require the prosodic break associated with lexical DPs. The pattern that is more commonly found is one where the pronoun is the subject and the free relative DP is the predicate. This is shown in (62) below.
To recap, this section has advanced an analysis whereby the structure assigned to the cleft clause is the same as that of free relatives in Arabic, namely, a DP structure. The analysis is based on properties of the complementiser *illi* in Arabic, its relationship with the null operator in its Spec position and with the resumptive pronoun inside the clause. Granting the DP-analysis for the cleft clause, the following section provides an analysis for the overall structure of (wh-) clefts.

### 4.5 (Wh-) clefts are equative copular (Wh-) constructions

In the previous section, I have shown that the cleft clause has the same properties as definite relatives in Arabic. Both have also been shown to appear in the same contexts where DPs occur, hence, the structure assigned for the cleft clause is a DP structure. This section puts forward an analysis for Arabic clefts based on the analysis of copular constructions presented in the first chapter. First, I present evidence from the syntax and semantics of such constructions which suggests a parallelism with cleft constructions. The structure and derivation of cleft constructions in the Arabic dialects presented here will follow afterwards.

#### 4.5.1 Evidence from the syntax and semantics

The data presented so far show that Arabic clefts may have either the, more common, format [DP PRON *illi*-RC] or the, less common, format [DP *illi*-RC]/[*illi*-RC DP]. Putting aside the latter for the moment, a consequence of
the DP analysis of *illi*-clauses advanced in the previous section is that Arabic clefts have the format [DP PRON DP]. This is precisely the format for LD’ed (left-dislocated) equatives in Arabic discussed earlier in this study. This entails that cleft sentences must be treated on a par with LD’ed equative constructions\(^{27}\). Such a treatment is not only syntactically, but also semantically, supported.

As far far as the syntax is concerned, there are several arguments that can be used as evidence for treating clefts as a species of equative copular construction, more precisely, as LD’ed equatives. Empirically, the fact that the cleft sentences contain two DPs separated by a subject pronoun can only be explained on the grounds that they are LD’ed equative constructions. Compare, for instance, the LD’ed equative sentences in (63) with the cleft sentences in (64).

(63)  a. Nada hii l-muddarisch
     Nada she the-teacher
     ‘Nada is the teacher’

     b. Nali hu l-mudarris
     Ali he the-teacher
     ‘Ali is the teacher’

(64)  a. Nada hii illi Majdi tjwaaz-ha
     Nada she that Majdi married-her
     ‘It is Nada that Majdi married’

     b. Nada hii illi tjwaazat Majdi
     Nada she that married Majdi
     ‘It is Nada that married Majdi’

As might be recalled, the set of properties that characterise LD’ed equatives such as (63) are as follows: the absence of a verbal copula, the definiteness of the two DPs and the presence of a subject pronoun i.e., PRON. Granting that *illi*-clauses are definite DPs, the properties displayed by the cleft sentences in

\(^{27}\)For a different view see Ouhalla 1999 analysis of Moroccan clefts.
(64) are the same as those in (63). As such, there is a clear sense in which the sentences in (64) must be treated on a par with those in (63). They are left-dislocation structures that derive from a basic subject-predicate clause, much in the same way as LD'ed equatives were said to derive from the basic two-DP equative constructions. The paradigms below illustrate the case at hand\(^{28}\).

\[(65)\]

a. Nada, illi Majdi tjwaaz-ha
   Nada that Majdi married-her
   ‘It is Nada that Majdi married’

b. Nada, illi tjwaazat Majdi
   Nada that married Majdi
   ‘It is Nada that married Majdi’

(66) a. Nada, l-muddariseh
   Nada the-teacher
   ‘Nada is the teacher’

b. Ya\‘li, l-mudarris
   Ali the-teacher
   ‘Ali is the teacher’

As a strategy, left-dislocation in cases such as (63) and (64) above identifies the predicate and eliminates possible ambiguity arising in contexts where the DPs forming the predication are lexical such as (65) and (66). Cowell 1964 makes a similar observation with regards to Syrian Arabic. For instance, the sentences in (67) are said to be the result of extraposing the subject DP in (68)\(^{29}\).

\[(67)\]

a. l-\textasciitilde{aa}Di huwwe yalli byehkom
   the-judge he that decide.3ms
   ‘The judge is the one who makes the decision’

b. yalli byehkom huwwe l-\textasciitilde{aa}Di
   that decide.3ms he the-judge
   ‘The one who makes the decision is the judge’

\(^{28}\)Recall that indefinite DPs cannot be clefted, a fact that was also observed in equatives and LD’ed equatives.

\(^{29}\)Contra Cowell 1964, the current analysis base-generates the left-dislocated DP in its surface position, rather than arriving via movement (see Chapter 1).
According to Cowell, extraposition in this case of equational predication, is commonly used not so much to emphasise the extraposed subject, but simply to identify the predicate as such, as we have seen in the case of LD’ed equatives in the first chapter of this study. To clarify further, the predication in (68a), for example, might in some circumstances be confused with the modified, i.e., relative, noun phrase *the judge who makes the decision*. The predication, thus, tends to be replaced by a topical sentence like that in (67a) when no special emphasis is intended.

The second type of argument supporting the LD’ed equative analysis for clefts lies in their interpretive properties, i.e., presupposition. More specifically, the type of presupposition exhibited in LD’ed equatives and clefts can be characterised as exhaustive (see Kiss 1998). The cleft sentence in (64a), for instance, presuppose that *Nada* is the only contextually relevant person that *Majdi* married. Likewise, the LD’ed equative sentence in (63a) presupposes that *Nada* is the only contextually relevant individual of which the property denoted by the internal predicational sentence *hii l-muddariseh* ‘she is the teacher’ holds.

The view that the semantics of clefts is similar to the semantics of (LD’ed) equatives is not particular to Arabic clefts only. English clefts, too, are said to pattern interpretively with specification sentences (Reeve 2010). Consider,

Accordingly, clefts like (69a) are taken to be a species of copular sentences (69b) where the post-copular DP, i.e., *John* answers a question represented by the pre-copular complex DP.
for instance, the cleft sentence in (69a) and it specificational counterpart in (69b) (Reeve’s (17a and b)).

(69) a. It was John that Mary saw.
    b. The one that Mary saw was John

Like their Arabic counterparts, both the sentences in (69) involve exhaustive presupposition where the DP John is presupposed to be the only contextually relevant person that Mary saw. Such a view is advanced in Higgins 1973, Percus 1997 and Hedberg 2000. In fact, these authors argue that not only do English clefts pattern semantically with specificational copular sentences, but they also pattern syntactically (contra Reeve 2010). That is, the sentences in (69) are said to be underlyingly syntactically parallel. The same argument holds of Arabic, consider the specificational sentences below that are counterparts to the cleft sentences in (64).

(70) a. illi Majdi tjwaaz-ha hii Nada. (Jordanian)
    that Majdi married-her she Nada
    ‘The one that Majdi married is Nada’
    b. illi tjwaazat Majdi hii Nada
    that married.3fs Majdi she Nada
    ‘The one that married Majdi is Nada’

In (70), while the pre-PRON constituent represents a question, the post-PRON DP specifies the answer to that question. In this respect, the sentences in (70) constitute additional evidence for the view that clefts are equative constructions. Recall that (LD’ed) equatives have a property whereby the two DPs can be reversed in order with no apparent effect on meaning (Ouhalla 1999 and Edwards 2006). This is precisely the case for the clefts in ((64) and (70)) in that the order of the two DPs is reversed with the cleft clause appearing clause-initially followed the clefted phrase with no change in the propositional content\(^\text{31}\).

\(^{31}\)Such ordering gives rise to what is known in the literature as *pseudo-clefts* (Schachter 1973 and Heycock and Kroch 1996) as in the English sentences: *What this country needs is a five-cent cigar* and *A five-cent cigar is what this country needs.*
4.5.2 The structure of (wh-) clefts

If this analysis is on the right track, clefts must be assigned a structure similar to that of LD’ed equatives. The structure for Arabic LD’ed equatives that has been advanced in this study is the one based on the PredP framework (Bowers 1993, 2001; Chomsky 2000 and Adger and Ramchand 2003) in which PredP is the predicational shell where thematic roles are licensed. The structure for LD’ed equative sentences like (71a) is shown in (71b).

(71) a. ʕali hu(wa) l-mudarris
    Ali he   the-teacher
    ‘Ali, he is the teacher’

b. TP
   /|
  DP T'
  |
ʕali T PredP
   /|
Ali DP Pred'
  |
hu Pred DP
  |
he l-mudarris
   the-teacher

Under this structure, the LD’ed DP is base generated in Spec,TP, an A’-position (Plunkett 1993, Soltan 2007 and Aoun, Benmamoun and Choueiri 2010). The structure also involves a complete predication whose subject is the resumptive strong pronoun, i.e., PRON, which is also base-generated in Spec,PredP. The relationship between LD’ed DP in Spec,TP and the subject pronoun in Spec,PredP is a coreference relationship. Finally, since equatives considered here are present-tense (nominal) sentences, T has, in addition to the (nominal) feature [+D], an abstract tense [+Present] (Benmamoun 2000 and Aoun, Benmamoun and Choueiri 2010).
Concerning the structure of clefts, it must be the same as the structure of LD’ed equatives except that the second DP is free relative DP. Compare, for instance, the structure of clefts roughly represented in (72) below with that of LD’ed equatives shown in (71b).

(72) The structure of clefts

The difference between the structure in (71b) and the one in (72) lies in the nature of the DP predicate. In LD’ed equatives, the DP is simple; whereas in clefts, the DP is more complex in that it contains a relative clause modifying it.
Turning to Arabic wh-clefts, we are now in a position to provide an account that can straightforwardly capture all the relevant facts observed so far. Staying with wh-clefts of the form [(Wh-) DP [PRON DP]], the analysis just outlined for clefts entails that the wh-clefts must be treated as wh-counterparts of cleft sentences, which are, in turn, analysed as instances of LD’ed equatives. Structurally, then, wh-clefts would have the same structure as cleft sentences except that the subject in wh-clefts is a Wh-DP. Compare the initial structure for the cleft sentence in (72) with its wh-counterpart in (73).

(73) a. miin hii illi Majdi tjawwaz-ha?
   who she that Majdi married-her
   ‘Who is it that Majdi married’

   b. 
      \[ TP \\
       \quad \downarrow \\
       \quad \text{DP} \quad T' \\
       \quad \downarrow \\
       \quad \text{miin} \\
       \quad \text{who} \\
       \quad \downarrow \\
       \quad [+D] \quad \text{PredP} \\
       \quad \downarrow \\
       \quad \text{hii} \\
       \quad \text{Present} \\
       \quad \downarrow \\
       \quad \text{Pred} \\
       \quad \downarrow \\
       \quad \text{DP} \\
       \quad \downarrow \\
       \quad \text{pro} \\
       \quad \text{CP} \\
       \quad \downarrow \\
       \quad \text{OP} \quad C' \\
       \quad \downarrow \\
       \quad \text{ill} \\
       \quad \text{that} \\
       \quad \text{Majdi tjawwaz-ha} \\
       \quad \text{Majdi married} \]

In this structure, the wh-DP is base-generated in Spec,TP which is, as argued here, an A’-position. This is the case for LD’ed DPs in cleft sentences and equa-
tives. The wh-phrase in Spec,TP binds the resumptive pronoun in Spec,PredP. The latter is also base-generated in its surface position as subject for the lower DP predicate, i.e., the free relative DP. Recall that the free relative also involves a binding relationship between the operator in Spec,illi and the resumptive pronoun in the argument position inside the clause. The equative nature of such constructions is reflected in the shared agreement between all the DPs in the structure.

Before discussing the derivation of wh-clefts in various Arabic dialects, there are two related issues that I would like to address at this stage. The first issue is that there is a crucial difference between clefts and wh-clefts that mirrors the difference between equatives involving lexical DPs and their wh-counterparts. To be more precise, I have shown that equatives involving two lexical DPs are likely to be ambiguous. Because it involves insertions of a pronoun, left-dislocation eliminates such ambiguity. This situation does not arise in the wh-counterparts of equatives, although the two DPs are also lexical. The reason why ambiguity does not arise is because one of the DPs is a Wh-DP. Consequently, the appearance of a pronoun is less common in equative copular wh-constructions. The same reasoning extends to clefts and wh-clefts with the consequence that wh-clefts of the form [Wh-DP [PRON DP]] are less common than those wh-clefts that have the form [Wh-DP DP]. Because ambiguity is not at play, Wh-clefts involving PRON differ from cleft sentences in that they have an emphatic interpretation associated with the presence of PRON.

The second related issue concerns the long-standing assumption that wh-clefts may optionally involve PRON, i.e., the so-called copula pronoun (Eid 1983, 1991, 1992; Doron 1983, 1986; Ouhalla 1999; Shlonsky 2002; Al-Horais 2006 and Aoun, Benmamoun and Choueiri 2010). This assumption is made in Shlonsky 2002 for Palestinian wh-clefts (74a); in Soltan 2009 for Egyptian wh-clefts (74b); and in Aoun, Benmamoun and Choueiri 2010 for Lebanese wh-clefts.
(74c).

(74)  

a. miin (hu) ỉlli hall l-muşkele? (Palestinian)
who (he) that solved.3ms the-problem
‘Who is it that solved the problem?’

b. miin (huwwa) ỉlli Darab boš (Egyptian)
who (he) that hit.3ms Bush
‘Who is it that hit Bush?’

c. miin (huwwe) ỉlli ɿaft-o b-l-maTŶam? (Lebanese)
who (h)e that saw.2sm-him in-the-restaurant
‘Who is it that you saw in the restaurant?’

Under the analysis developed in this study, the format [(Wh-) DP [PRON DP]] is a derived version of the more basic format [(Wh-) DP DP]. Whereas the latter is a subject-predicate clause, the former is analysed as a left-dislocated structure that involves base-generation of both the initial DP, in Spec,TP, and the subject pronoun, in Spec,PredP. Within this analysis, then, the so-called *copula pronoun* is in fact a resumptive subject pronoun bound by its antecedent DP in the left-periphery.

On this base, I would like to argue that optionality of PRON in cases like (74a-c) is only apparent. More precisely, it is not the case that PRON is optional in wh-clefts; rather, the case is that wh-clefts involving PRON are different from wh-clefts without PRON. Whereas the latter are instances of LD’ed equatives, the former are basic subject-predicate equatives. The two, therefore, have different structures. The structure for wh-clefts involving PRON is the one shown in (73b). As far as the structure of wh-clefts without PRON is concerned, it is similar to the structure that has been assigned for PRON-less equatives earlier on in this study. Consider, for instance, the structural representation of the Lebanese case (74c) as reproduced in (75) below.
As it stands, the proposed analysis has many advantages over the one advanced in Shlonsky 2002 and Cheng 1991. For instance, the current analysis distinguishes between two types of (wh-) cleft; a fact that both alternative analyses fail to capture. One type is a basic subject-predicate construction and has the same structure as PRON-less equatives. The second type is a left-dislocation construction and has the structure of LD'ed equatives. Moreover, the current analysis accounts in a systematic way for the distribution of PRON, not only in wh-clefs, but also in cleft sentences in Arabic. Meanwhile, the analysis avoids the problems associated with the assumption that PRON is the realisation of agreement features in Infl, and takes PRON to be what it is, i.e., as a subject pronoun, thereby providing a more plausible and natural account for its occurrence in such constructions.
Moreover, the structure assigned for wh-clefts involving PRON contains two predication relationships that are correctly represented, syntactically and semantically. The first relationship is the internal relationship, i.e., between the free relative DP as a predicate and the DP in Spec,PredP as (an external) subject, which is mediated through the functional head Pred. This predication as a whole is itself predicated of the (external) subject located in Spec,TP that is mediated through the functional head T.

4.5.3 Consequences

The proposed analysis has several consequences. We have seen already how the analysis captures all the facts related to PRON, be it in (wh-) clefts or in (equative) copular constructions. Concerning the restriction of clefts and wh-clefts on indefinite argumental DPs, it can be explained on the grounds that Spec,TP is inherently a position that can only host definite DPs and free relatives are always definite.

Furthermore, the analysis accounts for wh-cleft constructions in various Arabic dialects in a systematic way. First, in wh-movement dialects such as Jordanian and Palestinian (Shlonsky 2002), wh-clefts are derived by wh-movement on a par with canonical wh-questions. Upon Agree between C, the probe, and the wh-phrase in Spec,TP, the goal, the wh-phrase undergoes movement to Spec,CP to satisfy the [EPP] property on C. Questions like (73a) above might be given the representation in (76) below after wh-movement has taken place.
Secondly, in wh-in-situ languages like Egyptian (Wahba 1984, Cheng 1991 and Soltan 2009), wh-clefts receive the same treatment as wh-in-situ questions in which wh-movement is not involved. Instead, the wh-phrase is bound in situ by a [WH] operator that is base-generated in Spec,CP. As such, the structure for a sentence like (77a) would look like that in (77b).

\[(77)\] a. miin illi Mona ṣaafit-uh?
   who that Mona saw.3fs-him
   ‘Who is it that Mona saw’?

\[\]
Under this analysis, then, Egyptian wh-clefts are wh-in-situ questions. The analysis here radically differs from Cheng’s analysis and is, in fact, more desirable in that it provides a uniform approach to Egyptian wh-questions. In both types of wh-question, i.e., wh-in-situ questions and wh-cLEFTs, a [+WH] operator is directly merged in Spec,CP and binds the wh-phrase in its base position. Compare, for instance, the structure for wh-in-situ questions in (78) below with (77) above.

(78)  a. Mona šaafit miin ?mbara?h?
    Mona saw.3fs who yesterday
    ‘Who did Mona see yesterday?’
Finally, in Lebanese, the situation is slightly different. This is because Lebanese is both a wh-movement and a wh-in-situ language where wh-in-situ is of the d-linked type (Aoun and Choueiri 1999, Aoun and Li 2003 and Aoun, Bennamoun and Choueiri 2010). As far as wh-clefts are concerned, they present an interesting case. The reason is that wh-clefts are by default d-linked wh-questions by virtue of the fact that they involve presupposition, as we have seen earlier. In addition, the cleft clause in wh-clefts has also been shown to be an instance of relativisation and resumption, both may be subsumed under d-linking. It follows that Lebanese wh-clefts are cases of d-linked wh-in-situ. As such, they receive similar treatment to Egyptian wh-clefts. The wh-phrase
remains in its original position and is licensed in situ by a null wh-operator. The structure for Lebanese wh-clefts may, thus, be roughly represented in (79) below.

(79) a. miin yalli Majdi t₃awwaz-a?
    who that Majdi married-her
    ‘Who is it that Majdi married’

b. 

As concerns wh-clefts with PRON, the wh-phrase remains in situ in Spec,TP, as shown in the following structure.

(80) a. miin hiyye yalli Majdi t₃awwaz-a?
    who she that Majdi married-her
    ‘Who is it that Majdi married’
Within the present analysis, Lebanese wh-clefts are interpreted via unselective binding (Baker 1970 and Pesetsky 1987) in a similar fashion to d-linked wh-in-situ questions (Aoun and Choueiri 1999), as we have seen in earlier parts of this study. This way, the analysis provides a unified approach to wh-clefts not only in Lebanese but also in other dialects like Egyptian, Jordanian and Palestinian. In wh-movement languages, wh-clefts pattern with wh-movement questions. In wh-in-situ languages, wh-clefts pattern with wh-in-situ questions.
Another consequence of the current analysis is that it captures cases where two strong pronouns forms co-occur in this type of question. To illustrate, consider the following example from Jordanian with its representation.

(81) a. miin hu huu illi šaft-o b-l-maTīyam?  
who he he that saw.2s in-the-restaurant  
‘Who is the one you saw in the restaurant’

b. 

The structure (81b) captures the two occurrences of the subject pronoun. The lower subject pronoun is the head of the relative clause located in Spec,DP of
the relative noun phrase. Meanwhile, the higher pronoun is the resumptive subject which appears in LD’ed structures, i.e., PRON, located in Spec,PredP.

To recap, this section has advanced an analysis whereby Arabic clefts have been shown to be a species of equative copular constructions. As such, there are two types of cleft in Arabic: the subject-predicate type of cleft which involves no PRON separating the two and the LD’ed type which involves PRON. Similarly, wh-clefts have been shown to pattern with equative copular wh-constructions which may either be simple subject-predicate clauses or LD’ed constructions involving PRON. The structures proposed for each type of construction have been shown to account for a number of facts about (wh-) clefts in several Arabic dialects.

4.5.4 English Clefts

The LD’ed equative analysis advanced in this study has implications for the analysis of English clefts. As might be recalled, English clefts have much in common with Arabic clefts. The structure advanced in this chapter for Arabic clefts, I would like to propose, can be also extended to English clefts. Thus, under the current analysis, a cleft sentence like (82a) is assigned the structure in (82b).

(82)  a. It was John that Mary saw
Within the current analysis, the copula is located under Pred and is attracted to T to check the [+Past] tense feature. The pronoun It is base-generated in Spec,TP like definite DPs in Arabic LD’ed equatives. The fact that it is a definite DP also accounts for the semantics of such constructions. The difference between English and Arabic resides in the nature of Spec,TP. While in Arabic Spec,TP is an A'-position, it is an A-position in English.

The analysis proposed here which assimilates Arabic clefts to LD’ed equatives presents a perfect case for the match between syntax and semantics in that the cleft clause both semantically and syntactically modifies the clefted DP. Such a view recalls the analysis proposed in Reeve 2010 for English clefts. Based on Higgins 1973, Percus 1997 and Hedberg 2000, Reeve maintains that English clefts are semantically parallel to equative sentences. In (83), for in-
stance, the two XPs that are (semantically) equated are the DP *it* in Spec,IP and the clefted DP *John*.

(83)

\[
\begin{array}{c}
\text{IP} \\
\text{DP} \quad \text{I'} \\
\text{it} \quad \text{I+V}_i \quad \text{VP} \\
\text{was} \quad \text{t}_i \quad \text{DP} \\
\text{DP} \quad \text{CP} \\
\text{John} \quad \text{that Mary saw}
\end{array}
\]

Reeve argues, however, that the cleft clause *that Mary saw* syntactically modifies, i.e., is adjoined to, the clefted DP, not the DP *it*, as can be seen in (83); a case of syntax-semantics mismatch. Such a view is challenged, however, by the current analysis for Arabic clefts. Arabic clefts provide evidence against Reeve’s view of the syntax-semantics mismatch that English clefts present. The current analysis differs in that it takes the cleft clause in Arabic clefts as well as English clefts to semantically and syntactically modify the clefted DP in Spec,TP through a sequence of agreement relations established between the constituents of the cleft sentence.

Before bringing this section to a close, there remains one issue that needs further scrutiny. The analysis so far has been dealing with wh-clefts (as well as equatives) involving only the argument Wh-DP ‘who’, which gives rise to the question of how wh-clefts involving the wh-expression *šu* are derived, given the CP-analysis proposed in the previous chapter. This will be the subject of the next section.
4.6  šu-lli-questions

As might be recalled, we started this chapter with a question carried over from the previous chapter about the conflict between properties of the Lebanese wh-expression šu. This wh-expression has been shown to display a mixed pattern of distributional properties that distinguishes it from other argumental Wh-DPs. For instance, it was shown that the wh-expression šu cannot stay in situ in argument position, whether in single wh-in-situ questions or whether in multiple wh-questions, it cannot be d-linked nor be related to a resumptive pronoun. To account for these properties, I have proposed that the wh-expression šu is a CP that can only be generated clause initially, i.e., Spec,CP. Following Koster 1978 and Alrenga 2005, I have argued that šu binds a null DP in the argument position. It was concluded that šu-questions are base-generated questions and, hence, differ syntactically from canonical wh-movement questions.

Meanwhile, one property that has been identified, yet remains to be explained, is the ability of šu to occur in contexts that are known to be restricted to DPs, such as (LD’ed) equative constructions and wh-clefts. The relevant sentences are repeated below.32

(84) a. ʔeyš hu l-soʔaal? (Jordanian)
   what it the-question
   ‘What is the question?’

   b. ʔeyš hu illi baddak tesʔahu-h?
   what (it) that want.2ms ask-it
   ‘What is it that you want to ask?’

(85) a. šu huwwa l-ʔakl l-yom? (Lebanese)
   what it the-food the-day
   ‘What is on the menu today?’

32As might be recalled from the previous chapters, Moroccan škun and Iraqi šeno have been analysed as DPs. However, unlike šu, both Moroccan škun and Iraqi šeno have DP distributional properties. This is a case of parametric variation.
b. šu huwwe illi rah teekhu-u l-yom?
   what it that Fut. eat.2p-it the-day
   ‘What is it that you are going to eat today?’

Like the Wh-DP ḥeyš ‘which thing’ in (84a) and (84b), the Wh-CP šu may appear in LD’ed equative sentences (85a) as well as in wh-clefts (85b). This property of šu is in conflict with its other properties and seems to challenge the CP-analysis proposed for canonical wh-questions. To address this conflict, this chapter embarked on an investigation of properties of wh-clefts and proposed an analysis whereby wh-clefts are treated on a par with equative constructions. The analysis is based on the view that the two DPs forming an equative sentence are: the cleft clause, and the Wh-DP. As such, in wh-movement languages like Jordanian and Palestinian, wh-clefts are derived by wh-movement of the wh-phrase from Spec,TP to Spec,CP; whereas, in wh-in-situ languages like Egyptian, wh-clefts are treated as wh-in-situ questions in which the wh-phrase is licensed in situ via unselective binding.

One interesting case was wh-clefts in Lebanese, a language which uses wh-movement and d-linked wh-in-situ in the formation of wh-questions. I have proposed that wh-clefts in this language are wh-in-situ questions on the grounds that cleft constructions are d-linked wh-questions. However, as may have been noticed, the wh-clefts that we have been dealing with so far involve the argument Wh-DP miin ‘who’, but not the Wh-CP šu. Given the asymmetries between these wh-phrases and given that Spec,TP of equatives and clefts can host only definite DPs, the proposed analysis gives rise to the question of whether wh-clefts involving the wh-expression šu have the same derivation as other argumental Wh-DPs. Although the CP-analysis of the wh-expression šu appears to be challenged by this property, i.e., its occurrence in DP contexts, closer examination reveals that it provides further support for the CP-analysis.
As far as ʔeγš-questions are concerned (84a,b), they are syntactically similar to the questions that involve the argument Wh-DP miin ‘who’. That is, in both cases, the wh-phrase originates in Spec,TP and undergoes movement to Spec,CP. Thus, a question like (84b) can be assigned the structure in (86).

As far as ʔeγš-questions are concerned (84a,b), they are syntactically similar to the questions that involve the argument Wh-DP miin ‘who’. That is, in both cases, the wh-phrase originates in Spec,TP and undergoes movement to Spec,CP. Thus, a question like (84b) can be assigned the structure in (86).

Concerning ʔə-u-questions, two alternatives may be explored. One alternative is to postulate some mechanism that transforms the wh-expression ʔə-u into a DP, or simply treat it as a DP, that therefore enables it to occur in such questions. Wh-clefts involving the wh-expression ʔə-u would then be derived in the same manner as other wh-clefts involving argumental Wh-DPs like miin ‘who’.

Concerning ʔə-u-questions, two alternatives may be explored. One alternative is to postulate some mechanism that transforms the wh-expression ʔə-u into a DP, or simply treat it as a DP, that therefore enables it to occur in such questions. Wh-clefts involving the wh-expression ʔə-u would then be derived in the same manner as other wh-clefts involving argumental Wh-DPs like miin ‘who’.

Concerning ʔə-u-questions, two alternatives may be explored. One alternative is to postulate some mechanism that transforms the wh-expression ʔə-u into a DP, or simply treat it as a DP, that therefore enables it to occur in such questions. Wh-clefts involving the wh-expression ʔə-u would then be derived in the same manner as other wh-clefts involving argumental Wh-DPs like miin ‘who’.

Concerning ʔə-u-questions, two alternatives may be explored. One alternative is to postulate some mechanism that transforms the wh-expression ʔə-u into a DP, or simply treat it as a DP, that therefore enables it to occur in such questions. Wh-clefts involving the wh-expression ʔə-u would then be derived in the same manner as other wh-clefts involving argumental Wh-DPs like miin ‘who’.
Based on the assumption that Lebanese wh-clefts are wh-in-situ questions, the wh-expression ˇs u would originate in Spec,TP and get bound in-situ by a [+WH] operator in Spec,CP.

However, the DP-analysis just outlined would give rise to several problems. Recall that the wh-expression ˇs u cannot stay in situ in single wh-in-situ questions and multiple wh-questions. Admitting that the wh-expression ˇs u stays in-situ in wh-clefts but not in other types of wh-question amounts to a contradiction. This also means that we have to give up the CP analysis of ˇs u-questions and search for other ways to account for the distributional asymmetries between the wh-expression ˇs u and other argumental Wh-DPs not only in single wh-in-situ questions and multiple wh-questions, but also in resumptive wh-questions and d-linked wh-questions.

In the context of wh-clefts, the DP-analysis gives rise to another set of problems. First, recall that there is a relationship between the resumptive pronoun in the cleft clause and the initial wh-expression, that is mediated through the null operator in Spec,illi. This relationship would not be possible with the wh-expression ˇs u which is resistant to resumption, as we have seen in resumptive wh-questions. The second problem concerns the restriction of wh-clefts to definite argument Wh-DPs. A DP-analysis grants the wh-expression ˇs u the status of a definite argumental Wh-DP. However, as we have seen, the wh-expression ˇs u does not behave like argumental Wh-DPs in other types of questions.

If, on the other hand, we extend the CP-analysis of the wh-expression ˇs u to wh-clefts in which it occurs, the problems just outlined will not arise. Put differently, by maintaining that ˇs u is indeed a CP that is base-generated in

\[33\] The problems associated with a DP analysis outlined above are based on the assumption that Lebanese wh-clefts are d-linked wh-in-situ questions. If, however, we treat wh-clefts as wh-movement questions and continue to assume that the wh-expression ˇs u is a DP, the conflict in its properties and the asymmetries with other argument wh-phrases would still remain unexplained.

251
Spec,CP of the wh-clause and that it binds a null DP in Spec,TP, the analysis captures the apparent conflict in the CP and DP properties of ŝu in the two types of wh-question. This way, the CP-analysis of ŝu offers a unified treatment for ŝu-questions, regardless of question types. Thus, the structure for a sentence like (85b) can be roughly represented in (87).

Notice that the presence of PRON in the structure also supports the CP-analysis. As we have seen before, the wh-expression ŝu cannot be related to pronominal forms since it is not a DP. In sentences like the above, it must be the case that PRON is related to the empty DP in Spec,TP, which is, in turn, bound by the wh-expression ŝu in Spec,CP. Recall that there is a coreference
relationship between PRON and the DP in Spec,TP which is what we see in this case.

In questions where no left-dislocation is involved, the structure involves a subject-predicate relationship between the null DP in Spec,TP and the DP predicate complement of T. The null DP gets bound by the wh-expression šu in Spec,CP. The structure might be represented as follows.

(88)

Under this analysis, the wh-expression šu behaves like a wh-operator binding a null DP in the subject position of the wh-cleft, as we have seen for verbal wh-questions in the previous chapter.

To sum up, this section was concerned with wh-clefts that involve the wh-expression šu and the question of whether the derivation of such questions is
the same as wh-clefts involving argumental Wh-DPs like *miin* ‘who’, given the asymmetries between these wh-expressions. The section has argued that a DP-analysis is unlikely to be an ideal solution in that it fails to capture the asymmetries in the two types of wh-question we have been dealing with. A DP-analysis is thus abandoned in favour of the CP-analysis advanced in the previous chapter. Under the CP-analysis, all questions involving the wh-expression *šu* are base-generated questions in which the argument position is occupied by a null DP. The null DP is bound by the wh-expression *šu* acting as an operator that is directly merged in Spec position of the matrix wh-clause.

4.7 Summary and Conclusion

In this chapter, I have advanced an approach for clefts and wh-clefts in Arabic whereby they are treated as equative copular constructions. The chapter commenced with a question concerning the argument properties that the wh-expression *šu* displays, which appeared to challenge the CP analysis developed in the previous chapter. More specifically, it was shown that despite its CP status, it occurs in contexts that only admit DPs, such as equative constructions of the form [Wh-DP DP]/[Wh-DP PRON DP] and the so-called reduced cleft wh-questions of the form [Wh-DP RC]/[Wh-DP PRON RC].

To resolve this conflict, I began an investigation about Arabic cleft constructions and their wh-counterparts. I have shown that such constructions display an interesting set of properties that distinguish them from English clefts, in spite of certain syntactic and semantic similarities between the two languages. The main differences were shown to reside in the restriction of Arabic (wh-) clefts to definite DPs; the occurrence of resumptive pronouns in object relatives but not in subject relatives; and the obligatory appearance of PRON in clefts vs. its optional appearance in wh-clefts. Two previous approaches were subsequently discussed: the reduced cleft analysis of Cheng 1991 and the equative analysis
of Shlonsky 2002. I have shown that although both are correct on their own terms, some refinement was still necessary.

A more elaborated proposal was then spelled out, central to which is the view that clefts are a species of the equative copular construction as analysed in the first chapter of the present thesis. Several arguments from the syntax and semantics of both constructions have also been presented in support for the analysis. This was followed by positing a more appropriate structure based on the PredP framework (Bowers 1993, 2001; Chomsky 2000 and Adger and Ramchand 2003). The facts from various languages were shown to follow from the proposed analysis.

The last part of this chapter returned to the question of whether the proposed analysis captures the asymmetries observed between argument wh-phrases like miim ‘who’ and ?ayya-NP ‘which’-NP phrase vs. the wh-expression šu in canonical wh-questions. Two alternatives were discussed, i.e., a DP-analysis and the CP-analysis developed in the preceding chapter. It was concluded that the CP-analysis is more plausible, and, in fact, more desirable, in that not only does it avoid the problems associated with a DP-analysis, but it also offers a unified account for the properties of wh-phrases in both types of wh-question.
Chapter 5

Summary and Conclusion

The central goal of this thesis has been to address the question of how the internal syntax of Arabic wh-phrases affects their external syntax. In particular, wh-questions in languages such as Iraqi and Lebanese have been shown to have more hybrid syntax than we ever thought. The hybrid nature of the syntax of wh-questions in the languages under consideration lies in the fact that they employ for the formation of wh-questions more than one strategy, such as the movement strategy, the in-situ strategy, the base-generation strategy and the resumptive strategy. Furthermore, each one of these strategies is associated with an array of properties that sets it apart from other strategies within each dialect and across. I have shown, for instance, that wh-in-situ in both Iraqi and Lebanese has a different distribution from wh-movement in each and between the two dialects. The table below summarises how these strategies are applied in the dialects examined in this work.
(1) Interrogative Strategies in Arabic Dialects

<table>
<thead>
<tr>
<th></th>
<th>Iraqi</th>
<th>Lebanese</th>
<th>Jordanian</th>
<th>Egyptian</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Wh-movement:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. ([+_Q]+)EPP on C</td>
<td>-</td>
<td>miin</td>
<td>miin</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>'who'</td>
<td>'who'</td>
<td>'who'</td>
<td></td>
</tr>
<tr>
<td>b. ([+_WH]) on Wh-phrase</td>
<td>-</td>
<td>men</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>'who'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>šen</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>'what'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Wh-in-situ:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Binding by ([+_Q]) C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a- D-linked</td>
<td>meno</td>
<td>miin</td>
<td>-</td>
<td>miin</td>
</tr>
<tr>
<td></td>
<td>'who'</td>
<td>'who'</td>
<td>'who'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>šeno</td>
<td>?ayya-NP</td>
<td>-</td>
<td>?eh</td>
</tr>
<tr>
<td></td>
<td>'what'</td>
<td>'which'-NP</td>
<td>'what'</td>
<td></td>
</tr>
<tr>
<td>b- Non-D-linked</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>miin</td>
</tr>
<tr>
<td></td>
<td>'who'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>?eh</td>
</tr>
<tr>
<td></td>
<td>'what'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Base-generation</td>
<td>-</td>
<td>šu</td>
<td>šu</td>
<td>-</td>
</tr>
<tr>
<td>([_CP Wh-Op_i\ldots_DP_\emptyset_i])</td>
<td>'what'</td>
<td>'what'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Resumption</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. [Wh-DP...[illi-RC RP]]</td>
<td>men</td>
<td>miin</td>
<td>miin</td>
<td>miin/?eh</td>
</tr>
<tr>
<td></td>
<td>'who'</td>
<td>'who'</td>
<td>'who'</td>
<td>'who'/what'</td>
</tr>
<tr>
<td>b. [Wh-DP ...RP]</td>
<td>-</td>
<td>miin</td>
<td>miin</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>'who'</td>
<td>'who'</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>?ayya-NP</td>
<td>?ayya-NP</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>'Which'-NP</td>
<td>'Which'-NP</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

257
There was a clear sense in which the type of variation in wh-constructions in these dialects is not solely parametric. A number of questions were put forward to address the source of the variation observed here. The first question was concerned with the notion of parametric variation and whether the facts from Arabic wh-constructions can be captured within current approaches to cross-linguistic variation that are mainly couched within the P&P framework (Chomsky 1981, 1995, 2000). Within the P&P framework and its recent reformulations in the Minimalist Program, languages are said to be typologically divided into either wh-movement languages or wh-in-situ languages. This typological divide is essentially motivated by the assumption that languages either have an EPP feature on C, the sole motivator of wh-movement, or they don’t. However, the facts from various dialects show that languages do not comply with such a strict typological divide, i.e., whether the wh-parameter is set to either a +EPP or a -EPP on C. At first, it appeared that we had to abandon this strict typological division between languages. However, rather than rejecting such an approach to parametric variation outright, and without delving into complicated philosophical issues about which theory for parametric variation one should, or should not, adopt, I have shown that the wh-parameter in its current guise is, as a matter of fact, too rigid and must therefore allow other factors into play.

The second question brought forward as a result was what factors must be considered to explain the variation facts? The answer to this question gradually materialised in the four chapters that made up this thesis. Chapter one proposed a unified syntactic analysis for Arabic copular constructions. Under the analysis proposed, constructions which have the schematic structure [DP XP] are treated as subject-predicate clauses. The chapter has also proposed a L(eft)-D(islocation) analysis for the so-called equative sentences that systematically display the schematic structure [DP [PRON DP]]. Based on observations from LD in verbal predications, such as CLLD and LD in SV constructions in
Arabic, I have argued that the latter is a derived construction from the simple [DP DP] subject-predicate construction and is, therefore, best analysed as an LD structure. A number of consequences were shown to follow from the LD analysis of equatives. The first consequence is that LD is also possible in simple predicational sentences on a par with equatives, but only in cases where the subject is a definite DP. More importantly, the LD analysis offered a natural solution to the problems associated with PRON which have been a thorny issue for analyses. The solution lies in the treatment of PRON as a subject pronoun, which is what it is, and one that functions as such. As a consequence, the LD analysis solved the problem of the putative optionality of PRON in such constructions and their wh-counterparts by showing that cases with PRON and those without represent two different structures, hence, optionality is only apparent. Equally important, the analysis warranted promising avenues for syntactic studies of Arabic copular wh-constructions including the so-called reduced cleft wh-questions (Cheng 1991 and Ouhalla 1996) or Class II interrogatives (Shlonsky 2002 and Aoun, Benmamoun and Choueiri 2010).

Chapter two dealt with Iraqi wh-asymmetries and proposed an analysis based on the morpho-syntax of wh-phrases. I have argued that Iraqi argument wh-phrases are [WH] free relatives that are bound in situ in a similar fashion to d-linked wh-phrases. As a start, the chapter highlighted the core problem with wh-in-situ constructions in Iraqi and Lebanese. I have shown that wh-in-situ displays an asymmetric variation with regards to the distribution of argument wh-phrases. Moreover, I have shown that current approaches to parametric variation are incompatible with the facts reported from Iraqi and Lebanese wh-constructions. In particular, I discussed two main approaches, namely, the LF-movement approach advanced in Huang 1982 and the binding approach developed in Pesetsky 1987. A number of arguments from both Iraqi and Lebanese have been presented to show that approaches exploiting LF movement alone are simply inadequate and fail to capture the variation observed. Meanwhile, bind-
ing by an operator seemed more likely to account for the data, but only when the
morpho-syntax of wh-phrases is taken into account. Hence, a morpho-syntactic
analysis was called for.

In developing our morpho-syntactic analysis, I have argued that Iraqi ar-

gument wh-phrases have complex internal structure, identical to copular wh-

clauses analysed in the preceding chapter, which involves internal movement of

the wh-phrase to the Spec position of its own clause, i.e., Spec,CP. The analysis

was based on the view that Iraqi wh-phrases contain a wh-phrase and a con-

tracted version of a third person subject pronoun yielding a [DP DP] structure.

Such a structure has been shown to be characteristic of equative copular con-

structions of the type discussed in chapter one of this thesis.

I then moved on to discuss the consequences of this analysis for wh-in-situ

in Iraqi. It was argued that the overall structure of meno is identical to the

structure of free relatives. However, this posed a syntax-semantics mismatch

problem and the question of how a CP like meno occurs in DP distributional

contexts in Iraqi wh-in-situ questions. The problem was overcome by considera-

tions of analyses of free relatives which present a similar type of conflict. I have

shown that the common argument in the majority of studies on free relatives is

that these are DPs and must be structurally represented as such. To this effect,

I have followed Citko’s 2008 DP analysis of free relatives in which movement of

the wh-phrase to Spec,CP is said to result in projecting the label of the moved

element. Based on observations from Arabic free relatives, it was concluded that

Iraqi meno is indeed a free relative DP bearing an interrogative feature and that

it is this feature that sets it apart from standard, i.e., non-interrogative, free

relatives. Reasoned as such, the facts from Iraqi wh-in-situ followed and a sec-

ond conclusion was drawn to the understanding that Iraqi uses relativisation as

a means of d-linking. The latter was said to be responsible for the licensing of

wh-in-situ in this language.
Chapter three solved the problem posed by the asymmetric variation in Lebanese wh-constructions. This variation lied in the conflict between argument wh-phrases *miin* ‘who’ and *?agga*-NP ‘which’-NP and the wh-phrase *šu* ‘what’. Whereas the former were shown to display properties characteristic of argument wh-phrases in the various wh-contexts, the latter was shown to present yet another conflict in its non-argument and argument distributional properties. This had led me to put forward the hypothesis that the wh-phrase *šu* ‘what’ is not a single lexical item, but rather, a complex wh-expression. Support for this hypothesis was, then, obtained from observations about the morphology of similar forms in various Arabic dialects. Indeed, it was shown that this wh-element is a composite of the three-word string *?ay ši hu* ‘which thing it’. A syntactic analysis was proposed on the basis of this finding with the conclusion that the wh-phrase *šu* ‘what’ is a copular clause with a CP structure. The CP analysis was, then, shown to account for all the non-argument properties that this wh-expression displays and explained the asymmetries in various wh-constructions. In particular, I have argued that this wh-phrase can only be base-generated clause initially on a par with sentential subjects and objects of the type analysed in Koster 1978 and Alrenga 2005. It was further concluded that wh-questions in which the wh-phrase *šu* occurs syntactically differ from canonical wh-questions in terms of binding for the former and wh-movement for the latter. Meanwhile, the argument properties that wh-phrase *šu* displays appeared to constitute a challenge for the CP analysis; a challenge that chapter four rose to.

Chapter four explored another type of wh-question in Arabic, namely, copular wh-constructions and reduced cleft wh-questions, which are restricted to DPs and the question of how the wh-phrase *šu*, analysed as a CP, appears in such constructions. After presenting properties associated with this type of questions, it was shown that past attempts have failed to account for all the
facts presented here. I have provided a unified analysis that was essentially based on the analysis developed in chapter one. In particular, after presenting evidence that the cleft clause is a free relative DP, I have argued that clefts pattern with equative copular constructions of the form [DP DP]; whereas, clefts of the form [DP PRON DP] pattern with LD’ed equatives. The analysis proposed was shown to account for a number of related facts from various Arabic dialects. Returning to the foreground was the question concerning the DP distribution of the wh-phrase šu in such constructions. Given the conclusion that copular constructions allow only DPs, the CP analysis seemed more challenged than ever before. I have explored two alternatives. The first alternative was the assumption that some mechanism transforms the CP into a DP, like that proposed for iraqi meno in the preceding chapter, so that we account for the distribution of the wh-phrase šu in copular constructions. However, this alternative gave rise to more problems than it apparently solved. The second alternative was to maintain the CP analysis and treat copular wh-questions in which the wh-phrase šu occurs as base-generated questions on a par with verbal wh-questions. The CP analysis proved more desirable in that it captures all the asymmetries associated with the wh-phrase šu while providing a uniform account for the different types of questions in which it appears.

To sum up, the analysis proposed here for the wh-asymmetries in the spoken dialects provides compelling evidence for a micro-parametric approach to the syntax of wh-questions and simultaneously shows that the parametric approach of the P&P framework (Chomsky 1981, 1995) is too broad to account for the facts reported in this work. After careful examination of the morphological makeup of Arabic argument wh-phrases, I have shown that their concealed internal syntax is the only way to connect to their external syntax. The approach developed here strongly suggests that syntax operates at micro-levels starting from single morphemes ending with larger structures in the course of derivation of any given configuration.
Extensions, Implications and Questions for Future Research

There are a number of issues and phenomena that have not received full attention in the course of this study, for reasons to do with the scope of this study, space and time. That should not mean, however, that such issues are less important than those discussed here. These issues, a summary of which is given below, raise a number of interesting questions which can be taken up further in future research.

One of advantages of the analysis advocated in this thesis is that it can be neatly extended to various Arabic dialects that either have been touched upon or have not been considered here, especially, the western Arabic dialects like Moroccan and Libyan. These dialects seem to have much in common with Iraqi in terms of the morphological composition of wh-expressions. It would be interesting to see what role morpho-syntax plays in the distribution of wh-phrases in these dialects and whether such distribution differs from that of Iraqi and Lebanese. This will give more substance to the view that syntax operates at the micro-level in the derivation of larger structures.

Equally interesting is another area which we have not addressed here concerning the syntax of Arabic adjunct wh-phrases and the question of whether their internal syntax is as decisive as that of argument wh-phrases. Recall that the spoken dialects developed various wh-forms which have no counterparts in Standard Arabic, as we have seen in the case of argument wh-phrases. Like argument wh-phrases, adjunct wh-phrases are complex wh-expressions and have distributional patterns that distinguish them from their counterparts in other languages. In fact, the role of adjunct wh-phrases in the syntax of wh-questions in Arabic is almost absent. The analysis provided here certainly opens the way for future research on this very topic.
Another area which has not been capitalised on is the semantics of copular wh-phrases and the questions in which they occur. The question that can be raised in this context is whether or not the syntax and semantics of such questions stand on a strict one-to-one relationship. The data seem to provide evidence against the widely held compositional view between the two. However, this should not be taken as such. This is a particularly interesting subject given the propositional nature of Arabic wh-phrases. A more in-depth analysis will be able to tackle these issues further and perhaps reconcile the type of mismatch that might be observed.

The other phenomenon that deserves further investigation is agreement in LD’ed equatives which have been shown to display the format [DP [PRON DP]]. As presented in the first chapter, the patterns of agreement that the subject pronoun displays in such constructions are intriguing. It was observed that such patterns are apparently inconsistent ranging over full agreement, partial agreement or no agreement at all. Even in cases where agreement holds, it was not clear whether agreement is between PRON and its predicate or whether it is between PRON and its antecedent topic. The same also holds for cleft constructions in Arabic. This gives rise to a handful of questions about these issues that can be pursued in a separate study.

Last, but not least, the analysis offered in this study has crucial implications for recent approaches to parametric variation that seek to address typological variation from a new perspective while departing away from the P&P approach. This study has shown that the external syntax of wh-phrases operates at a morphemic level, and in some cases, at a sub-morphemic level, as is the case for Iraqi and Lebanese wh-phrases. While further work is still needed on Spell-Out strategies, the data presented here would provide a good testing ground and, I
believe, a valuable contribution to new trends in syntactic theory such as the Nanosyntax framework, currently being developed by an ever increasing number of researchers in the (sub)fields of linguistics.
Bibliography


Boeckx, Cedric. 2010. What Principles and Parameters Got Wrong. ICREA/UAB.


