Acquisition of Modern Standard Arabic by Speakers of Different Arabic Colloquial Varieties

Sami Saleh Alresaini

Submitted for the degree of PhD in Linguistics

The University of York

Department of Language & Linguistic Science

June 2012

Acquisition of Modern Standard Arabic by Speakers of Different Arabic Colloquial Varieties

Sami Saleh Alresaini

Submitted for the degree of PhD in Linguistics

The University of York

Department of Language & Linguistic Science

June 2012

Abstract

This thesis conducts an empirical investigation of Arabic speakers' underlying ultimate knowledge of Modern Standard Arabic (MSA). The goal is to determine whether these speakers' end-state MSA grammar can be classed as a native language (L1) type grammar or a second language (L2) type grammar. The motivation for this research comes from the frequent claim in the literature that there are no native speakers of MSA (e.g., Kaye, 1970; Maamouri, 1998). This claim has been made because MSA is not spoken at home and it is acquired through literacy and formal schooling which does not start for most children until age 5 or 6. Another factor that can support such a claim is that children who start acquiring MSA at school already speak their colloquial varieties which were acquired naturally from birth. Because there are differences between the colloquial varieties and MSA in all linguistic domains (e.g., Altoma, 1969; Ayari, 1996; Maamouri, 1998), this has been assumed to have possibly affected the ultimate attainment of MSA and resulted in different end-state grammars; a characteristic of a second language. However, the claim that there are no native speakers of MSA has not been investigated empirically. In fact, there is a clear scarcity of empirical works in the literature that investigate and discuss Arabic speakers' end-state underlying knowledge of MSA. The current thesis aims to fill in this gap.

The current research examines the impact of age of first exposure (AoE) and knowledge of the previously acquired dialect (L1) on the ultimate attainment of MSA acquired by Arabs across three regions of the Arabic-speaking world: Egypt, the Levant and the Gulf regions. The main objective is to explore how MSA's end-state underlying grammar is represented by native speakers of different colloquial varieties whose first exposure to MSA also varied between exceptionally 'early' exposure (from age two or three in MSA immersion nurseries and kindergartens) and typically 'late' exposure to MSA (from age five or six in normal primary schools). Based on results of a corpus study of colloquial varieties, the syntactic variables for investigation were identified as: i) resumption in definite object relative clauses, ii) collective subject-verb agreement in SV sentences, and iii) word order preference in conversational and narrative contexts.

147 adolescent participants were recruited to take part in an experiment designed to examine their underlying knowledge of the three syntactic phenomena in MSA. The participants completed two tasks: an acceptability judgement task and a conversation role-play task. Using ANOVA and planned comparisons, the differences in participants' performance in these tasks were evaluated across five groups corresponding to different AoE and the colloquial varieties the participants speak. The analysis of the data showed no significant effect of AoE or of the L1, and post hoc tests showed no significant differences between the groups of participants. These results were discussed in relation to previous views on L1 influence and effect of age of onset on the ultimate attainment of a second language.

Uniform performance in the tasks of the study, despite variant AoE and L1s, was interpreted to indicate success of typical Arabic speakers in attaining a native-like competence in at least the three MSA syntactic phenomena investigated. Further research that involves investigation of end-state knowledge of more grammatical phenomena in different linguistic domains is required to make a comprehensive assessment of typical Arabic speakers' end-state knowledge of MSA.

List of Contents

List of Tables

List of Figures

Transliteration Scheme Used for Arabic

Arabic Script	IPA	Symbol used here
أ	?	?
ب	b	b
ت	t	t
ث	θ	θ
ج	ʒ - dʒ	j
ح	ħ	H
خ	x	x
د	d	d
ذ	ð	ð
ر	r	r
ز	z	z
س	s	s
ش	ʃ	š
ص	sˤ	S
ض	dˤ	D
ط	tˤ	T
ظ	Δˤ	Z
ع	ʕ	9
غ	ɣ	G
ف	f	f
ق	q	q
ك	k	k
ل	l	l
م	m	m
ن	n	n
ه	h	h
و	w	w
ي	j	y

ا	a:	aa
ي	i:	ii
و	u:	uu
Ó	a	a
Ọ	ɪ	i
Ọ	ʊ	u

- Modified from Hellmuth (2006:7)

10

Acknowledgements

It would not have been possible to write this thesis without the help and support of many people around me, to only some of whom it is possible to give particular mention here.

A special and sincere 'thank you' goes to my family: my parents, for giving me life in the first place, for educating me with aspects from both arts and sciences, for missing me the most, for their unconditional support and encouragement to pursue my interests, even when interests went beyond boundaries of language and geography. I would like to thank my dear wife, for being beside me throughout all the stages of this work, for listening to my complaints and frustrations, and for believing in me. Thanks to my brothers and sisters who kept calling me offering their help of all kinds and who made my visits to my home country more enjoyable and less costly.

This thesis would not have been possible without the help, support and patience of my two supervisors: Dr. Heather Marsden, to whom I am sincerely and heartily grateful for her guidance throughout completing this thesis, and Dr. Bernadett Plunkett, to whom I am indebted and thankful for her insightful advice and comments on this work.

I would like to show my gratitude to people who kindly have been members in my thesis advisory panel: Prof. Helen Goodluck, Dr. George Tsoulas, and Dr. Sam Hellmuth along with my two supervisors. I also would like to thank other academic staff members for offering their comments on my work on several occasions including Prof. Susan Pintzuk, Dr. Hidekazu Tanaka, and Dr. Tamar Keren-Portnoy.

Acknowledgements

Declaration

This thesis contains original work completed by Sami Alresaini under the supervision of Dr. Heather Marsden and Dr. Bernadette Plunkett.

This work has not previously been submitted for a degree at the University of York or any other university or institute of learning.

Chapter 1

Introduction

This thesis investigates ultimate attainment in acquiring Modern Standard Arabic (MSA) by speakers of different Arabic colloquial varieties. The goal is to dermine whether these speakers' end-state MSA grammar can be classed as a native language (L1) type grammar or a second language (L2) type grammar. This introductory chapter explains what the research is about and describes the linguistic contexts within which two groups of participants in an experimental study have acquired MSA. The first section defines the research focus of this thesis. Section two lays out the research questions and the thesis hypothesis. Section three describes the diglossic situation in the Arabic-speaking countries including issues such as how MSA is normally acquired by Arabic speakers compared to acquiring the local colloquial varieties of Arabic. The fourth section introduces an MSA Immersion Program which is designed to provide natural MSA input to Arab children during the pre-school period and aims to make them grow into native speakers of this language. The last section summarizes this chapter and outlines the organization of the rest of this thesis.

1   Focus of the Thesis

This thesis aims to make an empirical investigation of Arabic speakers' underlying ultimate knowledge of MSA to see whether this knowledge is of an L1 or of an L2 type. The motivation for this research comes from the frequent claim in the literature that there are no native speakers of MSA. As will be explained later in this chapter, this

claim perhaps can be supported by the fact that Arab children do not acquire MSA naturally at home and their exposure to MSA starts only at the primary school when their innate ability to acquire language natively is assumed to be no longer available or to have started to deteriorate. Another support for such a claim comes from the fact that Arab children have already acquired and speak the local colloquial form by the time they start learning MSA at primary school. Because there are some differences between the colloquial forms and MSA across all linguistic domains, this may lead different speakers to develop different MSA end-state grammars; which is a characteristic of a second language. However, as far as the researcher knows, there is a clear scarcity of empirical works in the literature that discuss Arabs' acquisition of MSA, investigate their underlying knowledge of it, and confirm or disconfirm claims about whether there are native speakers of MSA.

This research was conducted to fill this gap and examine the ultimate attainment of MSA acquired by Arabs across three regions of the Arabic-speaking world: Egypt, the Levant and the Gulf regions. The main objective is to explore how MSA's underlying grammar is represented by native speakers of different colloquial varieties across these regions. Basically, certain syntactic phenomena which are represented differently in the native grammars of colloquial varieties will be examined in the underlying end-state grammar of MSA in the three regions mentioned above. If MSA is acquired natively, then the underlying steady-state grammar of MSA is expected to be uniform in these different regions despite the syntactic differences between the colloquial varieties. This is based on the widely accepted assumption that all acquirers of the same first language or dialect achieve the same steady-state grammar (Guasti, 2002:4; White, 2003:241; Meisel, 2011:22). On the other hand, if MSA is learned or acquired as an L2, then the

underlying steady-state grammar is expected to be represented differently, based on the assumption that L2 speakers differ from each other in their ultimate attainment, even in the case of speakers with the same L1 who have acquired the same L2 (White, 2003:241; Meisel, 2008:57).

The primary goal of this thesis will be to investigate knowledge of a number of syntactic variables in MSA, in two contrasting groups: participants whose first 'proper' exposure to MSA was not before the age of entrance to formal primary schools; and participants who were, exceptionally, exposed to MSA from as early as the age of two. The latter group are the graduates of nurseries and kindergartens which apply the MSA immersion program to provide MSA input for children from age two up to age six. As will be explained later, these nurseries and kindergartens were deliberately founded to enable children to acquire MSA naturally as a first language before the age of six. In these institutions, MSA is the only language permitted to be used for both teaching and communication throughout the whole school day. The results of such a comparison should shed light on the issue of age of first exposure effects on ultimate attainment and the issue of the critical period hypothesis.

The syntactic aspects under study are represented differently in the grammars of the colloquial varieties. This is verified by the results of a corpus based study that will be presented in chapter 3 of this thesis. The syntactic phenomena are i) resumption in object relative clauses, ii) collective subject-verb agreement in SV sentences, and iii) word order preference in conversational and narrative contexts. Thus, the participants in each group will be speakers of different colloquial dialects as their L1. This will provide

16

a chance to test if these speakers have transferred their L1 syntactic properties to MSA grammar. If such a finding is reached, this asserts that MSA grammar is not represented natively.

2 Research Questions & Hypothesis

This thesis should lead to answers to the following interrelated questions:

I.   Do Arabs across the Arab world represent 'uniform' (L1-type) or 'variant' (L2-type) versions of MSA end-state grammar?

II. If variation in attainment exists, is it an effect of different age of first exposure to MSA?

III. If variation in attainment exists, is it an effect of the differences between the colloquial varieties (the L1s)?

The first question seeks to determine whether or not Arabs have a native competence of MSA, given that most of them do not start learning MSA until they reach the age of primary school entrance, and that the participants of the main study of this research speak syntactically different colloquial dialects as their L1s. If the participants of the study are found to represent uniform versions of MSA, this may indicate that those participants are native speakers of MSA. However, if the participants are found to represent variant versions of MSA end-state grammar, this means that they represent an L2-type of MSA competence. The questions in (2) and (3) address the cause of variation in ultimate attainment among the speakers of MSA (if it exists).

The working hypothesis behind this thesis is that only participants who have been exposed to natural MSA input during their preschool years will demonstrate a 'uniform'

representation of MSA end-state grammar, regardless of the colloquial form of Arabic they speak at home. This is based on the assumption that these participants have started acquiring MSA within the critical period which is claimed here to start to decline towards its end from age 6. On the other hand, it is expected that participants who were first exposed to MSA in primary school will show variant representations of MSA end-state grammars due to the differences between the colloquial varieties they speak as their L1s and to their late exposure to MSA.

The remaining part of this chapter presents background information to describe, in one section, the linguistic situation in the Arabic-speaking countries in terms of how MSA coexists with another local colloquial variety of Arabic. This section compares MSA and the local colloquial varieties in terms of their linguistic functions, Arabic speakers' attitudes toward each form, grammatical differences between the two forms, how Arabic speakers are normally exposed to and acquire each form, and the negative impact of this diglossic situation. The fourth section introduces the MSA Immersion Program which is designed as a proposal to alleviate the problem of diglossia by providing natural MSA input to Arab children during the pre-school period and make them grow into native speakers of this form of the language as well as the colloquial form. The content of these two sections serves as a general description of the linguistic contexts within which the two main groups of participants in the main experimental study of this thesis have acquired MSA. The final section summarizes the chapter and outlines how the rest of this thesis is organized.

3   Arabic Diglossia & Acquisition of MSA

The linguistic phenomenon of diglossia is the coexistence of two divergent language varieties in the same speech community to serve different functions (see Ferguson, 1959). This applies exactly to the linguistic situation in all Arabic speaking countries (e.g., Al-Batal, 1992; Haeri, 2000). In each country of the Arab world, Modern Standard Arabic (MSA) is used mainly for writing and in formal situations along with the local colloquial variety of that country which is mainly used for everyday speech. The local colloquial varieties vary greatly in all linguistic domains across the Arabic speaking countries and the intelligibility between one colloquial variety and another varies depending on the geographical distance between the regions where they are spoken (Aoun et al., 2010). MSA, however, is considered as 'uniform' across the Arabic-speaking world (Altoma, 1969), and thus, it serves as a lingua franca whenever needed to bridge the poor intelligibility between the colloquial varieties.

Ferguson (1959), among many others, provided lists to exemplify appropriate situations in which MSA and the local colloquial variety are used; and pointed out that it is essential to use the correct variety in the appropriate situation to avoid being an object of ridicule. MSA is appropriately used in writing and in formal situations in general. Such situations include religious sermons, personal letters, speech in parliament or in formal political speech in general, university lectures, news broadcasts, newspaper editorials, news stories, and poetry (Ferguson, 1959). MSA is also used in formal schooling as the language of instruction and as the content of Arabic literature and history (Maamouri, 1998:31).

The local colloquial varieties of Arabic, on the other hand, are used in informal situations such as instructions to servants, waiters, workmen, and clerks; in ordinary conversations with family, friends and colleagues at home or mostly everywhere else outside schools or formal settings; in radio and TV 'soap operas'; and in folk literature (Ferguson, 1959; Maamouri, 1998). Colloquial varieties can also be used in plays, advertisements, health messages, common political speeches and meetings, and in court discussions and related activities (Maamouri, 1998). In short, the spoken local colloquial varieties are used mainly as the languages of informal daily life communications (Zughoul, 1980; Holes, 1995).

Ferguson (1959) explained that code-switching between the two varieties can occur in a single situation where different linguistic functions are needed. An example of such a situation would be giving a formal university lecture in MSA and conducting informal drills, explanations, and section meetings by using the spoken colloquial variety. Teachers in schools, for example, usually spend a considerable amount of their time speaking in the colloquial variety to explain the meaning of material which has been presented in lectures and books in MSA.

Despite the fact that each variety has a distinct function in the speech community, Arabic speakers have different attitudes toward MSA as compared to the local colloquial varieties of Arabic. MSA always has a special and superior status amongst speakers of Arabic (e.g., Maamouri, 1998) and it is the official form of Arabic in all Arabic speaking countries (Altoma, 1969). To most Arabs, 'Arabic language' refers to MSA alone as the only 'real' form of the language and the other colloquial varieties are

considered as not real or even do not exist as languages (Ferguson, 1959; Maamouri, 1998). MSA is considered to be prestigious, more beautiful, more logical, and more capable of expressing abstract and complex thoughts (e.g., Ferguson, 1959; Ayari, 1996).

In contrast, Arabs' attitudes toward the spoken local varieties are less appreciative.[1] These colloquial varieties of Arabic are looked at by many Arabs not only as inferior to MSA but as distortions of the 'real' form and thus to be disregarded (Zughoul, 1980). These local varieties are usually associated with ignorance, illiteracy, and vulgarity. To many Arabs, the colloquial varieties are only suitable for simple daily life communications and cannot be capable of expressing abstract or complex thoughts (Zughoul, 1980; Ayari, 1996). The low status of these colloquial varieties among Arabic speakers may have brought about the scarcity of studies on colloquial Arabic dialects; traditional grammarians viewed them as unworthy of analysis (Khamis-Dakwar, 2011).

Shouby (1951), among others, explained that the higher status bestowed on MSA was due to its sacred value gained by the fact that it is the language of Quran. MSA also is the language of the vast Arabic literature produced over many centuries which is highly appreciated and respected by most Arabic speakers. Being claimed 'uniform' across Arabic speaking countries and the only form that bridges the intelligibility problems between some speakers of different local colloquial varieties, MSA is considered as a symbol of unity among the Arabic nations. It is also usually connected to the pride of

---

[1] Ibrahim (1986) claims that some varieties of the local dialects in each Arabic speaking community can be regarded as more prestigious than others despite the fact that all are considered as non-standard. See Ibrahim (1986) for elaboration on this point.

the past history of Arabs being the form of the language used by people in earlier eras. Also, MSA is the language of literacy and education and due to this it is believed to be superior to the spoken forms even by illiterate or less educated people.

As is the case in any other diglossic situation, there are many differences between the grammatical structures of the two coexistent varieties of Arabic (Ferguson, 1959). In general, the colloquial varieties are described by many authors (not in generative grammar terms) as showing simpler grammatical structures than MSA (e.g., Ferguson, 1959; Zughoul, 1980; Suleiman, 1986; Holes, 1995; Stevens, 2006) as exemplified below. Although there are many similarities between MSA and the spoken varieties, the differences between the two forms of Arabic are observed as occurring in all linguistic domains (e.g., Altoma, 1969; Ayari, 1996; Maamouri, 1998; Khamis-Dakwar et al., 2012).

Maamouri (1998), among others, provided a list of general differences between MSA and the spoken varieties. In the domain of phonology, there are 28 consonants in MSA, 3 short vowels, and 3 long vowels, whereas the vocalic structures in the colloquial varieties seem to be more complex than this. The lexicon used in both coexistent varieties of Arabic, standard and colloquial, is described to be rich (Maamouri, 1998) but greatly divergent (Suleiman, 1986). Speakers of MSA tend to use its almost unlimited derivation system to make more use of the available lexicon, and when speaking the colloquial variety, they seem to feel less conservative towards borrowing lexicon from foreign sources (Maamouri, 1998). Most of the lexicon used in the colloquial forms of Arabic, except for foreign loan words, has its origins in MSA, but

the lexical items themselves usually have different morphological or phonetic properties (Altoma, 1969).

In terms of syntax, MSA has a complex case marking system to distinguish grammatical functions with overt case inflections at the end of each word. For example, the final inflection in a word such as الولد ?al-walad-u 'the-boy-Nom.' indicates that the case is nominative, whereas it is accusative in الولد ?al-walad-a 'the-boy-Acc.' in which a different final inflection is used. This has been completely abandoned in the colloquial varieties making distinction between the grammatical or thematic functions highly dependent on the context (e.g., Holes, 1995; Maamouri, 1998; Mansouri, 2000). Moreover, MSA has different morphological inflections to mark differences in number (singular, dual, and plural) and gender (masculine and feminine). None of the colloquial varieties, on the other hand, have the dual form and many other forms have been lost in one colloquial variety or another. An example of this is the loss of the plural feminine in colloquial Tunisian Arabic which seems to be still used in other colloquial varieties (Maamouri, 1998:36). The final example of the general differences between MSA and the colloquial varieties is that the common word order in MSA is VSO, whereas in the colloquial varieties, it is claimed to be SVO (Shlonsky, 1997; Maamouri, 1998).[2]

Given that MSA is considered by the majority of Arabic speakers to be the only form of Arabic that is 'real' language, most Arabs indicate that MSA is their 'mother tongue' or 'native' language and never the colloquial dialect they speak (Maamouri, 1998). However, researchers in diglossia claim that it is the colloquial Arabic varieties that

---

[2] Detailed discussion of the difference in word order will come in chapters 2 and 3.

form the true native languages of the people across the Arabic-speaking world (e.g., Ferguson, 1959; Kaye, 1970; Zughoul, 1980). This is obviously due to the fact that these varieties of Arabic are what people speak at home and in everyday situations and to which children are extensively exposed from birth and, hence, acquired naturally.

Indeed, it is often claimed that there are no native speakers of MSA (e.g., Kaye, 1970; Maamouri, 1998). This is apparently because MSA is acquired through literacy and formal schooling which does not start for most children until age 5 or 6 when the children's innate ability to acquire languages natively may no longer be available or may have started to deteriorate (Al-Dannan, 2010). Another factor that may support such a claim is the fact that children who start acquiring MSA at school already speak their colloquial varieties which were acquired naturally from birth. Because there are differences between the colloquial varieties and MSA in all linguistic domains (e.g., Altoma, 1969; Ayari, 1996; Maamouri, 1998; Khamis-Dakwar et al., 2012), this may affect the ultimate attainment of MSA and may result in different end-state grammars; a characteristic of a second language (L2).

To this last point, Kaye (1970) describes performance in MSA by Arabic speakers as always fraught with inconsistency and showing characteristics of the colloquial variety.[3] To him, this is an outcome of the fact that MSA is not learned natively in a natural setting like home and, hence, the linguistic system of MSA is 'ill-defined' consisting of many variant versions that all can be lumped under the designation 'MSA'. Thus, due to

---

[3] Kaye (1970) did not back up his claim of inconsistency in MSA among speakers of Arabic with empirical findings of his own or any others'. However, he used many examples of MSA lexicon and pointed out that they would be pronounced differently by speakers of different colloquial varieties when speaking MSA due to using phonological features of the colloquial variety that they acquired natively.

the ill-defined nature of MSA, Kaye (1970) argued that consistency in speaking this variety cannot be expected, be it across coherent groups of speakers who share similar levels of education and speak the same colloquial variety or within individual speakers of MSA.

Since MSA is not a spoken variety that can be used in everyday informal settings, exposure to MSA at home, if there is any, is very poor and usually comes from sources like certain TV programmes and cartoon films, not from the family (Maamouri, 1998). Ayari (1996) pointed out that the fact that children reach the age of schooling without being exposed to MSA is very common in the Arab world. Iraqi (1990; cited in Ayari 1996 & in Feitelson et al, 1993) reported in a study conducted on Palestinian kindergarten children that only 1.8% of 290 surveyed families said that they buy books and read from them in MSA to their children before they reach the age of schooling. Also, even when families read to their children from books, they often translate the standard form used in books to the colloquial form for their children (Doake, 1989). The phenomenon of not reading to children in MSA before the age of schooling was justified by parents in Iraqi's study (1990) in terms of MSA being too difficult and complex for their children to understand and enjoy at this age.

Some Arab children go to nurseries and kindergartens, between the age of 2 to 5 years, where exposure to MSA should be expected, given that these are educational settings; but the number of these does not exceed 15% of the total number of children in the Arab

world (Al-Dannan, 2010).[4] In addition and contrary to what is expected in these institutions, communication and activities like singing, describing pictures, narrating stories, etc., are all carried out in the colloquial form, not in MSA (Al-Dannan, 2010). In fact, Feitelson et al. (1993) reported that the surveyed kindergarten teachers in their study, for example, strongly believed that kindergarten children are too young to be exposed to MSA and that children would not be able to understand the content of the stories if read in this form of the language.

Arab children, then, start their 'real' exposure to MSA when they start going to primary schools at age 5 or 6 where MSA is the language of writing and literacy. The grammar of MSA is learned at school in terms of explicit explanation of its rules, and there are usually no frequent chances to practice speaking MSA even at school (Al-Dannan, 2010). According to him, using MSA at schools in the Arab world is usually limited to writing and reading from books. The communication language in schools in the Arabic-speaking countries is the local colloquial form of Arabic; the pupils use this form to talk to each other, and to talk to or discuss the material with the teacher. The pupils use MSA as spoken only when reading from books or when asked to answer a question in class. The teachers, on the other hand, communicate with pupils and usually explain the lessons in the colloquial form of the language as well. Teachers usually do not use MSA at school unless when reading from books or writing on the class board (Al-Dannan, 2010:46). Based on the above description of means of exposure to MSA, it can be concluded that the main source of MSA input for most Arab children is the written form and this does not start until they learn literacy from age 5 or 6.

---

[4] The percentage of children who go to nurseries and kindergartens mentioned here seems to be based on Al-Dannan's impression as he did not report on a survey-based study.

The diglossic situation in the Arab world has been described as a problem that needs to be resolved, especially in the field of education (e.g., Ayari, 1996; Maamouri, 1998; Al-Dannan, 2010). Ayari (1996:243) states that the diglossia is one main factor behind the noted spread of illiteracy in the Arabic-speaking countries. Al-Dannan (2010) explained that an Arab child starts schooling with a common shock that the language of literacy is different from the language that he or she previously acquired and mastered at home. This leaves the child to face a double challenge of acquiring literacy and the language of literacy at the same time. According to Al-Dannan, this challenge usually leads to a negative attitude towards books from the start, and this in turn results in weak skills in reading and writing for many Arab children. The problem becomes more complicated by the fact that children do not acquire MSA naturally by being exposed to spoken MSA, but they learn this form via extensive sessions of explaining its rules without specifying sufficient time to practice speaking it in school or anywhere else. Ayari (1996:245) indicated that many literate Arabs who managed to satisfy academic requirements of reading and writing in MSA well enough confess that they do not feel confident or skilled when reading or writing in this form of the language.

A number of proposals have been offered and discussed in the literature to alleviate the problem of diglossia in the field of education. One of which, for example, is a call to use the local colloquial variety as the language of literacy and instruction in each Arabic-speaking country. Ayari (1996) maintained that this call has been strongly resisted by the common belief that the colloquial forms of Arabic are in themselves a result of illiteracy and ignorance and are not equipped with the expressive power that is

necessary for knowledge acquisition. Another argument used by the opponents of this call is the fact that this decision will eventually deprive future generations from access to the large quantity of works written in MSA over the past centuries (Zughoul, 1980). In addition, such an act would be against the political policies of many Arabic-speaking countries which seek to enhance unity among their nations.

Another proposal that considers the need for MSA to stay as the language of instruction was postulated by Al-Dannan (2010). Al-Dannan suggested to apply what is called an 'MSA Immersion Program' in preschools to provide natural MSA input to children before they start their primary school education and during the period when their innate ability to acquire languages is still active. This proposal will be presented separately in the following section due to its close relevance to the focus of this thesis.

4 MSA Immersion Program (Al-Dannan, 2010)

In order to alleviate the negative effect of diglossia on the acquisition of literacy in the Arab world, Al-Dannan (2010) proposed that Arab children need to acquire MSA natively, through immersion in an MSA environment, by investing their innate ability before it starts to deteriorate after age of 6.[5] To be able to do this, Arab children need to be exposed to spoken MSA and practice speaking it during the pre-school period in a distinct place where MSA is the only medium of communication or with a distinct person who speaks only MSA to them all the time. This way, Arab children can grow into bilinguals who speak both the local colloquial variety and MSA natively. According to Al-Dannan, children who acquire MSA naturally and natively are not

---

[5] Al-Dannan (2010) assumes that the innate ability of children to acquire languages is active from birth till age 6 and starts to deteriorate afterwards till age 10 when it ceases to function. Detailed discussion of the critical period for acquiring language will come in the following chapter.

expected to have difficulties when they start learning literacy at primary school as they will be prepared for acquiring knowledge in MSA, which they are already native speakers of.

Al-Dannan's proposal was based on a conclusion of a long-term experiment he conducted on his son, Basil, in 1977. Al-Dannan in this experiment talked to everybody in his house in the local colloquial form except for Basil to whom he talked only in MSA from age 4 months.[6] However, Al-Dannan asked everybody in the house to talk to Basil in the local colloquial form of Arabic and never try to mix it with MSA. By doing this, Al-Dannan was trying to make a clear distinction for Basil that the two varieties are not the same and each can only be used with specific persons.

Al-Dannan stood by his plan till Basil became an adult and grew up a bilingual. Eventually, Basil became able to speak to the members of his family in the colloquial form and to his father in a perfect MSA without the need for explicit instruction on its formal rules. Al-Dannan noted that his son did not encounter difficulties when he started learning literacy in the primary school. In fact, Al-Dannan indicated that Basil's skill of reading and his ability of comprehending MSA were remarkable;[7] he managed to read more than 360 children's books in one year when he was in the second grade. This long-term experiment was replicated in 1981 by Al-Dannan himself with his

---

[6] Al-Dannan commented that although he was using the local colloquial form with everybody else in the house, this did not affect Basil's acquisition of MSA who kept using MSA when speaking to his father. See Al-Dannan (2010) for further details on this experiment.

[7] While the statement that Basil could develop a 'remarkable' skill of reading and comprehension in MSA due to acquiring MSA before primary school is interesting, this observation of Al-Dannan about his own child cannot be considered scientific as a study of a parent on his own children is inevitably biased. However, Al-Dannan's description of Basil's skill of reading and comprehension in MSA is supported by empirical findings of Jenkins (2001) who tested this skill among children who acquired MSA via immersion during pre-school period. Details of this study will be presented later in this section.

daughter, Luna. The outcome of this experiment was similar to that of the previous experiment with Basil.

Based on the results of these two 'case-studies', Al-Dannan decided to found nurseries and kindergartens in which MSA is the only form of Arabic permitted for teaching and communication inside school. This came to reality when Al-Dannan launched *dar lHaDanah ?l9arabiyyah* 'the Arabic Nursery House' in 1988 in Kuwait. For this project and before starting to accept children into the nursery, Al-Dannan first trained the nursery teachers for 3 months to speak MSA fluently with no code-switching to the local colloquial form. Then, the teachers were asked to use MSA strictly and exclusively when speaking to children and even when speaking to each other inside the nursery. When a child makes a request to the teacher in the colloquial form, the teachers were asked to reply, repeating what the child has said in MSA with a question tone. The teachers were also asked to narrate stories in MSA without translating the meaning to the local colloquial form, and use pictures, gestures, etc, to help children understand the meaning of the story. Whenever a child utters a word or a sentence in the colloquial form, the teachers were asked to repeat what is said in MSA to provide positive input. When a child had spent about two months in the nursery, the teachers were asked to pretend that they did not understand if talked to in the colloquial form, but when a child started to try to speak in MSA, the teachers were asked to help at the beginning and encourage continuing speech in MSA. Families at home were told that it was not necessary to speak to their children in MSA at home, and if they were spoken to in MSA by the child, they should respond to their requests or reply in the colloquial form.

Al-Dannan (2010) reported that most children accepted into this nursery started to respond with MSA in a month's time, and they started speaking MSA preserving the vocalic case markings at the end of each word when they had spent two months in the nursery with occasional use of colloquial vocabulary. After four months in the nursery, Al-Dannan added, all children were taking the initiative to speak in MSA with their teachers without being asked to do so.

Al-Dannan described the outcome of applying the MSA immersion program in this nursery as successful in bringing up a different generation of Arab children who speak MSA natively and thus survive the negative effect of diglossia when acquiring literacy and knowledge at primary school. This nursery had to close, unfortunately, in 1990 upon the Iraqi invasion of Kuwait. However, Al-Dannan founded another kindergarten in 1992 that applies this immersion program in Syria and called it _rawDat l?azhaar ?al9arabiyyah_ 'The Arabic Flowers Kindergarten'. Later on, a number of nurseries and kindergartens across the Arab world followed the example of these two pre-schools and applied the MSA immersion program in their educational institutions. These include, for example, _?almadaaris ?al9aSriyyah_ 'The Contemporary Schools' in Jordan, _HaDaanatii_ 'My Nursery' in United Arab Emirates, _madaaris ?albassaam_ 'Albassam Schools' in Saudi Arabia, among many others across the Arab world.

Two academic studies have been conducted to evaluate how helpful this MSA immersion program to the students when they go to the primary school. One is a masters dissertation (Jenkins, 2001) to test the impact of the preschool MSA immersion program on the student's skills of reading and composition when they are at the primary

school. Part of this dissertation is a quantitative study that checked score records for 2797 students in Harasta primary schools in Syria. The aim was to compare two groups of students in terms of the Arabic language GPA, reading, composition, and recitation scores for students in grades 1 to 6. The two sides of comparison were students who previously attended the MSA immersion program in the Arabic Flowers Kindergarten in Harasta, as one group, and students who either attended traditional preschools or did not attend any preschool program, as the second group. The results indicated that there was a statistical significant difference between the two groups in favor of the MSA immersion program students. Jenkins concluded that attending the Arabic Flowers Kindergarten which applies the MSA immersion program has a significant effect on Arab children's scores in reading and composition in Harasta primary schools.

The second academic work is also a masters dissertation (Alomari, 2009). Alomari aimed to compare two groups of students in 3rd grade in terms of their achievement in subjects of Arabic Language and Mathematics. One group included students of Albassam schools which applies MSA immersion program in its preschool levels and the second group consisted of students of another traditional primary school that does not apply the MSA immersion program. The subjects of Arabic Language and Mathematics were taught in MSA for the students in the former group and in the local colloquial form for the students in the latter group (the books and materials used for both groups used MSA as the literacy language). The researcher conducted two post teaching tests; one right after the end of the course to measure the students' achievement in both subjects, and another delayed post test after two weeks to measure whether the students were able to retain what they have learned from the teaching sessions. The results of this study indicated that the students of Albassam Schools significantly

outperformed the students in the traditional school in both post tests in terms of their achievement in both subjects. Therefore, the researcher concluded that teaching in MSA (the language of the books) for students who are able to use it and understand it (those who acquired it naturally via MSA immersion) has a positive effect on school achievement and acquisition of knowledge.

5 Summary & Organization of the Thesis

This chapter has outlined the research focus of the present thesis, laid out the research questions, and stated the hypothesis. Also, it presented a general description of the diglossic situation in the Arab world and introduced and described the MSA immersion program. This program was proposed as a possible solution to the problem of Arabic diglossia as it provides natural MSA input to preschool children so that they grow up native speakers of this form of the language as well as the colloquial form. Graduates of preschools that apply MSA immersion program together with typical speakers of Arabic who start being exposed to MSA from age 6 in primary schools will compose the two main groups of participants in the MSA experimental study.

The rest of the thesis is organised as follows. The next chapter provides background about relevant issues of language acquisition and the syntactic variables of the study. Chapter three presents two corpus based studies: one is based on a set of corpora of colloquial Arabic varieties and the other is based on MSA corpora obtained from different regional sources. The purpose of conducting the colloquial Arabic study is to pinpoint syntactic differences between the colloquial varieties that are suitable to be taken as variables for the main experimental study of MSA. The MSA corpus study was

conducted as a preliminary test to determine if the differences found between the colloquial varieties also exist to differentiate between the grammatical representations of MSA across the regions from which the MSA data obtained.

Chapter four describes the methodology of the main experimental study in this thesis. Chapter five lays out the results of the study, which are going to be discussed in chapter six. Chapter seven will be the conclusion of the thesis.

Chapter 2

Background

This research aims to examine the end-state grammar of Modern Standard Arabic (MSA) acquired by Arabs from different parts of the Arab world. The main objective is to determine whether these learners represent MSA natively or as a second language. If MSA is represented as an L1, then all the speakers are expected to have uniform representations of MSA grammar. However, if MSA is acquired as an L2, there should be differences concerning how MSA grammar is represented by variant speakers. If MSA speakers are found to be representing an end-state grammar of the L2-type, this is thought to be probably an effect of late exposure to MSA input, L1 influence, or both.

This chapter provides background information for the present work. The first section reviews the literature in relation to language acquisition issues. These include (i) how first language end-state grammar is different from second language grammar; (ii) the Critical Period Hypothesis and its effects on both L1 and L2 acquisition; and (iii) the impact of the first language on second language acquisition. The second section presents background information about the syntactic aspects under investigation in MSA. These include word order and subject-verb agreement, adverb placement, and resumption in object relatives.

1 Language Acquisition

The assumption that acquiring a first language (L1) is remarkably different from learning a second language (L2) is well established in the literature. To start with first language acquisition, children acquire their native languages in an amazingly predictable manner. With the exception of some types of neural pathology or impairment, all children start their acquisition process with no lexicon, pragmatic competence, phonemic inventory or overt grammatical knowledge, and finish it with virtual mastery of these domains within a few short years (Chomsky, 1965).

Findings of research about first language development reveal some generalizations that apply to acquisition of different languages. One such generalization is that children acquiring their native language, regardless of what language they acquire, which culture they come from, or what linguistic environment they live in, seem to have similar rates of acquisition, and start producing language and complete the learning process approximately at the same age (e.g., Slobin, 1982). Another generalisation is that L1 learners with different cognitive abilities appear to produce common errors and sequences, and achieve parallel levels of attainment (e.g., Guasti, 2002). Also, children manage to develop the bulk of their L1 grammar by age 4-6 (Karmiloff-Smith, 1986; Gleitman and Newport, 1995; Guasti, 2002; Hyltenstam and Abrahamsson, 2003; Schwartz, 2004; Unsworth, 2004).

Chomsky (1965) asserts that the linguistic knowledge that L1 learners eventually acquire cannot be the result of exposure to linguistic data alone; this data to which

children are exposed is of deficient quality and fraught with performance errors and inconsistencies.[8] The fact that children succeed in learning more than they have experienced constitutes what is known as 'the logical problem of language acquisition'. A proposed solution to this problem states that L1 learners must have an innate language faculty which guides them to arrive at a perfect and uniform grammar:

> A consideration of the character of the grammar that is acquired, the degenerate quality and narrowly limited extent of the available data, the striking uniformity of the resulting grammars, and their independence of intelligence, motivation and emotional state, over wide ranges of variation, leave little hope that much of the structure of language can be learned by an organism initially uninformed as to its general character.

(Chomsky, 1965:58)

Chomsky's statement informs that something must aid L1 acquisition. This refers to what became known later in the literature as Universal Grammar (UG). The theory of Universal Grammar holds that all human languages are constrained by a set of innate invariable universal principles as well as a specified number of varying parameters to allow for the growth of different languages (Chomsky, 1981; 1988).

White (2003) explains the role of UG in first language acquisition by saying that it provides the innate knowledge which L1 learners need along with exposure to linguistic input. This knowledge is important to help L1 learners determine the precise form of the grammar they will eventually attain. Based on the input available in their linguistic environment, children build up a language-specific lexicon and set the UG parameters to values which represent the grammatical properties of the

---

[8] See, e.g., Hornstein and Lightfoot (1981) for discussion of L1 input deficiency.

language in question; then, finally, arrive at a steady state grammar for their native language.

There are two obvious differences between first language and second language acquisition mentioned in the literature. The first is that the initial state in L2 acquisition is not the same as in L1 acquisition, given that L2 learners already know at least one other language (Schwartz and Sprouse, 1994; Schwartz and Eubank, 1996). The second is that children acquiring their L1 reach perfect mastery of whatever language they are exposed to, whereas adult learners display varying degrees of imperfection despite long periods of exposure to the second language (Bley-Vroman, 1989; Sorace, 2003). Indeed, Sorace (1993), among others, argues that even native-like performance in the second language does not necessarily reflect native-like competence. Thus, L1 and L2 learners of the same language appear to go through different initial states and end up also with different final state grammars.

Bley-Vroman (1990) declares that, in comparison with L1 acquisition, second language acquisition is marked with incompleteness, variation of success, and no equal potentiality for the learners. Also, age of onset in second language acquisition seems to have effects on how successful the process will be; younger learners seem to be more successful language learners in terms of ultimate attainment (e.g., Patkowski, 1980; Bley-Vroman, 1989; Hyltenstam, 1992; DeKeyser, 2000; Hyltenstam and Abrahamsson, 2003; DeKeyser, 2010).

There are many approaches toward explaining the differences between first language and second language acquisition. One line of argument attributes the differences to the claim that adult second language learners have no access (e.g., Bley-Vroman, 1989), or have only partial access to Universal Grammar (e.g., Schachter, 1989; Schachter, 1996), probably due to maturational reasons. Another view attributes these differences to the influence of the first language (e.g., Schwartz and Sprouse, 1996). Both of these accounts are considered later in this thesis on the context of the experimental study findings. Therefore, the following two sections will review the literature on these two accounts: age of first exposure and L1 influence.

1.1 The Critical Period Hypothesis (CPH)

The critical period, generally, is a time of life during which some specific experience affects the development of an organism more than at other times (Colombo, 1982). The notion of the critical period can be traced back to studies in embryology and ethology to account for some kind of innate development and other instinctive behaviours (Oyama, 1979). Penfield & Roberts (1959) and Lenneberg (1967) were probably the first to propose that there is a critical period for language acquisition. These scholars, among others, claim that the first few years of life constitute the crucial time in which an individual can acquire a first language if presented with adequate stimuli; if language input does not occur until after this time, the individual will never achieve a full command of language.

Penfield & Roberts (1959) concluded from their observation of aphasics' language behaviours that there is a biological constraint on language learning. They noted that children who experienced damage to the speech areas of the brain before puberty were

able to recover and redevelop normal language; but those who had the same experience after puberty failed to attain a full recovery. The researchers maintained that the younger patients' recovery was due to their brains' neurological plasticity which made transfer of the speech functions from the damaged hemisphere to the undamaged one possible; the human brain, they argued, becomes progressively more stiff and rigid after the age of nine.

The notion of the critical period, however, became popular in the field of language acquisition after Lenneberg's work, *Biological Foundations of Language*, in 1967. Lenneberg states that language emerges between the ages of two and three by an interaction of maturation and self-programmed learning. He proposes that the chance for first language acquisition becomes good between the ages of three and the early teens; during this time, the individual appears to be most sensitive to stimuli and to preserve some innate flexibility of the brain to organize its functions and carry out the complex integration of sub-processes necessary for the smooth elaboration of speech and language. After puberty, the brain's ability for self-organization declines quickly and its lateralization process becomes complete; the primary and basic linguistic skills not acquired by that time usually remain deficient for life (1967:158). This conceptualization of the critical period has been considered as the original formulation of the Critical Period Hypothesis (CPH) in language acquisition.

1.1.1 The Critical Period in First Language Acquisition (L1A)

The Critical Period Hypothesis has been one of the most fiercely debated issues over the past few decades. In the field of first language acquisition, several empirical studies

yielded support to the hypothesized existence of a critical period for the acquisition process.

There are several cases reported to show that if the first exposure to linguistic input was later than the offset of the critical period, the ultimate attainment will be incomplete and deficient. One very famous case that supports this conclusion is the tragic case of Genie. Curtiss (1977) has reported that Genie was deprived from any linguistic input by almost total isolation from when she was about 20 months old until she was discovered at the age of 13;7. At the time she was discovered, Genie was not able to speak or understand a language. Despite the fact that Genie received intensive language teaching for almost seven years right after she was found, she was not able to develop a normal native-like attainment. Genie was able to learn much of the lexicon and semantic notions, but she was less successful in learning morpho-syntactic structures. Curtiss states that Genie's case can be taken to indicate that first language learning is possible after puberty, but that learning will be irregular and incomplete.

Another case of isolation was reported by Mason (1942). A girl named Isabelle was imprisoned with her mute uneducated mother until age 6½. Isabelle, unlike Genie, was able to develop a native linguistic competence within 18 months following her release. Similarly, Jim is a son of deaf non-signing parents who brought him up in a fairly isolated rural area. Television was almost the only source of linguistic input for Jim until he was found at age 3½ to be studied by Sachs, Bard & Johnson (1981). The researchers reported that Jim had succeeded in learning about 50 words just from watching TV. Following his exposure to normal conversations, Jim was able to show

rapid development in his language. The case of Genie in contrast to the cases of Isabelle and Jim has been taken to indicate the validity of proposing a critical period for primary language development. Jim and Isabelle have gained proper exposure to the language before the proposed offset of the critical period, and hence, were able to develop native competence of their languages. Genie, on the other hand, unfortunately was deprived of exposure to language and linguistic interaction until after the offset of the critical period, hence, failed to develop a perfect and complete competence of the language.

One could say, however, that cases of children like Genie who experienced traumas or abusive treatments cannot be taken solely to indicate the existence of the critical period; failure of children like these to develop a language later in life may be due to some reason (probably psychological) related to the difficult experience they have been through, rather than being specifically a result of the lack of exposure to language during early age. More clear evidence of incomplete language acquisition resulting from late exposure to linguistic input comes from the work of Newport (1984) and Newport & Supalla (1987). They conducted a large-scale project on congenitally deaf individuals who were fluent users of American Sign Language (ASL). All had been exposed to this language as their first language for equivalent periods of 10 years. The subjects were divided into three groups depending on age of onset of language learning. The performance of those who started being exposed to ASL between birth and 6 years was native-like. The performance of the second group who were exposed to the language from age 7 onwards was fraught with some errors in grammatical morphology. Finally, the ultimate attainment of the third group, who were exposed to input after the age of 12, was clearly non-native and showed some individual variation; their use of grammatical morphemes was described as sporadic, irregular and often incorrect.

Similarly, Newport (1990) analysed the performance of 30 users of ASL in 7 tests related to ASL morphology. This study found that participants who were exposed to ASL since birth outperformed those whose first exposure to ASL as a primary language was between age 4 and 6. These two groups of participants, in turn, did better and more systematically than participants who were exposed to ASL for the first time at age 12 or after. Also, this study found a correlation between the test scores and the age of first exposure to ASL. Newport concluded that these results constitute strong evidence for age of acquisition effects on ultimate attainment in L1 acquisition: "the later the language is learned, the less its use is native ... in character" (1990:18).

1.1.2 The Critical Period in Second Language Acquisition (L2A)

Researchers of second language acquisition have a dispute about i) whether or not the critical period applies in L2 acquisition as well; and if it does, ii) whether or not it is the same notion that applies to L1 acquisition, i.e., resulting from maturational reasons. The source of these queries is probably due to some conflicting research findings. Some findings seem to show that there are no age-related effects on L2 acquisition; adults, like L1 learners, are able to succeed in acquiring a native-like competence. Other findings claim that this is not possible after the offset of the critical period. Also, some researchers concluded that age effects exist in second language acquisition, but they are different from those noted for L1 acquisition. Example studies that represent these findings will be presented in the remaining part of this section.

A number of studies have suggested that second language learners differ in their attainment depending on their age when they started the learning journey. These studies attributed the different levels of attainment to the existence of a critical period in second language acquisition. In studying acquisition of L2 phonology, Oyama (1976), for example, examined the English accent of 60 Italian subjects whose ages varied from 6-20 when they first arrived to the USA. All the subjects having been residents in the English speaking country for no less than 5 years at the time of the study, they were considered to be at the final stage of acquisition. Oyama found that the results of judging those who arrived before the age of 12 for the degree of accent in English were in the range of the results achieved by the native controls; but those who arrived after the age of 12 had detectable foreign accents.

Similar findings were reported by Seliger, Krashen, and Ladefoged (1975). They divided their subjects (394 immigrants to the USA) into three groups depending on their age of arrival; at the age of 9 or younger, between 10 and 15, or at the age of 16 or older. The results showed that participants who had arrived younger than 9 had a native-like accent; those who arrived after the age of 16 had clear foreign accents; and a variation in performance was noted for the rest.

Further support for the validity of proposing a critical period for second language acquisition comes from studies designed to examine L2 syntactic knowledge. Patkowski (1980), for example, conducted a study in which 67 non-native speakers of English and a number of native controls were interviewed for the purpose of assessing their English syntactic knowledge. Trained judges were asked to evaluate written transcripts of the

participants' interviews. The results showed that participants who started learning English after puberty were given low scores, whereas both native speakers and younger non-natives achieved high scores.

In another attempt to examine L2 learners' syntactic knowledge in relation to age of onset, Johnson and Newport (1989) have conducted a study which involved 46 Korean and Chinese speakers of L2 English, whose age of arrival to the United States was between 3 and 39. The subjects were asked to judge the grammaticality of 276 spoken English sentences. The researchers found a negative linear correlation between age and performance up to the age of puberty; the performance of those who arrived after puberty, however, was characterized as low and highly variable. Johnson (1992) has repeated the same study with all its procedures and materials preserved. One difference, though, was that the subjects were asked to make judgments on written sentences this time. Johnson's study revealed similar but weaker negative correlation between age of arrival and performance.

For the same purpose of examining age effects on L2 syntactic attainment, Shim (1993) tested the judgments of Korean speakers of L2 English on sentences which involved issues related to Subjacency and Anaphoric Binding principles. A number of native speakers of English were asked to do the same tasks as controls for the study. The results showed that performance of those who were exposed to English before the age of 5 was not very different from that of native controls. Also, the results showed some age-related effects on the performance of those who began learning English between the age

of 6 and 11. The performance of those who started later than age 11 was found not consistent.

As of the L2 studies outlined so far suggest that a critical period applies in L2 acquisition, however, this proposal has been empirically challenged. A number of studies have found that native-like proficiency is apparently attainable even when L2 acquisition begins later than age 12, in other words, after the proposed closure of any critical period. To begin with research on phonology, Bongaerts et al. (1997) have conducted a study which involved adult Dutch learners of English to assess their accents. The subjects were divided into two groups of non-natives; those with excellent command of English and others with various proficiency levels, plus one group of native controls. Thirteen native speakers of English were asked to rate these subjects' accents, based on samples of their reading of English sentences. In spite of the fact that all of the non-native participants had started being exposed to English after the age of 12, the judges could not differentiate between at least some of the highly proficient learners' and the native speakers' accents.

A study conducted by Ioup et al. (1994) also appears to controvert the idea of a critical period in second language acquisition. The researchers tape-recorded spontaneous speech of two English speakers of L2 Egyptian Arabic, and played samples of the speech to 13 native teachers of Arabic as a foreign language. Both participants acquired Arabic as adults; yet, 6 of the judges rated them as native speakers, and another 2 judges rated one of them as a native speaker.

Further cases with native-like performance continue to be found by studies conducted to assess L2 syntactic knowledge. White and Genesee (1996), for example, asked their French participants, who started learning English after 12, to judge 60 English sentences with wh-movement structures. The researchers compared the non-native subject's performance with that of native speakers of English and found no significant differences. Cranshaw (1997), as another example, reported similar results. In this study, 20 L1 French and 20 L1 Chinese speakers of L2 English were asked to do grammaticality judgment tasks on sentences with issues related to English tense-aspect features. Despite the fact that all the participants in the study started learning English after 12, Cranshaw found that the judgments of a significant number of participants were within the range of those given by the native controls.

In general, nonetheless, these studies and the others of the like (e.g., Neufeld, 1977; 1979) have been criticised as not sufficient for refuting the idea of a critical period in second language acquisition. Hyltenstam and Abrahamsson (2003), for example, argue that the cases of successful late starters of L2 in attaining native-like accents or knowledge are very limited in number, and cannot be considered as counterevidence for critical period existence in second language acquisition. In addition, they maintain that to count learners as 'native-like', they should behave in a native-like manner in all the domains of the language, not only in one particular domain like phonology. Thus, further research is needed to ascertain the validity of such claims.

Another challenging argument to the critical period supporting studies, though, is that many of the results indicate a general linear negative correlation between the onset age and L2 proficiency; there appears to be no identifiable offset of this correlation to point to the critical period terminus point. However, this argument cannot be sustained when studies like Johnson and Newport's (1989) are considered. This particular study has identified a cut-off point at which the correlation between onset age and decline of proficiency comes to an end; when the onset age is about 17 or older, they maintained, the distribution of performance was essentially random.

More recent studies also claimed that there is an even earlier age before which acquisition must begin if ultimate attainment is not to be affected. Meisel (2009), for example, supports the claim that there are several sensitive periods during early development and argued that some of the optimal phases during which learners can be successful in learning and developing grammatical properties by being merely exposed to linguistic input start to waste away as early as around age 4 (2009:7). Meisel (2008) analysed recorded data of 10 German participants whose first exposure to French was between age 2;8 and 4. Meisel found that participants who were exposed to French at or before age 3;6 used L1 type constructions whereas data of participants who were exposed to French at age 3;7 or after were fraught with adult L2 type characteristics. Furthermore, Kroffke et al. (2007) studied acquisition of German subject-verb agreement and verb placement by two Turkish children who started acquisition at age 3 and age 6. The researchers found that the boy who started acquiring German at age 6 resembled adult L2 developmental patterns whereas the other participant who started acquisition at 3 was found developing these two structures in the same way as L1 children. It was concluded that if exposure to the second language in a successive

bilingual acquisition happens at or after age 6, this can be characterised as child L2 acquisition rather than acquiring two L1s.

In contrast, several other studies have found that for both early and late starting L2 learners, earlier is better across the lifespan. Birdsong and Molis (2001) did a replication of Johnson and Newport's (1989) study. They used the same methods, materials, and tasks, but more participants (61) who speak Spanish as their L1 (AoA range = 3-44 years old). In contrast to the findings of Johnson and Newport, the distribution of performance of those arrived at the age of 17 or older was not random this time; they found a significant correlation between age and performance among the 32 participants who arrived at age 17 or more. Similar findings were reported by Stevens (1999), for instance. Bialystok and Hakuta's study (1999) also reached similar findings and the researchers concluded that factors other than an innate biological program were behind the performance decline.

Birdsong (1999) maintained, however, that results, such as these, which challenge the predictions of having a critical period in second language acquisition should not be taken to deny maturational factors all together; it might be possible to explain the age-related effects on L2 performance over the lifespan as being "the product of different causal mechanisms along the way, that is, the result of developmental factors up to the end of maturation, and of non-developmental factors thereafter" (Birdsong, 1999:12).

Many of the critical period proponents have also suggested that there may be interacting factors with maturation that cause age-related differences. Moyer (2004), for example, claimed that an integration of neurobiological, cognitive, and social-psychological perspectives would yield a better explanation for 'younger is better' in second language learning. Hyltenstam and Abrahamasson (2003), as another example, posited that maturation can account for the negative correlation between age and language proficiency, but it cannot on its own account for the variability between exceptionally successful and non-exceptional L2 learners of the same starting age; there might be non-maturational factors in the picture.

Other accounts of different types also have been proposed in the literature. One such account takes maturation of cognitive abilities as a reason for 'younger is better' in language attainment. This has been proposed by Newport (1991) who argues that children have the advantage of having a low capacity short-term memory when acquiring the language. When children start their acquisition process, this type of memory does not allow them to extract and hold more than a few morphemes from the input. Working within these processing limits helps children to focus on analysing these small portions of data successively. In contrast, due to their greater memory capacity which allows for extracting more of the input, adults are faced with larger strings of data right from the outset.

Meier (1995) states that as language learners mature, their increasingly sophisticated cognitive abilities hinder the function of the language acquisition capacity they have. The roots of Meier's proposal can be found in the work of Felix (1985). Felix proposed

that adults' decline in language learning potential is due to the development of advanced general cognitive system. This mature cognitive system competes with the still functioning language acquisition system, and usually wins the task of analysing the language data. Because this general cognitive system is ill-suited for such a specific task, it entails the lack of success associated with adults' language learning.

Other critical period opponents have also made attempts to account for the correlation between age and language attainment. These accounts include, for example, social or psychological factors (e.g., Schumann, 1978), or Input factors (e.g., Hatch, 1977).

To summarize section 1.1 of this chapter, existence of a critical period for acquiring the first language is evident from the results of studying cases of L1 development in abnormal circumstances (i.e., Genie compared with Isabelle and Jim). The work of Newport (1984), Newport & Supalla (1987), and Newport (1990) indicates that native competence in L1 acquisition of a sign language is guaranteed to those who were exposed to the language before the age of 4; a decline in ability to perform in a native manner was noted for those who were exposed to the language after 7 and a non-native performance was guaranteed for those who started after 12.

Most of the studies on second language acquisition support the claim that age of first exposure to the acquired language has effects on how complete the outcome grammar will be. These studies, however, can be divided into four groups depending on their findings of whether or not there is a cut-off point before which native-like attainment is

possible or guaranteed. Studies in the first group (e.g., Oyama, 1979; Patkowski, 1980) suggested that native-like attainment is possible for those who started learning any time during the critical period. The second group includes studies like Selinger et al. (1975), Shim (1993), and Meisel (2008) which reported that native-like attainment was only possible before a certain age during the critical period (9, 5, and 3;7 respectively), and a decline or variation of performance was noted thereafter until the end of the critical period. The third group of studies did not specify an age before which native-like attainment is possible or guaranteed; but they suggested that there is a general negative correlation between age and attainment that ends at puberty pointing to the end of the critical period (e.g., Johnson & Newport, 1989; Johnson, 1992). Finally, the fourth group consists of studies like Birdsong and Molis (2001), Stevens (1999) and Bialystok & Hakuta (1999) which found that the rule of 'younger is better' is true across the life span.

As mentioned before, the main study of this thesis will compare between participants whose first real exposure to MSA was not before the age of entrance to formal primary schools (5-6 years old) with those who were exceptionally exposed to MSA from as early as the age of two. The findings of the studies which support age-related effects have different predictions for the current study. Because all the subjects have been exposed to MSA before puberty, studies in groups 3 and 4 have the same prediction: there will be individual differences in performance and attainment among all the subjects depending on age of onset.

Studies in group 2 predict that at least those who were exposed to MSA before primary school will have native-like attainment, but those who started learning MSA at primary school may or may not reach such a level of attainment: Shim (1993), for example, showed some age-related effects (younger is better) on the performance of those whose age of first exposure was between 6 and 11. Meisel (2008), also, reported that native-like attainment was possible only for those who started acquisition before age 3;7.

Studies in group 1 predict that all the subjects under study have the opportunity to develop a native-like competence of MSA if exposed to an adequate amount of input. This last prediction is also supported by the findings of another group of studies which affirmed that native-like proficiency is attainable at least by some learners even after the end of the critical period (e.g., White & Genesee, 1996; Cranshaw, 1997).

1.2 L1 Influence

The issue of L1 influence on the grammar of a second language has been the topic of much research throughout the past few decades. L1 influence has been studied from a variety of perspectives and within a variety of approaches, and it has undergone significant reconceptualization over the years (for an overview, see, e.g., Odlin, 1989; 2003; Ellis, 2008). Lado (1957), for example, formed the Contrastive Analysis Hypothesis which considered L1 influence as playing a negative role in second language learning. This hypothesis was formulated based on an area of inquiry prevailing in 1960s and early 70s which aimed to describe differences and similarities across languages. The Contrastive Analysis Hypothesis claimed that difficulties in second language learning derive from the differences between the new language and the learner's first language; errors in these areas of difference derive from first language

interference. By doing a contrastive analysis, i.e. identifying the similarities and differences between the L1 and the target L2, these errors can be explained and even predicted. However, the results of some studies conducted thereafter indicate that this claim cannot be sustained (see, e.g., Jackson and Whitnam, 1971; Dulay and Burt, 1974).

The issue of L1 influence has also been examined within the generative framework, particularly the L1 influence-UG access relationship and its determining role of shaping learners' mental representations at different stages of L2 acquisition (see White, 2000). Although the focus of the present investigation is the end-state grammar, it is necessary to understand first the claims about whether or not L1 grammar is transferred to L2 acquisition, as these claims make different predictions regarding the final outcome of L2 acquisition.

According to White (2000), L2 researchers have come with five different theoretical perspectives on the nature of L2 initial state grammar and the availability of UG. First, L2 initial state grammar is the L1 grammar and L2 learners have no or partial access to UG; only those UG principles that have been activated in first language acquisition are accessible via the L1, (e.g., Schachter, 1989; 1996). Second, L2 acquisition starts with some parts of L1 grammar and learners have partial access to UG (e.g., Eubank et al., 1997). Third, L2 learners' access to UG is similar to that they had when acquiring their first language, but their L2 initial state grammar represents some L1 grammatical properties, (e.g., Eubank, 1994; Vainikka and Young-Scholten, 1996). Fourth, L2 acquisition is just like L1 acquisition; L2 initial state is free from any previously

54

acquired grammatical properties and the learners have full access to Universal

Grammar, (e.g., Flynn, 1996). According to this fourth assumption, L1 properties do not

affect the second language grammatical representation at any stage of the acquisition

process. Finally, the fifth perspective is that L2 learners continue to have full access to

UG, but their initial state grammar is the mature complete grammar acquired based on

their L1 input (e.g., Schwartz and Sprouse, 1996). According to this assumption, L2

learners will gradually restructure their initial state grammar based on L2 input; L1

properties that contradict evidence provided by L2 data will be reset, and new

properties, which are required by the L2 data and not supplied by L1 grammar, will be

generated via access to UG.

As evident from the above, one common position that four of the five perspectives take

is that the L1 has a role to play at least in the initial stage of second language

acquisition. Indeed, many studies during the past two decades have shown clear

evidence of L1 transfer in second language acquisition (e.g., Hulk, 1991; Vainikka and

Young-Scholten, 1994; 1998; Bohnacker, 2006; Gil and Marsden, 2010; Gil et al.,

2011).

Assuming L1 effects to be extant in the L2 initial stages, the question that remains is

whether these transferred L1 grammatical properties will continue to be found

throughout the subsequent stages of L2 acquisition. In other words, we still need to

know whether the previous L1 knowledge will also play a role in shaping the L2 end-

state grammar. Han (2004), for example, argues that the L1 preprograms L2 learning; it

determines the developmental as well as the ultimate outcomes of L2 acquisition. Han

holds the same view established by Slobin's work (1996), that is, L2 learners transfer their L1-based conceptual system to L2 acquisition, and hence, different L1-based conceptual systems lead to varying levels of L2 ultimate attainment.

Hawkins & Chan (1997), in their 'failed functional features hypothesis'[9], attribute the observed common impairment in L2 ultimate attainment to L1 influence. They assume transfer of at least the values of the features in functional categories in L2 acquisition. According to them, learners who start learning the L2 after the critical period will not be able to reset these values despite the availability of the UG principles, which constrain building of the L2 grammar. Instead of resetting the values, L2 learners will map L2 morphological forms onto L1 transferred values at early stages. Then, with continued exposure to L2 data, their performance in the L2 will gradually come closer to that of its native speakers and away from their L1. This happens not by resetting the L1 values, as they are no more accessible, but by establishing "grammatical representations which diverge from those of native speakers, as well as from their own L1s, but which are nevertheless constrained by the principles of UG" (Hawkins and Chan, 1997:216).

The idea that the first language will potentially have an impact on L2 ultimate attainment is supported by some of the UG-oriented models presented earlier. Schachter (1990), for example, concluded from her study that "native language has a significant effect on knowledge of one principle of Universal Grammar [i.e., Subjacency] in post-puberty-acquired second language grammars", (116). She maintained that only UG

---

[9] According to Hawkins & Chan (1997), this hypothesis was posited as following on theoretical proposals given in Tsimpli and Smith (1991) and in Smith and Tsimpli (1995).

principles that have been instantiated in first language acquisition constitute an internal knowledge source for L2 learners. This internal knowledge (i.e., L1-based instantiations of UG) together with the external evidence (i.e., L2 input) will guide the process of acquiring a second language and constrain much of the form of its grammar. According to Schachter, this predicts that L2 learners will fail to acquire the features of L2 grammar which are not provided by the L1 grammar and are not made evident by L2 positive input.

In a similar vein, Bley-Vroman (1989), in his Fundamental Difference Hypothesis, argues that L1 knowledge defines the grammar that is possible for adult L2 learners. According to him, the difference between the attainments of children versus adults in acquiring a second language is due, in part, to the fact that UG is no longer available to adults after the closure of the 'putative' critical period, and, hence, L1 knowledge, together with general problem-solving systems, takes the role of guiding adults' acquisition process. Bley-Vroman posits that adult L2 learners reconstruct much of the original scheme of Universal Grammar, which is not available after the critical period, by observing the native language. He added that because of the L1 providing an indirect knowledge of UG, plus being 'incomplete' and 'accidental', "one can expect some partial success, little chance of perfect success, and some considerable individual variation" (1989:53).

The Full Transfer/Full Access model, developed by Schwartz and Sprouse (1994), affirms the view that the native language has an impact on L2 ultimate attainment; but not by taking the place of UG. According to this model, L2 learners start their L2

acquisition with the final state grammar of L1 acquisition. Then, based on L2 input, and by having access to UG, the learners will start restructuring all the features of L1 which fail to represent L2 data, and keep those which apply to both languages as they are. However, this model does not guarantee a full convergence on the grammar of the target language. This is because L2 learners, unlike L1 acquirers, start with previously set values; it may be that the data needed to force restructuring simply do not exist, or that the needed positive L2 data are highly obscure, complex, or rare (Schwartz and Sprouse, 1996:42). Under this model, the starting point of acquiring a language as an L1 or an L2 differs, and the outcome of these two acquisition processes are likely to differ: "the final states of L2 acquisition do not systematically replicate the final state of L1 acquisition", and this is due to distinct initial states of the two processes (Schwartz and Sprouse, 1996:42).

Sorace (1993) concluded that the L1 determines what type of competence L2 learners will reach at the end of the acquisition process. According to her study, even near-native speakers of the second language who are at the most advanced stage of second language acquisition will not, most likely, attain a native-like competence; their underlying competence may be either incomplete (lacking some L2 grammatical properties) or divergent (certain L2 properties are presented differently) from the native competence of the target language. This feature of incompleteness or divergence is argued to be due to L1 influence.

The focus of research on L1 transfer has shifted recently from investigating transfer of knowledge of a single linguistic domain or part of it to investigating transfer at the

linguistic interfaces (e.g., Slabakova, 2008; Sorace and Serratrice, 2009; White, 2009). To this point, structures at external interfaces, such as syntax-pragmatics, have been claimed subject to transfer more than structures at interfaces between internal domains of the grammar like syntax-semantics or syntax-morphology (e.g., Tsimpli and Sorace, 2006; Sorace and Serratrice, 2009). While transfer at external and internal interfaces has been attested in the literature (e.g., Gabriele, 2010; Haznedar, 2010; Oh, 2010), the issue of which is more vulnerable to transfer is not resolved yet (White, 2009). Ionin & Zubizarreta (2010), in an introductory paper to selected articles discussing the issue of L1 transfer at the linguistic interfaces, pointed out that these articles provided evidence for the following claims. L1 knowledge of structures at both external and internal interfaces is subject to transfer more than knowledge of purely syntactic structures (see, Montrul, 2010). Also, it is possible to recover from negative transfer of knowledge of at least structures at the syntax-semantics interface (see, Oh, 2010).

To summarize section 1.2 of this chapter, the influence of the first language is not a new topic in the context of second language acquisition, but has been presented and discussed in various ways. What is more relevant to the present work is the treatment of the L1's role within the framework of generative grammar. It is important here to digress and recall that the participants of the main study in this thesis are of two different types: the graduates of MSA immersion nurseries and kindergartens and those who started learning MSA at primary school. At least the latter group of subjects can be safely considered as acquiring MSA as an L2 based on the assumption that they had already acquired the basics of their L1 grammar (the grammar of the colloquial dialect) by age 5 or 6 when they started learning MSA (Karmiloff-Smith, 1986; Guasti, 2002; Schwartz, 2004; Unsworth, 2004). One of the research questions is about whether these

subjects have succeeded in acquiring a native-like competence of MSA or whether their end-state competence is fraught with imperfection or incompleteness that is indicated by the presence of L1 properties.

Most of the models reviewed above assume that L2 learners (the second group of our participants) will transfer all or some of their L1 grammatical properties to their L2 initial state. However, these models have different perspectives on the role of the L1 in determining the final outcome of the second language acquisition process. Researchers like Schachter (1990) and Bley-Vroman (1989) assume that knowledge of the L1 will play the role of UG in the L2 acquisition process for those who no longer have access to UG due to closure of the 'putative' critical period. Others like Schwartz and Sprouse (1994) argue that all L2 learners have full access to UG anyway, yet do not guarantee full convergence to native-like competence of the L2 due to their starting out with the L1 grammar. If this model is correct, then our participants who acquire MSA as a second language may not succeed in attaining a native-like competence, if not because of the closure of the critical period or the declining innate ability after the age of 5 or 6, then due to the fact that they started out with the previously-set values of their colloquial dialects' grammars. The current study will test their representation of some grammatical properties of MSA that are represented differently in their colloquial dialects. These aspects of MSA grammar will be introduced in the following section.

2   MSA Syntax

As has been mentioned before, this thesis is intended to test the ultimate attainment of Modern Standard Arabic (MSA) grammar acquired by Arabs who speak different local dialects of colloquial Arabic (CA). MSA grammar is different in some aspects from the grammars of the colloquial dialects. One of the main purposes of this thesis is to determine whether or not these speakers represent a native-like 'uniform' competence of MSA despite the grammatical differences between MSA and their different local colloquial varieties of Arabic. In order to do this, certain aspects of MSA grammar will be tested to see if these speakers have uniform representations (native-like competence) or distinct ones (L2-like competence), and if differences are found, whether they are due to the differences between the CA varieties (L1 influence).

This section will present general background knowledge on the syntactic aspects of MSA grammar which will be tested in this research. These aspects include word order & subject-verb agreement, adverb placement and resumption in relative clauses. Section 2.1 encompasses several general issues in relation to clause structure, word order, and subject-verb agreement in MSA. It lays out the word orders permitted in MSA and determines the basic underlying order from which the other variations are derived. Moreover, it explains how at least the two main orders (VSO and SVO) are syntactically derived from the underlying structure. Also, it introduces the optionality of plural agreement in sentences with pre-verbal collective nouns. Finally, it presents some pragmatic and non-pragmatic factors which trigger word order variation in Arabic. Section 2.2 defines where adverbs are placed in the clause structure of MSA. Section 2.3 discusses resumption in different types of relative clauses in MSA. In chapter 3,

these syntactic phenomena will be investigated in a corpus-based study of three colloquial varieties to pinpoint differences among these colloquial varieties, on one hand, and between the varieties and MSA, on the other hand. Differences will be used as syntactic variables for the experimental study that will be presented in the subsequent chapters of this thesis.

2.1 Word Order & Subject-Verb Agreement

Modern Standard Arabic (MSA) allows two possible main word orders: VSO and SVO, as in (1) and (2):[10]

1. <u>VSO</u> (this is the common order in MSA (e.g., Bakir, 1981))

   a. اشترى الطلاب كتباً
      ?ištaraa   l-Tullaab-u   kutub-an
      bought 3m.s   the-students-nom-3m.p   books-acc
      'The students bought books'

   b. بدأت التحقيقات منذ أيام
      bada?at   ?a-taHqiqaat-u   munðu   ?ayaam
      started-3f.s   the-investigations-nom-3f.p since   days
      'The investigations started days ago'

2. <u>SVO</u>

   a. الطلابُ اشتروا كتباً
      ?al-Tullaab-u   ?ištara-uu   kutub-an
      the-students-nom-3m.p   bought 3m.p   books-acc

      'The students bought books'

   b. التحقيقاتُ بدأت منذ أيام
      ?a-taHqiqaat-u   bada?at   munðu   ?ayaam
      the-investigations-nom-3f.p   started-3f.s   since   days
      'The investigation has started since days'

---

[10] Transliteration of Arabic script follows the transliteration scheme given at the beginning of this document.

The contrast between the examples in (1) and (2a) exhibits a subject-verb agreement asymmetry: when the verb follows the NP subject *?al-Tullaab-u* 'the-students', it agrees with it in person, gender, and number, as in (2a), but when the verb precedes the NP subject, it agrees with it in person and gender only; number agreement is lost in such a case. In addition, another subject-verb agreement asymmetry can be noted between the examples in (2a) and (2b) which are both in SVO order: the verb in this order shows full agreement with the preceding NP only if it has a human reference as in (2a); when the pre-verbal NP is inanimate or non-human, the verb takes the singular form regardless of the number feature of the preceding noun, as in (2b). This latter asymmetry can be explained by the fact that non-human plurals are treated in Arabic as a singular group (Badawi et al., 2004:93), and, hence, the lack of plural agreement.

MSA also allows all the other permutations of the basic sentence constituents; Subject, Verb and Object. These variant orders are used to add pragmatic functions to the communicated message like emphasis or contrastive focus (Bakir, 1980); see the examples below:

3. <u>VOS</u>

اشترى كتاباً الرجلُ
?ištaraa   kitaab-an   ?r-rajul-u
bought   (a) book-acc   the-man-nom
'The man bought a book'

4. <u>SOV</u>

الرجلُ كتاباً اشترى
?r-rajul-u   kitaab-an   ištaraa
the-man-nom (a) book-acc   bought
'The man bought a book'

63

5. **OSV**

الكتابُ الرجلُ اشتراهُ

?l-kitaab-u   ?r-rajul-u   ištaraa-hu
the-book-nom   the-man-nom   bought-**it**
'The man bought the book'

6. **OVS**

تفاحةً أكلَ الرجلُ

tuffaHat-an ?akal-a ?r-rajul-u
apple-Acc   ate3m.s   the-men-Nom
'The man ate an apple'

The MSA behaviour of allowing these different varieties of word order and exhibiting subject-verb agreement asymmetry has led to the investigation of issues which have engaged linguists for a long time and on which they have reached little agreement (e.g., Mahfoudhi, 2002). The following subsections will determine the basic underlying word order in MSA, review syntactic analyses of how VSO and SVO orders are derived, introduce the relevant aspect of subject-verb agreement to the current work, and present some pragmatic and non-pragmatic factors which trigger word order variation in MSA.

2.1.1 Basic Underlying Word Order

It is widely accepted that VSO order is the common pragmatically unmarked order in MSA (see, e.g., Bakir, 1980; Fassi-Fehri, 1993). This is supported by the observation that this order is the only one used in discourse-initial sentences and in answering questions about general state of affairs like 'what happened?' (Bakir, 1980). Such observations suggest that VSO is the neutral order which does not serve pragmatic functions like topicalisation or focus; evidently, introductory sentences do not usually topicalise old information, and answers to such general questions do not supply

emphatic information. VSO being the most common unmarked order does not necessarily mean that it is the basic order in the sense that it is the underlying structure from which other variations are derived (Fassi-Fehri, 1993). So, the question that remains is what is the basic word order in MSA?

From the six possible word orders, three have been nominated in the literature to be candidates as underlying structures from which other variations are derived; VOS (e.g., Majdi, 1990; Homeidi, 1991), VSO (e.g., Bakir, 1980), and SVO (e.g., Bolotin, 1995). Bakir (1980) and Fassi-Fehri (1993) exclude the possibility that VOS is the basic word order in MSA. They maintain that the S(ubject) should always precede the O(bject) in the underlying structure, regardless of the position of the verb. They argue that this is evident from the fact that only **SO** interpretation is possible in contexts where the morphological case is not overt, whereas a shift to **OS** is possible only when the morphological case is overt; compare the example in (7) with (8):

7. ضرب عيسى موسى

Daraba	Issa	Mussa
hit 3s.m	Issa	Mussa

'Issa hit Mussa'

8. زيداً ضرب موسى

Zayd-an	Daraba	Mussa
Zayd-acc.	hit 3s.m	Mussa

'Mussa hit Zayd'

In (7), where the morphological case marking is not overt, the proper noun 'Issa' has to be interpreted as the subject, due to the fact that it came first in the linear order. In (8), on the other hand, the proper noun 'Zayd', though it came first in the linear order, is

interpreted as the object, as the grammatical relation borne by 'Zayd' is recoverable through its case marking. Thus, it is argued that in non-overtly Case-marked contexts, NP objects have to remain *in situ*, due to the fact that their abstract Case is recoverable only from the position in which they originate.

Bakir (1980) proposed that examples like (7) not only rule out **VOS**, but also suggest that **VSO** is the basic underlying structure in MSA. According to him, a basic underlying word order in a language should be the least marked, pragmatically and syntactically, and consequently the most frequently used in that language; VSO, for him, satisfies these conditions in MSA. In addition to the fact that this particular order is naturally selected in such ambiguous contexts, VSO, in comparison to SVO, is used dominantly, and sometimes exclusively, in many other syntactic structures. In questions, for instance, the order of the constituents after the wh-element has to start with a verb followed by other constituents (Bakir, 1980:8). Bakir also provided a comprehensive explanation of how other word orders could be derived from VSO using the mechanisms of Transformational Grammar.

More analyses of MSA word order, however, (e.g. Fassi-Fehri, 1993; Plunkett, 1993; Aoun et al., 1994) adopt the proposal that the subject is generated as Spec of VP, and O as complement of V (Koopman and Sportiche, 1991). If this proposal is, crosslinguistically, correct, Bakir's proposal of VSO being the underlying word order cannot be maintained; leaving the one possible option, i.e. SVO, to be considered the underlying word order in MSA. The structure in (9) illustrates the basic underlying structure in MSA from which all the variations are derived, including the surface SVO.

9.

```
              VP
            /    \
         Spec     V'
          |      /  \
          S     V    NP
                     |
                     O
```
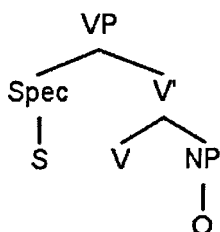
Derivation of the two main word orders in MSA (VSO and SVO) will be addressed in the following subsection.

2.1.2 Derivation of VSO/SVO Word Orders

Many of the analyses of clause structure in MSA take the derivation of VSO and SVO sentences into consideration. This is because, first, they are the two main and most common orders in Arabic in general, and second, the agreement asymmetry between the subject and the verb is evident by contrasting these two orders. The issue of whether or not the verb shows full agreement is sensitive to where it stands in the sentence in relation to the position of the subject. This subsection will review briefly some of the proposals about how these two orders are derived from the basic underlying structure defined above.

2.1.2.1 VSO Order

Most of the analyses of MSA word order assume that VSO order is derived by raising the verb to the I position, with the subject remaining in [Spec, VP]; its base position (e.g., Koopman and Sportiche, 1991; Fassi-Fehri, 1993; Plunkett, 1993; Mahfoudhi, 2002). Fassi-Fehri (1993) states that this operation involves an adjunction of a head to

another head, and it is subject to general principles of head movement. According to him, the motivation for raising the verb to I, which contains T(ense) and AGR, is to support the bound morphemes in AGR at surface structure. This kind of support for the bound morphemes is required by Lasnik's (1981:164) Stranded Affix Filter.[11]

Fassi-Fehri (1993) and Plunkett (1993), among others, argue that at least in declarative sentences of Arabic, the verb does not move beyond I to C. Fassi-Fehri supported his position by the fact that negation and modality occur after the complementiser and before the verb. The sentence in (10) is an example:[12]

10. زعَمَ أن قد لا يأتي زيدٌ

Za9am-a	?an	qad	laa	y-a?tii	Zayd-un
claimed-3.s.m	that	may	not	3-come	Zayd-Nom

'He claimed that Zayd may not come'

(Fassi-Fehri, 1993:26)

Plunkett, on the other hand, based her analysis on the fact that the verb may occur before the subject in both root and embedded clauses, and that verb initial order and complementizers are not in complementary distribution as shown in example (11) below. According to her, the verb in Arabic does not move beyond the Mood position.

11. أريدُ أن يخرجَ الطلابُ

?uriidu	?an	yaxruja	a-Tullaab-u
want-1s	that	leave-3sm	the-students-nom

'I want the students to leave'

(Plunkett, 1993:240)

---

[11] Lasnik's 'Stranded Affix Filter' states: "A morphologically realized affix must be a syntactic dependent of a morphologically realized category, at surface structure" (Lasnik, 1981:164).
[12] The transliteration of these examples and any others taken from external sources is reformulated to match the transliteration scheme followed in this work.

Aoun et al. (1994), however, proposed that in interrogative sentences, the verb might

move to C, based on the fact that the order in Wh-questions is always VS where the

question particle sometimes cliticizes to the verb as in (12):

12. أقرأتَ الكتابَ؟
    ?a-qara?ta   l-kitab-a?
    Q-read-2ms   the-book-Acc

'Did you read the book?'

(Aoun et al., 1994:204 fn 8)

2.1.2.2 SVO Order

Proposals about the derivation of SVO order in MSA depend very much on whether the

preverbal NP is considered a subject or a topic. Some linguists (e.g., Bolotin, 1995)

treat preverbal NPs as regular subjects. According to this view, the SVO order is

derived by moving V to I, and the internal VP subject to [Spec, AgrP].

In contrast, Plunkett (1993), among others, argued that the preverbal NP in this order is

not the subject but a topic. This type of analysis considers every sentence that begins

with a noun as a 'nominal' sentence which is composed of a *topic* 'al-mubtada' and a

comment 'al-xabar'. The topic is the preverbal nominal about which a statement is

made, while the comment may be either sentential or non-sentential, and whose

function is to modify the topic. The sentential comment may either contain a verb or be

verbless. This assumption sounds plausible especially when considering the fact that

MSA does not allow non-specified indefinite nouns (non-referential NPs) to occur

before the verb (Fassi-Fehri, 1993).

Ouhalla (1997) posits that the preverbal NP is either a left dislocated topic or a focus phrase. According to him, if the preverbal NP is a topic, it originates in a left-peripheral Top position or adjoined to the highest projection of the clause, but if it is a focus phrase, it moves from its base position to [Spec, F(ocus)P], a position below CP, to get focus or to check the feature [+f].

Plunkett (1993) treats preverbal NPs as a case of Left Dislocation to [Spec, TP] or [Spec, MoodP]. These preverbal NPs are coreferential with resumptive *pro* in [Spec, VP]; where subjects in Arabic are generated and remain throughout the derivation process. According to her, this is analogous to the case in Arabic of Left Dislocated topics that are coreferential with overt resumptive pronouns in object positions.

Fassi-Fehri (1993), however, proposes that the preverbal NP can be interpreted as a topic or a subject, depending on the intended sense and on other referential properties. He asserts that for the preverbal NP to be interpreted as a topic, it needs to be a 'strongly referential' definite NP. Preverbal subjects, on the other hand, can be definite NPs, like topics, or specific or quantificational indefinite NPs. Thus, according to Fassi-Fehri's analysis, sentences with a definite preverbal NP, like the one in (13) below, are ambiguous: they can have a topic reading or a subject reading. On the other hand, any sentence with an indefinite preverbal NP, like the ones in (14) and (15), must have a subject reading. According to him, sentences with a non-specific indefinite preverbal NP should not be possible in MSA.

13. الأولادُ جاؤوا
    ?al-?awlaadu   jaa?-uu
    The-boys-3mp   came-3mp
    'The boys came'

14. جاسوسٌ أقبلَ علينا
    jaasuws-un   ?aqbal-a   9alaynaa
    (a) spy-nom   appeared-3ms on-us
    'A spy came toward us'

15. كلُّ رَجُلٍ يحترمُ هذا
    kull-u   rajul-in   yaHtarimu   haðaa
    every-nom   man-gen   respect-3ms this
    'Every man respects this'

(Fassi-Fehri, 1993:27-28)

Fassi-Fehri states that deriving the SVO order in MSA depends on whether the preverbal NP is a topic or a subject. If it is a topic, it will move to a position external to IP to adjoin to CP. If it is considered to be a subject, then it moves to [Spec, IP] or [Spec, AGR]; if I is split. The following diagrams in (16), from Fassi-Fehri (1993:28), show the structure of both topic and subject readings of an SVO sentence like the one in (13) above:
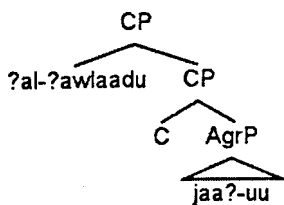
16.
   a. Topic reading   b. Subject reading
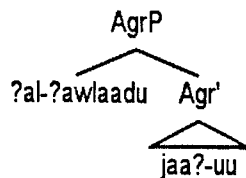
The issue of whether the preverbal NP in SVO order is a topic or a genuine subject in MSA will be re-visited and discussed in light of the results of the experimental study in Section 4 of Chapter 6.

2.1.3 Subject-Verb Agreement

There have been several analyses that attempted to account for the subject-verb agreement asymmetry in SVO/VSO variation in MSA introduced at the start of this section. Mahfoudhi (2002) presented a thorough, yet concise, review of different proposals. These analyses varied between considering agreement asymmetry as an effect of the presence or absence of a c-commanding thematic subject (Mohammad, 1990), the presence or absence of Spec-Head relation (Akkal, 1993), Agr-Criterion (Fassi-Fehri, 1993), *pro*-drop (Plunkett, 1993), or an effect of different feature values of Tense and Agr in these two word orders (Bolotin, 1995; for a detailed summary of these different syntactic analyses, see Mahfoudhi, 2002).

One relevant aspect of subject-verb agreement to the current thesis is the optionality of plural agreement in SVO sentences with collective nouns of a certain type being in pre-verbal position. Collective nouns of this type denote rational beings (humans) and include, for example, nouns like ناس naas 'people', شعب ša9b 'nation', and وفد wafd 'delegation'. The sentences in (17) and (18) have an SVO word order with a human collective in pre-verbal position, and both are grammatical in MSA:

17. الناسُ يخافون الخروج من المنزل ليلا" في هذا البلد

    l-naas-u   yaxaaf-uwna l-xuruwja mina l-manzili laylaan fii haðaa l-balad

    the-people-nom fear-3pm going out from the-house at night in this the-territory

    'People fear going out at night in this territory'

18. الناسُ تخافُ الخروج من المنزل ليلاٰ في هذا البلد

l-naas-u   taxaafu l-xuruwja mina l-manzili   laylaan fii haðaa l-balad

the-people-nom fear-3sf going out from the-house at night in this the-territory

'People fear going out at night in this territory'

In SVO sentences, human collective nouns are not like regular human plurals which require strict plural agreement, nor are they like inanimate and non-human plurals which require strict singular agreement. The verbal agreement with pre-verbal human collectives in these sentences can be plural or singular (Badawi et al., 2004), as shown in (17) and (18) above, respectively. Plural agreement is with the plural denotation of these collectives and singular agreement is with their singular grammatical forms. As will be discussed in chapter 3, some of the colloquial varieties of Arabic treat human collective nouns as regular human plurals which require strict plural agreement when being in a pre-verbal position. Thus, this phenomenon will be taken as a syntactic variable in the experimental study.

Having presented the possible word orders in MSA, defined the basic underlying structure from which word order variations are derived, and introduced the phenomenon of human collective noun-verb agreement in MSA, the following subsection will address what pragmatic functions the different orders serve and what other factors would trigger word order variation in MSA.

2.1.4 Pragmatic Functions and Factors of Word Order Variation

Bakir (1980:6-8) argues that all the variations of word order in MSA, except for VSO which he considers the unmarked neutral order, serve a certain pragmatic function. These are listed below:

19. **SVO** order is used to make a statement which constitutes new information about a preverbal NP topic which is old information and known to both the speaker and the addressee probably from previous sections of the discourse.

20. **OSV** order is used to topicalise the object about which the subject and the verb supply new information. In this order, the initial object is assigned nominative case and is coreferential with a resumptive pronoun in the original object position.

21. **OVS** order is used when the object needs to be focused and supplies new information. In such a case, the initial object keeps its accusative case and does not refer back to a resumptive pronoun.

22. **SOV** order is used to topicalise the subject which provides old information and about which the object supplies the new information.

23. **VOS** order is used to fulfil a contrastive function. In this order, the object is focused to supply new information in contrast with information previously mentioned in the discourse.

Based on Bakir's (1980) presentation, topicalisation and emphasis seem to be the main factors triggering word order variation in MSA. The topic serves the function of establishing "a special, temporal, or individual framework within which the main predication holds" (Chafe, 1976:50-51); hence, it occupies an initial position. Focus, on the other hand, "presents what is relatively the most important or salient information in the given setting" (Dik, 1978:19). According to Moutaouakil (1989), there are two types of focus in Arabic: 'new information focus', which presents new information to the addressee, and 'contrastive focus', which occurs in contexts where the speaker gives information conflicting with extant information in the discourse. Moutaouakil asserts that, in Arabic, only arguments belonging to the latter type are preposed to an initial position of the sentence, whereas those elements of the former type may be focused in situ by phonological means.

Bakir's (1980) presentation also implies a correlation between the topic/focus dichotomy and the old/new information distinction. One common organizing principle of information structure states that new information gravitates toward the end of the sentence following old information which comes first (Brustad, 2000). Brustad states that Arabic is among those languages which apply this discourse principle. Therefore, the correlation between the topic/focus dichotomy and the given/new information distinction becomes clear when considering the type of information that topics and focused arguments represent. Topics are always definite and contain old information that is given in previous discourse or generally known by both participants; thus, they are found in initial positions. In addition, focused arguments can be divided based on the type of information they represent: those of the 'new information' type tend to occupy sentence-final position; whereas arguments that represent contrastive focus do

75

not provide new information, since only known entities can be contrasted, and hence, they are usually preposed to sentence-initial position (Brustad, 2000). This given followed by new principle also entails that definite arguments should precede indefinite ones. So, for example, if we have a definite object which is not topicalised and an indefinite subject, the order of the sentence should be VOS.

Another factor that may be associated with word order variation in Arabic is the type of the general context or discourse. Brustad (2000), for example, states that VSO order should be dominant in event narration contexts, whereas SVO should be used in descriptive or conversational contexts where the discourse topic shifts around or is taken as a frame within which a main sentence predication holds. Dahlgren (1998) reached conclusions that support Brustad's claims when he analysed what he called 'early Arabic' (Classical Arabic) data. His results indicated that VSO is the common order in narrative discourse, whereas SVO is more common in dialogue discourse. However, he could not reach a definite conclusion about which order is more common in descriptive discourse due to insufficient relevant data. Owens (2009), also, argues that verb-subject order is used to present events. Following Myhill (1992), he points out that verb-subject word order is universally correlated with temporal sequencing. Also, he found that subject-verb order is used to signal available referentiality. According to the results of his corpus analysis of Arabic spoken in the Arabian Peninsula, he found that subjects of the type pronoun, pronominal,[13] personal noun, and definite noun with general referentiality tend to occur in subject-verb order; whereas indefinite subjects are more common in verb-subject order.

---

[13] Owens (2009:43) includes in the classification, 'pronominal', quantifiers, numerals, demonstrative pronouns and demonstratives when used alone as subjects.

Factors other than pragmatic ones have also been mentioned in the literature as causing the word order to vary in Arabic. Holes (1995:205-6), for example, pointed out that individual speakers of MSA may choose a particular order due to influence of the local spoken variety or probably a second language he/she speaks. Furthermore, Holes asserts that there is a rhythmic-based principle involved in regulating word order in Arabic, and which he calls 'end-weighting'. According to this, 'heavier' arguments which contain more lexical words appear to the right of 'lighter' arguments. Holes said that this principle should operate in harmony with the given followed by new principle; clauses which contain new information usually contain more lexical items than those which contain given information. However, the rhythmic-based principle should not apply if it conflicts with the information structure principle: given information always comes first.

The issues of word order and subject-verb agreement will be first investigated, in the following chapter, in a study based on corpora of three different colloquial dialects of Arabic. The purpose is to determine how similar or different from MSA these aspects are in the dialects (the L1s), and to pinpoint differences between these varieties. Then, these differences will be used as one of the syntactic variables when examining MSA end-state grammar acquired by speakers of different colloquial dialects. The general aim is to see whether these speakers transfer their L1 grammatical properties to MSA.

2.2 Adverb Placement

The typical function of adverbs in general is to "add specific information about time, manner, or place to the meanings of verbs or whole clauses", (Hurford, 1994:10).

Unlike adverbs in Germanic or Romance languages, Arabic adverbs have not been adequately studied within the framework of generative grammar, or even described in basic terms (Fassi-Fehri, 1998). This is thought to be partly due to the fact that adverbs, in general, are additives to the core structure of a clause; they are elements which are somehow incorporated into a sentence for modificational rather than argumental purposes, and hence, can be easily dispensed with. Also, adverbs in Arabic have received less attention from linguistic research probably because their class is heterogeneous in terms of its composition; constituents which function as adverbs in Arabic may be composed of categories like adjectives and nouns in addition to the few adverbs in and of themselves (Fassi-Fehri, 1998; Ryding, 2005). The following are examples of adverbs composed of different categories:

24. درستُ الكتابَ جيداً

| darastu | l-kitaaba | *jayyid-an* | | (Adjective) |
| studied | the-book | perfect-acc | | |

'I studied the book perfectly'

25. نسيَّ الولدُ طريقَه تماماً

| nasiya | l-waladu | Tareeq-a-hu | *tamaam-an* | (Noun) |
| forgot | the-boy | way-acc-his | completeness-acc | |

'The boy lost his way completely'

26. أكلَ الولدُ تفاحتينِ فقط

| ?akal-a | l-waladu | tuffaaHatayni | *faqaT* | (Adverb) |
| ate | the-boy | two apples | only | |

'The boy ate only two apples'

Adverbs in Arabic can be single words, like the ones in the examples above, or phrases, like the prepositional phrase *bi-sur9at-in* 'with-rapidity-gen = rapidly' (Ryding, 2005). As in other languages, most Arabic adverbs fall into four groups depending on the type

of their modification: degree, manner, place, and time. Adverbs of degree modify the meaning of the verb in terms of intensity and quantity. Manner adverbs are those which specify, for example, the state, the condition, or the way in which an event has happened. Place and time adverbs obviously denote where and when a certain event has happened (Ryding, 2005). The following are examples of these four types:

27. اكلتُ قليلاً

| ?akaltu | qaliyl-an | | (degree) |
| (I) ate | little-Acc | | |

'I ate a little'

28. سقطتُ فجأةً

| saqattu | faj?at-an | | (manner) |
| (I) fell | sudden-Acc | | |

'I fell down suddenly'

29. بدأ الحدثُ هنا

| bada?-a | al-Hada θ-u | huna | (place) |
| started | the-event-Nom | here | |

'The event started here'

30. غداً ستفتحُ الأبواب

| Gad-an | sa-tuftaHu | l-?abwaab | (time) |
| tomorrow | will- be opened | the-doors | |

'The doors will be opened tomorrow'

According to Ryding (2005), placement of Arabic adverbs in the clause structure is flexible to a certain extent; only some adverbs prefer certain positions. However, Al-Shurafa (2005) restricts this flexibility to the semantic scope of the adverbs. Based on her analysis of Arabic data, she asserted that adverbs should be classified into VP-adverbs and S-adverbs, depending on their functional properties in the sentence. VP-adverbs are those whose function in the sentence is limited to modifying the mode of

the verb. S-adverbs, on the other hand, have a wider scope and specify the action of the whole sentence. Compare the following two examples:

31. بسرعةٍ أنهيتُ واجبي

bi-sur9at-in	?anhaytu	waajib-ii	(VP-adverb)
with-rapidity-gen	(I) finished	duty-my	

'I finished my duty quickly'

32. طبعاً أحبُ أكلَ التفاح

Tab9-an	?uHibbu	?akl-a	l-tuffaaH-i	(S-adverb)
evidence-acc	(I) love	eating-acc	the-apple-gen	

'Evidently, I love eating apples'

Al-Shurafa (2005) claims that the structural syntactic position of an adverb in Arabic depends on its semantic function and scope in the sentence. Given that adverbs are not core syntactic elements in the structure (adjuncts), she proposed that VP-adverbs are adjoined to the VP projection whereas S-adverbs are adjoined to the S projection. The apparent flexibility of adverb placement in Arabic results from having a choice to adjoin adverbs either to the right or to the left side of the syntactic projection.[14] The adverbs in examples (31) and (32) above can also be placed in different positions; as shown in (33-34):

33. (بسرعةٍ) أنهيتُ واجبي (بسرعةٍ)

(bi-sur9at-in)	?anhaytu	waajib-ii	(bi-sur9at-in)
with-rapidity-gen	(I) finished	duty-my	with-rapidity-gen

'I finished my duty quickly'

---

[14] Because the position of the preverbal NPs is debated in the literature (subjects are placed inside IP and topics outside IP), Al-Shurafa (2005) perhaps used the term 'S projection', rather than IP or CP, to indicate that S-adverbs are adjoined to the highest projection in the structure. This means that such adverbs occur in the leftmost position of the sentence if they are left adjoined to that projection, and hence will always appear before the preverbal NP, whether it is a subject or a topic. Al-Shurafa, though, did not state this in her paper.

34. (طبعاً) أحبُّ أكل التفاح (طبعاً)

    (Tab9-an)   ?uHibbu   ?akl-a   l-tuffaH-i   (Tab9-an)

    evidence-acc (I) loveeating-acc   the-apple-gen evidence-acc

    'Evidently, I love eating apples'

Although it appears that the adverbs in both examples (33) and (34) are placed initially or finally, the adverbs in (33) occupy different syntactic positions from those occupied by the adverbs in (34). In (33), *bi-sur9at-in* 'quickly' is of the VP-adverb type; and hence, adjoined to the left or the right side of the VP-projection. In (34), however, the scope of *Tab9-an* 'evidently' involves the whole sentence; therefore, it is adjoined to the left or the right side of the S projection. This should be evident by comparing the above examples with the following ones:

35. أصدقائي (بسرعة) أتوا (بسرعة) إلى الدرس

    ?aSdiqaa?-ii (*bi-sur9at-in*) ?atuw (*bi-sur9at-in*) ?ilaa   l-dars-i

    friends-my with-rapidity-gen came with-rapidity-gen to   the-class-gen

    'My friends came quickly to class'

36. (طبعاً) أصدقائي أتوا إلى الدرس (طبعاً)

    (*Tab9-an*)   ?aSdiqaa?-i ?atuw ?ila l-dars-i   (*Tab9-an*)

    evidence-acc friends-my   came to   the-class-gen evidence-acc

    'Evidently, my friends came to class'

Since *bi-sur9at-in* 'quickly' is adjoined to the VP projection, it comes in (35) either between the subject and the verb or immediately following the verb. In (36), on the other hand, *Tab9-an* 'evidently' is adjoined to the S projection, and hence, had the option of occurring either initially or finally.

In the following chapter, adverb placement will be investigated in the corpus data of the colloquial dialects of Arabic mentioned before. The differences between the dialects will be used as one of the variables when examining MSA produced by speakers of the three dialects under study.

2.3 Resumption in Relative Clauses

Resumption is one of the strategies which Arabic, among other languages, employs to form unbounded dependency constructions. Relative clauses constitute one such type of construction. One common way to form a relative clause is by moving a wh-phrase (or a null operator) from the extraction site to [Spec-CP], leaving a bound gap behind:

37. العونُ الذي قدمتَ للعراق ...

   ?al-9awn-nu   [CP l-aðii$_t$ [IP qaddamta $\varnothing_t$ li-l-9iraaq-i]] ...
   the-aid-nom   which   (you) offered   to-the-Iraq-gen

   'The aid that you have offered to Iraq ...'

An alternative strategy is the resumption strategy. In relatives formed by this strategy, a *resumptive* pronoun is inserted in the extraction site as a variable bound by the wh-phrase, which is also directly inserted in [Spec-CP] position:

38. العونُ الذي قدمتَهُ للعراق ...

   ?al-9awn-u   [CP l-aðii$_t$ [IP qaddamta-hu$_t$ li-l-9iraaq-i]] ...
   the-aid-nom   which   (you) offered-it   to-the-Iraq-gen

   'The aid that you have offered to Iraq ...'.

Relative clauses formed this way are claimed to be immune to locality constraints (see, e.g., Borer, 1984). This is because, by assumption, the resumption strategy does not involve movement; both the relativised element and the resumptive pronoun are claimed to be directly generated in their relevant positions.

In Arabic, it is not always the case that both strategies are available when forming a relative clause. The availability of such optionality depends, for example, on whether or not the relative clause in question is definite or indefinite. Definite relatives are those which always occur with a complementiser such as *?allaðii* 'that' as in (39) below, and the indefinite ones are those which cannot occur with a complementiser, as in (40) below. The optionality of choice between the resumption or movement strategies is available only for forming definite relatives; indefinite relatives must have a resumptive pronoun in their relativisation site (Aoun et al., 2010). The following two examples from Aoun et al. (2010:10) should make this distinction clear:

Definite Relative Clause

39. ضـاعَ الكتابُ *(الذي) اشتريتهُ البارحةَ

    Daa9a   l-kitaab-u   *(l-aðii) štaraytu-(hu) l-baariHata

    be-lost-3ms   the-book-nom   that (I) bought-(it) the-past night

    'The book that I bought last night is lost'

Indefinite Relative Clause

40. أفتّشُ عن كتابٍ (*الذي) أضعتـ *(ـه)

    ?ufattišu 9an kitaab-in   (*l-aðii)   ?aDa9tu-*(hu)

    (I) look   for (a) book-gen   that   (I) lost- *(it)

    'I am looking for a book that I lost'

The choice between the two strategies also depends on the grammatical position of the relativisation site. For example, resumptive pronouns do not appear in the highest

subject position of the relative clause.[15] By contrast, gaps are not licit in prepositional and genitive object positions (e.g., Galal, 2004). Compare the following examples:

41. رأيتُ الولدَ الذي (*هو) اشترى الكتاب

ra?aytu	l-walad-a	l-aðii	(*huwa)	štara	l-kitaab
(I) saw	the-boy-acc	who	(*he)	bought	the-book

'I saw the boy who bought the book'

42. اشتريتُ الكتابَ الذي قرأتُ عنـ(ـه)* كثيراً

?ištaraytu	l-kitaab-a	l-aðii	qara?tu	9an-*(hu)	kaθiir-an
(I) bought	the-book-acc	which	(I) read	about-it	a lot

'I bought the book which I read about a lot'

43. رأيتُ الرجلَ الذي احترقَ بيتـ(ـه)*

ra?aytu	l-rajul-a	l-aðii	?iHtaraqa	baytu-*(hu)
(I) saw	the-man-acc	who	burnt	house-his

'I saw the man whose house has burnt'

As shown from (41), relative clauses with extraction from the subject position cannot be formed with a resumptive pronoun in the extraction site. On the other hand, the examples in (42) and (43) indicate, respectively, that the resumption strategy seems to be the only option available for forming relatives with extraction from prepositional and genitive object positions; gaps in such places would render the sentence ungrammatical. However, both strategies appear to be available when extracting from the object position; a resumptive pronoun seems to vary with gaps in filling this extraction site (e.g., Aoun et al., 2010). The following sentence is an example of this case:

---

[15] The observed gap in subject position is argued by Galal (2004) to be a resumptive *pro*; not a trace. This is because the gaps in this position are licit within locality islands.

44. (هـ)احترقَ البيتُ الذي بنيتُ

?iHtaraqa	l-bayt-u	l-aðii	banaytu-**(hu)**
burnt	the-house	which	(I) built-**(it)**

'The house which I built has burnt'

MSA uses relativisers of two types; relativisers of the first type denote specific referents in terms of gender and number, and relativisers of the other type can be used with any referents of any gender or number (Ryding, 2005). Example relativisers of the former type include الذي *?allaðii* for masculine singular, التي *?allatii* for feminine singular, اللذان *?allaðaani* for masculine dual, اللتان *?allataani* for feminine dual, اللذين *?allaðiina* for masculine plural, and اللاتي *?allaatii* for feminine plural. The second type relativisers are من *man* for animate referents of any number or gender, and ما *maa* for inanimate referents of any type and number.

In contrast, as we shall see in the following chapter, the colloquial dialects use only one relativiser *illi* 'that' for all types of referents. Also, these dialects seem to differ in relation to allowing gaps in object relatives. Shlonsky (1992), for example, claimed that resumptive pronouns are essential in object relatives in the Palestinian dialect of Arabic. Other dialects may not be like Palestinian Arabic; they may allow resumptive pronouns or gaps to fill in the extraction site in objects relatives, just like MSA. The colloquial corpora will be examined in the following chapter for this issue and the differences between the dialects will be used as one of the variables when examining MSA.

Finally, it is argued that resumptive pronouns add a portion of meaning to the sentence (Galal, 2004); they add specificity and more clarity to the reference of the relativised element (Alresaini, 2007). Consider the following examples:

45. سيجدُ عادلُ المرأةَ التي يُحِب

    sayajidu Adel-u ?al-mar?at-a   allatii yuHib
    will-find Adel-nom the-woman-acc that (he) love
    'Adel will find the woman that he loves'

46. سيجدُ عادلُ المرأةَ التي يُحِبُها

    sayajidu Adel-u ?al-mar?at-a   allatii yuHibbu-ha
    will-find Adel-nom the-woman-acc that (he) love-her
    'Adel will find the woman that he loves'

Doron (1982) and Sells (1984) pointed out (using equivalent examples from Hebrew) that only in sentence (46), the reference of the object relativised element, ?al-mar?ah 'the-woman', must refer to a particular individual woman that Adel loves. However, in (45), where the resumptive pronoun is lacking, this relativised element can refer to a particular woman, as in (46), or have a more generic reference to unspecified woman of particular properties. Also, the proposal that resumptive pronouns add clarity to the meaning of the sentence can be confirmed when considering examples like (47) below where resumption helps for disambiguation:

47. هذا الولدُ الذي ضَرَبَ-*(هُ) موسى

    haðaa l-walad-u   ?allaðii Daraba-*(hu) Musaa
    this the-boy-nom that   hit-3ms-him Musaa-nom
    'This is the boy that Musaa hit'

The sentence in (47) is ambiguous without the resumptive pronoun: the relativized element, ?al-walad 'the-boy', could be missinterpreted as the agent due to lack of overt nominative case marking on Musaa and due to the shared φ features in ?al-walad 'the-

boy' and <u>Musaa</u>. In such examples, the resumptive pronoun in the relativisation site is required to disambiguate the meaning of the sentence by making it clear that the agent is <u>Musaa</u> and the theme is <u>?al-walad</u> 'the-boy'.

2.4 Summary

Section 2 aimed to provide a general background to the syntactic issues of MSA covered in this research, including word order & agreement, adverb placement, and resumption in relative clauses. Section 2.1 presented the possible orders in MSA and defined the basic underlying order of MSA clause as being SVO. This section also presented proposals on how the main two orders in MSA (VSO and SVO) are derived, which when contrasted show subject-verb agreement asymmetry as agreement depends on where the subject is placed in relation to the verb. Some of these proposals claimed that SVO sentences have a Topic-Comment structure, others proposed that preverbal NPs are regular grammatical subjects, and a third view assumed that preverbal NPs could be interpreted as topics or subjects depending on the intended sense. These analyses will be discussed in light of the results of the experimental study in Section 4 of chapter 6.

This section also introduced the phenomenon of human collective noun-verb agreement in SVO sentences. In contrast to regular human plurals, which require strict plural agreement when preceeding the verb, human collective nouns can be in plural or singular agreement with the following verb in SVO sentences.

Discussion of word order also encompassed presenting the pragmatic factors behind its variation in subsection 2.1.4. It has been proposed that word order varies in MSA to

serve mainly two pragmatic functions: topicalisation and focus. It has also been proposed that different types of discourse entail preference of one order over the others.

Section 2.2 presented a review of adverbs and their placement in MSA clause structure. As one of the rare studies on Arabic adverbs, Al-Shurafa (1995) proposed that adverbs should be divided into two groups according to the semantic scope of their modification: VP-adverbs are those which modify the verb, and S-adverbs are those which modify the action of the whole sentence. Adverbs of the former type should be adjoined to VP and those of the latter type should be adjoined to S. Adjunction of either type could occur to either side of the projection, resulting with four possible places for adverbs: before or after the S projection, and before or after the VP projection.

The final section (section 2.3) addressed resumption in MSA relative clauses. It was shown that resumptive pronouns are obligatory in indefinite relatives, and their presence in definite relatives depends on the relativisation position: they do not show up in the highest subject position of the relative clause, always appear in genitive and prepositional object positions, and may or may not be used in object positions.

The following chapter will present a study based on corpora from different colloquial varieties of Arabic. The purpose of this study is to determine how similar or different these syntactic phenomena are in the grammars of the colloquial varieties. Defining similarities and differences between the colloquial varieties and MSA will help in

determining if MSA grammar represented by the subjects of the main study in this thesis is affected by different properties of the colloquial grammar (L1 influence).

Chapter 3

Methodology: Corpus Studies

One of the main aims of the current thesis is to observe and define the characteristics of the steady-state grammar of Modern Standard Arabic (MSA) spoken in three different regions of the Arab world: Levantine countries, Gulf countries and Egypt. If the underlying grammar of MSA varies across these regions, this might have implications on whether or not Arabs acquire MSA natively. If speakers of MSA are found to represent its grammar differently, it is important, and interesting at the same time, to know why this is the case and where differences come from. Two main factors are suggested in this work that may cause different representations of MSA grammar (if they are found to be different): L1 influence and age of first exposure to MSA.

This chapter presents two corpus based studies. The first one is mainly to pinpoint syntactic differences between the local non-standard dialects of Arabic spoken in Levantine countries, Gulf countries and Egypt. These local non-standard dialects will be referred to as Colloquial Levantine Arabic (CLA), Colloquial Gulf Arabic (CGA), and Colloquial Egyptian Arabic (CEA), respectively. The syntactic differences between these Colloquial Arabic (CA) varieties will be used as variables when testing MSA spoken in the three regions mentioned above. Four syntactic phenomena have been investigated: word order, subject-verb agreement in specified contexts, adverb placement, and resumption in object relatives. The second study investigates the same syntactic issues in a set of corpora of MSA data produced in the three Arabic regions of

Levantine, Gulf, and Egypt. The aim of conducting this study is to test whether the differences found between the local varieties also exist to differentiate grammatical representation of MSA spoken in those different regions. If such a result is reached, then L1 influence exists, and this, in turn, indicates that MSA grammar has the characteristics of a second language grammar.

This chapter is divided into two main sections. The analysis of the CA corpora will be presented in the first section whereas the analysis of the MSA corpora will be presented in the second section. Under each section, the corpus data will be described as well as the method used in analysing the data. Also, the results of investigating the syntactic issues will be given and discussed.

1   Colloquial Arabic Corpus Study

In order to discover whether discrete grammars of MSA are there due to influence of the grammatical differences between the unequivocally natively acquired local spoken non-standard varieties of Arabic (the L1s), we need first to determine these grammatical differences. Unfortunately, grammatical differences between the colloquial varieties of Arabic have received little attention in the literature (Cuvalay-Haak, 1997:19). In general, works comparing CA varieties usually find more grammatical similarities than differences between these varieties (see, e.g., Brustad, 1991; 2000). Also, grammatical differences found by such studies are usually of a very salient type exemplified by the contrast between (48) and (49) in terms of the syntactic position of question words:

48. حتروحي فين؟

    HatruwHi   **fiyn?**   (Colloquial Egyptian Arabic)

    Will-(you) go-2fs   **where**

    'Where will you go?'

49. وين بتروحين؟

    wayn bitroHiyn?   (Colloquial Gulf Arabic)

    where will(you)go-2fs

    'Where will you go?'

Colloquial Egyptian Arabic and Colloquial Gulf Arabic apply different strategies when forming wh-questions; it is possible in the former to leave the wh-question word in situ whereas in Gulf Arabic, wh-questions are formed by applying the movement strategy. Such a difference, however, would be very clear and obvious for Egyptians, for instance, to avoid when speaking MSA, given that they know from input that wh-questions in MSA begin consistently with a wh-word. Therefore, a study was needed to pinpoint less salient syntactic differences between the colloquial Arabic varieties which may continue to appear in MSA spoken in different regions of the Arab world.

In this corpus based study, data from three colloquial Arabic varieties were analysed to investigate four syntactic phenomena which were thought of as good candidates to reveal grammatical differences suitable for the purposes of this thesis. As mentioned above, the colloquial Arabic varieties from which data was studied are Colloquial Levantine Arabic (CLA), Colloquial Gulf Arabic (CGA), and Colloquial Egyptian Arabic (CEA). These three varieties of Arabic represent main geographical linguistic groupings in the Arab world in addition to those varieties spoken in North African Arabic countries (Aoun et al., 2010:2). The syntactic aspects investigated include word order, adverb placement, resumption in object relatives, and, finally, subject-verb

agreement when the sentence is in SV order and the subject is in 3rd person plural form.

Selection of these particular aspects will be justified when each is discussed in sections

that follow.

All the data examined in this study was obtained from a set of corpora developed by the

Linguistic Data Consortium (LDC).[16] The colloquial Levantine Arabic data examined in

this study is a part of an LDC corpus entitled 'Levantine Arabic Conversational

Telephone Speech, Transcripts' (Appen Pty Ltd, 2007).[17] This corpus contains a total of

495 files of transcribed natural unscripted telephone conversations between native

speakers of CLA. The colloquial Gulf Arabic data was obtained from corpus entitled

'Gulf Arabic Conversational Telephone Speech, Transcripts' (Appen Pty Ltd, 2006).[18]

This corpus contains a total of 526 files of transcribed natural unscripted telephone

conversations between native speakers of CGA. The colloquial Egyptian Arabic data

was obtained from an LDC corpus that has the title 'CALLHOME Egyptian Arabic

Transcripts' (Gadalla et al., 1997).[19] This corpus contains a total of 120 files of

transcribed natural unscripted telephone conversations between native speakers of CEA.

The data analysed for the current study was extracted from the above described corpora

by using software called 'CLAN'. It is designed specifically to analyse data transcribed

in the format of the Child Language Data Exchange (Mac Whinney, 2000). The

software allows a large number of automatic analyses to be performed on transcript

---

[16] The LDC is an open consortium of several research supporting institutions that is hosted by the University of Pennsylvania. It basically creates, collects and distributes speech and text databases, lexicons, and other resources for research and development purposes.
[17] The LDC catalogue number for the Levantine corpus is (LDC2007T01).
[18] The LDC catalogue number for the Gulf corpus is (LDC2006T15).
[19] The LDC catalogue number for the Egyptian corpus is (LDC97T19).

data, such as frequency counts, word searches, co-occurrence analyses, interactional analyses, etc. For the current study, one of CLAN applications for searching the data; namely, 'COMBO', was given lists of key words, like common verbs, adverbs, etc., to search for and extract the relevant data which provide instances of the syntactic aspects under investigation. The following sections will tackle each syntactic aspect in all three colloquial varieties of Arabic. Each section will first define how similar or different the CA varieties are from MSA in relation to the syntactic aspects under study.[20] Then, they will provide details of how CLAN was instructed to extract the relevant data, how the data was analysed, and the results of the analysis.

1.1 Word Order

Like MSA, CA varieties allow all the permutations of the basic constituents of the sentence; Subject, Verb, and Object, as long as the meaning is clear. One difference between MSA and CA varieties, however, is that VSO is the common unmarked order in the former whereas SVO is claimed to be the unmarked order in CA varieties (Cuvalay-Haak, 1997:19). Unmarkedness of word order is defined by Fassi-Fehri's (1993:19) as "the order found in so-called pragmatically neutral contexts, i.e. in sentences which require fewer mechanisms of interpretation or derivation". Bakir (1980:6) points out that if a certain order is marked the least, pragmatically and syntactically, it should be used most frequently in that language or variety.

Moreover, as detailed in section 2.1.4 of the previous chapter, people like Brustad (1991; 2000), Dahlgren (1998) and Owens et al. (2009), among others, claim that there

---

[20] See Section 2 of the previous chapter for description of MSA in relation to the syntactic aspects under study.

may be more than one common order in Arabic. They maintained that different contexts and discourse types tend to determine which word order should be prominent. Therefore, having all word order variations possible in both MSA and CA varieties with the possibility that MSA is different from CA varieties in favouring particular orders in certain contexts, it is pretty much possible too that each CA variety is distinct from other varieties in this regard. This corpus based study aims to explore word order variation in Colloquial Levantine Arabic, Colloquial Gulf Arabic, and Colloquial Egyptian Arabic. If it turns out that each or some varieties have their distinct preference for a particular order, these differences will be used as variables when testing MSA spoken in those three Arabic regions.

1.1.1 Data Extraction & Coding

The data examined in this study was extracted from bigger corpora collected and transcribed by the LDC. For all the CA varieties under study, an application called 'COMBO' was used to search for a list of common verbs and extract the relevant data which contained sentences with those verbs. This list was developed as follows. The LDC corpora of Levantine and Gulf Arabic each came with a file that lists all the complete Arabic words used by the speakers in those telephone conversations. By going through these two long lists, common verbs like 'eat', 'drink', 'walk', 'talk', 'say', 'buy', etc., that occurred in both corpora, and which are normally expected to appear in everyday speech, were picked to create a list of 140 common verbs. The 'CALLHOME' corpus of Egyptian Arabic, on the other hand, did not contain such a file. Thus, the list developed for the other two CA varieties was adapted for searching the CEA data. Some verbs were modified where necessary to match the way they are used by Egyptians. An example of this case is the verb جا jaa 'came-3ms' from CGA data which appeared as

اجإ?ijaa in CLA data and was modified to جه qih in the list of verbs adapted for CEA. The final list of verbs for searching CEA data was checked by a native speaker of Egyptian Arabic to make sure that verbs are in the correct forms that would be used by native Egyptians (see Appendix 1 for the complete lists of verbs).

In addition, the transcription of the data was in Arabic script, and because Arabic verbs may be inflected with prefixes and/or suffixes to indicate person, gender and number, verbs were prefixed and/or suffixed with a wildcard character (*), where applicable, to instruct 'COMBO' to find all the possible forms.[21] An example of these verbs is أكل ?akal 'ate-3ms' from CGA data which may occur in different forms like أكلوا ?akal-aw 'ate-3mp', أكلت ?akal-at 'ate-3fs', أكلن ?akal-an 'ate-3fp', ياكل yaa-kil 'eat-3ms', ياكلون yaa-kl-uwn 'eat-3mp', تاكل taa-kil 'eat-3fs', and ياكلن yaa-kl-in 'eat-3fp'. It is sufficient to give the searching application the verb *أكل* ?akal, prefixed and suffixed with a wildcard, to find all these variant forms if they appeared in the data.

'COMBO' was instructed to search for sentences with these verbs in a 100 files of each corpus. There was no specific reason behind choosing this particular number of files except that it was available in all the three corpora. Each file contains one transcribed unscripted conversation between native speakers of each variety of colloquial Arabic. Although the number of conversations searched in each corpus was the same, the size of each conversation may vary. The LDC reported that each speaker in CLA and CGA data has contributed with an average of 5 to 6 minute speech. Assuming at least two

---

[21] Some verbs, like those which end with a long vowel, could not be suffixed or prefixed with a wildcard because deriving the other forms of the verb would require changes on the stem of the verb. For example, the final long vowel in صلى Sallaa 'prayed-3ms' needs to be deleted or changed in other inflected forms of the verb like صلت Sall-at 'prayed-3fs' or تصلي tSall-ii 'pray-3fs'.

speakers in each conversation, the average duration of each conversation may be 10 to 12 minute long. On the other hand, the duration of the recorded telephone conversations in the 'CALLHOME' corpus was reported as variable with each lasting up to 30 minutes.

Furthermore, 'COMBO' was instructed to output all the result sentences in a context of no less than 10 conversational turns before and after the sentence of interest. This was instructed to control for the type of context within which the sentence of interest has appeared. This also helped to find more sentences to analyse with verbs other than those contained by the searching lists.

As Arabic, in general, is known as a *pro*-drop language, and because word order cannot be defined if the subject is missing, only declarative sentences with overt subjects were analysed and counted.[22] Similarly, sentences without verbs were not considered for the same reason; word order cannot be defined if the verb is missing.[23]

---

[22] Brustad (2000:318) considered sentences that have no overt subjects as sentences in a VSO order.
[23] According to Aoun et al. (2010), among many others, Arabic in general (standard and non-standard) allowes present tense sentences to have predicates without verbs. Such a predicate could be a noun phrase, an adjective phrase, or a prepositional phrase. The following MSA sentences from Aoun et al. (2010:35-36) exemplify such sentences:

   1. عمرُ معلّمٌ
      9umar-u   mu9allim-un
      Omar-nom   teacher-nom
      'Omar is a teacher'
   2. البيتُ كبيرٌ
      ?al-bayt-u   kabiir-un
      the-house-nom big-nom
      'The house is big'
   3. الكتابُ على المكتبِ
      ?al-kitaab-u   9alaa   ?al-maktab-i
      The-book-nom on   the-desk-gen
      'The book is on the desk'.

The output data was coded for word order and the type of context within which the sentences appeared in the data. As suggested in the previous chapter (section 2.1.4), type of context, among other factors, may have an impact on the choice of a certain order. It was claimed in some works (e.g., Brustad, 1991; Dahlgren, 1998) that VSO, for example, is prominent in narrative contexts whereas SVO occurs more frequently in conversational contexts. Therefore, the data for this study was coded to indicate whether the sentence in question occurs in conversational or narrative context as defined below.

The criterion followed to determine type of context is as follows. Bearing in mind that the examined data represent transcribed telephone 'conversations', the default type of the general discourse was considered conversational, which, according to Brustad (1991:116), usually expresses social relations and personal attitudes. Narrative contexts were determined when a sequence of events or actions were narrated about a constant topic or incident by one speaker to another. To have an example, what follows is an extract from a telephone conversation between two speakers of CLA who were talking about their business and the market in general. The extract is divided into two parts: the first part in (50) was considered conversational, and the second part in (51) was considered narrative:[24]

50. Conversational Context:

B: يعني قديه عم يعمل – قديه عم يعمل المحل بالشهر هيك بدك تحسب

'Ok, how much does it make –how much does the shop make in a month— you should calculate this way'

A: ايه لا هلق يعني هلق من عندك من أول الشهر لهلق يعني علنا خ- فوق الخمسطعشر صار يعني هالعشرة ايام فوق الخمسطعشر

'yah, no, now, yah now, since the beginning of the month till now, we made more than fifteen, yah in these ten days more than fifteen'

---

[24] Translation to English of this extract is mine.

A: إيه يعني بم- بمعدل كان المعدل مبارح عندي ألف وستمية باليوم

'yah, last night, the average was sixteen hundreds a day'

B: لعشرة ايام، طيب ممتاز أم

'for ten days, ok this is excellent, umm'

A: اليوم على السوق التعبان نزل للألف واربعمية وخمسة وسبعين هيك شي

'today, with the poor activity in the market, it went down to a fourteen hundreds and seventy five, something like that'

B: أم

'Umm'

B: هاي ممتاز

'this is excellent'

A: يالله المهم خلصنا من الأحد يعني ظل لك هاليومين بيجي اربعا خميس جمعة بيغطوا

'OK, what is important is that we are done with Sunday, two days remaining till Wednesday comes, with Thursday and Friday, which will make it up for me'

B: برافو

'Bravo'

A: عرفت علي، أنا عم باحاول قدر الإمكان إنه اغطي بالثلاثة ايام الثانية بحيث انه يعوضوا الأيام النايمة

'you know, I am trying my best to make use of the coming three days so that they make up for the dead days'

B: أكيد أكيد ها منيح

'sure, sure, this is good'

A: إيه

'yah'

B: يالله ياسيدي الله يرزقكم ويوفقكن يارب

'Ok, may Allah give you more and help you'

In (50), speaker A and B are having a conversation about speaker A's shop: speaker B asked a question, speaker A gave a fairly long answer, and speaker B comes in every now and then with a comment about what speaker A says: a continuous interaction between the speakers exists. When this conversation continued in (51) below, speaker A started narrating the events of an incident that happened in his shop between him and a customer who was asking for discounts:

99

51. Narrative Context:

A: هي هيك حركتنا مثل اليوم يامعلم زبونة جايتني من قبل الخصم أنا الخصومات بلشت اعملها من خمسة ايام أوكي

'this is what our business is like, like today, a customer who came to me before the sale-- I started making reductions five days ago, by the way, ok'

B: أم

'Umm'

A: بالضبط اجتني هي من قبل أسبوع شافت قطعة قلت لها مية وخمسة وسبعين

'precisely, she came to me a week ago, she saw a piece (of clothes, probably$_{SA}$), I said to her (the price is$_{SA}$) a hundred and seventy five'

B: أم

'umm'

A: قالت لي ايمتى رح تعمل تنزيلات قلت لها بعدني يومين ثلاثة ميلي علي ميلت علي اليوم

'She said when are you going to make price reductions, I said to her, two to three days yet then come back, and she came back today'

B: أم

'umm'

A: قالت لي قديه هاي قلت لها كانت سعرها مية وخمسة وسبعين هلق عشرين بالمية خصم صارت مية واربعين قال ايها بدي اياها بمية وعشرين
قلت لها نعم

'She said to me, how much is this? I said this was a hundred and seventy five, but now with twenty percent discount, it became a hundred and forty, she said, I want it for a hundred and twenty, and I said, 'what'?'

B: هاي هاي هاي

'this this this...'

A: يعني حتى بعد الخصم بدها تخصم

'even after discount she wants to pay less!'

B: هاي

'this...'

A: قالت لي ليش هذيك الزبونة حسبت لها اكثر من عشرين بالمية قلت لها خذي هذيك القطعة وباحسب لك اياها بخمسين بالمية

'she said to me, why for that customer did you give more than twenty percent discount?, and I said, take that piece and I will give you fifty percent'

B: والله

'really'

A: قال لا بدي هي قلت لها شو أنا حسب القطعة أنا ما بقدر إنه اضغط عحالي أي نقص بادفعه أنا من جيبي قال ادفع

'she said, no, I want this one, and I said to her, what?, it depends on what piece you take, I cannot press on myself, if I give you more discount I will pay the difference from my own pocket, and she said, ok, pay'

A: لا والله قلت لها قالت أيش قلت لها هن مية واربعين قالت انا بادفع مية وعشرين وانت ادفع عشرين قلت لها وانتي بتستخدميها يوم وأنا
يوم

'I said, really, she said, what?, I said, this is for a hundred and forty, she said, I will pay a hundred and twenty and you pay twenty, then I said, you use it for one day and I use it for the other'

B: والله

'really'

A: (ضحك)

'(laughter)'

B: (ضحك) حلوة هاي ملعوبة

'(laughter) nice, well said'

B: انت يوم وانا يوم

'One day for you and one day for me'

A: عن جد قال آخر شي زعلت قالت لي خلص أوكي ما رح تحسب لي اياها بمية وعشرين قلت لها لا

'seriously, at last she got upset and said to me, so you won't make it for a hundred and twenty? And I said, no'

A: مافي مجال أنا أعطيتك اياها سعرها بعد الخصم بعد هيك شو بدي اخصم لك بدي ادفع من جيبي

'I can't, I gave it to you already discounted, and I can't give you more discounts, I would pay from my pocket'

B: والله

'really'

A: قال طيب خلص شكرا قلت لها أهلا وسهلا

'She said, ok thank you, and I said, you are welcome'

In (51), speaker A took the role of the narrator and speaker B turned to a listener with minimal participation or addition to the content of the extract. All what has been said in (51) was about a single incident that took place in speaker A's shop. So, the context of passages similar to this in which one speaker talks to the other about an incident or a sequence of events or actions has been considered narrative when analysing the data; otherwise the context would have been considered conversational. The following section presents the results of data analysis for word order variation in CA varieties.

1.1.2 Results

The coded data contained a total of 712 CLA sentences; 823 CGA sentences; and 1096 CEA sentences, all in context. Table 1 below compares the frequency of using different word orders in the three colloquial varieties under study; the frequency of using word orders in separate contexts will be given thereafter. It is clearly indicated by the analysis

of the data that the distribution of word order in all the colloquial varieties of Arabic follows the same pattern when it comes to the general ranking of favourite orders. The majority of the sentences in the data were in SVO order, with VSO coming second, and VOS favoured the least. The other three variations of word order; OVS, SOV, and OSV, were rarely used by the speakers of CLA and CGA, and never used by the CEA speakers.
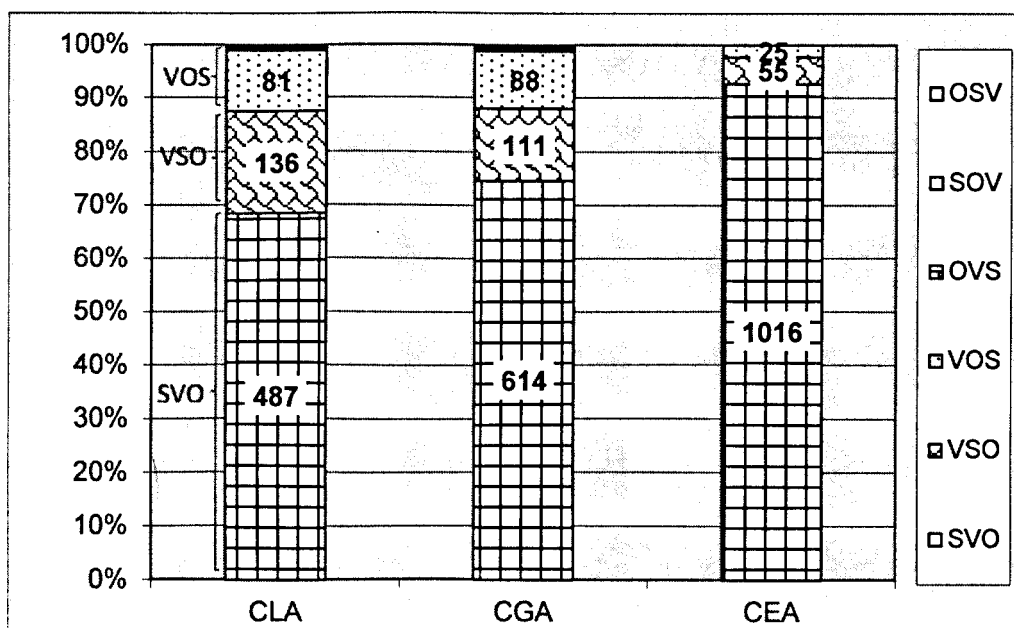
Word Order	SVO	VSO	VOS	OVS	SOV	OSV
CA Varieties	%	%	%	%	%	%
CLA	68	19	11	0	0	0
CGA	75	13	11	1	0	0
CEA	93	5	2	0	0	0

Table 1 : The frequency of using different word orders in CA varieties

Although ranking of word orders appears to be the same in all the CA varieties, the frequency of occurrence of these orders is found to be different when the CA varieties are compared, especially in CEA (see Table 1 , above). Speakers of this variety used SVO order in 1016 sentences which comprise 93% of the sentences found in the data.

On the other hand, speakers of CGA and CLA varieties used this order in 614 and 487 sentences comprising 75% and 68% of the sentences found in the data of these varieties, respectively. Furthermore, speakers of CEA seem to be different in using the other two orders too. They used VSO order in 55 sentences which are only 5% of the data compared to 13% (111) and 19% (136) in CGA and CLA, respectively. VOS order was used equally in CGA and CLA in 11% of the time, whereas in CEA, it was used only in 2% of the sentences. Other orders were not used even once in CEA data, which may indicate that this variety does not allow these orders, and were rarely used by the speakers of the other two CA varieties: in CGA data, OVS was used in 5 sentences, SOV in 2, and OSV in 3 sentences; and in CLA data, OVS and SOV were used 3 times and OSV was used twice.

The frequency data of using SVO and VSO in the three colloquial varieties were entered into 3 Pearson Chi-Square tests using IBM SPSS 19. The first test compared the frequency of using these two word orders in CLA and CGA, and the test revealed a significant difference, $\chi^2(1, N= 1348) = 9.52, p = .002$. The second test compared the frequency of using these two orders in CLA and CEA, and again the difference was significant, $\chi^2(1, N= 1694) = 109.74, p = .001$. The third test compared the frequency of using these word orders in CGA and CEA, and the difference is significant, $\chi^2(1, N= 1796) = 53.36, p = .001$.

As the source of the data was telephone conversations, most of the sentences were produced in conversational discourse. Therefore, the results do not vary that much if distribution of word order was considered in conversational contexts only. Table 2

below compares the frequency of word order variation used in this context. Obviously, the same ranking of favourite orders pertains: SVO comes out as the most common order in the three CA varieties, followed by VSO, then VOS. When the CA varieties are compared in terms of word order frequency in this particular type of context, the same similarities and differences are sustained. Most of the CEA sentences were put in SVO order (93%), with the rest being in VSO (5%) more than in VOS (2%), which differentiate CEA from the other two CA varieties. On the other hand, CGA speakers still use more SVO sentences (75%) and fewer VSO sentences (13%) than CLA speakers; 70% and 18%, respectively. All the instances of other orders were found in this type of discourse.
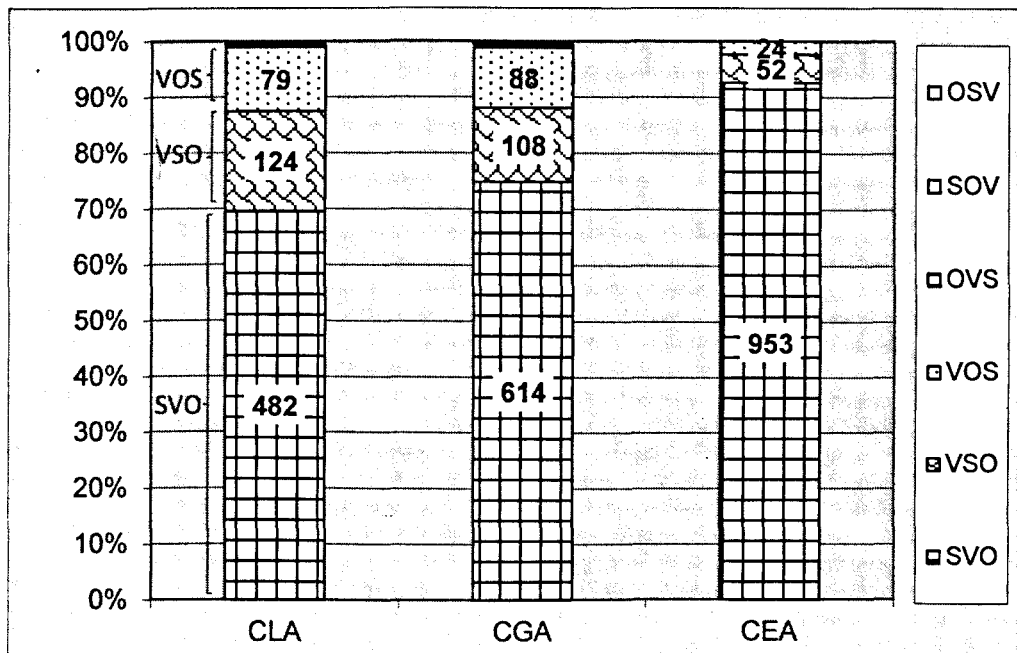
Word Order	SVO	VSO	VOS	OVS	SOV	OSV
CA Varieties	%	%	%	%	%	%
CLA	70	18	11	0	0	0
CGA	75	13	11	1	0	0
CEA	93	5	2	0	0	0

Table 2 : The frequency of using different word orders in 'Conversational Context' in CA varieties

In contrast, the distribution of word order in narrative contexts appears to reveal interesting differences between the CA varieties. Unfortunately, though, the instances found in this type of discourse are very few (89 sentences in total), and this is obviously due to the type of data analysed. Table 3 , below, shows the results of word order distribution in this type of context.
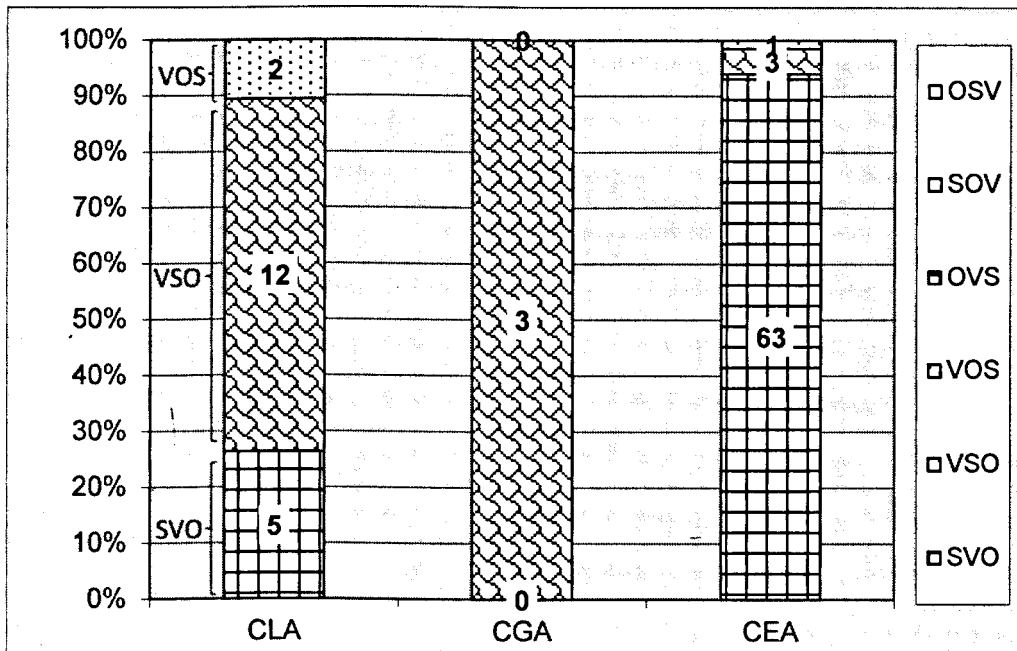
| Word Order | SVO | VSO | VOS | OVS | SOV | OSV |
CA Varieties	%	%	%	%	%	%
CLA	26	63	11	0	0	0
CGA	0	100	0	0	0	0
CEA	94	4	1	0	0	0

Table 3 : The frequency of using different word orders in 'Narrative Context' in CA varieties

In the CGA data, only three sentences were found in this discourse type, and they all were in VSO order. In the CLA data, the preference between SVO and VSO is reversed

in this type of discourse; 5 sentences were in SVO order compared to 12 in VSO, with VOS remaining the least favoured. In the CEA data, SVO continues to be the favourite order despite changing discourse type. This order was used 63 times out of 67 (94%); 3 sentences were in VSO and one was in VOS. None of the CA varieties' data contained instances of the other three word orders in this type of discourse.

To summarize what has been presented so far, CEA was found to be distinct from the other two CA varieties. In CEA, SVO is the most common order and is used more than 90% of the time regardless of the type of the discourse. However, in CGA and CLA, type of discourse seems to make a difference. SVO order is used the most in conversational discourse, followed by VSO, then by VOS; whereas in narrative discourse VSO becomes the most frequently used order, followed by SVO, then by VOS. The other three word orders were never or rarely used by the speakers of the three colloquial varieties.

These differences found between colloquial varieties in terms of word order frequencies will be taken as one of the variables when testing MSA. Existence of L1 influence would be confirmed if these differences reflect on speakers' selection of word orders when speaking MSA. Egyptians, for example, are expected to use SVO most frequently than the native speakers of the other colloquial varieties when they speak MSA. Having determined the first syntactic variable for studying MSA, the following section will compare 3rd person plural subject-verb agreement in the three colloquial varieties in an attempt to find differences that are suitable to be taken as a second variable.

1.2 Plural Subject-Verb Agreement

Like MSA, CA varieties exhibit full subject-verb agreement in person, gender, and number when the subject is pre-verbal. Unlike MSA, however, CA varieties continue to show the same type of agreement even when the subject is post-verbal (Ouhalla and Shlonsky, 2002:13). Compare the following examples from the data examined for this study:[25]

VS Order

تجاوز العراقيون مرحلة مهمة. 52

tajaawaza	l-9iraaqiyuwna	marHalatan	muhimmah	(MSA)
passed-**3ms**	the-Iraqis-**3mp**	(a) stage	important	

'The Iraqi people overcame an important stage'

طلعوا الهنود. 53

Til9a-**wu**	l-ihnwud	(CGA)
went out-**3mp**	the-Indians-**3mp**	

'The Indians went out'

عم يجوا السعوديين مامعاهن فلوس. 54

9am	yij-**wu**	l-su9wudiyiin	maa ma9-hun	flwus	(CLA)
prog.	come-**3mp**	the-Saudis-**3mp**	Neg with-them	money	

'The Saudis come with no money'

فا جم بتوع الحي. 55

fa	qum	bitwu9	l-Hay	(CEA)
then	came-**3mp**	those-**3mp**-for	the-area	

'Then, those (engineers assigned) for the area came'

SV Order

السوريون وصلوا. 56

?as-suuriywuna	waSal-**wu**	(MSA)
the-Syrians-**3mp**	arrived-**3mp**	

'The Syrians arrived'

---

[25] Transliteration, glosses, and translation to English of all examples taken from the data examined in this study are mine.

57. اربعة شباب أخذوا الكمبيوتر

?arba9ah	šabaab	?axaða-**wu**	l-kumbyuwtar	(CGA)
4	guys-**3mp**	took-**3mp**	the-computer	

'Four guys bought the computer'

58. كلهن عملوا ذات المشكلة

kil-hun	9iml-**wu**	zat	l-mišklih	(CLA)
all of-them	did-**3mp**	the-same	the-problem	

'They all made the same trouble'

59. ربعمية نفر ماتوا

rub9umiyyit	nafar	maat-**wu**	(CEA)
400	person	died-**3mp**	

'Four hundred persons died'

MSA and CA varieties exhibit different agreement patterns when the subject follows the verb; the verb in MSA always takes the singular form regardless of the number feature of the following subject, whereas in CA varieties, the verb takes the appropriate number inflection which agrees with the number feature of the following subject when the subject is human. This is evident by comparing the example in (52) from MSA with examples (53)-(55) from different CA varieties. On the other hand, when the subject precedes the verb, agreement in number is expected to appear in MSA and CA varieties as the examples in (56)-(59) demonstrate. Furthermore, when the pre-verbal NP is non-human, the verb takes the feminine singular form at all times in all the varieties of Arabic, standard and non-standard (Moawad, 2006:42). The following are examples of such cases taken from our data:

60. التحقيقات بدأت منذ أيام

?a-taHqiqaat	bada?at	munðu	?ayaam	(MSA)
the-investigations-**3mp**	started-**3fs**	since	days	

'The investigations started days ago'

61. الأيام مرت بسرعة

    ?il-ayaam   marrat   ?ibsir9ah   (CGA)

    the-days-**3mp** passed-**3fs**   quickly

    'Days have gone by quickly'

62. الأيام بتفوت بسرعة

    ?il-ayaam   bitfwut   ?ibsur9ah   (CEA)

    the-days-**3mp** pass-**3fs**   quickly

    'Days pass quickly'

63. مشاكل هالحياة ما بتخلص

    mašaakil   hal-Hayat   maa   btixlaS   (CLA)

    problems-**3fp** this-life   Neg   finish-**3fs**

    'The problems of this life do not end'

One common feature between the subjects in these examples is that they all have non-human referents. On the other hand, the subjects in examples (56)-(59) have human referents; hence, the verb in those examples takes the plural inflection in agreement with the plural feature of the subject. Therefore, CA varieties in general show similar patterns of subject-verb agreement in SV order: only when the plural pre-verbal subject has human reference, does it agree with the verb in number; but if it has non-human reference, verb agreement in number with the plural form of the preceding subject is lost in all the Arabic varieties; standard and non-standard.

Having described generally the common pattern for subject-verb agreement in CA varieties in comparison to MSA, Brustad (1991) notes some exceptions to this description in sentences with pre-verbal human collective nouns in some colloquial varieties. As explained for MSA in section 2.1.3, when the order is SV with a pre-verbal collective noun, some CA varieties allow the verb not to have plural agreement with the subject (even when it has a human referent); plural agreement and singular agreement

become available options. Compare the following two examples which Brustad

provided from Syrian Colloquial Arabic data:[26]

64. فيه يعني ناس بتفكر إنه شغلة عادية

fii	ya9ni	naas	bitfakkar		?inn-u	šaGla	9aadiyya
there	mean	**people**	think-3sf		that-it	thing	normal

'There are people who think that it is a normal thing'

65. بعرف ناس بيسافروا كل سنة

ba9rif		naas	bisaafr-wu	kil	sinih
(I) know		**people**	travel-3pm	every	year

'I know people who travel every year'

(Brustad, 1991:31)

Brustad (1991:29), following Cowell (1964), argues that the 'singular agreement' option

becomes available only when the pre-verbal human collective has a general reference

rather than a heterogeneous or specific one. In example (64), the subject ناس naas

'people' refers to a homogeneous, generic, abstract group, thus, the following verb takes

the singular form. In contrast, the subject ناس naas 'people' in (65) refers to people who

are known to the speaker, and thus, the verb takes the plural form to agree with the

plural number of the members of that specific group. Regardless of whether this

explanation of the contrast between (64) and (65) is correct, the distinction in agreement

between collective versus non-collective plural subjects may not be present in other CA

varieties data.[27] Therefore, the data of all the CA varieties were examined for such a

distinction. Speakers of any variety which does not allow the 'no agreement' option with

pre-verbal collective nouns may continue to use only the 'agreement' option when

---

[26] The transliteration of these examples and any others taken from external source is reformulated to match the transliteration scheme followed in this work.
[27] Based on the researcher's native intuition, SVO sentences that start with a human collective noun and have no plural agreement between the collective noun and the following verb are not possible in gulf colloquial varieties of Arabic. This will be verified by the results of the current colloquial Arabic corpus study.

speaking MSA; thus, a difference of this type between the CA varieties is interesting to look for.

1.2.1 Data Coding & Results

The data used to examine subject-verb agreement was the same data extracted to examine word order variation. All the coded sentences in that data were with overt subjects and verbs in various orders; nonverbal sentences and those with a null subject were not included in the data. Having both subjects and verbs overt is important to examine subject-verb agreement. These sentences were coded to indicate the number features of the subject, i.e., whether the subject was singular or plural.

For the purpose of examining subject-verb agreement, only sentences in SV order which have 3rd person plural subjects were coded for such agreement since these are the only ones where such variation could be theoretically possible. Unfortunately, because the data under study is of a conversational type, most of the sentences had singular subjects in 1st or 2[ed] person; speakers in a telephone conversation usually talk to each other about each other and rarely about a group of absent people. This made it difficult to find instances with 3rd person plural subjects, and therefore, the relevant tokens found in the data were relatively few. The data contained a total of 34 relevant sentences in CLA, 57 in CGA, and 67 in CEA. Table 4 below shows the general distribution of plural subject-verb 'agreement' and 'no agreement' in these sentences, regardless of the type of the subject (or the pre-verbal NP).

CA varieties Number Agreement?	CLA	CGA	CEA
NO	13 (38%)	2 (4%)	10 (15%)
YES	21 (62%)	55 (96%)	57 (85%)
Total	34	57	67

Table 4 : The general distribution of plural subject-verb agreement in CA varieties

As shown in Table 4 , it is confirmed that the verb does not have to agree with the plural subject in SV order. This is evident by having instances of SV sentences with no agreement between the plural subject and the verb in all the CA varieties. However, it is important to take into account the type of the subject to determine whether or not CA varieties have different agreement patterns. Specifically, we need first to check whether the subject in those sentences is human or non-human, and if it is of human reference, whether it is a collective noun or not. All the sentences were re-checked for the type of subject and these specific details. The results are shown in Table 5 below.

Subject Type	Human				Non-human		Pronouns		Total
	Collective		Non-collective						
Agreement?	Yes	No	Yes	No	Yes	No	Yes	No	
CLA	2	6	7	0	0	7	12	0	34
CGA	10	0	29	0	0	2	16	0	57
CEA	8	5	14	0	0	5	35	0	67

Table 5: The distribution of verbal plural agreement with different types of plural subjects in SV word order

112

Despite the small number of tokens, the results in Table 5 confirm that the verb does not agree with a non-human plural subject in all the varieties. Also, all the varieties show the same agreement pattern with plural pronouns: agreement always occurs in such a case. However, the results indicate that only CEA and CLA allowed 'no agreement' with human plural subjects. In CEA, 5 instances of no agreement were with a subject of human reference. All of these 5 instances have the word ناس naas 'people', which has a collective meaning, as the subject of the sentence. Similarly in CLA, 6 sentences with a plural human subject did not show subject-verb agreement. The subjects in these sentences also denote collectivity. To have examples, the following two sentences contain subjects that signify the collective meaning of people; ناس naas and عالم 9aalam. The following examples are from CLA and CEA data:

66. الناس هنا بتخاف

    ?in-naas   hina   bitxaaf   (CEA)
    The-people   here   be scared-3fs
    'People here get scared'

67. العالم كلها صارت تبكي

    ?il-9aalam   kullaha   Saarat   tibkii   (CLA)
    the-world   all-of-it   became-3fs   cry-3fs
    'All people started crying'

The same type of subject was also found in sentences with subject-verb agreement in all the CA varieties. Compare the examples in (66) and (67) with the following ones from the same CA varieties; CEA and CLA, respectively:

68. خد هدى معاك تشوف الناس والناس يشوفوها

    xud   Huda   ma9-aak   tishwuf   ?in-naas   wi   n-naas yišwuf-wu-ha
    take   Huda   with-you to see-3sf the-people and people see-3mp-her
    'Take Huda with you, so that she sees people and people see her'

العالم اتطوروا واحنا لسا محل ما احنا 69.

?il-9aalam	iTawar-wu		wi	?iHnaa	lissa	maHal	maa	?iHnaa
the-world	developed-**3mp**		and	we	still	place	where	we

'Other people have improved and we are still where we are'

Colloquial Gulf Arabic data also contained subjects of collective type. However, this variety does not seem to differentiate between subjects denoting collectivity and those which denote individuation; all plural subjects of human reference were found agreeing with the following verb which has a plural inflection. The following examples are sentences from CGA data which used the word ناس naas as the subject:

الناس هنا بيضحكون عليه 70.

?an-naas	hina	bi-yiDHak-wun	9alayh
The-people	here	will-laugh-**3mp**	on-him

'People here will laugh at him'

الناس ما انطونا قميص 71.

?an-naas	maa	?anTa-wu-na	qamiiS
The-people	Neg	gave-**3pm**-us	(a) shirt

'The people did not give us a shirt'

These two examples and others which have plural subjects denoting collectivity always take verbs with plural inflections in CGA; the type of distinction found in CEA and CLA was not found in CGA data.

To summarize, the agreement pattern was found to be the same in SV sentences in all the Arabic varieties when the plural subject is inanimate or non-human: the verb usually takes the singular form at all times. However, when the plural subject is of human reference, CA varieties seem to have different agreement patterns. CEA and CLA appear to differentiate between subjects with collective meaning from those with non-collective meaning. Sentences with subjects of the latter type always have agreement,

whereas sentences with subjects of the former type do not necessarily show subject-verb agreement; the verb in such sentences may take a singular feminine form. CGA, on the other hand, does not differentiate between subjects of human reference; all sentences with subjects of this type show subject-verb agreement.

This difference in verbal agreement with collective nouns in preverbal position will be taken as a second variable for examining MSA. If L1 different grammars have an effect on the way MSA grammar is represented by Arab speakers; CLA and CEA speakers may be expected to continue following the same pattern of subject-verb agreement in their L1 when speaking MSA. CGA speakers, on the other hand, may be expected not to make distinctions between human subjects; they should show full agreement between the verb and human subjects of all types. The following section will present analysis of the data for the purpose of determining differences between the CA varieties in relation to resumption in object relative clauses. If such differences exist, these will be used as a third variable when examining MSA.

1.3 Resumption in Object Relative Clauses

The term 'Object relatives' here refers to restrictive definite relatives which are formed by extraction from the direct object position. As has been explained before (chapter 2 section 2.3), MSA allows two strategies to form this type of relative: the resumption strategy and the movement strategy; gaps alternate with resumptive pronouns to fill in the extraction site. All the CA varieties under study use the resumption strategy to form this type of relative, but they seem to differ in whether or not they allow gaps to fill in the extraction site as well. Shlonsky (1992) pointed out that resumptive pronouns are necessary for object relatives in Colloquial Palestinian Arabic to be grammatical; gaps are not licit in this variety of Arabic. Similarly, Aoun et al. (2010) asserted that gaps are

not allowed in object relatives in Colloquial Lebanese Arabic either. However, other CA varieties are not like Palestinian or Lebanese Arabic; some colloquial varieties allow gaps and resumptive pronouns alternately to fill in the extraction site in object relatives (Alresaini, 2007).[28] Therefore, the data of all the CA varieties were examined for resumption in object relatives. The speakers of those varieties which use the resumption strategy exclusively to form object relatives may continue to avoid relatives with gaps in MSA. Thus, a difference of this type between the CA varieties is interesting to look for.

1.3.1 Data & Results

The data examined for resumption were extracted by CLAN from the set of LDC corpora described above. All the CA varieties form the definite restrictive relatives by using one relativiser, and that is the word اللي *illi* 'that'. Therefore, the relevant data was extracted by instructing 'COMBO' to search for sentences which contain this relativiser in a 100 files of transcribed unscripted data of telephone conversations.

The output data contained a total of 271 relatives in CLA, 323 relatives in CGA, and 778 relatives in CEA of different extraction positions. This data was coded for the grammatical position from which relativisation has taken place and whether or not the relative clause contained a resumptive pronoun.

Out of the overall number of relatives found in the data, object relatives occurred 61 times in CLA, 74 times in CGA, and 314 times in CEA data. The following table shows

---

[28] Alresaini (2007) looked at resumption in relative clauses in Qassimi dialect (one of the local dialects in Saudi Arabia) and pointed out that resumption is optional in object relatives.

the distribution of resumptive pronouns in object relatives in the CA varieties under study.

Resumption? CA Varieties	Yes	No
CLA	59	2
CGA	66	8
CEA	314	0

Table 6 : The distribution of resumption in object relatives in CA varieties

The results of this particular study are straightforward. Using resumptive pronouns seems to be preferred in all the CA varieties when forming object relatives. However, speakers of CGA did not use a resumptive pronoun in 8 of the 74 object relatives they produced. This indicates that having a gap replacing the resumptive pronoun is possible and grammatical in this variety. On the other hand, speakers of CLA formed only two object relatives without resumptive pronouns out of 61. In both instances, the position where the resumptive pronoun should have occurred was at the end of the utterance. So, it is very much possible that the speakers who formed these two relatives were interrupted by the other participant in the conversation. Furthermore, Shlonsky's (1992) and Aoun et al.'s (2010) claims that colloquial Palestinian and Lebanese Arabic, which are Levantine dialects, do not allow such relatives, would support the proposal that these two relatives were not produced in their complete forms, but possibly something prevented the speakers from completing their utterances; speech overlapping is a strong possibility. In contrast to CGA and CLA, the results strongly suggest that CEA does not allow gaps to replace resumptive pronouns; in all of the 314 instances of object relatives, the speakers of this variety used resumptive pronouns only all the time.

These results lead to the conclusion that the CA varieties differ in whether they allow gaps to replace resumptive pronouns in object relatives. A clear contrast is found between CGA and CEA in relation to resumption in object relatives; resumptive pronouns are crucial in the latter whereas they can be replaced by gaps in the former. Since CGA allows gaps in such relatives, the speakers of this variety may not have a problem with MSA relatives which are formed with gaps instead of resumptive pronouns. On the other hand, speakers of CEA may consider MSA relatives of this type as ungrammatical. Finally, CLA may be considered like CEA; speakers of this variety may find object relatives with gaps ungrammatical. This is indicated by their tendency to use resumptive pronouns in almost all the relatives they produced. Thus, these differences will form the third variable when testing MSA grammar for L1 influence. The following section will present analysis of the data in relation to placement of adverbs in CA varieties.

1.4 Adverb Placement

The idea behind choosing the issue of adverb placement as a candidate to reveal differences between the CA varieties is that placement of adverbs seems to be flexible in all the varieties of Arabic including MSA. This means that it is possible that each variety may have its distinct pattern of preferring one place over the others. The speakers of those varieties may also continue to keep the same preference patterns when speaking MSA.

Al-Shurafa's analysis (2005) of adverb placement in Colloquial Hijazi Arabic (one of the CGA dialects), Palestinian Arabic (one of the CLA dialects), and MSA, restricts the number of syntactic positions where adverbs can occur to four places depending on the semantic function and scope of the adverb in use: those modifying the verb are adjoined

to VP and those specifying the action of the whole sentence are adjoined to S projection; adjunction in both cases can be to the right or the left side of the syntactic projection. Al-Shurafa (2005:85) states that her analysis to a certain extent may explain adverb placement in all varieties of Arabic .[29] Therefore, the data of all the CA varieties under study were analysed to determine if there is a certain pattern of placing adverbs in each variety and if that pattern differs among these varieties. The speakers of these varieties may keep their distinct patterns of preferring one or more particular positions when speaking MSA.

1.4.1 Data Extraction & Coding

The data for this analysis was extracted from the LDC corpora by using CLAN. The application 'COMBO' was given lists of common adverbs of manner, place, and time to search for in the corpora. The long LDC lists of complete lexical words that were actually used by speakers of CLA and CGA were checked and 50 adverbs were selected to develop the lists for searching the data. Since 'CALLHOME' Egyptian corpus did not come with a file that lists lexical items used by Egyptian, the same list developed from the other two corpora was adapted to search the CEA data. However, some adverbs were modified where necessary to match the way they are used by Egyptians. An example of this case is the adverb of time اليوم ?il-yuwm 'today' from CGA and CLA data was modified to النهاردة ?in-nahaardah in the list of adverbs adapted for CEA. Again, a native speaker of CEA was asked to check the final list of adverbs adopted for searching CEA data to make sure that adverbs are in the correct form that would be used by native Egyptians (see Appendix 2 for the complete lists of adverbs).

---

[29] See chapter 2 (section 2.2) for a background description of adverb placement in MSA and more details about Al-Shurafa's (2005) analysis of adverbs.

'COMBO' was instructed to search for sentences which contain these adverbs in a 100 files of transcribed unscripted conversations in each corpus.[30] The sentences in the output data were coded for word order, adverb type, and place of adverb. Word order was controlled for to see if the pattern of placement changes when the word order of the sentence is changed. Similarly, the type of adverb was coded for to check if certain types of adverbs are preferred in certain places. The place of the adverb was determined according to its function and scope in the sentence. To clarify this further, if the adverb modifies the whole sentence (S-adverb) and occurred at the beginning of the sentence, that place is classified as 'pre-sentential'. If the adverb specifies the action of the whole sentence (S-adverb) and occurred at the end of the sentence, that place is classified as 'post-sentential'. The following two sentences from CGA data exemplify these two situations, respectively:

72. Pre-sentential:

أصلاً أنا متنت

?aSlaan   ?anaa   mitant
basically   I   gained weight
'Basically, I gained weight'

— S-adverbs

73. Post-sentential:

أنا مازرتها صراحة

?anaa maa   zirta-haa   SaraaHah
I   Neg   visited-her   frankly
'Frankly, I didn't visit her'
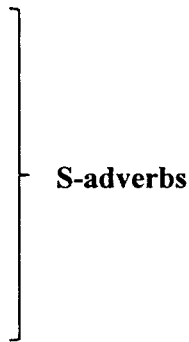
The adverbs in these two sentences did not modify the action of 'gaining weight' or 'visiting', but their scope was wider to modify the whole sentence. Such S-adverbs are assumed to be adjoined to the highest projection of the structure; either to its left side, as in (72), and hence, that position is called 'pre-sentential'; or to its right side, as in (73), and hence, that position is called 'post-sentential'. However, if the adverb modifies the

---

[30] Again, this particular number of files was not determined for a particular reason, except that it is available in all three corpora from which the data is extracted.

verb of the sentence (VP-adverb) and occurred before it, that place is called 'pre-verbal'.

Adverbs of this type can come at the beginning of the sentence in cases where the order

is VS, and between the subject and the verb if the order is SV. Finally, if the adverb is

placed after the verb and it is of the verb-modifying type, that place is referred to as

'post-verbal'. These adverbs can come following the verb or at the end of the sentence.

The following sentences are examples of these situations from CLA data:

74. Pre-verbal:

   a. دغري ركضت على الباب وفتحت

     diGrii   rkiDit 9a l-baab u   fataHit
     directly ran-1s to the-door and opened-1s
     'I ran directly to the door and opened it'

   b. أنا كتير بحب السينما

     ?anaa **ktiir** bHib l-sinamaa
     I   a lot like   the-cinema
     'I like the cinema a lot'

75. Post-verbal:

   c. بتعرفي كتير باللبس الحلو

     bti9rif-iy   **ktiir**   bi-l-libs   l-Hiluw
     (you) know-2fs a lot   on-the-clothing the-nice
     'You know a lot about nice clothing'

   d. السعر بينزل بسرعة

     ?isi9ir   biyinzil   **bsir9ah**
     the-price   drop   quickly
     'The price drops quickly'

VP-adverbs

The scope of adverbs in (74) and (75) is limited to modifying the verb. In (74), the

adverbs دغري diGrii 'directly' and كتير ktiir 'a lot' modify the manner of 'running' and

degree of 'liking'. Similarly, in (75), كتير ktiir 'a lot' and بسرعة bsir9ah 'quickly' modify

the degree of 'knowing' and manner of 'dropping'. These VP-adverbs are assumed to be

adjoined to VP; either to its left side, as in (74), and hence, that position is called 'pre-

verbal'; or to its right side, as in (75), and hence, that position is called 'post-verbal'.

The pre-verbal adverb in (74a) occurred initially because the sentence has a null subject, whereas in (74b), it occurred between the pre-verbal NP and the verb. In (75a), the post-verbal adverb occurred following the verb, whereas in (75b), it occurred finally.

1.4.2 Results

The data examined for adverb placement contained a total of 127 sentences in CLA, 210 in CGA, and 275 sentences in CEA. Table 7 below compares the general distribution of adverbs in four possible places in each CA variety: pre-sentential, pre-verbal, post-verbal, and post-sentential positions; distribution of adverb placement in sentences with different word orders will be given thereafter.
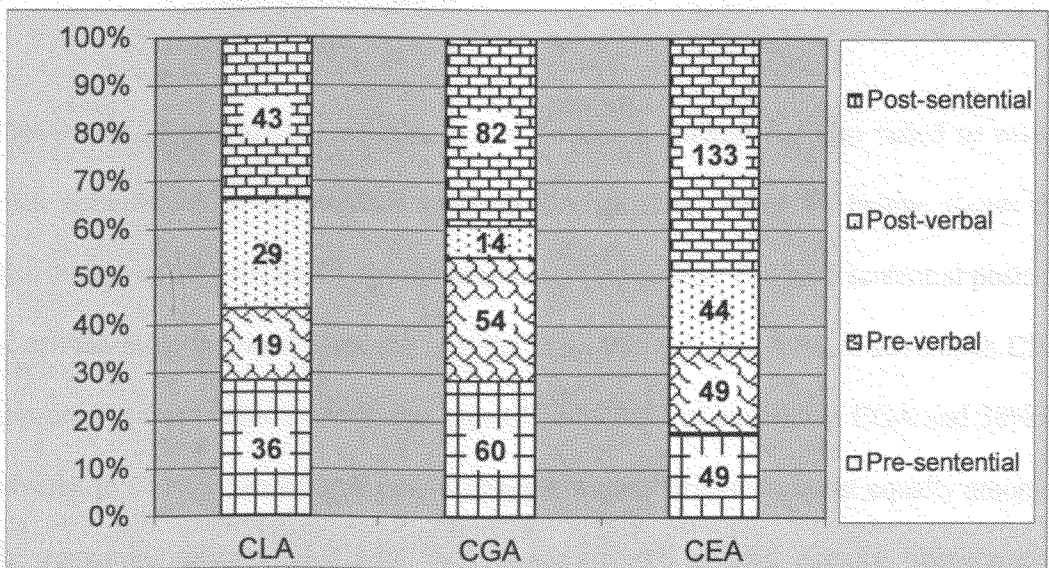
Adverb Place	Pre-sentential	Pre-verbal	Post-verbal	Post-Sentential
CA Varieties	%	%	%	%
CLA	28	15	23	34
CGA	28	26	7	39
CEA	18	18	16	48

Table 7 : The general distribution of adverb placement in CA varieties

All the CA varieties seem to show similar patterns when it comes to favouring one place over the others. In all the CA varieties, the post-sentential position is the most favoured

place followed by the pre-sentential position. However, the distribution of adverbs in pre-verbal and post-verbal positions is slightly different in CA varieties. In CLA, 34% of adverbs occurred in post-sentential position and 28% in pre-sentential position. In this variety, post-verbal is preferred more than pre-verbal; 23% of adverbs occurred in the former and only 15% occurred in the latter position. CGA is like CLA in favouring post-sentential followed by pre-sentential places for positioning adverbs; 39% of adverbs occurred post-sententially and 28% appeared pre-sententially. Unlike CLA, though, more adverbs in CGA appeared in pre-verbal (26%) than in post-verbal places (7%). In CEA, about half of the adverbs (48%) occupied post-sentential positions; the rest of the adverbs were distributed almost equally amongst the remaining three places.

The results of considering adverbs in SV sentences exclusively also failed to reveal clear patterns which differentiate between CA varieties. Table 8 , below, shows the distribution of adverbs in SV sentences in each CA variety. The post-sentential position is still the most favoured place for adverbs in all the varieties. 49% of adverbs in CEA appeared in this position which also was filled by 39% of adverbs in CGA and 38% of adverbs in CLA. However, the rest of adverbs were distributed almost equally amongst the remaining places. CLA had 40% of its adverbs divided equally between pre-sentential and post-verbal positions and 22% were placed in pre-verbal positions. CEA is not that different from CLA. 19% of adverbs in this variety were placed in pre-verbal, 17% in post-verbal and 16% in pre-sentential positions. In contrast, few adverbs in CGA were placed in post-verbal position (3%); the rest were placed in pre-verbal (31%) and pre-sentential (27%) positions.
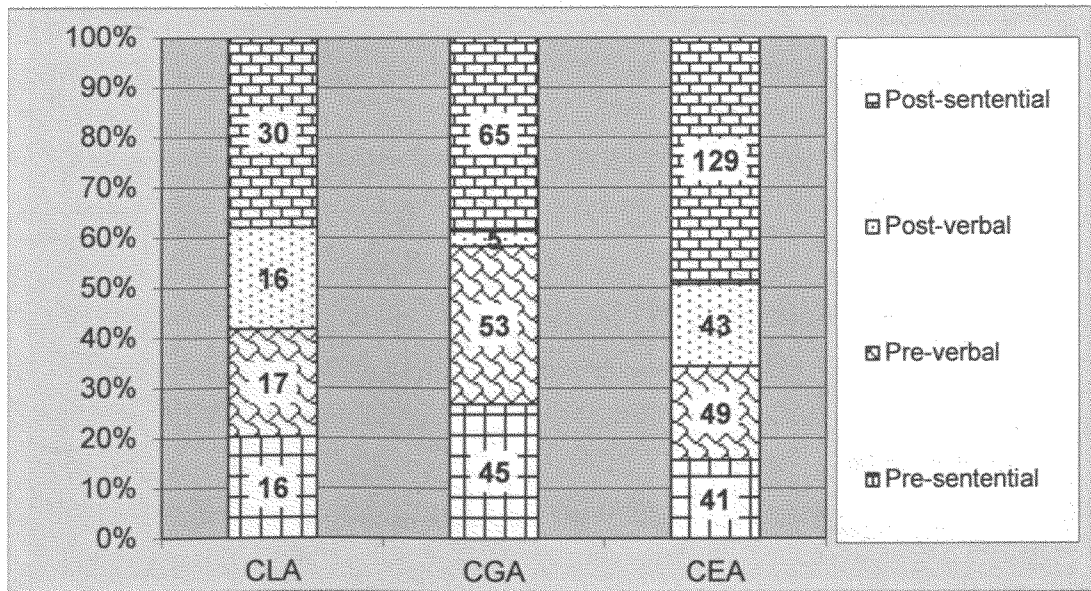
Adverb Place	Pre-sentential	Pre-verbal	Post-verbal	Post-sentential
CA Varieties	%	%	%	%
CLA	20	22	20	38
CGA	27	31	3	39
CEA	16	19	17	49

Table 8 : The distribution of adverb placement in CA varieties when the sentence order is SV

The distribution of adverbs is quite different in VS sentences from that in SV sentences, yet, CA varieties still show similar patterns. In general, all the CA varieties have fewer or no adverbs in pre-verbal positions, few adverbs in post-verbal positions, with the majority divided between pre-sentential and post-sentential positions (see Table 9 below). In CLA, more adverbs were placed pre-sententially (42%) and two adverbs only were placed pre-verbally; post-verbal and post-sentential positions shared the rest of adverbs equally. In CEA, VS sentences which contain adverbs are quite few. All adverbs in VS sentences were placed pre-sententially (8 adverbs) or post-sententially (6 adverbs) except for one adverb which was placed post-verbally; the speakers of CEA did not use the pre-verbal position when the order was VS. In CGA, the speakers placed 17 adverbs post-sententially, 15 pre-sententially, 9 post-verbally, with only 1 adverb in pre-verbal position.
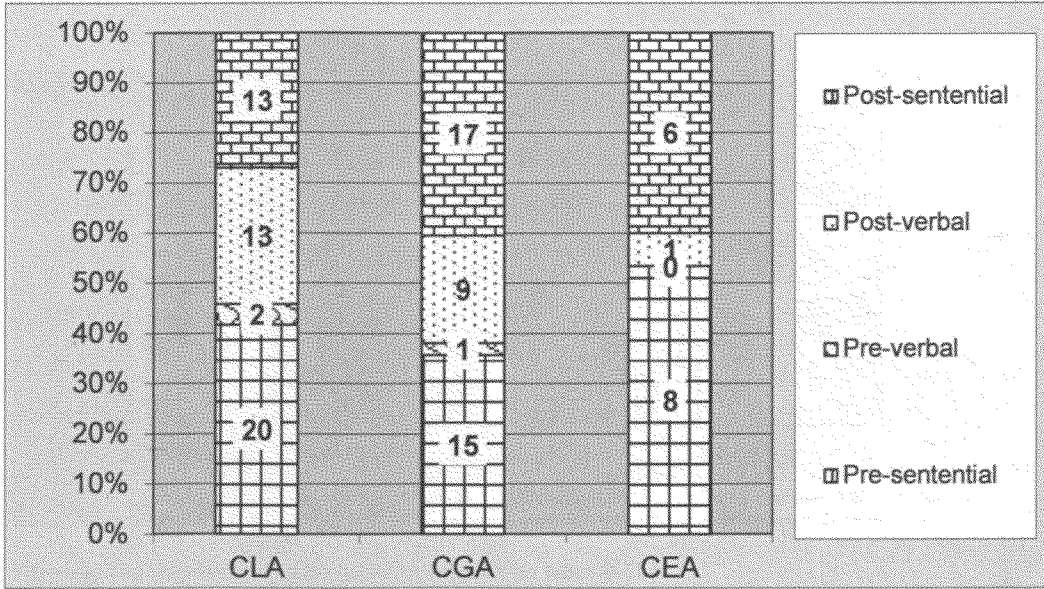
Adverb Place	Pre-sentential	Pre-verbal	Post-verbal	Post-sentential
CA Varieties	%	%	%	%
CLA	42	4	27	27
CGA	36	2	21	41
CEA	53	0	7	40

Table 9 : The distribution of adverb placement in CA varieties when the sentence order is VS

The data also were examined for whether certain types of adverbs are placed in certain places and whether that would reveal differences between the CA varieties. In general, the results do not show clear different patterns of distribution in the CA varieties examined (see Figures 1-3 below). Temporal adverbs were found everywhere in all the varieties with more adverbs occurring pre-sententially in CLA and CGA and post-sententially in CEA. Adverbs of place clearly prefer post-sentential positions in all the CA varieties. Finally, adverbs of manner occurred in all the possible places with slightly greater numbers in post-verbal positions in CLA and CEA and in post-sentential positions in CGA.
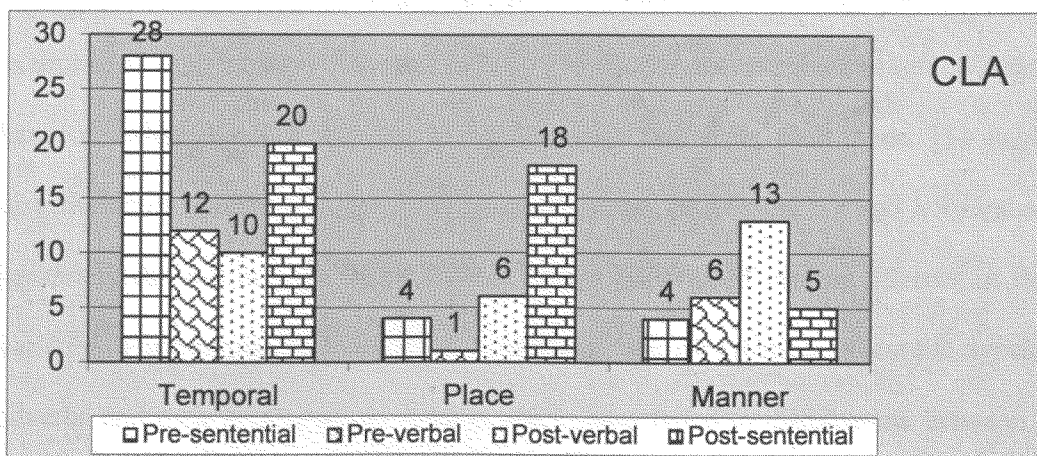
Figure 1 : The distribution of adverbs according to their types in CLA
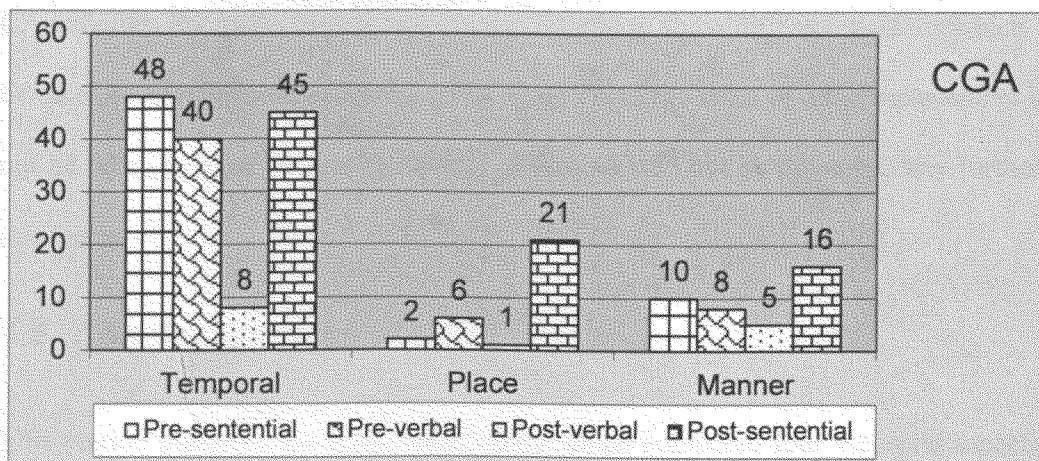
Figure 2 : The distribution of adverbs according to their types in CGA
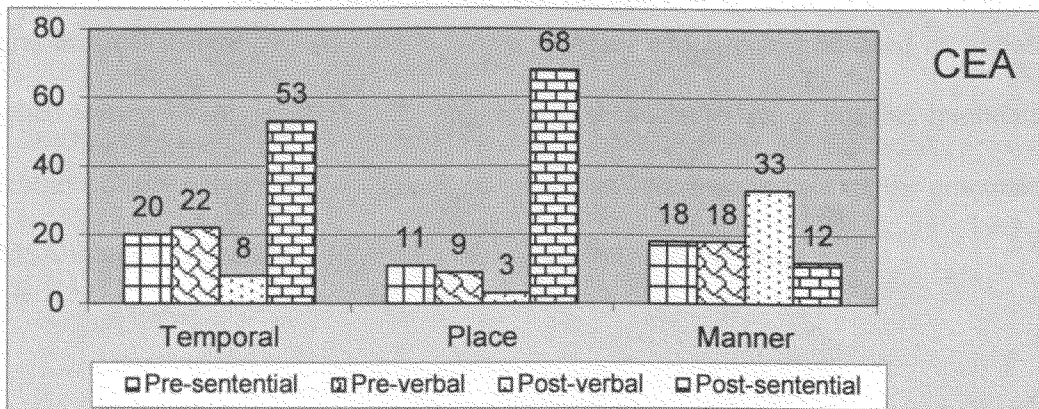
Figure 3 : The distribution of adverbs according to their types in CEA

To sum up, the data was analysed to determine if there is a certain pattern of placing adverbs in each CA variety. The frequency of occurrence in four possible places was examined in each variety. Unfortunately, the results did not reveal important differences based on which the CA varieties can be distinguished from each other. The general distribution of adverbs was similar in all the varieties. Examining SV and VS sentences separately did not help to reveal clear different patterns of placing adverbs in these varieties. Furthermore, the results of examining the distribution of certain types of adverbs and their preference for certain places also failed to differentiate between the CA varieties. Although the results of placing adverbs are not exactly the same across the CA varieties, the differences found are not enough to show different patterns which can be taken as variables when testing MSA data. Thus, the issue of adverb placement will not be taken as a variable when testing MSA and will not be discussed any further in this thesis.

1.5 Conclusion

This section presented a corpus based study which compared four syntactic aspects in CLA, CGA, and CEA. The main purpose of conducting this study was to find grammatical differences between the CA varieties which could be used as variables when testing MSA end-state grammar. This is to test whether or not the grammatical features of the unequivocally L1s (the CA varieties) have an impact on the grammar of MSA that is produced by native speakers of different CA varieties.

The syntactic aspects which this study investigated were word order, 3rd person plural subject-verb agreement, resumption in object relatives, and adverb placement. These particular issues were chosen to be studied as they were thought of as good candidates

to reveal grammatical differences of the type which is not salient enough for the speakers to avoid when speaking MSA.

Word order was chosen to be studied to test if speakers of different CA varieties follow different patterns of preferring certain orders. Because all possible word orders in CA varieties are allowed in MSA, speakers of CA varieties may continue using their native patterns when speaking MSA. The results of this study showed that CEA has a distinct pattern from the other two CA varieties. The speakers of CEA were found to use SVO persistently regardless of discourse type. In contrast, word order pattern in CLA and CGA was found sensitive to discourse type despite the few instances of narrative context. The speakers of these two varieties tended to use SVO more than VSO in conversational discourse, whereas in narrative discourse, VSO became the common order. This means that if speakers use MSA as a second language that is influenced by the L1, we can expect at least the speakers of CEA to use SVO most frequently than the speakers of the other two varieties when they speak MSA.

The results of studying 3rd person plural subject-verb agreement in SV order also indicated that CA varieties are different in this regard when the subject has a human referent. CLA and CEA were found to differentiate between collective and non-collective plural subjects; sentences with subjects of the latter type were found to have agreement all the time, whereas sentences with collective plural subjects did not necessarily show subject-verb agreement in SV order; the verb in such sentences may take the singular feminine inflection regardless of the plural feature of the subject. In contrast, CGA did not show this kind of distinction between plural subjects of human

reference; all the sentences with this type of subjects were found showing full subject-verb agreement. This leads us to conclude that if L1 influence exists in MSA acquisition, the speakers of CLA and CEA may be expected to continue producing SV sentences of the type specified above with no subject-verb agreement; whereas CGA speakers may reject sentences with no agreement between the collective subject and the verb.

By studying resumption in object relatives, a clear contrast was found between CGA and CEA. Resumption was found to be crucial in CEA whereas it can be forsaken in CGA. Since CGA allows gaps in such relatives, the speakers of this variety will not have a problem with MSA relatives which are formed with gaps instead of resumptive pronouns. On the other hand, speakers of CEA can be expected to mistakenly consider MSA relatives of this type as ungrammatical. Finally, CLA may be considered like CEA; speakers of this variety may find object relatives with gaps ungrammatical. This is indicated by their tendency to use resumptive pronouns in almost all the relatives they produced.

Finally, the results of examining adverb placement failed to reveal clear different patterns across the CA varieties. Although some differences were noted between the CA varieties, these differences were not enough to declare distinct patterns which can be taken as variables when testing MSA data.

2   Modern Standard Arabic Corpus Study

The previous section presented a corpus based study of four syntactic aspects in three CA varieties. The results of that study showed some differences between the CA varieties in word order, 3rd person plural subject-verb agreement in SV order, and in resumption in object relatives. Studying word order led to discover that CEA is different from the other two CA varieties. In CEA, SVO was found, being used dominantly regardless of the context type. This made CEA distinct from CLA and CGA which were found sensitive to the type of discourse. In these two CA varieties, SVO was used dominantly in conversational discourse, whereas in narrative discourse, VSO was used more than SVO.

Furthermore, these three CA varieties were found to show different patterns of 3rd person plural subject-verb agreement in SV order, especially when the subject has a human referent. CGA did not differentiate between plural subjects with human reference; all were found to agree with the following verb in number at all times. CLA and CEA, on the other hand, allowed the 'no agreement' option to occur when such subjects denoted collectivity.

Studying resumption in object relatives also revealed differences between the CA varieties. The results of analysing the data showed that resumption is optional in CGA and obligatory in CEA; the results of CLA were disputable. Finally, adverb placement did not reveal any clear distinct patterns followed by the three CA varieties, and hence,

it was decided that this aspect of the grammar will not be taken as a variable for studying MSA.

This section will present another corpus based study. This study investigated the same syntactic issues but in MSA data. The data was produced in three regions of the Arab world: Levantine countries, Gulf Countries, and Egypt. MSA spoken in these three regions will be referred to here as Standard Levantine Arabic (SLA), Standard Gulf Arabic (SGA), and Standard Egyptian Arabic (SEA), respectively. The aim of this study is to test whether the differences found between the CA varieties also exist to differentiate grammatical representation(s) of MSA across these regions. If such a finding is reached, this indicates that MSA grammar is influenced by the speakers' L1 grammar which leads to the conclusion that the grammar of the standard variety is of an L2 type.

2.1 MSA Data & Data Coding

The data examined in this study were obtained from two sets of corpora which were also developed by the Linguistic Data Consortium (LDC). The first set has the LDC title 'GALE Phase 1 Arabic Broadcast News Parallel Text- Part 1' (Ma and Zakhary, 2007).[31] These corpora contain speech transcripts of 17 hours of Arabic broadcast news selected from six different sources. The sources include Al-Hurra, Voice of America, Dubai TV, Nile TV, and Lebanese Broadcast. The second set of corpora has the title 'GALE Phase 1 Arabic Broadcast News Parallel Text- Part 2' (Ma and Zakhary, 2008).[32] These corpora contain 10.7 hours of speech transcripts of broadcast news taken

---

[31] The LDC catalogue number for this set of corpora is (LDC2007T24).
[32] The LDC catalogue number for this set of corpora is (LDC2008T09).

from a variety of sources. These include Radio Sawa, Voice of America, Dubai TV, and Nile TV.

The data examined in the current study were a part of the above described corpora. Only data obtained from Dubai TV, Nile TV, and Lebanese Broadcast were included in this study to represent SGA, SEA, and SLA, respectively; data from the other sources were excluded because they were not related to any of the regions under study. Dubai TV data contained speech transcripts of 3.9 hours of news broadcasting which were recorded in 2005. Similarly, Lebanese Broadcast data were recorded in 2005, and contained a total of 2.0 hours of transcribed speech. The data obtained from Nile TV were recorded in 2000 and contained speech transcripts of 1.1 hours of news broadcasting.

Because the relevant data were not as much as that examined for CA varieties, CLAN was not needed for extraction this time; all the data in hand was coded for the syntactic issues under investigation. Coding for word order and agreement, however, excluded sentences with dropped subjects or those without verbs. As in the colloquial corpus study, only sentences with verbs and overt subjects were coded in this study for word order, type of context, and whether the subject is singular or plural. Sentences in SV order with plural subjects were coded for subject-verb agreement.

The criterion for coding the context was as follows. Since the data was transcripts of broadcast news which evidently involve narration of events and stories, the general

discourse type was considered to be narrative. Conversational contexts were determined when the presenter hosts a guest to discuss with some of the current events or situations. These contexts were easy to notice plus that they were clearly marked in the transcription.

All definite relative clauses were analysed and coded for the grammatical position from which relativisation has taken place and for whether or not these relatives contained resumptive pronouns.

2.2 Results

2.2.1 Word Order

The coded data contained a total of 172 sentences in SLA, 387 sentences in SGA, and 197 sentences in SEA, all in context. Table 10 below compares the general distribution of word order in the three Standard Arabic (SA) varieties under study; distribution of word order in different contexts will be given thereafter. Based on the analysed data, VSO is the most frequently used order in SLA and SEA. The speakers of these two varieties used this order respectively in 63% and 72% of the sentences which appeared in the data under study. SVO came in the second rank in terms of frequency; it was used in 33% of SLA sentences and in 24% of SEA sentences. In contrast, SVO appeared more than the other orders in the data of SGA; the speakers of this variety used this order in 205 sentences which comprise 53% of the data. VSO came in the second rank and was used in 40% of the total number of SGA sentences. VOS was the least favoured in all the three SA varieties. The speakers of SLA and SEA used this order in 7 sentences only, which comprise 4% of the data. Speakers of SGA used this order in 28

sentences comprising 7% of the data. The other three variations of word order; i.e., OVS, SOV, and OSV, were not used in this data.
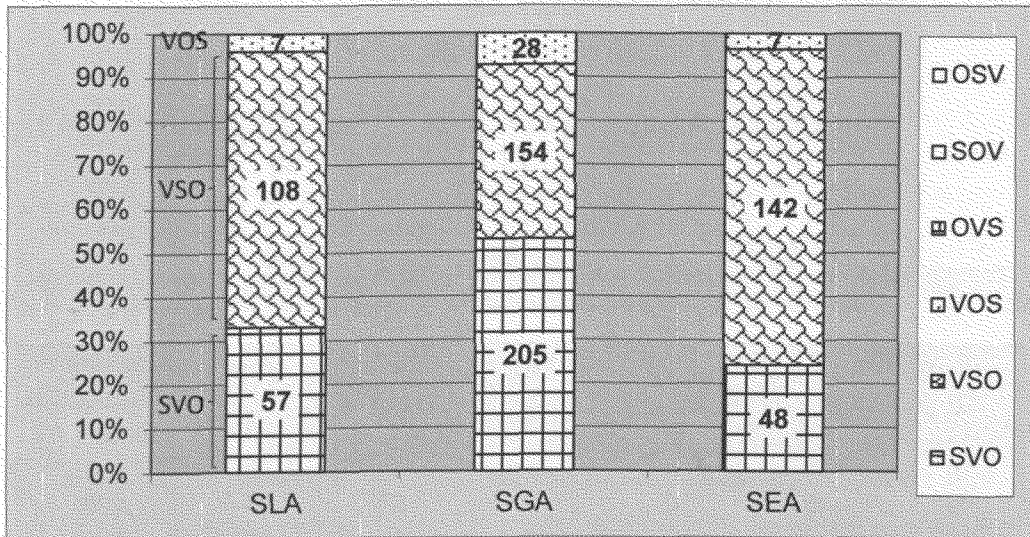
Word Order	SVO	VSO	VOS	OVS	SOV	OSV
SA Varieties	%	%	%	%	%	%
SLA	33	63	4	0	0	0
SGA	53	40	7	0	0	0
SEA	24	72	4	0	0	0

Table 10 : The frequency of using different word orders in SA varieties

Recalling the fact that the data studied here is taken from news broadcast, most of the sentences occurred in narrative discourse. Therefore, the results did not vary that much when the frequency of word orders was considered in narrative discourse only (see Table 11 below). VSO is still dominantly used by the speakers of SLA and SEA, followed by SVO, then VOS. The speakers of SLA and SEA used VSO in 64% and 72%; SVO in 32% and 24%; and VOS in 4% of the analysed sentences, respectively. SGA is still different when the narrative discourse was analysed separately; the speakers of this variety preferred using SVO in 51% of the analysed sentences; VSO was used in 41% and VOS in 7% of the relevant data.
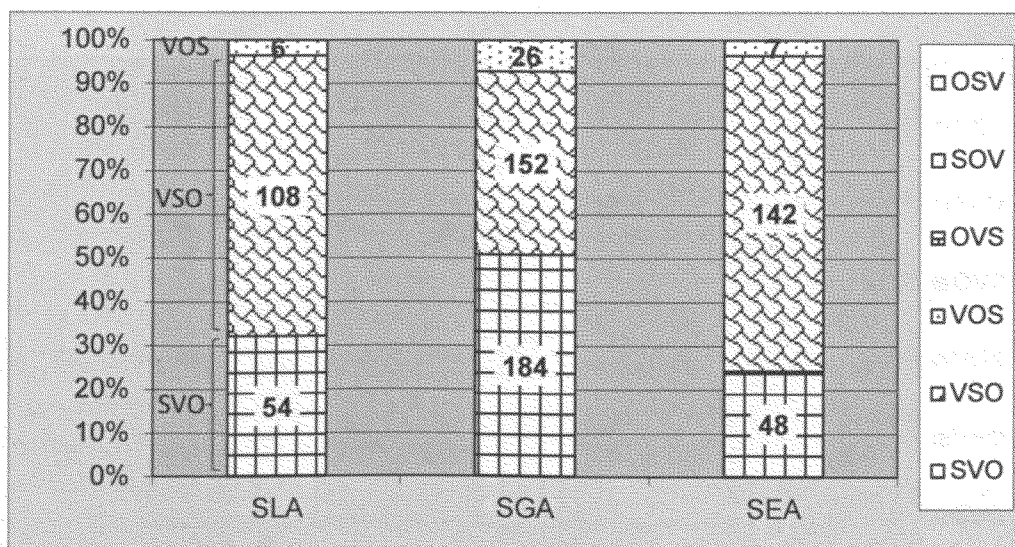
Word Order	SVO	VSO	VOS	OVS	SOV	OSV
SA Varieties	%	%	%	%	%	%
SLA	32	64	4	0	0	0
SGA	51	42	7	0	0	0
SEA	24	72	4	0	0	0

Table 11 : The frequency of using different word orders in 'Narrative Context' in SA varieties

In contrast, very few sentences were found in conversational discourse (see Table 12 below). This is obviously due to the type of the analysed data. Only 4 sentences were found in such contexts in SLA; three of which were in SVO order and the remaining one was in VOS order. The data of SGA contained 25 sentences in this type of discourse. The speakers of this variety used SVO in 21 sentences; VSO and VOS were used in two sentences each. Finally, the data of SEA contained no instances at all of this type of discourse.
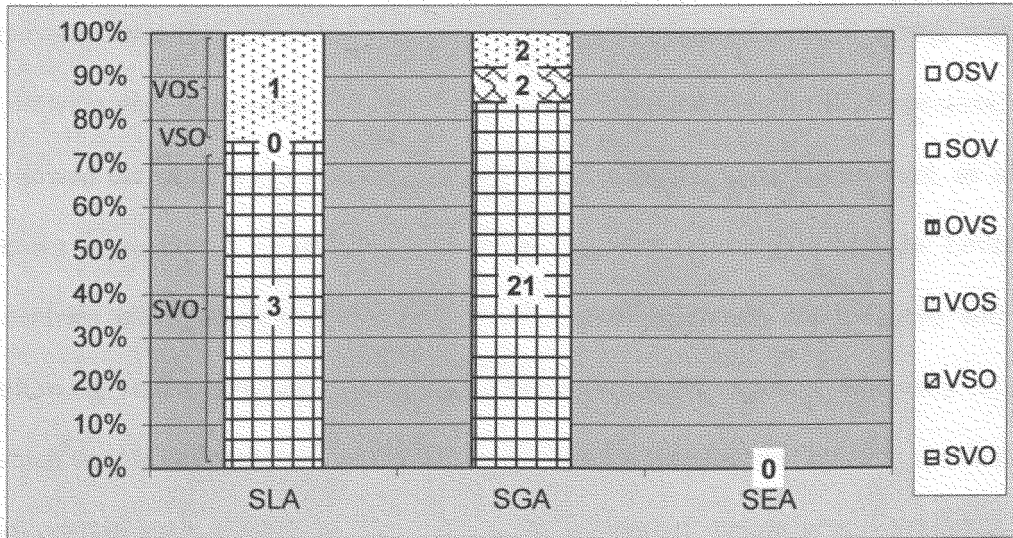
Word Order	SVO	VSO	VOS	OVS	SOV	OSV
SA Varieties	%	%	%	%	%	%
SLA	75	0	25	0	0	0
SGA	84	8	8	0	0	0
SEA	0	0	0	0	0	0

Table 12 : The frequency of using different word orders in 'Conversational Context' in SA varieties

2.2.2 Plural Subject-Verb Agreement

The analysed data of all the SA varieties contained only 65 sentences in SV order which have 3rd person plural subjects. Only 2 of these sentences were found in SLA data, 11 in SEA data, and 52 in SGA data. Table 13 below compares the SA varieties in relation to subject-verb plural agreement in such sentences.

Agreement?	Yes		No		Total
Subject Type	Human	Non-human	Human	Non-Human	
SLA	0	0	0	2	2
SEA	4	0	0	7	11
SGA	29	0	0	23	52

Table 13: Subject-verb plural agreement in SV sentences in SA varieties

Based on these results, one common pattern of subject-verb agreement appears to be that whenever the pre-verbal plural subject has a human referent, it agrees with the following verb in number, and vice versa. This is not surprising for SGA speakers; they followed the same pattern when speaking their colloquial variety. However, as it was shown in section 1.2.1 of this chapter, the speakers from Egypt and the Levantine countries allowed some plural subjects of human reference not to agree with the following verb when they were speaking their colloquial varieties, especially subjects which denote collectivity. Unfortunately, the MSA data produced by such speakers did not contain such subjects to check if the 'no agreement' option is still available. The data of SEA contained only 4 sentences with human subjects, and none of these subjects denoted collectivity. The data of SLA did not contain sentences with plural human subjects. Finally, there were 3 sentences in the data of SGA which have human subjects denoting collectivity; all of which agreed with the following verb in number. To have an example, one of these sentences is the following:

76. السنة العرب سيشكلون دائرة القضبان

?as-sunnat-u   l-9arab   sa-yušakil-uwna   daa?irat-a   l-quDbaan
the-Sunni-nom   the-Arab   will-form-3mp   circle-acc   the-bars
'The Sunni Arabs will form the circle of bars (to prevent going over the limit)'

2.2.3 Resumption in Object Relative Clauses

The data examined for resumption contained a total of 115 relative clauses in SLA, 235 relatives in SGA, and 92 relatives in SEA of different extraction positions. This data was coded for the grammatical position from which relativisation has taken place and whether or not these relatives contained a resumptive pronoun. Object relatives occurred 34 times in the data of SLA, 64 times in the data of SGA, and 29 times in the data of

SEA. The following table compares the SA varieties in terms of whether object relatives contained resumptive pronouns.

Resumption? SA Varieties	Yes	No
SLA	30	4
SGA	62	2
SEA	29	0

Table 14 : the distribution of resumption in object relatives in SA varieties

The results in Table 14 show that resumptive pronouns were generally preferred over gaps in object relatives in all SA varieties under study. Speakers of SLA used a gap in 4 object relatives out of 34 whereas speakers of SGA used gaps only twice out of 64 tokens. The most interesting result here is that Egyptian speakers of MSA continued to use resumptive pronouns in all of the 29 object relatives they produced; they never used a gap despite the fact that, as detailed in chapter 2, section 2.3, MSA grammar allows them in object relative. However, ideally more instances of object relatives in SEA are needed to confirm that this is due to L1 influence.

2.3 Discussion & Conclusion

The analysis of the current data of the MSA corpus study does not seem to provide clear-cut findings about whether the speakers of Standard Arabic (SA) varieties represent different grammars of MSA due to the grammatical differences between the CA varieties they speak as L1s. This is because of some limitations that are related to the type and amount of the examined data. This section will first discuss the results of

the MSA corpus study (presented in section 2 of this chapter) in light of the findings of the colloquial corpus study (presented in section 1), and present the limitations of the MSA corpus study.

The results of examining word order in the colloquial corpus study indicated that the speakers of CGA generally used SVO order more than VSO or VOS. They used SVO in 75% of the sentences they produced in conversational contexts whereas only 3 sentences were found in narrative context and they were all in VSO order. The speakers of this variety apparently continued to favour SVO over the other two orders when they spoke MSA despite the fact that VSO has been claimed to be the common unmarked order in MSA (see chapter 2, section 2.1.1). The current analysis of MSA corpus data showed that more than half of the sentences in the SGA data were in SVO order regardless of type of discourse. This may be taken to suggest that the grammar of MSA represented by the speakers of L1 CGA is influenced by the grammar of CGA.

Like CGA, CLA speakers used SVO order to form most of the sentences in their colloquial variety. However, they used VSO order in 12 sentences out of 19 which occurred in narrative context. In contrast, the results of the current analysis of MSA corpus data showed that the SLA speakers preferred VSO more than the other two orders. This order was used generally in 63% of the sentences that appeared in the data and in 64% of the sentences which occurred in narrative discourse. SVO, on the other hand, was used to form 3 sentences out of 4 that were found in conversational discourse; the remaining one was in VOS order. The contrast between preferring SVO when speaking the colloquial variety and VSO when speaking the standard can be

explained by the different types of data examined in each study. The colloquial data was taken from telephone conversations whereas the MSA data was taken from news broadcast in which narrative discourse is dominant. It is conceivable that the speakers of Levantine varieties of colloquial and standard Arabic prefer SVO in conversational discourse and VSO in narrative discourse regardless of whether they are speaking colloquial or standard Arabic, and hence, such a contrast emerged. However, this possibility cannot be conclusive without examining more conversational data in SLA to see if its speakers prefer SVO in that type of discourse too. Otherwise, these speakers might have preferred VSO due to the fact that it is the common unmarked order in MSA regardless of their L1 colloquial preferences.

The data from Egyptian varieties of Arabic, standard and colloquial, presented a clear contrast in terms of word order selection. SVO was dominantly used in CEA data in both types of discourse. However, VSO order appeared more frequently (72%) than the other two orders in the data of SEA which was all narrative in nature. This may suggest that MSA grammar that is represented by the Egyptian speakers has its own word order preferences and is not influenced by the grammar of the colloquial Egyptian variety.

Studying 3rd person plural subject-verb agreement was not successful in revealing differences in how MSA grammar is represented by speakers of different CA varieties. According to the results of analysing colloquial Arabic data, all the varieties follow the same pattern of plural subject-verb agreement: if the subject has a human referent, it agrees with the following verb and vice versa. The exception from this was noted in CLA and CEA data: when the subject has a human referent and denotes collectivity, it

may fail to agree with the following verb. Unfortunately, the data of SLA and SEA in the MSA corpus study simply did not contain sentences with subjects of this type to check if the 'no agreement' option was made available in MSA by the speakers of these two varieties due to L1 influence.

The results of analysing the data for resumption were straightforward and easy to interpret. As detailed in chapter 2, section 2.3, MSA in general was described as allowing both resumptive pronouns as well as gaps to fill in the extraction site in object relatives (e.g., Aoun et al., 2010). The speakers of the gulf varieties of Arabic exhibited this optionality of 'resumption' and 'no resumption' in both CGA and SGA with a clear preference for using resumptive pronouns. Similarly, the speakers of the Levantine varieties also allowed gaps to replace resumptive pronouns in some object relatives in both CLA and SLA. There were only 2 relatives with gaps out of 61 in CLA and those were disputable, whereas in SLA, the data contained 4 object relatives with gaps out of 34. Egyptian speakers, on the other hand, did not allow gaps to replace resumptive pronouns even once in all of the 314 object relatives in CEA and the 29 object relatives in SEA. This may suggest that the grammar of MSA represented by these speakers may have been influenced by the grammar of CEA when it comes to resumption in object relatives.

Ultimately, the findings of the MSA corpus study cannot be taken as final and conclusive due to certain limitations. These limitations include that the amount of data analysed in this study is small compared to that examined in the colloquial corpus study to define the variables. In the MSA corpus study, the data contained, for example, a

total of about 756 relevant sentences which were examined for word order and agreement. There is a big difference between this number and the 2631 sentences which were examined in the colloquial corpus study. Consequently, this had an impact on, for example, finding the relevant data for testing 3rd person plural subject-verb agreement in MSA. While it was possible to find about 101 sentences suitable for testing agreement in CLA and CEA data, only 13 sentences of such type were tested in SLA and SEA data, and none of these had a plural human subject denoting collectivity. This limitation of small amount of data in the MSA corpus study could not be avoided by the attempt to combine data from two sets of corpora; the relevant tokens were few despite the fact that all the available data was analysed.

Another limitation of the MSA corpus study which entailed few numbers of relevant comparable tokens is the source of the data. The data of this study was transcripts of broadcast news which evidently involved narration of events and stories. This data is different in nature from the data examined for determining the variables of the study; the colloquial data involved transcripts of telephone conversations. As the type of discourse is important in examining word order, for example, this difference had a great impact on the number of comparable sentences in each data. While most of the sentences in the colloquial data occurred in conversational discourse (about 2,524 sentences), only 29 sentences were found in such discourse in the standard data. Also, despite the small amount of standard data in general, it contained far more sentences in narrative discourse than the colloquial data; 727 and 89 sentences, respectively. This limitation could not be avoided due to the fact that MSA is not often used in conversations or in natural speaking situations like telephone calls, and, thus, it was difficult to find MSA data of this type.

142

Other limitations of the MSA corpus study are related to the validity of the examined standard data. These include the possibility that news data are carefully written in advance and probably checked more than once before being broadcasted. This, in fact, eliminates that chance that the data can be considered as natural data, based on which one can define the grammatical representation of MSA. In addition, obtaining data from an Egyptian channel like Nile TV, for example, does not necessarily mean that the presenter of the news broadcasted in that channel is Egyptian; it is possible that the channel has employed somebody from Lebanon, for instance, to read the news. Furthermore, the presenter may not necessarily be the author of the data he/she presents; the news may have been written by another person from a different region. These possibilities together shed doubts on the origin of the examined data in the MSA corpus study; what has been considered as SEA may have been produced by a non-Egyptian speaker.

These limitations will be overcome in the following more controlled experimental study which will provide specific information about the speakers and test exactly the relevant variables. The experimental study will form a key component of this thesis and the methodology of which will be presented in the following chapter.

Chapter 4

Methodology: Experimental Study

The primary goal of this experimental study was to test the research questions that relate to the effect of the colloquial varieties and different ages of first exposure on the ultimate attainment of MSA in the Arab world as stated in Chapter 1. Separate instruments to measure the effect of these variables were utilized to this end. The methodology employed to test the research questions is presented in this chapter. The chapter is organized into four sections: (1) participants of the study, (2) tasks of the study, (3) data collection, and (4) data analysis.

1   Participants of the Study

This study is intended to compare two subsets of the population of MSA speakers; those who began being exposed to MSA as early as at the age of 2 up to the age of 6 in nurseries and kindergartens, and those who started being exposed to MSA at the age of 6 in normal primary schools. Age of starting school was used to indicate age of first exposure to MSA as usage of MSA is almost limited to schooling environment in the Arabic speaking countries (see Chapter 1 for more details of using MSA in the Arabic-speaking countries). All the participants were native speakers of one of the colloquial varieties which were examined in chapter 3; Colloquial Levantine Arabic (CLA), Colloquial Gulf Arabic (CGA), and Colloquial Egyptian Arabic (CEA). Also, all the participants were exposed to MSA for at least 5 years to ensure that we are testing the end-state grammar. This criterion of requiring a minimum 5 years of immersion in the

target language to ensure testing the end-state grammar has been commonly adopted by researchers in the field (e.g., Patkowski, 1980; Johnson and Newport, 1989; Johnson, 1992; Shim, 1993).

The sample recruited for this study was comprised of 147 participants. The participants' age ranged between 12;0 and 18;9, with an average of 14;6. The participants were first divided into two main groups depending on their age of first exposure to MSA. *Group 1* included 60 participants whose first exposure to MSA was before the age of 6. During their childhood, these participants went to nurseries and kindergartens that apply what is called 'the MSA immersion program' (MSA-IP) which provides a restricted MSA only speaking environment.[33] *Group 2* consisted of 87 participants whose first 'proper' exposure to MSA started in primary school at the age of 6.[34]

The members of these two main groups were also distributed across five subgroups depending on the colloquial variety of Arabic (CV) they speak and the age of first exposure (AoE) to MSA. These groups are: (i) E-CLA group – 30 speakers of CLA with early exposure to MSA (before 6), (ii) L-CLA group – 27 speakers[35] of CLA with late exposure to MSA (after 6), (iii) E-CGA group – 30 speakers of CGA with early exposure to MSA, (iv) L-CGA group – 30 speakers of CGA with late exposure to MSA, and (v) L-CEA group – 30 speakers of CEA with late exposure to MSA. Unfortunately,

---

[33] See chapter 1 for a thorough description of this type of school.
[34] Most Arabic-speaking people start being exposed to MSA properly in primary schools as MSA is adopted to be the literacy language in all Arabic speaking countries. Before primary school, children may hear MSA on TV or Mosques, for example, but they usually do not have a chance to communicate by using it; see chapter 1 for more details on this issue.
[35] This group was originally planned to include 30 participants like the other groups, but the researcher had to leave Syria leaving behind the remaining 3 participants due to the growing political conflict in the country in April 2011.

there are no pre-schools in Egypt that apply the MSA immersion program. Thus, it was not possible to include speakers of CEA whose exposure to MSA was before the age of 6 as if they exist, they would be exceptional cases which cannot be easily identified. However, including the L-CEA group in this study is useful to test for L1 influence by comparing the performance of its participants to that of the L-CLA and L-CGA participants.

All the participants in this experimental study were students of 6 different schools in Saudi Arabia and Syria. The E-CGA participants were students in a secondary school in Dammam City of Saudi Arabia who, when they were children, had attended Al-Bassam kindergarten which applies the MSA Immersion Program. The L-CGA participants were students of a secondary school in Riyadh City of Saudi Arabia who did not attend a nursery or a kindergarten before primary school. The E-CLA participants had attended, when they were children, Al-Azhaar Al-Arabiyyah kindergarten in Harasta City of Syria which applies the MSA Immersion Program. The members of this group were recruited from three different local secondary schools near the kindergarten because the required number for the group could not be found in one school. The L-CLA participants were students of one secondary school in Harasta city of Syria. These students were not in a kindergarten before the primary school. Finally, the participants of the L-CEA group were students of a high school in Saudi Arabia. The members of this group were Egyptians who had come with their families who work for a temporary time in Saudi Arabia, but they had had their primary and secondary education in

Egypt.[36] These students were not in a kindergarten before they started their primary school education. Table 15 below summarizes what has been mentioned so far about the participants of the study (for demographic details of the individual participants, see Appendix 3).

	Subgroups	No.	CV	AoE	School	MSA-IP
Group 1	E-CLA	30	CLA	Early	3 secondary schools in Harasta City of Syria	Al-Azhaar Al-Arabiyyah Kindergarten
	E-CGA	30	CGA	Early	A secondary school in Dammam City of Saudi Arabia	Al-Bassam Kindergarten
Group 2	L-CLA	27	CLA	Late	A secondary school in Harasta City of Syria	N/A
	L-CGA	30	CGA	Late	A secondary school in Riyadh, Saudi Arabia	N/A
	L-CEA	30	CEA	Late	A high school in Riyadh, Saudi Arabia	N/A

Table 15 : Information about the participants of the experimental study

All the participants of the study did a proficiency test in MSA and passed the advanced level. This test was developed by a non-profit organization in Saudi Arabia called 'Arabic for All', which, among its several activities, works on designing Arabic language teaching curricula and Arabic language courses and tests. The proficiency test used here was designed to be given to students in advanced level and its passing mark is 65% of the total mark. The test is composed of four sections: listening comprehension, reading comprehension, appropriate use of vocabulary, and composition. Table 16 below shows the groups' mean and range of the participants' results in this test.

---

[36] It was planned to collect data for this group from Egyptian students in Egypt in January 2011, but the data collection trip was hindered first by delay of sponsoring funds and then the Egyptian Revolution started and made it impossible to go to Egypt at that time.

Group	E-CGA	L-CGA	E-CLA	L-CLA	L-CEA
Mean	82.4	82.2	80.6	80.9	83.4
Range	70.4 - 91.6	71.2 - 92.8	70.0 - 91.6	70.8 - 93.6	70.4 - 93.6

Table 16 : The groups' mean and range of the participants' results in the MSA proficiency test

The participants of the study were identified and invited through the schools where they study. The key contact in each school identified the potential participants who fit the definition of each group and invited them to take part in the study. Then, the school developed a list of potential participants who came back with positive response to our invitation. All those agreed to take part did the proficiency test on the first day of the researcher's visit to each school.

The number of students who agreed to take part in the study exceeded the required number for each groups except for E-CLA where participants were invited from three different schools due to the fact that only few people in each school had attended an MSA Immersion Program before the age of 6. Then, the key contact in each school started to choose students from the list, depending on factors related to organization of their school day, and send the potential participants in small groups of three to each data collection session. This selection procedure continued until the required number for the group was reached, i.e., 30 participants. Then, the researcher apologized to the rest who accepted our invitation but we couldn't recruit them for the study due to time limitation.[37] The following section will describe the tasks that these participants were asked to complete.

---

[37] More details about how the data was collected will be presented in section 3 of this chapter.

2   Tasks of the Study

The participants in this study were administered two tasks: an acceptability judgment task and an elicited production task. The former was designed to assess the participants' underlying knowledge of two of the MSA syntactic phenomena under study: resumption in object relative clauses and plural agreement between a human subject of the collective type and the verb in declarative sentences in SV order.

The acceptability judgment task has been a common measure in the literature of L2 acquisition. It is used to tap into learners' underlying knowledge of specific linguistic structures by asking the learners to decide, based on their metalinguistic intuitions, whether a given sentence or a context is acceptable or not. Many researchers in the field have used this tool in various formats to investigate knowledge of the acquired language. Examples of these researchers include Johnson & Newport (1989), Johnson (1992), Shim (1993), Flege (1999), DeKeyser (2000), McDonald (2000), Birdsong & Molis (2001), Marsden (2008; 2009), Montrul (2010), among many others.

An acceptability judgment task (AJT) could not be used to investigate the third syntactic phenomenon under study, i.e. word order, due to the fact that all variations of word order are accepted in all varieties of Arabic; standard and non-standard. Thus, since the question about word order variations is not whether they are accepted or not, but which variation is preferred in which context, this was assessed in an elicited production task which I will call 'Conversation Role-Play' (CR-P).

Using two different measures to investigate the same research questions adds an advantage to the design of this experimental study. This has been considered important in the literature to validate the research findings. Han (1996), for example, pointed out that it is essential to employ multiple tasks that test various linguistic structures when investigating the Critical Period Hypothesis. This is to strengthen the validity and generalizability of the research findings. Han argued that the results of an AJT can be more reliable if its results are supported by the results of other measures on the same participants or items. The following sections will present more details about these two tasks.

2.1 Acceptability Judgment Task

In this task, each group of three participants viewed single sentences on a laptop screen and listened to these sentences simultaneously. The participants were asked to judge individually whether the sentence they saw and heared were grammatically acceptable or not by choosing one option for each sentence from a four-point scale ranging from 'Very strange, Unacceptable' to 'Perfectly good, Perfectly acceptable'. A fifth option of 'Can't decide' was also available. The scale used was presented in the following form (see Appendix 4 for the AJT cover sheet & answer sheet):

Very strange. Unacceptable.	A bit strange. Not really acceptable.	Fairly good. Acceptable.	Perfectly good. Perfectly Acceptable		Can't decide
–2	–1	+1	+2		X

Figure 4: The scale used to judge sentences in the AJT

Each sentence in this task showed on the screen for 10 seconds only and then the next sentence appeared. Also, the participants were not allowed to go back and change their decisions about previous sentences. This procedure of attempting to obtain quick

responses without allowing a great deal of thinking time was followed to ensure assessing the participants' knowledge of the grammar and not their knowledge of MSA formal rules (Gass and Mackey, 2007).

The task was composed of 60 sentences in total which included three types of sentences: i) 10 sentences with object relative clauses repeated twice, once with resumptive pronouns and once without them; ii) 10 sentences with a human subject of the collective type in SV order repeated twice, once with verbs showing plural agreement and once with verbs in a singular form; iii) 20 sentences as distractors, 10 of which are clearly grammatical and the other 10 are clearly ungrammatical. Repeating the same sentence with a sole difference related to the investigated syntactic issue was done to compare the learners' judgments and make sure that if judgments are different then this is caused by manipulating the sentence (Schutze, 1996). Examples of the test sentence types are shown in Table 17. Appendix 5 lists all the test sentences used in this task.

The number of test items in this task was decided based on some recommendations found in the literature that relate this issue to the reliability of the test. Cowan & Hatasa (1994), for example, pointed out that longer tests result in greater reliability. According to these researchers, an AJT should include between 60-72 test items. Also, Gass & Mackey (2007) recommended that test items in an AJT should not be more than an approximate number of 50-60 sentences.

Type	Variations	Examples
1a	Object Relatives + resumption (10 sentences)	قرأ ماجدٌ الكتابَ الذي اشتراه الأسبوع الماضي qara?a maajid-un l-kitaab-a l-aðii štaraa-hu l-?usbuw9-a l-maaDii read Majid-nom the-book-acc that-s.m. (he)bought-it the-week the-past 'Majid read the book that he bought last week'
1b	Object relatives – resumption (10 sentences)	قرأ ماجدٌ الكتابَ الذي اشترى الأسبوع الماضي qara?a maajid-un l-kitaab-a l-aðii štaraa-Ø l-?usbuw9-a l-maaDii read Majid-nom the-book-acc that-s.m. (he)bought-Ø the-week the-past 'Majid read the book that he bought last week'
2a	Collective Subject-Verb + plural agreement (10 Sentences)	الناسُ يخافون الخروجَ من المنزلِ ليلاً في هذا البلد l-naas-u yaxaaf-uwna l-xuruwja mina l-manzili laylaan fii haðaa l-balad the-people-nom fear-3pm going out from the-house at night in this the-country 'People avoid going out at night in this country'
2b	Collective Subject-Verb – plural agreement (10 Sentences)	الناسُ تخافُ الخروجَ من المنزلِ ليلاً في هذا البلد l-naas-u taxaafu l-xuruwja mina l-manzili laylaan fii haðaa l-balad the-people-nom fear-3sf going out from the-house at night in this the-country 'People avoid going out at night in this country'
3a	Distractor + grammatical (10 sentences)	عليٌ اشترى طعامَ العشاء من مطعم مجاور لبيته 9aliyy-un štaraa Ta9aama l-9ašaa?i min maT9amin mujaawirin li-bayt-ih Ali-nom bought-3sm food the-dinner from restaurant next to-house-his 'Ali bought dinner from a restaurant next to his house'
3b	Distractor – grammatical (10 Sentences	*أخي كتبت واجبه المدرسي بإتقانٍ تام *?ax-ii katab-at waajiba-hu l-madrasiya bi-?itqaanin taam brother-my wrote-3sf homework-his the-school with-perfection completely 'My brother wrote his school homework perfectly'

Table 17: Types & examples of test sentences in AJT

All the test sentences were declaratives except for two sentences of the ungrammatical distractor type which were questions with the question word ungrammatically remaining in situ. The length of the sentences was between 6-9 words long with the majority being 7 words long. The declarative sentences alternated between using two word orders, SV and VS, except for those of type 2 (collective subject-verb agreement sentences) which were all in SV order due to the specification of the variable under study. This was done on purpose to minimize possible modeling for the participants'

production of declarative sentences in the following task. The sentences in general were straightforward using simple vocabulary and structures. Also, all the test sentences were checked for clarity and simplicity by a specialist in modern standard Arabic from the College of Arabic in Al-Imam Muhammed Ibn Saud Islamic University in Riyadh, Saudi Arabia.

The test started with two distractor sentences, one grammatical and one ungrammatical, to give the participants a chance to familiarize themselves with the rating system and the test format.[38] The test also ended in the same way to avoid random answers due to possible boredom or lack of concentration which may be caused by, for example, rushing to finish. Having each sentence with a specific number during the design stage, a random number generator was used to distribute the rest of the sentences randomly in the test. Half of the participants in each group were shown the sentences in this randomly generated order. This order, however, was reversed for the other half of the group. This was done in such a way that if, for example, a sentence Y occurred with resumption as the third and without resumption as the tenth sentence in the first order, these sentences were revered in the second order. The purpose of having two different orders is to control for the possibility that the participants may judge the following sentences based on their judgments of the sentences which appeared first. Ordering the test items this way followed Schutze's recommendations (1996) which state that a researcher should control for order effects by randomizing test items and counterbalance orders across different participants. Schutze also advised to consider nervousness at the beginning of the session and fatigue at the end when distributing the items of the test.

---

[38] As will be explained later, the participants also did a pre-test training session for the same purpose.

Before the test starts, the participants were given the following instructions in a written

form in Arabic (see Appendix 4: A.1 for the Arabic Version):

> "In this task, you will see a number of single sentences in Standard Arabic.
> Please judge whether the sentence is grammatically acceptable to you.
> Indicate your answer by circling one of the options on the scale on your
> answer sheet."

Given the fact that many participants find the idea of rating sentences for

grammaticality or acceptability as something new and unusual, additional oral

instructions were given to participants to ensure that they know what they should

do exactly in this task. First, the researcher explained to the participants that

speakers of any language seem to have a certain feeling about whether a certain

sentence is possible or not in that language. It was stated to the participants clearly

that their job is to concentrate on how they feel about the sentences they are going

to see and hear, without trying to remember the rules that they had studied in

class. They were told to read and listen to each sentence carefully and concentrate

on the structure of the sentence before they decide about it. Also, the participants

were told that they may think that some sentences have appeared before, but the

fact is that none is repeated in the test and that they should treat each sentence in

its own right without relating it to other sentences they may have seen or heard in

the test. Furthermore, it was mentioned to them that sometimes they may have no

clear feeling about a particular sentence, and in this case, they should choose 'can't

decide'. Finally, the participants were instructed that when answering, they should

ignore any problems with spelling, punctuation, pronunciation, and other problems that are not related to the structure of the sentence.[39]

.Before starting the test, the participants did a pre-test training session to acquaint themselves with the test format and the rating scale. This session included three sentences that are irrelevant to the syntactic variables under study in this task. After doing this session, the researcher had a short discussion with the participants based on their judgements to make sure they understood the procedure of the test. Also, the training session was repeated for some participants who seemed still not confident about using the rating scale, but they were told that the sentences of the test that they are about to see and hear will not be repeated. When the test started, the participants had a one minute break after judging half of the sentences to turn the page over and re-gain their concentration.

The actual test lasted for approximately 10 minutes. After completing the task, the participants had a short break in the test room and then started the following task. The details of the conversation role-play task are given in the following section.

2.2 Elicited Production Task: Conversation Role-Play

This task was designed to prompt production of declarative sentences in two different contexts: conversational and narrative. The purpose is to check which word order participants would use in each context.

---

[39] The instructions given in Bley-Vroman, Felix, and Ioup (1988: 32) were used as a model to form the instructions given here, but it was preferred to keep the written instructions to a reasonable length and give the rest orally to avoid the problem that impatient participants may not read them completely.

The participants completed this task in the same data collection session when they did the AJT. In this task, they were asked to play the role of one of the two speakers in a written interactive conversation that was pre-programmed in a software implemented specifically for this task.[40] Basically, the participants were asked to form complete sentences that are suitable for the context of the ongoing conversation. The participants could form their sentences by dragging words into blanks from three available boxes which are labelled with specific instructions like 'choose one word; two words; as many as you need; choose any if you need to, etc.'. Figure 5 is a printed screen that shows two examples of conversational turns for each speaker from the actual test, and Figure 6 provides the English translation of these examples (see Appendix 6 for the content of this task).
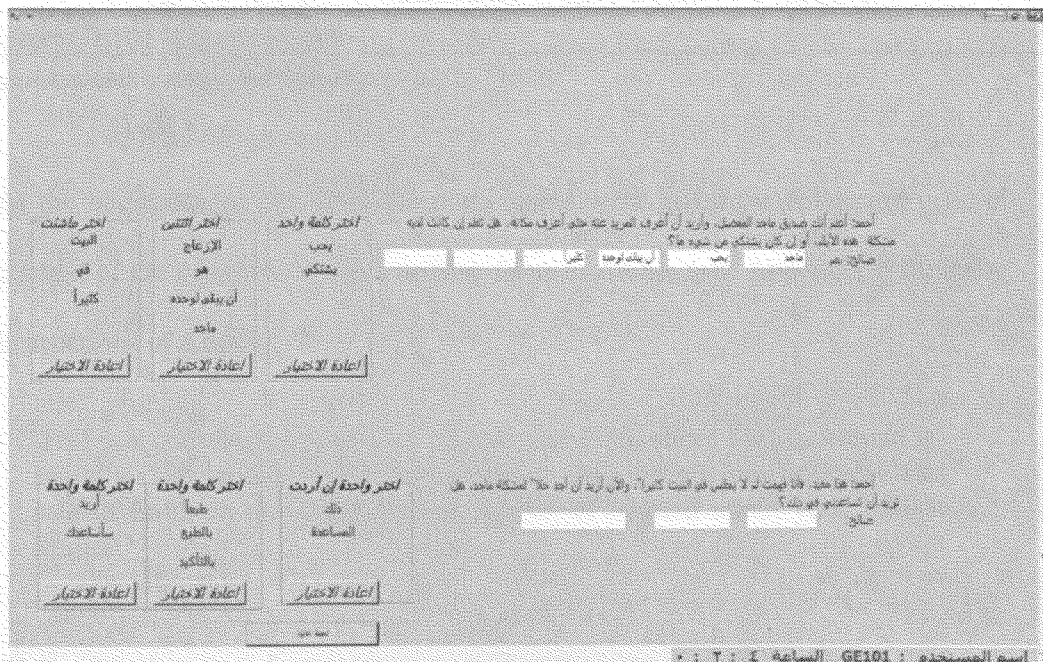
Figure 5: Two example conversational turns for each speaker from the actual test

---

[40] This software was kindly implemented for me by my friend programmer Mr. Ahmed Sulaiman in 2010.

Ahmed: I know you are Majid's best friend, and I want to know more about him so that I can know where he is, do you know if he has a problem these days or if he complains about something?

Salih: Yes, | Majid | | likes | | to be alone | | a lot | | | | |

Use one word	**Use two words**	**Use as many as you want**
likes complains (about)	noise he to be alone Majid	the house in alone
Reset Choices	Reset Choices	Reset Choices

Ahmed: This helps to understand why he does not like to stay at home a lot, now I want to find a solution for Majid's problem, do you want to help me on this?

Salih: | | | | | | |

Use one word, if you want	**Use one word**	**Use one word**
that helping	sure certainly definitely	(I)want (I)will-help-you
Reset Choices	Reset Choices	Reset Choices

User: GE101 Clock: 0:2:4

Figure 6: English translao n of Figure 5

At the beginning of this task, the participants were given the following general context of the conversation in a written form in Arabic (see Appendix 6 for these instructions in Arabic):

> 'This is a conversation between Ahmed and Salih about Majid. Ahmad is Majid's elder brother, and Salih is Majid's close friend. You will play the role of Salih in this conversation talking to Ahmad about your friend Majid.'

Right after the general context, the participants were given the following detailed written instructions in Arabic:

'Please do the following to complete this task:

First: Read Ahmed's speech.

Second: Read all the words given in the available boxes, and the
    instructions for selection above each box.

Third: Use the words available in the boxes according to the given
    instructions to form a sentence in your mind first that suits the
    context of the conversation.

Finally: drag the words of the sentence that you have formed in your
    mind successively to fill in the blanks.'

The participants were told to follow these steps for each conversational turn that they are about to take. They were asked to read the other speaker's speech first to keep up with the context of the ongoing conversation. Then, they were asked to read all the available words and the instructions above the word boxes to see what words they can use to form their sentence. Words in boxes can be combined in different ways to form different sentences that fit the context of the conversation. For example, in Figure 5 & Figure 6, instead of *'Majid likes to be alone a lot'*, a participant could have chosen other words to form a different sentence like *'He complains about noise in the house'* and still obey the instructions of the boxes and keep up with the context. The third and fourth steps of the general instructions were given to ensure that the participants form complete sentences in their minds before dragging words and not drag one word then think what to drag next. This is to prevent possible confusion and several changes and help to reduce thinking time.

The software screen shows two conversational turns at a time; one for the first speaker filled in by the software and one for the participant to fill in. When the

participants finish filling in the blanks with their sentence, they press 'next', and the screen scrolls down to show the following two new turns while keeping the previous two turns visible on the screen for the participant to look at, when needed, to check the context. Also, the participant can always scroll up the screen to see all the previous conversational turns.

The software does not allow the participant to proceed to the following conversational turn before finishing the current turn according to the instructions given above the word boxes. For example, if a participant did not choose two words from a box that states this and pressed 'next', a small window will pop up saying that he/she should fill in the blanks according to the instructions before proceeding to the next turn. Also, word boxes do not allow taking out more words than were specified by the instructions above them. If the instructions above one box state, for instance, 'choose one word', the participant will not be able to take more than one word from that box. In each word box, there is a 'reset-choice' button. This button when pressed will bring back all the words taken from that box. The participants, however, were told not to press this button at all to minimize thinking time and encourage participants to drag the first sentence they think of. This button was put there for the researcher to use in cases such as when a participant drops a word in a wrong place by mistake. Finally, there is a running clock always visible at the bottom of the screen. This is to indicate to the participants that the task is timed and encourage them to form sentences as quickly as possible and to drag the first sentence they manage to form in their minds.

By taking part in this conversation, a participant would form 20 sentences. By controlling the task in the way described above, a participant would have to produce 10 of these sentences as declarative sentences all with an overt subject and a verb and the participant would have a choice of making these sentences in either an SV or a VS order. The participant would see three word boxes with each conversational turn in this task. With the conversational turns which are meant to elicit this type of sentence (the token sentences), these three boxes were as follows. One box contained transitive verbs and was labelled with 'choose one word'. Another box contained nouns and/or pronouns, suitable as subjects and/or objects. This box was labelled with 'choose two words'. These two boxes with these specified instructions guaranteed that the elicited sentences would have an overt subject, a transitive verb, and an overt object. The third box contained lexical items of other types like adjectives, adverbs, prepositions, etc., that might be needed to complete the sentences and make them more comprehensible. This last set of boxes were labelled with various instructions like 'choose as many as you want' or 'choose one if you need to'.

The other 10 sentences that a participant would produce in this conversation were distractors. These sentences were of various types like greetings, questions, requests, verbless sentences, and null subject sentences. The content of word boxes for these conversational turns and the instructions above each box varied and were not as systematic as those used for the non-distractor sentences.

Because the word boxes were put in a linear order next to the blank conversational turn, it is possible that the order of the boxes might have influenced the order of the sentence of the non-distractor type that the participant was asked to form. In other words, if a participant, for example, found the first box from the right containing verbs, he/she might start the sentence with a verb given that Arabic writes from right to left. To control for this possibility, the boxes were organized in such a way that in 5 random conversational turns of non-distractor type, boxes of nouns preceded boxes of verbs, and in the remaining 5 conversational turns, boxes of verbs preceded boxes of nouns. In addition, there were two versions of this task. The only difference between the two versions was that if, in version one, the word boxes for a conversational turn Y were ordered in a way that boxes of verbs preceded boxes of nouns, this order was reversed for the same conversational turn in version two. Half of the members of each group did a different version from that completed by the other half.

To elicit declarative sentences in a narrative context, the other speaker invites the participant near the end of the conversation to watch a silent film (10 minutes long) and then, when they had finished watching, to write 10 sentences to narrate the events of that film. The other speaker also requested that each sentence should mention at least one name of the characters in the film, and the names of all characters were provided. This was requested to make sure that the participant uses overt subjects in his/her sentences in order to be able to check the order of the sentences produced. The film was of the comedy type (Mr. Bean) to serve as a good break between filling in the conversational blanks and writing the narrative sentences.

When the participants complete the task, they press on the 'Save' button and all the data entered become saved in a separate Microsoft Access file. The saved data shows the code of the participant and the version of the task that he/she has completed. The participants spent between 35-45 minutes to complete the whole task.

Before starting the task, the participants did a pre-test training session to familiarize themselves with the software functions and the test format. In this training session, the researcher first explained to the three participants in each data collection session the way they should complete the task. Then, each participant having his/her separate laptop took part in a short written conversation (three conversational turns only) with exactly the same procedures as in the main task. Then they watched a very short film clip (about 1 minute long), then they were asked to write 2 practice sentences mentioning at least one name of the characters of the film in each sentence. After doing this session, the researcher had a short discussion with the participants to make sure they understood the procedure of the task, and answer any further queries.

2.3 The pilot study

The aim of conducting a pilot study was to check the reliability of the test format and the test items in both tasks. To this end, a total of 12 adult participants (age mean 26,9; range 18,2-36,9) were recruited in this pilot study. The participants were four native speakers of colloquial Gulf Arabic (CGA); four native speakers of colloquial Levantine Arabic (CLA), and four native speakers of colloquial Egyptian Arabic (CEA) who all had their primary education in an Arabic speaking

country. The purpose of including speakers of these three colloquial varieties was to assess the clarity and suitability of the test items for the participants in the main study who speak these different colloquial varieties of Arabic.

The participants were postgraduate students from different departments at the University of York (CLA & CEA participants) or students from different English Language schools in York, UK (CGA participants). None was exposed to MSA before the primary school and started to be exposed to MSA in schools after the age of 6.

The participants in this pilot study completed the tasks individually in order for the researcher to concentrate on observing how the participants completed the tasks and note any problematic issues. The tasks of this pilot study were the same as described in the previous two sections of this chapter, 2.1 & 2.2, except for few differences which, as will be mentioned below, were modified later based on the outcome of this pilot study. These differences include that the participants in this pilot study were given less detailed instructions on how to complete the conversation role-play task. Compare the following instructions which were given in the pilot study to those given in the main study which were stated in section 2.2 of this chapter:

'Fill in the blanks by forming comprehensive sentences that fit the conversation context by using the appropriate words according to the given instructions'

Also, the pre-test training for the conversation role-play task in this pilot study did not include a film clip and did not ask the participants to practice writing narrative sentences.

After completing the tasks of the study, the participants were asked to complete a short questionnaire about the clarity of the test format and the test items (see Appendix 7). This questionnaire included seven questions about usefulness of pre-test training (2 questions), clarity of instructions on how to complete the two tasks (2 questions), simplicity of the test sentences in the AJT (1 question), suitability of pausing time between the sentences in the AJT (1 question), and the simplicity and clarity of the context and flow of the conversation in the conversation role-play task (1 question). The participants were asked to give a mark out of 10 to answer each question. The researcher discussed with each participant his/her answers to these questions and noted any issues that needed to be considered. Also, the researcher asked the participants if there were any lexical items in the two tasks that they could not understand. Such lexical items were replaced with more common ones in the main study.

The results of the questionnaire were as follows. The mean rate given to usefulness of pre-test training for the AJT (question 1) was 9.42 out of 10 (range 8-10). The usefulness of pre-test training for the conversation role-play task (question 2) had a mean rate of 8.92 (range 8-10). Clarity and details of instructions in the AJT (question 3) were rated with a mean of 9.5 (range 8-10), where in the conversation role-play task (question 4), the mean rate was 8 (range

6-10). Simplicity of the AJT test sentences (question 5) was rated a mean of 9.58 (range 8-10). Length of pauses between test sentences in the AJT (question 6) had a mean rate of 9.75 (range 9-10). Finally, clarity of the context and the flow of the conversation in the role-play task (question 7) was rated a mean of 9.33 (range 8-10). The following table shows the mean rate given by the four speakers of each colloquial variety to these seven questions:

	Q1	Q2	Q3	Q4	Q5	Q6	Q7
CGA	9.25	9.0	9.25	8.0	9.75	9.75	9.5
CLA	9.5	9.0	9.5	8.0	9.75	10	9.25
CEA	9.5	8.75	9.75	8.0	9.25	9.5	9.25
TOTAL	9.42	8.92	9.5	8.0	9.58	9.75	9.33

Table 18 : Mean rates given by pilot participants to answer the post-tasks questionnaire

Based on the results of this questionnaire, the discussions with the participants, and the researcher observation, a few changes were applied to the design of the tasks. First, instructions on completing the role-play task were changed to be clearer and more detailed (the modified version is given in section 2.2 of this chapter). Second, the pre-test training for the role-play task was changed to include a short film clip and the participants were asked to practise writing sentences to narrate its events. Also, a few technical software faults were noted when observing the participants completing the role-play task and these were corrected. Finally, one lexical word مكامن *makaamin* 'places of' was mentioned by two of the participants as difficult to understand in the AJT and this was replaced by a more common word أماكن *?amaakin* that has the same meaning.

The distractor items in the AJT being 10 clearly grammatical versus 10 clearly ungrammatical were used to check the reliability of this task. Table 19 shows the number of sentences accepted and rejected by each group of speakers and the total amount of acceptance and rejection of each type of the distractor sentences:

	Grammatical Distractors		Ungrammatical Distractors	
	Acc. No.	Rej. No.	Acc. No.	Rej. No.
CGA Speakers	39 (% 97.5)	1 (%2.5)	9 (%22.5)	31 (%77.5)
CLA Speakers	40 (%100)	0 (%0)	5 (%12.5)	35 (87.5)
CEA Speakers	30 (%75)	10 (%25)	9 (%22.5)	31 (%77.5)
TOTAL	109 (%90.83)	11 (%9.17)	23 (%19.17)	97 (%80.83)

Table 19 : Groups' acceptance and rejection of distractor items in the pilot study

The lower rates of accepting grammatical distractors in the CEA group or rejecting the ungrammatical sentences in the CEA and the CGA groups appear to be caused by individual behaviours in the test. For example, one CEA participant incorrectly rejected 6 grammatical sentences out of 10. This same speaker incorrectly accepted 3 ungrammatical distractors. Another speaker of the same group rejected 4 grammatical sentences and accepted 3 ungrammatical sentences, whereas the remaining two participants of this group accepted all the grammatical and rejected most of the ungrammatical distractors. Similarly, one CGA participant accepted 5 of the 9 incorrectly accepted ungrammatical distractor sentences by the members of that group.

However, by analysing the results of the distractor sentences by test item, there was no specific test item that received incorrect answers except for sentence 59

which was accepted by 7 speakers out of 12. This perhaps was due to the fact that this sentence is at the end of the test when the participants are expected to lose their concentration rushing to finish the test.

Overall then, the fact that grammatical sentences were correctly accepted for an average of 91% of the time and the ungrammatical sentences were correctly rejected for an average of 81% of the time was considered to indicate adequate reliability of the AJT.

2.4 Scoring the Tasks of the Main Study

2.4.1 The Acceptability Judgment Task

In the AJT, any choice of +1 or +2 on the rating scale was considered to indicate acceptance, and any choice of -1 or -2 was considered to indicate rejection of that test sentence.[41] Responses of 'Can't decide' did not occur in this data.

Once all the items were scored following this scoring scheme, responses for sentences with object relative clauses (type 1) were checked to compare each participant's response to a sentence of the type **a** (with resumption) with his/her

---

[41] The scores of 1 and 2 (positive and negative) were conflated because positive choices were clearly labelled with the word 'acceptable' and the negative ones were clearly labelled with the word 'not acceptable/unacceptable' (see the AJT scale in Figure 4 in section 2.1 of this chapter). The purpose of this task was to check whether the examined sentences were acceptable or not acceptable; the aim was not to check gradient judgments. However, the four-point scale was used to consider possible different personalities of the participants: some learners may feel reluctant or less confident to commit themselves to a definite judgment for personality reasons (Sorace, 1996:397-398).

response to the same sentence of the type **b** (without resumption). If both sentences were accepted, then these are given the value **1** to indicate that resumption is considered to be *optional* in this item. If the sentence with resumption is accepted and the same sentence without resumption is rejected, then these are given the value **2** to indicate that resumption is *obligatory* in this test item. The third possible scenario is that the sentence with resumption is rejected and the same sentence without resumption is accepted. In this case, this pair of sentences was given the value **3** to indicate that resumption is not accepted in object relative clauses in that item. The fourth possible scenario is when both sentences are rejected, and in this case, these were given the value **4** to indicate that rejection has occurred for reasons other than resumption.

All pairs of sentences which received values bigger than two were excluded from the analysis. This is because value **3** indicates that at least overt resumption is not possible where the fact is that it is possible in all the varieties of Arabic, standard and nonstandard (see section 2.3 in chapter 2). Also, value **4** indicates that rejection is for an irrelevant reason and thus data of this type were irrelevant to the test. It is also possible that one member of the pair be not decided upon by choosing a 'can't decide' option. If this case had happened, these pairs would have been excluded from the analysis as it would not be clear whether resumption is optional or obligatory in that sentence. However, there were no instances of this scenario. Finally, values of **1** or **2** were counted for each group to indicate how much resumption is perceived as optional or obligatory in object relative clauses.

The same procedure of sentence pair comparison was followed for the sentences of type 2 where number agreement between a human subject of the collective type and the following verb is checked. If both sentences (with and without plural agreement) were accepted, then these were given the value **1** to indicate that plural agreement in this context was *optional*. If the sentence with plural agreement was accepted and the same sentence without plural agreement was rejected, then this was given the value **2** to indicate that plural agreement was *obligatory*. Also, if the sentence with plural agreement was rejected and the other sentence without plural agreement was accepted, then this was given the value **3** to indicate that plural agreement was not possible in that sentence. The value **4** was given to cases where both sentences were rejected to indicate that rejection was for some other irrelevant reason.

Again, only pairs with values **1** and **2** were included in the analysis as they indicate whether plural agreement is optional or obligatory. Data with the value **3** were excluded because plural agreement is accepted in all varieties of Arabic, standard and non-standard (see section 2.1.3 in chapter 2 & section 1.2 in chapter 3). Data with the value **4** were excluded because rejection occurred for reasons other than violation of the rule of number agreement between the subject and the verb in this type of sentence. Moreover, any pair with a sentence that was not decided upon would have been removed from the analysis as no information about whether plural agreement is optional or obligatory can be detected from such a comparison. Finally, values of **1** and **2** were counted for each group to see how much plural agreement was considered optional or obligatory in each group.

2.4.2 The Conversation Role-Play Task

Scoring for this task was simple and straightforward. Each participant formed 20 relevant sentences in this task, 10 in a conversational context and the other 10 in a narrative context. The software used for this task saves all the input data into a Microsoft Access file, each sentence in a separate line, conversational sentences first followed by the narrative sentences.

All the relevant sentences formed by each participant were checked for order. Sentences formed with an SV order were given the value **1**, and sentences formed in VS order were given the value **2**. These values were counted in each context to check how many sentences were formed in a particular order and in what context. The distractor sentences in the conversational context were not included in the analysis.

3   Data Collection

The data collection stage started with contacting schools in Saudi Arabia and Syria where potential participants could possibly be found. These schools include, in Saudi Arabia, Albassam Schools for E-CGA participants, Al-Maali Schools for L-CGA participants, and Prince Mensour School for L-CEA participants. In Syria, the schools contacted were Al-Hawraani Secondary school in Harasta city for L-CLA participants, and Harasta Al-Oula secondary school, Harasta Alsadisah secondary school and Al-Hawraani secondary school for E-CLA participants.

As mentioned in section 1 of this chapter, the key contact person in each school identified the potential participants who fit the description of each group and invited them to take part in this study. As most of the participants were under the age of 18, each participant was handed an information sheet about the study and a consent form to participate for parents or guardians to sign (see appendix 8 for information sheets and consent forms). Then, the key contact person in each school developed a list of those students who came back with a positive response to our invitation.

The researcher conducted several data collection sessions in each school during the school day. Each school kindly dedicated a quiet room for data collection like the school library, a free classroom, or a computer lab. For each data collection session, the key contact person chose three participants from the developed list depending, as mentioned before, on factors related to organization of their school day. This way of choosing participants was found useful for increasing the chance that the participants of the study were selected in a random way.

At the beginning of each data collection session, the researcher gave brief information about the study and made sure that each participant had read the information sheet and brought back a consent form signed by his or her parent or guardian if the participant was under 18 years of age. The researcher gave some time for the participants to ask questions about the study and informed the participants that they could withdraw at any time without being asked for reasons. Then, each participant was given a unique code which showed the group, the

version of the test taken and a unique serial number for that participant. After that, each participant was asked to fill out a form which asked about age, gender, native language(s), the colloquial variety of Arabic spoken at home, age of first time going to school including kindergarten, and information about whether kindergarten or primary school attended applies the MSA Immersion Program, and if yes, the names of schools attended (see Appendix 9 for this form).

When these forms were filled out, the researcher distributed the answer sheets for the AJT and asked the participants to write their codes at the top of the page. Then, the researcher asked the participants to read the instructions of the task and gave additional oral instructions as described in section 2.1 of this chapter. When the participants were ready, the participants were asked to sit in front of a lap-top screen and the pre-test training session started, followed by a short time for discussion and questions to make sure that the participants understood what they should do to complete the tasks. After that, the AJT started and participants read and listened to 60 sentences with a short 1 minute break halfway through before turning the page over to re-gain their concentration.

When this task was completed, the participants had a short break in the test room and then the researcher explained the conversation role-play task and asked the participants to move and use separate lap-tops to do the pre-test training session. Before the participants start the main task, the researcher gave time for a brief discussion and questions about the task then the participants started the main task. The participants were asked to enter their codes to start the software, read the

general context of the conversation and the instructions of the task before they start (these are given in section 2.2 of this chapter). When the task was completed, the participants were thanked for their participation and a new data collection session started with different participants.

4   Data Analysis

Several group comparisons and statistical tests were conducted to address the research questions and the hypotheses presented in chapter 1. Data from the Acceptability Judgment Task (AJT) were analysed in the following way. After excluding the irrelevant data with values 3 and 4, as explained in section 2.4 of this chapter, tokens with values 1 and 2 were counted for each group to see if resumption in object relatives or plural agreement between the human collective subject and the following verb were considered optional or obligatory.

Several independent samples t-tests were first conducted to compare the optional resumption and the optional plural agreement means of the two subgroups within each group to check if the different orders of the test items had affected the results. As will be reported in the following chapter, no significant differences were found between most of these pairs of subgroups and it was thus concluded that order had no effect, hence, data from the subgroups were merged.

The general effect of the two independent variables, the colloquial variety (the L1) and the age of first exposure to MSA, was measured in the data of this task by running a Univariate Analysis of Variance using the software IBM SPSS Statistics

19. Then, the optional resumption and optional plural agreement means of the five groups were analysed separately by using One Way ANOVA tests in SPSS, taking the groups as the sole factor and optional resumption or optional plural agreement as the dependent variable.

Further statistical tests were conducted afterwards to measure the effect of the two independent variables separately. For testing the effect of the AoE, the data of groups with the same colloquial variety but different AoE were entered into statistical comparisons. Two independent samples t-tests in SPSS were conducted to compare the optional resumption means of the group of E-CGA with the L-CGA and the group of E-CLA with the L-CLA. Another two independent samples t-tests were conducted the same way to compare the optional plural agreement means. Also, two further separate independent samples t-tests were run to compare the data of all the participants who had *early* exposure to MSA with the data of all participants who had *late* exposure to MSA including speakers of the Egyptian colloquial Arabic; one test to compare the means of optional resumption and the other test to compare the means of optional plural agreement.

To test the effect of the colloquial variety (the L1), the data of groups with the same AoE but different L1s were entered into statistical comparisons. Two independent samples t-tests in SPSS were conducted to compare between the E-CGA and the E-CLA in terms of their means of optional resumption and optional plural agreement. Also, two One Way ANOVA statistical tests were conducted to compare the means of optional resumption and optional plural agreement achieved

by the three remaining groups with *late* exposure to MSA; L-CGA, L-CLA and L-CEA. Furthermore, the optional resumption data and the optional plural agreement data of all participants were divided into three sets each depending on the colloquial varieties these participants speak regardless of the age of first exposure to MSA. These three sets of data of each dependent variable, i.e. resumption and plural agreement, were compared using a One Way ANOVA in SPSS. Post hoc tests were conducted following all ANOVA tests mentioned above, and only when such tests revealed significant differences, the results are reported in the next chapter.

Finally, the test sentences of the AJT with relative clauses or collective subject-verb agreement were analysed in terms of acceptance and rejection. The rate of acceptance and rejection was calculated for each token and presented across the five groups of participants. The total rates for all groups are also reported in the next chapter.

Data of the Conversation Role-Play task were analysed in the following way. First, data of the subgroups of each group who completed different versions of the conversational part of the task were analysed via five independent samples *t*-tests to compare the mean rate of use of VSO order in conversational context. This was to check if the order of the word boxes had affected the participants' choice of word order. As will be reported in the following chapter, no significant differences were found between most of these pairs of subgroups and it was thus concluded

that the order of the word boxes had no effect and, hence, data from the subgroups were merged.

The data realting to use of VSO order in conversational and narrative contexts were entered into a Repeated-Measures ANOVA to check for effect of context, the colloquial variety (the L1), and the age of first exposure to MSA. Then, two One Way ANOVA statistical tests were conducted to compare the rate of VSO use across the five groups in conversational and narrative contexts separately. This test was followed by post hoc tests to check group to group comparisons.

Finally, the conversational turns in which the participants were supposed to form their token sentences in this task were analysed in terms of which order they used in each turn. The rates of using VSO and SVO orders were calculated for each turn and presented across the five groups of participants in the results chapter. The total rates for all groups are also reported in the next chapter.

5 Summary

This chapter described the methodology employed to conduct the experimental study which is a main component of this thesis. The content of this chapter was presented in four main sections. Section one described the participants of the study in terms of their number, age, age of first exposure to MSA, duration of exposure to MSA at the time of the test, the colloquial varieties of Arabic they speak, the groups they were divided into, how they were identified and recruited, and their other demographic details. In section two, the Acceptability Judgment task and the

Conversation Role-Play were thoroughly described in terms of their design, content, administration, and how they were scored. This section also included a subsection that presented details of the pilot study which was conducted to check the reliability of the format of the tasks and the clarity as well as the suitability of the test items. In addition, this subsection showed how the design of these two tasks of the study was amended based on the outcome of the pilot study. Section three explained how the data were collected; how recruited participants were invited to data collection sessions, and how these sessions were administered. Finally, the fourth section explicated the method by which the collected data were analysed to provide answers to the current research questions. The following chapter will lay out the results of this study.

Chapter 5

Results of the Experimental Study

This chapter presents the results of the experimental study which was described in the previous chapter. The results will be presented in separate sections according to the three dependent variables under study. Section one will lay out the results of judging object relative clauses with and without resumption. Section two will be about the results of judging sentences with and without number agreement between the collective human subjects and the following verbs in sentences with subject-verb order. The third section will present the results of producing declarative sentences in SV or VS order in two different contexts: conversational and narrative. The final section will summarize this chapter.

1   Results of Resumption in Object Relative Clauses

Testing the participants' underlying knowledge of resumption in MSA object relative clauses was part of the Acceptability Judgment Task (AJT). In this task, each participant used his/her metalinguistic intuition to judge 10 sentences with object relative clauses. As a reminder, these sentences were repeated twice in the test; once with resumption and once without resumption.

As explained in chapter 4 (section 2.4.1), these 10 pairs of sentences were scored by comparing judgments of each sentence in each pair. Such comparisons revealed one of

four possible scenarios: i) resumption is optional; when both sentences were accepted, ii) resumption is obligatory; when only the sentence without resumption was rejected, iii) resumption is not allowed; when only the sentence with resumption was rejected, and iv) resumption is not relevant to the judgment, when both sentences were rejected.[42] Table 20 below shows the results of the 147 participants' judgments of these 10 pairs of sentences in terms of resumption in MSA object relative clauses.

Resumption →	Optional	Obligatory	Not Allowed	Irrelevant	Total No. of Judged Pairs
Sentence Pairs N (%)	721 (49%)	579 (39.4%)	47 (3.2%)	123 (8.4%)	1470

Table 20 : Judgments on sentence pairs in terms of resumption in MSA object relatives (all groups collapsed together)

1.1 Excluded Data

From Table 20 above, it can be calculated that 88.4% of the judged pairs revealed that resumption in MSA object relative clauses was perceived as either optional or obligatory. As shown in chapters 3 and 4, resumption in object relatives is allowed in all varieties of Arabic under study; standard and non-standard. Therefore, it was decided to exclude from the analysis the 3.2% of judgments which revealed that resumption is not allowed in MSA. Similarly, data in which both sentences of one pair were rejected (labeled as 'irrelevant' in Table 20) were excluded from the analysis. This is because the participants who made such judgments seem to have based their decisions on reasons irrelevant to resumption in object relatives; otherwise, they would have accepted at least one of the sentences. Thus, responses that indicate that resumption is not allowed or rejection of both sentences of a pair seem to have occurred as random errors and were treated as noise.

---

[42] There is a fifth possible scenario and that is when one or both of the sentences were not decided upon by choosing 'Can't decide' option from the scale. This scenario did not occur in this data.

1.2 Effect of Test Items Orders & Results of Merged Subgroups

As explained in the previous chapter, each group of participants was split into two subgroups who saw the AJT test items in two different orders. This was done to control for the possibility that the participants might judge the sentence appearing second in the test based on their judgment of the sentence that appeared first in each pair. Table 21 below shows the subgroup results of judging pairs of sentences with and without resumption in object relatives after exclusion of the above mentioned data.

Resumption →		Optional	Obligatory	Total Number of Judged pairs
Groups	Order of Test Items	No. (%)	No. (%)	
E-CGA	Order 1	80 (57.6)	59 (42.4)	139
	Order 2	59 (43.1)	78 (56.9)	137
L-CGA	Order 1	76 (59.8)	51 (40.2)	127
	Order 2	72 (55.8)	57 (44.2)	129
E-CLA	Order 1	73 (57.5)	54 (42.5)	127
	Order 2	59 (46.5)	68 (53.5)	127
L-CLA	Order 1	85 (64.4)	47 (35.6)	132
	Order 2	32 (32.0)	68 (68.0)	100*
L-CEA	Order 1	99 (70.2)	42 (29.8)	141
	Order 2	86 (61.0)	55 (39.0)	141
All Groups Total		721 (55.5)	579 (44.5)	1,300

Table 21 : Subgroups' judgments on resumption in MSA object relatives after exclusion of irrelevant data

* the number of participants in this sub-group is only 12 compared to 15 in all other sub-groups

To check if there was an effect of test item order in the test, five independent samples *t*-tests were conducted on separate data, using the software IBM SPSS Statistics 19, to compare the means of optional resumption scores achieved by the two subgroups in each group. The results of the *t*-tests showed there were no significant differences found for any group except L-CLA. The descriptive statistics and the results of these five *t*-tests are in Table 22 below.

t-test	Groups	Order of Test Items	No.	Mean	sd	*t*	df	*p*	95% CI for the difference in means
1	E-CGA	Order 1	15	56.39	29.50	1.51	23.87*	.145	-5.05, 32.33
		Order 2	15	42.75	18.95				
2	L-CGA	Order 1	15	61.01	22.81	.572	28	.572	-13.31, 23.63
		Order 2	15	55.85	26.45				
3	E-CLA	Order 1	15	58.80	26.20	1.46	28	.155	-5.62, 33.54
		Order 2	15	44.84	26.16				
4	L-CLA	Order 1	15	63.88	24.11	3.699	25	.001	14.51, 50.98
		Order 2	12	31.13	21.15				
5	L-CEA	Order 1	15	70.13	19.09	1.027	28	.313	9.14, 27.53
		Order 2	15	60.94	28.94				

Table 22 : t-tests descriptive statistics for the order effect on judging optional resumption

* Using Welch's procedure

Since the two subgroups in all the other groups were not found significantly different from each other due to taking the test in different orders, it seems unlikely that the difference found between the two subgroups in the L-CLA group was caused by the order of the test items. Thus, it was concluded that order of the test items had no effect on the way the participants judged the test sentences. Table 23 below shows the results of judging resumption after merging the two subgroups in each group.

It was expected by the hypotheses stated in chapter 1 (section 2) that groups with early exposure to MSA would treat resumption in MSA object relatives as optional regardless of the colloquial variety they speak. Also, it was expected that the colloquial variety (the L1) would have some effect on how the groups with late exposure to MSA would treat resumption in MSA object relatives. Based on this, it was expected that L-CLA and L-CEA participants would judge resumption in this study as obligatory given that it is obligatory in the colloquial varieties they speak, and the L-CGA participants would continue to treat resumption as optional in MSA object relatives as it is optional in their colloquial variety (see chapter 3, section 1.3, for the results of the corpus-based study of the colloquial varieties).
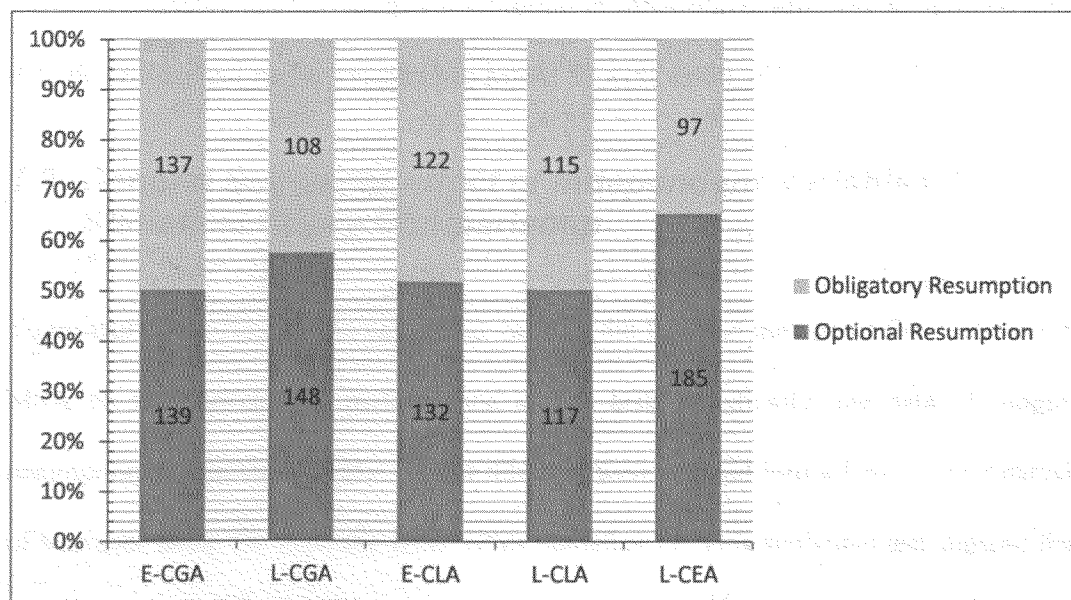
Resumption →	Optional	Obligatory	Total Number of Judged pairs
Groups	No. (%)	No. (%)	
E-CGA	139 (50.36)	137 (49.64)	276
L-CGA	148 (57.81)	108 (42.19)	256
E-CLA	132 (52.0)	122 (48.0)	254
L-CLA	117 (50.43)	115 (49.67)	232
L-CEA	185 (65.60)	97 (34.40)	282
All Groups Total	721 (55.46)	579 (44.54)	1300

Table 23 : Groups' judgments on resumption in MSA object relative clauses

By looking at the figures in Table 23 above, these expectations do not seem to be supported. The Egyptian speakers, for example, achieved the highest rate of treating resumption in MSA object relatives as optional (65.6%) despite the fact that speakers of the same colloquial variety did not produce a single object relative clause without resumption in the colloquial corpus analysed in chapter 3. Also, there seems to be little difference between the judgments given by colloquial Levantine Arabic speakers with Early and Late exposure to MSA; the participants in both groups have judged resumption in half of the relevant test items as optional and as obligatory in the other half. Finally, although resumption in object relatives is optional in the colloquial Gulf variety of Arabic, as it is in MSA, the judgments given by the speakers of this variety do

not seem to be different from the CLA speakers' judgments. In fact, the CGA speakers with *late* exposure to MSA treated resumption in MSA object relatives as optional more than those who had earlier exposure to MSA; 57.81% vs 50.36%, respectively.

1.3 General Statistical Effect of Independent Variables & Statistical Group Comparisons

To measure the effect of the colloquial variety (the L1) and the age of first exposure to MSA on the participants' performance in this task statistically, the data of judging resumption as optional in MSA object relatives were entered into a Univariate Analysis of Variance using the software IBM SPSS Statistics 19. This statistical test showed that there was no significant effect of either the colloquial varieties (CV) or the age of first exposure to MSA (AoE) on judging resumption in MSA object relatives. For CV: F $(2,142) = 2.237$, $p = 0.111$; for AoE: F $(1,142) = 0.447$, $p = 0.505$. The interaction of CV and AoE was not significant either (F $(1,142) = 1.418$, $p = 0.236$).

Further statistical analyses were conducted to compare the means of the five groups on judging resumption in MSA object relatives as optional. This was done using a One Way ANOVA in SPSS Statistics 19 with the groups as the factor and optional resumption as the dependent variable. The ANOVA showed no statistical effect of different groups on data (F $(4,142) = 2.163$, $p = 0.076$). The descriptive statistics are given in Table 24 below.

Groups	No of Participants	Mean	Sd	95% CI for Mean		Minimum	Maximum
				Lower	Upper		
E-CGA	30	49.6	25.3	40.1	59.0	11.1	100.0
L-CGA	30	58.4	24.4	49.3	67.6	20.0	100.0
E-CLA	30	51.8	26.7	41.9	61.8	.0	100.0
L-CLA	27	49.3	27.9	38.3	60.4	.0	90.0
L-CEA	30	65.5	24.5	56.4	74.7	20.0	100.0
Total	147	55.1	26.2	50.8	59.3	.0	100.0

Table 24: Descriptive statistics for groups' scores on judging resumption as optional

The statistical tests conducted so far seem to suffice to answer the research questions and controvert the research hypotheses related to resumption. However, further statistical analyses were conducted to measure the effect of AoE and CV (the L1) separately by making various specific comparisons and manipulating groups of participants. This was done to make sure that every relevant analysis was conducted. These analyses are presented in the following two subsections.

1.4 Further Statistical Analysis: Effect of AoE

1.4.1 Comparing Groups with the Same CV but Different AoE

To check further if early AoE has helped the participants to acquire a more native-like knowledge of resumption in MSA object relatives, data of groups with the same CV but different AoE were tested by two separate independent sample t-tests. This included comparing E-CGA with L-CGA and E-CLA with L-CLA participants' scores on optional resumption. The results of this test showed that there was no significant difference between the E-CGA and the L-CGA participants (E-CGA mean = 49.6, sd = 25.3, N = 30; L-CGA mean = 58.4, sd = 24.4, N = 30; the 95% CI for the difference in means is -21.72, 3.99; t = -1.38, p = .173, df = 58). Comparison between E-CLA and L-CLA also did not disclose a significant difference (E-CLA mean = 51.8, sd = 26.7, N =

30; L-CLA mean = 49.3, sd = 27.9, N = 27; the 95% CI for the difference in means is -

12.00, 16.98 ; t = .344, p = .732, df = 55).

1.4.2 All *Early* vs All *Late* AoE

Further statistical analyses were run to compare the optional resumption scores of those

participants who had *early* exposure to MSA as one group and those who had *late*

exposure to MSA as another group regardless of the colloquial variety. An independent

sample t-test, in SPSS Statistics 19, showed that there was no significant difference

between these two groups (Early Exposure group mean = 50.7, sd = 25.8, N = 60; Late

Exposure group mean = 58.1, sd = 26.1, N = 87; the 95% CI for the difference in means

is -15.98, 1.26; t = -1.69, p = .094, df = 145).

1.5 Further Statistical Analysis: Effect of CV (L1)

1.5.1 Comparing Groups with the Same AoE but Different CVs

To verify if the colloquial variety (the L1) has an effect on the participants' knowledge of

resumption in MSA relative clauses, data from groups with the same AoE but different

CVs were entered into statistical tests. An independent sample t-test was conducted in

SPSS Statistics 19 to compare between the E-CGA and the E-CLA scores in optional

resumption. This test showed no significant difference between these two groups (E-CGA

mean = 49.6, sd = 25.3, N = 30; E-CLA mean = 51.8, sd = 26.7, N = 30; the 95% CI for

the difference in means is -15.70, 11.19; t = -.335, p = .739, df = 58).

This result was expected by the hypothesis here as it was assumed that earlier age of first

exposure to MSA would help the participants to acquire a more native-like knowledge of

resumption in MSA relatives and reduce the effect of the L1 if it exists. In fact, the L1 influence was expected to be found with the participants who had late exposure to MSA. A one way ANOVA was conducted in SPSS Statistics 19 to compare between L-CGA, L-CLA, and L-CEA groups in terms of their scores in optional resumption. The test showed that groups with late exposure to MSA but different CVs had no significant effect on data $(F (2,84) = 2.857, p = .063)$. Tukey HSD and LSD post hoc tests showed that L-CLA and L-CEA groups are significantly different $(p = .050 \& .019$, respectively). This difference, however, cannot be due to L1 influence given that resumption in object relatives in both colloquial varieties is obligatory (see Chapter 3). The descriptive statistics for this test are given in Table 25 below.

Groups	No of Participants	Mean	Sd	95% CI for Mean		Minimum	Maximum
				Lower	Upper		
L-CGA	30	58.4	24.4	49.3	67.6	20.0	100.0
L-CLA	27	49.3	27.9	38.3	60.4	.0	90.0
L-CEA	30	65.5	24.5	56.4	74.7	20.0	100.0
Total	87	58.1	26.1	52.5	63.6	.0	100.0

Table 25: Descriptive statistics for optional resumption scores of groups with Late exposure to MSA

1.5.2 Different CVs Regardless of AoE

Further statistical analyses were run using a One way ANOVA in SPSS Statistics 19 to compare the optional resumption scores achieved by the participants with different CVs regardless of the time of their first exposure to MSA. For this test, the scores of all the participants in this study were divided into 3 groups depending on the colloquial variety they speak: Colloquial Gulf Arabic (CGA), Colloquial Levantine Arabic (CLA) and Colloquial Egyptian Arabic (CEA). The test results showed that these groups with different CVs had a significant effect on the data $(F (2,144) = 3.373, p = .037)$. Table 26 shows the descriptive statistics for these groups.

Groups	No of Participants	Mean	Sd	95% CI for Mean		Minimum	Maximum
				Lower	Upper		
CGA	60	54.0	25.1	47.5	60.5	11.1	100.0
CLA	57	50.6	27.0	43.5	57.8	.0	100.0
CEA	30	65.5	24.5	56.3	74.7	20.0	100.0
Total	147	55.1	26.2	50.8	59.3	.0	100.0

Table 26: Descriptive statistics for groups by CV regardless of AoE in optional resumption

A Tukey HSD post hoc test showed that that there is a significant difference between the CLA and CEA groups ($p = .030$). As explained above, this difference cannot be considered as a result of L1 influence as resumption in both colloquial varieties is obligatory in object relatives (see chapter 3). The difference between CGA and CLA, on one hand, and the difference between CGA and CEA, on the other hand, did not emerge as significant in this post hoc test ($p = .761$ & .115, respectively). Nevertheless, an LSD post hoc test revealed a significant difference between the CGA and CEA groups ($p = .047$). However, this again cannot be considered due to L1 influence given that CEA speakers, who treat resumption in their colloquial variety as obligatory, judged resumption in MSA object relatives as optional more than speakers of CGA in which resumption is optional in object relatives (mean = 65.5 & 54.0, respectively).

1.6 Results of Acceptance & Rejection by Test Item

To check if there were certain test sentences that got accepted or rejected more than the other sentences, the rate of acceptance and rejection was calculated for each test sentence and presented in Table 27 below. As explained in the previous chapter, any sentence judged as +2 or +1 was considered to indicate acceptance whereas choices of -2 or -1 were considered to indicate rejection of that particular test sentence. Responses of 'Can't Decide' did not occur in this data. Refer to Appendix 5 to see the test sentences.

Type	Items	Judgments	E-CGA	L-CGA	E-CLA	L-CLA	L-CEA	Total
Relative Clauses with Resumptive Pronouns	T1A01	Acc. No. (%)	28 (93.3)	29 (96.7)	29 (96.7)	25 (92.6)	30 (100)	141 (95.9)
		Rej. No. (%)	2 (6.7)	1 (3.3)	1 (3.3)	2 (7.4)	0 (0)	6 (4.1)
	T1A02	Acc. No. (%)	29 (96.7)	29 (96.7)	28 (93.3)	26 (96.3)	29 (96.7)	141 (95.9)
		Rej. No. (%)	1 (3.3)	1 (3.3)	2 (6.7)	1 (3.7)	1 (3.3)	6 (4.1)
	T1A03	Acc. No. (%)	26 (86.7)	21 (70.0)	24 (80.0)	20 (74.1)	22 (73.3)	113 (76.9)
		Rej. No. (%)	4 (13.3)	9 (30.0)	6 (20.0)	7 (25.9)	8 (26.7)	34 (23.1)
	T1A04	Acc. No. (%)	23 (76.7)	15 (50.0)	15 (50.0)	18 (66.7)	26 (86.7)	97 (66.0)
		Rej. No. (%)	7 (23.3)	15 (50.0)	15 (50.0)	9 (33.3)	4 (13.3)	50 (34.0)
	T1A05	Acc. No. (%)	29 (96.7)	28 (93.3)	24 (80.0)	25 (92.6)	29 (96.7)	135 (91.8)
		Rej. No. (%)	1 (3.3)	2 (6.7)	6 (20.0)	2 (7.4)	1 (3.3)	12 (8.2)
	T1A06	Acc. No. (%)	27 (90.0)	22 (73.3)	26 (86.7)	24 (88.9)	28 (93.3)	127 (86.4)
		Rej. No. (%)	3 (10.0)	8 (26.7)	4 (13.3)	3 (11.1)	2 (6.7)	20 (13.6)
	T1A07	Acc. No. (%)	30 (100)	25 (83.3)	27 (90.0)	25 (92.6)	30 (100)	137 (93.2)
		Rej. No. (%)	0 (0)	5 (16.7)	3 (10.0)	2 (7.4)	0 (0)	10 (6.8)
	T1A08	Acc. No. (%)	29 (96.7)	27 (90.0)	27 (90.0)	23 (85.2)	29 (96.7)	135 (91.8)
		Rej. No. (%)	1 (3.3)	3 (10.0)	3 (10.0)	4 (14.8)	1 (3.3)	12 (8.2)
	T1A09	Acc. No. (%)	29 (96.7)	30 (100)	25 (83.3)	23 (85.2)	30 (100)	137 (93.2)
		Rej. No. (%)	1 (3.3)	0 (0)	5 (16.7)	4 (14.8)	0 (0)	10 (6.8)
	T1A10	Acc. No. (%)	27 (90.0)	30 (100)	29 (96.7)	23 (85.2)	29 (96.7)	138 (93.9)
		Rej. No. (%)	3 (10.0)	0 (0)	1 (3.3)	4 (14.8)	1 (3.3)	9 (6.1)
Relative Clauses without Resumptive Pronouns	T1B01	Acc. No. (%)	10 (33.3)	13 (43.3)	10 (33.3)	7 (25.9)	16 (53.3)	56 (38.1)
		Rej. No. (%)	20 (66.6)	17 (56.7)	20 (66.6)	20 (74.1)	14 (46.7)	91 (61.9)
	T1B02	Acc. No. (%)	14 (46.7)	18 (60.0)	17 (56.7)	17 (63.0)	18 (60.0)	84 (57.1)
		Rej. No. (%)	16 (53.3)	12 (40.0)	13 (43.3)	10 (37.0)	12 (40.0)	63 (42.9)
	T1B03	Acc. No. (%)	11 (36.7)	11 (36.7)	12 (40.0)	11 (40.7)	19 (63.3)	64 (43.5)
		Rej. No. (%)	19 (63.3)	19 (63.3)	18 (60.0)	16 (59.3)	11 (36.7)	83 (56.5)
	T1B04	Acc. No. (%)	12 (40.0)	8 (26.7)	6 (20.0)	10 (37.0)	17 (56.7)	53 (36.1)
		Rej. No. (%)	18 (60.0)	22 (73.3)	24 (80.0)	17 (63.0)	13 (43.3)	94 (63.9)
	T1B05	Acc. No. (%)	10 (33.3)	14 (46.7)	12 (40.0)	10 (37.0)	16 (53.3)	62 (42.2)
		Rej. No. (%)	20 (66.6)	16 (53.3)	18 (60.0)	17 (63.0)	14 (46.7)	85 (57.8)
	T1B06	Acc. No. (%)	6 (20.0)	12 (40.0)	15 (50.0)	7 (25.9)	16 (53.3)	56 (38.1)
		Rej. No. (%)	24 (80.0)	18 (60.0)	15 (50.0)	20 (74.1)	14 (46.7)	91 (61.9)
	T1B07	Acc. No. (%)	12 (40.0)	11 (36.7)	9 (30.0)	12 (44.4)	19 (63.3)	63 (42.9)
		Rej. No. (%)	18 (60.0)	19 (63.3)	21 (70.0)	15 (55.6)	11 (36.3)	84 (57.1)
	T1B08	Acc. No. (%)	15 (50.0)	12 (40.0)	14 (46.7)	8 (29.6)	12 (40.0)	61 (41.5)
		Rej. No. (%)	15 (50.0)	18 (60.0)	16 (53.3)	19 (70.4)	18 (60.0)	86 (58.5)
	T1B09	Acc. No. (%)	28 (93.3)	30 (100)	26 (86.7)	25 (92.6)	30 (100)	139 (94.6)
		Rej. No. (%)	2 (6.7)	0 (0)	4 (13.3)	2 (7.4)	0 (0)	8 (5.4)
	T1B10	Acc. No. (%)	28 (93.3)	28 (93.3)	26 (86.7)	20 (74.1)	28 (93.3)	130 (88.4)
		Rej. No. (%)	2 (6.7)	2 (6.7)	4 (13.3)	7 (25.9)	2 (6.7)	17 (11.6)

Table 27: Rate of acceptance and rejection of test sentences with object relative clauses

The figures in Table 27 show no big differences between the groups of participants in terms of accepting or rejecting a certain test sentence. Also, these figures show that object relative clauses with resumptive pronouns were generally more accepted than those without resumptive pronouns. The average of accepting object relatives with resumptive pronouns by all the participants is 88.5% (range: 66.0% - 95.9%) compared to 52.2% (range: 36.1% - 94.6%) as the average of accepting these sentences without resumptive pronouns. This reflects the expected preference for resumption in object relatives noted in chapter 2 (section 2.3).

Although the general rate of acceptance of relative clauses without resumption is low, the last two sentences (T1B09 & T1B10) achieved a remarkably higher rate of acceptance (94.6% & 88.4%, respectively, compared with <58% on the other T1B tokens). The relative clauses in these two sentences were with the relativisers _maa_ and _man_, respectively, which have different characteristics from the other relativisers used in the rest of the test sentences. As explained in chapter 2, MSA uses different relativisers depending on the gender and the number of the referent like _ʔallaðii_ for masculine singular, _ʔallatii_ for feminine singular, _ʔallaðiina_ for masculine plural, etc. However, _maa_ and _man_ do not have these gender and number specifications and can be used with referents of any gender and number. One difference between these two relativisers though is that _maa_ is used with inanimate referents whereas _man_ is used with animate referents. The test items results here seem to suggest that relative clauses using these two relativisers are more accepted in MSA than with other relativisers when the resumptive pronoun is absent.

189

The next section will present the judgment results of SV sentences with and without plural agreement between the collective human subject and the following verb.

2 Results of Collective Subject-Verb Agreement in SV Sentences

One section of the test items in the Acceptability Judgment Task (AJT) included declarative sentences in subject-verb order with human subjects of the collective type with and without subject-verb agreement in number. The aim of including such sentences in this task was to test the participants' underlying knowledge of this syntactic phenomenon in MSA and see whether AoE and\or the CV have an effect on their ultimate attainment of the MSA grammar. The participants were asked in this task to use their metalinguistic intuitions to judge 10 sentences. These sentences were repeated twice in the test; once with plural agreement between the human collective subject and the following verb, and once without agreement.

As explained in the scoring section of chapter 4, these 10 pairs of sentences were scored the same way the resumption sentences were scored; scoring was based on comparing judgments of each sentence in each pair. Such comparisons resulted in one of four possible scenarios: i) plural subject-verb agreement is optional, when both sentences were judged acceptable; ii) plural subject-verb agreement is obligatory, when only the sentence without agreement was rejected; iii) plural subject-verb agreement is not allowed, when only the sentence with agreement was rejected; and iv) plural subject-

verb agreement is not relevant to judgment, when both sentences were judged unacceptable.[43] Table 28 below shows the results of such comparisons.

Plural Agreement →	Optional	Obligatory	Not Allowed	Irrelevant	Total No. of Judged Pairs
Sentence Pairs N (%)	980 (66.7)	311 (21.2)	116 (7.9)	63 (4.3)	1470

Table 28: Judgments on sentence pairs in terms of collective subject-verb plural agreement in SV order in MSA (all groups collapsed together)

2.1 Excluded Data

Based on the figures shown in Table 28, 87.9% of the judged pairs show that plural agreement between the human collective subject and the following verb is either optional or obligatory. Data that show such agreement is not allowed were excluded from the analysis because, in fact, it is allowed in all varieties of Arabic under study, standard and non-standard (see chapter 3, section 1.2). Similarly, data in which both sentences of one pair were rejected (labeled as 'irrelevant' in Table 28) were excluded from the analysis. This is because the participants who made these judgments should have accepted at least one of the sentences in these pairs if their judgments were based on whether or not plural agreement between the human collective subject and the following verb is permissible. Since both sentences in these pairs were rejected, it seems that these judgments were based on something irrelevant to the syntactic phenomenon under study. The types of response that indicated that agreement is not allowed or rejected both test items of one pair seem to have occurred as random errors and, hence, were treated as noise.

---

[43] There is a fifth possible scenario and that is when one or both of the sentences were not decided upon by choosing 'Can't decide' option from the scale. This scenario did not occur in this data.

2.2 Effect of Test Items Orders & Results of Merged Subgroups

As explained in chapter 4, to control for the possibility that the participants might judge the sentences appearing second of each pair based on their judgments of the sentence that appeared first in the test, the participants in each group were divided into two subgroups and saw the AJT test items in two different orders. Table 29 below shows the subgroups' judgments after exclusion of the above mentioned data.

Plural Agreement →		Optional	Obligatory	Total Number of Judged pairs
Groups	Order of Test Items	No. (%)	No. (%)	
E-CGA	Order 1	105 (73.9)	37 (26.1)	142
	Order 2	91 (68.4)	42 (31.6)	133
L-CGA	Order 1	108 (79.4)	28 (20.6)	136
	Order 2	87 (70.2)	37 (29.8)	124
E-CLA	Order 1	114 (83.8)	22 (16.2)	136
	Order 2	97 (77.0)	29 (23.0)	126
L-CLA	Order 1	97 (77.6)	28 (22.4)	125
	Order 2	78 (78.8)	21 (21.2)	99*
L-CEA	Order 1	101 (75.9)	32 (24.1)	133
	Order 2	102 (74.5)	35 (25.5)	137
All Groups Total		980 (75.9)	311 (24.1)	1291

Table 29: Subgroups' judgments on plural agreement after exclusion of irrelevant data

* The number of participants in this sub-group is only 12 compared to 15 in all other sub-groups.

To see if there was an effect of test items order in the test, five independent samples t-tests were conducted on separate data, using the software IBM SPSS Statistics 19, to compare the means of optional agreement scores achieved by the two subgroups in each group. The results of these t-tests came out as there are no significant differences found for any group. The descriptive statistics and the results of these five t-tests are in Table 30 below.

t-test	Groups	Order of Test Items	No.	Mean	sd	t	df	p	95% CI for the difference in means
1	E-CGA	Order 1	15	74.20	19.67	.924	28	.363	-9.13, 24.16
		Order 2	15	66.69	24.57				
2	L-CGA	Order 1	15	78.80	19.02	1.243	28	.224	-6.39, 26.11
		Order 2	15	68.94	24.13				
3	E-CLA	Order 1	15	83.63	13.72	1.147	28	.261	-5.41, 19.20
		Order 2	15	76.74	18.80				
4	L-CLA	Order 1	15	76.48	23.17	-.037	25	.970	-16.52, 15.93
		Order 2	12	76.78	16.02				
5	L-CEA	Order 1	15	76.80	20.54	.415	28	.681	-12.10, 18.26
		Order 2	15	73.72	20.05				

Table 30: t-tests results for the order effect on judging optional collective subject-verb agreement

Having no significant difference between the two subgroups in each group, it was concluded that the order of the test items in the test did not affect the participants' judgments. Table 31 below shows the group results after merging the two subgroups in each group.
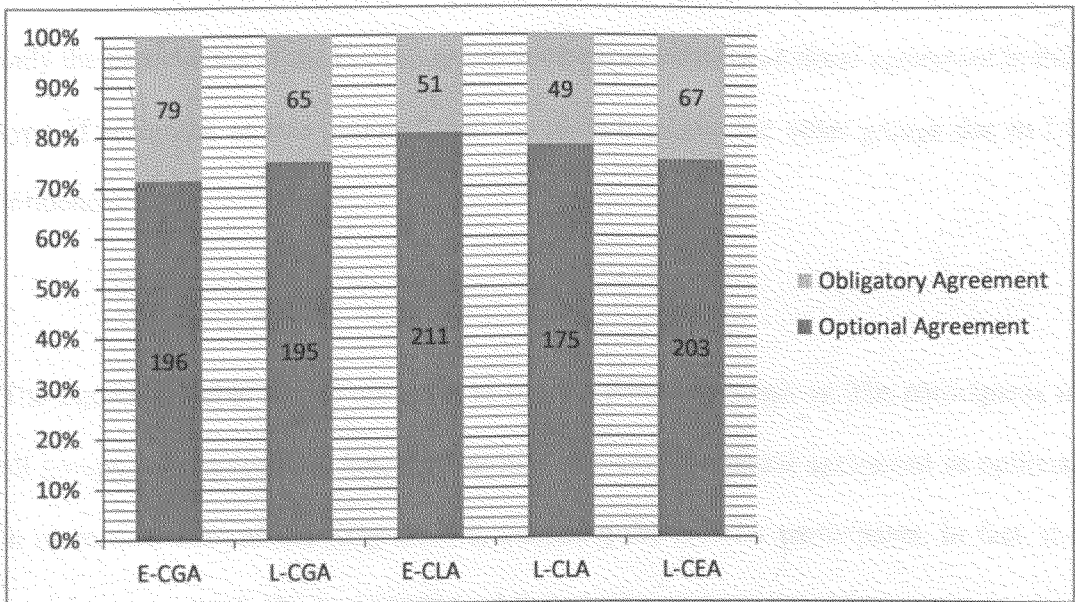
Agreement →	Optional	Obligatory	Total Number of Judged pairs
Groups	No. (%)	No. (%)	
E-CGA	196 (71.3)	79 (28.7)	275
L-CGA	195 (75.0)	65 (25.0)	260
E-CLA	211 (80.5)	51 (19.5)	262
L-CLA	175 (78.1)	49 (21.9)	224
L-CEA	203 (75.2)	67 (24.8)	270
All Groups Total	980 (75.9)	311 (24.1)	1291

Table 31 : Groups' judgments on human collective subject-verb plural agreement in SV sentences

It was expected by the hypothesis of this thesis that early exposure to MSA will help to reduce L1 influence and increase the chance of obtaining a more native-like knowledge of MSA grammar. Thus, the participants in E-CGA and E-CLA groups were expected to treat plural agreement between the human collective subjects and the following verbs as optional in MSA regardless of whether it is optional or obligatory in their colloquial varieties. Also, it was expected that the grammar of the colloquial variety (the L1) will affect how the participants with late exposure to MSA will judge the agreement sentences. Based on this, the participants in the L-CGA group were expected to judge the human collective subject-verb agreement as obligatory as it is in their colloquial variety (see section 1.2.1 in chapter 3). On the contrary, the participants in the L-CLA and the L-CEA groups were expected to continue considering this agreement as optional as it is in their colloquial varieties (see section 1.2.1 in chapter 3). In short, only the participants in the L-CGA group were expected to treat plural agreement in this type of sentences as less optional than the participants in the other groups due to L1 influence.

The figures in Table 31 do not seem to support these expectations. The participants in all groups have judged the human collective subject-verb plural agreement as optional in more than 70% of the judged pairs, including the L-CGA participants. In fact, the participants in the L-CGA group judged the agreement as optional in more pairs than the participants in the E-CGA group who had earlier exposure to MSA (75.0% vs 71.3%, respectively). The rate of judging this agreement as optional by the two CLA groups and the L-CEA group was as high as expected due to the fact that plural agreement is optional in the Egyptian and the Levantine colloquial varieties.

2.3 General Statistical Effect of Independent Variables & Statistical Group Comparisons

Using the IBM SPSS Statistics 19, a Univariate Analysis of Variance test was conducted to check statistically the effect of the colloquial varieties and the age of first exposure to MSA on the data of optional agreement. The results of this test showed that there is no significant effect of these two independent variables (for CV: $F_{(2,142)} = 1.388$, $p = 0.253$; for AoE: $F_{(1,142)} = .000$, $p = 0.983$). The interaction of CV and AoE was not significant either ($F_{(1,142)} = 0.872$, $p = 0.352$).

To compare the means of the five groups on judging plural agreement as optional in MSA, a One Way ANOVA was conducted using the SPSS Statistics 19 with the groups as the factor and the optional agreement as the dependent variable. This test showed no significant effect of different groups on data ($F_{(4,142)} = 0.934$, $p = 0.446$). The descriptive statistics for this test are given in Table 32 below.

Groups	No of Participants	Mean	Sd	95% CI for Mean		Minimum	Maximum
				Lower	Upper		
E-CGA	30	70.5	22.2	62.2	78.7	12.5	100.0
L-CGA	30	73.9	21.9	65.7	82.1	28.6	100.0
E-CLA	30	80.1	16.5	74.0	86.4	33.3	100.0
L-CLA	27	76.6	19.9	68.7	84.5	12.5	100.0
L-CEA	30	75.3	20.0	67.8	82.7	25.0	100.0
Total	147	75.250	20.2172	71.9	78.5	12.5	100.0

Table 32: Descriptive statistics for groups' scores on judging plural agreement as optional

It seems that the statistical tests conducted so far suffice to answer the research questions and controvert the research hypotheses about knowledge of collective subject-verb agreement in SV sentences. However, further statistical analyses were conducted to measure the effect of AoE and CV (the L1) separately by making various specific

comparisons and manipulating groups of participants. This was done to make sure that every possible analysis was conducted. These analyses are presented in the following two subsections.

2.4 Further Statistical Analysis: Effect of AoE

2.4.1 Comparing Groups with the Same CV but Different AoE

To check further if AoE has an effect on the participants' underlying knowledge of MSA plural agreement between a human collective subject and the following verb, optional plural agreement data of groups with the same CV but different AoE were analysed by conducting two separate independent samples t-tests; one comparing the E-CGA with the L-CGA group, and the other comparing the E-CLA with the L-CLA group. The results of the test comparing E-CGA with L-CGA showed no significant difference between the two groups (E-CGA mean = 70.5, sd = 22.2, N = 30; L-CGA mean = 73.9, sd = 21.9, N = 30; the 95% CI for the difference in means is -14.8, 7.9; t = .600, p = .551, df = 58). The difference between the E-CLA and the L-CLA also was not significant according to this test (E-CLA mean = 80.2, sd = 16.5, N = 30; L-CLA mean = 76.6, sd = 19.9, N = 27; the 95% CI for the difference in means is – 6.11, 13, 26; t = .739, p = .463, df = 55).

2.4.2 All *Early* vs All *Late* AoE

Further statistical analysis were run to compare the scores of optional plural agreement of those participants who had *early* exposure to MSA as one group with the scores of those who had *late* exposure to MSA as another group regardless of the colloquial varieties they speak. An independent samples t-test showed no significant difference between these two groups (Early Exposure group mean = 75.3, sd = 20.02, N = 60; Late

Exposure group mean = 75.2, sd = 20.46, N = 87; the 95% CI for the difference in means is -6.61, 6.84; t = .034, p = .973, df = 145).

2.5 Further Statistical Analysis: Effect of CV (L1)

2.5.1 Comparing Groups with the Same AoE but Different CVs

Optional plural agreement data of groups with the same AoE and different CVs were analysed statistically to verify if the colloquial variety (the L1) had an effect on the participants' underlying knowledge of MSA grammar. An independent samples t-test was run in SPSS Statistics 19 to compare the means of the E-CGA and the E-CLA groups. As expected due to early exposure to MSA, the difference between these two groups came out as not significant despite the difference between the two colloquial varieties in terms of plural agreement between the human collective subject and the following verb being obligatory in CGA and optional in CLA (E-CGA mean = 70.5, sd = 22.20, N = 30; E-CLA mean = 80.2, sd = 16.5, N = 30; the 95% CI for the difference in means is -19.85, .38; t = -1.926, p = .059, df = 58).

If a group difference due to L1 influence exists, it is expected to be found between groups with late exposure rather than between groups with early exposure to MSA. Therefore, optional plural agreement data of L-CGA, L-CLA and L-CEA groups were analysed by a One Way ANOVA in SPSS Statistics 19 to see if the L-CGA participants did significantly different from the participants in the other two groups due to L1 influence. Despite the fact that the participants in the L-CGA group achieved the lowest mean in optional plural agreement (see Table 33 below), the test showed that groups with late exposure to MSA but different CVs had no significant effect on data (F $(2,84)$ = .125, p =

.883). Post hoc tests like Tukey HSD, LSD, and Bonferroni showed no significant difference between the three groups ($p = <.05$). The descriptive statistics for these groups are given in Table 33 below.

Groups	No of Participants	Mean	Sd	95% CI for Mean		Minimum	Maximum
				Lower	Upper		
L-CGA	30	73.9	21.9	65.68	82.05	28.6	100.0
L-CLA	27	76.6	19.9	68.72	84.50	12.5	100.0
L-CEA	30	75.3	20.0	67.79	82.73	25.0	100.0
Total	87	75.2	20.5	70.84	79.56	12.5	100.0

Table 33: Descriptive statistics for optional agreement scores of groups with *Late* exposure to MSA

2.5.2 Different CVs Regardless of AoE

One more statistical test was conducted using SPSS Statistics 19 to compare the optional plural agreement means achieved by the participants with different CVs regardless of the time of first exposure to MSA. For this test, the scores of all the participants in this study were divided into three groups depending on the colloquial variety the participants speak: CGA, CLA, and CEA. A One Way ANOVA showed that these groups with different CVs had no significant effect on data (F (2, 144) = 1.443, $p = .240$). Post hoc tests results showed no significant differences between the groups ($p = >.05$). The descriptive statistics for these groups are in Table 34 below.

Groups	No of Participants	Mean	Sd	95% CI for Mean		Minimum	Maximum
				Lower	Upper		
CGA	60	72.2	21.94	66.49	77.83	12.5	100.0
CLA	57	78.5	18.15	73.67	83.31	12.5	100.0
CEA	30	75.3	20.01	67.79	82.73	25.0	100.0
Total	147	75.3	20.21	71.95	78.54	12.5	100.0

Table 34 : Descriptive Statistics for groups by CV regardless of AoE in optional plural agreement

2.6 Results of Acceptance & Rejection by Test Item

To check if there are certain test sentences that get accepted or rejected more than the other sentences, the rate of acceptance and rejection was calculated for each test sentence and presented in Table 35 below. As explained in the previous chapter, any sentence was judged as +2 or +1 was considered to indicate acceptance whereas choices of -2 or -1 was considered to indicate rejection of that particular test sentence. Responses of 'Can't Decide' did not occur in this data. Refer to Appendix 5 to see the test sentences.

The figures in Table 35 show that the rate of acceptance or rejection of a certain test sentences does not change dramatically across the five groups. Also, the sentences with plural agreement between the collective subject and the following verb seem to be more accepted than those without subject-verb plural agreement. The average of accepting sentences with plural agreement by all participants is 87.9% (range: 72.1% - 95.2%) whereas the average of accepting the sentences without plural agreement is 74.5% (range: 49.0% - 90.5).

Type	Item	Judgments	E-CGA	L-CGA	E-CLA	L-CLA	L-CEA	Total
Collective Subject-Verb with Plural Agreement	T2A01	Acc. No. (%)	30 (100)	28 (93.3)	28 (93.3)	19 (70.4)	27 (90.0)	132 (89.8)
		Rej. No. (%)	0 (0)	2 (6.7)	2 (6.7)	8 (29.6)	3 (10.0)	15 (10.2)
	T2A02	Acc. No. (%)	29 (96.7)	28 (93.3)	27 (90.0)	24 (88.9)	26 (86.7)	134 (91.2)
		Rej. No. (%)	1 (3.3)	2 (6.7)	3 (10.0)	3 (11.1)	4 (13.3)	13 (8.8)
	T2A03	Acc. No. (%)	25 (83.3)	22 (73.3)	22 (73.3)	16 (59.3)	21 (70.0)	106 (72.1)
		Rej. No. (%)	5 (16.7)	8 (26.7)	8 (26.7)	11 (40.7)	9 (30.0)	41 (27.9)
	T2A04	Acc. No. (%)	29 (96.7)	29 (96.7)	29 (96.7)	24 (88.9)	29 (96.7)	140 (95.2)
		Rej. No. (%)	1 (3.3)	1 (93.3)	1 (3.3)	3 (11.1)	1 (3.3)	7 (4.8)
	T2A05	Acc. No. (%)	27 (90.0)	28 (93.3)	27 (90.0)	24 (88.9)	29 (96.7)	135 (91.8)
		Rej. No. (%)	3 (10.0)	2 (6.7)	3 (10.0)	3 (11.1)	1 (93.3)	12 (8.2)
	T2A06	Acc. No. (%)	29 (96.7)	26 (86.7)	29 (96.7)	26 (96.3)	28 (93.3)	138 (93.9)
		Rej. No. (%)	1 (3.3)	4 (13.3)	1 (93.3)	1 (3.7)	2 (6.7)	9 (6.1)
	T2A07	Acc. No. (%)	29 (96.7)	27 (90.0)	25 (83.3)	26 (96.3)	28 (93.3)	135 (91.8)
		Rej. No. (%)	1 (3.3)	3 (10.0)	5 (16.7)	1 (3.7)	2 (6.7)	12 (8.2)
	T2A08	Acc. No. (%)	29 (96.7)	26 (86.7)	26 (86.7)	22 (81.5)	30 (100)	133 (90.5)
		Rej. No. (%)	1 (3.3)	4 (13.3)	4 (13.3)	5 (18.5)	0 (0)	14 (9.5)
	T2A09	Acc. No. (%)	19 (63.3)	25 (83.3)	24 (80.0)	21 (77.8)	26 (86.7)	115 (78.2)
		Rej. No. (%)	11 (36.7)	5 (16.7)	6 (20.0)	6 (22.2)	4 (13.3)	32 (21.8)
	T2A10	Acc. No. (%)	29 (96.7)	22 (73.3)	25 (83.3)	22 (81.5)	26 (86.7)	124 (84.4)
		Rej. No. (%)	1 (3.3)	8 (26.7)	5 (16.7)	5 (18.5)	4 (13.3)	23 (15.6)
Collective Subject-Verb without Plural Agreement	T2B01	Acc. No. (%)	25 (83.3)	27 (90.0)	27 (90.0)	23 (85.2)	24 (80.0)	126 (85.7)
		Rej. No. (%)	5 (16.7)	3 (10.0)	3 (10.0)	4 (14.8)	6 (20.0)	21 (14.3)
	T2B02	Acc. No. (%)	26 (86.7)	25 (83.3)	25 (83.3)	21 (77.8)	25 (83.3)	122 (83.0)
		Rej. No. (%)	4 (13.3)	5 (16.7)	5 (16.7)	6 (22.2)	5 (16.7)	25 (17.0)
	T2B03	Acc. No. (%)	18 (60.0)	17 (56.7)	25 (83.3)	22 (81.5)	15 (50.0)	97 (66.0)
		Rej. No. (%)	12 (40.0)	13 (43.3)	5 (16.7)	5 (18.5)	15 (50.0)	50 (34.0)
	T2B04	Acc. No. (%)	28 (93.3)	26 (86.7)	27 (90.0)	22 (81.5)	26 (86.7)	129 (87.8)
		Rej. No. (%)	2 (6.7)	4 (13.3)	3 (10.0)	5 (18.5)	4 (13.3)	18 (12.2)
	T2B05	Acc. No. (%)	23 (76.7)	23 (76.7)	29 (96.7)	24 (88.9)	27 (90.0)	126 (85.7)
		Rej. No. (%)	7 (23.3)	7 (23.3)	1 (3.3)	3 (11.1)	3 (10.0)	21 (14.3)
	T2B06	Acc. No. (%)	10 (33.3)	19 (63.3)	20 (66.7)	15 (55.6)	18 (60.0)	82 (55.8)
		Rej. No. (%)	20 (66.7)	11 (36.7)	10 (33.3)	12 (44.4)	12 (40.0)	65 (44.2)
	T2B07	Acc. No. (%)	13 (43.3)	15 (50.0)	18 (60.0)	15 (55.6)	11 (36.7)	72 (49.0)
		Rej. No. (%)	17 (56.7	15 (50.0)	12 (40.0)	12 (44.4)	19 (63.3)	75 (51.0)
	T2B08	Acc. No. (%)	17 (56.7)	14 (46.7)	16 (53.3)	18 (66.7)	17 (56.7)	82 (55.8)
		Rej. No. (%)	13 (43.3)	16 (53.3)	14 (46.7)	9 (33.3)	13 (43.3)	65 (44.2)
	T2B09	Acc. No. (%)	25 (83.3)	24 (80.0)	26 (86.7)	23 (85.2)	28 (93.3)	126 (85.7)
		Rej. No. (%)	5 (16.7)	6 (20.0)	4 (13.3)	4 (14.8)	2 (6.7)	21 (14.3)
	T2B10	Acc. No. (%)	29 (96.7)	24 (80.0)	25 (83.3)	26 (96.3)	29 (96.7)	133 (90.5)
		Rej. No. (%)	1 (3.3)	6 (20.0)	5 (16.7)	1 (3.7)	1 (3.3)	14 (9.5)

Table 35 : Rate of acceptance & rejection of test sentences with\out collective subject-verb agreement

The following section presents the results of producing declarative sentences with VSO or SVO orders in conversational and narrative contexts.

3   Results of Word Order in Conversational & Narrative Contexts

The participants' underlying knowledge of word order in MSA was assessed in this experimental study by completing the 'conversation role-play' task. As explained in chapter 4, the participants in this task formed and produced 20 declarative sentences with verbs and overt subjects; 10 in a conversational context and the other 10 in a narrative context. The aim was to check which order the participants would use in which context and if the L1 would influence their choice of word order due to time of first exposure to MSA.

As explained in chapter 4, the participants in each group were divided into two subgroups and completed two different versions of the conversation role-play task. The only difference between the two versions was the order of word boxes from which the participants should choose their vocabulary to form sentences in the conversational context; if the box of nouns preceded the box of verbs in a certain conversational turn in version 1, this order was reversed in version 2. This was done to control for the possible effect of the linear order of the word boxes on the participants' choice of word order (see section 2.2 of the previous chapter for a thorough description of this task). Table 36 below shows the distribution of word orders used by the subgroups participants in the conversational context in CR-P task.

Word Order →		VSO	SVO	Total Number of Formed Sentences
Groups	Order of word boxes	No. (%)	No. (%)	
E-CGA	Order 1	94 (62.7)	56 (37.3)	150
	Order 2	101 (67.3)	49 (32.7)	150
L-CGA	Order 1	105 (70.0)	45 (30.0)	150
	Order 2	105 (70.0)	45 (30.0)	150
E-CLA	Order 1	100 (66.7)	50 (33.3)	150
	Order 2	101 (67.3)	49 (32.7)	150
L-CLA	Order 1	107 (71.3)	43 (28.7)	150
	Order 2	89 (74.2)	31 (25.8)	120*
L-CEA	Order 1	113 (75.3)	37 (24.7)	150
	Order 2	85 (56.7)	65 (43.3)	150
All Groups Total		1000 (68.0)	470 (32.0)	1470

Table 36: Distribution of word orders used by the subgroups' participants in conversational context in CR-P task

* The number of participants in this sub-group is only 12 compared to 15 in all other sub-groups.

3.1 Effect of Order of Word Boxes in Conversational Context

To check if the order of the word boxes had an effect on the participants' choice of word order, five independent samples t-tests were conducted on separate data using the SPSS Statistics 19 to compare between the two subgroups in each group in terms of the means of using VSO order in the conversational context. The results of these tests came out as there are no significant differences found for any group except L-CEA. The descriptive statistics and the results of these five t-tests are in Table 37 below.

t-test	Groups	Order of Word Boxes	No.	Mean	sd	t	df	p	95% CI for the difference in means
1	E-CGA	Order 1	15	62.7	25.5	-.564	28	.577	-21.62, 12.28
		Order 2	15	67.3	19.4				
2	L-CGA	Order 1	15	70.0	19.3	.000	28	1.000	-13.85, 13.85
		Order 2	15	70.0	17.7				
3	E-CLA	Order 1	15	66.7	26.4	-.072	28	.943	-19.64, 18.31
		Order 2	15	67.3	24.3				
4	L-CLA	Order 1	15	71.3	13.0	-.514	25	.612	-14.18, 8.52
		Order 2	12	74.2	15.6				
5	L-CEA	Order 1	15	75.3	15.1	3.004	28	.006	5.94, 31.39
		Order 2	15	56.7	18.8				

Table 37 : t-tests results for the boxes order effect on choosing VSO in conversational context

However, this difference between the two subgroups of the L-CEA group seems unlikely to be caused by the order of the word boxes given the fact that different orders of word boxes did not cause differences between the subgroups of the other four groups. Thus, it was concluded that order of the word boxes had no effect on how the participants formed their sentences in the conversational part of the conversation role-play task.

3.2 Results of Merged Subgroups in Conversational & Narrative Contexts

Table 38 below shows the distribution of word order in both conversational and narrative contexts after merging the two subgroups in each group. Recall from chapter 2 that VSO order is the common order in MSA (e.g., Bakir, 1980; Fassi-Fehri, 1993). Also, context has an effect on which order to use making VSO more common in narrative contexts whereas SVO becomes more common in conversational contexts (e.g., Brustad, 2000; Dahlgren, 1998). The results of the colloquial Arabic corpus study in chapter 3 suggested that at least colloquial Egyptian Arabic (CEA) is different from MSA in terms of word order preference. The Egyptian speakers used SVO order in more than 90% of the sentences they produced in that corpus regardless of the context (see section 1.1.2 of chapter 3).

The hypothesis here expected that only those participants who had *early* exposure to MSA would not show L1 effects when producing MSA sentences. Thus, based on the results of the colloquial Arabic corpus study in chapter 3, it was expected that only the participants of the L-CEA group would show L1 effects by using SVO order in MSA more than the other orders regardless of the context, due to L1 influence. On the other

hand, the participants of the CGA and the CLA groups were expected to use word orders depending on the context as is the case in MSA; SVO more in conversational contexts than in narrative contexts and VSO more in narrative contexts than in conversational contexts.
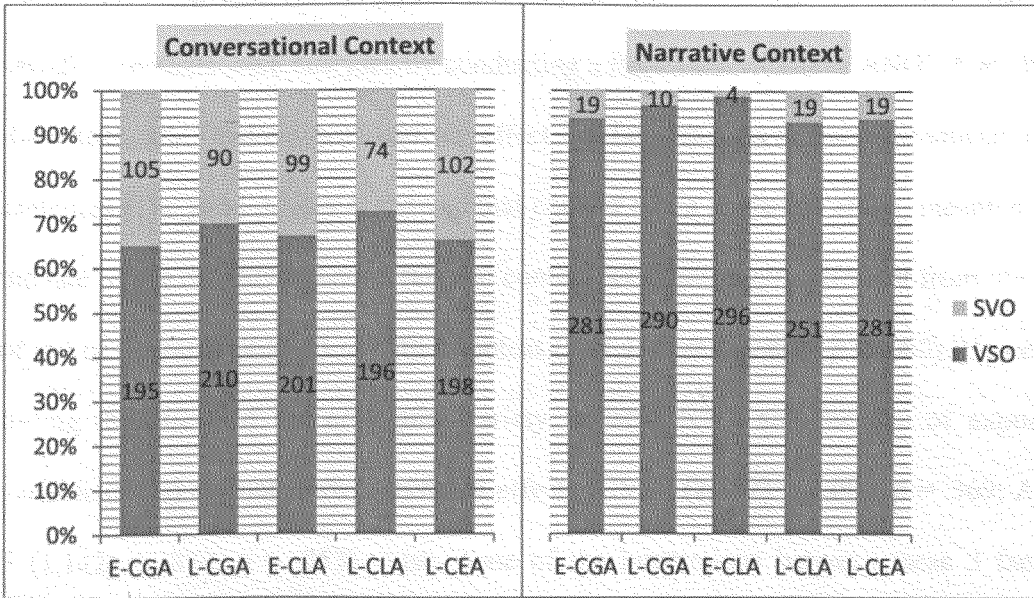
Context →	Conversational		Total Sentences	Narrative		Total sentences
	VSO	SVO		VSO	SVO	
Groups	No. (%)	No. (%)	No.	No. (%)	No. (%)	No.
E-CGA	195 (65.0)	105 (35.0)	300	281 (93.7)	19 (6.3)	300
L-CGA	210 (70.0)	90 (30.0)	300	290 (96.7)	10 (3.3)	300
E-CLA	201 (67.0)	99 (33.0)	300	296 (98.7)	4 (1.3)	300
L-CLA	196 (72.6)	74 (27.4)	270	251 (93.0)	19 (7.0)	270
L-CEA	198 (66.0)	102 (34.0)	300	281 (93.7)	19 (6.3)	300
Total	1000 (68)	470 (32)	1470	1399 (95.2)	71 (4.8)	1470

Table 38 : Distribution of word order by groups in conversational and narrative contexts

The figures in Table 38 show that VSO was used more than SVO order in both contexts by all the participants of this study (68% vs 32% in conversational; 95.2% vs 4.8% in narrative context). Also, SVO order was used by all the participants more in conversational context than in narrative context (32% vs 4.8%, respectively). Contrary to the expectations of the hypothesis here, the participants of the L-CEA group did not

use SVO order more than VSO in either context in this task despite their *late* exposure to MSA.

3.3 General Statistical Effect of Context, CV, and AoE

Using SPSS Statistics 19, the data concerning use of VSO order in conversational and narrative contexts were analysed by conducting a Repeated-Measures ANOVA to check if context, CV, and/or AoE had an effect on how the participants produced their sentences. The R-M ANOVA showed that context had a statistical effect, meaning that the rate of use of VSO in conversational context was statistically different from the rate of use of the same order in narrative context ($F (1,142) = 219.000$, $p = .000$). However, testing the effect of the colloquial variety variable (CV) and the age of exposure variable (AoE) did not show significant effects (CV: $F (2,142) = 1.005$, $p = .369$; AoE: $F (1,142) = .870$, $p = .353$). Also, none of the interactions between these 3 factors revealed any significant effects; all came out with p values greater than .05. The descriptive statistics for this test are given in Table 39 below.

Conversational Context					
Groups →	E-CGA	L-CGA	E-CLA	L-CLA	E-CEA
VSO Mean	65.0	70.0	67.0	72.6	66.0
sd	22.4	18.2	24.9	14.0	19.2
No.	30	30	30	27	30
Narrative Context					
Groups →	E-CGA	L-CGA	E-CLA	L-CLA	E-CEA
VSO Mean	93.7	96.7	98.7	93.0	93.7
sd	10.7	6.6	4.3	13.5	11.6
No.	30	30	30	27	30

Table 39: Descriptive statistics for Repeated-Measures ANOVA

3.4 Statistical Group Comparisons

To compare the means of the five groups when using VSO order in conversational contexts, a One Way ANOVA was conducted in SPSS Statistics 19 with the groups

being the factor and VSO means in conversational context being the dependent variable. This test showed no significant effect of different groups on data (F (4,142) = .679, p = .608). Another One Way ANOVA was conducted to compare the means of using VSO in narrative context. This test also showed no significant effect of groups on data (F (4,142) = 1.808, p = .131). Post hoc tests showed no significant differences between the five groups in both conversational and narrative contexts (p = >.05 in all comparisons).

3.5 Results of Word Order in Conversational Context by Conversational Turns

To check if the participants of the study preferred to use a certain word order in a particular conversational turn in the conversational context, the rate of using VSO and SVO orders was calculated for each conversational turn and presented in Table 40 below.

Based on Table 40, the rate of use of a certain order in a particular conversational turn did not vary dramatically across the five groups. The rate of use of VSO order by the participants of all groups was always higher than the rate of use of SVO order in all conversational turns except for Conv01 where SVO was used more than VSO. This should be expected as in this particular conversational turn, the other speaker was asking the participant a question that required information about a person. The question asked was 'Has anybody seen Majid today?'. The answer to such a question would normally start with the subject to provide information about who saw Majid that day (See Appendix 6 for the content of this task).

Item	Order	E-CGA	L-CGA	E-CLA	L-CLA	L-CEA	Total
Conv01	VSO No. (%)	8 (26.7)	8 (26.7)	8 (26.7)	5 (18.5)	7 (23.3)	36 (24.5)
	SVO No. (%)	22 (73.3)	22 (73.3)	22 (73.3)	22 (81.5)	23 (76.7)	111 (75.5)
Conv02	VSO No. (%)	20 (66.7)	25 (83.3)	24 (80.0)	22 (81.5)	22 (73.3)	113 (76.9)
	SVO No. (%)	10 (33.3)	5 (16.7)	6 (20.0)	5 (18.5)	8 (26.7)	34 (23.1)
Conv03	VSO No. (%)	14 (46.7)	19 (63.3)	19 (63.3)	21 (77.8)	19 (63.3)	92 (62.6)
	SVO No. (%)	16 (53.3)	11 (36.7)	11 (36.7)	6 (22.2)	11 (36.7)	55 (37.4)
Conv04	VSO No. (%)	12 (40.0)	15 (50.0)	18 (60.0)	14 (51.9)	17 (56.7)	76 (51.7)
	SVO No. (%)	18 (60.0)	15 (50.0)	12 (40.0)	13 (48.1)	13 (43.3)	71 (48.3)
Conv05	VSO No. (%)	30 (100)	25 (83.3)	25 (83.3)	26 (96.3)	29 (96.7)	135 (91.8)
	SVO No. (%)	0 (0)	5 (16.7)	5 (16.7)	1 (3.7)	1 (3.3)	12 (8.2)
Conv06	VSO No. (%)	18 (60.0)	23 (76.7)	22 (73.3)	24 (88.9)	25 (83.3)	112 (76.2)
	SVO No. (%)	12 (40.0)	7 (23.3)	8 (26.7)	3 (11.1)	5 (16.7)	35 (23.8)
Conv07	VSO No. (%)	23 (76.7)	25 (83.3)	16 (53.3)	21 (77.8)	21 (70.0)	106 (72.1)
	SVO No. (%)	7 (23.3)	5 (16.7)	14 (46.7)	6 (22.2)	9 (30.0)	41 (27.9)
Conv08	VSO No. (%)	28 (93.3)	25 (83.3)	27 (90.0)	22 (81.5)	19 (63.3)	121 (82.3)
	SVO No. (%)	2 (6.7)	5 (16.7)	3 (10.0)	5 (18.5)	11 (36.7)	26 (17.7)
Conv09	VSO No. (%)	18 (60.0)	20 (66.7)	18 (60.0)	19 (70.4)	15 (50.0)	90 (61.2)
	SVO No. (%)	12 (40.0)	10 (33.3)	12 (40.0)	8 (29.6)	15 (50.0)	57 (38.8)
Conv10	VSO No. (%)	24 (80.0)	25 (83.3)	24 (80.0)	22 (81.5)	24 (80.0)	119 (81.0)
	SVO No. (%)	6 (20.0)	5 (16.7)	6 (20.0)	5 (18.5)	6 (20.0)	28 (19.0)

Table 40: the rate of using VSO & SVO in each conversational turn in the CR-P task

4 Summary

This chapter presented the results of the main experimental study in this thesis described in the previous chapter. The results were detailed according to the three dependent syntactic variables under study: resumption in object relative clauses, collective subject-verb number agreement in declarative SV sentences, and word order in conversational and narrative contexts.

The results of the acceptability judgment task (AJT) were as follows. Total judgments of resumption in object relatives and number agreement between collective subjects and the following verbs showed that resumption and agreement were judged as optional more than obligatory. This was also true of participants' judgments on resumption and

agreement when results of the five groups were considered separately. However, comparison across the five groups in terms of judging resumption and agreement as optional did not reflect the expectations of the hypotheses of this thesis. An SPSS Univariate Analysis of Variance revealed that there was no significant effect of either AoE or CV on the collected data ($p = < .05$). Also, statistical comparison between the data of the five groups using a One Way ANOVA revealed no significant results due to group difference ($p = < .05$). Further statistical analyses were conducted to measure the effect of AoE and CV (the L1) separately by making various specific comparisons and manipulating groups of participants. Most of these analyses revealed no significant results, and when a significant result was found, none could be interpreted as due to the effect of the independent variables.

The results of acceptance or rejection of each test item were presented as well. These results showed no big differences across the five groups in terms of accepting/rejecting a certain test item. Across all groups, object relatives with resumption and sentences with plural agreement between the collective subject and the following verb were found to be accepted more than those without resumption or agreement. One exception was the relatives with the relativisers _maa_ and _man_ which were accepted more that rejected even when the resumptive pronoun was not overt.

The results of the Conversation Role-Play Task were as follows. In total, the rate of using VSO order was more than the rate of using SVO order in both conversational and narrative contexts. Also, SVO order was used more in conversational context than in narrative context. These preferences remained the same when the results of the five

groups were considered separately. Nevertheless, the expectations of the hypotheses that variant AoE and/or CV (the L1) should have effects on data were not supported by analysing the results of the five groups of participants. An SPSS Repeated-Measure ANOVA revealed no significant effect of either AoE or CV (p = <.05). However, this statistical test showed that context had a significant effect on which word order to use (p = .000); the rate of using VSO order in conversational context was found to be significantly different from the rate of using the same order in narrative context. Two 'One Way ANOVA' tests showed no statistical differences across the five groups in terms of using VSO order in either conversational or narrative contexts (p = < .05).

The results of word order distribution in each conversational turn were also presented in this chapter. These results showed no big differences across the five groups of participants in terms of using a particular word order in a particular conversational turn. Also, VSO order was used more than SVO order in all conversational turns except in the first turn where SVO order was used more. This was expected as the participant in this turn was asked to provide an answer to a question that starts with 'who'. Answers to such questions normally start with the subject. The results of this experimental study will be discussed in the following chapter.

Chapter 6

Discussion

As motivated by the frequent claim in the literature that there are no native speakers of MSA, this thesis aimed to investigate the ultimate underlying knowledge of MSA represented by speakers of different Arabic colloquial dialects across three regions of the Arab world. The research questions laid out in chapter 1 seek to answer whether Arabs across the Arab world represent 'uniform' (L1 type) or 'variant' (L2 type) versions of MSA end-state grammar, and if variant, whether this is due to age of first exposure to MSA (AoE), influence of the colloquial dialect (the L1), or both.

In the preceding chapter, the presentation and analysis of the experimental study data were reported. Briefly, the results of this study showed that participant groups of distinct AoE and L1 were found not significantly different in terms of performance in the two distinct tasks of the study which aimed to assess the participants' MSA competence (AJT) and their use of it (CR-P). Also, the results showed that there was no statistical effect of the two independent variables: AoE and the colloquial dialect.

This chapter consists of discussion of the experimental study findings in relation to the research questions and related literature. The first section discusses the main findings of the study related to the type and nature of the ultimate attainment of MSA. The second section discusses the critical period and the effect of age of first exposure on the

ultimate attainment of MSA. The third section discusses the effect of L1 knowledge on the ultimate attainment of MSA. Section four adds implications for the syntactic analyses of the investigated phenomena based on the results of the experimental study. The final section summarises the chapter.

1   Is MSA Competence of L1 or L2 type?

There is a general consensus in the language acquisition literature that all acquirers of the same first language or dialect are expected to achieve the same end-state grammar (e.g., Chomsky, 1965; Bley-Vroman, 1989), whereas acquirers of the same second language are likely to differ from each other in their ultimate attainment (e.g., Bley-Vroman, 1989; Schachter, 1990; Schwartz and Sprouse, 1994; White, 2003) due to a variety of factors including effect of AoE and L1 influence.

It was assumed in this thesis that E-CLA and E-CGA participants had acquired MSA as a second L1 along with the colloquial dialect. This is because they started acquiring MSA naturally via immersion when they were 2 or 3 years old while their acquisition of the colloquial dialect was still in process. In contrast, participants of the other groups with relatively *late* exposure to MSA (at age 6) were assumed to have acquired MSA as an L2 at least due to their knowledge of the previously acquired dialect when they started acquiring MSA.

The main question of this thesis was whether the participants of this latter group, who represent the majority of Arabic speakers across the Arab world, succeeded in acquiring a native-like (uniform) MSA competence or not, despite their late exposure to MSA and

the variant colloquial dialects they speak. Contrary to the hypothesis of this thesis, the data presented in the previous chapter strongly point to an affirmative answer to this question at least in relation to their ultimate knowledge of the three MSA syntactic structures under study. According to this empirical data, L-CLA and L-CEA participants have managed to learn that resumption in MSA object relatives is optional not obligatory, in contrast to their colloquial dialects; L-CGA participants similarly have managed to learn that number agreement between human collective subjects and the following verb is optional in MSA though it is not in their colloquial variety; and all the participants of these 3 groups have learnt that VSO is the common order in MSA. Moreover, the performance of the late-exposed participants in the tasks of this study was not significantly different from the performance of those who acquired MSA earlier and in more naturalistic settings.

These empirical findings do not support Kaye's (1970) expectations of inconsistent performance in MSA due to what he called the ill-defined system of MSA. Also, they do not support the acquisition theory-based expectation that L2 acquirers will not achieve native-like competence due to lack of access to UG (if it is assumed that the critical period had declined or come to its end before age 6) and/or due to L1 influence (e.g., Bley-Vroman, 1989; Schachter, 1990; Schwartz and Sprouse, 1994; White, 2003).[44] If it was true that L2 acquirers had no or only partial access to UG due to end of the critical period (Bley-Vroman, 1989; Schachter, 1990), the current findings could perhaps be explained by the fact that even the *late* acquirers of MSA in this study started acquisition at age 6 which is still early and may be within the age boundaries of

---

[44] The issue of the critical period and AoE effect will be discussed separately in the following section. Also, the issue of L1 influence will be discussed later in section 3.

the critical period during which access to UG is still available and their innate ability is still active. If, however, full access to UG is available for L2 acquirers anyway regardless of AoE, as the FT/FA model by Schwartz and Sprouse (1994) assumes, then the participants in this study may have managed to restructure their transferred L1 grammar to converge on native knowledge of the target language before reaching the end-state stage of acquiring MSA.

The syntactic phenomena investigated were not expected to be salient in the L2 input due to the fact that they are optional in MSA and one of the two options is always available in the L1 grammar. Also, exposure to input was almost entirely limited to the written form of MSA for the *late* acquirers. Yet, if the participants are to be assumed to have managed to restructure their L1 grammars, they must have encountered clear and obvious cues in the MSA input during their acquisition process, which helped them to acquire the three MSA syntactic structures under study.

This, however, does not necessarily mean that speakers of MSA with *late* AoE have managed to restructure all the properties of their L1 grammar and acquire a full native-like competence of MSA. In fact, speakers of MSA may have failed to restructure other grammatical properties of their L1 that were not covered by this study, due to obscure or insufficient input. Therefore, further research is needed to investigate acquisition of other syntactic structures as well to confirm or disconfirm the present findings.

Moreover, there are further reasons why this finding that speakers of different dialects represent native-like competence in MSA cannot be considered to be conclusive based solely on the data presented here. As Hyltenstam & Abrahamsson (2003) suggested, it is essential to assess the learners' ultimate attainment in other sub-components of the target language as well, to come to a final conclusion about native-like attainment. The empirical data presented in this thesis assessed knowledge and use of grammatical competence within certain linguistic domains, whereas further empirical evidence must be sought in other linguistic domains (e.g., phonology) to heighten the validity of the current finding and reach a final conclusion about whether or not speakers of MSA have actually attained a comprehensive native-like proficiency.

Sorace (1993; 2003) argues that even L2 learners whose performance in the target language is characterised as native-like may have divergent or incomplete representations that differ in non-obvious ways from native speakers' grammatical representation. In addition, Hyltenstam & Abrahamsson (2003) claimed that close investigation of early L2 acquisition may reveal at the level of subtle details that the subjects are different from monolingual native speakers. Such claims imply that differences between observed native-like and actual native competence may be missed when investigating general attainment of specific structures in a single linguistic domain. This possibility becomes even more probable with cases like the case of acquiring MSA for which there are no monolingual native speakers to compare with.

Although, in general, resumption and plural agreement between human collective subjects and the following verb were treated as optional in the present experimental

study; the rates of obligatory resumption and plural agreement were still relatively high: up to 50% in some cases. This contrasts with descriptive grammars of MSA. According to these descriptive grammars, these two syntactic phenomena are optional not obligatory in MSA (see sections 2.1.3 & 2.3 of chapter 2). Moreover, such judgements cannot be attributed to L1 influence as they occurred in data of participants who speak dialects that are not different from MSA with regard to these syntactic phenomena (CGA in resumption, and CLA & CEA in agreement). Also, participants of groups with early as well as late exposure to MSA produced this type of data, which factors out the effect of AoE. Although performance of the study's participants was uniform across the variant groups, existence of such data suggests that the MSA competence developed by these participants may be different from the competence of monolingual native speakers of MSA if they exist. Another explanation of the existence for such data could be that MSA has undergone some change and descriptive grammars simply do not reflect the current version of it. The latter explanation though is less appealing, as analysis of judgments by test items did not reveal clear patterns of judgment regarding specific test items as unacceptable without resumption or agreement. Unfortunately, it is not possible to investigate these explanations further because (by hypothesis) monolingual speakers of MSA do not exist.

To summarise, although the results of the current experimental study point to the conclusion that speakers of different colloquial dialects with variant AoE represent a uniform L1 type knowledge of MSA, further research is required for a comprehensive assessment of their knowledge of it. Future research should involve investigation of knowledge of more grammatical phenomena in the same linguistic domain and in other domains as well. Whether the MSA knowledge represented by the participants in this

study is similar to the native knowledge of monolingual speakers of MSA cannot, unfortunately, be confirmed due to the lack of monolingual MSA speakers.

2   The Critical Period and the Effect of AoE

As reviewed in chapter 2 (section 1.1), studies investigating the issue of a critical period in second language acquisition came with different age ranges during which native-like attainment is possible or guaranteed. According to studies like Oyama (1979) and Patkowski (1980), if acquisition takes place at any time during the critical period which ends at puberty (or around age 12 for Oyama), then native-like attainment is possible. Other studies reported results that indicate several phases of the critical period during which AoE has different effects on ultimate attainment of the target language (e.g., Seliger et al., 1975; Shim, 1993; Meisel, 2008).   Native-like attainment is only guaranteed if acquisition started during or before the optimal phase of the critical period. The optimal phase is followed by a declining phase during which some L1-L2 attainment differences can be noted. This phase ends with a cut-off point that marks the end of the critical period. After the end of the critical period, ultimate attainment should be characterised by random variation and imperfection.

Meisel's (2008) results, for instance, marked age 3;7 as the end of the optimal phase and the start of the declining phase for morphosyntax.[45] Shim (1993) reported that AoE before age 5 revealed native-like attainment whereas AoE from age 6 to 11 showed age related effects. The results of Seliger et al. (1975) showed that participants with AoE before 9 performed in a native-like manner whereas the declining phase was represented

---

[45] As will be explained later in this section, Meisel assumes that there are several sensitive periods for different domains of grammar.

by the performance of participants with AoE of 10 to 15 years. The data of the current study showed that participants with AoE of 6 or before have managed to attain a uniform L1-type knowledge of at least the syntactic phenomena under investigation. If a critical period exists for second language acquisition, this suggests that the participants of the current study were at the optimal phase of the critical period when they started acquiring MSA. Since all the participants in this study started acquiring MSA at age 6 or earlier, and they performed in a uniform manner, the present data cannot point to the end of the optimal phase. Future research may include participants with later AoE and make correlation analyses between AoE and attainment to contribute with specifications on when the optimal phase of the critical period starts and comes to an end.

The results of the current study contrast with Meisel's (2008) and Shim's (1993) results in relation to the end of the optimal phase. This is because the participants with late AoE in this study started acquiring MSA later than age 3;7 and age 5 and, yet, they managed to perform in a uniform manner that is comparable to that of the participants with early AoE to MSA. Although it is not clear from analysing this data alone when the optimal phase comes to its end with the case of speakers of Arabic dialects acquiring MSA, this extended optimal phase, compared to what Meisel and Shim reported, can be explained by the typological closeness between the Arabic dialects and MSA. This is in analogy to Birdsong & Molis's (2001) explanation of the extended optimal phase effects on their participants' performance compared to the results of Johnson & Newport (1989). Birdsong & Molis (2001) argued that because there is a great deal of similarity between Spanish (L1) and English (L2), compared to Korean and Chinese (L1) and English (L2) in Johnson & Newport (1989), even participants who arrived relatively late in an English speaking community could successfully master L2

properties. Similarly, in a partial replication of Johnson & Newport (1989) but with Spanish and Chinese learners of English, Bialystok & Miller (1999) found that the typological similarities between Spanish and English could be a reason for the slightly delayed age effects noted for the Spanish learners, compared to the Chinese learners. This applies even more so to the situation of the participants of the current study. Although there are differences between the colloquial dialects of Arabic and MSA, there are, at the same time, a large number of similarities between them across all the linguistic domains, compared to L1-L2 pairings in Meisel's (German-French) and in Shim's (Korean-English) studies. This may have helped the participants of the current study to have extended the period of optimal sensitivity to MSA input that shows specifications on the syntactic phenomena under investigation helping them to aquire these phenomena successfully.

Meisel (2008; 2009; 2011), following Seliger (1978) and Schachter (1996), argues that there are several sensitive periods within the critical period. According to him, the grammatical domains should not be expected to be affected during a single age period. Because past research has pointed out that different areas of grammar do not develop simultaneously, Meisel argues that "the critical period is better understood as a *cluster of sensitive phases* during which the LAD [the language acquisition device] is optimally prepared to integrate new information into developing grammars" (2011:205). In fact, he argues that even acquisition of properties within a single grammatical domain might not be affected during a single age period, as development of properties of a single grammatical domain may occur at different age periods. Moreover, Mesiel (2011) assumes that each sensitive phase during the critical period should be viewed as starting with a relatively short onset followed by an optimal period followed by a declining

period toward a gradual offset. If this assumption is correct, then it is safe to claim that at least the sensitive period(s) for acquiring the three syntactic phenomena examined in the current study had not faded out by the age of 6. In fact, it is safe to claim that the optimal period of this sensitive phase (or phases) had not passed by this age. This is evident by the fact that participants with AoE at age 6 or before managed to acquire these syntactic phenomena and perform in the tasks of the study in a uniform manner. By adopting this assumption about critical periods, the possibility of finding age related effects on the acquisition of other syntactic phenomena or the acquisition of other phenomena in different grammatical domains is not excluded; it is possible that the sensitive phase(s) for acquiring those phenomena could be over by the age of 6. Thus, further research is needed to confirm or disconfirm the major finding of this study in relation to the ultimate attainment of acquiring MSA starting at age 6.

The present data did not yield support for the 'younger is better' effect reported by Johnson & Newport (1989) and Johnson (1992) during the critical period and by Birdsong & Molis (2001), Stevens (1999), and Bialystok & Hakuta (1999) across the life span. This is because there was no significant difference between participants who were first exposed to MSA at age 2 or 3 and participants who started being exposed to MSA at age 6. A yet wider range of starting age of acquiring MSA is needed to substantiate this finding. Johnson & Newport (1989), for example, had participants with AoE ranging from age 3 to age 39 to come up with their conclusion. It may be necessary to have a comparable range of AoE to MSA to confirm or disconfirm their finding. Having such a wide range of AoE may help to contribute with specifications on the age boundaries of the optimal phase and the declining phase for acquiring the present syntactic phenomena, and specifications on the age boundaries of the critical

period as a whole. Also, data with such a wide range of AoE may yield better findings in relation to the issue of 'younger is better' either during the critical period or across the life span.

Finally, the results of the present experimental study are limited to the effect of age on acquisition of grammatical knowledge. These results do not support the hypothesis that learners with earlier exposure to MSA have an advantage over those who start acquiring MSA at age 6 in terms of attainment of grammatical knowledge. However, exposure to MSA before age 6, as Jenkins (2001) and Alomari (2009) reported, is advantageous to the learners in educational means such as acquiring school skills of reading and composition in MSA, or being prepared to start acquiring literacy in MSA at the age of 6.

To summarise, starting to acquire MSA at age 6 does not affect the ultimate attainment of the three linguistic phenomena investigated. Learners who had started learning MSA at this age did not seem to miss the optimal phase of the sensitive period(s) for acquiring the MSA syntactic structures under study. The results, though, do not show specifications on the age boundaries of these sensitive periods or the critical period as a whole. Further research with a wider range of AoE is required to identify these age boundaries.

3   The Effect of the Colloquial Dialects (the L1s)

One common position that researchers take about the L2 initial state is that L2 learners transfer their L1 grammar either fully or partially when they start acquiring the L2 (see,

Section 1.2 in Chapter 2). These researchers though have different claims in relation to how L2 learners restructure their L1 knowledge to converge on the target language grammar and whether they will be completely successful in this task or not when reaching the end-state of L2 acquisition. Bley-Vroman (1989) and Schachter (1990), for example, explained the failure to acquire native-like competence in L2 acquisition by lack of full and direct access to UG after the closure of the critical period. Schwartz & Sprouse (1994; 1996), on the other hand, claimed that full access to UG is available to L2 learners regardless of any critical period, but full convergence on the grammar of the target language is not guaranteed because the L2 input needed to force restructuring of L1 knowledge may not be available or may be complex, rare, or not clear.

The data presented in the previous chapter showed that the variant colloquial varieties of Arabic which the participants of this study speak as their L1 had no significant effect on the results. Keeping in mind the conclusion that participants who started acquiring MSA at age 6 did not miss the optimal phase of the sensitive period(s) for acquiring the investigated syntactic phenomena, it might be appropriate to claim that the participants of this study all had full access to UG when acquiring MSA, regardless of which model of L2 acquisition is considered (see section 1.2 of chapter 2, for the models). In addition, the fact that participants performed in a uniform manner despite the variant colloquial dialects they speak suggests that they were successful in restructuring their variant L1 knowledge to converge on a uniform MSA grammar. This also leads to the conclusion that despite the fact that exposure to MSA was almost limited to the written form for the late acquirers, the required MSA input for restructuring L1 knowledge of the syntactic phenomena under study must have been adequately available to these participants together with their full access to UG.

Montrul (2010) claimed that L1 knowledge of structures at both internal and external linguistic domain interfaces is subject to transfer more than L1 knowledge of purely syntactic structures. This claim places further emphasis on the assumption here that the participants with *late* AoE in the present study had at least transferred their L1 knowledge of the syntactic phenomena under study when they started acquiring MSA, but they also managed to restructure this knowledge to a uniform representation of MSA. This is because acquiring the investigated syntactic phenomena in the present study can be claimed to involve knowledge from more than one linguistic domain as well. Word order, for example, involves knowledge of syntax and pragmatics at the same time. The choice of agreement or no agreement between subjects (or pre-verbal NPs) and the following verb involves integration of knowledge of both syntax and semantics. Also, resumption in object relatives can be claimed to be at the syntax-semantics interface if resumptive pronouns are viewed as adding more specificity to the reference of the relativised element (see Alresaini, 2007).

Oh (2010) concluded that it is possible to recover from negative transfer of L1 knowledge at the syntax-semantics interface. The results of the current study support Oh's (2010) conclusion. As mentioned earlier in this chapter and the previous chapter, L-CLA and L-CEA participants were successful in acquiring MSA properties of resumption in object relative clauses. If it is assumed that these learners had transferred their dialects' grammars, then they seem to have managed to restructure from these grammars that allow only base-generation strategy to form object relatives with resumptive pronouns to a grammar that allows both this base-generation strategy and a movement strategy to form this type of relatives with a gap filled with a trace instead of a resumptive pronoun depending on whether more specifity is needed.

Also, L-CGA participants were successful in acquiring MSA properties of agreement between human collective subjects (or preverbal NPs) and the following verb. These participants succeeded in acquiring these properties despite the difference between their L1 knowledge of these structures and the knowledge they acquired for MSA. These participants seem to have managed to restructure from a grammar that treats preverbal human collective nouns as typical human plurals to a grammar that recognises the different possible interpretations of the collective nouns (generic/non-generic) in the way number agreement is expressed (singular/plural; repectively).

Moreover, the results of this study suggest that it is also possible to recover from transfer at the syntax-pragmatic interface. This is evident by the results of the participants with late AoE in the Conversation Role-Play task which aimed to investigate their knowledge of word order in MSA. These participants succeeded in learning that VSO order is the common order in MSA, and the L-CEA participants learnt successfully that context and intended pragmatic functions affect the choice of word order in MSA.

To summarise, at least the participants with *late* AoE are assumed to have transferred their L1 knowledge of the investigated phenomena when they started acquiring MSA. However, they managed to recover from negative transfer of knowledge related to these phenomena due to their full access to UG when they started acquiring MSA and to the availability of the required MSA input that enforces restructuring to the target language grammar. Thus, the effect of L1 influence was not significant on the results of the current study.

4 Syntactic Implications

The main focus of this thesis is to investigate and discuss the ultimate attainment of MSA in light of language acquisition theories. However, the results of the experimental study presented in the previous chapter can be interpreted to imply support to or oppose certain analyses of the MSA syntactic phenomena reviewed in section 2 of chapter 2.

The results of the experimental study concerning the word order task, for example, showed that VSO order was used more frequently than SVO order in both conversational (68% vs. 32%) and narrative (95.2% vs. 4.8%) contexts by the participants when speaking MSA (see, Table 38 in chapter 5). This adds support to Bakir's (1980) claim that VSO is the most common word order in MSA. Bakir (1980) explained that the high frequency of using VSO order is due to the fact that it is the neutral order in MSA, that is, least marked syntactically and pragmatically; this order does not serve pragmatic functions like topicalisation or focus. This observation is supported by the fact that the participants in this study used VSO more frequently than SVO regardless of the change of contexts which usually entails the requirement of using different pragmatic functions.

SVO order, on the other hand, seems to be used in MSA to serve pragmatic functions like topicalisation and focus. This is evident from the results of the current experimental study which showed that the rate of using SVO order increases in conversational context where subject-oriented discourse is expected. In contrast, this order is hardly used in narrative context where event-oriented discourse is expected. These results

contrast with, for example, the results of analysing colloquial Egyptian Arabic presented in section 1.1 of chapter 3. Speakers of colloquial Egyptian Arabic used SVO order in more than 90% of the sentences they produced, regardless of the context of discourse (see Table 1 -Table 3 in chapter 3). Based on these results, SVO order in colloquial Egyptian Arabic does not seem to be used to serve specific pragmatic functions as it does in MSA, but rather it seems that this order is the neutral least marked order in colloquial Egyptian dialect and, hence, the high rate of using it in Egyptian data.

Considering SVO order as used in MSA for pragmatic purposes may add support to Plunkett's (1993) and Ouhalla's (1997) analyses of preverbal NPs in MSA. These researchers claimed that preverbal NPs in MSA should be considered as left dislocated topics; not subjects. Using SVO order less than VSO order in general and the higher rate of using SVO in conversational contexts can be considered as empirical evidence supporting these analyses. Following the same line of reasoning, preverbal NPs in colloquial Egyptian dialect can be considered as subjects, not topics. This is because their occurrence in this position does not seem to depend on the need for employing pragmatic functions in the discourse as this order is used most frequently in both types of examined discourse; narrative and conversational. The diagrams in (77) and (78) demonstrate the structure of an SVO sentence in MSA and in colloquial Egyptian Arabic, repectively:

77. SVO sentence in MSA: 78. SVO sentence in colloquial Egyptian Arabic:
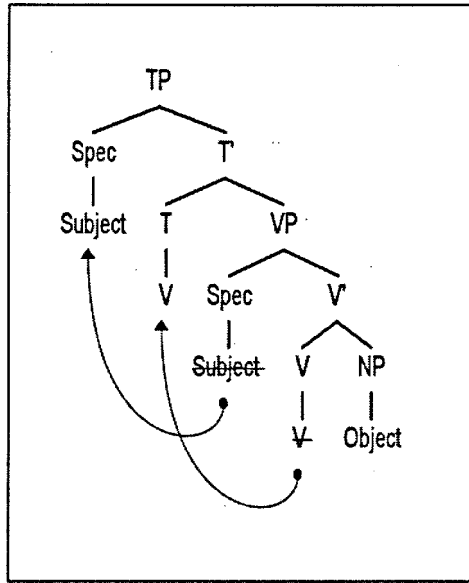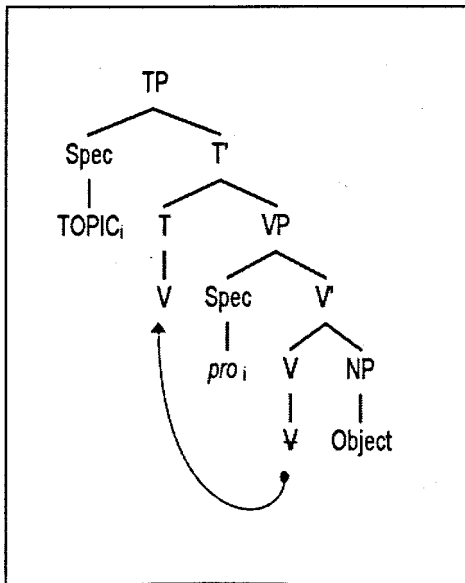
Another syntactic implication is related to the results of resumption in object relative clauses. These results showed that resumption was treated as optional in 55.5% of the data whereas it was treated as obligatory in 44.5% of the data (see Table 23 in chapter 5). When the data were analysed by test item, the results showed that sentences without resumptive pronouns were accepted more if the relativiser is *man* or *maa*, neither of which has gender or number specifications; otherwise, using resumptive pronouns was preferred by the participants in sentences with other relativisers that have gender and number specifications (see Table 27 in chapter 5).

These results may imply that optionality of using resumptive pronouns in MSA object relative clauses is not free as it appears to be. In accordance with Alresaini (2007), at least in the varieties where they are optional, resumptive pronouns seem to add more specificity and clarity to the reference of the relativised element. Thus, when a specified

relativiser like *?allaðii* for masculine singular is used, it seems that using a resumptive pronoun is preferred in this specified context. A less specific reading of the reference of the relativised element is not excluded by using this type of relativiser: it can be expressed by not using a resumptive pronoun when the reference is clear from the context. On the other hand, when a generic relativiser like *man* is used, this encourages a less-specific reading of the reference of the relativised element and, hence, not using a resumptive pronoun becomes more acceptable. This again does not exclude using generic relativisers when a more specified reference of the relativised element is required when ambiguity is possible; this can be expressed by inserting a resumptive pronoun.

If this analysis is correct, this may also explain treatment of resumption as obligatory in part of the examined data. By analysing the data by test item, treatment of resumption as obligatory was random as there was no clear pattern of requiring resumptive pronouns with certain test items. This may be explained by the fact that the test items of the AJT were shown to participants out of context. This may have caused the participants to be unsure whether they ought to use resumptive pronouns to add more specificity or not. Thus, treating resumption as obligatory in a random manner may have resulted in different participants' decisions about whether more specificity is needed or not. To confirm this analysis, further research should aim to elicit judgements on object relative clauses within context where the requirement of more specified reference is controlled for.

Another syntactic implication is related to the results of number agreement between human collective nouns and the following verb. As explained in section 2.1 in chapter 2, lack of plural agreement with inanimate plurals in preverbal position was suggested to be due to treating them as a singular group rather than a group of separate members. Therefore, the verb occurring after these inanimate plurals in SVO order always takes the singular form. The results here of agreement between preverbal human collective nouns and the verb confirm that these collective nouns can be treated as a group of individuals like animate plurals, hence, full agreement in number, or as a singular generic group like inanimate plurals, hence, the verb takes the singular form.

The participants from all groups in the present study treated plural agreement between human collective nouns and the following verb as obligatory in 24% of the collected data (see Table 31 in chapter 5). This cannot be interpreted as an effect of L1 because there was no significant difference between the participants who speak different colloquial dialects in relation to this phenomenon. Treating this type of agreement as obligatory in part of the data may indicate that optionality of agreement is not free as it appears to be. The test items used different human collective nouns like _l-naas_ 'the-people', _al-qawm_ 'the-folk', etc (see Appendix 5 for the test items). Yet, by examining judgments by test item, there was no clear pattern of judging plural agreement as obligatory with certain collective nouns; judging plural agreement as obligatory was, in fact, random (see Table 35 in chapter 5). Random judgments of plural agreement as obligatory may be due to the fact that the test items were shown in the AJT out of context. As explained in section 1.2 of chapter 3, Brustad (1991) claimed that optionality of plural agreement with collective nouns is possible only when these nouns are referring to a generic homogenous abstract group. Thus, context may be necessary

for the participants to decide whether to consider the collective noun referring to a heterogeneous specific group, hence, plural agreement is required, or homogeneous abstract group for which no agreement is possible. Since context was lacking, participants seem to have had varied judgments due to random guessing in relation to the reference of these collective nouns. Further research should aim to elicit judgments on these test items being shown within context that clarifies the reference of the collective nouns.

5   Summary

This chapter discussed the findings of the experimental study in relation to the research questions and related literature. The first section discussed whether participants with exposure to MSA at age 6 succeeded in acquiring a native-like competence in MSA. It was concluded in this section that these participants have managed to attain native-like end-state knowledge of at least the three syntactic phenomena investigated. However, further investigation of other phenomena in the syntactic domain and other linguistic domains is required for a comprehensive assessment of the ultimate attainment of MSA by learners who start acquisition at age 6. Section 2 discussed the critical period and effect of the AoE on the ultimate attainment of the target language. It was argued in this section that starting to acquire MSA at age 6 does not affect negatively the ultimate attainment of at least the three linguistic phenomena investigated. It seems that participants who start at this age do not miss the optimal phase of the relevant sensitive period(s). Further research with a a wide range of AoEs is needed to determine specifications of the age boundaries of the sensitive period(s) or the critical period as a whole. Section 3 discussed the effect of the colloquial dialects on the end-state grammar of MSA. The participants seem to have managed to restructure their transferred L1

knowledge due to full access to UG and the adequate availability of the relevant MSA

input. Thus, L1 knowledge was found not to have a significant effect on the results.

Section 4 added implications based on the results of the experimental study in relation

to the syntactic analyses of the three linguistic phenomena investigated.

Chapter 7

Conclusion

The previous chapter discussed the findings of the experimental study in relation to the research questions and related literature. The present chapter concludes this thesis in three sections. The first section summarises the main findings of this research. The second section lays out the limitations of the experimental study. The final section presents recommendations for relevant future research.

1 Summary of Findings

This thesis was conducted to present an empirical investigation of the ultimate attainment of MSA acquired by speakers of different colloquial varieties of Arabic. The main query was whether typical learners of MSA who had started their acquisition at age 6 in primary schools and who spoke different colloquial dialects natively had been able to acquire a native-like competence of MSA or not. Other interrelated queries were about the effect of the relatively late age of exposure (at age 6) and the effect of the previously acquired colloquial dialect on the ultimate attainment of MSA. The findings of this research are presented in the following paragraphs.

First, the colloquial Arabic corpus study revealed grammatical differences between the three colloquial varieties of Arabic (the L1s) under study, and these differences were taken as the syntactic variables in the experimental study of MSA. CGA was found

different from CLA and CEA in terms of resumption in object relatives and in terms of preverbal collective noun-verb agreement in number. In CGA, resumption was optional, whereas it was obligatory in CLA and CEA, and agreement was obligatory in CGA, whereas it was optional in CLA and CEA. Examining word order in this study also revealed differences between the colloquial varieties. CEA speakers used SVO order persistently regardless of the change of the discourse type. In contrast, word order pattern in CLA and CGA was found sensitive to discourse type. The speakers of these two varieties used SVO more than VSO in conversational discourse, whereas in narrative discourse, VSO became the common order. The MSA corpus study did not disclose reliable findings due to limitations relevant to the type, amount, and validity of the MSA corpus data. Therefore, an experimental study was designed and conducted.

Second, the results of the experimental study point strongly to the conclusion that speakers of Levantine, Egyptian, and Gulf colloquial dialects who had started their acquisition of MSA at age 6 in primary schools were successful in attaining a native-like competence of at least the three syntactic phenomena: word order, resumption in object relatives, and number agreement between collective nouns and the following verb. This conclusion was reached based on the following:

(a) The performance of these participants was uniform in both experimental tasks of this study despite the grammatical differences between the colloquial dialects they had previously acquired in relation to the syntactic phenomena investigated.

(b) The performance of these participants was not found significantly different from the performance of those who had started acquiring MSA earlier at age 2 or 3 and who could be considered native speakers of MSA.

232

(c) Uniform performance was attested in results of all the three MSA syntactic phenomena elicited by employing two different tasks which assessed the participants' MSA competence (AJT) and their use of it (CR-P).

Third, statistical analysis of the results showed that difference in age of first exposure had no significant effect; starting to acquire at least the three investigated MSA syntactic phenomena at age 6 did not have a negative effect on the ultimate attainment. This was interpreted as showing that, despite starting to acquire MSA at age 6, these learners did not miss the sensitive period(s) of the critical period, or even the optimal phase(s) of these periods, for acquiring a native-like knowledge of the investigated linguistic phenomena. However, the results do not show specifications on the age boundaries of any relevant sensitive periods or the age boundaries of the critical period as a whole.

Fourth, the previously acquired knowledge of the variant colloquial dialects of Arabic did not have a significant effect on the ultimate attainment of at least the three linguistic phenomena that were investigated. If it is assumed that L2 learners start with the grammar of their L1, the participants who had started acquiring MSA at age 6 have managed to successfully restructure this knowledge to converge on the MSA grammar. That this was possible can be assumed to be due to their full access to UG and the availability of adequate exposure to MSA input in relation to the investigated syntactic phenomena. Therefore, it was concluded that recovery from negative transfer of L1 knowledge at the syntax-semantics and the syntax-pragmatics interfaces is possible for child L2 learners.

2 Limitations of the Experimental Study

Examining the ultimate attainment of MSA by conducting this carefully designed experimental study, the limitations of the MSA corpus study (e.g., small amount of data, validity of data, etc.) explained in section 2.3 of chapter 3 were successfully overcome. However, and in spite of the above-mentioned findings and implications, the experimental study could not help suffering some limitations related to data collection and AJT design.

First, it was not possible to collect data from speakers of colloquial Egyptian Arabic who had early exposure to MSA via immersion. This limitation could not be overcome because there are no preschools in Egypt that apply the MSA immersion program. It might not be impossible to find exceptional families who intentionally expose their children to MSA before the age of 6, but if these families really exist, it would be very difficult to find and make contact with them. It would be better for comparisons to have two sets of participants who speak the same colloquial dialect; one with early AoE and the other with late AoE. This was done with the speakers of colloquial Levantine Arabic (E-CLA vs. L-CLA) and with the speakers of colloquial Gulf Arabic (E-CGA vs. L-CGA); but could not be done with the speakers of colloquial Egyptian Arabic (no E-CEA group).

Second, the MSA data collected for this study from speakers of colloquial Egyptian Arabic who had been first exposed to MSA at age 6 in primary schools in Egypt was collected in Saudi Arabia; not Egypt. Although these participants were first exposed to

and learnt MSA in Egypt as they had their primary and secondary education there, their temporary stay in Saudi Arabia may have affected their MSA knowledge, especially knowledge of word order and resumption. During their stay in Saudi Arabia, these participants were exposed to MSA in Saudi Arabia at school and to colloquial Saudi Arabic in everyday life. This may have alleviated the influence of colloquial Egyptian Arabic on their performance in the tasks of the study.[46] Further research may be conducted later that collects similar data from participants in Egypt to verify the results of the current study.

Third, the test items in the AJT were shown to participants out of context. As discussed in section 4 of the previous chapter, context may be essential for participants to decide whether a resumptive pronoun is needed to add more specifity to the reference of the relativised element. Showing the test items in context also would help the participants to know whether the reference of the collective noun is generic or not, so that they make accurate judgments in terms of preverbal collective noun-verb agreement in number.

3 Recommendations for Further Research

The above findings and limitations of the current study have led to identify some research gaps that deserve further attention and examination in the future of language acquisition research and research on acquisition of MSA by speakers of colloquial Arabic dialects.

---

[46] As explained in section 1 of chapter 4, it was planned to collect these data in Egypt in January 2011, but the data collection trip was hindered first by delay of sponsoring funds then the Egyptian Revolution started in Egypt which made it impossible for the researcher to go to Egypt at that time.

First, in order to have a comprehensive and conclusive answer to the question of whether or not Arabs achieve native-like ultimate attainment in MSA, further research is required. Empirical data of investigating end-state knowledge of one or few linguistic phenomena which show differences between learners of variant AoE and/or CVs would be sufficient to provide a negative answer to this question. In contrast, the data of the current study showed that the participants of all variant groups performed in a uniform manner in relation to the three syntactic phenomena that were investigated. Although the current data point strongly to the conclusion that speakers of different colloquial dialects who start acquiring MSA at age 6 or younger achieve an L1-type attainment, this data on its own cannot provide a comprehensive and conclusive positive answer to the question mentioned above. As explained in the previous chapter, speakers who learn MSA as an L2 due to starting acquisition relatively late at age 6 may not have encountered adequate data in MSA input for acquiring other linguistic phenomena that is required for restructuring their L1 knowledge to converge on MSA grammar. Also, it is possible that sensitive periods for acquiring other linguistic phenomena have already passed or faded out by the age of 6, if the claim of several separate sensitive periods is assumed. Therefore, and as Hyltenstam & Abrahamsson (2003) suggested, it is essential to assess the learners' ultimate attainment of other linguistic phenomena in the same sub-component of MSA grammar and in other domains as well to come to a final conclusion about native-like attainment. Thus, only with a large body of research that investigates ultimate attainment of an adequate number of linguistic phenomena in different linguistic domains is it possible to provide an affirmative conclusion that speakers of different colloquial Arabic dialects who start MSA acquisition at age 6 achieve native-like ultimate attainment.

Second, to have better understanding of the elusive differences between the ultimate attainments of native bilinguals versus child L2 learners of MSA, the linguistic phenomena to be investigated need to be decided upon carefully. Hyltenstam & Abrahamsson (2003) noted that a study aiming to detect such differences needs to include investigation of not only the core linguistic phenomena of the target language, but also other peripheral language-specific phenomena such as idiomatic expressions and unusual structures that are usually expected to be difficult for advanced L2 learners to master. The current study investigated three syntactic phenomena with optionality that were expected not to be salient enough in MSA input for the participants of the study due to availability of one of the options in the colloquial dialects; the participants who acquired MSA as an L2 were expected to keep using their L1 option when speaking MSA as nothing in MSA input was expected to contradict using it. Yet, these participants seem to have encountered positive evidence in MSA input that helped them to acquire the MSA representation of these phenomena. Therefore, a more challenging task that targets investigating more difficult or subtle MSA phenomena needs to be designed to tap into elusive discrepancies between native ultimate attainment and child L2 learners' ultimate attainment of MSA.

Third, as there is a scarcity of works on colloquial Arabic varieties, it was essential to conduct the colloquial Arabic corpus study to find grammatical variables for the experimental study. This is because the grammatical differences between colloquial varieties of Arabic had received little attention in the literature (Cuvalay-Haak, 1997:19). In general, works comparing CA varieties usually find more grammatical similarities than differences between these varieties (see, e.g., Brustad, 1991; 2000). Therefore, in order to conduct more profound studies on the acquisition of MSA by

native speakers of the colloquial varieties, there seems to be a great need for more background corpus research on different varieties and styles of Arabic. Such research can provide a finer basis on which to rely for the choice of best syntactic phenomena to investigate in the future.

Fourth, the findings of the present experimental study emphasise the necessity for further research about the distinct types of bilingual knowledge attained by simultaneous bilinguals and early child L2 learners in comparison to monolingual speakers' knowledge of their primary language. The fact that the participants in this study with variant AoE and CVs produced some data that did not reflect the described grammar of MSA might suggest that their ultimate attainment in MSA was not identical to native attainment of monolingual speakers of MSA (if they existed), despite the participants' uniform performance in the tasks of the present study. Although it is not possible to examine the ultimate attainment of monolingual speakers of MSA as they do not exist, further research efforts might be necessary to detail the nature of knowledge attained by these apparently distinct types of acquisition in different languages.[47]

Fifth, the current research investigated the ultimate attainment of MSA by adolescent speakers of colloquial Arabic dialects who had early versus late AoE and found no significant difference between these two populations. Future long-term research may investigate younger participants of these two populations, who are still acquiring MSA,

---

[47] Sorace (2012) mentioned in a talk that she and collegues (Unsworth, Cornips, Hulk, Argyri, and Tsimpli) have already completed a project called 'The Early Child Bilingualism' with the aim to investigate knowledge attained by different types of acquisition such as acquisition of two languages from birth (simultaneous bilingualism), compared to acquiring the second language from age 1 to 4 (early successive bilingualism), and from age 4 to 8 (child L2 acquisition). Their work is not published yet.

to assess whether they develop their knowledge in the same way as well during the process of acquisition. Such a comparasion will inform whether or not late starters of acquiring MSA at age 6, who represent the majority of Arabic speakers, actually develop and reach an L1-type end-state knowledge of MSA despite the relatively late age of onset and the previously acquired L1 knowledge of the colloquial dialect they speak.

Sixth, future research on MSA acquisition that recruits participants with a wide range of AoE may help to contribute with specifications on the age boundaries of the different phases of the sensitive period for acquiring the investigated linguistic phenomena or the different phases of the critical period as a whole. The data of the present experimental study suggested that learners of MSA who started at age 6 or younger did not miss the optimal phase(s) for acquiring the investigated syntactic phenomena. However, the current research did not aim to identify when this phase starts to decline or comes to its end. In addition to these participants, future research may include late literates who started acquiring MSA at school at different times after the age of 6. A correlation analysis between AoE and ultimate attainment of MSA by these participants, for example, may yield findings on the age boundaries of the different phases of the sensitive periods for acquiring the investigated phenomena. The need for such research is suggested by the different findings of this study compared to, for instance, the findings of Meisel (2008) and Shim (1993). As Birdsong & Molis (2001) suggested, a large number of similarities between the first and the target language may affect the length of the optimal phase(s) for acquiring the different target language properties. Therefore, investigating the acquisition of MSA by speakers of Arabic dialects may

come up with important findings on the critical period in relation to this kind of acquisition of a relatively similar target language.

Seventh, examining knowledge of resumption in MSA object relatives and plural agreement between human collective nouns and the following verb may yield better findings if examined within context. The present experimental study examined these two syntactic phenomena out of context in an AJT. However, because context may play a role for the participants to decide whether resumption or agreement is needed or not, future research on acquiring these two MSA syntactic structures is advised to include context to control for whether a specific reference of the collective nouns or an emphasised reference of the relativised item is needed or not.

Eighth, re-examining the ultimate knowledge of MSA attained by speakers of colloquial Egyptian Arabic in Egypt may show different results from the results presented here. As explained in the limitations section of this chapter, the data by these speakers was necessarily collected for the present study in Saudi Arabia; not in Egypt. Therefore, to obtain more precise results in relation to the effect of the colloquial Egyptian Arabic on the ultimate attainment of MSA by speakers of this dialect, such data should be collected in Egypt.

Finally, the data of the present study was quantitative in nature. In addition to this data, obtaining qualitative data is recommended to have better understanding of the age effects on attainment of MSA properties by speakers of the colloquial dialects.

McDonald (2000) and Hyltenstam & Abrahamsson (2003), for example, recommended that age-related investigations should be based on a balance of quantitative and qualitative data. DeKeyser & Larsen-Hall (2005) pointed out that qualitative analysis would be better to have more accurate understanding of age effects on attainment of some L2 properties such as processing difficulties and semantic complexity. Therefore, qualitative analysis of MSA data produced by speakers of the colloquial dialects may lead to the exposure of subtle differences between the ultimate attainments of bilinguals versus child L2 learners of MSA.

Appendices

Appendix 1: List of Verbs for Searching Colloquial Arabic Corpora

NOTE:

- As described in chapter 3 (seco n 1.1.1), wildcard characters (*) indicate possible prefixes (if put to the right of the Arabic script) or suffixes (if put to the left). This is to instruct 'COMBO' find all the possible forms of the verb.

Colloquial Gulf Arabic Verb List

No.	Verb in Arabic Script	Transliteration	English Translation
1	*أكل*	?akal	eat
2	*شرب*	šarab	drink
3	راح*	raaH	go
4	جا*	jaa	come
5	*حاول*	Haawal	try
6	*تكلم*	tikallam	talk
7	صلى	Sallaa	pray
8	*تفرج*	tifarraj	watch
9	مشى	mišaa	walk
10	*توقع*	w aqqa9	expect
11	*جرب*	jarrab	try
12	*جلس*	jalas	sit
13	*حس*	hass	feel
14	*حط*	HaT	put
15	*أخذ*	?axað	take
16	*درس*	daras	study
17	*دفع*	difa9	pay
18	اشترى	?ištaraa	buy
19	*دق*	daq	call (by phone) or knock
20	اتصل*	?ittiSal	call (by phone)
21	*دور*	dawwar	look for
22	*ذاكر*	ðaakar	study
23	*رد*	rad	reply
24	*رقد*	riqad	sleep
25	*نام*	naam	sleep
26	*أرسل*	?arsal	send
27	*أزعج*	?az9aj	bother
28	*زعل*	za9al	be upset
29	*زار*	zaar	visit
30	*سأل*	sa?al	ask
31	*سمع*	sima9	hear
32	سوى	sawwaa	make

242

33	اشتاق*	?ištaaq	miss
34	اشترك*	?ištarak	subscribe
35	*تعلم*	ta9allam	learn
36	باع*	baa9	sell
37	تريا	tarayaa	wait
38	انتظر*	?intaZar	wait
39	*رقص*	riqaS	dance
40	*تمنى	timannaa	wish
41	*كلم*	kallam	talk
42	*جاب*	jaab	bring
43	*حب*	Hab	love
44	حكى	Hakaa	narrate
45	*تذكر*	tiðakkar	remembered
46	*رمس*	rimas	call (somebody's name)
47	سار*	saar	go
48	اشتغل*	?ištaGal	work
49	*شكر*	šikar	thank
50	شاف*	šaaf	see
51	*صبر*	Sibar	be patient
52	*ضحك*	DaHak	laugh
53	*ضرب*	Darab	hit
54	*ضيع*	Dayya9	loose
55	*طلب*	Talab	request
56	*طلع*	Tala9	leave (a place)
57	عطى	9aTaa	give
58	*عرف*	9araf	know
59	*غسل*	Gasal	wash
60	*فكر*	fakkar	think
61	*قصد*	qiSad	mean
62	*قعد*	qa9ad	sit
63	قال*	qaal	say
64	*كتب*	kitab	write
65	*كذب*	kiðab	tell a lie
66	*كمل*	kammal	complete
67	*لاحظ*	laaHaZ	notice
68	*لبس*	libas	wear
69	*مسك*	misak	hold (by hand)
70	نادى	naadaa	call (somebody's name)
71	*نظف*	naZZaf	clean
72	*وضح*	waDDaH	clarify
73	*بدا*	bidaa	start
74	*حضر*	HaDar	attend
75	*خلص*	xallaS	finish
76	*درش*	dardaš	chat
77	استعمل*	?ista9mal	use
78	استفاد*	?istafaad	benefit
79	*سخن*	saxxan	warm (food)
80	*سهر*	sahar	stay up (at night)
81	*حرق*	Haraq	burn
82	تغدى	taGadda	lunch

Appendix 1: List of Verbs for Searching Colloquial Arabic Corpora

83	تعشى	ta9aššaa	dine
84	*تدلع*	d alla9	dally
85	*بالغ*	baalaG	exaggerate
86	*زهق*	zahaq	be bored
87	*ساعد*	saa9ad	help
88	*تعب*	ta9ab	be tired
89	*تحمل*	taHammal	endure
90	*دخن*	daxxan	smoke
91	*رجع*	rija9	get back
92	*بدل*	baddal	exchange
93	*غير*	Gayyar	change
94	*فتح*	fitaH	open
95	*قفل*	qifal	lock
96	*قطع*	qiTa9	cut
97	*جرح*	jaraH	wound
98	*وقف*	wiqaf	stop
99	*هون*	hawwan	change (mind)
100	*تزوج*	tizawwaj	marry
101	*خطب*	xaTab	propose (marriage)
102	*نزل*	nizal	get down
103	*رفع*	rifa9	raise
104	رقى	riqaa	get up (e.g., the stairs)
105	*وصل*	wiSal	arrive
106	*أسرع*	?asra9	speed up
107	*لعب*	la9ab	play
108	رمى	rimaa	throw
109	*جدع*	jida9	throw
110	*عاقب*	9aaqab	punish
111	*عاتب*	9aatab	admonish
112	*عالج*	9aalaj	treat
113	*حاسب*	Haasab	Pay (money)
114	*رتب*	rattab	put in order
115	*جمع*	jima9	collect
116	*راجع*	raaja9	revise
117	*كشف*	kišaf	uncover
118	*فحص*	faHaS	examine
119	*شبع*	šiba9	be full (food)
120	*مل*	mall	be bored
121	*طفش*	Tifaš	be bored
122	*هرب*	harab	flee
123	*سافر*	saafar	travel
124	*طبخ*	Tibax	cook
125	*صام*	Saam	fast (from food)
126	تمشى	timaššaa	stroll
127	بكى	bikaa	cry
128	*ناقش*	naaqaš	negotiate
129	*خاف*	xaaf	fear
130	استلم	?istalam	receive
131	*سكت*	sikat	stop talking
132	*ذاق*	ðaaq	taste

244

Appendix 1: List of Verbs for Searching Colloquial Arabic Corpora

133	*ترك*	tarak	leave (something)
134	لغى	laGaa	cancel
135	*فهم*	faham	understand
136	*نجح*	nijaH	succeed
137	*رسب*	risab	fail
138	*أقنع*	?aqna9	convince
139	*رفض*	rifaD	refuse
140	استأجر *	?istaajar	rent

Colloquial Levantine Arabic Verb List

No.	Verb in Arabic Script	Transliteration	English Translation
1	*أكل*	ʔakal	eat
2	*شرب*	širib	drink
3	*راح*	raaH	go
4	*إجا*	ʔijaa	come
5	*حاول*	Haawil	try
6	*حكي*	Hikii	talk
7	صلى	Salla	pray
8	*تفرج*	ʔitfarraj	watch
9	*مشي*	mišii	walk
10	*توقع*	ʔitwaʔa9	expect
11	*جرب*	jarrab	try
12	*قعد*	ʔi9id	sit
13	*حس*	Hass	feel
14	*حط*	HaT	put
15	*أخد*	ʔaxad	take
16	*درس*	daras	study
17	*دفع*	dafa9	pay
18	اشترى	ʔištaraa	buy
19	*دق*	daʔ	knock
20	اتصل*	ʔittaSal	call (by phone)
21	*دور*	dawwar	look for
22	انتبه*	ʔintabah	be aware
23	*رد*	rad	reply
24	*غفي*	Gifii	nod
25	*نام*	naam	sleep
26	*بعت*	ba9at	send
27	*أزعج*	ʔaz9aj	bother
28	*زعل*	zi9il	be upset
29	زار*	zaar	visit
30	*سأل*	saʔal	ask
31	*سمع*	simi9	hear
32	سوى	sawwaa	make
33	اشتاق*	ʔištaaʔ	miss
34	اشترك*	ʔištarak	subscribe
35	*تعلم*	ʔit9allim	learn
36	باع*	baa9	sell
37	استنى	ʔistannaa	wait
38	انتظر *	ʔintazar	wait
39	*رقص*	riʔiS	dance
40	*تمنى	ʔitmannaa	wish
41	*كلم*	kallim	talk
42	جاب*	qaab	bring
43	*حب*	Hab	love
44	*عمل*	9imil	make
45	*تذكر *	ʔitzakkar	remembered
46	*صرخ*	Sarax	shout
47	*بين*	bayyan	clarify

Appendix 1: List of Verbs for Searching Colloquial Arabic Corpora

48	اشتغل*	?ištaGal	work
49	*شكر	šakar	thank
50	شاف*	šaaf	see
51	*صبر	Sabar	be patient
52	*ضحك*	DiHik	laugh
53	*ضرب*	Dirib	hit
54	*ضيع*	Dayya9	loose
55	*طلب*	Tilib	request
56	*شلف*	šalaf	throw
57	عطى	9iTii	give
58	*عرف*	9irif	know
59	*غسل*	Gasal	wash
60	*فكر*	fakkar	think
61	*قصد*	?aSad	mean
62	دري*	dirii	know
63	قال*	?aal	say
64	*كتب*	kitib	write
65	*كذب*	kizib	tell a lie
66	*كمل*	kammil	complete
67	*لاحظ*	laaHiz	notice
68	*لبس*	libis	wear
69	*مسك*	misik	hold (by hand)
70	نادى	naadaa	call (somebody's name)
71	*نظف*	nazzif	clean
72	*وضح*	waDDaH	clarify
73	*بلش*	ballaš	start
74	*حضر*	HaDar	attend
75	*خلص*	xallaS	finish
76	*دردش*	dardaš	chat
77	استعمل*	?ista9mal	use
78	استفاد*	?istafaad	benefit
79	*سخن*	saxxan	warm (food)
80	*سهر*	sihir	stay up (at night)
81	*حرق*	Hara?	burn
82	*تغدى	?itGadda	lunch
83	*تعشى	?it9ašaa	dine
84	*تدلع*	?itdalla9	dally
85	*بالغ*	baalaG	exaggerate
86	*زهق*	zihi?	be bored
87	*ساعد*	saa9id	help
88	*تعب*	9ib	be tired
89	*تحمل*	?itHammal	endure
90	*دخن*	daxxan	smoke
91	*رجع*	riji9	get back
92	*بدل*	baddil	exchange
93	*غير*	Gayyar	change
94	*فتح*	fitiH	open
95	*قفل*	?afal	lock
96	*قطع*	?aTa9	cut
97	*جرح*	jiriH	wound

247

Appendix 1: List of Verbs for Searching Colloquial Arabic Corpora

98	وقف*	wi?if	stop
99	*تردد*	?itraddad	hesitate
100	*تجوز*	?itjawwaz	marry
101	*خطب*	xiTib	propose (marriage)
102	*نزل*	nizil	get down
103	*رفع*	rafa9	raise
104	*طلع*	Tili9	leave (a place)
105	*وصل*	wiSil	arrive
106	*أسرع*	?asra9	speed up
107	*لعب*	li9ib	play
108	*رمى*	rimii	throw
109	*زت*	zatt	throw
110	*عاقب*	9aa?ib	punish
111	*عاتب*	9aab	admonish
112	*عالج*	9aalij	treat
113	*حاسب*	Haasib	pay (money)
114	*رتب*	rattib	put in order
115	*جمع*	jama9	collect
116	*راجع*	raaji9	revise
117	*كشف*	kašaf	uncover
118	*فحص*	faHaS	examine
119	*شبع*	šibi9	be full (food)
120	*مل*	mall	be bored
121	*طفش*	Tifiš	be bored
122	*هرب*	hirib	flee
123	*سافر*	saafir	travel
124	*طبخ*	Tabax	cook
125	*صام*	Saam	fast (from food)
126	*تمشى*	?itmašaa	stroll
127	*بكي*	bikii	cry
128	*ناقش*	naa?iš	negotiate
129	*خاف*	xaaf	fear
130	*استلم*	?istalam	receive
131	*سكت*	sikit	stop talking
132	*داق*	daa?	taste
133	*ترك*	tarak	leave (something)
134	لغى	laGaa	cancel
135	*فهم*	fihim	understand
136	*نجح*	nijiH	succeed
137	*رسب*	rasab	fail
138	*قنع*	?ana9	convince
139	*رفض*	rafaD	refuse
140	*استأجر*	?ista?jar	rent

248

Colloquial Egyptian Arabic Verb List

No.	Verb in Arabic Script	Transliteration	English Translation
1	*أكل*	?akal	eat
2	*شرب*	širib	drink
3	*روح*	rawwaH	go
4	جه	qih	come
5	*حاول*	Haawil	try
6	*تكلم*	?itkallim	talk
7	صلى	Sallaa	pray
8	*تفرج*	?itfarraq	watch
9	*مشي*	mišii	walk
10	*توقع*	?itwa?a9	expect
11	*جرب*	qarrab	try
12	*جلس*	qalas	sit
13	*حس*	Hass	feel
14	*حط*	HaT	put
15	خد*	xad	take
16	*درس*	diris	study
17	*دفع*	dafa9	pay (money)
18	اشترى	?ištaraa	buy
19	*دق*	da?	knock
20	اتصل*	?ittaSal	call (by phone)
21	*دور*	dawwar	look for
22	*ذاكر*	zaakir	study
23	*رد*	rad	reply
24	*نعس*	ni9is	sleep
25	*نام*	naam	sleep
26	*بعت*	ba9at	send
27	*أزعج*	?az9aq	bother
28	*زعل*	zi9il	be upset
29	زار*	zaar	visit
30	*سأل*	sa?al	ask
31	*سمع*	simi9	hear
32	*عمل*	9amal	make
33	اشتاق*	?ištaa?	miss
34	اشترك*	?ištarak	subscribe
35	*تعلم*	?it9allim	learn
36	باع*	baa9	sell
37	استنى	?istannaa	wait
38	انتظر*	?intazar	wait
39	*رقص*	ra?aS	dance
40	اتمنى	?itmannaa	wish
41	*كلم*	kallim	talk
42	جاب*	qaab	bring, buy
43	*حب*	Hab	love
44	*حكي	Hikii	narrate
45	*تذكر*	?itzakkar	remembered
46	*نده*	nadah	call (somebody's name)
47	*بين*	bayyin	clarify

Appendix 1: List of Verbs for Searching Colloquial Arabic Corpora

48	اشتغل*	?ištaGal	work
49	*شكر*	šakar	thank
50	شاف*	šaaf	see
51	*صبر*	Sabar	be patient
52	*ضحك*	DiHik	laugh
53	*ضرب*	Darab	hit
54	*ضيع*	Dayyi9	loose
55	*طلب*	Talab	request
56	*طلع*	Tili9	leave (a place)
57	ادى	?iddaa	give
58	*عرف*	9irif	know
59	*غسل*	Gasal	wash
60	*فكر*	fakkar	think
61	*قصد*	?aSad	mean
62	*قعد*	?a9ad	sit
63	قال*	?aal	say
64	*كتب*	katab	write
65	*كذب*	kizib	tell a lie
66	*كمل*	kammil	complete
67	*لاحظ*	laaHiz	notice
68	*لبس*	libis	wear
69	*مسك*	misik	hold (by hand)
70	نادى	naadaa	call (somebody's name)
71	*نظف*	nazzif	clean
72	*وضح*	waDDaH	clarify
73	بدا*	badaa	start
74	*حضر*	Hadar	attend
75	*خلص*	xallaS	finish
76	*دردش*	dardaš	chat
77	استعمل*	?ista9mal	use
78	استفاد*	?istafaad	benefit
79	*سخن*	saxxan	warm
80	*سهر*	sihir	stay up
81	*حرق*	Hara?	burn
82	*تغدى*	?itGadda	lunch
83	*تعشى*	?it9ašaa	dine
84	*تدلع*	?itdalla9	dally
85	*بالغ*	baalaG	exaggerate
86	*زهق*	zihi?	be bored
87	*ساعد*	saa9id	help
88	*تعب*	9ib	be tired
89	*تحمل*	?itHammil	endure
90	*دخن*	daxxan	smoke
91	*رجع*	riqi9	get back
92	*بدل*	baddil	exchange
93	*غير*	Gayyar	change
94	*فتح*	fataH	open
95	*قفل*	?afal	lock
96	*قطع*	?aTa9	cut
97	*جرح*	qaraH	wound

Appendix 1: List of Verbs for Searching Colloquial Arabic Corpora

98	وقف*	wi?if	stop
99	*انتبه	?intabah	be aware
100	*تجوز*	?itqawwiz	marry
101	*خطب*	xaTab	propose (marriage)
102	*نزل*	nizil	get down
103	*رفع*	rafa9	raise
104	*صعد*	Si9id	get up (e.g., the stairs)
105	*وصل*	wiSil	arrive
106	*أسرع*	?asra9	speed up
107	*لعب*	li9ib	play
108	رمى	ramaa	throw
109	*حذف*	Hadaf	throw
110	*عاقب*	9aa?ib	punish
111	*عاتب*	9aab	admonish
112	*عالج*	9aaliq	treat
113	*حاسب*	Haasib	pay (money)
114	*رتب*	rattib	put in order
115	*جمع*	qama9	collect
116	*راجع*	raaqi9	revise
117	*كشف*	kašaf	uncover
118	*فحص*	faHaS	examine
119	*شبع*	šibi9	be full (of food)
120	*مل*	mall	get bored
121	*طفش*	Tifiš	get bored, flee
122	*هرب*	hirib	flee
123	*سافر*	saafir	travel
124	*طبخ*	Tabax	cook
125	*صام*	Saam	fast (from food)
126	*تمشى*	?itmašaa	stroll
127	*بكي*	bikii	cry
128	*ناقش*	naa?iš	negotiate
129	*خاف*	xaaf	fear
130	*استلم*	?istalam	receive
131	*سكت*	sikit	stop talking
132	*داق*	daa?	taste
133	*ترك*	tarak	leave (something)
134	لغى	laGaa	cancel
135	*فهم*	fihim	understand
136	*نجح*	niqiH	succeed
137	*رسب*	rasab	fail
138	*قنع*	?ana9	convince
139	*رفض*	rafaD	refuse
140	*أجر*	?aqqar	rent

Colloquial Gulf Arabic Adverb List

No.	Adverb in Arabic Script	Transliteration	English Translation
1	اليوم	?ilyuwm	today
2	أمس	?ams	yesterday
3	باكر	baaki	tomorrow
4	البارح	?ilbaariH	last night
5	البارحة	?ilbaarHah	last night
6	عادة	9aadatan	usually
7	أحيانا	?aHyaanan	occasionally
8	دايما	daayman	always
9	بدري	badrii	early
10	الليلة	?illylah	this night
11	أبدا	?abadan	never
12	أيضا	?ayDan	also
13	غالبا	Gaaliban	often
14	فعلا	fi9lan	indeed
15	حقيقة	Haqiqatan	truly
16	جد	jid	seriously
17	كثير	kθiir	a lot
18	تماما	tamaaman	completely
19	بوضوح	biwDuwH	clearly
20	بسرعة	bsir9ah	rapidly
21	بسهولة	bishuwlah	easily
22	بصعوبة	biS9uwbah	with difficulty
23	بالتحديد	bittaHdiid	precisely
24	خصوصا	xuSuwSan	especially
25	بالضبط	biDabT	precisely
26	الآن	?al?aan	now
27	عموما	9umuwman	generally
28	جدا	jiddan	very much
29	فجأة	faj?atan	suddenly
30	يوميا	yawmiyan	daily
31	فقط	faqaT	only
32	طبعا	Tab9an	evidently
33	شهريا	šahriyan	monthly
34	هالحين	halHiin	now
35	الحين	?alHiin	now
36	هنا	hinaa	here
37	هناك	hinaak	there
38	أسبوعيا	?isbuw9iyan	weekly
39	سنويا	sanawiyan	annually
40	كمان	kamaan	also
41	بالقوة	bilquwah	with force
42	قريبا	qariyban	soon
43	أخيرا	?axiyran	at last
44	أولاً	?awalan	firstly
45	ثانياً	θaaniyan	secondly
46	بشويش	bišwyiiš	softly
47	بس	bas	only

Appendix 2: List of Adverbs for Searching Colloquial Arabic Corpora

48	صراحة	SaraHah	frankly
49	أصلاً	?aSlan	basically
50	أساساً	?asaasan	basically

Appendix 2: List of Adverbs for Searching Colloquial Arabic Corpora

Colloquial Levantine Arabic Adverb List

No.	Adverb in Arabic Script	Transliteration	English Translation
1	اليوم	?ilyuwm	today
2	أمس	?ams	yesterday
3	باكر	baakir	tomorrow
4	مبارح	?imbaariH	last night
5	مبارحه	?imbaarHih	last night
6	عادة	9aadatan	usually
7	أحيانا	?aHyaanan	occasionally
8	دايما	daayman	always
9	بكير	bakkyiir	early
10	الليلة	?illylah	this night
11	أبدا	?abadan	never
12	تماما	tamaaman	completely
13	غالبا	Gaaliban	often
14	فعلا	fi9lan	indeed
15	حقيقة	Ha?i?atan	truly
16	جد	jad	seriously
17	كتير	ktiir	a lot
18	خصوصا	xuSuwSan	especially
19	بوضوح	buwDuwH	clearly
20	بسرعة	bsir9ah	rapidly
21	بسهولة	bishuwlih	easily
22	بصعوبة	biS9uwbih	with difficulty
23	بالضبط	biDDabiT	precisely
24	فورا	fawran	at once
25	يوميا	yawmiyan	daily
26	الآن	?il?aan	now
27	فقط	faqaT	only
28	جدا	jiddan	very much
29	فجأة	faj?ah	suddenly
30	طبعا	Tab9an	evidently
31	اسبوعيا	?isbuw9iyan	weekly
32	شهريا	šahriyan	monthly
33	سنويا	sanawiyan	annually
34	هلا	halla?	now
35	الحين	?ilHiin	now
36	هون	huwn	here
37	هونيك	huwniik	there
38	بالقوة	bil?uwih	with force
39	دغري	diGrii	directly
40	كمان	kamaan	also
41	بشويش	bišwyiiš	softly
42	قريبا	?ariyban	soon
43	أخيرا	?axiyran	at last
44	أولاً	?awalan	firstly
45	ثانياً	θaaniyan	secondly
46	بطلاقة	biTalaa?ah	fluently
47	بس	bas	only

Appendix 2: List of Adverbs for Searching Colloquial Arabic Corpora

48	بالتحديد	bittaHdiid	precisely
49	أصلاً	?aSlan	basically
50	أساساً	?asaasan	basically

Colloquial Egyptian Arabic Adverb List

No.	Adverb in Arabic Script	Transliteration	English Translation
1	النهارده	?innahardah	today
2	بكير	bakkyiir	early
3	بكرا	bukraa	tomorrow
4	مبارح	?imbaariH	last night
5	مبارحه	?imbaarHih	last night
6	عادة	9aadatan	usually
7	أحيانا	?aHyaanan	occasionally
8	دايما	daayman	always
9	بدري	badrii	early
10	الليلة	?illylah	this night
11	أبدا	?abadan	never
12	تماما	tamaaman	completely
13	غالبا	Gaaliban	often
14	فعلا	fi9lan	indeed
15	حقيقة	Ha?i?atan	truly
16	جد	qad	seriously
17	كتير	ktiir	a lot
18	طبعا	Tab9an	evidently
19	بوضوح	buwDuwH	clearly
20	بسرعه	bsur9ah	rapidly
21	بسهوله	bishuwlah	easily
22	بصعوبة	biS9uwbah	with difficulty
23	ببساطة	bibaSaTah	simply
24	بالتحديد	bittaHdiid	precisely
25	خصوصا	xuSuwSan	especially
26	الآن	?al?aan	now
27	عموما	9umuwman	generally
28	جدا	qiddan	very much
29	فجأة	faq?ah	suddenly
30	بالضبط	biDDabT	precisely
31	يوميا	yawmiyan	daily
32	اسبوعيا	?isbuw9iyan	weekly
33	شهريا	šahriyan	monthly
34	دلوقت	dilwa?ti	now
35	سنويا	sanawiyan	annually
36	هنا	hinaa	here
37	هناك	hinaak	there
38	دغري	duGrii	directly
39	بس	bas	only
40	قريبا	?ariyban	soon
41	أخيرا	?axiyran	at last
42	أولاً	?awalan	firstly
43	ثانياً	θaaniyan	secondly
44	فقط	fa?aT	only
45	أصلاً	?aSlan	basically
46	أساساً	?asasan	basically
47	كمان	kamaan	also

Appendix 2: List of Adverbs for Searching Colloquial Arabic Corpora

48	بالقوة	bil?uwah	with force
49	بشويش	bišwyiiš	softly
50	فورا	fawran	at once

Appendix 3: Demographic Details of Participants in the Experimental Study

		Participants ID	Age	Gender	Colloquial Variety	Age of First Exposure to MSA	Nursery & Kindergarten Type	Current Educational Level
1		GE101	13,1	Male	Gulf Arabic	Early	MSA I-P	Secondary School
2		GE102	14,11	Male	Gulf Arabic	Early	MSA I-P	Secondary School
3		GE103	13,9	Male	Gulf Arabic	Early	MSA I-P	Secondary School
4		GE104	13,2	Male	Gulf Arabic	Early	MSA I-P	Secondary School
5		GE105	13,5	Male	Gulf Arabic	Early	MSA I-P	Secondary School
6		GE107	13,8	Male	Gulf Arabic	Early	MSA I-P	Secondary School
7		GE108	13,8	Male	Gulf Arabic	Early	MSA I-P	Secondary School
8		GE109	13,7	Male	Gulf Arabic	Early	MSA I-P	Secondary School
9		GE110	14,7	Male	Gulf Arabic	Early	MSA I-P	Secondary School
10		GE112	14,5	Male	Gulf Arabic	Early	MSA I-P	Secondary School
11		GE113	13,8	Male	Gulf Arabic	Early	MSA I-P	Secondary School
12		GE114	13,0	Male	Gulf Arabic	Early	MSA I-P	Secondary School
13	E-CGA GROUP	GE115	13,11	Male	Gulf Arabic	Early	MSA I-P	Secondary School
14		GE116	14,3	Male	Gulf Arabic	Early	MSA I-P	Secondary School
15		GE117	14,9	male	Gulf Arabic	Early	MSA I-P	Secondary School
16		GE201	14,5	male	Gulf Arabic	Early	MSA I-P	Secondary School
17		GE202	13,3	male	Gulf Arabic	Early	MSA I-P	Secondary School
18		GE203	14,1	male	Gulf Arabic	Early	MSA I-P	Secondary School
19		GE204	13,5	male	Gulf Arabic	Early	MSA I-P	Secondary School
20		GE205	13,2	male	Gulf Arabic	Early	MSA I-P	Secondary School
21		GE206	13,1	male	Gulf Arabic	Early	MSA I-P	Secondary School
22		GE207	14,6	male	Gulf Arabic	Early	MSA I-P	Secondary School
23		GE210	13,11	male	Gulf Arabic	Early	MSA I-P	Secondary School
24		GE211	13,6	male	Gulf Arabic	Early	MSA I-P	Secondary School
25		GE212	13,2	male	Gulf Arabic	Early	MSA I-P	Secondary School
26		GE213	13,10	male	Gulf Arabic	Early	MSA I-P	Secondary School
27		GE214	15,0	male	Gulf Arabic	Early	MSA I-P	Secondary School
28		GE215	13,8	male	Gulf Arabic	Early	MSA I-P	Secondary School

Appendix 3: Demographic Details of Participants in the Experimental Study

		Participants ID	Age	Gender	Colloquial Variety	Age of First Exposure to MSA	Nursery & Kindergarten Type	Current Educational Level
29		GE216	13,5	male	Gulf Arabic	Early	MSA I-P	Secondary School
30		GE217	13,0	male	Gulf Arabic	Early	MSA I-P	Secondary School
31		GL101	14,11	male	Gulf Arabic	Late	N/A	Secondary School
32		GL102	13,5	male	Gulf Arabic	Late	N/A	Secondary School
33		GL103	14,8	male	Gulf Arabic	Late	N/A	Secondary School
34		GL104	14,2	male	Gulf Arabic	Late	N/A	Secondary School
35		GL105	14,10	male	Gulf Arabic	Late	N/A	Secondary School
36		GL106	13,0	male	Gulf Arabic	Late	N/A	Secondary School
37		GL107	15,9	male	Gulf Arabic	Late	N/A	Secondary School
38		GL108	12,0	male	Gulf Arabic	Late	N/A	Secondary School
39		GL109	13,7	male	Gulf Arabic	Late	N/A	Secondary School
40		GL110	14,7	male	Gulf Arabic	Late	N/A	Secondary School
41		GL111	14,0	male	Gulf Arabic	Late	N/A	Secondary School
42	L-CGA GROUP	GL112	14,3	male	Gulf Arabic	Late	N/A	Secondary School
43		GL113	13,8	male	Gulf Arabic	Late	N/A	Secondary School
44		GL114	14,9	male	Gulf Arabic	Late	N/A	Secondary School
45		GL115	14,9	male	Gulf Arabic	Late	N/A	Secondary School
46		GL201	14,7	male	Gulf Arabic	Late	N/A	Secondary School
47		GL202	13,0	male	Gulf Arabic	Late	N/A	Secondary School
48		GL203	13,10	male	Gulf Arabic	Late	N/A	Secondary School
49		GL204	14,2	male	Gulf Arabic	Late	N/A	Secondary School
50		GL205	15,2	male	Gulf Arabic	Late	N/A	Secondary School
51		GL206	14,11	male	Gulf Arabic	Late	N/A	Secondary School
52		GL207	14,0	male	Gulf Arabic	Late	N/A	Secondary School
53		GL208	14,2	male	Gulf Arabic	Late	N/A	Secondary School
54		GL209	14,1	male	Gulf Arabic	Late	N/A	Secondary School
55		GL210	14,0	male	Gulf Arabic	Late	N/A	Secondary School
56		GL211	14,6	male	Gulf Arabic	Late	N/A	Secondary School

Appendix 3: Demographic Details of Participants in the Experimental Study

		Participants ID	Age	Gender	Colloquial Variety	Age of First Exposure to MSA	Nursery & Kindergarten Type	Current Educational Level
57		GL212	14,0	male	Gulf Arabic	Late	N/A	Secondary School
58		GL213	13,2	male	Gulf Arabic	Late	N/A	Secondary School
59		GL214	14,0	male	Gulf Arabic	Late	N/A	Secondary School
60		GL215	13,10	male	Gulf Arabic	Late	N/A	Secondary School
61		LE101	13,02	male	Levantine Arabic	Early	MSA I-P	Secondary School
62		LE102	14,02	male	Levantine Arabic	Early	MSA I-P	Secondary School
63		LE103	13,05	male	Levantine Arabic	Early	MSA I-P	Secondary School
64		LE104	13,03	male	Levantine Arabic	Early	MSA I-P	Secondary School
65		LE105	13,07	male	Levantine Arabic	Early	MSA I-P	Secondary School
66		LE106	13,09	male	Levantine Arabic	Early	MSA I-P	Secondary School
67		LE107	13,10	female	Levantine Arabic	Early	MSA I-P	Secondary School
68		LE108	14,03	female	Levantine Arabic	Early	MSA I-P	Secondary School
69		LE109	14,03	female	Levantine Arabic	Early	MSA I-P	Secondary School
70		LE110	14,00	female	Levantine Arabic	Early	MSA I-P	Secondary School
71		LE111	14,00	female	Levantine Arabic	Early	MSA I-P	Secondary School
72	E-CLA GROUP	LE112	14,00	female	Levantine Arabic	Early	MSA I-P	Secondary School
73		LE113	13,07	male	Levantine Arabic	Early	MSA I-P	Secondary School
74		LE114	14,01	male	Levantine Arabic	Early	MSA I-P	Secondary School
75		LE115	14,02	male	Levantine Arabic	Early	MSA I-P	Secondary School
76		LE201	13,08	male	Levantine Arabic	Early	MSA I-P	Secondary School
77		LE202	14,03	male	Levantine Arabic	Early	MSA I-P	Secondary School
78		LE203	13,11	male	Levantine Arabic	Early	MSA I-P	Secondary School
79		LE204	15,03	male	Levantine Arabic	Early	MSA I-P	Secondary School
80		LE205	13,09	male	Levantine Arabic	Early	MSA I-P	Secondary School
81		LE206	13,10	male	Levantine Arabic	Early	MSA I-P	Secondary School
82		LE207	13,10	male	Levantine Arabic	Early	MSA I-P	Secondary School
83		LE208	13,08	male	Levantine Arabic	Early	MSA I-P	Secondary School
84		LE209	14,10	male	Levantine Arabic	Early	MSA I-P	Secondary School

		Participants ID	Age	Gender	Colloquial Variety	Age of First Exposure to MSA	Nursery & Kindergarten Type	Current Educational Level
85		LE210	13,03	female	Levantine Arabic	Early	MSA I-P	Secondary School
86		LE211	14,00	female	Levantine Arabic	Early	MSA I-P	Secondary School
87		LE212	14,01	female	Levantine Arabic	Early	MSA I-P	Secondary School
88		LE213	14,00	male	Levantine Arabic	Early	MSA I-P	Secondary School
89		LE214	13,07	male	Levantine Arabic	Early	MSA I-P	Secondary School
90		LE215	14,01	female	Levantine Arabic	Early	MSA I-P	Secondary School
91		LL101	14,00	male	Levantine Arabic	Late	N/A	Secondary School
92		LL102	13,03	male	Levantine Arabic	Late	N/A	Secondary School
93		LL103	16,00	male	Levantine Arabic	Late	N/A	Secondary School
94		LL104	13,08	male	Levantine Arabic	Late	N/A	Secondary School
95		LL105	13,07	male	Levantine Arabic	Late	N/A	Secondary School
96		LL106	14,01	male	Levantine Arabic	Late	N/A	Secondary School
97		LL107	13,10	male	Levantine Arabic	Late	N/A	Secondary School
98		LL108	14,03	male	Levantine Arabic	Late	N/A	Secondary School
99		LL109	14,00	male	Levantine Arabic	Late	N/A	Secondary School
100	L-CLA GROUP	LL110	13,09	male	Levantine Arabic	Late	N/A	Secondary School
101		LL111	13,09	male	Levantine Arabic	Late	N/A	Secondary School
102		LL112	13,11	male	Levantine Arabic	Late	N/A	Secondary School
103		LL113	13,11	male	Levantine Arabic	Late	N/A	Secondary School
104		LL114	13,11	male	Levantine Arabic	Late	N/A	Secondary School
105		LL115	13,04	male	Levantine Arabic	Late	N/A	Secondary School
106		LL201	13,10	male	Levantine Arabic	Late	N/A	Secondary School
107		LL202	13,04	male	Levantine Arabic	Late	N/A	Secondary School
108		LL203	14,00	male	Levantine Arabic	Late	N/A	Secondary School
109		LL204	14,03	male	Levantine Arabic	Late	N/A	Secondary School
110		LL205	13,08	male	Levantine Arabic	Late	N/A	Secondary School
111		LL206	14,01	male	Levantine Arabic	Late	N/A	Secondary School
112		LL207	13,08	male	Levantine Arabic	Late	N/A	Secondary School

		Participants ID	Age	Gender	Colloquial Variety	Age of First Exposure to MSA	Nursery & Kindergarten Type	Current Educational Level
113		LL208	13,04	male	Levantine Arabic	Late	N/A	Secondary School
114		LL209	13,09	male	Levantine Arabic	Late	N/A	Secondary School
115		LL210	13,06	male	Levantine Arabic	Late	N/A	Secondary School
116		LL211	14,00	male	Levantine Arabic	Late	N/A	Secondary School
117		LL212	13,03	male	Levantine Arabic	Late	N/A	Secondary School
118		EL101	15,9	male	Egyptian Arabic	Late	N/A	High School
119		EL102	16,6	male	Egyptian Arabic	Late	N/A	High School
120		EL103	16,9	male	Egyptian Arabic	Late	N/A	High School
121		EL104	17,8	male	Egyptian Arabic	Late	N/A	High School
122		EL105	18,9	male	Egyptian Arabic	Late	N/A	High School
123		EL106	17,11	male	Egyptian Arabic	Late	N/A	High School
124		EL107	18,5	male	Egyptian Arabic	Late	N/A	High School
125	L-CEA GROUP	EL108	16,0	male	Egyptian Arabic	Late	N/A	High School
126		EL109	17,9	male	Egyptian Arabic	Late	N/A	High School
127		EL110	17,6	male	Egyptian Arabic	Late	N/A	High School
128		EL111	16,3	male	Egyptian Arabic	Late	N/A	High School
129		EL112	17,1	male	Egyptian Arabic	Late	N/A	High School
130		EL113	15.5	male	Egyptian Arabic	Late	N/A	High School
131		EL114	15,10	male	Egyptian Arabic	Late	N/A	High School
132		EL115	16	male	Egyptian Arabic	Late	N/A	High School
133		EL201	16,7	male	Egyptian Arabic	Late	N/A	High School
134		EL202	16,11	male	Egyptian Arabic	Late	N/A	High School
135		EL203	16,3	male	Egyptian Arabic	Late	N/A	High School
136		EL204	17,5	male	Egyptian Arabic	Late	N/A	High School
137		EL205	17,3	male	Egyptian Arabic	Late	N/A	High School
138		EL206	17,4	male	Egyptian Arabic	Late	N/A	High School
139		EL207	17,0	male	Egyptian Arabic	Late	N/A	High School
140		EL208	18,0	male	Egyptian Arabic	Late	N/A	High School

Appendix 3: Demographic Details of Participants in the Experimental Study

		Participants ID	Age	Gender	Colloquial Variety	Age of First Exposure to MSA	Nursery & Kindergarten Type	Current Educational Level
141		EL209	16,2	male	Egyptian Arabic	Late	N/A	High School
142		EL210	15,1	male	Egyptian Arabic	Late	N/A	High School
143		EL211	16,7	male	Egyptian Arabic	Late	N/A	High School
144		EL212	15,0	male	Egyptian Arabic	Late	N/A	High School
145		EL213	17,7	male	Egyptian Arabic	Late	N/A	High School
146		EL214	16.5	male	Egyptian Arabic	Late	N/A	High School
147		EL215	15.2	male	Egyptian Arabic	Late	N/A	High School

<div dir="rtl">

الإرشادات (المهمة الأولى)

في هذه المهمة، سترى عدداً من الجمل على الشاشة من اللغة العربية الفصحى؛ من فضلك قرر ما إذا كانت هذه الجمل مقبولة لديك من حيث التركيب. حدد إجابتك برسم دائرة على الخيار المناسب في الجدول الآتي:

لا أستطيع القرار		جيدة جداً مقبولة تماماً	تبدو جيدة مقبولة	غريبة بعض الشيء لاتبدومقبولة لدي	غريبة جداً غيرمقبولة لدي
X		+2	+1	-1	-2

أمثلة

(هل الجملة تبدو لك غريبة أم جيدة؟)

لا أستطيع القرار		جيدة جداً مقبولة تماماً	تبدو جيدة مقبولة	غريبة بعض الشيء لاتبدو مقبولة لدي	غريبة جداً غير مقبولة لدي	
X	١	+2	+1	-1	-2	المثال الأول
X		+2	+1	-1	-2	المثال الثاني
X		+2	+1	-1	-2	المثال الثالث

</div>

264

Instructions (Task 1)

In this task, you will see a number of single sentences in Standard Arabic. Please judge whether the sentence is grammatically acceptable to you. Indicate your answer by circling one of the options on the scale on your answer sheet. The scale is as follows:

Very strange. Unacceptable.	A bit strange. Not really acceptable.	Fairly good. Acceptable.	Perfectly good. Perfectly Acceptable		Can't decide
−2	−1	+1	+2		X

Examples

(Does the sentence sound strange or good to you?)

	Very strange. Unacceptable.	A bit strange. Not really acceptable.	Fairly good. Acceptable.	Perfectly good. Perfectly acceptable.		Can't decide
Ex. 1	−2	−1	+1	+2		X
Ex. 2	−2	−1	+1	+2		X
Ex. 3	−2	−1	+1	+2		X

Appendix 4: B.1 The AJT Answer Sheet (in Arabic)

(هل الجملة تبدو لك غريبة أم جيدة؟)

لا أستطيع القرار		جيدة جداً مقبولة تماماً	تبدو جيدة مقبولة	غريبة بعض الشيء لاتبدومقبولة لدي	غريبة جداً غيرمقبولة لدي	
X		+2	+1	-1	-2	1
X		+2	+1	-1	-2	2
X		+2	+1	-1	-2	3
X		+2	+1	-1	-2	4
X		+2	+1	-1	-2	5
X		+2	+1	-1	-2	6
X		+2	+1	-1	-2	7
X		+2	+1	-1	-2	8
X		+2	+1	-1	-2	9
X		+2	+1	-1	-2	10
X		+2	+1	-1	-2	11
X		+2	+1	-1	-2	12
X		+2	+1	-1	-2	13
X		+2	+1	-1	-2	14
X		+2	+1	-1	-2	15
X		+2	+1	-1	-2	16
X		+2	+1	-1	-2	17
X		+2	+1	-1	-2	18
X		+2	+1	-1	-2	19
X		+2	+1	-1	-2	20
X		+2	+1	-1	-2	21
X		+2	+1	-1	-2	22
X		+2	+1	-1	-2	23
X		+2	+1	-1	-2	24
X		+2	+1	-1	-2	25
X		+2	+1	-1	-2	26
X		+2	+1	-1	-2	27
X		+2	+1	-1	-2	28
X		+2	+1	-1	-2	29
X		+2	+1	-1	-2	30

266

Appendix 4: B.1 The AJT Answer Sheet (in Arabic)

لا أستطيع القرار		جيدة جداً مقبولة تماماً	تبدو جيدة مقبولة	غريبة بعض الشيء لاتبدومقبولة لدي	غريبة جداً غيرمقبولة لدي	
X		+2	+1	-1	-2	**31**
X		+2	+1	-1	-2	**32**
X		+2	+1	-1	-2	**33**
X		+2	+1	-1	-2	**34**
X		+2	+1	-1	-2	**35**
X		+2	+1	-1	-2	**36**
X		+2	+1	-1	-2	**37**
X		+2	+1	-1	-2	**38**
X		+2	+1	-1	-2	**39**
X		+2	+1	-1	-2	**40**
X		+2	+1	-1	-2	**41**
X		+2	+1	-1	-2	**42**
X		+2	+1	-1	-2	**43**
X		+2	+1	-1	-2	**44**
X		+2	+1	-1	-2	**45**
X		+2	+1	-1	-2	**46**
X		+2	+1	-1	-2	**47**
X		+2	+1	-1	-2	**48**
X		+2	+1	-1	-2	**49**
X		+2	+1	-1	-2	**50**
X		+2	+1	-1	-2	**51**
X		+2	+1	-1	-2	**52**
X		+2	+1	-1	-2	**53**
X		+2	+1	-1	-2	**54**
X		+2	+1	-1	-2	**55**
X		+2	+1	-1	-2	**56**
X		+2	+1	-1	-2	**57**
X		+2	+1	-1	-2	**58**
X		+2	+1	-1	-2	**59**
X		+2	+1	-1	-2	**60**

267

Appendix 4: B.2 The AJT Answer Sheet (English Version)

(Does the sentence sound strange or good to you?)

	Very strange. Impossible.	A bit strange. Not really possible.	Fairly good. Possible.	Perfectly good. Perfectly possible		Can't decide
1	−2	−1	+1	+2		X
2	−2	−1	+1	+2		X
3	−2	−1	+1	+2		X
4	−2	−1	+1	+2		X
5	−2	−1	+1	+2		X
6	−2	−1	+1	+2		X
7	−2	−1	+1	+2		X
8	−2	−1	+1	+2		X
9	−2	−1	+1	+2		X
10	−2	−1	+1	+2		X
11	−2	−1	+1	+2		X
12	−2	−1	+1	+2		X
13	−2	−1	+1	+2		X
14	−2	−1	+1	+2		X
15	−2	−1	+1	+2		X
16	−2	−1	+1	+2		X
17	−2	−1	+1	+2		X
18	−2	−1	+1	+2		X
19	−2	−1	+1	+2		X
20	−2	−1	+1	+2		X
21	−2	−1	+1	+2		X
22	−2	−1	+1	+2		X
23	−2	−1	+1	+2		X
24	−2	−1	+1	+2		X
25	−2	−1	+1	+2		X
26	−2	−1	+1	+2		X
27	−2	−1	+1	+2		X
28	−2	−1	+1	+2		X
29	−2	−1	+1	+2		X
30	−2	−1	+1	+2		X

	Very strange. Impossible.	A bit strange. Not really possible.	Fairly good. Possible.	Perfectly good. Perfectly possible		Can't decide
31	−2	−1	+1	+2		X
32	−2	−1	+1	+2		X
33	−2	−1	+1	+2		X
34	−2	−1	+1	+2		X
35	−2	−1	+1	+2		X
36	−2	−1	+1	+2		X
37	−2	−1	+1	+2		X
38	−2	−1	+1	+2		X
39	−2	−1	+1	+2		X
40	−2	−1	+1	+2		X
41	−2	−1	+1	+2		X
42	−2	−1	+1	+2		X
43	−2	−1	+1	+2		X
44	−2	−1	+1	+2		X
45	−2	−1	+1	+2		X
46	−2	−1	+1	+2		X
47	−2	−1	+1	+2		X
48	−2	−1	+1	+2		X
49	−2	−1	+1	+2		X
50	−2	−1	+1	+2		X
51	−2	−1	+1	+2		X
52	−2	−1	+1	+2		X
53	−2	−1	+1	+2		X
54	−2	−1	+1	+2		X
55	−2	−1	+1	+2		X
56	−2	−1	+1	+2		X
57	−2	−1	+1	+2		X
58	−2	−1	+1	+2		X
59	−2	−1	+1	+2		X
60	−2	−1	+1	+2		X

Appendix 5: Test Sentences in the Acceptability Judgment Task

Sentence Type 1A: Object Relatives with Resumptive Pronouns:

1. قرأ ماجدٌ الكتابَ الذي اشتراهُ الأسبوعَ الماضي

qara?a	maajid-un	l-kitaab-a	l-aðii	štaraa-hu	l-?usbuw9-a	l-maaDii
read	Majid-nom	the-book-acc	that-s.m.	(he)bought-it	the-week	the-past

'Majid read the book that he bought last week'

2. حَفِظَ نوّافٌ القصيدةَ التي سَمِعَها في المِذياع

HafiZa	nawaaf-un	l-qaSiidat-a	l-atii	sami9a-ha	fii	l-miðyaa9
memorized	Nawaaf-nom	the-poem-acc	that-s.f.	(he) heard-it	on	the-radio

'Nawaaf learnt by heart the poem that he heard on the radio'

3. قاد صالحٌ السيّارةَ التي استأجرها من المطار

qaada	SaaliH-un	l-sayyaarat-a	l-atii	sta?jara-ha	mina	l-maTaar
drove	Salih-nom	the-car-acc	that-s.f.	(he) rented-it	from	the-airport

'Salih drove the car that he hired from the airport'

4. هاتفتْ فاطمةُ النسوةَ اللاتي قابلتهُنّ في السوق

haataf-at	faaTimat-u	l-niswat-a	l-aatii	qaabalat-hunna fii l-suwq
called-s.f.	Fatimah-nom	the-women	that-p.f.	(she) met-them on the-mall

'Fatimah called the women that she met with on the mall'

5. سلمانُ وجذ القلمينِ اللَذينِ أضاعَهُما الليلةَ الماضية

salmaan-u	wajada	l-qalam-ayni	l-aðayni	?aDa9a-humaa	l-laylata	l-maaDiyah
Salman-nom	found-3sm	the-pen-d.m.	that-d.m.	(he)lost-them(d.)	the-night	the-past

'Salman found the two pens that he lost last night'

6. فهدٌ قرأ الرسالتين اللتين كتبهما أخوهُ

fahad-un	qara?a	l-risaalat-ayni	l-atayni	kataba-humaa	?axuw-hu
Fahad-nom	read	the-letter-d..f.	that-d.f.	wrote-them(d.)	brother-his

'Fahad read the two letters that his brother wrote'

7. خالدٌ تحدّث مع الرجالِ الذين قابلهم في المسجدِ

xaalid-un	taHaddaθa ma9a	l-rijaali	l-aðiina	qabala-hum	fii l-masjid
Khalid-nom	talked	with the-men	that-p.m.	(he) met-them	on the-Mosque

'Khalid talked with the men that he met at the Mosque'

8. ثامرٌ أغلقَ البابَ الذي فتحه في الصباح

θaamir-un	?aGlaqa	l-baab-a	l-aðii	fataHa-hu	fii l-SabaaH
Thamir-nom	shut	the-door-acc	that-s.m.	(he) opened-it	in the-morning

'Thamir shut the door that he opened in the morning'

9. حَرِصْتُ أنْ أذاكِرَ ما تعلّمتُهُ في المدرسة

HariStu	?an ?uðaakira	maa	ta9allamtu-hu	fii l-madrasati
(I was keen)-1s.	that (I) study	what	(I) learnt-it	on the-school

'I was keen to study what I have learnt at school'

10. اردتُ أنْ أرحب بمن رأيتهم في حفل الزواج

?aradtu	?an ?uraHHiba	bi-man	ra?aytu-hum	fii Hafli l-zawaaji
(I) wanted	that (I) welcome	with-who	(I) saw-them	on party the-wedding

'I wanted to welcome whom I saw at the wedding party'

Sentence Type 1B: Object Relatives without Resumptive Pronouns:

1. قرأ ماجدٌ الكتابَ الذي اشترى الأسبوعَ الماضي

qara?a	maajid-un	l-kitaab-a		l-aðii	štaraa	l-?usbuw9-a	l-maaDii
read	Majid-nom	the-book-acc		that-s.m.	(he)bought	the-week	the-past

'Majid read the book that he bought last week'

2. حَفِظَ نوَّافٌ القصيدةَ التي سَمِعَ في المذياع

HafiZa	nawaaf-un	l-qaSiidat-a	l-atii	sami9a	fii	l-miðyaa9
memorized	Nawaaf-nom	the-poem-acc	that-s.f.	(he)heard	on	the-radio

'Nawaaf learnt by heart the poem that he heard on the radio'

3. قادَ صالحٌ السيّارَةَ التي استأجرَ من المطار

qaada	SaaliH-un	l-sayyaarat-a	l-atii	sta?jara	mina	l-maTaar
drove	Salih-nom	the-car-acc	that-s.f.	(he)rented	from	the-airport

'Salih drove the car that he hired from the airport'

4. هاتفتْ فاطمةُ النسوةَ اللاتي قابَلتْ في السوق

haataf-at	faaTimat-u	l-niswat-a	l-aatii	qaabalat	fii l-suwq
called-s.f.	Fatimah-nom	the-women	that-p.f.	(she)met	on the-mall

'Fatimah called the women that she met with on the mall'

5. سلمانُ وجدَ القلمينِ اللَّذينِ أضاعَ الليلةَ الماضية

salmaan-u	wajada	l-qalam-ayni l-aðayni	?aDa9a	l-laylata	l-maaDiyah
Salman-nom	found-3sm	the-pen-d.m. that-d.m.	(he)lost	the-night	the-past

'Salman found the two pens that he lost last night'

6. فهدٌ قرأ الرسالتين اللتين كتبَ أخوهُ

fahad-un	qara?a	l-risaalat-ayni l-atayni	kataba	?axuw-hu
Fahad-nom	read	the-letter-d..f. that-d.f.	wrote	brother-his

'Fahad read the two letters that his brother wrote'

7. خالدٌ تحدثُ مع الرجالِ الذين قابلَ في المسجدِ

xaalid-un	taHaddaθa ma9a l-rijaali	l-aðiina	qabala	fii l-masjid
Khalid-nom	talked with the-men	that-p.m.	(he)met	on the-Mosque

'Khalid talked with the men that he met at the Mosque'

8. ثامرٌ أغلقَ البابَ الذي فتحَ في الصباح

θaamir-un	?aGlaqa l-baab-a		l-aðii	fataHa	fii	l-SabaaH
Thamir-nom	shut the-door-acc		that-s.m.	(he)opened	in	the-morning

'Thamir shut the door that he opened in the morning'

9. خَرَصْتُ أنْ أذاكرَ ما تعلّمتُ في المدرسة

HariStu	?an ?uðaakira maa ta9allamtu fii l-madrasati
(I was keen)-1s.	that (I) study what (I) learnt on the-school

'I was keen to study what I have learnt at school'

10. أردتُ أنْ أرحّبَ بمن رأيتُ في حفل الزواج

?aradtu	?an ?uraHHiba bi-man ra?aytu fii Hafli l-zawaaji
(I) wanted	that (I) welcome with-who (I) saw on party the-wedding

'I wanted to welcome whom I saw at the wedding party'

Sentence Type 2A: SV sentences with human collective subjects + plural agreement

1. الناسُ يخافون الخروج من المنزل ليلاً في هذا البلد

l-naas-u   yaxaaf-uwna l-xuruwja mina l-manzili   laylaan fii haðaa l-balad
the-people-nom   fear-3pm   going out from the-house at night in this   the-country
'People avoid going out at night in this country'

2. الناسُ هنا يُفَضّلونَ أكل الفاكهة قبلَ قدومِ المساءِ

l-naas-u   hunaa yufaDDil-uwna ?akla l-faakihata qabla quduwmi   l-masaa?i
the-people-nom here prefer-3pm   eating the-fruits   before the-coming-of the-evening
'People here prefer eating fruits before evening time'

3. الناسُ يعيشونَ يومهم دونَ قلقٍ حولَ مُستقبلِ أيامهم

l-naas-u   ya9iiš-uwna yawma-hum duwna qalaqin Hawla mustaqbali ?ayaami-him
the-people-nom   live-3pm   day-their   without worry about future   days-their
'People live the day without worries about their future days'

4. الناسُ يقضونَ الإجازاتِ في زيارةِ أماكنَ مُختلفةٍ من العالم

l-naas-u   yaqD-uwna l-?ijaazaati fii ziyaarati ?amaakina muxtalifatin mina l-9aalam
the-people-nom   spend-3pm the-holidays on visiting places   different   from the-world
'People spend holidays on visiting different places of the world'

5. النساءُ يقضين الكثيرَ من الوقتِ في اهتمامهنّ بأولادهن

l-nisaa?-u   yaqD-iina l-kaθiira mina l-waqti fii ?ihtimaami-hinna bi-?awlaadi-hinna
the-women-nom   spend-3pf the-lot   from the-time on care-their   with-children-their
'Women spend a lot of time on taking care of their children'

6. العربُ اتفقوا على موقفٍ واحدٍ تجاه القضيةِ الفلسطينية

l-9arab-u   ?itafaq-uw 9alaa mawqifin waHidin tijaaha l-qaDiyyati l-filisTiiniyyah
the-arabs-nom   agreed-3pm on   stance   one   toward the-case   the-Palestinian
'Arabs agreed to take one position toward the Palestinian case'

7. البشرُ تمكّنوا من استكشافِ مُعظمِ مناطقِ الكرةِ الأرضية

l-bašar-u   tamakkan-uw   min ?istikšaafi mu9Zami manaaTiqi l-kurati   l-?arDiyyah
the-humans-nom   (could)-3pm   from explore   most of areas   the-global   the-earth
'Humans succeeded in exploring most of the global earth'

8. القومُ أجمعوا على مغادرةِ أماكنِ الخطرِ في الباديةِ

l-qawm-u   ?ajma9-uw 9alaa muGaadarati ?amaakini l-xaTari   fii l-baadiyyah
the-folk-nom   agreed-3pm on   leaving   places of   the-hazard   on the-desert
'The folks agreed on leaving hazardous places of the desert'

9. الوفدُ وصلوا إلى مقرّ إقامتهم هذا الصباحِ سالمين

l-wafd-u   waSal-uw   ?ilaa maqarri   ?iqaamati-him haðaa l-SabaHi   saalimeen
the-delegation-nom   arrived-3pm to   the place of   stay-their   this the-morning safe
'The delegation arrived to their place of residence safely this morning'

10. الشعبُ طالَبوا المسئولينَ بدراسةِ قضايا البطالةِ في البلد

l-ša9b-u   Taalab-uw   l-mas?uwliina   bi-diraasati qaDaayaa l-baTaalati   fii l-balad
the-nation-nom demanded-3pm the-responsible people with-studying issues the-unemployment in the-country
'The nation asked officials to study the unemployment issues in the country'

Sentence Type 2B: SV sentences with human collective subjects + no plural agreement

1. الناسُ تخافُ الخروج من المنزل ليلا" في هذا البلد

l-naas-u taxaafu l-xuruwja mina l-manzili laylaan fii haðaa l-balad
the-people-nom fear-3sf going out from the-house at night in this the-country
'People avoid going out at night in this country'

2. الناسُ هنا تفضّلُ أكل الفاكهة قبلَ قدومِ المساء

l-naas-u hunaa tufaDDilu ?akla l-faakihata qabla quduwmi l-masaa?i
the-people-nom here prefer-3sf eating the-fruits before the-coming-of the-evening
'People here prefer eating fruits before evening time'

3. الناسُ تعيشُ يومها دونَ قلقٍ حولَ مُستقبلِ أيامهم

l-naas-u ta9iišu yawma-hum duwna qalaqin Hawla mustaqbali ?ayaami-him
the-people-nom live-3sf day-their without worry about future days-their
'People live the day without worries about their future days'

4. الناسُ تقضي الإجازاتِ في زيارةِ أماكنَ مُختلفةٍ من العالم

l-naas-u taqDii l-?ijaazaati fii ziyaarati ?amaakina muxtalifatin mina l-9aalam
the-people-nom spend-3sf the-holidays on visiting places different from the-world
'People spend holidays on visiting different places of the world'

5. النساءُ تقضي الكثيرَ من الوقتِ في اهتمامهنّ بأولادهن

l-nisaa?-u taqDii l-kaθiira mina l-waqti fii ?ihtimaami-hinna bi-?awlaadi-hinna
the-women-nom spend-3sf the-lot from the-time on care-their with-children-their
'Women spend a lot of time on taking care of their children'

6. العربُ اتفقت على موقفٍ واحدٍ تجاة القضيةِ الفلسطينية

l-9arab-u ?itafaq-at 9alaa mawqifin waHidin tijaaha l-qaDiyyati l-filisTiiniyyah
the-arabs-nom agreed-3sf on stance one toward the-case the-Palestinian
'Arabs agreed to take one position toward the Palestinian case'

7. البشرُ تمكنت من استكشافِ مُعظمِ مناطقِ الكرةِ الأرضية

l-bašar-u tamakkan-at min ?istikšaafi mu9Zami manaaTiqi l-kurati l-?arDiyyah
the-humans-nom (could)-3sf from explore most of areas the-global the-earth
'Humans succeeded in exploring most of the global earth'

8. القومُ أجمعت على مغادرةِ أماكنِ الخطرِ في البادية

l-qawm-u ?ajma9-at 9alaa muGaadarati ?amaakini l-xaTari fii l-baadiyyah
the-folk-nom agreed-3sf on leaving places of the-hazard on the-desert
'The folks agreed on leaving hazardous places of the desert'

9. الوفد وصلَ إلى مقرٍّ إقامته هذا الصباحِ سالماً

l-wafd-u waSal-a ?ilaa maqarri ?iqaamati-him haðaa l-SabaHi saaliman [48]
the-delegation-nom arrived-3sm to the place of stay-their this the-morning safe
'The delegation arrived to their place of residence safely this morning'

10. الشعبُ طالَبَ المسئولينَ بدراسةِ قضايا البطالةِ في البلد

l-ša9b-u Taalab-a l-mas?uwliina bi-diraasati qaDaayaa l-baTaalati fii l-balad
the-nation-nom demanded-3sm the-responsible people with-studying issues the-unemployment in the-country
'The nation asked officials to study the unemployment issues in the country'

---

[48] Although the issue here is the agreement/disagreement in number between the preverbal collective noun and the following verb, there is a discrepancy in relation to agreement in gender: while in sentences 1-8, the gender feature of the verb is feminine, in sentences 9 & 10, it is masculine. I ckecked the literature to see why the gender changes with these particular collective nouns (الوفد 'the delegation' & الشعب 'the nation', in 9 & 10, respectively), but I could not find an explanation.

Sentence Type 3A: Distractor Sentences (Grammatical)

1. خالدٌ وأحمد تعلما مهارات الحاسوب في مدرسة واحدة

xaalid-un   wa   ?aHmad-u   ta9allam-aa mahaaraat l-Haasuwbi   fii madrasatin waaHidah
Khalid-nom and Ahmad-nom learnt-3dm skills   the-computer in school   one
'Khalid and Ahmad learnt their computer skills at the same school'

2. عليٌّ اشترى طعام العشاء من مطعم مجاور لبيته

9aliy-un   štaraa   Ta9aama l-9ašaa?i   min maT9amin mujaawirin li-bayt-ih
Ali-nom   bought-3sm food   the-dinner from restaurant next   to-house-his
'Ali bought dinner from a restaurant next to his house'

3. ابراهيمُ لم يذهب لحضور زواج أخيه الليلة الماضية

?ibraahiim-u lam yaðhab li-HuDuwri zawaaji ?axii-hi   l-laylata l-maaDiyah
Ibrahim-nom Neg go-3sm to-attend wedding brother-his the-night the-past
'Ibrahim didn't go to attend his brother's wedding last night'

4. فريقي المفضلُ فاز ببطولة كرة القدم هذه السنة

fariiq-ii l-mufaDDal faaza   bi-buTuwlati   kurati l-qadam haaðihi l-sanah
team-my the-favourite won-3sm with-tournament ball the-foot this   the-year
'My favourite team won the football tournament this year'

5. معلّمي قدِمَ إلى بيتنا هذا المساء للقاء أبي

Mu9allim-ii qadima   ?ilaa bayti-naa haaðaa l-masaa? li-liqaa?i ?ab-ii
Teacher-my came-3sm to   house-our this   the-evening to-meet   father-my
'My teacher came to our house to meet my father this evening'

6. سافرت أسرتي لقضاء إجازة الصيف في فرنسا

saafar-at   ?usrat-ii   li-qaDaa?i ?ijaazati l-Sayfi   fii faransaa
travelled-3sf family-my to-spend holiday   the-summer in France
'My family went to spend the summer holiday in France'

7. اشترى والدي لعبة جديدة لأختي الصغيرة

?ištaraa   waalid-ii lu9batan jadiidatan li-?uxt-ii   l-SaGiirah
bought-3sm   father-my toy   new   for-sister-my the-young
'My father bought a new toy for my little sister'

8. احتفل جميع أصدقائي بنجاحهم الدراسي هذه السنة

?iHtafala   jamee9u ?aSdiqaa?-ii bi-najaaHi-him   l-diraasiyi   haaðihi l-sanah
celebrated-3sm all   friends-my with-success-their the-educational this   the-year
'All my friends celebrated their studying success this year'

9. تناول أخي إفطاره باكراً هذا الصباح

tanaawala ?ax-ii   ?ifTara-hu   baakiran haaðaa l-SabaaH
ate-3sm   brother-my breakfast-his early   this   the-morning
'My brother ate his breakfast early this morning'

10. أعدت أمي لنا وجبة لذيذة هذا المساء

?a9add-at ?umm-ii   la-naa wajbatan laðiiðatan haaðaa l-masaa?
prepared-3sf mother-my for-us meal   delicious   this   the-evening
'My mother prepared a delicious meal for us this evening'

Sentence Type 3B: Distractor Sentences (Ungrammatical)

1. أخي كتبت واجبه المدرسي بإتقانٍ تام

?ax-ii   katab-at   waajiba-hu   l-madrasiya bi-?itqaanin   taam
brother-my   wrote-3sf homework-his the-school with-perfection completely
'My brother wrote his school homework perfectly'

2. ينصحن الأطباء جميع الطلاب والطالبات بتناول وجبة الإفطار

yanSaH-na ?al-?aTTibbaa?-u   jamii9 l-Tullaabi   wa l-Taalibaati   bi-tanaawuli wajbati l-?ifTaari
advise-3pf the-doctors (3pm)-nom all the-students.pm and the-students.pf with-eating meal the-breakfast
'Doctors advise all students to have their breakfast meal'

3. الطالبةُ أجاب على جميع الأسئلة بإجاباتٍ صحيحة

?aT-Taalibat-u   ?ajaabuw   ?alaa jamii9 l-?as?ilati   bi-?ijaabaatin SaHiiHatin
the-students(3sf)-nom answered-3pm on   all   the-questions with-answers correct
'The female students answered all the questions correctly'

4. أحمدُ استعددن جيداً لدخول امتحان اللغة العربية غداً

?aHmad-u ?ista9dad-na jayyidan li-duxuwli imtiHaani l-luGati   l-9arabiyyiati Gadan
Ahmad-nom prepared-3pf well   for-entering exam   the-language the-Arabic   tomorrow
'Ahmad prepared herself well for taking the Arabic Language exam tomorrow'

5. غداً فاطمة سافرت مع أخيها للدراسة في الخارج

Gadan   faaTimat-u   saafarat   ma9a ?axii-ha   li-l-diraasati fii lxaariji
tomorrow Fatima-nom   travelled-3sf with brother-her for-the-studying on abroad
'Fatima travelled with her brother for studying abroad tomorrow'

6. في الليلة الماضية سوف يلتقي عصام أصحابه في ساحة المدرسة

fii l-laylati l-maaDiya sawfa yaltaqii   9iSaam-u   ?aSHaaba-hu fii saaHati l-madrasati
on the-night the-past   will meet-3sm   Essam-nom friends-his at yard the-school
'Last night, Essam will meet his friends in the schoolyard'

7. كثيرٌ من الطلاب زرعوا العديد من الأشجار السنة القادمة

kaTiirun mina l-Tullaabi zara9-uw   l-9adiida mina l-?ašjaari l-sanata l-qaadimah
a lot of the-students planted-3pm the-several from the-trees the-year the-coming
'A lot of students planted several trees next year'

8. غداً حضّرنا جميع مستلزمات رحلة الصيد

Gadan   HaDDar-naa jamii9a mustalzamaati riHlati l-Saydi
tomorrow prepared-1p all   stuff   trip   the-hunting
'Tomorrow, we have prepared all what we need for the hunting trip'

9. ذهب المعلم وطلاب الصف هذا الصباح أين؟

ðahaba   l-mu9allimu wa l-Tullabu   l-Saffi haaðaa l-SabaaHi ?ayna?
went-3sm   the-teacher and the-students the-class this   morning where
'Where did the teacher and his class students go this morning?'

10. تعلمت جميع حيوانات الحديقة تنظيف حظائرهن كيف؟

ta9allam-at jamii9a Hayawaanaati l-Hadiiqati tanZiifa HaZaa?iri-hinna kayfa?
learnt-3sf all   animals   the-garden cleaning barn-their   how
'How did the zoo animals learn cleaning their barns

Please note:

- This task was completed by participants in Arabic only.
- English content is provided here for the sole purpose of translating the Arabic content.
- The film shown in this task is the first clip 'at the barber shop' of an episode titled 'Hair by Mr. Bean of London'. This can be found in Mr Beans' official YouTube channel by the title at: http://www.youtube.com/show/mrbean?s=1

سياق المحادثة:

هذه محادثة بين أحمد وصالح حول ماجد. أحمد هو الأخ الأكبر لماجد، أما صالح فهو من أصدقاء ماجد المقربين إليه. أنت ستلعب دور صالح في هذه المحادثة؛ تتحدث مع أحمد حول صديقك ماجد.

Conversation's Context:

'This is a conversation between Ahmed and Salih about Majid. Ahmad is Majid's elder brother, and Salih is Majid's close friend. You will play the role of Salih in this conversation talking to Ahmad about your friend Majid.'

إرشادات أداء المهمة:

قم بالآتي لأداء هذه المهمة:

أولاً: قم بقراءة حديث أحمد.

ثانياً: اقرأ جميع الكلمات في الصناديق المتاحة، وإرشادات الاختيار فوق الصناديق.

ثالثاً: استخدم الكلمات المتاحة في الصناديق حسب الارشادات لتكوّن جملةً مفهومةً في ذهنك أولاً تناسب سياق المحادثة.

أخيراً: اسحب الكلمات تباعاً لملء الفراغات بالجملة التي كونتها.

Task's Instructions:

'Please do the following to complete this task:

First: Read Ahmed's speech.

Second: Read all the words given in the available boxes, and the instructions for

selection above each box.

Third: Use the words available in the boxes according to the given

instructions to form a sentence in your mind first that suits the

context of the conversation.

Finally: drag the words of the sentence that you have formed in your

mind successively to fill in the blanks.'

محتوى المحادثة:

The Dialogue's Content:

أحمد: مساء الخير يا صالح

Ahmad: Good evening, Salih

اختر ماشئت، إن أردت	اختر كلمة واحدة	اختر كلمة واحدة	صالح: ☐ ☐ ☐ ☐
يا أحمد	مرحباً مساء	النور الخيرات بك	

Salih: ☐ ☐ ☐ ☐	choose 1 word	choose 1 word	choose as you like
	the-light the-good-things to-you	hello evining-of	oh Ahmed

أحمد: كيف حالك اليوم؟

Ahmad: How are you today?

اختر ما شئت	اختر كلمة واحدة	اختر كلمة واحدة	؟ ☐ ☐ ☐ ☐ ☐ :صالح
يا أحمد وأنت	أنا حالي	طيبة بخير	

Salih: ☐ ☐ ☐ ☐ ☐ ?	choose 1 word	choose 1 word	choose as you like
	fine okay	I my-state	Oh-Ahmed and-you

أحمد: أنا بحالة جيدة اليوم، شكراً لسؤالك، أريد أن أسألك: هل رأى أحدٌ ماجد اليوم؟

Ahmad: I am good today, thanks for asking. I want to ask you: did anybody see Majid today?

Appendix 6: Content of the Conversation Role-Play (CR-P) Task

صالح: نعم، ☐ ☐ ☐ ☐ ☐ ☐ ☐

اختر كلمتين	اختر كلمة واحدة	اختر ما شئت
انا ماجد علي دروسه	رأيت حضر رأى	المدرسة في الصباح هذا

Salih (Conv01): Yes, ☐ ☐ ☐ ☐ ☐ ☐ ☐

choose as you like	choose 1 word	choose 2 words
The-school in The-morning this	saw-1s attended-3ms saw-3ms	I Majid Ali his-lessons

أحمد: هذا جيد، كنت قلقاً عليه لأنه لم يأتِ إلى البيت هذا المساء، هل تعلم كيف يقضي ماجد وقت فراغه عادةً؟

Ahmad: this is good, I was worried about him as he didn't come home this evening, do you know how he usually spends his free time?

صالح: نعم، عادة ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐، عند فراغه من أداء واجباته المدرسية،

اختر كلمة واحدة	اختر كلمتين	اختر ما شئت
يحب يفضل يقرأ يزور	الذهاب ماجد هو البقاء القصص أصدقائه	السوق في البيت إلى الحديقة

Salih (Conv02): Yes, usually ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐, when he finishes his school homework

choose as you like	choose 2 words	choose 1 word
the-market in the-house to the-park	the-going Majid he the-staying the-stories his-friends	like-3ms prefer-3ms read-3ms visit-3ms

ولكن ☐ ☐ ☐ ☐ ☐ ☐

اختر ماشئت	اختر كلمة واحدة	اختر كلمة واحدة
عن عنه ذلك هذا ؟	لم لماذا ماسبب	السؤال سؤالك تسأل

But ☐ ☐ ☐ ☐ ☐ ☐ ☐

choose 1 word	choose 1 word	choose as you like
the-question your-question ask-2ms	why why what-reason	about about-him that this ?

أحمد: أعلم أنك صديق ماجد المفضل، وأريد أن أعرف المزيد عنه حتى أعرف مكانه، هل تعلم إن كانت لديه مشكلة هذه الأيام؟ أو إن كان يشتكي من شيء ما؟

Ahmad : I know that you are a close friend of Majid, and I want to know more about him to find out where he is, do you know if he has a problem these days, or if he is complaining about something?

صالح: نعم، ☐ ☐ ☐ ☐ ☐ ☐

اختر ما شئت	اختر كلمتين	اختر كلمة واحدة
البيت في ا كثيراً	الإزعاج هو أن يبقى لوحده ماجد	يحب يشتكي

Salih (Conv03): Yes, ☐ ☐ ☐ ☐ ☐ ☐

choose 1 word	choose 2 words	choose as you like
like-3ms complain-3ms	the-noise he to-be-alone Majid	the-house in a lot

أحمد: هذا مفيد، فأنا فهمت لم لا يجلس في البيت كثيراً، والآن أريد أن أجد حلاً لمشكلة ماجد، هل تريد أن تساعدني في ذلك؟

Ahmad: this helps, I understand why he doesn't stay home, and now I'd like to solve Majid's problem, would you like to help me on this?

اختر كلمة واحدة   اختر كلمة واحدة   اختر واحدة إن أردت

صالح: ☐ ☐ ☐

أريد سأساعدك	طبعاً بالطبع بالتأكيد	ذلك المساعدة

Salih: ☐ ☐ ☐

choose one, if you want to   choose 1 word   choose 1 word

that helping	sure definitely certainly	want-1s (I)-will-help-you

أحمد: عظيم، إذن أريدك أن تخبرني بالمزيد عن ماجد لأتعرف أكثر على شخصيته واهتماماته، هل كنت معه بالأمس؟

Ahmad: Great, then I'd like you to tell me more about Majid so that I know more about his personality and interests. Were you together yesterday?

صالح: نعم، ☐ ☐ ☐ ☐ ☐ ☐ ☐

اختر كلمتين   اختر كلمة واحدة   اختر ماشئت إن أردت

كرة القدم نحن أنا وماجد طعام العشاء	تناولنا لعبنا	سوياً الماضية معاً الليلة

Salih (Conv04): Yes, ☐ ☐ ☐ ☐ ☐ ☐ ☐

choose as you like, if you want to   choose 1 word   choose 2 words

together the-past together this-night	ate-1p played-1p	football we Majid-and-I the-dinner

أحمد: ممتاز، هل تحدث إلى أحدٍ عندما كنتما سوياً؟

Ahmad: excellent, did he talk to anybody when you were together?

Appendix 6: Content of the Conversation Role-Play (CR-P) Task

صالح: نعم، [ ] [ ] [ ] [ ] [ ] [ ]

اختر ماشئت	اختر كلمتين	اختر كلمة واحدة
ونحن كنا عندما معاً	جاركم سلمان ماجد صديقك عمر	هاتف التقى

Salih (Conv05): Yes, [ ] [ ] [ ] [ ] [ ] [ ]

choose as you like	choose 2 words	choose 1 word
as-we (we)-were when together	your-neighbour Salman Majid your-friend Omar	phoned-3ms met-3ms

أحمد: وماذا دار بينهما من حديث؟

Ahmad: and what did they say?

صالح: في الحقيقة، [ ] [ ] [ ] [ ] [ ] [ ]

اختر ماشئت	اختر كلمة واحدة	اختر كلمة واحدة
شيئا حديثهما شيء؛ حوارهما	لا أتذكر لم أستمع لم أسمع	إلى من

Salih: actually, [ ] [ ] [ ] [ ] [ ] [ ]

choose 1 word	choose 1 word	choose as you like
to from	(I)-don't-remember (I)-didn't-listen (I)-didn't-listen	a-thing-acc. their-talk a-thing-dat. their-dialogue

أحمد: حسناً، هل أخبرك ماجد عن أي شيء سيفعله اليوم؟

Ahmad: well, did Majid tell you about anything he is planning to do tody?

صالح: لا، لم يفعل؛ بل تذكرت، [ ] [ ] [ ] [ ] [ ]

اختر ماشئت	اختر كلمة واحدة	اختر كلمتين
اليوم المكتبة في	سيشتري سيكتب	هو كتبه ماجد واجبه المدرسي

Salih (Conv06): no, he didn't; oh I remembered now, ☐ ☐ ☐ ☐ ☐

choose 2 words	choose 1 word	choose as you like
he his-books Majid his-homework	will-buy-3ms will-write-3ms	today the-library in

أحمد: هذا جيد، الآن استطعنا تحديد مكان محتمل لوجوده، ولكن أريد معرفة المزيد لو سمحت.

Ahmad: this is good, now we know a possible place where he could be, but I want to know more if you don't mind.

صالح: ولكن، ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐! فنحن تقريباً عرفنا أين يوجد الآن.

اختر ماشئت	اختر كلمة واحدة	اختر كلمة واحدة
المزيد معرفة بعد من الأسئلة ؟	ألم لم لماذا	ننتهي تريد

Salih: But, ☐ ☐ ☐ ☐ ☐ ☐ ☐ ! It seems that we know now where he is.

choose 1 word	choose 1 word	choose as you like
finish-1p want-2ms	didn't-(we) why why	the-more knowing yet from the-questions ?

أحمد: لتحديد احتمالات أخرى، فقد لا يكون في ذلك المكان الذي حددنا قبل قليل، والآن أخبرني... هل عملتما أيَّ شيء آخر الليلة الماضية؟

Ahmad: to determine other possibilities; he might not be necessarily at that place you have mentioned, now tell me...did you two do anything else last night?

صالح: نعم، ☐ ☐ ☐ ☐

اختر كلمة واحدة	اختر كلمة واحدة	اختر كلمتين
مشوقة رائعاً مثيراً حزينة	قرأنا شاهدنا	نحن فلماً قصة أنا وماجد

Appendix 6: Content of the Conversation Role-Play (CR-P) Task

Salih (Conv07): Yes, ☐ ☐ ☐ ☐

choose 2 words	choose 1 word	choose 1 word
we film-acc story-acc Majid-and-I	read-past-1p watched-1p	interesting terrific thrilling sad

ولكن، ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

اختر ما شئت	اختر كلمة واحدة	اختر كلمة واحدة
هذا من في شيء مشكلة ماجد ؟	هل ما كيف	سيفيد الفائدة

but, ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

choose 1 word	choose 1 word	choose as you like
will-benefit-3ms the-benifit	Q-word what how	this from in a-thing problem Majid ?

أحمد: لا تدري، أي معلومة قد تدلنا على شيء مفيد، والآن هل تتذكر الأحداث في القصة؟ عن ماذا تتحدث؟

Ahmad: you never know, anything may lead to something useful, now, do you remember what the story was about?

صالح: بالطبع، ☐ ☐ ☐ ☐

اختر كلمتين	اختر كلمة واحدة	اختر كلمة واحدة
أحداثاً القصة الأحداث خيالاً	تحكي تحبك	حزيناً علمية حزينة علمياً

Salih (Conv08): definitely, ☐ ☐ ☐ ☐

choose 1 word	choose 1 word	choose 2 words
sad-3ms scientific-3fs sad-3fs scientific-3ms	narrate-3fs weave-3fs	events-acc the-story the-events imagination-acc

283

Appendix 6: Content of the Conversation Role-Play (CR-P) Task

أحمد: جيد، وهل هذا النوع من القصص هو ما تحبان عادةً؟

Ahmad: good, is this the type of stories that you usually like?

صالح: [ ] [ ] [ ] [ ] [ ] [ ] !

اختر كلمة واحدة	اختر كلمة واحدة	اختر ماشئت
ما تقصد غريب	سؤالك لا أعلم	الشيء هذا بعض بسؤالك

Salih: [ ] [ ] [ ] [ ] [ ] [ ] !

choose as you like	choose 1 word	choose 1 word
the-thing this some with-your-question	your-question (I)-don't-know	what-you-mean strange

أحمد: قد يكون لهذا علاقة بما يشعر به ماجد وعدم بقاءه في البيت مؤخراً.

Ahmad: this may have something to do with how Majid feels these days and, hence, not staying home that much lately.

صالح: [ ] [ ] [ ] [ ] [ ] [ ] [ ]

اختر كلمة واحدة	اختر ماشئت	اختر كلمة واحدة
كيف فسّر	لو فضلك من سمحت ؟	ذلك اكثر

Salih: [ ] [ ] [ ] [ ] [ ] [ ] [ ]

choose 1 word	choose as you like	choose 1 word
that more	if your-coutesy from allow-past-2ms ?	how explain

284

أحمد: قد يكون لمحتوى القصص تأثيرٌ في حالة ماجد وعدم تناسب أجواء البيت معها.

Ahmad: the content of the story may have an impact on Majid's state, and which keeps him out of home.

صالح: نعم، ☐ ☐ ☐. أجل، نحب أن نقرأ أو نشاهد هذا النوع من القصص.

اختر كلمة واحدة	اختر كلمة واحدة	اختر كلمة واحدة
الآن بهذا	فهمت أفهم اتضح	قصدك ماتقصد السبب

Salih: Yes, ☐ ☐ ☐. Yes, we like to read or watch this type of stories.

choose 1 word	choose 1 word	choose 1 word
your-aim what-you-mean the-reason	understood-1s understand-1s became-clear-3ms	now with-this

أحمد: هل تريد أن تخبرني شيئاً آخر عن ماجد؟

Ahmad: do you want to tell me anything else about Majid?

صالح: نعم، قد يكون ما تفكر به صحيحاً.

Salih: Yes, what you are thinking might be correct.

أحمد: ماذا تقصد، لم تقول هذا؟

Ahmad: what do you mean? Why do you say this?

صالح: ☐ ☐ ☐ ☐ ☐ ☐

اختر كلمتين	اختر ماشئت	اختر كلمة واحدة
القصائد هو الشعر ماجد	مؤخراً الأيام هذه كثيراً	يكتب يسمع

Salih (Conv09): ☐ ☐ ☐ ☐ ☐ ☐

choose 1 word	choose as you like	choose 2 words
write-3ms listen-to-3ms	lately the-days these a-lot	the-poems he the-poetry Majid

أحمد: وهل المحتوى مماثل لمحتوى القصص التي ذكرت سابقاً؟

Ahmad: and is the content similar in type to that you mentioned earlier?

صالح: تقريباً؛ لا أتذكر تماماً، بل تذكرت الآن؛ ☐ ☐ ☐ ☐

اختر كلمة واحدة	اختر كلمة واحدة	اختر كلمتين
حزينة واسعاً	تحكي تحوي يحيك	القصائد خيالاً المحتوى قصصاً

Salih (Conv10): I don't remember exactly, oh I think I remembered now;

☐ ☐ ☐ ☐

choose 2 words	choose 1 word	choose 1 word
the-poems imagination-acc the-content stories-acc	narrate-3fs contain-3fs weave-3ms	sad-fs broad-ms

أحمد: أعتقد أنه لابد من الجلوس مع ماجد والتحدث معه عن مشكلته، فدعنا نذهب لنبحث عنه، ولكن قبل ذلك أريدك أن تعمل شيئاً إن لم تمانع.

Ahmad: I think I need to sit with Majid and speak to him about his problem, so let's go to look for him, but before we do this, I want you to do something if you don't mind.

صالح: وما هو؟

Salih: what is it?

أحمد: لاحظت نسيانك المتكرر، وأريد أن أختبر ذاكرتك بمهمة بسيطة، فما رأيك؟

Ahmad: I have noticed that you forget things repeatedly, and I'd like to test your memory with a simple task, what do you think?

صالح: لا مانع لدي.

Salih: no problem.

أحمد: إذن شاهد هذا المقطع من الفلم الصامت، ثم اكتب ما حصل من أحداث بالترتيب في عشر جمل تامة، كل واحدة منها تستقل بذاتها عن الأخرى، وتستخدم فيها جميعاً اسم واحد أو أكثر من أسماء شخصيات الفلم، وهي (السيد بين، الحلاق، الزبون الشاب، الزبون المسن، أم الطفل، الطفل).

Ahmad: Okay, please watch this clip from a silent film, and then write what happened in 10 complete sentences; all the sentences should use one name or more of the characters in the film. The names are (Mr. Bean, The barber, the young costumer, the old costumer, the mother, the child).

صالح:

Salih:

1.
2.
3.
4.
5.
6.
7.
8.
9.
10.

أحمد: ممتاز ياصالح، فقد أثبت أن ذاكرتك قوية، وشكراً لحديثك معي اليوم.

Ahmad: excellent, Salih. You have proved that your memory is good, and thanks for talking to me today.

صالح: شكراً لك أنت يا أحمد، فأنا أتعلم منك أشياءً كل يوم.

Salih: thanks to you, Ahmad; I learn from you everyday.

287

Dear Participant,

Please choose from 1 to 10 to rate your answer to the following questions based on your experience of completing the two tasks of this study:

1. Did you find the pre-test training useful and helping to understand how to complete task one (AJT)?

   Not useful   Very useful
   1   2   3   4   5   6   7   8   9   10

2. Did you find the pre-test training useful and helping to understand how to complete task two (conversation role-play)?

   Not useful   Very useful
   1   2   3   4   5   6   7   8   9   10

3. Did you find written and oral instructions on how to complete task one (AJT) clear and detailed enough to understand what you should do in this task?

   Not clear   Very clear
   1   2   3   4   5   6   7   8   9   10

4. Did you find written and oral instructions on how to complete task two (conversation role-play) clear and detailed enough to understand what you should do in this task?

   Not clear   Very clear
   1   2   3   4   5   6   7   8   9   10

5. Did you find the test sentences in task one easy to understand?

   Very difficult   Very easy
   1   2   3   4   5   6   7   8   9   10

6. Was the length of the pause between sentences in task one enough for you to make a judgment?

   Very short pause   long enough
   1   2   3   4   5   6   7   8   9   10

7. How easy and clear did you find the context and the flow of the conversation in task two?

   Very difficult   Very clear
   1   2   3   4   5   6   7   8   9   10

عنوان البحث: اكتساب متحدثي اللهجات العربية للغة العربية الفصحى.

معلومات عن الدراسة للمشاركين

تشرفنا دعوتك للمشاركة في هذه الدراسة العلمية التي يقوم بها الباحث سامي بن صالح الرسيني، طالب الدكتوراه في قسم اللغة والعلوم اللغوية في جامعة يورك في بريطانيا.

الهدف من الدراسة: الهدف من هذه الدراسة هو محاولة الاستزادة من المعرفة حول أنماط استخدام اللغة العربية الفصحى التي يتحدثها العرب في مختلف البلدان العربية.

كيفية المشاركة: في هذه الدراسة، سيطلب منك فضلاً أن تشارك/تشاركي في مهمتين مختلفتين لهما علاقة باللغة العربية الفصحى. في المهمة الأولى، سترى/سترين جملاً تعرض على الشاشة، ثم سيطلب منك أن تقرر/تقرري ما إذا كانت كل واحدة منها مقبولة لديك من حيث التركيب والمعنى. سيمكنك تحديد إجابتك برسم دائرة حول الخيار المناسب من عدة خيارات متاحة في ورقة الإجابة. في المهمة الثانية، سيطلب منك أن تلعب/تلعبي دور أحد المتحادثين في محادثة تفاعلية مكتوبة. مهمتك ستقتصر على تعبئة الفراغات في المحادثة بجملٍ تامةٍ من خلال اختيار كلمات هذه الجمل من عدة صناديق متاحة. بعد ذلك، سيدعوك المتحادث الآخر لمشاهدة عرض قصير، ثم سيطلب منك أن تروي ماحدث في العرض بكتابة عشر جملٍ تامة. سيكون هنالك أمثلة واضحة تشرح كيفية إتمام المهام عندما تحضر/تحضرين للمشاركة. إتمام مهام الدراسة مع الإيضاحات لن يستغرق أكثر من ساعة واحدة.

الفائدة المرجوة من الدراسة: نتائج هذا البحث ستساهم في فهم أفضل لكيفية اكتساب اللغات، وهذا قد يقود إلى تطوير طرق تدريس وتعلم اللغات في المستقبل.

خصوصية المشارك: ستبقى جميع المعلومات الشخصية التي يذكرها المشارك/المشاركة في هذه الدراسة في سرية تامة ولن تتاح لأي طرف آخر.

المشاركة والانسحاب من المشاركة: لك كامل الحق في اختيار المشاركة من عدمها دون أية ضغوط، وإن تفضلت/تفضلتي بالمشاركة في هذه الدراسة، فلك أن تعتذر/تعتذري عن المواصلة في أي وقت من دون إبداء الأسباب أو وجود أية عواقب من أي نوع.

الاستفسارات حول الباحث أو الدراسة: إذا كان لديك أي سؤال أو استفسار عن الدراسة أو مشروع البحث عموماً، أرجو التواصل مع الباحث مباشرة: سامي بن صالح الرسيني (البريد الإلكتروني: ssa501@york.ac.uk).

Research Project: Acquisition of Modern Standard Arabic by Speakers of Different Colloquial Varieties

Information for Participants

You are invited to participate in a research project conducted by Sami Alresaini, a PhD student in the Department of Language and Linguistic Science, University of York.

- **Purpose of the project**

This research project is part of my PhD studies. The purpose of the project is to try to find out more about the patterns of usage of Modern Standard Arabic that is spoken across the Arab World.

- **Procedures**

You will be asked to participate in two tasks related to Arabic language. In the first task, you will see sentences presented on a screen, and you will be asked about each sentence whether it is acceptable to you. You will be able to indicate your judgement on an answer sheet, by circling a number on a scale. In the second task, you will be asked to take part in a play-role game. You will be asked to play the role of one of the speakers in a written conversation. Your job will be simply to fill in blanks with complete sentences by choosing words from available boxes. Then, you will be invited by the other speaker to watch a short film, and then you will be asked to narrate the story of the film in 10 complete sentences. Clear examples of how to complete the tasks will be given on the day. Completion of the two tasks, along with instructions, will take less than one hour.

- **Potential benefit of research**

The results of this research will contribute towards a better understanding of how we acquire languages. This may lead to new teaching and learning methods in the future.

- **Confidentiality**

All data obtained via your participation in these tasks will be anonymized and will not be associated to your identity. This anonymous data will be used and discussed in my PhD dissertation. If you chose to provide your e-mail address, you will receive a summary of the research output.

Any other information that may lead to your identity, like your participation consent form and your e-mail address, will be kept separately and will not be associated to your participation in the research tasks. Your e-mail address will be deleted once used to notify you with the research output and will not be shared with anybody.

- **Participation and withdrawal**

You can choose whether to participate in this project or not. If you volunteer to participate in this project, you may withdraw at any time without giving any reasons and without any consequences of any kind. If you decided to withdraw at any time, all data related to your participation will be removed from the research project.

- **Questions about the investigators or the research**

If you have any questions or concerns about the research, please feel free to contact the researcher: Sami Alresaini, ssa501@york.ac.uk

دعوة للمشاركة في دراسة علمية بعنوان "اكتساب متحدثي اللهجات العربية للغة العربية الفصحى"

معلومات عن الدراسة

ممكن أن تساعدني؟ اسمي سامي الرسيني، وأدرس في مرحلة الدكتوراه في جامعة يورك في بريطانيا. تشرفني دعوتك للمشاركة في هذه الدراسة عن اللغة العربية، هل يمكنك مساعدتي؟

ما هذه الدراسة؟ هذه الدراسة هي جزء من رسالة الدكتوراه التي أعدها. أنا أعمل هذه الدراسة لأني أريد أن أعرف أكثر عن اللغة العربية الفصحى وكيف يتحدثها الناس في الوطن العربي.

ما المطلوب من المشارك/المشاركة؟ أريدك أن تشارك/تشاركي في مهمتين، كلاهما عن اللغة العربية. في المهمة الأولى، سترى/سترين جملاً بسيطة على الشاشة ثم ستخبرني/ستخبرينني رأيك عنها؛ فقط تخبرني/تخبرينني إن وجدتها جملاً جيدة أو غير مقبولة في اللغة العربية. ستخبرني/ستخبرينني رأيك بتحديدك على الاختيار المناسب من عدة اختيارات متاحة في ورقة الإجابة. في المهمة الثانية، ستلعب/ستلعبين دور أحد المتحدثين في محادثة مكتوبة. سترى/سترين محادثة مكتوبة على الشاشة، كلام أحد متحدثيها غير متاح. مهمتك أن تتخيل/تتخيلي نفسك ذلك المتحدث وتملأ/تملأي الفراغات الموجودة بجمل تامة وذلك باختيار كلمات هذه الجمل من صناديق متاحة أيضاً. بعد الحديث قليلاً مع المتحدث الآخر، سيدعوك لمشاهدة عرض طريف ثم رواية ما حصل في العرض له بكتابة عشر جمل. سأريك كيف تتم هذه المهام بشكل أوضح عندما تأتي/تأتين للمشاركة. مشاركتك لن تستغرق أكثر من ساعة على الأرجح؟

ما الفائدة من الدراسة؟ هذه الدراسة عندما تنتهي، ستساعد لنفهم بشكل أفضل كيف نكتسب اللغات. أيضاً، من الممكن أن تساعدنا لتجربة طرق جديدة لتدريس وتعلم اللغات.

خصوصية المشارك: تأكد/تأكدي تماماً بأني لن أتعرف على هويتك من إجاباتك في مهام الدراسة، لأني لن أطلب منك كتابة اسمك على ورقة الإجابة. سأطلب منك كتابة اسمك على ورقة الموافقة على المشاركة في هذه الدراسة، لكن هذه الورقة ستكون غير مرتبطة بورقة الإجابة أبداً. إذا كنت تريد/تريدين معرفة نتائج هذه الدراسة فيما بعد، اترك/اتركي لي عنوان بريدك الإلكتروني عندما تحضر/تحضرين للمشاركة، وسأرسل لك خلاصة النتائج لاحقاً. أيضاً، سأحذف عنوان بريدك الإلكتروني بعد إرسال الرسالة مباشرة، ولن أمنحه لأي شخص آخر قبل ذلك.

هل تريد/تريدين المشاركة؟ أتمنى أن ترغب/ترغبي في المشاركة في هذه الدراسة، لكن القرار هو قرارك وحدك. إذا كنت لا ترغب/ترغبين في المشاركة، فبإمكانك أن ترفض/ترفضي. أيضاً إذا رغبت/رغبتي في المشاركة ثم غيرت قرارك فيما بعد، فبإمكانك أن تنسحب/تنسحبي في أي وقت ولن أسألك عن الأسباب. وأيضاً، إن قررت الانسحاب، فسأحذف جميع إجاباتك من بيانات الدراسة.

عندك استفسار؟ إن كنت تريد/تريدين الاستفسار أكثر عن هذه الدراسة، فبإمكانك التواصل معي عبر البريد الإلكتروني التالي: البريد الإلكتروني: <u>ssa501@york.ac.uk</u>

شكراً جزيلاً

Research Project: Acquisition of Modern Standard Arabic by Speakers
of Different Colloquial Varieties

Information Sheet for Young Participants

Can you help me?

My name is Sami Alresaini. I am a student at the University of York in
the United Kingdom. I would like to invite you to take part in this
research about Arabic language, so can you help me?

What is it all about?

This research is part of my doctorate studies. I am doing this because I
want to know more about Standard Arabic and how it is spoken generally
across the Arab world.

What would you have to do if you take part?

I want you to do two tasks for me, both related to Arabic Language. In
the first one, you will see simple sentences on the screen and tell me
what you think of them; you just need to tell me whether you find them
good sentences of Arabic or not. You do this by drawing a circle on an
answer sheet around the number that best matches your opinion. In the
second task, you take part in a role-play game. You will see a written
conversation between two speakers on the screen. The speech of one of
these two speakers is in the form of blanks. Your job is to imagine
yourself being that person and fill in the blanks with complete sentences
by choosing words from available boxes. After a little while of talking to
the other speaker, he will invite you to watch a short funny film and then
tell him what happened by writing 10 sentences. I will show you how this
works more clearly on the day. This should not take more than an hour of
your time.

What is the benefit of this?

When this research is done, it will hopefully help in having better
understanding about how we learn languages. Also, this may help to try
new ways of teaching and learning languages.

It's private and confidential

I would like to assure you that neither I nor anybody else will know who you are by your answers because I will not ask you to write your name on your answer sheets. However, you will write your name on a separate paper to tell me that you agree to take part in this research, but this will be kept separate from your answer sheets and they will not be connected to each other. Also, if you want to know about the results of this research, you can write for me your e-mail address on a separate paper and I will send you a message later on to tell you about that, and I will delete your e-mail address right away and will not give it to anybody else.

Do you want to take part?

I hope you will want to take part in this research, but it is completely your choice. If you do not want to take part, it is OK to say no. Also, if you decide to take part now and changed your mind later, you can withdraw at any time and I will not ask you why. If you decide to withdraw at anytime I will also not include your answers in my research.

Any questions?

If you would like to talk to me more about this research, or if you have questions about it, you can always contact me using the following e-mail address: ssa501@york.ac.uk.

Thank you very much

دعوة للمشاركة في دراسة علمية بعنوان "اكتساب متحدثي اللهجات العربية للغة العربية الفصحى"

تشرفنا دعوة ابنك/ابنتك للمشاركة في هذه الدراسة العلمية التي يقوم بها الباحث سامي بن صالح الرسيني، طالب الدكتوراه في قسم اللغة والعلوم اللغوية في جامعة يورك في بريطانيا.

معلومات عن الدراسة

<u>الهدف من الدراسة:</u> هذه الدراسة هي جزء من رسالة الدكتوراه التي أعدها. الهدف من هذه الدراسة هو محاولة الاستزادة من المعرفة حول أنماط استخدام اللغة العربية الفصحى التي يتحدثها العرب في مختلف البلدان العربية.

<u>كيفية المشاركة:</u> في هذه الدراسة، سيطلب من المشارك فضلاً أن يقوم بمهمتين مختلفتين لهما علاقة باللغة العربية الفصحى. في المهمة الأولى، سيرى المشارك جملاً تعرض على الشاشة، ثم سيطلب منه أن يقرر ما إذا كانت كل واحدة منها مقبولة لديه لغوياً. سيحدد المشارك إجابته برسم دائرة حول الخيار المناسب من عدة خيارات متاحة في ورقة الإجابة. في المهمة الثانية، سيطلب من المشارك أن يلعب دور أحد المتحادثين في محادثة مكتوبة. مهمة المشارك ستقتصر على تعبئة الفراغات في المحادثة بجمل تامة من خلال اختيار كلمات هذه الجمل من عدة صناديق متاحة. بعد ذلك، سيقدم المتحادث الآخر دعوة للمشارك لمشاهدة عرض قصير، ثم سيطلب منه أن يروي ماحدث في العرض بكتابة عشر جمل تامة. سيكون هنالك أمثلة واضحة تشرح كيفية إتمام المهام عندما يحضر المشارك للمشاركة. إتمام مهام الدراسة مع الإيضاحات لن يستغرق أكثر من ساعة واحدة على الأرجح، كما يمكن لثلاثة مشاركين أداء مهام الدراسة في ذات الوقت.

<u>الفائدة المرجوة من الدراسة:</u> نتائج هذا البحث ستساهم في فهم أفضل لكيفية اكتساب اللغات، وهذا قد يقود إلى اقتراح طرق جديدة لتدريس وتعلم اللغات في المستقبل.

<u>خصوصية المشارك:</u> ستبقى جميع الإجابات التي يذكرها المشارك في هذه الدراسة مرمزة لايمكن ربطها بهوية المشارك. هذه المعلومات المرمزة ستستخدم وتناقش في رسالة الدكتوراه. إن أردت الحصول على خلاصة نتائج هذه الدراسة، فاترك لي عنوان بريدك الإلكتروني لأرسلها لك لاحقاً.

أي بيانات أخرى قد تدل على هوية المشارك، مثل ورقة الموافقة على المشاركة أو عنوان البريد الإلكتروني، ستحفظ في منأى عن المعلومات المرمزة، ولن تربط بإجابات ابنك/ابنتك في مهام الدراسة. إن اخترت ترك عنوانك بريدك الإلكتروني، فسأحذفه مباشرة بمجرد إرسال الرسالة التي تحوي خلاصة الدراسة، ولن أمنحه لأي شخص آخر.

<u>المشاركة والانسحاب من المشاركة:</u> للمشارك أو ولي أمره كامل الحق في اختيار المشاركة من عدمها دون أية ضغوط، وإن تفضل المشارك بالمشاركة في هذه الدراسة، فله أن يعتذر عن المواصلة في أي وقت من دون إبداء الأسباب أو وجود أية عواقب من أي نوع. وفي حال الانسحاب من المشاركة في أي وقت، فكل البيانات المتعلقة بالمشاركة ستحذف تماماً من بيانات الدراسة.

<u>الاستفسارات حول الباحث أو الدراسة:</u> للسؤال أو الاستفسار عن الدراسة أو مشروع البحث عموماً، يمكن التواصل مع الباحث مباشرة: سامي بن صالح الرسيني (البريد الإلكتروني: ssa501@york.ac.uk).

Research Project: Acquisition of Modern Standard Arabic by Speakers of Different Colloquial Varieties

Your son/daughter is invited to participate in a research project conducted by Sami Alresaini, a PhD student in the Department of Language and Linguistic Science, University of York.

- **Purpose of the project**

This research project is part of my PhD studies. The purpose of the project is to try to find out more about the patterns of usage of Modern Standard Arabic that is spoken across the Arab World.

- **Procedures**

Your son/daughter will be asked to participate in two tasks related to Arabic language. In the first task, the participant will see sentences presented on a screen, and he/she will be asked about each sentence whether it is acceptable to him/her. The participants will be able to indicate their judgement on an answer sheet, by circling a number on a scale. In the second task, the participants will be asked to take part in a play-role game. The participants will be asked to play the role of one of the speakers in a written conversation. The participants' job will be simply to fill in blanks with complete sentences by choosing words from available boxes. Then, they will be invited by the other speaker to watch a short film, and then they will be asked to narrate the story of the film in 10 complete sentences. Clear examples of how to complete the tasks will be given on the day. Completion of the two tasks, along with instructions, will take less than one hour.

- **Potential benefit of research**

The results of this research will contribute towards a better understanding of how we acquire languages. This may lead to new teaching and learning methods in the future.

- **Confidentiality**

All data obtained via your son/daughter's participation in these tasks will be anonymized and will not be associated to his/her identity. This anonymous data will be used and discussed in my PhD dissertation. If you or your son/daughter chose to provide an e-mail address, you will receive a summary of the research output.

Any other information that may lead to participant's identity, like the participation consent form and the e-mail address, will be kept separately and will not be associated to your son/daughter's participation in the research tasks. E-mail addresses will be deleted once used to notify you with the research output and will not be shared with anybody.

- **Participation and withdrawal**

You can choose for your son/daughter whether to participate in this project or not. If you agree that your son/daughter volunteer to participate in this project, you may decide for him/her to withdraw at any time without giving any reasons and without any consequences of any kind. If you decided for your son/daughter to withdraw at any time, all data related to his/her participation will be removed from the research project.

- **Questions about the investigators or the research**

If you have any questions or concerns about the research, please feel free to contact the researcher: Sami Alresaini, ssa501@york.ac.uk

Appendix 8: D.1 Consent Form for Adult Participants (in Arabic)

<div dir="rtl">

موافقة على المشاركة في الدراسة

عنوان البحث: اكتساب متحدثي اللهجات العربية للغة العربية الفصحى

الباحث: سامي بن صالح الرسيني؛ حاصل على درجة الماجستير في العلوم اللغوية من جامعة يورك في بريطانيا، وهو الان طالبٌ في مرحلة الدكتوراه في قسم اللغة والعلوم اللغوية في الجامعة ذاتها.

أرجو الإشارة على الآتي:

1. قرأت وفهمت ورقة المعلومات الخاصة بالدراسة المذكورة أعلاه وأتيحت لي الفرصة للسؤال والاستفسار، وتمت الإجابة على استفساراتي بشكل مرضي. ☐

2. أعلم أن مشاركتي في هذه الدراسة اختيارية وأن لي حق الانسحاب متى شئت دون إبداء الأسباب. كما أعرف أيضاً أنه في حال الانسحاب من الدراسة، جميع البيانات المتعلقة بمشاركتي ستحذف من بيانات الدراسة. ☐

3. أعلم أن إجاباتي في مهام الدراسة ستكون مرمزة، وهوية المشارك لن تكون معروفة في نتائج الدراسة. كما أعرف أيضاً أن تركي لعنوان بريدي الإلكتروني هو فقط لإخطاري فيما بعد بخلاصة نتائج الدراسة، وأنه سيتم حذفه مباشرة بمجرد الانتهاء من ذلك. ☐

4. أوافق على مشاركتي في هذه الدراسة. ☐

التوقيع	التاريخ	اسم المشارك
_____	_____	_____

سامي بن صالح الرسيني

التوقيع	التاريخ	الباحث
_____	_____	_____

</div>

Consent to Participate in Research

Working title of Project:

Acquisition of Modern Standard Arabic by Speakers of Different Colloquial Arabic Varieties

Investigator: Sami Alresaini, MA; PhD student, Department of Language and Linguistic Science, University of York

Please initial box

1. I confirm that I have read and understand the information sheet for the above study and have had the opportunity to ask questions. Also, I confirm that my questions have been answered satisfactorily by the researcher. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason. Also, I understand that if I decided to withdraw at any time, all data that I provided will be removed from the research project. ☐

3. I understand that my answers to the tasks of this research will be anonymous and I am not going to be identifiable as an individual in any of the research project outputs. Also, I understand that providing my e-mail address is for the sole purpose that I will be contacted to be given a summary of the research output and it will be deleted once used for this purpose. ☐

4. I agree to take part in the above study. ☐

Name of Participant   Date   Signature

Sami Alresaini   _____   _____

Researcher   Date   Signature

Appendix 8: E.1 Consent Form for Young Participants (in Arabic)

<div dir="rtl">

<u>موافقة على المشاركة في الدراسة</u>

<u>عنوان البحث:</u> اكتساب متحدثي اللهجات العربية للغة العربية الفصحى

<u>الباحث:</u> سامي بن صالح الرسيني؛ حاصل على درجة الماجستير في العلوم اللغوية من جامعة يورك في بريطانيا، وهو الان طالبٌ في مرحلة الدكتوراه في قسم اللغة والعلوم اللغوية في الجامعة ذاتها.

أرجو الإشارة على الإجابة المناسبة:

1. هل قرأت، أو أحدٌ آخر قرأ لك ورقة المعلومات حول هذه الدراسة؟   نعم ☐ لا ☐

2. هل فهمت موضوع الدراسة وماتعنيه المشاركة فيها؟   نعم ☐ لا ☐

3. هل منحت لك الفرصة للسؤال عن الدراسة، وهل الباحث أجاب على أسئلتك بشكل جيد؟ نعم ☐ لا ☐

4. هل تعلم أنه بإمكانك أن ترفض المشاركة أو تنسحب من الدراسة في أي وقت دون إبداء السبب؟

نعم ☐ لا ☐

5. هل تعلم أن جميع إجاباتك في مهام الدراسة لن تربط باسمك حتى لا يعرف أحد إجاباتك؟ نعم ☐ لا ☐

6. هل تعلم أنك إذا رغبت بترك عنوان بريدك الإلكتروني، فسيستخدم فقط لإخبارك بحصيلة الدراسة لاحقاً وسيحذف بعدها مباشرة ولن يعطى لأي شخص آخر؟   نعم ☐ لا ☐

7. هل توافق على المشاركة في هذه الدراسة العلمية؟   نعم ☐ لا ☐

التوقيع	التاريخ	اسم المشارك
_____	_____	_____

إذا كان عمرك 17 سنة أو أقل، فأرجوا أن يقرأ ولي أمرك ورقة معلومات الدراسة المرفقة وخطاب الموافقة على المشاركة.

سامي بن صالح الرسيني

التوقيع	التاريخ	الباحث
_____	_____	_____

</div>

<u>Consent Form for Young People</u>

Working title of Project:

Acquisition of Modern Standard Arabic by Speakers of Different Colloquial Arabic Varieties

<u>Investigator:</u> Sami Alresaini, MA; PhD student, Department of Language and Linguistic Science, University of York

Have you read, or has someone read to you, the 'Information Sheet for Young People' about this research? Yes ☐ No ☐

Do you understand what the research is about and what taking part involves? Yes ☐ No ☐

Did you have a chance to ask questions about this research and did the researcher give you clear answers to your questions? Yes ☐ No ☐

Do you know that you can choose not to take part, and that if you decide to take part and later changed your mind, you can withdraw at anytime without giving reasons? Yes ☐ No ☐

Do you understand that all your answers to the tasks of this research will not be connected with your name, so that no one will be able to know who gave them? Yes ☐ No ☐

Do you know that if you chose to provide your e-mail address, it will be used only to inform you about the results of this research and will be deleted once used and will not be given to anybody else?

Yes ☐ No ☐

Would you like to take part in this research project? Yes ☐ No ☐

Please write your name here _____

Please Sign your name here _____ Date _____

If you are aged 17 or less, a parent/guardian also needs to read the attached sheet and give consent

Name of the researcher: Sami Alresaini

Signature _____ Date _____

Appendix 8: F.1 Consent Form for Parents/Guardians of Young Participants (in Arabic)

<div dir="rtl">

موافقة ولي الأمر على المشاركة في الدراسة

عنوان البحث: اكتساب متحدثي اللهجات العربية للغة العربية الفصحى

الباحث: سامي بن صالح الرسيني؛ حاصل على درجة الماجستير في العلوم اللغوية من جامعة يورك في بريطانيا، وهو الآن طالبٌ في مرحلة الدكتوراه في قسم اللغة والعلوم اللغوية في الجامعة ذاتها.

أرجو الإشارة على الآتي:

1. قرأت وفهمت ورقة المعلومات الخاصة بالدراسة المذكورة أعلاه وأتيحت لي الفرصة للسؤال والاستفسار، وتمت الإجابة على استفساراتي بشكل مرضي. ☐

2. أعلم أن مشاركة ابني/ابنتي في هذه الدراسة اختيارية وأن لنا حق الانسحاب متى شئنا دون إبداء الأسباب. كما أعرف أيضاً أنه في حال الانسحاب من الدراسة، جميع البيانات المتعلقة بمشاركة ابني/ابنتي ستحذف من بيانات الدراسة. ☐

3. أعلم أن إجابات ابني/ابنتي في مهام الدراسة ستكون مرمزة، وهوية المشارك لن تكون معروفة في نتائج الدراسة. كما أعرف أيضاً أن تركي لعنوان بريدي الإلكتروني هو فقط لإخطاري فيما بعد بخلاصة نتائج الدراسة، وأنه سيتم حذفه مباشرة بمجرد الانتهاء من ذلك. ☐

4. أوافق على مشاركة ابني/ابنتي في هذه الدراسة. ☐

التوقيع	التاريخ	اسم ولي أمر المشارك
_____	_____	_____

سامي بن صالح الرسيني

التوقيع	التاريخ	الباحث
_____	_____	_____

</div>

Consent to Participate in Research

Working title of Project:

Acquisition of Modern Standard Arabic by Speakers of Different Colloquial Arabic Varieties

Investigator: Sami Alresaini, MA; PhD student, Department of Language and Linguistic Science, University of York

Please initial box

1. I confirm that I have read and understand the information sheet for the above study and have had the opportunity to ask questions. Also, I confirm that my questions have been answered satisfactorily by the researcher. ☐

2. I understand that my son/daughter's participation is voluntary and that I am free to decide for them to withdraw at any time, without giving any reason. Also, I understand that if I decided for my son/daughter to withdraw at any time, all data that he/she provided will be removed from the research project. ☐

3. I understand that my son/daughter's answers to the tasks of this research will be anonymous and he/she will not be identifiable as an individual in any of the research project outputs. Also, I understand that providing any e-mail address is for the sole purpose that we will be contacted to be given a summary of the research output and they will be deleted once used for this purpose. ☐

4. I agree to that my son/daughter takes part in the above study. ☐

_____   _____   _____
Name of Participant's Parent/Guardian   Date   Signature

Sami Alresaini
_____   _____   _____
Researcher   Date   Signature

<div dir="rtl">

المعلومات الشخصية

1. تاريخ الميلاد ..../ ..... / .... م/هـ

2. الجنس: ☐ ذكر   ☐ أنثى

3. ماهي لغتك/لغاتك الأم (الأصلية)؟ ....

4. ماهي اللهجة التي تتحدثها/تتحدثينها في البيت؟
☐ العربية المصرية   ☐ العربية الخليجية   ☐ العربية السورية   ☐ لهجة عربية أخرى ....

5. كم كان عمرك عندما ذهبت إلى المدرسة لأول مرة (بما فيها الروضة)؟
☐ قبل ثلاث سنوات   ☐ ثلاث سنوات   ☐ أربع سنوات   ☐ خمس سنوات   ☐ ست سنوات (المدرسة الابتدائية)

6. ماهو مستواك التعليمي الآن؟   ☐ الإبتدائي   ☐ المتوسط   ☐ الثانوي   ☐ الجامعي   ☐ دراسات عليا

7. أرجوا الإشارة على كل ما ينطبق عليك من الآتي:
عندما كنت طفلا/طفلة:
☐ ذهبت إلى روضةٍ يتحدث الأطفال فيها باللغة العربية الفصحى فقط طوال اليوم المدرسي.
☐ ذهبت إلى روضةٍ يتحدث الأطفال فيها باللهجة المحلية.
☐ لم أذهب إلى الروضة قبل المدرسة الابتدائية.
☐ ذهبت إلى مدرسةٍ ابتدائيةٍ يتحدث الأطفال فيها باللغة العربية الفصحى فقط طوال اليوم المدرسي.
☐ ذهبت إلى مدرسةٍ ابتدائيةٍ يتحدث الأطفال فيها باللهجة المحلية.

8. إذا كنت ذهبت إلى روضةٍ يتحدث الأطفال فيها بالفصحى فقط،
اذكر من فضلك اسم الروضة هنا ....

9. إذا كنت ذهبت إلى مدرسةٍ ابتدائيةٍ يتحدث الأطفال فيها بالفصحى فقط،
اذكر من فضلك اسم المدرسة هنا ....

</div>

Appendix 9: B. Participants Personal Information Form (English Version)

Personal Details

1. Date of birth (dd/mm/yyyy):..........................

2. Gender: M   F

3. What is (are) your native language(s)? ....

4. What variety of Arabic do you speak at home?

   ☐ Egyptian Arabic   ☐ Gulf Arabic   ☐ Syrian Arabic   ☐ Other ....

5. How old were you when you started going to school (including kindergarten)?

   ☐ Before 3 years ☐ 3 years ☐ 4 years ☐ 5 years ☐ 6 years (primary school)

6. What is your current level of education?

☐ Elementary ☐ Intermediate (G 7-9) ☐ High school (G10-12) ☐ Undergraduate

☐ postgraduate

7. Please tick any of the following that apply to you:

   When I was a child:

   ☐ I went to a kindergarten in which children speak Standard Arabic only.

   ☐ I went to a kindergarten in which children speak the local colloquial variety of Arabic.

   ☐ I did not go to a kindergarten before primary school.

   ☐ I went to a primary school in which children speak Standard Arabic only.

   ☐ I went to a primary school in which children speak the local colloquial variety of Arabic.

8. If you went to a kindergarten in which children speak Standard Arabic only, what is the name of that school? ....

9. If you went to a primary school in which children speak Standard Arabic only, what is the name of that school? ....

Abbreviations

AJT	Acceptability Judgment Task
AoA	Age of Arrival
AoE	Age of first Exposure
ASL	American Sign Language
CA	Colloquial Arabic
CEA	Colloquial Egyptian Arabic
CGA	Colloquial Gulf Arabic
CLA	Colloquial Levantine Arabic
CR-P	Conversation Role-Play task
CPH	Critical Period Hypothesis
CV	Colloquial Variety
E-CEA	speakers of Colloquial Egyptian Arabic with Early exposure
E-CGA	speakers of Colloquial Gulf Arabic with Early exposure
E-CLA	speakers of Colloquial Levantine Arabic with Early exposure
L1	first language
L1A	first language acquisition
L2	second language
L2A	second language acquisition
LAD	the language acquisition device
L-CEA	speakers of Colloquial Egyptian Arabic with Late exposure
L-CGA	speakers of Colloquial Gulf Arabic with Late exposure

Abbreviations

L-CLA	speakers of Colloquial Levantine Arabic with Late exposure
LDC	Linguistic Data Consortium
MSA	Modern Standard Arabic
MSA-IP	Modern Standard Arabic Immersion Program
SA	Standard Arabic
SEA	Standard Arabic spoken in Egyptian region
SGA	Standard Arabic spoken in Gulf region
SLA	Standard Arabic spoken in Levantine region
UG	Universal Grammar

Bibliography

Akkal, A. (1993). *Syntactice Derivation of the Inflections of the Verb in Moroccan Arabic.* PhD Dissertation, University of Hassan II, Morocco.

Al-Batal, M. (1992). Diglossia Proc iency: The Need for an Alternav e Approach to Teaching. In A. Rouchdy (ed.), *The Arabic Language in America.* Detroit, MI: Wayne State Press. 284-304.

Al-Dannan, A. (2010). *Nathariatu Taaliimi Lughati Laarabiyyati Lfushaa Bilfitrati Wa Lmumaarasati: Tatbiiguhaa Wa Tagwiimuhaa Wa Ntishaaruhaa [Theory of Teaching Standard Arabic Innately and by Practice: Application, Evaluation and Its Spread].* Damascus: Dar Albashaer.

Al-Shurafa, N. (2005). Word Order, Funco ns, and Morphosyntacc Features of Adverbs and Adverbials in Arabic. *Journal of King Saud University (Arts)* 18(2): 85-100.

Alomari, M. (2009). *The Effect of the Use of the Standard Arabic Language on the Direct and Delayed Achievement of the First Three Graders in Dammam Schools in Saudi Arabia.* MA, Alyarmouk University.

Alresaini, S. (2007). *Qassimi Arabic Relative Clauses and Optionality of Resumption.* MA [not published], University of York, UK.

Altoma, S. J. (1969). *The Problem of Diglossia in Arabic: A Comparative Study of Classical Arabic and Iraqi Arabic.* Cambridge, MA: Harvard University Press.

Aoun, J., E. Benmamoun and L. Choueiri (2010). *The Syntax of Arabic.* Cambridge: Cambridge University Press.

Aoun, J., E. Benmamoun and D. Sporc he (1994). Agreement, Word Order, and Conjunco n in Some Varieties of Arabic. *Linguistic Inquiry* 25(2): 195-220.

Appen Pty Ltd, S., Australia (2006). *Gulf Arabic Conversational Telephone Speech, Transcripts.* Philadelphia: Linguistic Data Consortium.

Appen Pty Ltd, S., Australia (2007). *Levantine Arabic Conversational Telephone Speech, Transcripts.* Philadelphia: Linguistic Data Consortium.

Ayari, S. (1996). Diglossia and Illiteracy in the Arab World. *Language, Culture and Curriculum* 9(3): 243-253.

Badawi, E., M. G. Carter and A. Gully (2004). *Modern Written Arabic: A Comprehensive Grammar.* London: Routledge.

Bakir, M. J. (1980). *Aspects of Clause Structure in Arabic: A Study in Word Order Variation in Literary Arabic.* Bloomington: Indiana Univ. Linguistics Club.

Bialystok, E. and K. Hakuta (1999). Confounded Age: Linguisc and Cogniv e Factors in Age Difference for Second Language Acquisition. In D. Birdsong (ed.), *Second Language Acquisition and the Critical Period Hypothesis.* Mahwah, NJ: Erlbaum. 161-181.

Birdsong, D. (1999). The Introduco n: Whys and Why Nots of the Cric al Period Hypothesis for Second Language Acquisition. In D. Birdsong (ed.), *Second Language Acquisition and the Critical Period Hypothesis.* 2-22.

Birdsong, D. and M. Molis (2001). On the Evidence for Maturao nal Constraints in Second Language Acquisition. *Journal of Memory and Language* 44: 235-249.

Bley-Vroman, R. (1989). What Is the Logical Problem of Foreign Language Learning. In S. Gass and J. Schachter (eds.) *Linguistic Perspectives on Second Language Acquisition.* Cambridge, UK: Cambridge University Press. 41-68.

Bley-Vroman, R. (1990). The Logical Problem of Foreign Language Learning. *Linguistic Analysis* 20: 3-49.

Bley-Vroman, R., S. Felix and G. Ioup (1988). The Accessibility of Universal Grammar in Adult Language Learning. *Second Language Research* 4: 1-32.

Bohnacker, U. (2006). When Swedes Begin to Learn German: From V2 to V2. *Second Language Research* **22**: 443-486.

Bolon , N. (1995). Arabic and Parametric VSO Agreement. In M. Eid (ed.), *Perspectives on Arabic Linguistics VII*. Amsterdam, Philadelphia: John Benjamins Publishing Company. 7-27.

Bongaerts, T., C. V. Van Summeren, B. Planken and E. Schils (1997). Age and Ulm ate Attainment in the Pronunciation of a Foreign Language. *Studies in Second Language Acquisition* **19**: 447-465.

Borer, H. (1984). Restricv e Relav e Clauses in Modern Hebrew. *Natural Language and Linguistic Theory* **2**: 219-260.

Brustad, K. (1991). *The Comparative Syntax of Four Arabic Dialects: An Investigation of Selected Topics*. PhD Dissertation, Harvard University.

Brustad, K. (2000). *The Syntax of Spoken Arabic: A Comparative Study of Moroccan, Egyptian, Syrian, and Kwaiti Dialects*. Washington, D.C.: Georgetown University Press.

Chafe, W. L. (1976). Givennness, Contrasv eness, Den iteness, Subjects, Topics, and Point of View. In N. L. Charles (ed.), *Subject and Topic*. New York: Academic Press. 25-55.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1981). *Lectures in Government and Binding*. Dordrecht: Foris.

Chomsky, N. (1988). *Language and Problems of Knowledge*. Camridge, MA: MIT Press.

Colombo, J. (1982). The Cric al Period Concept: Research, Methodology, and Theorec al Issues. *Psychological Bulletin* **91**: 260-275.

Cowan, R. and Y. Hatasa (1994). Invesg an g the Validity and Reliability of Nav e Speaker and Second-Language Learner Judgments About Sentences. In E. Tarone, S. Gass and A. Cohen (eds.) *Research Methodology in Second Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates. 287-302.

Cowell, M. (1964). *A Reference Grammar of Syrian Arabic*. Washington DC: Georgetown University Press.

Cranshaw, A. (1997). *A Study of Anglophone Native and Near-Native Linguistic and Metalinguistic Performance*. PhD Dissertation, Universite de Montreal.

Curs s, S. R. (1977). *Genie: A Linguistic Study of a Modern Day "Wild Child"*. New York: Academic Press.

Cuvalay-Haak, M. (1997). *The Verb in Literary and Colloquial Arabic*. Berlin: Mouton de Gruyter.

Dahlgren, S.-O. (1998). *Word Order in Arabic*. Gothenburg, Sweden: Acta Universitatis Gothoburgensis.

DeKeyser, R. (2000). The Robustness of Cric al Period Ee cts in Second Language Acquisition. *Studies in Second Language Acquisition* **22**: 499-534.

DeKeyser, R. (2010). Cross-Linguistic Evidence for the Nature of Age Effects in Second Language Acquisition. *Applied Psycholinguistics* **31**: 413-438.

DeKeyser, R. and J. Larsen-Hall (2005). What Does the Cric al Period Really Mean? In J. E. Kroll and A. M. B. De Groot (eds.) *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford: Oxford University Press.

Dik, S. C. (1978). *Functional Grammar*. Amsterdam: North-Holland.

Doake, D. (1989). Learning to Be Literate in Arabic: Needs and New Direco ns. *Human Resource Developments in the Middle East and North Africa* **65**: 2-9.

Doron, E. (1982). On the Syntax and Semanc s of Resumpv e Pronouns. In R. Bley-Vroman (ed.), *Texas Linguisc Forum 19*. 1-48.

Dulay, H. and M. Burt (1974). You Can't Learn without Goon g. In J. Richards (ed.), *Error Analysis: Perspectives on Second Language Acquisition*. London: Longman Group Ltd.

Bibliography

Ellis, R. (2008). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Eubank, L. (1994). Towards an Explanao n for the Late Acquisio n of Agreement in L2 English. *Second Language Research* **10**: 84-93.

Eubank, L., J. Bischof, A. Hus tutler, P. Leek and C. West (1997). 'Tom Eats Slowly Cooked Eggs': Thematic-Verb Raising in L2 Knowledge. *Language Acquisition* **6**: 171-199.

Fassi-Fehri, A. (1993). *Issues in the Structure of Arabic Clauses and Words*. Dordrecht: Kluwer.

Fassi-Fehri, A. (1998). Layers in the Distribuo n of Arabic Adverbs and Adjecv es and Their Licensing. In E. Benmamoun, M. Eid and N. Haeri (eds.) *Perspectives on Arabic Linguistics XI*. Amsterdam: John Benjamins Publishing Co.

Feitelson, D., Z. Goldstein, J. Iraqi and D. Share (1993). Ee cts of Listening to Story Reading on Aspects of Literacy Acquisition in a Diglossic Situation. *Reading Research Quarterly* **28**(1): 70-79.

Felix, S. (1985). More Evidence on Compen g Cogniv e Systems. *Second Language Research* **1**: 47-72.

Ferguson, C. (1959). Diglossia. *Word* **15**: 325-340.

Flege, J. E. (1999). Age of Learning and Second Language Speech. In D. Birdsong (ed.), *Second Language Acquisition and the Critical Period Hypothesis*. Mahwah, NJ: Erlbaum. 101-131.

Flynn, S. (1996). A Parameter Sen g Approach to Second Language Acquisio n. In W. Ritchie and T. Bhatia (eds.) *Handbook of Second Language Acquisition*. San Diego: Academic Press. 121-158.

Gabriele, A. (2010). Deriving Meaning through Context: Interpren g Bare Nominals in Second Language Japanese. *Second Language Research* **26**: 379-405.

Gadalla, H., H. Kilany, H. Arram, A. Yacoub, A. El-Habashi, A. Shalaby, K. Karins, E. Rowson, R. MacIntyre, P. Kingsbury, D. Gra and C. McLemore (1997). *Callhome Egyptian Arabic Transcripts*. Philadelphia: Linguistic Data Consortium.

Galal, M. (2004). *A Minimalist Approach to Relative Clauses in Modern Standard Arabic*. PhD thesis, University of Kansas, US.

Gass, S. and A. Mackey (2007). *Data Elicitation for Second and Foreign Language Research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gil, K. H. and H. Marsden (2010). Semanc s before Syntax: L2 Knowledge of Anyone by Korean Speaking Learners. In M. Iverson, I. Ivanov, T. Judy, J. Rothamn, R. Slabakova and M. Tryzna (eds.) *Proceedings of the 2009 Mind/Context Divide Workshop*. Somerville, MA: Cascadilla Proceedings Project. 40-51.

Gil, K. H., H. Marsden and M. Whong (2011). L2 Acquisio n of Any: Negative Evidence, Negav e Implicature and Negav e L1 Transfer. In G. Granena, J. Koeth, S. Lee-Ellis, A. Lukyanchenko, G. Prieto Botana and E. Rhoades (eds.) *Selected Proceedings of the 2010 Second Language Research Forum: Reconsidering Sla Research, Dimensions, and Directions*. Somerville, MA: Cascadilla Proceedings Project. 29-39.

Gleitman, L. R. and E. L. Newport (1995). The Inveno n of Language by Children: Invironmental and Biological Influences on the Acquisition of Language. In L. R. Gleitman and M. Liberman (eds.) *An Invia o n to Cogniv e Science: Vol.1 Language* Cambridge, MA: MIT Press. 1-24.

Guas , M. (2002). *Language Acquisition: The Growth of Grammar*. Cambridge: MIT Press.

Haeri, N. (2000). Form and Ideology: Arabic Sociolinguisc s and Beyond. *Annual Review of Anthropology* **29**: 61-87.

Han, Z.-H. (1996). Fossilizao n: Can Grammac ality Judgment Be a Reliable Source of Evidence? In Z.-H. Han and T. Odlin (eds.) *Studies of Fossilization in Second Language Acquisition*. Buffalo, NY: Multilingual Matters LTD. 56-82.

Bibliography

Han, Z.-H. (2004). *Fossilization in Adult Second Language Acquisition*. Tonawanda, NY: Multilingual Matters LTD.

Hatch, E. M. (1977). Opm al Age or Opm al Learners? *Workpapers in TESL* 11: 45-56.

Hawkins, R. and C. Chan (1997). The Para l Availability of Universal Grammar in Second Language Acquisition: The 'Failed Functional Features Hypothesis'. *Second Language Research* 13(3): 187-226.

Haznedar, B. (2010). Transfer at the Syntax-Pragmatics Interface: Pronominal Subjects in Bilingual Turkish. *Second Language Research* 26: 355-378.

Hellmuth, S. (2006). *Intonational Pitch Accent Distribution in Egyptian Arabic*. Unpublished PhD Dissertation, School of Oriental and African Studies, University of London.

Holes, C. (1995). *Modern Arabic: Structures, Functions and Varieties*. New York: Longman Publishing.

Homeidi, M. (1991). Government and Binding and Word Order in Modern Standard Arabic. *Abhath Al-Yarmouk* 9.

Hornstein, N. and D. Ligho ot (1981). Introduco n. In N. Hornstein and D. Ligho ot (eds.) *Explanation in Linguistics: The Logical Problem of Language Acquisition*. London: Longman Group LTD.

Hulk, A. (1991). Parameter Sen g and the Acquisio n of Word Order in L2 French. *Second Language Research* 7: 1-34.

Hurford, J. (1994). *Grammar: A Student's Guide*. Cambridge: Cambridge University Press.

Hyltenstam, K. (1992). Non-Native Features of Near-Native Speakers. In R. J. Harris (ed.), *Cognitive Processing in Bilinguals*. Amsterdam: Elsevier.

Hyltenstam, K. and N. Abrahamsson (2003). Maturao nal Constraints on Second Language Acquisition. In C. Doughty and M. Long (eds.) *Handbook of Second Language Acquisition*. Oxford: Blackwell. 539-588.

Ibrahim, M. (1986). Standard and Presg e Language: A Problem in Arabic Sociolinguisc s. *Anthropological Linguistics* 28(1): 115-126.

Ionin, T. and M. Zubizarreta (2010). Introduco n to the Special Issue: Selecv e First Language Influence and Retreat from Negative Transfer. *Second Language Research* 26: 283-291.

Ioup, G., E. Boustagui, M. El g i and M. Moselle (1994). Reexamining the Cric al Period Hypothesis: A Case Study of Successful Adult SLA in a Naturalistic Environment *Studies in Second Language Acquisition* 16: 73-98.

Iraqi, J. (1990). *[Reading to Arabic-Speaking Kindergarten Children Compared to Alternative Enrichment Activities as a Means of Improving Reading Comprehension and Language Skills]*. MA, University of Haifa.

Jackson, K. and R. Whitnam (1971). *Evaluation of the Predictive Power of Contrastive Analysis of Japanese and English. Final Report*. Hawaii: University of Honolulu.

Jenkins, J. (2001). *The Effect of a Preschool MSA Immersion Program on Arab Children's Primary School Reading and Composition Scores*. MA, Brigham Young University.

Johnson, J. S. (1992). Cric al Period Ee cts in Second Language Acquisition: The Effect of Written Versus Auditory Materials on the Assessment of Grammatical Competence *Language Learning* 42: 217-248.

Johnson, J. S. and E. L. Newport (1989). Cric al Period Ee cts in Second Language Learning: The Influence of the Maturational State on the Acquisition of English as a Second Language. *Cognitive Psychology* 21: 60-99.

Karmiloff-Smith, A. (1986). Some Fundamental Aspects of Language Development ae r Age 5. In P. Fletcher and M. Garman (eds.) *Language Acquisition: Studies in First Language Development*. Cambridge: Cambridge University Press. 455-474.

Kaye, A. S. (1970). Modern Standard Arabic and the Colloquials. *Lingua* 24: 374-391.

Bibliography

Khamis-Dakwar, R. (2011). Early Acquisio n of SVO and VSO Order in Palestinian Arabic. In E. Broselow and H. Ouali (eds.) *Perpectives on Arabic Linguistics XXII-XXIII*. Amsterdam: John Benjamins Publishing Company. 281-292.

Khamis-Dakwar, R., K. Froud and P. Gordon (2012). Acquiring Diglossia: Mutual Inu ences of Formal and Colloquial Arabic on Children's Grammaticality Judgments. *Journal of Child Language* 39(1): 61-89.

Koopman, H. and D. Sporc he (1991). The Posio n of Subjects. *Lingua* 85: 211-258.

Kroe , Solveig, M. Rothweiler and E. Babur (2007). Turkish-German Successsive Bilingual Children: Age of Onset and the Development of German V2. Paper presented in *the 6th Internao nal Symposium on Bilingualism*, Hamburg.

Lado, R. (1957). *Linguistics across Cultures*. Ann Arbor: University of Michigan Press.

Lasnik, H. (1981). Restricn g the Theory of Transformation: A Case Study. In N. Hornstein and D. Lightfoot (eds.) *Explantion in Linguistics*. London: Longman. 152-173.

Lenneberg, E. H. (1967). *Biological Foundations of Language*. New York: Wiley.

Ma, X. and D. Zakhary (2007). *Gale Phase 1 Arabic Broadcast News Parallel Text - Part 1*. Philadelphia: Linguistic Data Consortium.

Ma, X. and D. Zakhary (2008). *Gale Phase 1 Arabic Broadcast News Parallel Text - Part 2*. Philadelphia: Linguistic Data Consortium.

Maamouri, M. (1998). *Language Education and Human Development: Arabic Diglossia and Its Impact on the Quality of Education in the Arab Region*. Discussion paper prepared for The World Bank, The Mediterranean Development Forum. Philadelphia: University of Pennsylvania, International Literacy Institute.

Mac Whinney, B. (2000). *The Childes Project: Tools for Analyzing Talk. 3rd Edio n*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mahfoudhi, A. (2002). Agreement Lost, Agreement Regained: A Minimalist Account of Word Order and Agreement Variation in Arabic. *Calfornia Linguistic Notes* 17.

Majdi, B. (1990). Word Order and Proper Government in Classical Arabic. *Current Issues in Linguistic Theory* 63: 127-153.

Mansouri, F. (2000). *Grammatical Markedness and Information Processing in the Acquisition of Arabic as a Second Language*. Munich: Lincom Europa.

Marsden, H. (2008). Pair-List Readings in Korean-Japanese, Chinese-Japanese, and English-Japanese Interlanguage. *Second Language Research* 24: 189-226.

Marsden, H. (2009). Distribuv e Quane r Scope in English-Japanese and Korean-Japanese Interlanguage. *Language Acquisition* 16(3): 135-177.

Mason, M. (1942). Learning to Speak ae r Six and One Half Years of Silence. *Journal of Speech Disorders* 7: 295-304.

McDonald, J. L. (2000). Grammac ality Judgments in a Second Language: Influence of Age of Acquisition and Native Language. *Applied Psycholinguistics* 21: 395-423.

Meier, R. P. (1995). Review of S. Pinker, the Language Insn ct: How the Mind Creats Language. *Language* 71: 610-614.

Meisel, J. (2008). Child Second Language Acquisition or Successive First Language

Acquisition? In B. Haznedar and E. Gavruseva (eds.) *Current Trends in Child Second Language Acquisition: A Generative Perspective*. Amsterdam: Benjamins. 55-80.

Meisel, J. (2009). Second Language Acquisio n in Early Childhood. *Zeitschrift für Sprachwissenschaft* 28: 5-43.

Meisel, J. (2011). *First and Second Language Acquisition: Parallels and Differences*. Cambridge, UK: Cambridge University Press.

Moawad, R. (2006). *The Acquisition of the Arabic Gender and Number Systems*. PhD Dissertation, University of Wales.

Mohammad, M. (1990). The Problem of Subject-Verb Agreement in Arabic: Towards a Solution. In M. Eid (ed.), *Perspectives on Arabic Linguistics I*. Amsterdam, Philadelphia: John Benjamins Publishing Company. 95-125.

Montrul, S. (2010). Dominant Language Transfer in Adult Second Language Learners and Heritage Speakers. *Second Language Research* **26**: 293-327.

Moutaouakil, A. (1989). *Pragmatic Functions in a Functional Grammar of Arabic*. Dordrecht: Foris Publications.

Moyer, A. (2004). *Age, Accent, and Experience in Second Language Acquisition*. Buffalo, NY: Multilingual matters LTD.

Myhill, J. (1992). Word Order and Temporal Sequencing. In D. Payne (ed.), *Pragmatics of Word Order Flexibility*. Amsterdam: Benjamins. 265-278.

Neufeld, G. (1977). Language Learning Ability in Adults: A Study on the Acquisio n of Prosodic and Articulatory Features *Working Papers on Bilingualism* **12**: 45-60.

Neufeld, G. (1979). Towards a Theory of Language Learning Ability. *Language Learning* **29**: 227-241.

Newport, E. L. (1984). Constraints on Learning: Studies in the Acquisio n of American Sign Language. *Papers and Reports on Child Language Development* **23**: 1-22.

Newport, E. L. (1990). Maturao nal Constraints on Language Learning. *Cognitive Science* **4**: 11-28.

Newport, E. L. (1991). Contrasn g Concepo ns of the Cric al Period for Language. In S. Carey and R. Gelman (eds.) *The Epigenesist of Mind: Essays on Biology and Cognition*. Hillsdale, NJ: Erlbaum.

Newport, E. L. and T. Supalla (1987). *A Critical Period Effect on the Acquisition of a Primary Language.*: University of Illinois.

Odlin, T. (1989). *Language Transfer: Cross Linguistic Influence in Language Learning*. Cambridge, UK: Cambridge University Press.

Odlin, T. (2003). Cross-Linguistic Influence. In C. Doughty and M. Long (eds.) *The Handbook of Second Language Acquisition*. Malden, MA: Blackwell. 436-486.

Oh, E. (2010). Recovery from First-Language Transfer: The Second Language Acquisition of English Double Objects by Korean Speakers *Second Language Research* **26**: 407-439.

Ouhalla, J. (1997). Remarks on Focus in Standard Arabic. In M. Eid and R. Ratclie (eds.) *Perspectives on Arabic Linguistics X*. Amsterdam, Philadelphia: John Benjamins Publishing Company. 9-45.

Ouhalla, J. and U. Shlonsky (2002). Introduction. In J. Ouhalla and U. Shlonsky (eds.) *Themes in Arabic and Hebrew Syntax*. Dordrecht: Kluwer Academic Publishers.

Owens, J., R. Dodsworth and T. Rockwood (2009). Subject-Verb Order in Spoken Arabic: Morpholexical and Event-Based Factors. *Language Variation and Change* **21**: 39-67.

Oyama, S. C. (1976). A Sensiv e Period for the Acquisio n of a Nonnav e Phonological System. *Journal of Psycholinguistic Research* **5**(3): 261-283.

Oyama, S. C. (1979). The Concept of the Sensiv e Period in Developmental Studies. *Merrill-Palmer Quarterly* **25**: 83-103.

Patkowski, M. (1980). The Sensiv e Period for the Acquisio n of Syntax in a Second Language Acquisition. *Language Learning* **30**: 449-472.

Pene ld, W. and L. Roberts (1959). *Speech and Brain Mechanisms*. New York: Atheneum Press.

Plunke , B. (1993). The Posio n of Subjects in Modern Standard Arabic. In M. Eid and C. Holes (eds.) *Perspectives on Arabic Linguistics V*. Amsterdam: John Benjamins Publishing Company. 231-260.

Ryding, K. (2005). *A Reference Grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.

Sachs, J., B. Bard and M. L. Johnson (1981). Language Learning with Restricted Input: Case Studies of Two Hearing Children of Deaf Parents. *Applied Psycholinguistics* **2**: 33-54.

Schachter, J. (1989). Tesn g a Proposed Universal. In S. Gass and L. Selinker (eds.) *Linguistic Perspectives on Second Language Acquisition.* Cambridge, UK: Cambridge University Press.

Schachter, J. (1990). On the Issue of Completeness in Second Language Acquisio n. *Second Language Research* **6**: 93-124.

Schachter, J. (1996). Maturao n and the Issue of Universal Grammar in Second Language Acquisition. In W. Ritchie and T. Bhatia (eds.) *Handbook of Second Language Acquisition.* San Diego: Academic Press. 159-193.

Schumann, J. H. (1978). Social and Psycholinguisc Factors in Second Language Acquisio n. In J. Richards (ed.), *Understanding Second and Foreign Language Acquisition.* Rowley, MA: Newbury House.

Schutze, C. (1996). *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology.* Chicago: The University of Chicago Press.

Schwartz, B. (2004). Why Child L2 Acquisio n? In J. Van Kampen and S. Baauw (eds.) *Proceedings of Generav e Approaches to Language Acquisio n 2003.* Utrecht: LOT Occasional Series. 47-66.

Schwartz, B. and L. Eubank (1996). What Is the "L2 Inia l State"? *Second Language Research* **12**: 1-6.

Schwartz, B. and R. Sprouse (1994). Word Order and Nominav e Case in Nonnav e Language Acquisio n: A Longitudinal Study of (L1 Turkish) German Interlanguage. In T. Hoekestra and B. Schwartz (eds.) *Language Acquisition Studies in Generative Grammar: Papers in Honor of Kenneth Wexler from the Glow 1991 Workshops.* Amsterdam: John Benjamins. 317-368.

Schwartz, B. and R. Sprouse (1996). L2 Cogniv e States and the Full Transfer/Full Access Model. *Second Language Research* **12**: 40-72.

Seliger, H. (1978). Implicao ns of a Mulp le Cric al Period Hypothesis for Second Language Learning In W. Ritchie (ed.), *Second Language Acquisition Research: Issues and Implications.* New York: Academic Press. 11-19.

Seliger, H., S. D. Krashen and P. Ladefoged (1975). Maturao nal Constraints in the Acquisition of Second Language Acquisition. *Language Science* **36**: 20-22.

Sells, P. (1984). *Syntax and Semantics of Resumptive Pronouns.* PhD Dissertation, University of Massachusetts at Amherst, U.S.

Shim, R. (1993). Sensiv e Periods for Second Language Acquisio n: A Reaco n-Time Study of Korean-English Bilinguals *Ideal* **6**: 43-64.

Shlonsky, U. (1992). Resumpv e Pronouns as a Last Resort. *Linguistic Inquiry* **23**: 438-468.

Shlonsky, U. (1997). *Clause Structure and Word Order in Hebrew and Arabic.* New York: Oxford University Press.

Shouby, E. (1951). The Inu ence of the Arabic Language on the Psychology of the Arabs. *Middle East Journal* **5**(3): 284-302.

Slabakova, R. (2008). *Meaning in the Second Language.* Berlin: Mouton de Gruyter.

Slobin, D. (1982). Universal and Parc ular in the Acquisio n of Language. In E. Wanner and L. R. Gleitman (eds.) *Language Acquisition: State of the Art.* Cambridge, UK: Cambridge University Press.

Slobin, D. (1996). From 'Thought and Language' to 'Thinking for Speaking'. In J. Gumperz and S. Levinson (eds.) *Rethinking and Linguistic Relativity.* Cambridge: Cambridge University Press. 70-96.

Smith, N. V. and I.-M. Tsimpli (1995). *The Mind of a Savant: Language Learning and Modularity.* Oxford: Blackwell.

Bibliography

Sorace, A. (1993). Incomplete Vs. Divergent Representao ns of Unaccusav ity in Nonnav e Grammars of Italian. *Second Language Research* 9: 22-47.

Sorace, A. (1996). The Use of Acceptability Judgments in Second Language Research. In W. Ritchie and T. Bhatia (eds.) *Handbook of Second Language Acquisition*. San Diego: Academic Press. 375-409.

Sorace, A. (2003). Near-Nativeness. In C. Doughty and M. Long (eds.) *The Handbook of Second Language Acquisition*. Oxford: Blackwell. 130-152.

Sorace, A. (2012, March). *The Bilingual Native Speaker.* Speech presented at the University of York, York, UK.

Sorace, A. and L. Serratrice (2009). Internal and External Interfaces in Bilingual Language Development: Beyond Structural Overlap. *The International Journal of Bilingualism* 13: 195-210.

Stevens, G. (1999). Age at Immigrao n and Second Language Proc iency among Foreign Born Adults. *Language in Society* 28: 555-578.

Stevens, P. (2006). Is Spanish Really So Easy? Is Arabic Really So Hard?: Perceived Dic ulty in Learning Arabic as a Second Language. In K. Wahba, Z. Taha and L. England (eds.) *Handbook for Arabic Language Teaching Professionals in the 21st Century*. New Jersey: Lawrence Erlbaum Associates. 35-63.

Suleiman, S. (1986). *Jordanian Arabic between Diglossia and Bilingualism: Linguistic Analysis*. Amsterdam & Philadelphia: John Benjamins.

Tsimpli, I.-M. and N. V. Smith (1991). Second Language Learning: Evidence from a Polyglot Savant. *UCL Working Papers in Linguistics* 3: 171-184.

Tsimpli, I.-M. and A. Sorace (2006). Die rena n g Interfaces: L2 Performance in Syntax-Semantics and Syntax-Discourse Phenomena. In *Proceedings of the 30th Boston University Conference on Language Development*, Somerville, MA: Cascadilla Press, 653-664.

Unsworth, S. (2004). Child L1, Child L2, and Adult L2 Acquisio n: Die rences and Similarities. In A. Brugos, L. Micciulla and C. E. Smith (eds.) *Proceedings of the 28th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press. 633-644.

Vainikka, A. and M. Young-Scholten (1994). Direct Access to X'-Theory: Evidence from Korean and Turkish Adults Learning German. In T. Hoekestra and B. Schwartz (eds.) *Language Acquisition Studies in Generative Grammar*. Amsterdam: John Benjamins. 265-316.

Vainikka, A. and M. Young-Scholten (1998). The Inia l State in the L2 Acquisio n of Phrase Structure. In S. Flynn, G. Martohardjono and W. O'Neil (eds.) *The Generative Study of Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates. 17-34.

Vainikka, A. M. and M. Young-Scholten (1996). The Early Stages in Adult L2 Syntax: Additional Evidence from Romance Speakers. *Second Language Research* 12: 140-176.

White, L. (2000). Second Language Acquisition: From Initial to Final State. In J. Archibald (ed.), *Second Language Acquisition and Linguistic Theory*. Oxford: Blackwell. 130-155.

White, L. (2003). *Second Language Acquisition and Universal Grammar*. Cambridge: Cambridge University Press.

White, L. (2009). Grammac al Theory: Interfaces and L2 Knowledge. In W. Ritchie and T. Bhatia (eds.) *The New Handbook of Second Language Acquisition*. Bingley, UK: Emerald Group Publishing. 49-70.

White, L. and F. Genesee (1996). How Nav e Is Near-Native?: The Issue of Ultimate Attainment in Adult Second Language Acquisition *Second Language Research* 12: 238-265.

Bibliography

Zughoul, M. R. (1980). Diglossia in Arabic: Invesg an g Soluo ns. *Anthropological Linguistics* **22**(5): 201-217.