CORPUS STUDY OF TENSE, ASPECT, AND MODALITY
IN DIGLOSSIC SPEECH IN CAIRENE ARABIC

BY

OLA AHMED MOSHREF

DISSERTATION
Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Linguistics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:
Professor Elabbas Benamoun, Chair
Professor Eyamba Bokamba
Professor Rakesh M. Bhatt
Assistant Professor Marina Terkourafi
ABSTRACT

Morpho-syntactic features of Modern Standard Arabic mix intricately with those of Egyptian Colloquial Arabic in ordinary speech. I study the lexical, phonological and syntactic features of verb phrase morphemes and constituents in different tenses, aspects, moods. A corpus of over 3000 phrases was collected from religious, political/economic and sports interviews on four Egyptian satellite TV channels. The computational analysis of the data shows that systematic and content morphemes from both varieties of Arabic combine in principled ways. Syntactic considerations play a critical role with regard to the frequency and direction of code-switching between the negative marker, subject, or complement on one hand and the verb on the other. Morph-syntactic constraints regulate different types of discourse but more formal topics may exhibit more mixing between Colloquial aspect or future markers and Standard verbs.
To the One Arab Dream that will come true inshaa' Allah!

Arab I am. My nation’s blood is the finest. As my father says

Iraqi Poet: Badr Shaker Elsayyab
ACKNOWLEDGMENTS

I’m sincerely thankful to my advisor Prof. Elabbas Benmamoun, who during the six years of my study at UIUC was always kind, caring and supportive on the personal and academic levels. I also acknowledge the support and encouragement of the kind-hearted and wise Prof. Eyamba Bokamba. The skepticism and scientific zeal of Prof. Marina Terkourafi have always gained my admiration and appreciation.

This work is also indebted to three colleagues at the department of Linguistics at UIUC. One long and very enlightening discussion with Bezza Ayalew helped me crystallize and pin down the procedure and methodology of my research. In the design and processing of the computational program, I benefited from the expertise of Tim Mahrt and Mahmoud Abu Nasser, who have ever and always been available to rescue in difficult times. Tim was also very kind to proof-read the thesis in a notably short period of time. I’m very thankful to the three of them.

As an international student away from my home country, I cherish the emotional, friendly, and family like atmosphere that my close circle of colleagues at the department of Linguistics and wider at UIUC has constantly established. I mention in particular my Arab colleagues: Noha Elsakka, Eman Saadah, Rania Al-Sabbagh, Fatemah Herms, and last but never least Abdelaadim Bidaoui, in addition to my semi-Arab friend and colleague beautiful Angela Selena Williams.
Table of Contents

Chapter I: Problem, motivation, and scope of study ................................................................. 1
  I-1 MSA and ECA ....................................................................................................................... 1
  I-2 Codeswitching in diglossic languages ................................................................................. 8
  I-3 Code-switching in distant languages ............................................................................... 11

Chapter II: Major theoretical approaches to code-switching .................................................... 14
  II-1 The Two Constraints Theory & the Government Principle ............................................ 14
  II-2 The Matrix Language Frame (MLF) hypothesis ............................................................. 20
  II-3 Optimality Theory and the typology of code-switching ............................................... 23
  II-4 The Dual Language Model (DLM) .................................................................................. 27

Chapter III: Codeswitching in Arabic ...................................................................................... 33
  III-1 Corpora and methodologies ......................................................................................... 33
  III-2 Approaches to Arabic code-switching ......................................................................... 37

Chapter IV: Tense, Aspect, Modality ....................................................................................... 57
  IV-1 The Arabic verb composition ......................................................................................... 57
  IV-2 Verb Negation in Arabic .............................................................................................. 64
  IV-3 Constituents of the Arabic verb phrase ......................................................................... 72
  IV-4 Tense and Aspect ........................................................................................................... 80
  IV-5 Modality ........................................................................................................................ 120

Chapter V: Methodology .......................................................................................................... 127
  V-1 Data collection ................................................................................................................ 127
  V-2 Classification of data ...................................................................................................... 128
V-3 Inter-phrase partitioning ................................................................. 130
V-4 Variety labeling of constituents ..................................................... 132
V-5 Procedure of data analysis ............................................................ 136

Chapter VI: Data analysis and results .................................................. 141
VI-1 Code-switching at phrase levels .................................................... 142
VI-2 Comparison by discourse topic ..................................................... 170

Chapter VII: Discussion and conclusion .............................................. 176

References .......................................................................................... 187

Appendix A: Transcription gloss .......................................................... 195
Appendix B: Abbreviations .................................................................... 196
Appendix C: CS comparison by Tense, Aspect, Modality ....................... 198
  Appendix C-1: Collective results for verb level ...................................... 198
  Appendix C-2: Collective results for subject-verb with no intervening markers ... 201
  Appendix C-3: Collective results for complement-verb with no intervening markers .... 205
  Appendix C-4: Collective results for case-marked complement-subject combinations .. 209

Appendix D: CS comparison by discourse .............................................. 211
  Appendix D-1: MSA verb stems with ECA aspectual prefix ...................... 211
  Appendix D-2: MSA Future markers in corpus........................................ 214
  Appendix D-3: MSA Negative markers in corpus....................................... 215
Chapter I: Problem, motivation, and scope of study

The structure of the verb phrase in Modern Standard Arabic (MSA) differs in some phonological, lexical, syntactic, and morphological aspects from Egyptian Colloquial Arabic (ECA). It is, thus, a good indicator of how the standard and colloquial varieties interact in spontaneous diglossic speech. The objective of this study is to explore the sentential constituents, ranging from the morphemic up to the lexical and phrasal levels, control standard-colloquial mixing. Verbal phrases include all different combinations of tense, aspect, and mood in spoken discourse. Data is collected from interviews in four Egyptian satellite TV channels and the topics of the interviews are religious, political/economic, and sport.

I-1 MSA and ECA

One of Ferguson’s nine features of diglossia is stability over centuries with consistent borrowing from the high variety H into the low variety L, resulting in the development of intermediate language forms (Ferguson, 1959:332). This is typical of the growing mix between ECA, and MSA. ECA is a vernacular that has urban and rural variations. Its tense/aspect system of affixes is regarded by some researchers as more complex than MSA (Owens, 2006:26). In this research ECA will refer to Cairene Arabic. Versteegh (2001) traces the current form of this variety to the end of the nineteenth century “when the flux of speakers from the countryside led to stigmatization of the rural dialects that has continued until today. As a result, new migrants to the capital tend to shift wherever possible to Cairene Arabic (p.197).
MSA is the modern form of Classical Arabic (CA). It attained its current position as a result of contact with western culture and consequent modernization in the Arab world since the 19th century and is now the effective formal language in education, media, literature and all government documentation in all Arab countries. In an attempt to integrate new political, technical, and scientific terms of western civilization, MSA is constantly coining and arabizing new terms, for example ديمقراطية [dīmoqrāṭīya] democracy, as well as introducing semantic shifts in classical terms. In addition, stylistic changes in sentence formation largely distinguish modern texts from classical writing styles; particularly in phraseology, syntactic calques, and prepositions (Holes, 1995:46-48). This is evident in newspaper styles that tend to translate from European languages, for example, the introduction of expressions like ما إذا [mā ?iḍa] whether, ل نهائي [lā nihāʔi] infinite, التقى مع [iṭitaqa maʔa] met with, and the extensive use of the dummy verb قام [qāma] took up and passive forms with تم [tamma] finished as in 1& 2 (Versteegh, 181:2001).

1. qām-a bi ʔamal-i igtimā-ʔin maʔa l-muʃāradg a
   take up.PRF-3sg.M by making-GEN meeting-GEN with the-opposition
   قام بعمل اجتماع مع المعارضة
   He held a meeting with the opposition

2. tam-at il-ʔamaliyyat-u bi nagāh
   finish-3sg.F the-operation-NOM by success
   تمت العملية بنجاح
   The operation was performed successfully
Regarding case markers that are totally absent in ECA, Anis (1960) maintains that they were not part of the linguistic intuition of all Arabians in classical times, but only of the literary elite. However, since later grammarians had as their reference the speech and judgments of Bedouin Arabians, these speakers could not have lacked the sound knowledge of Arabic grammar and the gap between their daily life language and the literary one must have been extremely limited. Socioeconomic and political factors have contributed to widening the gap between the two varieties in subsequent ages throughout the new Arabized territories. This is a natural outcome of the consistent effort on part of Arab linguists to preserve the classical variant, at the time when the daily life colloquial has been continuously changing (Ibrahim 1989:39-43).

It might, thus, seem that the colloquials that evolved in various Arabized regions continued to diverge away from their classical root. However, research traces many phenomena present in the colloquial back to pre-Islamic tribal dialects. One example out of many is the substitution of the prefix yi- for ya- in imperfect verbs, which is a feature in the dialect of Bahraa’ tribe known as تلتلة بهراء [taltalet bahraa?] (Abdel Tawaab 1988:264-275). Versteegh (1996) elaborates more on a similar example stating that “the pre-Islamic forms have not disappeared, but remain within the repertory of the speakers, even though nobody uses them anymore” (p.20). This means that speakers may intuitively “select” some archaic features, and neglect others. Versteegh also accounts on the concept of اتساع [ʔittisaaʔ] expansion that allowed speakers to use the language creatively without fear of violating the rules. This, as he puts it, “served to safeguard the essential stability of the language, while at the same time allowing for its adaptation to the needs of the speakers” (p.21).
The mixing of MSA and ECA was developed by educated speakers in formal and semi-formal occasions, hence called Educated Spoken Arabic (ESA). Badawi (1973) explains that ESA speakers have access to Western culture and can speak at least one foreign language, in addition to being educated in MSA and CA. Their acquaintance with both Arabic and Western cultures frees them from adherence to fixed norms and qualifies them to develop the standard language by introducing and coining new terms and expressions. Less educated speakers are far less influential in language development because of their limited access to foreign cultures and their lower status in society (pp.113-115).

There have been several attempts for characterizing the degree of the impact of MSA on the spoken language. Blanc (1960), Badawi (1973), and Meiseles (1980) identify a hierarchy of intermediate varieties. Their categorizations are compared in Table 1.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Semi-literary (elevated)</td>
<td>Colloquial</td>
<td>Oral Literary (Sub-standard) Arabic</td>
</tr>
<tr>
<td>Colloquial</td>
<td>Colloquial of the Cultured [Ŷāmiyyat al-muθaqqafīn]</td>
<td>Educated Spoken Arabic</td>
</tr>
<tr>
<td>Koineized Colloquial</td>
<td>Colloquial of the Enlightened [Ŷāmiyyat al-mutanawwīrīn]</td>
<td></td>
</tr>
<tr>
<td>Plain Colloquial</td>
<td>Colloquial of the Illiterate [Ŷāmiyyat al-ʔummiyyīn]</td>
<td>Basic/Plain vernaculars</td>
</tr>
</tbody>
</table>

Table 1 Hierarchy of spoken varieties in Blanc (1960), Badawi (1973), and Meiseles (1980)

Blanc’s categorization is linguistically based. He divides spoken forms into semi-literary colloquial, koineized colloquial, and plain colloquial. Badawi, on the other hand, proposes a
socially stratified classification and identifies three types of colloquial varieties: colloquial of the cultured (who are well educated), of the enlightened (who are partially educated), and of the illiterate. Meiseles combines the social-functional role of each variety with its linguistic features.

Highest on Meiseles’ classification is Oral Literary Arabic (OLA), a spoken counterpart of informal written Arabic. The latter is unedited writing that may violate MSA norms under the influence of the colloquial. OLA roughly corresponds to Blanc’s semi-literary Colloquial, which he describes as a koineized colloquial classicized beyond “mildly formal” (1960:85). OLA is only an approximation to the stringent descriptive rules of MSA and CA. Meiseles expresses this in Ferguson’s words as “an Arab’s attempt to speak classical Arabic” (1980:125). Functionally, OLA is used by the mass media and in formal settings. However, even in these situations, people may shift to more colloquial registers/varieties for the purpose of establishing a degree of intimacy with their interlocutors (Ferguson, 1959:235; Hary, 1996:76). Next on the spoken hierarchy is ESA, which being mildly classicized and leveled, corresponds to Blanc’s description of Koineized Colloquial. It also coincides with Badawi’s Colloquial of the Cultured and Enlightened, since it is spoken in certain registers and contexts by the cultural and societal elite.

The models of Blanc, Badawi, and Meiseles imply that the different levels of spoken Arabic fall within defined boundaries. Hary (1996) underscores the fact that these boundaries are only theoretical abstractions due to the frequent stylistic and functional shifts in the spoken discourse of Arabs. He borrows the terms acrolect and basilect to designate CA and the colloquial at the two extreme ends, and uses mesolect to capture aspects of the intermediate variation that falls between them. He uses these terms with reference to a set of variables that drive speakers to move back and forth along the continuum. These functional and stylistic variables determine the
degree of standardization in spoken discourse. For example, formal and intellectual situations like religious sermons, lectures, or news broadcast call for more of the classical variant. Even in these settings people may move to a certain level of colloquialism for realizing a certain degree of “intimacy” with their interlocutors. Style may vary in response to the person’s emotional state since the classical variant requires more concentration, while colloquial is more spontaneous (Ferguson 1959:235; Hary 1991:71-7; Badawi 1995). It also relates to the person’s skill in MSA as determined by the nature and frequency of contact; e.g. level of education and type of occupation. Based on experiments carried out by Parkinson (1991), Haeri (1997) attests that: “The kind of contact speakers have [with MSA/CA] and their frequency greatly affects what they do or do not perceive as fuSHa [MSA/CA] and what aspects of it they master enough to use actively in the right contexts” (pp.235-9).

Hary (1996) shows experimentally that the intermediate continuum is systematic and regular; i.e. it has ordered rules by which speakers select and combine features in their attempt to standardize colloquial forms. For example, I see him in ECA is شفته [ʃuf-t-u]. Level 1 standardization is to select the equivalent koineized standard lexeme: رأيت [raʔee-t]. In level 2, the verbal suffix is inflected and colloquial long vowels change to diphthongs: رأيته [raʔay-tu-h]. In level 3, the full MSA form is realized by inflecting the pronominal suffix: رأيته [raʔay-tu-hu] (pp.81). The ability of the subjects in his experiment to rank linguistic hybrid forms on a continuum is evidence that the transitional rules across the continuum are systematic. Thus, he claims that MSA, ECA, and intermediate varieties do not constitute independent systems, but rather they all share one core of a common underlying grammar (Hary, 1996:77). Hary’s proposition is equivalent to the notion that the diglossic CS between MSA and ECA is rule governed.
Variable degrees of mixing are constantly extending to all social classes and infringing on more situational contexts previously confined to either ECA or MSA. However, the principles and constraints that govern this kind of mixing provide an analytical challenge on all linguistic levels: phonological, lexical, morphological, and syntactic.

Hence, understanding the constraints that govern diglossic mixing in Arabic has a great potential value to linguistic theory. Besides, it is a pressing demand for pedagogical and pragmatic needs. In the study: “Children’s Attitudes towards the Diglossic Situation in Arabic and its Impact on Learning” (Dakwar, 2005), elementary students reported that MSA is important for purposes of reading, writing, and learning. However, they expressed low interest and joy in learning it, and their perception that learning MSA is easy decreased along grade level. It is ironical that although children employ the similarities between the two varieties as a learning tool, teachers tend to disconnect them (pp.82-3). In the same vein, current instructional material does not educate foreign learners about grammatical and sociolinguistic constraints of mixing; which native speakers extensively and spontaneously manipulate (Parkinson, 1996:91). Chomsky (1988), on discussing children’s acquisition of different languages, suggests that “the brain must have simultaneously several different switch settings” (p. 188). Diller (1993) adds that this assumption can be extended to the acquisition of different registers in diglossic languages and “might be affected by varying degrees of input from formal education, passive exposure to mass media, and other culturally situated language-related activities” (p.395). This places a duty on linguists to develop a descriptive grammar that deduces the rules governing the interplay between the two language variants in order for educators to design curricula that may satisfy the needs of native as well as foreign learners of Arabic.
1-2 Codeswitching in diglossic languages

Differences between MSA and ECA, and the mixing of their respective features in the verbal phrase resemble in some aspects other diglossic languages. For example, šadhu bhaša is the literary written variety of Bangali, and colit bhaša is the Standard colloquial for daily discourse. In very formal occasions and scholarly topics, speakers modify their language such that it would sound like šadhu bhaša (Dil, 1986). Dil observes that in a Bangla TV debate, the speaker overwhelmingly used literary lexical items characteristic of the classical variety, in conjunction with verb forms of the more contemporary language. The speaker’s text included pure literary nouns and adjectives together with Standard colloquial verbs. The words on the left hand side of 3, 4 & 5 are examples of the speaker’s text, and those on the right are their equivalent in the other variety (Dil, 1986:461). The present study will show the extent to which Egyptian speakers employ lexemes from one variety and adapt them to the other variety phonologically or morpho-syntactically.

3. [ɔporiharjo] (šadhu bhaša) = [dɔrkar](colit bhaša) \textit{indispensable}

4. [ɔnuprerona] (šadhu bhaša) = [utšaho] (colit bhaša) \textit{inspiration}

5. [bolte] (colit bhaša) = [bolbo](šadhu bhaša)

In mixing between traditional Gurindji and Gurindji Kriol, which is an English based creole spoken in North Australia, the classical variety provides case morphology on nouns and pronouns, as well as coverbs, while most of the syntax, and the tense, aspect, mood and transitivity morphology is drawn from the spoken creole. This “split pattern of language assignment … has stabilized in Gurindji Kriol… due to the most frequent and salient input to
child learners from adults in the 1960s-80s, combined with declining proficiency in traditional Gurindji among most young people” (McConvell, 2005:9). For example, in 6, the verb morphology [-bat] is from Kriol, but the locative and dative case are from Gurjindji. Content morphemes are from both varieties (McConvell, 2005:11).

6. **nyawa-ma** wan karu bin plei-bat pak-ta
   this-TOP One child PST play-CONT park-LOC

   **nyanuny** warlaku-yawung-ma
   3sg.DAT dog-having-TOP

   *This one kid was playing at the park with his dog*

ECA lacks case and mood marking that are present in MSA. They also differ in some tense, aspect, and mood morpho-syntactic features, such as negation and future forms, in addition to the progressive/habitual prefix which is an exclusive characteristic of ECA. Mixing of these features is expected to prevail in diglossic discourse as will be shown in this study.

The following examples illustrate the mixing of features between literary and spoken Sinhala. In spoken Sinhala, the verb has one invariant form, while the literary variety inflects the verb for person, number and gender. Besides, accusative case is realized only in literary Sinhala. Also, in equational sentences, spoken not literary Sinhala drops the copula verb (Paolillo, 2000:220-3). In 7, the subject has accusative case while the verb is non-agreeing, and in 8 the subject of an agreeing verb is nominative (Paolillo, 2000:237).
7. wesak pooyadaa mahaa maayaa deewiva kohi giyaa da?

*Vesak full-moon-day great Maya queen.acc where.loc go.pst q*

*Where did the great Queen Maya go on Vesak (month) full moon day?*

8. boosat ladaruwa kawuru kawuru waḍaagattoo da?

*Bodhisattva infant who.nom who.nom hold.pst.3pl q*

*‘Who all held the infant Bodhisattva?’*

There are some similar discrepancies between MSA and ECA in agreement, particularly as related to number and gender, e.g. the absence of the dual and the feminine plural in ECA. The use of copula verb is identical in MSA and ECA, but the latter alternates regular [kāna] with the dialectal verb [ba?a].

Formal Spoken Sinhala is in a sense similar to ESA. It “appears to have Colloquial grammar with Literary lexical items giving it its formal flavor” (Paolillo, 2000: 220). Grammatical variation of mixed forms in Sinhala has been shown to be motivated by sociolinguistic factors (Paolillo, 2000:257). Gair (1968:10) points to lack of proficiency in the high variety as an additional motivational factor for mixing. It is well known that proficiency in MSA is low among the majority of Arabic native speakers. A test was designed to assess the ability to understand and speak MSA with tasks including fluency, pronunciation, sentence construction and comprehension, as well as passive and active vocabulary use. The results showed that double the number of educated native speakers scored higher than uneducated ones, while the score among non-native learners varied in proportion to their ability level (Bernstein et al, 2009:20). The
speaker’s level of education, proficiency in MSA, and the topic of discussion are vital factors that are expected to have an impact on the level of mixing between MSA and ECA.

I-3 Code-switching in distant languages

It is also enlightening to compare the results of this study not only to mixing in other diglossic languages, but also to code-switching between separate languages. The Arabic verb stem and its affixes can make up a whole phrase by itself, thus mixing may involve morphemes from either variety. This often occurs in interlanguage CS, e.g. Dholuo-English, where an English verb stem may combine with a Dholuo tense or agreement prefix (Ochola, 2006: 212-3). In 9, the past tense is indicated not by the English [-ed], but by the Dholue morpheme [nè-] (Ochola, 2006: 212-3).

9. nè- wà- talk gí professor mòrô
   PST-1PL-talk with professor Adj. another
   *We talked with another professor*

Likewise, in Spanish-English CS, English verbs are morphologically adapted by incorporating Spanish morphemes. The frequency of morphologically-adapted English verbs “suggests that morphological adaptation begins in finite forms, then spreads to non-finite participial and infinitival forms” (Pfaff, 1979: 300). This is illustrated in examples 10, 11 & 12, where Spanish suffixed to English verb stems are underlined. The frequency of ECA features creeping on MSA forms and vice versa has not been appropriately addressed so far, despite their potential significance to a full understanding of language change in Arabic.
10. Los hombres **me trustearon**

*The men trusted me*

11. Ella va a ir bien **trainiada**

*She’s going to go well-trained*

12. Yo voy a **cuitar** ya

*I’m finally going to quit*

In Hebrew, as in MSA, a noun takes an accusative case marker when it is the object of a verb. This marker is dropped from the great majority of Hebrew nouns acting as verb objects in Spanish-based sentences. Only a highly balanced bilingual would realize the case marker [et] before a Hebrew noun in accordance with the grammar of literary Hebrew, even when the sentence is Spanish-based, as in 1313 (Berk-Seligson, 1986:329-30).

13. Sávez tu, **et hápidgdm**, "El rey es kon la jénte alderedór."

*Do you know the saying, "The king is with the people around him"*

All case on MSA nouns in subject and object positions, and on participles and some modals, as well as mood markers on verbs are absent in ECA. This is an important feature to observe especially that only speakers who are very well-trained in MSA can assign case markers correctly. Attempts to realize case and mood markers in order to sound professional and formal are often erroneous.
Word order is another characteristic that plays a role in CS. Welsh and English are syntagmatically incongruent, however, CS is possible in contexts like 14, where the VSO Welsh word order is maintained because a Welsh auxiliary [mae] precedes the subject, but a main verb follows it. In this case CS to an English main verb is possible (Deucher, 2005:265-6).

14. mae o’n fath- catching
   be.3S.PRES PRON.3S.M-PRT sort of catching

   It’s sort of catching

MSA displays both VSO and SVO word orders, while VSO is often awkward in ECA except in certain contexts. It is worth investigating whether the position and variety of the subject are correlated in diglossic CS of Arabic.

In this chapter, I presented the focus of the study, and its relevance to previous work on diglossic and bilingual CS. I also pointed to the main constituents that will be lexically, phonologically, and morpho-syntactically analyzed in the corpus; namely tense, aspect, and agreement affixes, negation, future, case and mood markers, in addition to the verb stem, subject and complement. I also discussed some sociolinguistic factors that may impact the results such as proficiency in MSA. Since Freguson (1959), it has been assumed that MSA is more frequently used in religious, political and economic topics, and ECA is used in non-intellectual topics such as sport. If the data reveals that this is not necessarily the case, then morpho-syntactic and sociolinguistic factors may prove to be more effective in this respect.
Chapter II: Major theoretical approaches to code-switching

This chapter reviews different approaches in the study of CS, some of which have been employed in analyzing diglossic speech in Arabic.

II-1 The Two Constraints Theory & the Government Principle

Code-switches are governed by social context, topic, and lexical need; and by syntactic constraints which are imposed by the grammars of the two languages under consideration. Poplack (1980) suggested two syntactic constraints to account for the results of English-Spanish CS of her Peurto Rican data. She defines them as follows:

a) The Free Morpheme Constraint: “Codes may be switched after any constituent in discourse provided that constituent is not a bound morpheme” (p.585).

b) The Equivalence Constraint: “Code switches will tend to occur at points in discourse where juxtaposition of L1 and L2 elements does not violate a syntactic rule of either language, i.e. at points around which the surface structures of the two languages map onto each other.” (p.586)

Affixation of a bound morpheme of L1 to another of L2 is inhibited by the Free Morpheme Constraint unless either morpheme is phonologically integrated into the language of the other. Hence, in deriving the present participle of [eat] in 15, a Spanish suffix cannot be attached to the
English stem. Likewise, constituent elements of an idiomatic expression are treated as bound morphemes, thus, the whole idiom in 16 must be monolingual (Poplack, 1980:586).

15. *eat-iendo
   eating

16. *Cross my fingers and hope to die and si dios quiere y la virgin
   Cross my fingers and hope to die and God and the virgin willing

The Equivalence Constraint means that syntactic categories can only be code-switched if their configurations within L1 and L2 sentences are equivalent. CS in the main clause of example 17 is acceptable. However, the subordinate clause is unacceptable for two reasons. First, the English verb wants subcategorizes an infinitive complementizer contrary to Spanish. CS in this sentence violates this requirement. Second, adjectival phrases in the two languages are configurationally unequal. [car nuevo] new car follows the Spanish order, in contradiction to English (Poplack, 1980:587).

17. *El man que came ayer wants John comprar a car nuevo
   The man who came yesterday wants John to buy a new car

Poplack’s (1980) experiment involved Spanish dominant speakers who are not proficient in English, and balanced bilinguals who have equal proficiency in the two languages across a range of contexts. Although none of the switches produced by either group of speakers violate
grammaticality, the complexity of CS from English to Spanish and vice versa reflects a clear distinction among them. Balanced bilinguals favor the ‘intimate’ type and the majority of their switches is between single nouns, whereas, CS among non-proficient speakers is mainly ‘emblematic’, particularly as tags, or interjections.

On the sociolinguistic level, Poplack (1980) singles out three social factors as most effective: age of L2 acquisition, work place, and gender. Women exhibit the highest and most complex switches. The inter-relationship of linguistic and social variables is evident because early age of L2 acquisition, and close association with the wider English speaking community at work serve to advance proficiency in L2, hence balanced bilinguals are the most equipped to code-switch without violating the equivalence constraint.

This study also showed a disparity between the two speaker groups in the direction of the switch. Spanish-dominant speakers switched mostly into Spanish, while the frequency of switches from and into Spanish was comparable among balanced bilinguals. In all cases, switches involved any constituent so long as they did not violate the equivalence constraint. However, the nature of the switch differed with regard to constituent length, i.e. inter-sentential vs. intra-sentential. The former is characteristic of non-bilinguals, while the latter is attempted only by proficient speakers of L1 and L2. Poplack, therefore, concludes that CS can be a measure of language proficiency. Constituent length is also evidence that CS has its own grammar which “is composed of the overlapping sectors of the grammars of L1 and L2” (Poplack 1980:615), because length is directly related to the equivalence of L1 and L2 surface structures. In other words, the more similar L1 and L2 grammars are, the longer the switching length can be, and the
more frequently it occurs. Due to the great similarity between MSA and ECA grammars, we may expect long stretches of switching and more frequent mixing of forms. Since proficiency in MSA depends on the speaker’s level of education, in addition to other attitudinal factors, it is very likely that these factors would affect the direction, length and frequency of diglossic switching between MSA and ECA.

In some other experiments such as CS in Welsh-English (Deuchar, 2005), and diglossic Arabic (Boussofar-Omar, 2003), either the Equivalence or the Free Morpheme constraint, or both are violated. Violations are also attested in switching between languages with different phrase structures like German-English (Gardner-Chloros & Edwards, 2004), for which a one to one mapping of syntactic order is not always possible.

Moreover, theoretically, the two constraint theory is criticized for overlooking hierarchical syntactic relations. In addition, it does not account for the absence of CS data at some allowable points, nor explain why “the strength of a syntactic boundary is directly proportional to the possibilities of switching” (Di Sciullo, Muysken & Singh, 1986:4). The Government Principle captures the structural dependency of code-switched elements. It is defined as follows:

c) The Government Principle “when a government relation holds between elements, there can be no mixing; when that relation is absent, mixing is possible” (Di Sciullo, Muysken & Singh, 1986:4).
Government is defined by:

“X governs Y if the first node dominating X also dominates Y, where X is a major category N, V, A, P and no maximal boundary intervenes between X and Y” (Di Sciullo, Muysken & Singh, 1986:5).

In other words, no CS is allowed within a maximal projection. This explains why the verb and its subject can be code-switched. Likewise, a complementizer and its complement may belong to different languages. On the other hand, CS cannot occur between the object and the verb or the conjunction and the element it conjoins (Di Sciullo, Muysken & Singh, 1986:8). The Government Principle constraint is supposed to subsume most cases predicted by the Equivalence Constraint. For example, the reason for the ungrammaticality of [car nuevo] new car in 17, according to the Government Principle, is that within the maximal projection of the noun phrase, the head noun and its modifier must come from the same language.

Data from Hindi-English confirms the predictions of Di Sciullo, Muysken & Singh (1986). Example 18 shows that a complementizer must be of the same language as its governing verb, but the embedded clause is free. If that is replaced by the Hindi [ki], the sentence is unacceptable:

18. I told him that rām bahut bimār hai

   I told him that Ram was very sick (Di Sciullo, Muysken & Singh, 1986:17)
Example 19 illustrates CS between a verb [diyā] *give* and its subject. In 20, however, the object is a noun phrase. Its specifier [apnī] *our* must be in the same language as the verb [becēge] *go*. The complement *laboratory* of the determiner is free.

19. kophī ne kamāl kar diyā

*The coffee did wonders*

(Di Sciullo, Muysken & Singh, 1986:20)

20. *ham* our laboratory becēge

ham apnī laboratory becēge

*The new mayor will go to Delhi tomorrow*

(Di Sciullo, Muysken & Singh, 1986:18)

The constraint on object-verb switching has counter examples in French-Arabic CS. The verb in 21 is Arabic, while its object is French. Alternatively, the French verb in 22 takes an Arabic object (Bentahila & Davies, 1983:313).

21. ʕateik une envelope

*I gave you an envelope*

22. Il ne faut pas changer ilwṣūl

*you must not change the receipt*
In later works, the constraints set forth by Poplack (1980) have been progressively modified and have gradually converged with pragmatics and cognitive linguistics. The notion of equivalence has been extended to include not only word order, but grammatical categories on the surface level (Deuchar, 2005), or lemmas which carry conceptual information on the abstract mental level (Myers-Scotton, 2006). In this way, languages in contact engage in different types of CS according to the degree of syntagmatic (word order), paradigmatic (grammatical categories), or abstract level congruency. In the absence of all three, CS is blocked (Deuchar, 2005; Myers-Scotton, 2006).

II-2 The Matrix Language Frame (MLF) hypothesis

CS results in a combination of matrix language (ML) and embedded language (EL) constituents. ML plays the dominant role in setting the morpho-syntactic frame of the sentence. It is defined as: “the language of more morphemes in interaction types including intrasentential CS” (Myers-Scotton, 1993:68). The relative frequency of L1 and L2 morphemes is a function of psycholinguistic and sociolinguistic factors, including proficiency and markedness (Myers-Scotton, 1993:66-7). Since the roles of L1 and L2 may alternate through the discourse, these factors are assumed to set the choice of ML (Gardner-Chloros & Edwards 2004). MLF model is based on two principles (Myers-Scotton, 1993:6-7):

The Morpheme Order Principle: “Morpheme order must not violate ML morpheme order.”

The System Morpheme Principle: “All syntactically relevant system morphemes must come from the ML.”
The first principle identifies the matrix language (ML) as the language whose structural order is dominant. ML provides system morphemes; i.e. functional words like demonstratives, definite articles, and prepositions. EL provides content morphemes, which are thematic assigners or receivers, e.g. nouns, verbs, and adjectives. In example 23, ML is Swahili because the morpheme order and system morphemes are Swahili. Two English content words *come* and *books* are embedded. The sentence illustrates intraword CS, where the verb phrase [si-ku-*come*] *I didn’t come*, combines the agreement [si-] and past tense [ku-] morphemes with the English stem *come*.

23. **leo si- ku-** come na books *z-angu*

   *Today I didn’t come with my books*

   (Myers-Scotton, 1993:80)

If CS is blocked, ML or EL islands are formed. Islands are entirely composed of either ML or EL morphemes. Blocking takes place if an EL content morpheme can be realized as an ML system morpheme, or if the thematic role or pragmatic function of the EL content morpheme and its ML counterpart are not congruent (Myers-Scotton, 1993:121). To illustrate, in 24, the prepositional phrase *for you* is an EL island. It is not acceptable to use the Swahili [wewe] *you* as a complement of the preposition, because Swahili has no counterpart for the English *for*.

24. **Nikamwambia anipe uhusa niende ni-ka**-check for you

   *And I told him he should give me permission so that I go and check for you*

   (Myers-Scotton, 1993:124)
MLF does not assume speakers to be highly proficient in EL. It suffices to know the content morphemes they embed and the morpho-syntax of EL islands if formed. They do, however, need to be familiar with the structural rules of ML at least at the level of a second language learner (Myers-Scotton, 1993:7-8).

The MLF hypothesis is modified by the 4-M model to provide an account for observed violations of the bound morpheme constraint (Boussofara-Omar, 2003; Myers-Scotton, 2006). The latter does not allow intraword CS such as [si-ku-come] in example 23. In the 4-M model, morphemes are subdivided into content morphemes and three other types of system/functional morphemes:

1- Early system morphemes are conceptually linked to content morphemes, e.g. plural affixes and determiners.

2- Bridge system morphemes conjoin larger constituents within a maximal projection, e.g. partitive ‘of’ or apostrophe ‘s’.

3- Outsider system morphemes depend on elements outside the constituent they conjoin to, e.g. morphemes marking case or subject/object-verb agreement.

The Differential Access Hypothesis suggests that these four types of morphemes are accessed in the abstract level at different stages of speech production. Content and early system morphemes are accessed first, followed by bridges and outsiders (Myers-Scotton & Jake, 2000).
Content morphemes in MSA and ECA may differ lexically or phonologically, whereas the differences in system morphemes are mostly phonological. The 4-M categorization of morphemes can prove useful in analyzing diglossic switching which occurs at all morphemic levels; lexical, phonological, and morpho-syntactic.

II-3 Optimality Theory and the typology of code-switching

Cross-linguistic studies contest all attempts to generalize grammatical constraints of CS. Bhatt (1997) resolves this conflict by employing the notion of ‘ranking’ in Optimality Theory (OT). He reformulates CS constraints that have been proposed in the literature as follows:

- **Linear Precedence Constraint (LPC):** Items of code-mixed clauses follow the word order of the language of the Infl…
- **Head-Syntax (HS):** Grammatical properties (e.g. Case, directionality of government, etc.) of the language of the head must be respected within its ‘minimal domain’…
- **Equivalence (EQUI):** Switched items follow the grammatical properties of the language to which they belong.
- ***SPEC:** Avoid switching Specifier of the maximal projection in a Case-position…
- **Complaisance (COMP):** A switched specifier of the maximal projection in a Case-position must accompany a switch of its head…” (Bhatt, 1997:236).

LPC is equivalent to the Morpheme Order Principle. HS requires that the head enforces the grammatical properties of its language on its minimal projection, e.g. if an L1 head verb assigns
a particular case marker to its direct object, and the latter is switched to L2, the L1 case marker must adjoin the L2 object. EQUI is tantamount to the Equivalence Constraint. *SPEC conforms to the Government Principle. If the specifier is code-switched, COMP requires that its head X also switches.

In the spirit of OT, languages are categorized according to the order of their ranking or ‘preference’ of these five constraints. When two constraints conflict, the higher ranking constraint wins and the lower ranking one is violated. Applying this approach to CS between Kashmiri, Hindi, Spanish, Swahili, or Adaŋme and English, and between Kashmiri-Hindi, three constraint rankings emerged:

a. LPC >> {HS, EQUI} & COMP >> *SPEC for Swahili/ Adaŋme-English
b. {HS, EQUI} >> LPC & COMP >> *SPEC for Hindi/ Kashmiri-English
c. {HS, EQUI} >> LPC & *SPEC >> COMP for Spanish-English

In (a), word order ranks higher than equivalence and head syntax, but the opposite is true in (b) and (c). In (a) and (b), *SPEC can be violated in favor of COMP contrary to (c). In this way, constraints may be considered universal and CS languages would be classified according to how they set their optimal well-formedness configuration.

To illustrate, consider the case in 25. The word order of Adaŋme is followed at the expense of HS, because in English the object [mī] me must follow not precede the head verb help.
25. a ṇe mī help-e (Adaŋme-English)

They are helping me

(Bhatt, 1997:241)

Hindi-English CS ranks the two constraints differently as illustrated by 26. English is the language of the Infl, but its word order is violated by the PP [is tebl pər] this table on. The head of the switched element is Hindi, and it enforces its grammar according to HS.

26. I left the book is tebl pər (Hindi-English)

I left the book on this table

(Bhatt, 1997:242)

The interaction between LPC and EQUIV is shown in 27 and 28. The word order of the switched element conflicts in L1 and L2. EQUIV requires that the adjective red in the first example precedes house according to English grammar. However, Adaŋme is the language of the inflection and requires the reverse order. The outcome is that LPC outranks EQUIV. The opposite is true in Hindi-English, where the word order of the English NP is maintained, and LPC is violated.

27. e hé house red ò (Adaŋme-English)

He/She bought the red house

(Bhatt, 1997:243)
28. **use aur ab bakaḍya** professor of linguistics **hai**  (Hindi-English)

   *And now he is a professor of linguistics*

   (Bhatt, 1997:244)

The following two examples show the conflict between *Spec and COMP. The two verbs *finished* and [uska] *read* assign case to the noun phrases [pocos estudiantes] *few students* and [uska critique] *his critique* respectively. In 29, the head noun is also switched according to COMP, but in 30, the specifier [uska] *his* agrees with the language of its case-assigning head; the verb [pɑṭʰa] *read*. Thus, COMP ranks higher in Spanish-English, and *SPEC is higher in Spanish-English.*

29. **pocos estudiantes** finished the exam  (Spanish-English)

   *Few students finished the exam*

   (Bhatt, 1997:245)

30. **maine uska critique pɑṭʰa**  (Hindi-English)

   *I read critique his*

   (Bhatt, 1997:246)

The **Optimality approach applied to MSA-ECA codeswitching**

Table 2 is a rough outline of some categories that differ in MSA & ECA verb phrase and would fall under the syntactic constraints summed up in this section. LPC applies when the verb precedes the subject; otherwise, both varieties share the same order. It also applies in the context...
of demonstratives in adjectival position, because they may precede or follow the noun depending on the variety. HS is relevant to the presence or absence of case and mood markers, to negative forms, complementizers that assign case, and to number/gender agreement between the verb and subject, especially with the dual and feminine plural; and in verb initial phrases when the subject is plural. EQUI will not apply because of the equivalent configurations of syntactic categories in MSA and ECA. *SPEC and COMP apply to pronominal demonstratives that differ lexically in MSA and ECA, and to the pronunciation of the definite article ([ʔal] vs. [ʔil]) and of some prepositions when they are pronominally suffixed (e.g. [ʕalay-ha] vs. [ʕalē-ha]).

| LPC: VSO order | Demonstratives in adjectival position |
| HS: case/mood | negation markers | complementizers |
| *SPEC/COMP: V-S agreement in VSO | definite article | Pronominal demonstratives | prepositional complement |

Table 2 Syntactic constraints and their relevant syntactic categories in MSA/ECA

**II-4 The Dual Language Model (DLM)**

Approaches presented in previous sections are all syntactic and sociolinguistic. DLM explains CS from a cognitive-pragmatic perspective. DLM and MLF models acknowledge that abstract conceptual information gets realized in syntactic structure. They differ, however, in the analytical approach. While MLF analyzes the surface structure by associating it with the abstract
level, DLM works in the opposite direction by associating the conceptual level with the syntax. DLM assumes that the bilingual possesses a dual language system called the Common Underlying Conceptual Base (CUCB) and two language channels for L1 and L2. The role of CUCB in speech production is explained as follows (Kecskes, 2006:260):

“… production begins with the speaker’s intention, which results in the preverbal message formulated and which is pre-structured in CUCB (conceptualizer). From the CUCB, the preverbal message gets into the language channels (formulator) where it gains its final form (articulator) by mapping conceptual representation onto linguistic representations and comes to the surface in a language mode required by the interplay of context and the speaker’s strategies.”

It is the “preverbal thought”, not ML, that selects an L1 or L2 grammatical frame, because the motivation for CS is primarily conceptual-pragmatic rather than syntactic. Hence, DLM is more concerned with content rather than system morphemes. CUCB contains concepts that are common to L1 and L2, language-specific concepts, and synergic concepts. The latter refers to “concepts that are lexicalized in both languages but have different socio-cultural load in each language” (Kecskes, 2006:263). The majority of concepts are shared between MSA and ECA, but there are some concepts that are variety specific especially since ECA is constantly evolving and coining new terms, e.g. روش [rewif] a cool person has no equivalent in MSA. Many concepts are synergic, e.g. راح [rāh] in ECA means to go. In MSA, it has an additional time
denotation of *going at night*. For this reason speakers find one variety more expressive of certain concepts than the other, especially idiomatically.

Like MLF, DLM acknowledges differential activation of the two languages, which gives rise to three types of CS; alternation, insertion, or congruent lexicalization. They are defined as follows (Muysken, 2000 cited in Kecskes, 2006:268-9):

“Insertion involves the incorporation of lexical items or entire constituents from one language into a structure of another. …

Alternations are distinguished from insertions by the size of the unit switched. They are usually larger than a single lexical item or phrasal constituent that usually encodes a single concept associated with a given language. … Syntactic relations do not extend over the conceptual units being conjoined as in the case of insertion.

Congruent lexicalization is defined on a surface level as the combination of items from different lexical inventories into a shared grammatical structure. ... The mapping of concepts onto linguistic form can often include function words that are attached to content words or expressions…[It] involves the sharing of grammatical structures and features between lexical items or expressions from different languages”.

Sometimes the inserted constituent is reduplicated by providing the L1 and L2 lexemes in succession, demonstrating the simultaneous activation of both language channels; as in 31.
31. **con el** sailors, **con los marineros**, sailors

*With the sailors, with the sailors, sailors*

(Kecskes, 2006:274)

In 32, the noun *nurse* is a single “inserted” concept, and *they were going to have a baby* is an “alternation” that strings together a series of concepts.

32. **había dos pacientes yanitos**, they were going to have a baby. **Ellas prefieren que esté**

*un* nurse **con ella que es de Gibraltar**

*There were two Yanito patients, they were going to have a baby. They preferred to have a nurse with them from Gibraltar*

(Kecskes, 2006:275)

By congruent lexicalization, the same concept has equivalent forms in L1 and L2, e.g. ‘acusar’ in Spanish and ‘accuse’ in English. In sentence 33, the English verb is used in a Spanish context. Due to the interaction between the two languages, Spanish grammar intervenes by attaching structural features of the Spanish equivalent to the English verb. This results in structural ill-formedness with respect to English grammar because *accuse* requires an NP complement, while Spanish [acusar] subcategorizes a PP.

33. He accused **a Mister Bigote de doble lenguaje**

*He accused *to Mister Moustache of double talk*

(Kecskes, 2006:271)
Kecskes applies DLM analysis to English-Spanish bilinguals in Gibraltar and finds that 20% of their switches are of the congruent lexicalization type. Since speakers may freely combine concepts from L1 and L2 at the preverbal level, then formulate them lexically and grammatically through the two “constantly interacting language channels”, the surface grammatical outcome may violate the structural rules of either language (Kecskes, 2006:279).

Although Kecskes states that DLM assumes equal proficiency in L1 and L2, he notes that language proficiency in L1 and L2 is closely linked to the underlying conceptual development and cultural competence in both languages. As a result, the three types of CS convey different levels of bilingual skill, as well as social or grammatical characteristics of the languages involved (Kecskes, 2006:266). Balanced bilinguals in Poplack’s (1980) study who are immersed in an L2 society apart from their L1 homeland, exemplify the alternation pattern, whereas Spanish-dominant speakers are described by the insertion pattern. Here L2 intervenes with L1 only sporadically and for short utterances. Finally, congruent lexicalization “is likely to occur between closely related languages, where their relative prestige is roughly equal, or where there is no tradition of overt language separation” (Gardner-Chloros & Edwards, 2004:121-2).

CS in diglossic Arabic may fit the congruent lexicalization type, which is characteristic of languages that have a common grammatical system but diverse vocabulary. However, it may not do so because the relative prestige of MSA and ECA differs. MSA is the formal language and is mastered only by educated speakers. It “enjoys very great prestige among Arabs” religiously and patriotically, because it is the offspring of CA, the language of the Qur’an, and is shared by
natives of all Arab countries. ECA, on the other hand, is “too often subject among Arabs to strangely unreasoning scorn” (Mitchell, 1986:8).

In this chapter, I gave a review of four main approaches to the analysis of CS, and pointed to how they may or may not apply to MSA-ECA mixing. Some of these approaches have been employed in studying Arabic CS. The next chapter discusses some of these studies.
Chapter III: Codeswitching in Arabic

In this chapter I review the methodologies and results of five previous studies on Arabic-French, Arabic-English, and Standard-Dialectal CS, in addition to one study that compares CS between Arabic and a different language and between Standard and Dialectal Arabic. The Arabic dialects involved are Egyptian, Moroccan, Tunisian, Levantine, and Gulf. This review will be of use in the analysis and discussion of my data. The examples in this chapter are from the studies reviewed. The standard variety or a foreign language in the examples is highlighted in bold, and dialectal Arabic is written in regular font.

III-1 Corpora and methodologies

Bentahila & Davies (1983) used a corpus of seven hours and a half of spontaneous conversations, in addition to elicitation judgments of constructed examples that are unavailable in the data. Elicitation is employed because, in Bentahila & Davies’s opinion, the absence of a structure may not be due to a CS constraint, but there may not be a sociolinguistic motivation for it (p.308). CS at all syntactic boundaries starting with the clause $\bar{S}$ down are examined.

Eid (1988) studied MSA-ECA codeswitching in radio/TV interviews and panel discussions with a university professor, a journalist, and some Cabinet members. Four syntactic structures are analyzed:

- Subordinate clause
- Relative clause
- Tense and verb constructions
- Negative and verb constructions

The relative and subordinate conjunctions; and markers of tense and negation are taken as focal points. These focal points differ in MSA and ECA, and occur in conjunction with other structural elements, namely a clause or a verb. The hypothesis was that these constituents would play a role in allowing or blocking CS. There are eight possible combinations of MSA and ECA immediately before and after the focal point for each structure. Data is classified according to these eight combinations. Combinations that never occurred in the data were tested for acceptability through constructed examples.

Eid’s criterion for judging elements as belonging to a certain variety is the presence or absence of an equivalent in the other variety. For example, رأيت [ra?ayt] I saw is clearly MSA and its ECA equivalent is شفت [juft]. The intermediate phonetic variant [ra?ēt] I saw has the same ECA equivalent. Therefore, it is also marked MSA, in spite of its deviation from the standard pronunciation. All data whose form is shared by MSA and ECA is disregarded.

Eid (1992) collected a corpus of five hours of spontaneous conversations among six Egyptian American bilinguals who are highly educated and fluent in both languages and had lived for at least ten years in the United States. Their ages range between twenty two and forty five. She focused on four types of clauses:
- Co-ordinate clause
- Subordinate clause
- Relative clause
- Complementary clause

These clauses share the structure: X-marker-Y; where the marker is the co-ordinate or complementary conjunction, relative marker, or complementizer. The method is the same as that of her Arabic-English study (Eid, 1988). For each clause, there are six possible combinations of English-Arabic switch patterns excluding monolingual combinations.

Boussofara-Omar (2003) works within the MLF framework. She uses a corpus of 17 public political speeches by the Tunisian President Bourguiba, in which styles vary from formal, semi-formal to informal. Two mixings that cannot be explained satisfactorily by MLF are discussed: the co-occurrence of standard and colloquial system morphemes in the same CP, and cases when CS results in subcategorization clashes between the two varieties. She examines the two structures using the 4-M and the Abstract level model that are modifications of MLF.

Bassiouney (2006) also works within the MLA framework. Data is composed of political speeches, mosque sermons, and a university lecture. Words in every monologue are tagged as either MSA or ECA. Words that are common to MSA and ECA are tagged ‘neutral’, and those that combine morphemes from both varieties are labeled ‘mixed’. According to the total count of each class, discourse is categorized as mainly MSA, mainly ECA, MSA with insertions from ECA, ECA with insertions from MSA, or a mixture of MSA and ECA. When the prominent code varies within one monologue and code variation maps to a transition in the subject of discourse, the text is broken into parts corresponding to this variation. Bassiouney’s analysis considers three mixed forms:
– Negation marker and verb or noun
– Demonstrative marker and noun
– Aspectual marker and verb

Albirini (2010) used audio and video recordings of religious lectures, political debates, and soccer commentaries in the media to compare the constraints on interlanguage CS on one hand and diglossic CS of Arabic on the other. Speakers were Egyptian, Gulf, and Levantine Arabs. Acceptability judgments were used to confirm the validity of the results. The analysis is divided in two stages: Stage 1 examines the constraints proposed in the literature of CS with respect to Standard-Dialectal mixing in the corpus. Stage 2 focuses on sentences that involve any of the following parameters:
– Pro-drop parameter
– Head directionality parameter
– Serial verb parameter

The acceptability of CS between Standard-Dialectal Arabic in these sentences is compared to equivalent Arabic -English, Spanish, French, Hebrew or Turkish sentences. The hypothesis is that CS between MSA and Dialectal Arabic is incompatible with CS between Arabic and a different language because Arabic varieties share a single syntactic system.
III-2 Approaches to Arabic code-switching

The Free Morpheme Constraint

Based on their data Bentahila & Davies state that “code-switching is not possible across word-internal morpheme boundaries” (Bentahila & Davies, 1983:317). This is a re-statement of the Free Morpheme Constraint. However, they do have intra-word switching, which they consider exceptions. Example 34 from Gulf Arabic, in Albirini, demonstrates the use of the dialectal aspect marker [bi-] with an MSA verb stem [taqūd] drive in violation of the Free Morpheme Constraint.

34. w ʔant bi-ta-qūd is-sayyāra

وأنت بتقود السيارة

While you are driving the car

(Albirini, 2010)

The Equivalence Constraint

French is strictly an SVO language, and Arabic displays both SVO and VSO. French word order is violated in sentence 35 because the verb [ja] came precedes the subject [le contrôle] the checking time. In 36, the Arabic noun [l waraqa] the paper is definite, but the French adjective [bleue] blue is not in violation to the rules of Arabic.

35. ja le contrôle

The checking-time came

(Moroccan-French)

(Bentahila & Davies, 1982:319)
Likewise, the Arabic-Turkish word-order differs. Albirini’s example 37 is unacceptable because Turkish requires the verb [jirib] *drank* to come last. In contrast, SVO word order is grammatical in MSA and ECA alike, and switching is possible, e.g. 38.

### 37. *köpek jirib mayya*  
*(ECA-Turkish)*  
(Albirini, 2010)

### 38. *al-kalbu jirib mayya*  
*(ECA-MSA)*

الكلب شرب مية  

l-kalb *jiriba māʔan*  
*(ECA-MSA)*

الكلب شرب ماء  

*The dog drank water*  
(Albirini, 2010)

In some other contexts, the linear order of the two varieties may differ. For example, in phrase 39, the ECA demonstrative pronoun [da] follows an MSA referent [t-taklīf] *responsibility*. Phrase 40, which is entirely MSA, shows that the linear order in MSA requires that the
demonstrative [hāūihi] precedes the referent. Hence, the equivalence constraint is violated in the first phrase.

39. *wa yurfaḥ ūannu t-taklīf* da 

ویرفع عنه التكليف ده

*And his responsibility is lifted off his shoulders*

(Bassiouney, 2006:119)

40. *hāūihi l-manatīq*

هذه المناطق

*These regions*

(Bassiouney, 2006:114)

The Government Principle:

The Government Principle does not constrain Arabic diglossic CS. All six studies have examples of CS within a maximal projection, for example, between the verb and its object as in 41, or within an adverbial phrase as in 42.

41. *ateik une envelope* 

*I gave you an envelope*

(Bentahila & Davies, 1982:313)

42. *amām wilād-u* 

 أمام أولاده
Directionality of CS:

Boussofara-Omar observes that MSA verb stems are inflected with MSA affixes only in EL islands, as in frozen expressions and Qur’anic quotations. Bassiouney also notes an uneven distribution of negative and demonstrative structures. MSA negatives [lā or laysa +PP] are absent from her data. Also ECA pronominal demonstratives [DEM + pronoun or noun], e.g. 43, are more frequent than their MSA equivalent. She gives no interpretation for this other than the ‘reluctance’ of speakers to use certain structures. Demonstratives in adjectival position precede the noun in MSA [DEM + definite noun], as in 39, and follow it in ECA [definite noun + DEM], e.g. 43. The MSA form occurs more frequently. Bassiouney accounts for this tendency by the markedness and saliency of the MSA form, which motivates speakers to implement them as a pragmatic tool.

43. hāḍa ragul-un ɣarīb-un

هذا رجل غريب

da rāgil ɣarīb

ئده راجل غريب

This is a strange man

When CS occurs, the direction of the switch may also have an uneven distribution. Bentahila & Davies found that in certain structures, CS in one direction tends to be more frequent than in the
other, e.g. an Arabic determiner or preposition with a French noun far exceeds the reverse. Boussofara-Omar observes that CS between a prefix and a verb is only from ECA to MSA. Albirini shows that CS is allowed between a dialectal demonstrative and an MSA noun. In 44, [hal] *this* is a demonstrative, and [kalām] speech is MSA according to the Levantine dialect. For the other direction, however, Bassiouney observes that an MSA demonstrative never precedes an ECA lexeme.

44. **an-nabiyy ... qāla hal-kalām** (Levantine-MSA)

النبي ... قال هالكلام

*The Prophet ... said this speech*

(Albirini, 2010)

An asymmetry in the direction of CS is also found in the four syntactic structures in Eid (1988). Switching before the focal point (relative marker, subordinating conjunction, NEG or tense marker) is free as shown in 45 and 46. The negative marker is MSA and ECA respectively, and the element just before NEG is switched in both cases (Eid, 1988:58-9). Accordingly, we may expect that CS between the negated verb and its subject is unconstrained in an SVO clause.

45. **bass lam takun bayn-i wa bayn-u șadāqa** (ECA-MSA)

بس لم تكن بيني وبينه صداقة

*but there were no friendship between me and him*

(Eid, 1988:58)
46. *at-tabaqa l-ṣāmiḥa fī-l-mādīhi ma-kat-ḥ bi-taṣamtiḥ bi-ḥurriyit-ha (ECA-MSA)

الطبقة العاملة في الماضي ماكتش بتستمتع بحريتها

*In the past, the work class did not enjoy its freedom*

(Eid, 1988:59)

Bassiouney & Eid (1988) find no examples of MSA negative marker followed by an ECA verb in their data. This result applies not only to NEG, but to the element after other focal points studied by Eid (1988). Hence, she posits the following Directionality Constraint:

“If the focal point is from SA [MSA], switching to EA [ECA] would not be permitted at the position immediately after that focal point” (p.74).

In Eid’s study sentence 47 was ranked unacceptable, which means that an MSA relative pronoun [allaḏi] *that* is an MSA cannot be followed by an ECA clause. However, the opposite in 48 is possible. Similarly, in subordinate clauses, such as the purpose clause in example 49, the head [yaḏān] *in order to* is ECA and S is MSA.

47. *fi-l-waqt ẓallaḏi bi-nāṭ-u dilwaṭti* (ECA-MSA)

وفي الوقت الذي نعيشه دلوقت

*at the time in which we are now living*

(Eid, 1988:60)
48. di illi **waqafat** hayat-ha ḳalē-na

*She is the one whose life is devoted to us*

(Eid, 1988:61)

49. ig-ġēj ʕām ʕaʃān **yuʔaddi fiyl**

*The army rose in order to perform an action*

(Eid, 1988:61)

Likewise, CS before the main clause in Arabic-English is free, but constrained after the English marker by DC, restated by Eid (1992:63) as:

“Switching after an English marker is not permitted. But after an Arabic marker it is free unless that marker is a relative marker.”

The same is attested in Arabic-French CS. CS is accepted after an Arabic wh-word in interrogative clauses, as in 50, but judged odd after a French wh-word.

50. ṣkun **a dit ça**

*Who said that*

(Moroccan-French)

(Bentahila & Davies, 1982:311)
According to DC, a relative marker is always followed by an element of the same variety or language. Contrary to Eid (1992), the analysis of Bentahila & Davies does not show asymmetry in direction of the switch between a main and embedded clause, whether it is adverbial, conditional, coordinate or relative. CS is free before or after a complementizer, relative pronoun, and conditional or coordination conjunction. In 51, the coordination marker is French followed by an Arabic clause. Also, Albirini precludes that Standard-Dialect CS between a functional head and its complement is constrained. In 52, the relative pronoun [illi] what is dialectal, and its complement is MSA.

51. ana tanxarj hadji kulu et tan dir l maa (Moroccan-French)

   *I take everything out and pour water over*

   (Bentahila & Davies, 1982:310)

52. tahqaq illi ana qultuh (Gulf-MSA)

   *what I said has happened*

   (Albirini, 2010)

Despite this disagreement, directionality is prevalent in various structures thus reviewed. To explain why one direction is favored over the other, Eid (1988) cites similar findings in other languages and attributes the phenomenon to “the manner of acquisition of each variety” (p.75). The directionality effect differs among languages. In Arabic, it is the non-native variety (MSA)
or the foreign language (English or French) that controls CS. But in Swedish-English, for example, the natively acquired Swedish is the controlling language (p.78).

**Contradictory Effect Constraint (CEC)**

Eid (1988) found that switching is not only constrained after an MSA negative marker as DC predicts, but an ECA negation always selects an ECA verb (e.g. 45 & 46). In an attempt to account for why CS is not allowed between NEG and verb, she suggests a Contradictory Effect Constraint (CEC):

> “Switching at some point, P, between two elements A and B is not permitted if the grammar of the two language varieties involved include contradictory conditions applicable to A and B-conditions that cannot be satisfied simultaneously” (Eid, 1988:74).

In MSA, tense is attached to the negative marker (mood-assigner type in section IV-5 below), where there is a distinct marker for every tense and the verb is always imperfective. ECA, on the other hand, has a shared marker for the negative, and tense is realized on the verb. Hence, if a colloquial NEG is followed by a standard verb, tense is not realized. Conversely, if a standard NEG is followed by a colloquial verb, tense is doubly marked on the NEG and verb. CEC resembles the Equivalence Constraint, but the latter focuses primarily on linear order, which is very similar in MSA and ECA verb phrases.

Data of Bassiouney and Boussofara-Omar, however, included MSA verbs that are often negated by ECA markers. The predominance of ECA system morpheme drives Bassiouney to conclude
that ML triggers the use of negation morphemes of its same code. The following example has the Tunisian negative [ma...] with the MSA [ʔa-ðunn] I think imperfective.

53. ma-ʔa-ðunnu-ʃ kēnu  (Tunisian-MSA)

ما اظننا كانوا

I do not think they were

(Boussofara-Omar, 2003:39)

This finding is confirmed by another study on CS between MSA and Hejazi dialect. Tense in example 54 is meant to be future, but neither the verb nor the Hejazi NEG [mā] is marked for future tense. According to Eid (1988), either [mā] must be followed by the Hejazi equivalent of the verb أفشل [ʔafaʃʃil-ik] I fail you, or the standard NEG [lan] should be used.

54. tāni mā-ʔaxōil-aki  (Hejazi-MSA)

تاني ما أخذك

Next time I won’t let you down

(Sabir & Safi, 2008:98)

Pro-drop in Albirini’s data provides support to CEC. In French, the inflection of the verb [vit] lives may denote third person singular masculine or feminine, hence the subject [elle] she must be overtly expressed. For this reason Moroccan-French CS in 55 is unacceptable. In contrast, pro-drop is possible in Standard-Dialectal CS, because the subject of an Arabic verb is denoted in the person affix of the verb.
Subcategorization requirement

In Eid (1992) no switch occurs after an English marker. However, it is allowed after a French marker on the condition that subcategorization is satisfied:

“All items must be used in such a way as to satisfy the (language-particular) subcategorization restrictions imposed on them.” (Bentahila & Davies, 1983:329)

For example, in 56, the switch is only accepted when the French infinitive [réussir] to succeed is marked by the Arabic tense prefix [n-], because [baʃ] in order to requires a finite verb.

56. nqra ¡wiya baʃ n- réussir á l’examen (Moroccan-French)

We work a bit in order that we may succeed in the examination

(Bentahila & Davies, 1982:323)

Albirini argues that dialectal and standard Arabic share one syntactic system. Consequently, subcategorization conflicts are not likely to occur in most contexts. For example, diglossic mixing in a serial verb structure is judged acceptable in his data in either direction, as in 57. In
contrast, Levantine-Spanish CS between serial verbs is not, because Spanish requires an agreement marker on the phrasal head, which is missing in the unacceptable sentence 58. The same holds for CS between an Arabic auxiliary and French infinitive. In 5859, CS is possible under the condition that the durative [tat-] prefixes the French verb [gratter] to scratch because [tatbqa] keep subcategorizes a finite verb.

57. taʕāla ʃūf / taʕa nṯur  
(Levantine-MSA)

تعالي شوف / تع انظر

Come see

(Albirini, 2010)

58. * venga ʃūf  
(Levantine-Spanish)

Come see

(Albirini, 2010)

59. tatbqa tatgratter  
(Moroccan-French)

You keep scratching

(Bentahila & Davies, 1982:315)

Although an Arabic prefix attaches to a French verb in order to satisfy subcategorization requirements, DC intervenes to prevent a French clitic pronoun or object pronoun from attaching to an Arabic verb, e.g. 60 & 61.
60. *je yadi/*ana vais   (Moroccan-French)

    I go

    (Bentahila & Davies, 1982:312)

61. *je vois hum   (Moroccan-French)

    I see them

    (Bentahila & Davies, 1982:314)

Pronoun doubling

Arabic complementizers are always followed by a nominal or pronominal subject. Eid (1992) observes that when the subject after an Arabic complementizer is an Arabic pronoun, the latter may be doubled by an English pronoun. For example, in 62, the subject pronoun [-i] I is suffixed to the complementizer [inn] that, and followed by its English equivalent. Eid (1992) notes that pronoun doubling has been observed in other studies cross-linguistically, e.g. Arabic-French and Spanish-Hebrew. The same phenomenon occurs in Arabic-French CS, e.g. in 63 [*ana] I is duplicated by the French [je] I.

62. What can I do huwwa inn-i I can join the air force   (ECA-English)

    What can I do is that I can join the air force

    (Eid, 1992:58)

63. il croyait bi *ana je faisais ça exprès   (Moroccan-French)

    He thought that I was doing that on purpose

    (Bentahila & Davies, 1982:311)
In an attempt to account for pronoun doubling, Eid (1992) refers to verb duplication in Japanese-English CS. In 64, the verb is duplicated in order to satisfy both the English SVO order, and the Japanese SOV one. This results in an SVOV sentence.

64. I saw Judy **mita** (Japanese-English)
   
   *I saw Judy I saw*  
   
   (Eid, 1992:66)

In a parallel account based on CEC, the difference between Arabic and English subject-verb agreement motivates duplication, because the agreement paradigm of Arabic differs from that of English and French. Hence, in 65, the Arabic pronoun [nta] *you* cannot precede the French verb [vas travailler] *will work* unless the latter includes an equivalent French clitic [tu] *you* that denotes person. This account, however, cannot explain why only pronominal, not nominal, subjects are duplicated.

65. nta **tu vas travailler** (Moroccan-French)

   *You, you are going to work*  
   
   (Bentahila & Davies, 1982:313)

**Dominance of L1 grammar**
Bentahila & Davies hypothesize that for a bilingual “the grammatical formators of the first language remain more basic even after the assimilation of the second language is also complete, and so tend to surface frequently even in L2 environments when the speaker is using this code-switching variety which pools the resources of both languages” (p.327). This is clear in some equivalence violations such as definiteness and agreement. For example, the feminine adjective [kulha] *whole* in 66 modifies a masculine French noun although the determiner is also masculine. Since the Arabic word for “journey” [riḥla] is feminine, the speaker is obviously influenced by his/her native language. This account reminds us with the Common Underlying Conceptual Base of the Dual Language Model discussed in the previous chapter. The preverbal message is conceptualized in Arabic, but formulated as [le trajet] *journey* in French. Gender agreement is, thus, mapped onto the Arabic concept.

66. dak *le trajet* kulha  
*MOROCCAN-FRENCH*  
*that whole journey*  
*(Bentahila & Davies, 1982:327)*

Example 66 also illustrates the subcategorization condition. The French definite marker *[le]* is inserted, because the demonstrative *[dak]* *that* subcategorizes a definite noun.

**Matrix Language Framework**

Bassiouney considers MLF the most appropriate model for analyzing MSA/ECA switches, because it does not rely on linear order, or on any particular theory of grammar. For example, phrase 67 is analyzed in MLF by identifying the matrix language as ECA, because system
morphemes (the definite article and demonstrative), and the word order are ECA. Content morphemes, on the other hand, [ʕaql] mind and [māda] substance are MSA.

67. it-ʕaql da māda (MSA-ECA)

This mind is a substance.

(Bassiouney, 2006:142)

However, system morphemes from both varieties mix. In 68, tense is marked for future by MSA [sa-], whereas negation [ma...] is dialectal and the verb stem is common between MSA and ECA. Besides, the ECA aspectual prefix [bi] may be dropped from ECA verbs, and may surface with MSA. For example, [bi-tunaffaḏ] were being applied, where [bi-] adjoins an MSA u-a passive verb. CS also involves content morphemes from both codes, as in 69, where the noun [kalām] talk is MSA or ECA, and [kāfiyan] enough is MSA.

68. ma-sa-ta-qif-ʃ (Moroccan-French)

ما ستفتتش

you won’t stand

(Boussofara-Omar, 2003:40)

69. hāḏa k-kalām laysa kāfiyan (MSA-ECA)

هذا الكلام ليس كافيا

52
Hence data does not always fit MLF especially that “it is almost impossible at times to say whether a certain morpheme belongs to ECA or MSA. Bassiouney determines ML statistically depending on the relative counts of MSA and ECA morphemes. But, it is not very easy to come up with one ML, since it is sometimes difficult to decide which code is being used in the first place” (Bassiouny, 2006:48). As a way out, Bassiouney adopts the composite ML. The composite ML is based on The Abstract Model, which divides the lexical structure in three levels: conceptual/pragmatic, predicate-argument/thematic, and morphological. These levels are parallel to the conceptualizer, formulator, and articulator of DLM presented in section II-4. In CS, structural levels may be formed by L1 or L2, allowing for the co-occurrence of content and system morphemes from both languages. For example, a lexeme may have thematic specifications of ML, but get realized morphologically as EL. Speakers are assumed to resort to the composite ML either because they do not have full access to the grammar of ML or they have divided loyalties towards the two languages (Myer-Scotton & Jake, 2000).

Boussofara-Omar (2003) questions the validity of the composite ML as well as the 4-M model by inquiring: “why are fuṣḥaa [MSA] tense/aspect markers not consistently and systematically activated along with fuṣḥaa verbs?” (p.41). Besides, it is “not possible to determine the ML in a CP if both languages supply the system morphemes” (p.41). She adds that electing an EL system morpheme should “constrain the choice of the late system morpheme that is necessary to structure this constituent” (p.41). In example 68, the dialectal negative controls the choice of the
future marker, such that if سوف [sawfa] will was selected instead of its equivalent [sa-], the negative form would be ungrammatical.

The alternative account proposed by Boussofara-Omar is that an EL system morpheme is called to rescue a conflict that may arise between ML system morphemes. In 68, the discontinuous negative marker [ma...f] rather than its alternative [mūf] is activated first because it is more salient. Since [ma...f] cannot be used with the ML future marker بيش [beeʃ] will, EL [sa-] is called. This account preserves the distinctive roles of ML and EL, and constrains the activation of EL system morphemes to the role of resolving possible conflicts (p.41).

Boussofara-Omar also addresses the case when system and content morphemes are from one variety but the word order and/or subcategorization follow the other variety. Ruling out the composite ML account, she doesn’t give any alternative explanation, but speculates that native speakers of different Arabic dialects share similar views of what constitutes spoken MSA. She states that “The lexical (i.e. content morphemes) and morphological (i.e. system morphemes) ‘flags’ seem to supersede the syntactic constraints in speakers’ judgment of what constitutes spoken fuṣḥaa [MSA] and the dialect” (p.44).

The data of the studies discussed in this chapter is at large drawn from political discourse. The Hejazi study (Sabir & Safi, 2008); which I referenced while discussing the Contradictory Effect Constraint, is unique in showing that diglossic CS is not restricted to political topics nor to adult and intellectual speakers. The subject of the study is a five-year-old Saudi child who has fully acquired the low variety, but is only exposed to MSA through cartoon films. The analysis shows
that the category most frequently switched is verbs although it is morphologically more complex than other categories. The Equivalence Constraint is never violated, indicating that “an underlying competence of the syntactic structures of both varieties [is attained] at a very young age” (Sabir & Safi, 2008, 91).

This chapter reviewed CS studies on Arabic and yielded several generalizations. First, the Two Constraints Theory and The Government Principle are inadequate for diglossic switching. This does not mean that Standard-Dialectal CS is unconstrained. The Contradictory Effect Constraint and subcategorization requirements play an important role in controlling CS. Second, MLF if applied requires a composite ML or other adaptations in order to account for the intricate mixing of system and content morphemes from either variety. Third, CS in many structures is more constrained and sometimes blocked when the first element is MSA. Fourth, although the standard and dialectal codes are actively interacting all the time, the colloquial language remains more dominant in formulating utterances. Finally, since there are phonological and agreement differences in MSA-ECA with respect to pronouns, pronoun reduplication may occur in CS between the two varieties.

If we try to relate these generalizations to Bhatt’s (1997) typology of code-switching (explained in section II-4), we look in the examples cited in this chapter for the constraints and categories listed in Table 2. The following tentative results are obtained:

- LPC is violated with demonstratives in adjectival positions. In 39, Infl [yurfاط] be lifted is MSA, but the demonstrative [دا] follows ECA order.
The generalization regarding directionality of CS is that an MSA marker restricts the element after it to its same variety. Examples 45 & 46 were accounted for by CEC, and 56 was explained in terms of subcategorization. According to HS constraint, an MSA head in these examples assigns case to its complement. Since case is never realized in ECA, CS is blocked in order not to violate HS. Hence, HS ranks over LPC.

As for *SPEC and COMP, relevant examples involve CS between the specifier and the head. In 44 and 66, the specifiers [hal] and [dak] are switched without switching the head noun. This means that *SPEC ranks over COMP.

We may, hence, expect that diglossic CS of Arabic and Arabic-English/French CS are of the type: HS >> LPC & *SPEC >> COMP. This tentative expectation will be tested by the findings of this study.
Chapter IV: Tense, Aspect, Modality

This chapter reviews what is meant by tense, aspect, and modality and the distinctions within each of these categories. It will show how “[v]erbal inflection, copular and auxiliary verbs, periphrastic verbal expressions, modal and negative particles, as well as temporal adverbs, contribute to locating situations in time” (Fassi Fehri, 1993:145) in both varieties; MSA and ECA. The analysis in chapter VI inspects these distinctions in the data. The list in Appendix B is the abbreviations that are used in this study for tagging Arabic sentences. All examples in this chapter are constructed by me. MSA morphemes are marked in the examples in bold, and ECA morphemes are in normal font.

IV-1 The Arabic verb composition

The Arabic verb is a dense morphological composite. Its root carries the lexical/semantic meaning. Tense, aspect, and agreement denotation of its affixes and/or vocalic melody is disputable. In traditional Arabic grammar, the verb inherently denotes a meaning in relation to one of the three tenses (ibn hifām, n.d.:14). It has three forms: past: what has elapsed, imperative: what is going to be and has not happened, and present: what is being and not having ceased (ibn hifām, n.d.:20; sibawayhi, 796 cited in Bahloul, 2008:40). The three forms differ morphologically, while the future, which is not included as a separate form, is morphologically identical to the present, and is often lexically marked by سوف [sawfa] will or its alternative prefix [sa-]. In this classification, the temporality of the imperative derives from the potential
occurrence of the event in future time. This view is adopted in modern Arabic grammar books and by some linguists (Bahloul 2008:40).

Others claim that the Arabic verb is not temporal, but aspectual. It has two forms, the first denotes “a finished act, one that is done and completed in relation to other acts (the Perfect); the other an unfinished act, one that is just commencing or in progress (the Imperfect)” (Wright 1996:52).

Based on the notion of abstract vs. relative tense, Comrie (1976) argues that Arabic verb forms have a combined Aspect/Tense meaning; where tense is relative to some process other than current time. Hence, the “perfect refers to past time (or completed action) and the imperfect to present or future time (an incomplete action)” (Bateson, 2003:23). This dual tense-aspect characterization of Arabic verbs is also forwarded by Fassi Fehri (1993) based on a view that verbal inflections have a combined grammaticalization of aspectual and tense oppositions.

From a different perspective, Aspect/Tense are considered abstract; i.e. not realized morphologically and affixes are only agreement morphemes. This view derives from the observation that the negative verb ليس [laysa] is not and the aspectual مازال [mā-zāla] is still denote present, but inflect with perfective suffixes. Similarly, the vocalic melody of the verb does not carry tense or aspect, since these negative and aspectual particles resemble perfective verb vocalic melody (Aoun, et al., 2010:22-3). Aspect/Tense of the imperfective are also abstract and their denotation varies widely as will be shown in IV-2 & IV-3 (Aoun et al., 2010:24-6). Additionally, Arabic dialects have independent aspectual markers that attach to the imperfective,
like [b-], [Ŷammāl], [baʔa], the participle, the copula, and adverbial complements (Mitchell & al-Hassan, 1994).

Regarding the vocalic melody, Bahloul (2008) in contrast to Aoun et al. (2010) argues that since in MSA the first vowel of the perfective is deleted in the imperfective and transformed from [a-] to [u-] upon passivization, it can express Aspect-Tense properties (pp.29-37).

For the purpose of this study, I will adopt Comrie’s (1976, 1985) terminology for the two morphological paradigms in Arabic perfective/imperfective, and the combined tense/aspect approach of Fassi Fehri (1993).

Table 3 presents the perfective and imperfective morphemic composition of the verb in MSA and ECA. As shown, the perfective never takes a prefix. ECA does not inflect for dual, and has a combined gender in the plural. Person appears in the prefix of the imperfective and suffix of the perfective. Comparing the suffixes, we see that the second person of the imperfective is masculine singular if the suffix is null [-ø], but feminine singular or masculine plural if it is [-t] or [-ū] respectively. In MSA, feminine plural has a suffix [-nna]. We, hence, assume a theory of null morphemes, whereby, the suffix always denotes number and gender. Accordingly, in the ECA paradigm of the perfective, the 3rd singular masculine is marked by a null suffix as shown in the table. First person is neutral with respect to gender in both varieties. The prefix vowel of the imperfective differs in quality from that of MSA, and internal vowels may also differ. In the perfective, final vowels of all singular persons are omitted in ECA, except 3rd person feminine whose vowel changes from [–a] to [–i]. Finally, the suffix of the 2nd plural masculine often
transforms in ECA from [–tum] to [–tū] and that of the 3rd plural masculine becomes [um] instead of [ū].

The MSA imperfective has three moods: indicative, subjunctive and jussive. It is marked by end morphemes as shown in Table 4. Mood reflects a modal meaning that correlates to other sentential constituents. Indicative mood is marked by a suffixal –u in the singular. The dual and plural, as well as the second person singular are marked by –n. The indicative expresses either factual propositions in declarative and interrogative main clauses, non-factual future or predictions (Holes, 1995:182). The three cases are exemplified in phrases 70, 71 & 72.
Table 4 Mood marking of the imperfective verb in MSA

<table>
<thead>
<tr>
<th></th>
<th>Indicative</th>
<th>Subjunctive</th>
<th>Jussive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; sg.</td>
<td>ء-کتوب</td>
<td>ء-کتوب</td>
<td>ء-کتوب ۰</td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; pl.</td>
<td>نا-کتوب</td>
<td>نا-کتوب</td>
<td>نا-کتوب ۰</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; sg.M</td>
<td>تا-کتوب--ع</td>
<td>تا-کتوب</td>
<td>تا-کتوب ۰</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; sg.F</td>
<td>تا-کتوب-ی-نا</td>
<td>تا-کتوب-ی-۰</td>
<td>تا-کتوب-ی-۰</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; sg. dl</td>
<td>تا-کتوب-ی-نا</td>
<td>تا-کتوب-ی-۰</td>
<td>تا-کتوب-ی-۰</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; pl.M</td>
<td>تا-کتوب-ی-نا</td>
<td>تا-کتوب-ی-۰</td>
<td>تا-کتوب-ی-۰</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; pl.F</td>
<td>تا-کتوب-ی-نا</td>
<td>تا-کتوب-ی-۰</td>
<td>تا-کتوب-ی-۰</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; sg.M</td>
<td>ئا-کتوب</td>
<td>ئا-کتوب</td>
<td>ئا-کتوب ۰</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; sg.F</td>
<td>تا-کتوب-ی-ض</td>
<td>تا-کتوب-ی-۰</td>
<td>تا-کتوب-ی-۰</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; dl.F</td>
<td>تا-کتوب-ی-ض</td>
<td>تا-کتوب-ی-۰</td>
<td>تا-کتوب-ی-۰</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; dl.M</td>
<td>تا-کتوب-ی-ض</td>
<td>تا-کتوب-ی-۰</td>
<td>تا-کتوب-ی-۰</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; pl.F</td>
<td>تا-کتوب-ی-ض</td>
<td>تا-کتوب-ی-۰</td>
<td>تا-کتوب-ی-۰</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; pl.F</td>
<td>تا-کتوب-ی-ض</td>
<td>تا-کتوب-ی-۰</td>
<td>تا-کتوب-ی-۰</td>
</tr>
</tbody>
</table>

70. Mona ta-drus-u

Mona 3-study.IMP.sg.F-IND

منی تدرسُ

*Mona is studying*

71. Mona sa-ta-drus-u

Mona FUT-3-study.IMP.sg.F-IND

منی ستدرسُ

*Mona will study*
Non-factual subordinate verbs have subjunctive mood. They express possibility, obligation, purpose, result or time that is future with respect to the main verb. Singular persons are marked by a final [-a]. The subordinate verb in these cases is preceded by the complementizer أنَّ [ʔan] that. On the other hand, factual subordinate verbs are in the indicative mood, and the complementizer أنَّ [ʔanna] that subcategorizes an NP. [ʔan] and [ʔanna] are contrasted in 73 & 74. Often the subject of the NP is a pronoun that cliticizes to the complementizer, as in 74 (Holes, 1995:182).

72. Mona qad ta-drus-u

Mona may 3-study.IMP.sg.F-IND

منى قد تدرس

Mona may study

73. mina al-mustaḥil-i ?an ya-natasm-r-a

from the impossible-GEN that 3-win.IMP.sg.M-SUB

من المستحيل أنْ تدرس

It is impossible that he wins

74. ?a-ðun-u ?anna-hu ya-skun-u huna

1sg-think.IMP-IND that-3sg.M 3-live.IMP.sg.M-IND here

أظن أنه يسكن هنا

I think he lives here
Jussive mood is confined to verbs negated with لم [lam], the imperative, and the conditional (Abdul-Massih, 1981:147). All three cases share the property that the action is not factual or not fulfilled (Holes, 1995:183). We will discuss the negative and the imperative in the following section and in section IV-5 respectively. The jussive is marked by a final [-٠] null vowel in the singular. In sentence 75, both verbs in the conditional and result clauses are jussive. Here the verbs are “hollow”, i.e. their second radical is a long vowel, so they are marked jussive by shortening these vowels. Hence, تري٠ [tu-rid] you want becomes ترد [tu-rid], and تنال [ta-nal] you get becomes تنل [ta-nal].

75. ﻰن ت رد خيرا تنله
   ﻰn تٰrid xayr-ٰn ta-nal-hu

   if 2-want.IMP.sg.M.JSS 2-get.IMP.sg.M.JSS-3sg.M

   إن تُرد خيرا تنله

   If you seek goodness, you achieve it

In ECA, the indicative, subjunctive and jussive all end with the latter’s [-٠] zero mark. Similarly, the indicative final [-١] of the dual and plural is lost. Hollow verb vowels are not shortened in the jussive. Hence, different moods are indistinguishable. However, the indicative is often indicated by a proclitic [ب-]; which disappears in the subjunctive non-factual statements, or the jussive imperatives. For example, in 76, the event is factual and the verb is prefixed by [ب-], but in 78, [ب-] does not cliticize to the imperative. Also for some non-factual actions such as an unrealized event in 77, the subjunctive is zero marked.
76. Ahmad bi-yi-ktrib-∅

Ahmad IND-3-write.IMP-sg.M

أحمد يكتب

*Ahmed writes/is writing*

77. ma-ʔand-ʔiʃ wāgif ∅-ʔa-ʔmil-uh

NEG-with-1sg-NEG homework SUB-1sg-do.IMP-3sg.M

ما عنديش واجب أعمله

*I don’t have homework to do*

78. ∅-ʔi-ktrib-∅

JSS-MPV-write-2sg.M

اكتب

*Write!*

In later sections, [b-] is referred to as an aspectual prefix, and annotated as PROG or HAB rather than an IND marker.

**IV-2 Verb Negation in Arabic**

As opposed to the affirmative, negative statements are semantically regarded as more complex. They are syntactically marked by particles or clitics that falsify the proposition. Due to their
markedness and complexity, negative sentences are less frequent than the affirmative (Ul-Haq, 1984:2-5,19).

Table 5 summarizes negation markers in MSA and ECA. In ECA, they all boil down to two markers: the discontinuous ما.. ش [ma.. f] with the verb inbetween, or مش [miʃ] followed by the verb. A detailed review of each marker follows.

<table>
<thead>
<tr>
<th>MSA</th>
<th>ECA</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mood Assigners</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lam yaktub</td>
<td>ma-katab-f</td>
<td>he didn’t write</td>
</tr>
<tr>
<td>lā yaktub-u</td>
<td>ma-(b)-yiktib-f</td>
<td>he doesn’t write</td>
</tr>
<tr>
<td>lā taktub-u</td>
<td>ma-tiktib-f</td>
<td>don’t write</td>
</tr>
<tr>
<td>lā katab-wa lā qara?-a</td>
<td>ma-katab-f wa la ūarâ</td>
<td>he neither wrote nor read</td>
</tr>
<tr>
<td>lā qaddar-a Allāhu</td>
<td>Allāh la-y-ūaddar</td>
<td>may Allah not destine [it]</td>
</tr>
<tr>
<td>lan yaktub-a</td>
<td>miʃ ha-yiktib</td>
<td>he will not write</td>
</tr>
<tr>
<td>?allā yu-hāwil-a</td>
<td>ūfān ma-yhūwil-f</td>
<td>that he won’t try</td>
</tr>
<tr>
<td>kaylā/likaylā yu-hāwil-a</td>
<td>ūfān ma-yhūwil-f</td>
<td>so that he won’t try</td>
</tr>
<tr>
<td>Case Assigners</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mā katab-a</td>
<td>ma-katab-f</td>
<td>he didn’t write</td>
</tr>
<tr>
<td>Non-mood Assigners</td>
<td></td>
<td></td>
</tr>
<tr>
<td>laysa kātib-an</td>
<td>miʃ kātib</td>
<td>will not write</td>
</tr>
<tr>
<td>las-tu ?aqṣ?u-du-ū dālik-a</td>
<td>ma-?s?ud-f</td>
<td>I don’t mean this</td>
</tr>
</tbody>
</table>

Table 5 Verbal negation in MSA and ECA

Negative markers can be classified in three types (Fassi Fehri, 1993:163-6,170-1):

a. Markers that assign mood to verbal projections: They have temporal specification; either past لم [lam], present ل [lā], or future لن [lan]. [lam] assigns jussive mood, and [lan] assigns subjunctive mood. Although ل [lā] keeps the verb in its unmarked indicative mood, and is thus not a mood assigner, it is classified as such because like the other two markers, it selects
an imperfective, e.g. 79, 80 & 81. In all three markers the subject cannot intervene between the marker and the verb, e.g. 82.

79. lā ya-ʔkul-∅-u

NEG.IMP 3-eat.IMP-sg.M-IND

miʃ bi-ya-kl-∅
NEG HAB-3-eat.IMP-sg

لا يأكل

ma-b-yākul-∅-ʃ
NEG-HAB-3-eat.IMP-sg.M-NEG

ما بياكلش

He does not eat

80. lam ya-ʔkul-∅-∅

NEG.PRF 3-M-eat.IMP-sg.M-JSS

لم يأكل

ma-ʔakal-∅-ʃ
NEG-eat.PRF-3sg.M-NEG

66
The examples above demonstrate that MSA negative markers host tense, because the imperfective in used with all tenses, while markers differ. In contrast, equivalent ECA negatives have different tensed verbs. According to Aoun et al. (2010), negation of perfective and imperfective in ECA is generated between TP and VP. In the perfective, e.g. 80, the verb is sandwiched between two negative morphemes [ma-] and [-ʃ]. It cannot move to T without
merging with the intervening head NEG in order not to violate the Head Movement Constraint\(^1\).

On the other hand, as shown in 79, ECA negation of the imperfective is either discontinuous or pre-verbal. If the verb moves to T, it must merge with NEG yielding the discontinuous \([ma..\tilde{\iota}]\) form. If it doesn’t, the form is \([mi\tilde{\iota}]\)+imperfective, and tense is not lexically supported. In 81, ECA future negation is only preverbal and tense is generated below not above negation. Movement places the perfective higher than its subject, which explains why person is suffixal in the perfective but prefixal in the imperfective and future (pp.24-6).

The position of the subject relative to the verb in this analysis conforms to idiomatic expressions of God wishes, such as 83 vs. 84, where the explicit subject of the perfective form is post-verbal while that of the imperfective is pre-verbal. (Aoun et al., 2010:29)

83. rahim-a-hu

\begin{verbatim}
have mercy.PRF-3sg.M-3sg.M
\end{verbatim}

\begin{verbatim}
Allâh
\end{verbatim}

رحمة الله

*May God have mercy on him*

84. Allâh yi-rham-u

\begin{verbatim}
Allah 3-have mercy.IMP-sg.M-3sg.M
\end{verbatim}

الله يرحمه

*May God have mercy on him*

\(^1\) Head Movement Constraint: “An \(X^o\) may only move into the \(Y^o\) which properly governs it” (Travis, 1984:131), i.e. if head \(Y\) c-commands both heads \(X\) and \(Z\), and \(Z\) c-ommands \(X\), the latter cannot move to \(Y\) across \(Z\). In other words, head movement may not skip intervening heads.
In MSA, when لم [lam] and لن [lan] are followed by a series of verbs, only the first verb is ascribed the jussive or subjunctive mood, and subsequent verbs are in the default indicative. ل [lā] may occur with modals denoting prohibition when it precedes a jussive imperfective or a wish when it precedes a perfective, as in 85. ECA prohibition is formed by an imperfective and a discontinuous negation. In a construction like 86, where لام [lā] precedes a perfective, its meaning is *neither .. nor* (Ul-Haq, 1984:150).

85. لام  تا-تختلف-ُ

NEG 2-disagree.IMP-pl.M

لا تختلفوا

ما تختلفوش

*Don’t disagree*

86. لام  كتب-ا  و  لام  قرأ-ا

NEG write.PRF-3sg.M  and  NEG read.PRF-3sg.M

لا كتب ولا قرأ

لام كتب ولام قرأ

NEG write.PRF.3sg.M  and  NEG read.PRF.3sg.M
لا كتب ولا قرأ

*He neither wrote nor read*

لا [؟اللَا], [كي/لكي] [كيلَة] and [لكيلا] are combinations of أن [؟ان] that, [كي/لكي] [كيلَة]/[لكيلا] so that; and ل [لَا] respectively. They come before a subjunctive imperfective and refer to future time (Ul-Haq, 1984:153-7). They are expressed differently in ECA, for example in 87, ECA [كيلَة] is substituted by [ياَفَانِ].

87. sa-؟ا-دَحَبُ-عِ ِلِ؟َانَا كَيْلَة ِ؟َا-تَا لَخَارْ-أ

FUT-1sg-go.IMP-IND now so that 1sg-be late.IMP-SUJ

سأذهب الآن كيلا أتأخر

h-a-رُح ّدِلْوَأَتِي ِيَاَفَانِ ْمَا-تَا لَخَارَ-ج

FUT 1sg-go.IMP now so that NEG-1sg-be late.IMP-NEG

حاروح دلوقت عشان ما أتأخرش

*I’ll go now so that I won’t be late*

b. **Markers that assign case to nominal projections:** This is the copular negative [لَسْا], classified by Arab grammarians as a perfective verb. It negates a nominal sentence and assigns accusative case to the predicate nominal, adjective, or participle. In literary styles, it may negate an imperfective stative or habitual verb (Ul-Haq, 1984:183-4). The ECA equivalent of [لَسْا] in 88 is [миِّفِ], which assigns no case to the predicate.
c. **Markers that are not mood-assigners (neutral):** Being neutral means that they keep the imperfective in its default indicative mood. They negate both the perfective and imperfective; hence, tense is specified by the verb selected. This type is represented by ﻦ [mā], which before the imperfective denotes the simple present (e.g. 89), and before the perfective denotes the near or absolute past (e.g. 90). In the former case, it is equivalent to ﻦ [lā], and in the latter, the construction is equivalent in meaning to ﻦ [lam]+ jussive imperfective (Ul-Haq, 1984:170-2).

89. ﻦ ﺎ-۴۰۰-۹-۰-۰

NEG 1sg-know.IMP-IND-3sg.M

ما أعرفه

ma ۴-۴۰-۹-۰

NEG 1sg-know.IMP-3sg.M-NEG
I do not know him

90. mā ʿraf-tu-hu

NEG know.PRF-1sg-3sg.M

I did not know him

ma ʿraf-ū-ʃ

NEG know.PRF-1sg-3sg-NEG

IV-3 Constituents of the Arabic verb phrase

The verb is the only obligatory constituent in the verb phrase, since the subject is recovered from its inflection. Figure 1 is a representation of one common order of all possible verb phrase constituents.

Figure 1 Phrase structure
As explained in the previous section, MSA negation is only pre-verbal and in ECA is either pre-verbal or discontinuous with one particle before the verb and the other after it. The future or aspect markers attach to the verb from its left, and always follow the negative marker.

Examples 79 & 81 illustrate the composition of levels 1, 2, 3 in MSA and ECA. These three levels are dependent on each other. For example, one particular negative form would implicate the use of one or another form of the verb; i.e. perfective, imperfective, participle or modal (Holes, 1995:176).

The complement may be an NP as in 91 & 92, PP as in 93, or a pronominal clitic as in 94. A verb phrase may not have a complement if the verb is intransitive. Case marking on the subject and complement allows flexible word order, but the most common is SVC in expository contexts both in MSA and ECA. The VSC is prevalent in narratives in MSA, e.g. in newspaper or bulletin reporting, because the context is event-oriented, as in 91. The subject may be free-standing (explicit noun, independent pronoun, or a clause). In 91, the free-standing subject follows the verb and in 92 it precedes it. Otherwise, it is figured out from the verb inflection as in 93 (Holes, 1995:204-5). The colloquial version of 94 has the five levels strung together in one word: NEG-SUB-ASP-V-COMP-NEG.

91. \texttt{hada\textsuperscript{\textgreek{a}}-a s\textsuperscript{\textgreek{d}}id\textsuperscript{\textgreek{\textgreek{m}}}-\text{un} musallah\textsuperscript{-un} bayana l-quww\textsuperscript{\textgreek{\textgreek{w}}}-\text{t-i}}

happen.PRF-3sg.M clash-NOM armed-NOM between the-forces-GEN

حدث صدام مسلح بين القوات

\texttt{has\textsuperscript{\textgreek{\textgreek{a}}}-al-\text{o} s\textsuperscript{\textgreek{d}}id\textsuperscript{\textgreek{\textgreek{m}}} musallah b\text{\textgreek{\textgreek{n}}} il-quww\textsuperscript{\textgreek{\textgreek{w}}}}
An armed clash occurred between forces.

92. **Aly-un** yu-hibb-u l-qahwat-a

Ali 3-like.IMP-sg.M-IND the-coffee-ACC

علي يحب القهوة

Aly bi-y-hibb il-ʔahwa

Ali HAB-3-like.IMP-sg.M the-coffee

علي يحب القهوة

*Ali likes coffee*

93. ḍahab-nā ʔilay-h

go.PRF-1pl to-3sg.M

ذهبنا إليه

ruh-nā l-u

go.PRF-1p. to-3sg.M

رحنا له

*We went to him.*

94. lā ʔu-kallim-u-ha
I'm not talking to her

VOS order is possible when the complement is a pronoun. In this case, the complement enclitics to the verb or its preposition. Also, when the subject is indefinite, it may follow the complement. Both cases rely on two factors: the informational principle and end-weighting. Informationally, elements that are less known in the context are highlighted by moving them away from their normal position. End-weighting positions lexically ‘heavier’ constituents to the right of ‘lighter’ ones, because the rhythmic break in Arabic comes near the middle of the sentence at a major constituent boundary. Since pronouns refer to elements previously mentioned in the context, they are higher on the scale of definiteness than a definite noun. Consequently, object pronouns in 95 & 96 precede the subject. In 97, the indefinite subject is highlighted by placing it sentence final. It is also lexically heavier than the complement by virtue of the qualifying adjective. In 98, the informational principle conflicts with end-weighting since it is the complement that has more lexical elements, and the subject is indefinite. In this case, word order follows the informational principle, and the subject is placed sentence final (Holes, 1995:205-7).
95. یمحد-ا-هم n-nās-u fi t-tilifāz-i

watch.PRF-3sg.M-3pl.M the-people-NOM in the-TV-GEN

شاهدهم الناس في التلفاز

96. یهاحب-a ییلای-him یومار-u

watch.PRF-3sg.M to-3pl.M Omar-NOM

ذهب إليهم عمر

97. قددم-at یددراسات-ا muwazzaf-at-un gadīd-at-un

present.PRF-3sg.F the-study employe-F-NOM new-F-NOM

قدمت الدراسات موظفة جديدة

People watched them on TV

Omar went to them
A new employee presented the study

A young employee presented the new economic study

One important anamoly of various word orders in verbal phrases is the asymmetry of subject-verb agreement. Examples 99-102 show that in MSA, subjects in pre-verbal position display full agreement with the verb, but post-verbal subjects agree with the verb in person and gender, not number. However, in ECA, full agreement is possible for both word orders.
The kids slept

The kids sleep

The kids sleep
All syntactic accounts for this behavior “fall short of accounting for all the agreement factors in Arabic while remaining empirically and conceptually consistent” (Aoun et al., 2010:83). In traditional Arabic grammar, SVO is analyzed as a nominal sentence, where the predicate is a verbal phrase and the dual or plural suffix of the verb is a subject pronoun. Hence, when a verb is in initial position (verbal sentence), the pronominal subject is suppressed if followed by an NP subject in order to avoid subject reduplication. This concept of redundancy is employed in the morphological account of agreement asymmetry, which postulates that one agreement relation holds for all word orders (Benmamoun 2000:128-132). They are derived according to the configuration represented in Figure 2. The verb and subject are generated in VP. Upon verb
raising to T, VS order obtains. If the subject also moves to Spec of TP, the order becomes SV. In both orders, the verb and subject are in Spec-Head relation. Since the number feature on the verb is expressed in its suffix, it becomes redundant when the verb is higher than the subject. A single prosodic unit forms and the number suffix of the verb merges with the subject.

IV-4 Tense and Aspect

The distinction between tense and aspect in Bhat’s words is that “Tense involves time that contains the event, aspect involves time that is contained in the event” (Bhat, 1999:93). The latter is termed “event time” (E), while the former is determined with respect to a “reference time” (R) that may or may not coincide with the “speech time” (S) (deictic or non-deictic reference time) (Eisele, 1999:60). In Comrie’s definition, tense is “the grammaticalization of location in time”
(Comrie, 1985:vii), which is attained either morphologically, lexically, or by means of lexical composite expressions (Comrie, 1985:10). Reichenbach’s schema describes tense and aspect jointly in terms of the relationship between R, S and E. The relationship between R and S is linear, denoting a deictic time reference \([\pm{\text{Anterior}}]\). When R to S is combined with the embedded non-deictic time reference E to R, the relationship becomes associative \([\pm{\text{Perfect}}]\) implying aspect. The tense-aspect schema in Table 6 ensues from different R, S and E temporal relations. For example if R=S (contemporal) and E < R (E is anterior to R), then the verb is in the present perfect, and if R>S (R is posterior to S) and E=R (E is contemporal with R), then it is in the simple future (Eisele, 1999:31-5 & Fassi Fehri, 1993:149).

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>simple present</td>
<td>(S = R = E)</td>
</tr>
<tr>
<td>simple past</td>
<td>(E = R &lt; S)</td>
</tr>
<tr>
<td>simple future</td>
<td>(S &lt; R = E)</td>
</tr>
<tr>
<td>present perfect</td>
<td>(E &lt; R = S)</td>
</tr>
<tr>
<td>past perfect</td>
<td>(E &lt; R &lt; S)</td>
</tr>
<tr>
<td>future perfect</td>
<td>(S &lt; E &gt; R)</td>
</tr>
</tbody>
</table>

*Table 6 Reichenbachian model of tenses*

Eisele combines Reichenbach’s framework with Carlson’s analysis of bare plurals\(^2\) to propose his notion of Specific vs. non-Specific time reference, and show that this distinction is particularly vital to the aspectual meaning (E to R) of the present (R=S). A present tense verb like يَكُتُب [ya-ktub] (MSA) / [(bi-)yi-ktib] (ECA) to write may denote a habitual, or iterative meaning. Eisele shows that the habitual is associated with non-specific time reference and the iterative is implied when the time reference is specific (Eisele, 1999:56-9 & 91-2).

\(^2\) For an overview of Carlson’s work, see Eisele (1999:50-6)
In Arabic, time location is also conveyed by time adverbials that are either single lexical items, such as [الآن] (MSA) / [دلوقت] (ECA) now, or lexical composite expressions, such as [الماضي الأسبوع الماضي] (MSA) / [الأسبوع اللي فات] (ECA) last week. In ECA, aspect is explicitly denoted by several constituents like [b-], [؟ammāl], [؟a?a], the participle, the copula, and adverbial complements (Mitchell & al-Hassan, 1994).

If verb affixes in Arabic denote tense, then their ordering with respect to aspect, and mood morphemes relative to the verb base agrees with the general consensus that “the aspect marking tends to occur closer to the verb, with tense markers next to them, and mood markers forming the outermost constituent” (Bhat, 1999:159). This hypothesis is based on the view that tense, aspect, and mood act like operators that modify three layers of the verb clause: (1) Nuclear: the inner temporality of the event or state modified by aspect. (2) Core: valency of the verb. (3) Periphery: all remaining constituents, e.g. spatio-temporal setting which is under the scope of tense. Due to different modal distinctions, mood is partially concerned with the core and partially with the periphery (Bhat, 1999:159-60).

The ordering of verbal morphemes in Arabic agrees to this proposition. In example 103, the closest morpheme to the verb base is [ya-]/[y],[-u] that denote imperfective aspect, number, and gender. The future tense marker [sa-]/[ha-] comes next. The outermost operator is that of the mood [bi-l-ta?kīd].

103.  

| bi-l-ta?kīd-i | sa-ya-sīl-ū-na | layl-an |
| with-the-certainty | FUT-3-arrive.IMP-pl.M -IND | night-ACC |
بالتأكيد سيصلون ليلا

bi-l-taʔkid  ha-y-wsʕal-u  bi-l-lēl

with-the-certainty  FUT-3-arrive.IMP-pl.M  in-the-night

بالتأكيد حيصلوا بالليل

Certainly the will arrive at night.

IV-4-A Tense

1- Absolute tense

Location of the event in time requires a reference point. Absolute tense specifies the present point as the reference/deictic center. The event of “eating” in examples 104, takes place prior to the time of speaking, and in 105, the event is simultaneous with the time of speech.

104. ʔakal-tu

ʔakal-t-∅

eat.PRF-1sg

أكلت

I ate.

105. ʕa-akul-u

1sg-eat.IMP-IND

أكل

b-a-akul
In certain contexts, the imperfective emphasizes the occurrence of the event rather than its
temporal interpretation. This is common in reporting and narrating events, as in the example 106.
The goal may have already been scored, and the commentator is re-describing how it took place.

106. wa yu-saddid-ø-u hadaf-a al-fawz-i
and 3-score.IMP-sg.M -IND goal-SBJ the-winning-GEN

wi y-saddid-ø hadaf il-fawz
and 3-score.IMP-sg.M goal the-winning

ويسدد هدف الفوز

And he scores the winning goal.

The imperfective with past tense interpretation

The imperfective can express past time, particularly in narrative, journalistic, and academic
writing. This is illustrated in example 107 of a journalistic headline, which says that Abbas did
already reject the warning, but the verb form is imperfective.
Abbas rejects Netanyaho's warnings.

The perfective with present tense interpretation

Sometimes, the perfective has a present tense interpretation. In 108, the time location of the event almost coincides with the moment of speaking, but the verb فهمت [fahim-ta] you understood/have understood is perfective. The fact that the perfective in this context collocates with the adverbial الآن [ʔalʔāna /dilwaʔti] now and not منذ قليل [mundu qalîlin /min ūwayyya] a while ago is evidence for this interpretation (Bahloul, 2008:51-4). Moreover, the perfective in this structure may be replaced by the imperfective in MSA, and the participle that denotes ongoing states in ECA, e.g. 109, without affecting the intended meaning.

108. ya-gib-⁰-u  ūn ta-ntahiy-⁰-a  min hāda l-⁰-amal-i

3M-should.IMP-sg.-IND  that 2-finish.IMP-sg.M-SBJ from this the-work-GEN

l-ʔāna / *mundu qalîl  hal fahim-ta

now / a while ago  Q understand.PRF-2sg.

يجب أن تنتهي من هذا العمل الآن / *من شوية ، هل فهمت ؟

lāzim ti-xallas⁵-⁰  iʃ-ʃuʃl  da dilwaʔti /min ūwayyya fihim-t

must 2-finish.IMP-sg.M the-work this now /*a while ago understand.PRF-2sg.M
You have to finish this work today, have you understood?

109. **ya-gib-š-u**  ذَٰلِكَ ۤتَنْتَهِيِّ-ْٰذَا ۤمِن ۤهَٰذَا ۤالْمَّعَامِلِ
3M-should.IMP-sg.IND that 2-finish.IMP-sg.M-SBJ from this the-work-GEN

ليشَّنَّا ۤهَٰلَ ۤتَفْهِمِّ-ْٰذَا
now Q 2-understand.IMP-sg.M -IND

لاَذِمَ ۤتُخَلَّصُ ۤالْمَهْلَ ۤتَنْتَهَٰذَا ۤمِن ۤهَٰذَا ۤالْمَّعَامِلِ؟

You have to finish this work today, do you understand?

110. **lāzīm ۤتَخَالَسُْٰ-ْٰذَا ۤيَفْعَلُ ۤذَا ۤدِلْوَٰا؟ۤتَفْهِمُ-ْٰذَا
must 2M-finish.IMP-sg the-work this now understand.PCP-sg.M

لاَذِمَ ۤتُخَلَّصُ ۤالْمَهْلَ ۤتَنْتَهَٰذَا ۤمِن ۤهَٰذَا ۤالْمَّعَامِلِ؟

You have to finish this work today, do you understand?

The perfective and imperfective with future tense interpretation

It is common to use the perfective with future interpretation in hypothetical and wish clauses as in 111 & 112 (Bahloul, 2008:60-1). The use of the perfective emphasizes the completeness (aspect) of the events considered, while locating them in future time (tense). In ECA, the verb in the result clause of conditional sentences, e.g. 111 is usually in the future form. Likewise, ECA wishful phrases use the imperfective, rather than the perfective, as in 112.
The future with modal value

In addition to their temporal function, future markers [sawfa]/[sa-]/[rāḥ]/[ha] have other modal values such as prediction, intention, certainty, and logical entailment (Bahloul, 2008:118). The colloquial future marker راح [rāḥ] in 113 is literally the verb to go in the perfective. According to Comrie (1976:106), this form incorporates motion from, similar to the English “he is going to
“eat”, as opposed to motion *towards*, as in the French “je viens de manger” *I have just eaten*.

Both, in his analysis, are directional expressions, the former having a prospective meaning, while the latter has a perfect one.

113. **sa/sawfa ʔa-akul-u**

<table>
<thead>
<tr>
<th>FUT</th>
<th>1sg-eat.IMP-IND</th>
</tr>
</thead>
<tbody>
<tr>
<td>سـ / سوف آكل</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FUT</th>
<th>1sg-eat.IMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>ح / راح آكل</td>
<td></td>
</tr>
</tbody>
</table>

*I will eat.*

**Unbound tense**

In proverbs, such as 114, the verb tense is unbound; hence the basic function of the imperfective is independent of its eventive value as well as of the time of speaking (Bahloul, 2008:106-28). Time in 115 & 116 is also unbound implying a degree of factuality. The two sentences are equivalent to a conditional structure without a conditional particle, where the main verb is imperative and the result is imperfective (Eisele, 1999:78). Similarly, the proverb in 117 is timeless, and the perfective (MSA) and imperfective (ECA) have a gnomic interpretation (Bahloul, 2008:57).
114. wa nahnu na-ḥtarim-u l-qānūn-a
and we 2pl-respect.IMP-IND the-law-SBJ

ونحن نحترم القانون

wi ḥna b-ni-ḥtirim il-qānūn
and we HAB-2pl-respect.IMP the-law

واحنا بنحترم القانون

And we respect the law.

115. xālif-∅-∅ tu-ḥraf-∅-∅
be different.MPV-2sg.M-JSS 2-know.IMP.PSS-sg.M-JSS

خالف تعرف

Be different, and you’ll be known.

116. fattah-∅ ūn-ak t-ākul-∅ malban

فتح عينك تأكل ملبین

Open your eyes, and you’ll eat ‘malban’.

117. man zaraḥ-∅ ḡasʿad-∅
who plant.PRF-3sg.M reap.PRF-3sg.M

من زرع حصد
2- **Relative tense**

The deictic center of relative tense is determined by the context (Comrie, 1985:58). Thus, in example 118, the reference point is in the future because the result event has a future form, and although the verb in the conditional clause has a perfective form, it is understood to take place in the future prior to the reference point.

118. ُ؟ِذَا ُتا-ُ؟ِاَخَار-ْتا ُسا-ُيَا-ُدُب-ُأ-ْع

<table>
<thead>
<tr>
<th>if</th>
<th>2-be late.PRF-sg.M</th>
<th>FUT-3-be angry.IMP-sg.M-IND</th>
</tr>
</thead>
</table>

إذا تأخرت سيغضب

If you are late, he’ll be upset

Example 119 shows that a different context makes the same verb form تأخرت [ta?axxarta]/[t?axxart] you were late refer to the habitual present prior to the reference point, which in this case, is the result event in the habitual present.
Examples 120, 121 & 122 illustrate the tense neutralization rule: “Only the first verb of a sequence within a sentence shows the expected tense, while all subsequent verbs are in a single tense category irrespective of the tense of the first verb” (Comrie, 1985:102). In 120, the form of نكلام[natakallamu]/[nitakallam] we speak is imperfective, but its progressive meaning is restricted to the past tense, because the first verb جلسنا[galasna]/[ašasna] we sat is perfective and located in past time. In contrast, time interpretation of the second verb in 121 is the present & in 122 it is the future, in accordance with the time of the first verb.

120.  

\textit{We sat talking(\textit{and talked})}
The category of tense in Arabic has evidence in negation distinctions. Bhat (1999) maintains that “there is a correlation between the prominence of verbal categories and the kind of morphemic distinction in the negation” (p.179). In Lewo, which is a mood prominent language, negative markers correlate with realis and irrealis moods, while in MSA, there are three negative markers لَا [lā], لَم [lam], and لِن [lan] for present, past, and future respectively.

IV-4-B Aspect

Given that aspect describes the internal structure of the event, then the first step in studying aspect is to realize that any situation may either be stative or dynamic. Verbs denoting states like أَعَرِف [ʔaˈrif] I know are static, i.e. they do not change except in response to an external factor. A static situation, however, may have a starting and ending point in its internal structure. For example, the presence of the interval time adverb ثَلَاث سَنَوَات [ʔalād sanawāt] three years in 123 forces the interpretation that the speaker no longer “knows them” (Eisele, 1999:66). Other verbs
are dynamic, e.g. [أتكلم] I speak and they denote a process internally composed of different phases (Comrie, 1976:13).

123. ُنارف-تُهم ٍلي-ّدالاد-ى ٍساناوات-ى

know.PRF.1sg-3pl.M for-three-GEN years-GEN

عرفتُهم لثلاث سنوات

I knew them for three years

As mentioned in section IV-1, Arabic verbs are either perfective or imperfective. In what follows, I will illustrate how various aspects are attained using these two forms individually, with the aid of an auxiliary, or in periphrasis.

1- The Imperfective

The habitual and iterative aspects of static or dynamic verbs are imperfective. The habitual indicates that the situation stretches over time. Intermissive occurrences are iterative (Comrie, 1976:26). Example 124 says that the state of “knowing” has started sometime in the past and continues to hold until the present time, possibly extending to the future. In MSA, states are denoted by the imperfective form, but in ECA, the participle is more natural. States that indicate a characteristic quality of the subject, as in 125, are denoted in ECA by a b-prefixed imperative (Eisele 1999:93-4). The use of the participle vs. the imperfective in ECA will be further discussed in a separate subsection below.
For imperfective dynamic situations, there is no morphological distinction between the habitual, iterative or progressive aspects in either MSA or ECA. The b- prefix is characteristic of ECA imperfectives. Time adverbials are added in examples 126 and 127 in order to mark the context more explicitly.

124. ئا-یریف-ع muṣkilat-a-ka
    1sg-know.IMP-IND problem-ACC-2sg.M
    أعرف مشكلتك

؟؟اریف miṣkilt-ak
    know.PCP.sg.M problem-2sg.M
    عارف مشكلتك

I know your problem

125. ئانا b-ا-یراف a-ىرا
    I HAB-1sg-know.IMP 1sg-read.IMP
    أنا بالعرف أقرأ

I know how to read.

أنا باعرف أقرا
b-a-tkallim  maʕā-ha  dilwaʔti
PRG-1sg-speak.IMP  with-3sg.F  now
باتكلم معاها دلوقت

I'm speaking with her now

127. ʕa-takallam-u  maʕa-ha  marrat-an  kull-a  ʔusbūʕ-in
1sg-speak.IMP-IND  with-3sg.F  once-ACC  every-ACC  week-GEN
أتيكلم معها مرة كل أسبوع

The b-prefix is dropped in ECA in several contexts, for example, when the time reference is unspecific as in 128, or when its presence is redundant because it comes in co-ordination with a preceding [b-] or [ha-] as in 129 (Eisele, 1999:81-2).

128. dāyman  ti-ʕmil-∅  muʃkila  min  mafīj
always  2-make.IMP- sg.M  problem  from  nothing
دايما تعمل مشكلة من مفيش

You always make a problem out of nothing
129. bi-yi-s̱ha is-sā́a saḇa w yu-xrug-Ø tamanya

HAB-3-wake up.IMP.sg.M the-hour seven and 3-go out.IMP-sg.M eight

بيصحى الساعة سبعة ويخرج تمانية

*He wakes up at seven o’clock and goes out at eight*

Moreover, the b-prefix may optionally be dropped with verbs that indicate a change, start or termination of a dynamic situation: بقى [yi-baʔa] *to be/* become, بئٌى [yi-btdi] *to start*, and بطل [yi-baṯa’al] *to stop*, or when the continuous or iterative occurrence of the event is emphasized by means of periphrasis with some verbs that mean *to keep*, e.g. بقعد [yu-ʔud]/[yu-ʔʔud] *to sit*, بفضل [yi-fḏal] *to remain*, بيتنه [yi-tannu] *to extend*³. Comrie (1976) considers that the function of such auxiliary verbs is locative. The colloquial version in example 130 resembles the Italian construction: “I stand-singing” *I’m singing*, or the older English use of pre-verbal locative prepositions: “He’s been a-singing” (pp.102-6). Some equivalent verbs in MSA are: بظل [ya-ḏa’al]/ لازال [lā ya-ẕal] *to remain*.

130. ya-ḏa’al-Ø yu-karrir-Ø-Ø ḥatta na-mall-a

3-keep.IMP-sg.M-IND 3-repeat.IMP-sg.M-IND until 1pl-be bored.IMP-ACC

بيظل يكرر حتى نمل

(b-)yu-ʔʔud-Ø yi-karrar-Ø lahad lamma ni-zhza?
(HAB-)3-sit.IMP-sg.M 3-repeat.IMP-sg.M until when 1pl-be bored.IMP

He keeps repeating until we get bored.

While some languages use complete or partial reduplication of the verbal base to convey the iterative and continued occurrence of the same event (Bhat, 1999:55), the Arabic verb form تفَعَّل (t V C₁V C₂C₂VC₃) expresses this meaning (?ismā‘īl, 1918:35). Examples 131 & 132 contrast the lexical meaning of this form to the basic form فعل (C₁V C₂V C₃).

131. ta-ftah--Ø-U  I-mahillāt-U  baṣḍa s’alāt-i  l-gumūṣat-i
3-open.IMP-sg.F-IND  the-shop.pl-NOM  after  prayer-GEN  the-Friday-GEN

تفتح المحال بعد صلاة الجمعة

il-mahallāt bi-ti-ftah-Ø  baṣḍ  s’estāt  ig-gumṣa
the-shop.pl  HAB-3  -open.IMP-sg.F  after  prayer  the-Friday

المحلات تفتح بعد صلاة الجمعة

Shops open after Friday prayer.

132. ta-tafattah-Ø-U  z-zuhūr-U  fi  r-rabī‘-i
3-blossom.IMP-sg.F-IND  the-flower.pl-NOM  in  the-spring-GEN

تتفتح الزهور في الربيع

?iz-zuhūr  bi-ti-tfattah-Ø  fi  r-rabī‘
the-flower.pl  HAB-3-blossom.IMP-sg.F  in  the-spring
Flowers blossom in the spring

Reduplication is present in Arabic in a few examples such as يتمتم [ya-tamalmal] fidget (MSA), whose basic form is يمل [ya-mal] to be bored, and يبلبع [yi-balba] gulp (ECA) from يبلع [yi-bla] to swallow.

2- The Perfective

The perfective emphasizes that the action is finalized regardless of the internal structure of the event. In Arabic, the internal structure is implied by the lexical meaning of the verb. The lexical meaning of the verb, i.e. the nature of the event, determines whether each E and R is a moment in time or an interval. Example 133 is unacceptable or awkward, because E and R of عرفت [Irif] to realize are essentially moments. Conversely, اشتغل [Itayal] to work in 134 can denote an E interval and R moment (Eisele, 1999:65-6).

133.  \(\text{Jaraf-tu} \quad \text{il-mu\\-jkilat-a} \quad \text{safat-ayni} \)
realize.PRF-1sg the-problem-ACC hour-dl.ACC

*\(\text{Irif-t} \quad \text{il-mu\\-jkila} \quad \text{sait-\-en} \)
realize.PRF-1sg the-problem hour-dl

عرفت المشكلة ساعتين

*I realized the problem two hours
I realized the problem for two hours

As noted before, the perfective usually relates to the past. The perfective is also used to refer to some past situation that resulted in the present situation, or is temporally close to the present time (Comrie, 1976:56-60). In non-deictic notation, the first case is $E<R$ and the second is $E=R$ (Eisle, 1999:65-6). The state of “happiness” in 135 is a result of the “success” of the negotiations. The event of coming in 136 took place in the recent past. Recency is made more explicit by means of time adverbials. Persistence and temporal closeness is more frequently denoted by the participle in ECA and by the verbal particle قد [qad] already in MSA.

I started working at two o’clock

As noted before, the perfective usually relates to the past. The perfective is also used to refer to some past situation that resulted in the present situation, or is temporally close to the present time (Comrie, 1976:56-60). In non-deictic notation, the first case is $E<R$ and the second is $E=R$ (Eisle, 1999:65-6). The state of “happiness” in 135 is a result of the “success” of the negotiations. The event of coming in 136 took place in the recent past. Recency is made more explicit by means of time adverbials. Persistence and temporal closeness is more frequently denoted by the participle in ECA and by the verbal particle قد [qad] already in MSA.
I'm happy that the negotiations have succeeded.

136. (qad)  giʔ-t-u  litawwi

(already)  come.PR-1sg-IND  just

جئت لتوي

?ana  lessa/ya dōb  gāy-ø

I  just  come.PCP.sg-M

أنا لسه/يادوب جاي

I’ve just come

Rather than denoting recentness, the function of قد [qad] may be emphatic. Alternatively, it may have a telic interpretation, whereby it emphasizes that the event has terminated (Bahloul, 2008: 72-7). Hence, three interpretations are possible for the sentence in 137 as shown.

137.  qad  waqqaʕ-na  l-ʕaqd-a

----  sign.PR-2pl  the-contract-ACC

قد وقعنا العقد

We have “already/just” signed the contract: recentness

We have “finished” signing the contract:  telicity

We have “indeed” signed the contract: emphasis
Examples 138 & 139 show that the perfective in Arabic does not distinguish between an event that occurred at some fixed time in the past (simple past) and an event that occurred in the indefinite past (perfect). The distinction is only inferred from the context or time adverbials.

138. zur-na ?aswān-a asfayf-a l-mādī
visit.PRF-1pl Aswan-ACC the-summer-ACC the-past

زرنا أسوان الصيف الماضي

zur-na ?aswān isfēf illi fāt
visit.PRF-1pl Aswan the-summer that pass.PRF.3sg.M

زرنا أسوان الصيف اللي فات (ECA)

We visited Aswan last summer.

139. zur-na ?aswān-a min qablu
visit.PRF-1pl Aswan-ACC from before

زرنا أسوان من قبل

zur-na ?aswān ?abl ke-da
visit.PRF-1pl Aswan before as-this

زرنا أسوان قبل كده

We visited Aswan before.
3- The Future

Prospective aspect predicts a future event based on a present state or event. As mentioned earlier, it is marked by the future lexemes سوف [sawfa] (MSA)/راح [rāh] (ECA) or prefixes [sa-] (MSA)/[ha-] (ECA) before the imperfective. In MSA, verbs of commencement like كاد [kāda] /أوشك [awška] be on the verge of, and the verb of desire عسي [āsā] may/perhaps are also prospective. Both the perfective and imperfective forms of this class of verbs denote the near future, but the verb that follows them (the predicted event) is always imperfective as in e.g. 140 (Abdul-Massih, 1981). The equivalent form of the first two verbs in ECA is قريب [arrab] to be close to.

140. ُyawšak-na / n-ūšik-u ?an na-gid-a
be about to.PRF-2pl /2pl-be about to.IMP-IND that 2pl-find.IMP-SBJ

ُhall-an li-hādhīhi l-muṣkilat-i
solution-ACC to-this problem-GEN

أوشكنا أن نجد حل لهذه المشكلة

ُ؟arrab-na n-lāʔi ُhall li-l-muṣkila di
be close.PRF-2pl 2pl-find.IMP solution to-the-problem this

 قريبنا نلقي حل للمشكلة دي

We are about to find a solution to this problem.
4- The Auxiliary verb [kāna]

More complex constructions combine the auxiliary [kāna] with the imperfective and perfective in different ways to express perfect or perfect progressive aspects in the past or future. Starting with the basic structure of the habitual in 141 & 142, other aspects 143-149 are constructed. The role of [kāna] in these examples is to place the subordinate verb in the past or future relative to the main verb. Also, note the role of [qad] already in MSA to denote perfect aspect (141, 144, 145, 147).

Habitual:

a. main V (imperfective) → AUX (imperfective) + qad + Event (perfective)

141. ḥīna ṭa-Ŷūd-u ya-kūn-ṣ-u qad nām-a

when 1sg-return.IMP-IND 3-be.IMP-sg.M-IND already sleep.PRF-3sg.M

حين أعود يكون قد نام

lamma b-a-rgaʕ b-y-kūn-Ø nām-Ø

when HAB-1sg-return.IMP HAB-3-be.IMP-sg.M sleep.PRF-3sg.M

لما بارجع بيكون نام

When I come back, he would have slept.

b. main V (imperfective) → AUX (imperfective) + Event (imperfective)
Past progressive

Past progressive of dynamic verbs is formed by the use of the perfective of the auxiliary verb [kāna] to be with the imperfective of the dynamic verb as in 143.

c. main V (perfective) → AUX (perfective) + Event (imperfective)

143. kun-tu  ?a-takallam-u maʃ̣a-ha  hīna daxal-a  sāmī
be.PRF-1sg  1sg-speak-IND with-3sg.F when enter.PRF-3sg.M Sami

I was speaking with her when Sami came in.
Past perfect:

d. main V (perfective) → AUX (perfective) + qad + Event (perfective)

<table>
<thead>
<tr>
<th>144. hīna</th>
<th>ūd-tu</th>
<th>kān-a</th>
<th>qad</th>
<th>nām-a</th>
</tr>
</thead>
<tbody>
<tr>
<td>when return.PRF-1sg</td>
<td>be.PRF-3sg.M</td>
<td>already sleep.PRF-3sg.M</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When you returned, he had already slept.

By the time I was free, he had been asleep.

Past perfect progressive:

As noted in section IV-3-a, the progressive meaning in 145 is established by periphrasis. The subordinate verb denoting the progressive event is preceded by a dynamic verb أخذ [ʔaxaʔ-a] to persist (MSA) and قعد [ʔaʔad] to sit (ECA).

e. main V (perfective) → AUX (perfective) + qad + Event [AUX (perfective) +imperfective]

<table>
<thead>
<tr>
<th>145. hīna</th>
<th>ūd-tu</th>
<th>kān-a</th>
</tr>
</thead>
<tbody>
<tr>
<td>when return.PRF-1sg</td>
<td>be.PRF-3sg.M</td>
<td></td>
</tr>
</tbody>
</table>

When you returned, he had been sitting.
Future progressive:

Here, note the use of [bi-] in ECA to denote the progressive. In MSA, the progressive aspect is implied by the semantics of the event.

f. main V (imperfective) → AUX (prospective) + Event (imperfective)

146. ḫīna ʔa-ʃūd-u   sa-ya-kūn-ø   y-drus-ø
when 1sg-return.IMP-IND FUT-3-be.IMP-sg.M-IND 3-work.IMP-sg.M-IND
حين أعود سيكون يدرس

lamma a-rgaʃ ḥa-y-kūn-ø bi-y-zākir-ø
when 1sg-return.IMP FUT-3-be.IMP-sg.M PRG-3-study.IMP-sg.M
لما ارجع حيكون بياذكر

*When I’m back, he will be studying.*
Future perfect:

g. main V (imperfective) → AUX (prospective) + qad + Event (perfective)

147. ʰīna ʔa-ʃūd-u sa-ya-kūn-∅-u qad nām-a
   when 1sg-return.IMP-IND FUT-3-be.IMP-sg.M-IND already sleep.PRF-3sg.M

حين أعود سيكون قد نام

lamma ʔ-rgaʃ-∅ ha-y-kūn-∅ nām-∅
   when 1-return.IMP-sg.M FUT-3-be.IMP-sg.M sleep.PRF-3sg.M

لما أرجع حيكون نام

When I'm back, he will have slept.

Future perfect progressive:

Like past perfect progressive, a dynamic verb is required to express the persistence of the event.

h. main V(imperfective) → AUX(prospective)+qad+Event [AUX (perfective)+imperfective]

148. ʰīna ʔa-ʃūd-u sa-ya-kūn-∅-u
   when 1sg-return.PRF-IND FUT-3-be.IMP-sg.M-IND

qad ʔaxaʃ-∅ ya-drus-∅-u
   already persist.PRF-3sg.M 3-study.IMP- sg.M-IND

حين أعود سيكون قد أخذ يدرس
When I came back, he will have been studying.

With certain MSA adverbials, e.g. [؟يلا هنين by the time when], either the imperfective subordinate verb or its verbal noun may be used to denote the event, as in the following example:

149. nu-gammid-u tašāmulāt-i-na ʔila hini

2pl-freeze.IMP-IND dealing.pl-ACC-2pl to when

ta-tahassan-ʔ-u 1-ʔawdāf-ʔ-u

3.F-improve-sg.F-IND the-state.pl-NOM

We freeze our dealings until the situation improves.

nu-gammid-u tašāmulāt-i-na ʔila hini

2pl-freeze.IMP-IND dealing.pl-ACC-2pl to when

tahassun-i 1-ʔawdāf-ʔ-i

3.F-improve-sg.IND the-state.pl-NOM

We freeze our dealings until the improvement of the situation.
5- The active participle

The participle in Arabic is termed اسم الفاعل [?ismu l-fā’il] doer of the action/agent nominal. The term reflects its composite nature with nominal, verbal, and adjectival attributes. There are deverbal participles that refer to states, and nominal participles that are agentive. A deverbal participle is made dual or plural by means of suffixes, but nominal participle plurals may either take suffixes or are broken with infixes. Only nominal participles occur in subject position, take a definite article or get modified by adjectives (Mughazy, 2004:25-33). Table 7 shows that deverbal participles agree with gender and number only, not person. As shown, there is no dual in ECA, no gender agreement for plurals, and no case marking. The participle كاتب [kātib] writer is derived from the root [k t b]. The stem is of the form CVVCV: كتب [katab]. This is the simplest of the ten forms of Arabic verbs, and all its participles are of the form CVVCV.

<table>
<thead>
<tr>
<th></th>
<th>MSA</th>
<th>ECA</th>
</tr>
</thead>
<tbody>
<tr>
<td>sg.M</td>
<td>kātib-Ø-u</td>
<td>kātib-Ø-a</td>
</tr>
<tr>
<td></td>
<td>kātib-at-u</td>
<td>kātib-at-a</td>
</tr>
<tr>
<td>sg.F</td>
<td>kātib-Ø-āni</td>
<td>kātib-Ø-āni</td>
</tr>
<tr>
<td></td>
<td>kātib-at-āni</td>
<td>kātib-at-āni</td>
</tr>
<tr>
<td>dl.M</td>
<td>kātib-Ø-ūna</td>
<td>kātib-Ø-ūna</td>
</tr>
<tr>
<td></td>
<td>kātib-Ø-Ūna</td>
<td>kātib-Ø-Ūna</td>
</tr>
<tr>
<td>dl.F</td>
<td>kātib-Ø-ūnā</td>
<td>kātib-Ø-ūnā</td>
</tr>
<tr>
<td></td>
<td>kātib-Ø-Ūnā</td>
<td>kātib-Ø-Ūnā</td>
</tr>
<tr>
<td>pl.M</td>
<td>kātib-Ø-ūna</td>
<td>kātib-Ø-ūna</td>
</tr>
<tr>
<td></td>
<td>kātib-Ø-ūna</td>
<td>kātib-Ø-ūna</td>
</tr>
<tr>
<td>pl.F</td>
<td>kātib-Ø-ūnā</td>
<td>kātib-Ø-ūnā</td>
</tr>
<tr>
<td></td>
<td>kātib-Ø-Ūnā</td>
<td>kātib-Ø-Ūnā</td>
</tr>
<tr>
<td>NOM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEN</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7 Active participle subject and case agreement in MSA & ECA

Like nouns, participles take case only in MSA. All case is word final. The nominative, accusative, and genitive markers are [-u], [-a], and [-i] respectively for the singular. Nominative
dual ends with [-ayni] and the plural with [-ūn]. The accusative and genitive both end with [-ayni] for dual, and [-īn] for plural.

Examples 150 & 151 contrast deverbal and nominal participles derived from the root عَمْلُ [ʕ-ml-] work/do. [ʕummāl] and [ʕāmil-āt] are nominal plural participles; while [ʕāmil-īn] is a deverbal plural participle. In ECA, deverbal participles differ aspectually from the b-prefixed imperfective. One difference is that the participle [ʕāml-īn] in 150 & 151 denotes an ongoing activity, but in 152 the b-prefix is habitual or iterative meaning that the strike is current, but that it is a customary and occasional.

150. al-ʕummāl-u  al-gudud-u   ya-ʕmal-ū-na  ʔidˤrāb-an
    nouvelles travailleurs se mettent en grève
    
151. al-ʕāmil-āt-u al-gādīd-āt-u ya-ʕmal-na  ʔidˤrāb-an
    the-worker-pl.F-NOM the-new-pl.NOM 3-do.IMP-pl.F strike-ACC
    les nouvelles travailleuses se mettent en grève
    
*The new male workers are going on strike.*
The new female workers are going on strike.

The new workers go on strike.

Deverbal participles copy the syntactic properties of their source verb. They inherit the two arguments of the verb and take similar sentential complements as in 153, prepositional complements as in 154, and non-finite complements as in 155 (Mughazy, 2004:34-48).

Sentential complement

153. huwa munkir-∅-un ?anna-hu qad ruʃiy-a

3sg.M deny.PCP-sg.M-NOM that-3sg.M already bribe.PSS.PRF-3sg.M

155. huwwa munkir-∅ inn-u rtaʃa-∅
He denies/is denying that he has been bribed.

Prepositional complement

Note that in this example, the semantic meaning of the participle and its verb differs. The prepositional verb means that the person “is giving his testimony”. The basic meaning of the participle is that the person is a “witness” of the incident. He may or may not be testifying at the present moment.

154. huwa  jahid-Ø-un  ġala  mā  ḥadaθ-Ø
3sg.M  witness.PCP-sg.M-NOM  on  what  happen.PRF-3sg.M

huwa  ya-ʃhad-Ø-u  ġala  mā  ḥadaθ-Ø
3sg.M  3- testify.IMP-sg.M-IND  on  what  happen.PRF-3sg.M

هو شاهد/يشهد على مث ما حدث

huwwa  jahid-Ø  ġala  lli  has’il-Ø
3sg.M  witness.PCP-sg.M  on  that  happen.PRF-3sg.M
He is a witness of what happened. / He testifies as a witness on what happened.

Non-finite complement

In this example, the aspects of the participle and verb differ. The participle denotes an ongoing process, while the verb in MSA and ECA is habitual.

155. **huwa xārig-Ø-un**  **ya-ṭari-Ø-va**  **ʔasyā?-a**


**huwa ya-xrug-Ø-u**  **ya-ṭari-Ø-va**  **ʔasyā?-a**


هو خارج/يخرج يشتري أشياء

**huwwa xārig-Ø**  **yi-ṭiri**  **hāq-āt**


**huwwa bi-yu-xrug-Ø**  **yi-ṭiri**  **hāq-āt**
3sf.M HAB-3-go out.IMP-sg.M 3.buy.IMP-3sg.M thing-pl

هو خارج/بيخرج يشتري حاجات

*He is going/goes out to buy something.*

Negation

In spite of the many similarities between verbs and their derived participles, they differ in negation because of the nominal character that distinguishes participles. The participle is negated like nouns and adjectives by ليس [laysa] (MSA)/مش [miʃ] (ECA), but the imperfective by لا [lā] (MSA)/ما...ش [ma..ʃ] (ECA)/مش [miʃ] (ECA).

In example 156, both the participle and verb have the same aspectual interpretation which is either habitual or iterative.

156. heya lays-at nāgih-at-an

3sg.F not to be.PRF-3sg.F succeed.PCP-sg.F-ACC

هي ليست ناجحة

heya là ta-ngah-š-u

3sg.F NEG 3-succeed.IMP-sg.F-IND

هي لا نتجح (MSA)

heyya miʃ nagh-a
3sg.F NEG succeed.PCP-sg.F

هي مش ناجحة

heyya ma-b-ti-ngah-o-ʃ

3sg.F NEG-HAB(ITR)-3-succeed.IMP-sg.F-NEG

هي مابتنجح

She isn’t successful/does not succeed.

It is clear from the above examples that the temporal and aspectual interpretation of verbs and their derived participles are not always equivalent in ECA and MSA. The present perfect interpretation of the main clause in 157 is allowed only with the participle. The imperfective in 158, on the other hand, describes a current state or event. Thus, modifying the complement by a relative clause that locates the event in the past renders the sentence ungrammatical (Mughazy, 2004:68).

157. zayd-un kātib-ø-un risālat-an ḥanū-ø-ha ḥams-i


زيد كاتب رسالة أنهاها أمس

zayd kātib-ø risāla xalasֿ-ha mbārih

zayd write.PCP-sg.M letter finish.3sg.M-3sg.F yesterday

محمد كاتب رسالة خلصها أمس
Zayd has written a letter that he finished yesterday.

Not only do participles differ from verbs in their tense and aspect, but time interpretation of the participle itself is ambiguous. Mitchell & al-Hassan (1994) state that “the participle in Egypt carries the implication of the current relevance of past acts that in general terms characterizes the perfect aspect” (p.78). Mughazy (2004) analyzes active participles as non-verbal complex predicates, and concludes that “they can be interpreted as describing a current state, whether a target or a resultant state, or as describing a proceeding or a future event that is relevant to the speech context” (p.221). The collocation of different adverbials with the same participle, [mrawwah] going home, in examples 159, 160 & 161 indicates that the participle by itself carries the meaning of a near or distant past action that is still in effect, or is about to start. For this reason, Arab Kufan grammarians considered active participles to have “permanent tense”, so that their tense is specified only by means of time adverbials (Mughazy, 2004:62).
159. ?ana mrawwah (min badri)

أنا مروح (من بدي)

I went home (long before).

160. ?ana mrawwah (lessa dilwa?ti)

أنا مروح (لسه دلوقت)

I’ve gone home (just now).

161. ?ana mrawwah (ba?de ðwayya)

أنا مروح (بعد شوية)

I’ll go home (after a while).

Fassi Fehri (1993) denies that participles have tense value. They only have an aspectual content that derives from the contingency of their source verb. In contrast to the examples given above, no participle can be derived from a non-contingent verb that is temporally unrestricted; such as مرض [marid’a] he became sick. That participles are essentially contingent and have no tense denotation is supported by the observation that, unlike imperfectives, they do not form complex tenses with [kāna] as shown in the contrast between 162 & 163. The perfective auxiliary puts the participle in a past time prior to the other action جاء [gāʔa] he came, which means that the participle has no tense of its own. On the other hand, the corresponding imperfective has an alternative interpretation that the process of ‘sitting’ was taking place simultaneously with the embedded verb (p.182-3). This distinction is evident in the ECA equivalents of both sentences.
While in MSA, the imperfective is ambiguous as to whether it denotes past perfect, or past progressive, the ECA equivalent in 162 only denotes the progressive meaning.

162. **kun-na**  
**na-glis-u**  
**hīna**  
**gāʔ-a**  
be.PRF-2pl  
2pl-sit.IMP-IND  
when  
come.PRF-3sg.M

كنا نجلس حين جاء

We were sitting when he came

163. **kun-na**  
**gālis-īna**  
**hīna**  
**gāʔ-a**  
be.PRF-2pl  
sit.PCP-pl.ACC  
when  
come.PRF-3sg.M

كنا جالسين حين جاء

We were sitting when he came

It is observed that the lexical class of the verb from which the participle is derived plays a key role in determining its tense/aspect meaning. Trans-stative verbs denote a transition from one
state to another. Among this class are cognitive verbs like [ya-ʔərif] (MSA)/[yi-ʔəraf] (ECA) to understand, where the state of ignorance is transformed to a state of knowledge. Participles of this class are “resultative”, where the change in state is initiated at a past moment and carries over to the present, as in example 164. Translocatives involve a spatial displacement and have future tense or present progressive aspect. An example of translocative verbs is [yu-sâfar] (MSA)/[yi-sâfir] (ECA) to travel as in sentence 165. But, not all motive verbs are translocative, e.g. the participle from [ya-grî] (MSA)/[yi-grî] to run (ECA) is meaningless in sentence 166 (Mitchell & al-Hassan, 1994:83; Eisele, 1999: 127-130).

164. ʕərif-ø-un  muʃkilat-a-ka

know.PCP-sg.M-NOM problem-ACC-2sg.M

I know/have known your problem

165. ʔana  musâfir-ø-un

1sg travel.PCP-sg.M-NOM

Ana مسافر
I'm/will be travelling

166. *?ana gārī

1sg run.PCP.sg.M

أنا جاري

IV-5 Modality

Traditional Arab grammarians distinguish two types of sentences. Predicative sentencesخبرية [xabariyya]; which may be true or false "يتقابق التصديق والتكذيب" (ibn hijām, n.d.:32) and its content
is not realized, i.e. no event takes place, by uttering a predicative sentence "مضمونها لا يتوقف عند النطق بها، أي لا يتحقق بنطقها حدوث فعل ما" (gomā‘a, 2005). Predicative or declarative sentences correspond to what Mitchell & al-Hassan (1994) describe as “unqualified or categorical assertions” (p.7). Sentence 167 is declarative:

167. sāfar-a

sāfir-∅

classification

travel.PRF-3sg.M

He travelled.

The other type of sentences, termed creative sentences إنشائية [inšā‘iyya], is neither true nor false. It is not denotative before it is uttered and the event did not occur prior to this moment لا"
Besides, it involves the speaker’s feelings and concepts (gom’a, 2005). They correspond to Austin’s “non-constative or performative utterances”; which “are not essentially true or false” (Austin, 1975:140). Mitchell & al-Hassan (1994) explain that modal sentences express the speaker’s “attitude or psychological stance towards a proposition” and are “qualified or tentative expressions as to the factuality of the … proposition” (p.7). In other words, they reflect the degree of “actuality” of the event in terms of “judgment”, “evidence”, or “need” (Bhat, 1999:63). Hence, modality may be defined as “the grammaticalization of speakers’ (subjective) attitudes and opinions” (Palmer, 1986:16)

Modality is expressed by modal verbs, mood, or particles and clitics (Palmer, 1986:33). In Arabic it includes interrogatives, imperatives, optatives, and vocatives, as well as conditionals and oath (Eisele 1999:70-80, 86 & gom’a, 2005). The modal meaning of mood in Arabic was discussed in section IV-1. Other Arabic modals govern finite verbs and express both modal and temporal meanings (Fassi Fehri, 1993:160). They are either participles, such as ممكن [mumkin] possible, imperfectives, e.g. أظن [a-ðun] I think, or other constructions like لابد [lā-budda] must; which is composed of a negative particle ل [lā] no and a noun ب [budda] way out.

The two major distinctions of modality in modern Linguistics are “epistemic”; which is knowledge-based, or “deontic”; which is action-based.

IV-5-A Epistemic modality

Epistemologically, languages employ different linguistic forms to distinguish between real and hypothetical events (realis/irrealis mood), to express varying degrees of certainty regarding the
actuality of the event (judgmental modality), and to validate such judgments (evidential modality). Some common epistemic modals (judgmental and evidential) in MSA and ECA are listed in Table 8.

<table>
<thead>
<tr>
<th>MSA</th>
<th>ECA</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>- mina l-lāzim من اللازم</td>
<td>- lāzim لازم</td>
<td>must</td>
</tr>
<tr>
<td>- lā-budda لابد</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- mina al- mu‘akkad المؤكد</td>
<td>- mu‘akkad / ?akīd</td>
<td>certainty indeed</td>
</tr>
<tr>
<td>- qad + PRF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- mina l-mumkin من الممكن</td>
<td>- mumkin / yi-mkin</td>
<td>possibly</td>
</tr>
<tr>
<td>- rubbama perhaps ربما</td>
<td>- yu-hayya? Li بيتىالي</td>
<td>perhaps</td>
</tr>
<tr>
<td>- qad + IMP</td>
<td>- ?a-šun I أظن</td>
<td>may be</td>
</tr>
<tr>
<td>- ya-bdū it appears يبدو</td>
<td>- ?a-fṭikir I think</td>
<td>it seems</td>
</tr>
</tbody>
</table>

Table 8 Epistemic expressions in MSA and ECA

The modals in the table include verbs, particles and phrases. In ECA, the majority of these modals have participial form. The complementizer أن [?an] (MSA)/إن [?en] (ECA) of the subordinate clause is obligatory in most MSA modal participles, but optional in most ECA ones; e.g. 168. The presence or absence of the b-/ha-prefixes of a subordinate imperative depends on the modal verb and the tense/aspectual meaning of its subordinate clause.


1-think.IMP-IND that-3sg.M FUT-3-agree.IMP-sg.M-IND

أظن أنه سيوافق
I think he will agree.

IV-5-B Deontic modality

Deontic modals express the kind of need or obligation to actualize an event in response to external (directive) or internal (commissive) factors, for example, necessity, request, order, besides wishes and oaths (Bhat, 1999:63-5, Palmer 1986:97, Eisele, 1986: 83-6). Palmer (1986) suggests a third type “dynamic” that concerns the subject’s ability and willingness, apart from his/her opinion or attitude (p.102-3). Arabic uses lexical verbs or participles that mean can/able to, e.g. [yastat'ī] in MSA, and [y ĭdar] in ECA to express ability. Table 9 lists some deontic modals.

ECA commissives are the only participles, whereas the majority of deontic modals in ECA and MSA are imperatives and imperfectives. Modal particles like [hayya/yalla] let’s and [balāṣ] let’s not precede imperfectives. The modal in 169 may express wish if the subordinate verb is perfective, or hope if the verb is imperfective. In MSA, the perfective may express a wish (Fassi Fehri, 1993:175), as in 170.
<table>
<thead>
<tr>
<th>MSA</th>
<th>ECA</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Directives</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- ّي-كتّب</td>
<td>- ّي-كتّب</td>
<td>write</td>
</tr>
<tr>
<td>- ّي-كتّب</td>
<td>- ّي-كتّب</td>
<td>write</td>
</tr>
<tr>
<td>- ّي-كتّب</td>
<td>- ّي-كتّب</td>
<td>write</td>
</tr>
<tr>
<td>- ّي-كتّب</td>
<td>- ّي-كتّب</td>
<td>write</td>
</tr>
<tr>
<td>- ّي-كتّب</td>
<td>- ّي-كتّب</td>
<td>write</td>
</tr>
<tr>
<td>- ّي-كتّب</td>
<td>- ّي-كتّب</td>
<td>write</td>
</tr>
<tr>
<td>- يرمي-حلا</td>
<td>- يرمي-حلا</td>
<td>wish you write</td>
</tr>
<tr>
<td>- يرمي-حلا</td>
<td>- يرمي-حلا</td>
<td>wish you write</td>
</tr>
<tr>
<td>- يرمي-حلا</td>
<td>- يرمي-حلا</td>
<td>wish you write</td>
</tr>
<tr>
<td>- يرمي-حلا</td>
<td>- يرمي-حلا</td>
<td>wish you write</td>
</tr>
</tbody>
</table>

| **Commissives** | | |
| - ّي-ريد | - ّي-ريد | wish you write |
| - ّي-ريد | - ّي-ريد | wish you write |
| - ّي-ريد | - ّي-ريد | wish you write |
| - ّي-ريد | - ّي-ريد | wish you write |
| - ّي-ريد | - ّي-ريد | wish you write |
| - ّي-ريد | - ّي-ريد | wish you write |

Table 9 Deontic modality expressions in MSA and ECA

169. *layta-hu*  

| wish-3sg.M | travel.PRF-3sg.M | /yu-sāfīr-∅-u |

ليته سافر/يسافر

| wish-3sg.M | travel.PRF-3sg.M | /yi-sāfīr-∅ |

يا ريته سافر/يسافر

*I wish he travelled / I hope he travels*

170. *rahim-a-hu*  

| have mercy.PRF-3sg.M-3sg.M | Allah-NOM |

رحمه الله

*May God have mercy on him*
The imperfective can be used in ECA to give orders as in sentence 171. In general, no [b-] is prefixed to ECA imperfectives that express deontic modality.

171. ti-xallas-Ø iʃ-ʃuyl da n-nahar-da

2-finish.IMP-sg.M the-work this the-day-this

تخلص العمل ده اليوم

_Do finish this work today._

The imperative paradigm in MSA and ECA for Form I فعل [faʕala] is presented in Table 10.

There are no dual and feminine plural in ECA, and vowelling may differ from MSA.

<table>
<thead>
<tr>
<th>Person</th>
<th>Number</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>2sg.M</td>
<td>?uktub</td>
<td>?iktib</td>
</tr>
<tr>
<td>2sg.F</td>
<td>?uktub-i</td>
<td>?iktib-i</td>
</tr>
<tr>
<td>2dl</td>
<td>?uktub-a</td>
<td></td>
</tr>
<tr>
<td>2pl.M</td>
<td>?ukub-u</td>
<td>?iktib-u</td>
</tr>
<tr>
<td>2pl.F</td>
<td>?uktub-na</td>
<td></td>
</tr>
</tbody>
</table>

Table 10 The imperative paradigm in MSA and ECA for verb form I

Epistemic and deontic moods are interdependent because there is a higher degree of certainty that an event would take place when the obligation to actualize it is stronger. This interdependence may reflect in using the same form to express both moods (Bhat, 1999:64). In example 172, the modals لازم [lāzim]/ لابد [lā-budda] mean strong obligation, while in example 173, they mean a high degree of certainty. In ECA, the verb خالي [xalli], is only disambiguated from context as to whether it means to let and permit, or to make and oblige (Mitchell & al-Hassan, 1994:44).
172. **lā-budda ʔan ya-ðhab-ø-a**

must that 3-go.IMP-sg.M-SBJ

لابد أن يذهب

lāzim yu-rūh-ø

must 3-go.IMP-sg.M

لازم يروح

*He must go*

173. **lā-budda ʔanna-hu qad ðahab-a**

must that-3sg.M already go.PRF-sg.M

لابد أنه قد ذهب

lāzim rāh-ø

must go.PRF-3sg.M

لازم راح

*He must have gone*
Chapter V: Methodology

V-1 Data collection

The corpus is collected from four Egyptian satellite TV channels. Three interview programs were selected from each channel. They were broadcast between the years 2008-2010, and recorded from youtube online. The total number of programs is 12, equally divided among three topics: political or economic, religious, and sports. The first fifteen minutes of each interview are recorded, hence a total of 3 hours are transcribed. The programs are selected from three privately owned channels: Elmehwar, Dream 2, and Elhaya, besides one state TV: Elmasriyya. The four of them present entertainment programs as well as talk shows on various social, religious, political, economic, art and sport topics. Data is drawn from the four programs listed in Table 11.

<table>
<thead>
<tr>
<th></th>
<th>Elmasriyya</th>
<th>Elmehwar</th>
<th>Dream 2</th>
<th>Elhaya</th>
</tr>
</thead>
<tbody>
<tr>
<td>Religious</td>
<td>masr in-naharda</td>
<td>?al-muslimün</td>
<td>Tarīq al-hidāya</td>
<td>kalām min il-?alb</td>
</tr>
<tr>
<td></td>
<td>مصر النهاردة</td>
<td>يطال المسلمين</td>
<td>طريق البداية</td>
<td>كلام من القلب</td>
</tr>
<tr>
<td></td>
<td>al-muslimūn yatasā?alūn</td>
<td>Muslims inquire</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>المسلمون يتساءلون</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Political/</td>
<td>dawām il-hāl</td>
<td>muntaha is-syāsa</td>
<td>na?sa? yala l?izāg</td>
<td>il-hayā l-yūm</td>
</tr>
<tr>
<td>Economic</td>
<td>دوام الحال</td>
<td>Ummost politics</td>
<td>Sorry for the disturbance</td>
<td>الحياة اليوم</td>
</tr>
<tr>
<td></td>
<td></td>
<td>منتهي السياسة</td>
<td>نأسف على الإزعاج</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sport</td>
<td>nugūm il-malā?yib</td>
<td>90 di?it?a</td>
<td>il-kōra fi drīm</td>
<td>ʃobër</td>
</tr>
<tr>
<td></td>
<td>نجوم الملاعب</td>
<td>90 minutes</td>
<td>Football in dream</td>
<td>Shbeir</td>
</tr>
<tr>
<td></td>
<td></td>
<td>٩٠ دقيقة</td>
<td>الكورة في دريم</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 11 Data sources: Selected programs in different discourse topics from different Egyptian satellite channels

In Elmasriyya, Dream 2, and Elhaya, the programs *Egypt today*, *Status quo*, *Sorry for the disturbance* and *Life today* are similar. Their hosts interview with public figures from all walks
of life to discuss current problems. *The path of guidance*, *Muslims inquire* and *A talk from the heart* are entirely religious. The first is an in depth religious program on principles of faith, interpretation of the Qur’an, and Islamic jurisdiction; while the other two are more concerned with viewers’ questions and problems. *Utmost politics* interviews prominent political figures, and all the sports programs listed host famous soccer players, commentators, or administrators.

Table 12 lists names and professions of program presenters and guest speakers. There are four female presenters and nine male ones. One interview is jointly presented by a man and woman. All speakers are male. Their age ranges from late thirties to mid-fifties and one is above sixty years old. Speakers in each separate discourse have similar backgrounds. In religion, three speakers are graduates of Al-Azhar; the world chief center of Islamic learning. The fourth speaker is a self-learner and researcher. In politics and economy, speakers are prominent political, syndicate, executive, or economic figures. In sports, all of the speakers had a long career in football and one holds a top position in the most successful football club in Egypt.

### V-2 Classification of data

The objective is to look for MSA-ECA mixing in verbal phrases from the morphemic level up to the phrasal and find out what constituent at each level is identified as the item that rules the switch. Table 13 relates verb categories and levels of the phrase structure to their candidate determiner constituents $C_i$. Data is annotated manually, and the counts are performed using a computational python program. Section V-3 demonstrates how sentences are broken into independent phrases, each with one verb and its arguments. Section V-4 explains the rules of variety labeling of every phrasal element and the annotation of the whole phrase. Section V-5
describes the computational program procedure for comparing the frequency and the direction of
the switch at consecutive levels and deducing determiner constituents at each of them.

<table>
<thead>
<tr>
<th>Presenter</th>
<th>Guest speaker</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mahmoud Saad</td>
<td><em>Sheikh Khaled El-Guindy</em> Azhar scholar and lecturer at Egypt University for Science and Technology</td>
<td>2010</td>
</tr>
<tr>
<td>Mohammad Said</td>
<td><em>Sheikh Mahmoud El-Masry</em> Religious preacher and writer</td>
<td>2010</td>
</tr>
<tr>
<td>Alaa Bassiouney</td>
<td><em>Dr. Mohammad Hedaya</em> Azhar scholar and member of the Supreme Council for Islamic Affairs</td>
<td>2009</td>
</tr>
<tr>
<td>Doaa Farouk</td>
<td><em>Sheikh Ramadan Abdel-Moez</em> Azhar scholar and ex-Imam of Islamic Center in New York</td>
<td>2010</td>
</tr>
<tr>
<td>Lamees El-Hadidi</td>
<td><em>Ahmad El-Bardaei</em> Ex-President of Cairo Bank</td>
<td>2010</td>
</tr>
<tr>
<td>Mahmoud Muslim</td>
<td><em>Mufeed Shehab</em> Minister of Parliamentary Councils and Legal Affairs</td>
<td>2008</td>
</tr>
<tr>
<td>Mona El-Hussainy</td>
<td><em>Ahmad El-Naggar</em> Editor in chief of Al-Ahram economic report</td>
<td>2009</td>
</tr>
<tr>
<td>Sherif Amer &amp; Lobna Asal</td>
<td><em>Hamdi Khalifa</em> President of the Lawyers Syndicate</td>
<td>2010</td>
</tr>
<tr>
<td>Muheeb Abdel-Hadi</td>
<td><em>Essam El-Hadari</em> Egyptian National team goalkeeper</td>
<td>2010</td>
</tr>
<tr>
<td>Moataz El-Demerdash</td>
<td><em>Ahmed Hassan</em> Egyptian National team footballer</td>
<td>2010</td>
</tr>
<tr>
<td>Mostafa Abdou</td>
<td><em>Hassan Hamdi</em> President of Al-Ahli sporting club and ex-director of Al-Ahli football</td>
<td>2010</td>
</tr>
<tr>
<td>Shobeir</td>
<td><em>Hazem Imam</em> Egyptian National team footballer</td>
<td>2010</td>
</tr>
</tbody>
</table>

Table 12 Program presenters and guest speakers
<table>
<thead>
<tr>
<th>Category</th>
<th>Constituent controlling the switch</th>
</tr>
</thead>
<tbody>
<tr>
<td>– Imperfective [PX-V-SX]</td>
<td>the subject affix or the verb</td>
</tr>
<tr>
<td>– Perfective, Participle, Modal [V-SX]</td>
<td></td>
</tr>
<tr>
<td>– Future [FUT-C₁]</td>
<td>C₁ or the functional pre-verbal element (i.e. future marker/aspectual prefix).</td>
</tr>
<tr>
<td>– Aspectual prefix [ASP-C₁]</td>
<td></td>
</tr>
<tr>
<td>– Negation level [NEG-C₂]</td>
<td>C₂ or the negative marker</td>
</tr>
<tr>
<td>– Subject level [S-C₃] or [C₃-S]</td>
<td>C₃ or the subject</td>
</tr>
<tr>
<td>– Complement level [C₄-C]</td>
<td>C₄ or the complement</td>
</tr>
</tbody>
</table>

Table 13 Controlling constituent in different levels and verb classes

V-3 Inter-phrase partitioning

Data is divided into five verb groups with respect to tense, aspect and modality:

1. Perfective (PRF)
2. Imperfective (IMP)
3. Future (FUT)
4. Participle (PCP)
5. Modals (md/)

Sentences in spontaneous speech are often too long with disfluencies and many broken or parenthetical phrases. The following excerpt illustrates how text is divided into phrases (refer to word order in section IV-3:

174. lākinahu ṣayd’an ʕindama ʔaqdam ʕala ʕīdāi ʔamwāl mutas ʕawiran ʔanahu sa-yatalaqqa misl hāzihi l-fawāyid w ʔanna hāza ʔamr tabīyi wa ʕuyāyyib ʕatru tamāman
لكنه أيضا عندما أقدم على إبداع أموال متصورا أنه سيتلقى مثل هذه الفوائد وأن هذا أمر طبيعي ويجيب عقله تماما ...

but when he decided to deposit money thinking that he would get interests like this and that this is an ordinary matter and he dismisses his mind completely ...

<table>
<thead>
<tr>
<th>Phrase</th>
<th>Verb form</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. lākina-hu ḥayd-an ẓindama qaḍam ẓala ẓīdāi ẓamwāl</td>
<td>PRF</td>
</tr>
<tr>
<td>but-3sg.M also when proceed on deposit money</td>
<td>but when he decided to deposit money</td>
</tr>
<tr>
<td>b. mutas'awwiran ?anna-hu sa yatalaqqay muṣl hāzihi l-fawāyid</td>
<td>PCP</td>
</tr>
<tr>
<td>think.PCP-ACC that-3sg.M FUT 3-get.IMP.sg.M like this the-interests</td>
<td>thinking that he would get interests like this</td>
</tr>
<tr>
<td>c. ?anna-hu sa yatalaqqay muṣl hāzihi l-fawāyid</td>
<td>FUT</td>
</tr>
<tr>
<td>that-3sg.M FUT 3-get.IMP.sg.M like this the-interests</td>
<td>that he would get interests like this</td>
</tr>
<tr>
<td>d. w ?anna hāza ?amr tabī'ī</td>
<td>non verbal</td>
</tr>
<tr>
<td>and that this matter natural</td>
<td>and that this is an ordinary matter</td>
</tr>
<tr>
<td>e. wa ṣu-yayyib-ā ẓaql-u tamāman</td>
<td>IMP</td>
</tr>
<tr>
<td>and 3-make absent.IMP.sg.M mind-3sg.M completely</td>
<td>and he give up his mind completely</td>
</tr>
</tbody>
</table>
The verb in phrase (a) is أقدم [؟aqdam] proceed, its subject is the pronoun [–hu], and its complement is the preposition على [؟ala] on. Phrase (b) has no overt subject. The complement of the participle is a complement clause whose complementizer is أن [؟anna] that. This clause makes up phrase (c), whose constituents are the subject [–hu], future verb سيتلقى [sa ya-talaqqa] will get, and complement مثل [misl] like. There is no verb in (d), so this phrase is omitted from the data. Finally, phrase (e) is composed of an imperfective يغيب [yu-؟ayyib] makes absent and NP complement عقل [؟aql].

V-4 Variety labeling of constituents

Each phrasal constituent is broken down into its component morphemes, and every morpheme is marked as MSA or ECA according to its phonological, lexical, and morpho-syntactic form:

- Lexically (lex)
- Phonologically including the definite article pronunciation (phn)
- Mood/case marking (syn)

The rules for marking are:

a. Constituents are marked [0] if they are common between the two varieties. Otherwise, if they are exclusively MSA, they are [1], and if they are exclusively ECA, they are marked [2]. For example, if the form is lexically common to the two varieties, but it is pronounced as ECA, it is marked lex(0)phn(2). If the lexical and phonological forms are both common to MSA and ECA, but the verb or noun carries a mood/case marker, the
item is lex(0)ph(0)syn(1). If case is absent, but the definite article pronunciation is characteristic of MSA, the item is phn(1)syn(2).

b. Prefixes are marked for their phonetic specification only because their lexical and syntactic specifications are identical in MSA and ECA.

c. Suffixes and mood/case markers are tagged as syntactic with the same notation SX.

d. The subject and complement are labeled lexically, phonologically, and syntactically after their head.

e. The ECA rendition of the MSA phonemes [ʒ], [θ], [ð], [ð̩] are [g], [s], [z], [z̩] respectively. They are overwhelmingly prevalent among all speakers in all types of discourse. Therefore, they are ignored in phonological marking in order not to conceal other shifts such as vowel variation.

Table 14 and example 175 demonstrate how marking is performed (see Appendix B for abbreviations).

<table>
<thead>
<tr>
<th>Constituent</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>lex</td>
<td>S1</td>
<td></td>
<td>IMPO</td>
<td>C0</td>
<td></td>
</tr>
<tr>
<td>phn</td>
<td>S2</td>
<td>PX1</td>
<td>IMPO</td>
<td>C0</td>
<td></td>
</tr>
<tr>
<td>syn</td>
<td>S2</td>
<td>ASP2</td>
<td></td>
<td>SX2</td>
<td>C0</td>
</tr>
<tr>
<td>Constituent label</td>
<td>s122</td>
<td>a2</td>
<td>px1</td>
<td>imp00</td>
<td>sx2</td>
</tr>
<tr>
<td>Phrasal annotation</td>
<td>c000 s122 a2 px1 imp00 sx2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 14 Marking phrase constituents

175. il- \( \text{ba} \) \( \text{d} \) \( \text{-} \) bi- \( \text{ya-lga} \) \( \text{-} \) -u-\( \text{o} \) l- il-raj\( \text{a} \) wa

the-some-NOM HAB-3- IMP-pl.M-IND to-bribery
Some resort to bribery

The subject [il- ba‘d] is a definite noun. The noun is lexically MSA, but has no case; hence it is syntactically ECA. The MSA definite article is [?al], but in the example, the glottal stop is dropped, and the vowel quality differs according to ECA pronunciation. Consequently, the subject is s122; i.e. lexically MSA [1], but phonologically and syntactically ECA [2]. In general, the case of MSA nouns that end with a long vowel is not realized, e.g. منى [monä] Mona.

The aspectual prefix is an exclusive ECA morphological marker [2]: a2.

The verb prefix is phonetically MSA. The ECA equivalent [yi] differs in the vowel quality, hence: px1.

The verb stem [lga?] is imperfective and MSA lexically and phonologically. It is tagged IMP11. Verb stems are marked only lexically and phonetically. Mood appears in the suffix. When the verb is hollow, i.e. its second root is a long vowel; jussive mood is expressed by shortening the internal vowel. However, in order to keep tagging consistent, mood is always labeled on the suffix of the verb. Hence, the verb stem [lga?] to resort is lexically and phonetically common [0]: imp00. An example of a verb stem that is lexically common but differs phonetically in the two varieties is قال [qāl-a] he said (MSA) vs. [?āl] (ECA). Another example of a stem that varies both lexically and phonetically is يذهب [ya-ðhab] (MSA) he goes vs. يروح [yi-rūh] (ECA). Passive form in MSA and ECA is lexically and phonetically different, e.g. قيل [qīl-a] was said (MSA) vs. أتقال [it-?āl] (ECA).
The verb suffix of يا-لِجَا [ya-lga?-u-∅] lacks the indicative marker [n], so it is ECA: sx2.

Perfective verbs have no mood, but variety differences of their suffixes is phonetic, e.g. خرجت [xarag-at] *she went out* (MSA) vs. [xarag-it] (ECA). The mood marker of imperfective verbs is suppressed in MSA when the verb ends with a long vowel, e.g. انتهى [ʔintaha] *it finished* (MSA/ECA), hence the suffix is marked [0].

The complement and the subject are marked after their heads. In 175, the complement is a prepositional phrase composed of a head [li-] to, a noun رشوة [riʃwa] *bribery*.

Prepositions and pronouns have no case. In this example, [li-] is common lexically and phonetically, hence, the complement is marked c000.

The overall phrasal annotation in the last row of Table 1 are compiled and entered to the computational programs. The computational analysis proceeds as outlined in section V-6 below.

Referring to Figure 1 and Table 13, there are 27 possible [0], [1], [2] combinations for the innermost level 1 when the verb is imperfective [PX-V-SX]. They are listed in 176:

176. [PX-V-SX]  
{0,0,0} {0,0,1} {0,0,2} {0,1,0} {0,1,1} {0,1,2} {0,2,0} {0,2,1} {0,2,2} {1,0,0} {1,0,1}  
{1,0,2} {1,1,0} {1,1,1} {1,1,2} {1,2,0} {1,2,1} {1,2,2} {2,0,0} {2,0,1} {2,0,2} {2,1,0}  
{2,1,1} {2,1,2} {2,2,0} {2,2,1} {2,2,2}

Each of the other verb categories and all higher levels have the following constituent relationships that form any of the nine combinations listed in 177:
Verb level 1  [PRF-SX], [PCP-SX], [MD-SX],
FUT/ASP level 2  [FUT-V], [ASP-V],
NEG level 3  [NEG-FUT], [NEG-ASP], [NEG-V],
SUB level 4  [SUB-NEG], [SUB-FUT], [SUB-ASP], [SUB-V],
COPM level 5  [COMP-SUB], [COMP-NEG], [COMP-FUT], [COMP-ASP], [COMP-V]

177.  [A-B]
{0,0} {0,1} {0,2} {1,0} {1,1} {1,2} {2,0} {2,1} {2,2}

In 176 & 177, the following combinations are non-switches:

0   {0,0,0}{0,0}
1   {0,0,1}{0,1,0}{0,1,1}{1,0,0}{1,0,1}{1,1,0}{1,1,1}{0,1}{1,0}{1,1}
2   {0,0,2}{0,2,0}{0,2,2}{2,0,0}{2,0,2}{2,2,0}{2,2,2}{0,2}{2,0}{2,2}

V-5 Procedure of data analysis

The input of the computational program is a compilation of the total number of annotated phrases. The analysis proceeds in the following steps:

1. The program divides the phrases into the five verb classes, and analyses each class separately.

2. It compares the specifications of the elements in 176 and 177, and counts the occurrence of possible [1], [2], [0] combinations. For example, if the frequency of {1,2} exceeds that of {2,1}, CS from MSA to ECA is preferred. If one of the two directions never occurs,
then the switch in this direction is forbidden. If both CS directions are missing, then the two elements under consideration must agree in variety.

3. **V**: First, lexical and phonological specifications of the verb stem are compared, from which we deduce which of the two determines the other.

4. **SX**: The variety of the suffix is then compared to the verb stem. For the imperfective, the suffix is compared to both the verb stem and the prefix. Permutations attested in the data for different tenses, aspects, and modalities are compiled and the frequency of switches MSA → ECA and ECA → MSA are compared in order to find out which constituent at this level forces one or more pattern. For example, whether one specific variety of the prefix allows or disallows the verb stem to come from the opposite variety, or it is the variety of the verb stem that determines that of the prefix. The label of the constituent determiner of the switch (referred to as C₁) is assigned to the whole category in the verb level: [PX-IMP-SX].

5. A verb is lexically marked as MSA or ECA depending on its context of use and its form. For example the participle in the phrase مقيّد في الجدول [muqayyad-u fi l gadwal] *enrolling him in the list* is MSA, but in another context: مقيّد حرّيته [muqayyid/mi’ayyid hurriyyat-u] *restricting his freedom* is common between MSA and ECA. In the first case it is a technical term, but in the second it is not.
6. **FUT/ASP:** The same analysis is replicated for the higher levels in the phrase structure.

   The variety of the future marker and the aspectual prefix are compared to that of the determiner constituent of the verb level. FUT/ASP has (syn) specification only: An MSA future marker is marked [1], and an ECA one is marked [2]. If the bi- prefix is absent from an ECA verb, because it is optional or forbidden, the form is similar to MSA, thus marked [0]. Otherwise, a bi-prefix is [2] and a null one is [1]. C₁ can also be [1], [2], or [0]. From this step we deduce whether it is the value of the verb level that selects the type of the future marker and/or the absence or presence of [bi-], or vice versa. The value of the element that selects is C₂.

7. **NEG:** The negative marker also has one specification: (syn). This step compares NEG to two elements:
   a. **NEG-V**
   b. **NEG- FUT/ASP**

   There are no negative forms common between MSA and ECA. MSA negative markers are always preverbal. ECA ones are either preverbal, or post-verbal, or both. The value C₂ of the preceding level is possibly [1], [2], or [0]. If the second level (FUT/ASP) is null, i.e. there is no future marker and the ECA form requires no bi-, then C₂ is replaced by C₁. From this step we deduce what the controlling constituent C₃ of the negation level is.

8. **SUB/COMP:** have (lex)(phn)(syn) specifications. In level 4 (SUB), the subject may precede or follow the verb, or it may be suppressed. Level 5 (COMP) includes the complement. Common complements are a noun or prepositional phrase, a relative clause,
or a phrase headed by a complementizer. The lexical or functional head of the subject or complement is the only element considered in marking. If the head is lexical, its variety is specified based on its context of use, e.g. استثمار [istiθmār] investment is MSA because it is an economic term. قانون [qānūn] law is common, i.e. [0], because it is not only a legislative term, but is also used in more familiar and popular contexts. This step includes the following analyses:

a. SUB/COMP-SX: If case marking on the suffix is present SX is [1], otherwise it is [2]. In some cases in MSA, case is not realized; hence, the suffix is [0]. For example, the accusative is not marked for nouns ending in a long vowel [ā]. Also, prepositional and complementizer complement heads are not marked for case, e.g. the MSA [ʔallaḏī] and its ECA equivalent are labeled c110 and c220 respectively, i.e. both are syntactically common.

b. SUB-COM: The varieties of the subject and the complement are compared to each other. The frequencies will show whether one of them affects the variety affiliation of the other or they are independent.

c. SUB/COM-NEG/FUT/ASP/V: The variety of the verb argument and that of the level preceding it are compared.

In this step we come up with C4; which is the constituent that governs the switch at the subject-verb interface. Similarly, C5 is the controlling constituent at the outermost complement level. As noted earlier, if any of the constituents is absent, the corresponding level inherits the value of the level immediately above it.
This chapter covered the methodology, detailing the sources from which data is drawn and the procedure followed in the analysis. Data is classified in five categories: Perfective, Imperfective, Future, Participle, and Modal. In accordance with the Arabic phrasal composition, sentential constituents make up a five level structure, starting from the verb stem and its affixes, then the future or aspectual marker, the negative marker, the subject, and the complement. Hence, a phrase is minimally composed of an affirmative single verb if no overt subject or complement is present. Every morpheme of each constituent is labeled MSA or ECA according to definite lexical, phonological and syntactic criteria. The frequencies and direction of CS between every two consecutive levels are compared, and the CS determiner component is thus deduced.
Chapter VI: Data analysis and results

The total number of phrases in the data is 3118. The counts are divided among the five verb classes, perfective, imperfective, future, participle, and modal as shown in the first column of Table 15. The second column shows the total number of verbs of the five categories in each discourse: religious, political, and sports. The majority of the verbs are imperfective and perfective. The future, participle, and modal verbs constitute a much fewer number of the total count.

<table>
<thead>
<tr>
<th>Verb class</th>
<th>Number of phrases</th>
<th>Topic</th>
<th>Number of phrases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfective</td>
<td>1058</td>
<td>Religious</td>
<td>1226</td>
</tr>
<tr>
<td>Imperfective</td>
<td>1364</td>
<td>Political</td>
<td>790</td>
</tr>
<tr>
<td>Future</td>
<td>171</td>
<td>Sports</td>
<td>1102</td>
</tr>
<tr>
<td>Participle</td>
<td>233</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modal</td>
<td>292</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>3118</strong></td>
<td><strong>Total</strong></td>
<td><strong>3118</strong></td>
</tr>
</tbody>
</table>

Table 15 Distribution of verb phrases among different verb classes and discourse topics

Data analysis is first performed for the three types of discourse conjointly, then contrastively. In section VI-1, the five phrase levels in Figure 1 are contrastively analyzed for the perfective, imperfective, future, participle, and modal. In section VI-2, the phrasal elements that play a role in CS as deduced from section VI-1 are compared among the three data sets of religious, political, and sports discourses.
VI-1 Code-switching at phrase levels

VI-1-A Verb level 1:

Collective results for verb level

The results for the verb level in perfective, imperfective, future, participle, and modal verb phrases is listed in Appendix C-1. The following analysis details the variety distribution of verb morphemes.

The suffix: The great majority of suffixes are either ECA or both MSA and ECA. Suffixes that are common between MSA and ECA are those of the 3rd person perfective [katab-u] they wrote, jussive negated forms when the mood marker is null [lam yaktub-0] he didn’t write, or verb stems that end with a long vowel, e.g. [؟intaha] finished. There are 252 MSA suffixes out of a total of 3118, i.e. about 8%. It is seen from Table 16 that none of the future forms have MSA suffixes. There are 14 MSA perfective passive verbs (prf pss) and five MSA imperfective passive ones (imp pss) making up about 8% of all verbs with MSA suffixes. When the suffix is MSA, the prefix and stem is also MSA or common.

<table>
<thead>
<tr>
<th>PRF = 149</th>
<th>IMP = 81</th>
</tr>
</thead>
<tbody>
<tr>
<td>prf pssl1 sx1</td>
<td>px0 imp00 sx1</td>
</tr>
<tr>
<td>prf00 sx1</td>
<td>px0 imp01 sx1</td>
</tr>
<tr>
<td>prf01 sx1</td>
<td>px1 imp pssl1 sx1</td>
</tr>
<tr>
<td>prf11 sx1</td>
<td>px1 imp00 sx1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PRF = 149</th>
<th>IMP = 81</th>
</tr>
</thead>
<tbody>
<tr>
<td>prf pssl1 sx1</td>
<td>px0 imp00 sx1</td>
</tr>
<tr>
<td>prf00 sx1</td>
<td>px0 imp01 sx1</td>
</tr>
<tr>
<td>prf01 sx1</td>
<td>px1 imp pssl1 sx1</td>
</tr>
<tr>
<td>prf11 sx1</td>
<td>px1 imp00 sx1</td>
</tr>
</tbody>
</table>

Table 16 MSA suffixes in various verb categories (Total# = 252)
Suffixes are, thus, considered independent of the prefix and the verb stem for CS purposes. Ignoring the suffix, the combination reduces to lexical and phonological specification of the prefix-stem of the imperfective, and the stem of the other verb classes.

Perfective, participle and modal results for verb level

Table 17 shows the lexical and phonological specification of the perfective, participle and modal verbs.

<table>
<thead>
<tr>
<th></th>
<th>PRF = 1058</th>
<th>PCP = 233</th>
<th>MD = 292</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA/ECA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf00</td>
<td>338</td>
<td>32%</td>
<td>pcp00</td>
</tr>
<tr>
<td>prf01</td>
<td>68</td>
<td>23.3%</td>
<td>pcp01</td>
</tr>
<tr>
<td>prf10</td>
<td>3</td>
<td></td>
<td>pcp11</td>
</tr>
<tr>
<td>prf pssl1</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf11</td>
<td>155</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf02</td>
<td>306</td>
<td>44.1%</td>
<td>pcp02</td>
</tr>
<tr>
<td>prf20</td>
<td>10</td>
<td></td>
<td>pcp20</td>
</tr>
<tr>
<td>prf pssl22</td>
<td>17</td>
<td></td>
<td>pcp22</td>
</tr>
<tr>
<td>prf22</td>
<td>134</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lex(MSA) phon(ECA)</td>
<td>prf12</td>
<td>6</td>
<td>0.6%</td>
</tr>
</tbody>
</table>

Table 17 Lexical and phonological specification of the perfective, participle, and modal verbs

Lexical and phonological realizations of the participle belong to the same variety. A few perfective verbs in the corpus are lexically MSA but phonetically ECA (prf12). In 178, the verb [waggih] *directed* has the internal vowel sequence [a-i], while in MSA pronunciation it must be [a-a]. The suffix is ECA, because the 1st singular suffix [-tu] is reduced to [-t]. Hence, the MSA verb stem is phonologically ECA, and there is a CS between its lexical specification and the suffix.
The few MSA perfective verbs with ECA pronunciation are listed in Table 18. This shows that MSA verbs are sometimes phonologically adapted to the colloquial pronunciation.

The only modal verb (md12) in Table 17 is given in phrase 179. It is actually an imperfective expressing belief [ba-ʕatibir] *I consider* and is increasingly used in ECA, to the extent that it may as well be marked lexically common, rather than MSA. The 1st person prefix in MSA is [ʔa-]. Here, the glottal stop is dropped because it is preceded by the ECA aspectual b-. Dropping the glottal stop and/or changing or deleting the vowel quality of an MSA prefix is a common
phonetic effect of the introduction of [b-] in ECA. The verb stem is, thus, analyzed as phonologically ECA.

179. SS254. w ana b -a -ṣtabir -∅ in and I HAB-1sg-consider.PRF-no mood that

S020 ASP2 MD12 SX2 C020

And I consider that

An example of vowel deletion due to the introduction of [b-] is presented in example 180. The 3rd singular masculine prefix in MSA is [ya-]. Due to the aspectual prefix the vowel of the subject prefix is deleted. The verb stem that is MSA lexically and phonologically is unaffected.

180. MS68. huwwa hilm b-y-rāwid-∅-∅ kul wāhid

it dream HAB-3-entice.IMP-sg.M-no mood every one

S002 ASP2 PX2 IMP11 SX2 C002

هو حلم بيراود كل واحد

It is a dream that entices every one

An ECA lexeme can only have an MSA pronunciation in limited contexts, either jokingly, or when a less educated speaker attempts to project an educated image of him/herself. One can, thus, conclude that in general the lexical specification supersedes the phonological form.

Consequently, when the perfective, participle, or modal are negated or combined with a subject
or complement, the verb level is marked [1] or [2] after its lexical specification. If it is lexically common, it is marked after its phonological specification.

Imperfective results for verb level

The total number of imperfective verbs in the three discourses is 1364. As shown in Table 19, the overwhelming majority of imperfective verbs is pure ECA (66%). Like the perfective, MSA stems can get realized with ECA pronunciation. Pure MSA and common combinations of person prefix and stem constitute 31% of the total number. CS occurs in 3% of the imperfectives only. They occur between an ECA prefix and an MSA lexical or phonological stem. There is no CS between an MSA prefix and ECA stem.

The absence of ECA verbs with an MSA prefix means that ECA stems are accessed only when the prefix is ECA, and CS is not allowed when the prefix is MSA. Hence, the prefix is the morpheme that allows CS when it is ECA and blocks it otherwise. The conclusion is that, at the verb level, the prefix is the dominant constituent.

The frequent CS between the verb stem and the suffix as well as the prefix violates the Free Morpheme Constraint in accordance with Albirini (2010). The frequency of these switches invalidates Bentahila & Davies’ (1983) claim that they are exceptions. The results also accord with CS in distant languages. It was mentioned in chapter I that English verb stems may combine with Dholuo tense and agreement prefixes (Ochala, 2006) as well as with Spanish suffixes (Pfaff, 1979).
<table>
<thead>
<tr>
<th>Source</th>
<th>Prefix</th>
<th>Imp</th>
<th>Occurrence</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA/ECA</td>
<td>px0 imp00</td>
<td>15</td>
<td>1%</td>
<td></td>
</tr>
<tr>
<td>MSA</td>
<td>px0 imp01</td>
<td>7</td>
<td>30%</td>
<td></td>
</tr>
<tr>
<td>ECA</td>
<td>px0 imp00</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECA</td>
<td>px0 imp02</td>
<td>26</td>
<td>66%</td>
<td></td>
</tr>
<tr>
<td>ECA</td>
<td>px0 imp11</td>
<td>166</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECA</td>
<td>px2 imp00</td>
<td>264</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECA</td>
<td>px2 imp02</td>
<td>411</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECA</td>
<td>px2 imp01</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECA</td>
<td>px2 imp11</td>
<td>34</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>ECA</td>
<td>px2 imp12</td>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 19 CS between the prefix and imperfective verb stem (Total # = 1356)

As an example, the verb in 181 is composed of a prefix, suffix and a stem. The 3rd person prefix in MSA is [ta-] and in ECA is [ti-]. Thus PX is [2]. The verb is lexically and phonologically MSA. The singular feminine part of the suffix is common in MSA and ECA, but the indicative mood marker [-u] is absent, hence SX as a whole is [2]. Since PX is the determiner element, the verb level is analyzed as colloquial. The complement head [ʕala] is a preposition that is common lexically and phonologically in the two varieties. Its suffix is also [0] because prepositions do not carry case.
that it extends even to

**VI-1-B** Future/Aspect level 2

Collective results for FUT/ASP level

The notation MSA/ECA in Table 20 means that there is no aspectual prefix because it is optional or forbidden in colloquial. In this case there is no CS, and the ECA imperfective as well as its MSA equivalent are unmarked for aspect. In the future, there are no common forms.

<table>
<thead>
<tr>
<th></th>
<th>ASP-PX = 136†</th>
<th>FUT-PX = 171</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA/ECA</td>
<td>a0 px0 47 53%</td>
<td>0 0 %</td>
</tr>
<tr>
<td>MSA</td>
<td>a0 px1 222</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a0 px2 458</td>
<td></td>
</tr>
<tr>
<td>ECA</td>
<td>a2 px0 4 10%</td>
<td></td>
</tr>
<tr>
<td>ECA &gt; MSA</td>
<td>a2 px1 40 3%</td>
<td>fut2 px1 12 7%</td>
</tr>
</tbody>
</table>

Table 20 Variety specification of aspectual/future marker (a/fut) and person prefix (px)

The results show that an aspectual prefix may precede an MSA verb, but MSA aspectual characteristics cannot be enforced on an ECA verb. In other words, if an ECA verb requires a b-prefix, the b- cannot be omitted. Hence, aspectual characteristics of ECA may be imposed on
MSA verb forms, but cannot be lifted from ECA verbs. This occurs in 3% of the forms. bi-
 prefixed MSA imperfectives that occur in the corpus are listed in Appendix D-1.

The 171 future forms are about 5% of the whole data set, and only 3.5% of them MSA. All
MSA future markers select MSA prefixes, which in turn select MSA verb stems. Therefore, ASP
and FUT dominate the verb level [PX V SX]. In other words, the three system morphemes
subject/tense prefix, future marker, and aspectual prefix disallow CS when they are MSA. They
are the constituents that influence CS at their respective levels.

In 182, the 2nd person prefix is MSA. It, therefore, selects an imperfective of its same variety as
predicted earlier. The verb level is [1], but the future marker [ha] is ECA. Since the future
marker determines the variety of the verb level, then the resultant of FUT and level 1 is [2] and
the future tense verb is considered ECA.

182. MR36. ha- ta-gid -ø -ø rabb-a -k
FUT 2 -find.IMP-sg.M-no mood God-ACC-2sg.M
FUT2 PX1 IMP1 SX2 C000

2 1
2 0

وحتجد ربك

And you will find your God

The same applies to ASP in 183, where CS takes place between ECA aspect [bi-] and the MSA
subject prefix [ya-] that is followed by an MSA verb stem [lga?]. In both cases, CS from
FUT/ASP level to Verb level is in the direction ECA $\rightarrow$ MSA. The opposite direction MSA $\rightarrow$ ECA is not allowed.

183. DP28. il-ba$\ddot{a}$d$^{f}$ bi- ya-lga$^{?}$ -$\emptyset$-$\emptyset$ l el-tawzif

the-some HAB-3-resort.IMP-sg-no mood to the-employment

S1 ASP2 PX1 IMP1 SX2 C000

2 1

1 2 0

Some resort to employment

VI-1-C Negation level 3

Collective results for NEG-V

In Table 21, the total number of negated verbs in the five verb categories is 292, i.e. 9% of the whole data. CS occurs in only 8% of the perfective negatives, and no switch is observed in other verb classes. CS occurs only between an ECA negative markers and an MSA perfective verb. The negative marker is thus considered the determiner constituent at this level.

As an example, the perfective verb تم [tamm] finished/done in 184 is MSA but negated by the discontinuous ECA marker [ma...]]. The resultant of such combination is equivalent to the variety of the negative marker, not the verb; i.e. the negated verb as a whole is ECA.
<table>
<thead>
<tr>
<th></th>
<th>NEG-PRF = 74</th>
<th>NEG-PX = 101</th>
<th>NEG-PCP = 18</th>
<th>NEG-MD = 18</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA/ECA</td>
<td>0 0%</td>
<td>0 0%</td>
<td>0 0%</td>
<td>3 17%</td>
</tr>
<tr>
<td>MSA</td>
<td>1 1%</td>
<td>49 4%</td>
<td>3 17%</td>
<td>3 17%</td>
</tr>
<tr>
<td>ECA</td>
<td>67 91%</td>
<td>52 12%</td>
<td>15 83%</td>
<td>12 66%</td>
</tr>
<tr>
<td>ECA &gt; MSA</td>
<td>n2 prf11 5%</td>
<td>0 0%</td>
<td>0 0%</td>
<td>0 0%</td>
</tr>
</tbody>
</table>

Table 21 Variety specification of negation marker and verb level (Total # = 292)

184. MP 34. wi fī kīr ma-tamm-i-ʃ il-istigāba-o

li-hum

and there many NEG-finish.PRF-v-NEG the-approving-no case to-3pl

NEG2 PRF11 SX2 NEG2 S122

2 1

2 12

وفي كثير ماتمش الاستجابة ليهم

And there were a lot of them that were not responded to with acceptance

Phrase 185 shows an MSA verb [istagāb] respond negated by the same discontinuous ECA marker. According to ECA phonology, the internal long vowel is shortened as a result of having a final [-ʃ]. Thus, the stem is phonetically [2], but the verb level is marked [1] because the lexeme is MSA. Like 184, the resulting combination of NEG2-V1 is ECA. This analysis suggests that although the verb is lexically standard, it is perceived by the listener as ECA. In other words, the impact of negation on perception is stronger than the verb form.

185. HP 46. wi lē ma-stagab-ū-ʃ

and why NEG-respond.PRF-3pl-NEG

NEG2 PRF12 SX0 NEG2
Eid (1988) observes no CS between after the negative marker. The few switches recorded in here for NEG-V shed doubt on this constraint. Also, Bassiouney (2006) claims that negative markers are drawn from ML in contrast to the present results, where up to 17% of the markers are MSA.

Collective results for NEG-ASP/FUT

Table 22 shows that no CS occurs between markers for negation and aspect or future. However, this can only be confirmed for NEG-ASP, because, in Table 20, it has been shown that only six MSA future forms are available in the data, and their context as shown in Appendix D-2 is affirmative. Therefore, it cannot be concluded whether an MSA future can be negated by an ECA negative marker or not.

<table>
<thead>
<tr>
<th>NEG-ASP</th>
<th>54</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA</td>
<td>14</td>
</tr>
<tr>
<td>ECA</td>
<td>40</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NEG-FUT</th>
<th>26</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA</td>
<td>0</td>
</tr>
<tr>
<td>ECA</td>
<td>26</td>
</tr>
</tbody>
</table>

Table 22 Variety specification of negation marker and FUT/ASP level

Phrase 186 illustrates the combination of ECA negative and future markers that are followed by an MSA imperfective. The passive form of the imperfective verb is MSA, where the vowel sequence is [a-a] and the prefix vowel is [u-]. In contrast, the ECA passive is أتحاسب [ʔa-t-hāsib]; where the stem is of the form [tCāCiC]. The verb level is MSA and the future marker is ECA.
Thus, the whole level preceding the negative is analyzed as ECA. Because of the clear distinction between passive forms in MSA and ECA, it is claimed that the switch in this sentence is strongly evident on the perception level. However, according to the analysis followed, the combination of the future, prefix and stem is considered ECA, so that no switch occurs at the negative level.

Similarly, in 187, no CS takes place between the negative marker and aspect. The equivalent of this negative verb in ECA requires a bi-prefix: [ma-b-ti-daxxal-], where [ma..] is the ECA pre- and post-verbal negative marker. ASP in the example is marked [1] because it is habitual, and the marker is null. Therefore, the verb level cannot be ECA. The PX, ASP, as well as NEG are all MSA.
It can be concluded that all negative markers agree in variety with the ASP level. If FUT/ASP is missing, the negative also agrees with the variety of the prefix or the verb stem for all categories except the perfective. However, CS between the negative marker and the perfective occurs only when the former is ECA.

VI-1-D Subject level 4 and Complement level 5

Case marking

Subjects and complements are often pronouns that do not carry case. This is clearly reflected in Appendix C-2 and Appendix C-3 for subjects and complements in all verb categories. Subjects and complements that carry case are highlighted in bold and have the label s(xy1) and c(xy1); where x is the variety of the lexical specification, y is the variety of the phonological specification, [1] is the MSA suffixal case. 3% of the subjects and 2% of the complements are marked for case, and are grouped in Table 23 and Table 24.

The percentages in the last row of the two tables are the ratios of case marked subjects/complements to the total number of complements in the verb class. The ratios of case marked arguments are similar except for subjects of modal verbs. Inspecting these cases, it is
found that they share some common features. For example, the accusative nunation [-an] on the complement of the copula [kāna] _be_ in 188 is emblematic of educated speech.

<table>
<thead>
<tr>
<th>SUB-PRF = 21</th>
<th>SUB-IMP = 23</th>
<th>SUB-FUT = 1</th>
<th>SUB-PCP = 0</th>
<th>SUB-MD = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>s001 prf pss11 1</td>
<td>s001 px1 1</td>
<td>s001fut1 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s001 prf01 6</td>
<td>s011 px1 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s001 prf11 6</td>
<td>s111 px0 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s011 prf01 2</td>
<td>s111 px1 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s011 prf11 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4% 4% 1% 0% 0%

Table 23 Subjects with MSA case marking (Total # = 45)

<table>
<thead>
<tr>
<th>COMP-PRF = 14</th>
<th>COMP-IMP = 22</th>
<th>COMP-FUT = 0</th>
<th>COMP-PCP = 2</th>
<th>COMP-MD = 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>c001 prf00 1</td>
<td>c001 px0 2</td>
<td></td>
<td>c111 pc01 1</td>
<td>c001 md11 1</td>
</tr>
<tr>
<td>c001 prf01 4</td>
<td>c001 px1 6</td>
<td></td>
<td>c111 pc11 1</td>
<td>c011 md11 1</td>
</tr>
<tr>
<td>c001 prf11 2</td>
<td>c011 px1 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c011 prf00 1</td>
<td>c101 px1 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c011 prf01 1</td>
<td>c111 px1 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c011 prf11 5</td>
<td>c111 px2 3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2% 2% 0% 1% 2%

Table 24 Complements with MSA case marking (Total # = 41)

188. DP 135. **rubbama l-ʔamr-° bi-y- kūn ........ ʕafwāʔiyy-an**

perhaps the-matter-no case HAB-3-be.IMP-sg.M-no mood random-ACC

```
MD110 S112 ASP2 PX2 IMP00 SX2 C111
  2 2

MD110 S112 2 2 C111
S112 2 C111
```

ربما الأمر في بعض الأحيان يكون عشوائنا

*Perhaps the matter is sometimes random*
Phrases 189 & 190 are both indirect speech narratives containing the verb يقول [y-ʔūl] he says. The complement is قال [qāl-a] he said is a quoted saying in 189, and a Qur’anic phrase مثل الجنة [maθal-u l-ṣanat-i] the similitude of Heaven in 190. MSA phonemes are realized in both cases: [q], [ʔ], and [ʕ] and the 3rd singular masculine suffix in 189, as well as nominative case [-u] in 190 are used.

189. DR 96. wi y-ʔūl qāl-a ...
   And 3-say.IMP-∅ say.PRF-3sg.M
       PX2 IMP02 SX2 PRF111
           2 C111

   و يقول قال ...

   And he says: “He said ...”

190. SR 250. y-ʔul-∅ l-ak maθal-u l-ṣanat-i
   3-say.IMP-∅ to-2sg.M similitude-NOM the-Heaven-GEN
       PX2 IMP02 SX2 C111
           2 C111

"يفكك: "مثل الجنة."

   He says to you: “The similitude of Heavens ...”

In 191, the ECA imperative شُوف [ʃūf ] see would rather be analyzed as an interjection followed by the NP Qur’anic verse مثل نوره [maθal-u nūr-i-ḥi] similitude of his light.
There are only six phrases where both the subject and complement are case marked. These are listed in Appendix C-4, from which we see that they are all Qur’anic verses.

Since the results confirm the rarity of case realization on verb arguments and the absence of any correlation between the subject and complement in terms of case marking. Therefore, the syntactic specification of the variety (case) is ignored in discussing CS at the subject and complement levels for the five verb categories.

The omission of case markers has been observed in CS in other languages. As stated in chapter I, accusative case of Hebrew is frequently dropped from verb objects in Hebrew-Spanish CS, when Spanish is the ML (Berk-Seligson, 1986).

Total results for COMP-SUB:

The correlation between the subject and complement with respect to their variety specification lexically and phonologically is shown in Table 25. The frequency of CS in either direction is 3%.

Example 192 illustrates the switch between an ECA complement and MSA subject. The subject is an MSA relative pronoun الذين [?allaðīna] who whose ECA equivalent is اللي [?illi], and the
complement is an ECA noun [yaftٰīt] *sign*, while the perfective stem [rafaٰy] *they raised* and its 3rd plural masculine suffix [-ū] are common in both varieties.

<table>
<thead>
<tr>
<th>COMP-SUB = 1103</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mixed</strong></td>
</tr>
<tr>
<td>Common</td>
</tr>
<tr>
<td>MSA</td>
</tr>
<tr>
<td>ECA</td>
</tr>
<tr>
<td>MSA &gt; ECA</td>
</tr>
<tr>
<td>ECA &gt; MSA</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Table 25 Correlation of subject and complement MSA case marking

192. DP 58. ْتالِدینا ِرضاٰی-ٓ sign-no case ِیُل-ٓیسلام

who raise.PRF-3pl the-Islam

<table>
<thead>
<tr>
<th>S110</th>
<th>PRF00</th>
<th>SX0</th>
<th>C222</th>
</tr>
</thead>
<tbody>
<tr>
<td>S11</td>
<td>0</td>
<td>C22</td>
<td></td>
</tr>
</tbody>
</table>

*Those who raised the sign of Islam*

In the following sentence, the subject بعض [‘il-baٰ`dٰ] *some* is mixed because it is a lexical MSA noun that is attached to a phonetically ECA definite article. The complement الحكومة [il-ٰیکِما-ٰ] *the government* is lexically common and its definite article is also ECA. The verb is an MSA perfective.
In Table 25, mixed forms constitute 2% and are all lexically MSA and phonologically ECA verb arguments. For example, the combination c12 s22 represents an ECA subject with a complement that is an MSA lexeme uttered in ECA. By the same argument forwarded for verb stems, an ECA noun does not have MSA pronunciation except in limited functions of use; the absence of c21 or s21 is expected. Hence, in analogy to the verb stem, the variety of a subject or complement that is mixed internally is determined by the lexical specification of the noun.

Based on the result that CS between the subject and complement is equal in either direction, neither of the two arguments controls the other. The SUB/COMP levels are better viewed as one external interface of the phrase. In 193, the subject and the verb agree in variety, and stand together in opposition to the complement. It is suggested that to a listener, CS is perceived to occur between the ECA complement and the MSA verb.

**Total results for SUB/COMP -Verb:**

The subject and complement behave alike in relation to the verb when there is no negative, aspect or future marker intervening between them. Switches in the variety of the subject or complement may occur in either direction. However, comparing the ratios of switches in the
third and fourth rows of Table 26 and Table 27, we see that CS between an ECA subject and MSA verb is twice to five times bigger than the reverse for various verb categories except the modal. CS of the complement, on the other hand, has closer ratios in both directions. Arguments of the modal are the least to switch and there is no instance where an MSA modal takes an ECA subject or complement.

<table>
<thead>
<tr>
<th></th>
<th>SUB-PRF = 426</th>
<th>SUB-PX = 654</th>
<th>SUB-PCP = 160</th>
<th>SUB-MD = 48</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA/ECA</td>
<td>45</td>
<td>6</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>MSA</td>
<td>129</td>
<td>143</td>
<td>25</td>
<td>4</td>
</tr>
<tr>
<td>ECA</td>
<td>225</td>
<td>414</td>
<td>115</td>
<td>40</td>
</tr>
<tr>
<td>MSA &gt; ECA</td>
<td>8</td>
<td>17</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>ECA &gt; MSA</td>
<td>15</td>
<td>63</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Mixed</td>
<td>s11 prf12 =1</td>
<td>s12 px1 =7</td>
<td>s12 pc1 =2</td>
<td>s02 md12 =1</td>
</tr>
<tr>
<td></td>
<td>s12 prf01 =1</td>
<td>s12 px2 =4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>s12 prf11 =2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2%</td>
<td>2%</td>
<td>0.4%</td>
<td>2%</td>
</tr>
</tbody>
</table>

Table 26 Variety specification of subject-verb combinations for the perfective, imperfective, participle and modal

<table>
<thead>
<tr>
<th></th>
<th>COMP-PRF = 626</th>
<th>COMP-PX = 1023</th>
<th>COMP-PCP = 158</th>
<th>COMP-MD = 30</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA/ECA</td>
<td>101</td>
<td>29</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>MSA</td>
<td>158</td>
<td>249</td>
<td>28</td>
<td>2</td>
</tr>
<tr>
<td>ECA</td>
<td>328</td>
<td>652</td>
<td>113</td>
<td>23</td>
</tr>
<tr>
<td>MSA &gt; ECA</td>
<td>10</td>
<td>44</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ECA &gt; MSA</td>
<td>20</td>
<td>42</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Mixed</td>
<td>c02 prf12 =1</td>
<td>c12 px1 =2</td>
<td></td>
<td>c02 md12 =1</td>
</tr>
<tr>
<td></td>
<td>c11 prf12 =1</td>
<td>c12 px2 =5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c12 prf02 =3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c12 prf11 =4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2%</td>
<td>0.7%</td>
<td>0</td>
<td>3%</td>
</tr>
</tbody>
</table>

Table 27 Variety specification of complement-verb combinations for the perfective, imperfective, participle and modal

The only recorded switch between an MSA subject and a modal is given in 194. The ECA modal لزيم [lāzim] must is lexically distinct from its MSA counterpart that comes in a prepositional
structure: من اللازم [min al-lāzim]. The subject إحالة [ʔihlāl] replacement is clearly MSA with no indicative case marker [-u].

There must be replacement and renewal

The verb and its arguments in the last row of the tables are internally mixed. For example, (s12 prf01) is a lexically MSA and phonetically ECA subject of a lexically common perfective that has an MSA pronunciation. In phrase 195, the complement is an MSA noun colloquialized by changing its internal glottal stop: [lāʔiha] to a glide: [lāyha]. It is also preceded by the ECA definite article [il], and is unmarked for case; hence it is (c122). The imperfective stem and its prefix are MSA and mood is dropped (imp112). According to the analysis of labeling, this phrase should be perceived as entirely MSA in spite of the noted deviations from the standard.

In order that I apply the regulations
The following example illustrates a CS between the subject and the verb. The verb is an MSA perfective, and the subject is the pronominal suffix of the complementizer إن [inn] that. In MSA, the pronominal suffix is [-hu] differing from its ECA equivalent [-u] only phonetically, and unmarked for case. Consequently, the subject is labeled S020.

```
196.  HR89. sayyid-na  ?abu-lubāba  mugarrad  ?inn-u  ?afār-ø
    master-2pl.  Abu-lubaba  merely  that-3sg.M  point.PRF-3sg.M
    S020  PRF11  SX2
```

سيدنا أبو لبابة مجرد إنه أشار

*Just that our master Abu-lubaba pointed*

CS between the complement and the verb in 197 involves an MSA complement and an ECA modal verb. The verb [y-nawwar] is imperfective in form but expresses wish. Its stem is both MSA and ECA, but since the internal vowel sequence [a-a] is ECA, while that of MSA is [a-i], the stem is labeled MD02. The prefix is also ECA, because its vowel is deleted. Therefore, the verb level is ECA. The complement, on the other hand, is an MSA feminine noun suffixed by a 2nd person possessive pronoun. Upon suffixation in ECA, the medial long vowel [ī] of the noun gets shortened; besides, the vowel of the peripheral feminine marker [-at] is also deleted. Thus, MSA [basfīrat-a-ka] becomes [basfīrt-ø-ak]. This makes the complement phonetically ECA. It is also syntactically labeled [0], because the accusative case marker [-a-] is concealed by the vowel of the ECA pronominal suffix [-ak]. Since lexical affiliation outranks phonetic specification, the complement is considered MSA.
197. SR4. Allah y-nawwar bas‘irt--ø-ak

Allah 3-enlighten.MD02-sg.M-no mood foresight-no case-2sg.M

<table>
<thead>
<tr>
<th>S002</th>
<th>MD02</th>
<th>SX2</th>
<th>C120</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

هو كده الله ينور بصيرتك

May God enlighten your foresight

In example 198, CS takes place between the two verb arguments and the verb. As in example 196, the subject is a pronominal suffix [-u] of the complementizer [?inn]. The complement is a complementizer clause and its head [inn] differs from MSA in vowel quality and the absence of the initial glottal stop [?anna]. Hence, both arguments are ECA. The participle, however, is MSA. As concluded above, the varieties of the subject and complement are independent and form a single level. Hence, CS occurs between the outermost ECA level and the MSA verb.

198. DR9. ?inn-u mudrik-ø inn il-ʔasbāb ʔand-u muʕatʕītala

that-3sg.M aware.PCP-no case that the-means at-3sg.M inoperative

<table>
<thead>
<tr>
<th>S020</th>
<th>PCP11</th>
<th>SX2</th>
<th>C020</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

إنه مدرك إن الأسباب عنده معطلة

That he is aware that the means on his side are inoperative

Total results for SUB/COMP–ASP/FUT:

The distribution of subjects and complements looks different in relation to the aspectual prefix and future marker. In Table 28 and Table 29, there are around 5% of the MSA subjects and
complements of bi-prefixed imperfectives, and a similar number of ECA arguments of MSA aspect level. In contrast, the future sustains only one direction of CS for subjects and complements alike. When an ECA future marker is used, the argument is usually MSA. There is only one CS between an ECA subject and MSA future marker.

![Table 28 Variety specification of subject and FUT/ASP levels](image)

In 199, the verb level is [1] and becomes [2] by the introduction of the bi-prefix. The subject [il-ba‘d] is mixed; lexically [1] and phonetically [2] because of the definite article. The complement is [1], because its head preposition [la] is phonetically MSA. It differs from the ECA preposition [li] in vowel quality and length, but both are identical lexically. They are also syntactically the same, because prepositions are not case marked. Thus, the complement is
labeled [1] after its phonetic specification. CS in this phrase occurs at an intermediate level from a standard core (V) to a colloquial (ASP) back to a standard periphery (SUB/COMP).

An example of a SUB-ASP switch in the opposite direction ECA→MSA is given in 200. The subject in this case is a noun phrase حضرتك [hadīr-ak] your presence which is functionally used as formal ‘you’ in Egypt. In MSA, there is no formal/informal distinction in addressing a second person. Hence, the subject is lexically ECA. The noun phrase in MSA is [hadūr-u-ka]; i.e. the two NPs differ phonologically in internal vowel quality of the head and syntactically in its case marking, besides the metathesis of the determiner [ak/ka] your.

البعض بينظر لها

Some look at it
A saying you know

The single CS case between an ECA subject and MSA future marker is given in phrase 201. The subject is again pronominal suffix [-u] of the complementizer [?inn] that differs from its MSA counterpart [-hu] phonetically. Hence, the subject is common lexically and syntactically and is ECA phonetically. Since the future marker is MSA, the prefix must also be MSA. Consequently, the lexically common verb is pronounced with the standard [q] sound rather than its colloquial glottal rendition.

201. DP 40. ?inn- u sa-u-?aqqi-q-0-0 ribh-0
that 3sg FUT-3-realize.IMP-sg-no mood profit-no case
S020 FUT1 PX1 IMP11 SX2 C000

 إنه سيحقق ربح

That he will realize profit

CS between the complement and the future is illustrated in 202. The complement is an MSA noun with no case, and the future marker is ECA. The suffix of the MSA verb is labeled [0], because indicative mood marker [-u] cannot appear on a verb that ends in an [i] vowel. CS in this phrase is similar to 199: from a standard core V through a colloquial FUT back to a standard COMP.
I will run elections

Total results for SUB/COMP-NEG:

Table 30 and Table 31 show the variety of subjects and complements in connection with the negative marker. The agreement between the two levels is between 60-90% in all verb categories. The distribution of CS subjects and complements in different verb classes is largely inconsistent. CS from an ECA subject or complement to an MSA negative is most frequent in the imperfective, and is absent in other verb classes. However, there are a few cases where the complement of a negative modal switches in this direction. CS in the other direction occurs in all classes except the future and participle.

As just noted, subjects and complements of the imperfective are the most commonly switched when the negative is MSA. In 203, both arguments are ECA and the negative is MSA. The subject ائنا [ihna] we is lexically distinct from its MSA counterpart نحن [naḥnu], and the complement ECA prepositional head [ʔalē-] on turns the MSA diphthong [ʔalay-] to a long vowel, so it is phonetically ECA. The verb, prefix, and negative marker لا نوافق [lā nu-wāfiq-u] do not agree are all MSA. Their ECA equivalent has the [q] is pronounced [ʔ], the prefix vowel is
dropped, and the negative marker is the discontinuous [ma..ʃ]: [ma-n-wāfiʔ-ʃ]. CS occurs only at the outermost level from ECA arguments to the MSA negative verb.

<table>
<thead>
<tr>
<th>SUB-NEG</th>
<th>PRF = 33</th>
<th>IMP = 62</th>
<th>FUT = 17</th>
<th>PCP = 12</th>
<th>MD = 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA</td>
<td>1</td>
<td>3%</td>
<td>17</td>
<td>27%</td>
<td>0</td>
</tr>
<tr>
<td>ECA</td>
<td>31</td>
<td>94%</td>
<td>36</td>
<td>58%</td>
<td>16</td>
</tr>
<tr>
<td>MSA &gt; ECA</td>
<td>1</td>
<td>3%</td>
<td>0</td>
<td>0%</td>
<td>1</td>
</tr>
<tr>
<td>ECA &gt; MSA</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>0</td>
<td></td>
<td>7</td>
<td>11%</td>
<td></td>
</tr>
<tr>
<td>s12 n1= 1</td>
<td>3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s12 n2= 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 30 Variety specification of subject and negative levels

<table>
<thead>
<tr>
<th>COMP-NEG</th>
<th>PRF = 46</th>
<th>IMP = 103</th>
<th>FUT = 17</th>
<th>PCP = 10</th>
<th>MD = 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA</td>
<td>1</td>
<td>2%</td>
<td>33</td>
<td>32%</td>
<td>0</td>
</tr>
<tr>
<td>ECA</td>
<td>35</td>
<td>93%</td>
<td>62</td>
<td>60%</td>
<td>17</td>
</tr>
<tr>
<td>MSA &gt; ECA</td>
<td>1</td>
<td>2%</td>
<td>3</td>
<td>3%</td>
<td>0</td>
</tr>
<tr>
<td>ECA &gt; MSA</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Mixed</td>
<td>c12 n2= 1</td>
<td>2%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 31 Variety specification of complement and negative levels

203. MP 138. ihna ka kanīsa lā nu-wāfiq ūalē-h

2pl. as church NEG 2pl-agree.IMP-no mood on-3sg.M

<table>
<thead>
<tr>
<th>S220</th>
<th>NEG1 PX1 IMP11 SX2</th>
<th>C020</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

1

2

1

2

احنا ككنيسة لا نوافق عليه

We, as a church, do not agree on it
In the following phrase the subject agrees and the complement switches. The imperfective [ya-kun] is common, but its prefix is MSA. Thus, the verb level is [1] and this requires an MSA negative marker لم [lam] not. The negation level is, therefore [1]. CS between an ECA negative and either argument is possible. But in this case the subject is lexically MSA and phonetically ECA because its medial glottal stop [ra'ismāl] capital is elided, so it is labeled after its MSA lexical specification. There is no CS between the subject and the negated verb, however, CS occurs between the combination of the two constituents and the phonetically ECA complement [haʔiʔi] because the latter has the [q] sound turned to [ʔ].

\[
\begin{array}{cccc}
\text{SP105. } & \text{rāsmāl-u} & \text{lam} & \text{ya-kun} \\
\end{array}
\]

\[
\begin{array}{cccc}
\text{capita-3sg} & \text{NEG} & \text{3-be.IMP.JSS} & \text{real-no case} \\
\end{array}
\]

\[
\begin{array}{cccc}
\text{S120} & \text{NEG1} & \text{PX1} & \text{IMP00} \\
\end{array}
\]

\[
\begin{array}{cccc}
\text{SX1} & \text{C022} & \text{1} & \text{1} & \text{2} \\
\end{array}
\]

\[
\begin{array}{c}
\text{Rasma}
\end{array}
\]

\[
\begin{array}{c}
\text{His capital was not real}
\end{array}
\]

Example 205 shows a similar switch in the other direction. According to Table 30 and Table 31, when the negative of the imperfective is ECA, the subject agrees, but the complement may switch. The subject in the example is a proper noun, hence common, but its definite article is ECA, so it is labeled [2] and agrees with the ECA negative verb. Since it was established that neither the subject nor the complement impact the variety of the other, the subject and verb, being of the same variety, stand together in opposition to the MSA complement. In other words, CS is sensed to occur between the negated colloquial verb and the MSA complement.
Zamalek has not yet seen the other side

From the analysis above, CS at the subject and complement differs with different preceding levels. It occurs in both directions with verb level when the FUT/ASP level is absent. It is most frequent between ECA subject and MSA verb, especially the imperfective. SUB/COMP-ASP switches are equally possible in both directions. As with the negative, no generalization can be reached for SUB/COMP-FUT due to the small number of standard future forms available. SUB/COMP-NEG switches vary among different verb classes. The inconsistency with respect to negation may indicate that subjects and complements are disconnected from the negative marker and are more closely tied to the verb level and its prefixes.

VI-2 Comparison by discourse topic

In this section, I compare how the topic of discourse impacts the features of verbal affixes and markers of future, aspect and negation, besides the realization of case and mood.
VI-2-A The suffix

Besides denoting number and gender, verbal suffixes in MSA also carry mood. They also denote person in the perfective. Nouns, verbal participles, and some modals express number and gender agreement in addition to case marking on their suffixes. Table 32 shows that the majority of MSA suffixes in the corpus come from religious speech, but more than half of them are quotations from the Qur’an, Prophet’s sayings, or poetry. There is no indication how subjects and complements compare in this respect, confirming the observation already made that they are independent of each other.

<table>
<thead>
<tr>
<th></th>
<th>Religious</th>
<th>Political</th>
<th>Sports</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spontaneous</td>
<td>Quoted</td>
<td></td>
</tr>
<tr>
<td>PRF</td>
<td>57</td>
<td>57</td>
<td>40</td>
</tr>
<tr>
<td>IMP</td>
<td>45</td>
<td>32</td>
<td>4</td>
</tr>
<tr>
<td>FUT</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MD</td>
<td>2</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>PCP</td>
<td>5</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>SUB</td>
<td>15</td>
<td>42</td>
<td>3</td>
</tr>
<tr>
<td>COMP</td>
<td>33</td>
<td>20</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>38%</td>
<td>44%</td>
<td>18%</td>
</tr>
</tbody>
</table>

Table 32 Distribution of MSA suffixes, case and mood among different discourses

Phrase 206 is an example of MSA suffixes in a phrase that is quoted from the Qur’an but incorporated in ordinary speech. The perfective suffix [-a] for the 3rd person singular masculine is always dropped in ECA, but realized in this quote. The subject is marked by the final vowel [-u] for nominative case. The speaker assumes the role of Zechariah and speaks in his voice. So he starts with أنا [ʔana] I followed by the Qur’anic quotation وَهَنَّ العَظَمَ مَنِي [wahan-a l-ʔaðm-u min-ni] my bones are weakened.
206. **DR 59. ʔana “wahan-a l-ʔaðm-u min-ni”**

<table>
<thead>
<tr>
<th>1sg. weaken.PRF-3sg.M</th>
<th>the-bones</th>
<th>from-1sg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRF11 SX1 S111</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

أنا وهن العظم مني

*I, my bones are weakened*

**VI-2-B Aspect prefix**

All bi-prefixed MSA imperfectives in the three discourses are listed in Appendix D-1. The distribution and effect of ECA aspectual prefix on the verb is summarized in Table 33. The aspectual prefix often affects the pronunciation of the person prefix of MSA imperfective rendering it colloquial. In other cases the prefix vowel sustains its MSA vowel. Half the number of verbs in Appendix D-1 are marked px1 and the other half px2. **Error! Reference source not found.** shows that the effect on person prefix is least in religious discourse and greatest in sports, however, the total number is very few. The bi-prefix may also result in changing the vowel sequence of the stem to ECA. There are two such cases in Appendix D-1 (marked imp12) and they occur in political discourse.

<table>
<thead>
<tr>
<th>bi + MSA imperfective</th>
<th>Religious</th>
<th>Political</th>
<th>Sports</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>12</td>
<td>13</td>
<td>34%</td>
</tr>
<tr>
<td>bi → ECA person prefix</td>
<td>3</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>bi → ECA internal vowelling</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 33 Effect of ECA aspect on MSA verb in different discourses**

172
The two effects happen in 207. The verb stem in MSA is [ta'addā], having the vowel sequence [a..a]. When prefixed by b-, the first [a] of the stem is deleted and the 3rd person prefix is changed from [ya] to [yi].

207. DP122. fa huna l-ʔamr-ø fi l-haʔiʔa b-yi-tʔaddā-ø masʔalit-ø …
so here the-subject in the-fact HAB-3-surpass.IMP-sg.M issue …

S112 ASP2 PX2 IMP12 SX2 C020

في هنا الأمر في الحقيقة بيتعدى مسألة أن هو يعني مجرد

VI-2-C Future marker

It has been shown that the vast majority of future markers are ECA. The only six MSA future phrases in the data are included Appendix D-2. Only one speaker in a religious program used the MSA future form, but all three speakers in political/economic programs used it, and it never occurred in sports talk. When an ECA future marker precedes an MSA verb, it may affect the person prefix in a similar way as the aspectual bi-, by changing or deleting its vowel. There are 68% of future forms that resist this change, i.e. maintain the MSA vowel of the prefix. The three discourses are comparable in this respect as shown in Table 34.

<table>
<thead>
<tr>
<th></th>
<th>Religious</th>
<th>Political</th>
<th>Sports</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA future marker</td>
<td>1</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>ECA future markers</td>
<td>44</td>
<td>60</td>
<td>65</td>
</tr>
<tr>
<td>ECA FUT → MSA person prefix</td>
<td>4</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 34 Effect of ECA future marker on MSA verb in different discourses
The verb in both examples 208 & 209 is MSA and the future marker is ECA. The MSA prefix vowel in 208 is deleted, while in 209, the prefix remains the same [ya].

208. HP 6. ha-y-tim-o

جوالن-ир فوزي البطل محمد

FUT-3-finish.IMP-sg.M announcing winning the-mister Mohammad

209. SS328. ha-ya-zall-o

ح يظل بيتك

FUT-3-remain.IMP-sg.M home-ر-2sg.M

VI-2-D Negative markers

<table>
<thead>
<tr>
<th></th>
<th>Religious</th>
<th></th>
<th>Political</th>
<th>Sports</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spontaneous</td>
<td>Quoted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRP</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>IMP</td>
<td>24</td>
<td>5</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>FUT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MD</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>PCP</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 35 Distribution of MSA negative marker among different discourses
As with the future, negation most often has ECA form. From Table 35, we see that MSA negation never occurs in sports data. Compared to MSA suffixes there is a wider gap between naturally spoken MSA negatives and quoted forms from religious and literary texts. Most MSA negation is in the imperfective.

Appendix D-3 lists MSA negated verbs by discourse and category. Five markers are used. The most frequent are [lā] with the imperfective and modal that have present tense interpretation, and لَمَ [lam] with the imperfective that refers to the past tense. Only one imperfective verb is negated by [mā] and the phrase is quoted for a Prophet’s saying. In political discourse, لا [‘allā] is also used a few times. The three participles are negated with غير [‘ēr] that is adapted phonetically to ECA by reducing the diphthong [ay] to a long vowel [ē]. It is considered MSA because participles in ECA are usually negated by مش [miʃ]. All these verbs are MSA because, as observed in section VI-1-C, no CS occurs after an MSA negative marker.
Chapter VII: Discussion and conclusion

By stratifying the verbal phrase according to levels and dividing every level into morpheme elements, one is able to discern how these elements may be drawn from MSA and ECA in order to recombine in diglossic speech. The aim was to find which element at every level is the decisive factor in this re-composition. The main finding is that functional elements are overwhelmingly ECA and that they play the key role in controlling CS between MSA and ECA. This conforms to Myers-Scotton’s (1993) MLF, as well as the dominance of L1 grammar account of Bentahila & Davies (1983). Various indications point that switching after an MSA element tend to be more constrained, in agreement with the Directionality Constraint that is suggested by Eid (1988).

I did not rely solely on Eid’s (1988) criterion for labeling the variety of a lexical term. Her criterion relies on the presence or absence of an equivalent in the other variety. Additionally, I take into consideration the phrasal and sociolinguistic/semantic context of the term. I also consider that mixing starts at the verb stem by phonological adaptation of an MSA lexeme. Phonological adaptation in CS has been an issue of debate as to whether the item is rather a borrowing from EL. However, by the same argument that the morpho-syntax of embedded EL material is “vulnerable to the ML because they are stranded in a sea of ML” (Myers-Scotton, 1993:181), it is natural that ML phonology would also impact EL elements. Taking the verb [raʔet] I saw as an example, Eid judges it as an MSA ‘intermediate phonetic variant’ because it has a lexically different ECA equivalent. Instead, I reach the same judgment by recognizing that it is a mixed form which is lexically MSA and phonetically ECA. Since the results show that
ECA phonetic characteristics fail to affect the majority of MSA verb stems, and that in normal contexts of use ECA lexemes do not have MSA pronunciation, the lexical identity of the verb precedes and the variety of the stem is traced to its lexical specification not its phonetic output.

In this respect, Saiegh-Haddad (2004) poses two mixed cases to consider, an MSA lexeme composed of phonemes common to both varieties; e.g.  وضع [wad\check{a}l-a] he puts, and an MSA lexeme that has at least one MSA phoneme, e.g. ذهب [\check{d}ahab-a] gold (p.498). In her discussion of the phonemic vs. lexical distance between MSA and the colloquial, she claims that the effects of the linguistic affiliation of the target phoneme and the lexical status of the stimulus word on language processing differ. Kindergarten children were more sensitive to both factors, phonological and lexical, than first graders, who “have internalized a rather robust phonological representation for MSA phonemes and MSA words” (p.505). On the other hand, when “a related competing phonological representation in the mental lexicon of children, as in the case of pseudo words” (p.506) is absent, kindergarten children and first graders performed alike. This is attributed to the fact that increased familiarity with the phonological representation of MSA facilitates the processing of words lexically affiliated to MSA (p.504-5). Hence, the phonological form of the word has a minor role in adult speakers’ awareness of the MSA/ECA status of a word as compared to its lexical affiliation. In other words, in diglossic speech, the native speaker considers an MSA lexeme standard irrespective of its pronunciation. Words that are lexically common to MSA and ECA are neutral and follow the variety of their phonetic specification.

Verb suffixes and mood are almost always ECA and are thus dropped from the analysis, and so is case marking on subjects and complements. Mejdell (2000) reasons that since suffixes are
usually unstressed phonologically, they are least noticed by listeners. In mixed speech, speakers attempt to sound more educated by incorporating MSA features in sentence structure. Hence, stylistically, suffixes “have little effect on the overall perception of more elevated style” (p. 19). Phonetically, Saiegh-Haddad (2004) observes that “the isolation of initial onset phonemes was consistently more difficult than the isolation of final phonemes” (p.507) and relates this to the more difficulty of isolating prevocalic consonants as compared to post vocal final consonants even when the vowel is embedded in a complex consonantal cluster rime coda. She notes that this result for Arabic speaking children is at odds with English speaking ones, and proposes that “language-specific prosodic constraints may affect the ease with which various intrasyllabic units are orally produced as the outcome of phonological analysis (p.507). The implication of her observation in relation to the present study is that suffixes are least attached to the stem, hence, more apt to be modified or dropped in diglossic speech.

In contrast, the link between the prefix and stem is stronger according to Saiegh-Haddad. Prefixes are also considered “the decisive linguistic feature in assigning the clause to level (2) (fuSha l-asSr) [MSA] or (3) (‘ammiyyat al-muthaqqafin) [ESA]” as Badawi claims (Mejdell, 2000:12). My data shows that MSA subject prefixes are always followed by MSA imperfective stems. Likewise, an MSA future marker is followed by an MSA prefix or stem. For example، ستبقى [sa-ta-bqā] will remain is a sequence of an MSA future marker [sa], MSA subject prefix [ta], and an MSA imperfective stem [bqā]. Only an ECA subject prefix may adjoin a verb from either variety, e.g. the verb to see may be expressed by an ECA prefix and an MSA stem: ينظر [yu-nō’ur], or an ECA prefix and an ECA stem يشوف [yi-ʃūf]. If the prefix is MSA, the second
option is not allowed and only \[\text{ينظر}\] [ya-n\text{ð}ur] is possible. This means that unless the higher functional element is ECA, CS does not occur.

Myers-Scotton (2010:94) states that “early system morphemes in SA [standard Arabic] may appear frequently in ML frames [ECA]”, and are activated at the same time with content morphemes. Hence, when a speaker selects an MSA verb, its MSA morphological specifications are activated concurrently. ECA prefixes are also available because system morphemes are normally provided by ML. In contrast, when an ECA verb stem is elected, only ECA, not MSA, prefixes are accessible. Thus, CS at this level only involves an ECA system morpheme and an MSA content morpheme as predicted by MLF. Similar mixings at the word-level occur in prefixing Bantu object clitics to English or French roots, or bound morphemes to nouns and verbs (Bokamba, 1987).

The aspectual bi- prefix may precede an imperfective from either variety. An aspectual bi- may precede an MSA verb, whether the subject/tense prefix is MSA or ECA. Using the verb *to see*, the subject prefix following an aspectual bi- may be MSA as in *\[\text{ينظر}\text{\ðūf}\]*, or it may be colloquialized as in *\[\text{يُنَظر}\text{\ðūf}\]*. In the first case, the verb stem cannot be colloquial *\[\text{يُنَظر}\text{\ðūf}\]*. The versatility of the aspectual prefix agrees with the findings of Bassiouney (2006).

Boussofara-Omar (2003) calls such co-occurrence of system morphemes from the two varieties “dual morphology”. Since ECA is the variety that sets the morpho-syntactic frame, it requires an explicit marker for aspect that MSA lacks. Hence, the occurrence of bi- may be forced by ML even when an MSA verb is elected.
The Directionality Constraint of Eid (1998) disallows switching only after an MSA negative marker. Bassiouney (2006) also found that the negative marker is drawn from ML, i.e. an ECA marker may negate an MSA verb. Her results do not distinguish between the elements that directly follow the negative marker. My results agree with these findings only with respect to the perfective verb, where the direction of CS is from ECA negative to MSA stem and not vice versa. For example, ما ماستثرش [ma-stuθir-ʃ] was not invoked is an MSA perfective verb surrounded by the ECA discontinuous negative marker [ma..ʃ]. At the interface between the negative marker and ASP, PX or verb stems other than the perfective, no CS is attested, i.e. the variety of the negative marker agrees with the variety of the verb, subject and/or aspectual prefix. This observation seems to require further investigation because forms like ما ميستثرش [ma-b-yu-staθar-ʃ] is not invoked would not sound ungrammatical to me as a native speaker. Here, the same ECA negative marker negates an MSA imperfective that has MSA subject and bi- prefixes. A wider corpus may demonstrate the possibility of CS between the negative and imperfective verbs. Also, negation of future verbs needs further analysis in more appropriate contexts that would show whether the same agreement between NEG and FUT is required.

Although the negative marker is a functional element like the prefix, future and aspectual markers, it behaves differently with respect to tense, aspect and mood. In section IV, it was argued that tense attaches to negation in MSA, but not in ECA. In the perfective, the ECA negative marker has a preverbal part, which is the same non-temporal MSA [mā], and a post-verbal part [-ʃ]. The MSA negated perfective gets tense by merging with the pre-verbal marker, then moving to T. Hence, CS between an ECA negative marker and MSA perfective is allowed. However, it is not possible to switch in the opposite direction; i.e. an MSA negative and ECA
perfective as in *لم راح [lam rāh] did not go, otherwise tense will be doubly realized on the MSA marker and the ECA verb.

The imperfective, on the other hand, moves to T when negated by [ma…ʃ] by the same reasoning as the perfective. However, it may stay in situ when the marker is [miʃ]. Since [miʃ] doesn’t carry tense, it cannot negate a tenseless MSA verb, e.g. مش يذهب [miʃ ya-ðhab] does not go, because tense will not be realized. Nor can an MSA marker [lā] negate an ECA imperfective, e.g. لا يروح [lā y-rūh] does not go because, in this case, tense will be doubly realized on the marker and the verb.

Like the imperfective, participle and mood also agree in variety with the negative marker. The only MSA negative markers used with participles in the data is the non-temporal [ɣēr]; which is phonologically adapted from [ɣayr]. Modal negation varies according to the modal verb form.

Negation is clearly linked syntactically to the verb rather than the subject and complement, and this explains why its behavior in connection to CS with respect to verb arguments varies among different verb classes. Also, the tendency to switch from ECA subject to MSA negative or affirmative verb is much stronger than the other direction. That this directionality is more notable with imperfective verbs is supported by the already mentioned constraint on CS between the subject prefix and verb stem. CS with the verb stem is only allowed when the subject prefix is ECA, and is preferred when the lexical subject is ECA. The constraint is relatively weaker in the latter case because a lexical subject is a free morpheme and more distant from the verb.
As for CS between the complement and the verb, it is rarer, possibly because of the government relationship that binds them. Although the validity of the Government Principle (Di Sciullo, Muysken & Singh, 1986) is still controversial and CS does occur between the verb and its complement in Arabic as well as other languages (Clyne, 2000; Bentahila & Davies, 1983; & Redouane, 2005), it appears that the different syntactic relationships between subjects and complements and their verb do influence the relative differences in their recorded CS frequencies. This agrees with the prediction made in section III-2 that Arabic is of the type HS >> LPC & *SPEC >> COMP. CS from an MSA verb to the complement is unfavorable because case assigned by the verb is never realized on ECA nouns. However, the variety of the subject in specifier position may switch without switching the head verb.

In general, the results confirm ECA > MSA directionality bias of CS and show that it does not only stem from the functional vs. content morpheme dichotomy as MLF predicts, but also pertains to the hierarchy of the phrase structure. It applies to PX → stem, FUT/ASP → PX, NEG → ASP/PX, as well as SUB → NEG. The switch between the negative and the future, and the subject and future when the FUT marker is MSA needs further investigation with a larger number of standard future forms in spontaneous speech.

The above results hold true for the three discourse types investigated in this study. The absence of MSA future and negation markers in sports discourse is not surprising. Poplack (1980) highlighted the impact of proficiency in EL on the frequency and quality of CS between English and Spanish. The same applies to CS between literary and spoken Sinhala (Paolillo, 2000). It is interesting that the frequency of MSA future marker in political/economic discourse exceeds its
frequency in religious speech contrary to the expectation that religious topics are more influenced by CA and MSA. However, MSA future markers are too rare in the data to make any generalizations. The interaction between the ECA aspectual prefix and the phonetic realization of the subject prefix and internal vowel change of the MSA verb is particularly enlightening, because the MSA verb is least affected in religious discourse, and is most yielding in sports, while political/economic talk is intermediate between the two. This is an indication that the tendency to abide by MSA phonology of the verb is strongest in religious discourse and least in sports. However, the result requires more substantiation from a larger amount of data. It is also clear that religious discourse gets its MSA flavor largely from the abundance of religious text quotes rather than from the input of the speakers themselves. This is attested most in the realization of case and mood markers because by extracting all quotations, the incorporation of MSA elements in religious discourse drops drastically.

In conclusion, the results of this study confirm previous findings about the role of aspectual prefix and negative marker. I extensively investigated all functional and content constituents in the verb phrase in relation to each other. Data from every category of tense, aspect, and modality is compared to the other, in order to find out whether the tendencies observed agree or differ among verb classes. In addition, I compared three different sociolinguistic discourses in order to test the consistency or variance of the findings among them. The results were largely consistent and point to the fact that diglossic speech in Egyptian Arabic relies on the variety of functional morphemes. Markers for negation, future, aspect, besides the subject prefix constrain the switch when they are drawn from MSA.
I used a novel procedure that encodes lexical, phonological, and syntactic characteristics of every element independently. This allowed me to inspect lexical, phonological and morpho-syntactic interactions in CS from the morphemic to the phrasal level. Addressing all these issues simultaneously by virtue of a simple yet all-inclusive coding procedure and computational analysis uncovered how rich and dynamic the mixing between MSA and ECA actually is. I stratified the clause into five levels on the assumption that one constituent is the determiner of CS in the level it belongs to, and passes over its variety specification to the whole level. Consecutive levels build up by incorporating every level with an additional constituent. This procedure has the advantage of combining linear and hierarchical structures in a morpho-syntactic analysis of CS, and the results obtained provide a systematic characterization of the mixed code in spoken Arabic.

Regarding CS in verbal clauses, there are five implications that can be drawn from this study. First, the ECA to MSA directionality bias is an outcome of factors that may equally apply to diglossic or separate languages that have the same word order. These factors are the dominance of L1 grammar particularly its functional categories as predicted by MLF, hierarchical structure of constituents in the clause, and the syntactic relationships between constituents. Second, suffixes including mood and case markers are the least constituent that may affect CS because they are more susceptible to detachment from the verb or noun. They have little impact on affiliating a constituent to one particular variety. This means that the least determiner of variety specification is syntactic, but the lexical followed by the phonological specifications are stronger. Third, temporal, aspectual, and modal specifications in Arabic do not influence the nature or frequency of CS, except when negation is present. Fourth, not all mixing is emblematic because
CS may occur within one constituent or between two or more constituents without affecting the overall designation of the phrase as MSA, ECA, or hybrid. Finally, the type of discourse does not impose any extra constraints on CS, but it does interfere in the interaction between the two varieties at the morphemic and phrasal levels, in a way that favors particular sets of CS combinations. This imparts a distinctive and noticeable coloring to the discourse whether religious, political or sport like.

The “distinctive and noticeable coloring” of the discourse needs a more in depth morpho-syntactic inspection, as well as sociolinguistic by taking a wider range of speakers of the same discourse and comparing their performance morph-syntactically. This would establish the findings of this study more firmly or modify them.

As an intuitive language behavior and a pragmatic language tool, the impact of CS within and across phrasal constituents is not even. The results point to how natives impart disproportionate weights on various morphemes and constituents in their manipulation or perception of MSA and ECA in a multi-layer framework of the verb clause. The experiment performed by Hary (1996) sheds some light on how natives perceive the mixed code of Arabic. Subjects in Hary’s experiment ranked all combinations of MSA-ECA morphemes in ‘I saw him’ on a scale of increasing standardization. Some combinations were proven unacceptable, while possible hybrid forms gave significant indications of some CS rules involved in the transformation from one code to the other. A follow up experiment that tests the results of this study by monitoring CS points as identified by listeners, and measuring the strength of their response will be a valuable addition to the implications of this study.
Maamouri (1998) maintains that “The linguistic spread of colloquial Arabic all over the Arab region will eventually lead to the necessary reform of the Arabic language standards used in the Arab educational structure” (p.68). This directly relates to the topic of this research since understanding the mutual impact of the colloquial and standard on each other is a pre-requisite to the production of a “Contemporary Grammar of Arabic” that “would include language patterns from the various literary periods and styles in written fusha [MSA], representative samples from the written media currently used in the Arab countries, and a recording of samples of representative and varied oral production from local and regional varieties of colloquial and fusha Arabic” (p.64). The coding procedure and hierarchical analysis employed in the current study is applicable to other morpho-syntactic and phrasal structures of Arabic, and is highly pertinent to computational corpus analyses.
References


*Lingua 59*: 301-30.


## Appendix A: Transcription gloss

<table>
<thead>
<tr>
<th>Consonants</th>
<th>Vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>ء</td>
<td>ء</td>
</tr>
<tr>
<td>ب</td>
<td>b</td>
</tr>
<tr>
<td>ت</td>
<td>t</td>
</tr>
<tr>
<td>ث</td>
<td>θ</td>
</tr>
<tr>
<td>ج</td>
<td>g / ڇ</td>
</tr>
<tr>
<td>ح</td>
<td>h</td>
</tr>
<tr>
<td>خ</td>
<td>x</td>
</tr>
<tr>
<td>د</td>
<td>d</td>
</tr>
<tr>
<td>ذ</td>
<td>δ / z</td>
</tr>
<tr>
<td>ر</td>
<td>r</td>
</tr>
<tr>
<td>ز</td>
<td>z</td>
</tr>
<tr>
<td>س</td>
<td>s</td>
</tr>
<tr>
<td>ش</td>
<td>ʃ</td>
</tr>
<tr>
<td>ص</td>
<td>sʰ</td>
</tr>
<tr>
<td>ض</td>
<td>d♯</td>
</tr>
<tr>
<td>ط</td>
<td>tʰ</td>
</tr>
<tr>
<td>ظ</td>
<td>ṣ̄ʰ / zʰ</td>
</tr>
<tr>
<td>ع</td>
<td>ʕ</td>
</tr>
<tr>
<td>غ</td>
<td>ɣ</td>
</tr>
<tr>
<td>ف</td>
<td>f</td>
</tr>
<tr>
<td>ق</td>
<td>q</td>
</tr>
<tr>
<td>ك</td>
<td>k</td>
</tr>
<tr>
<td>ل</td>
<td>l</td>
</tr>
<tr>
<td>م</td>
<td>m</td>
</tr>
<tr>
<td>ن</td>
<td>n</td>
</tr>
<tr>
<td>ه</td>
<td>h</td>
</tr>
<tr>
<td>و</td>
<td>w</td>
</tr>
<tr>
<td>ي</td>
<td>y</td>
</tr>
</tbody>
</table>
Appendix B: Abbreviations

Tense, Aspect and Modality

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRF</td>
<td>Perfective</td>
</tr>
<tr>
<td>IMP</td>
<td>Imperfective</td>
</tr>
<tr>
<td>MPV</td>
<td>Imperative</td>
</tr>
<tr>
<td>PSS</td>
<td>Passive</td>
</tr>
<tr>
<td>PCP</td>
<td>Participle</td>
</tr>
<tr>
<td>PRG</td>
<td>Progressive</td>
</tr>
<tr>
<td>HAB</td>
<td>Habitual</td>
</tr>
<tr>
<td>ITR</td>
<td>Iterative</td>
</tr>
<tr>
<td>MD</td>
<td>Modal</td>
</tr>
</tbody>
</table>

Number and Gender

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pl</td>
<td>Plural</td>
</tr>
<tr>
<td>sg</td>
<td>Singular</td>
</tr>
<tr>
<td>M</td>
<td>Masculine</td>
</tr>
<tr>
<td>F</td>
<td>Feminine</td>
</tr>
</tbody>
</table>

Case

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>Nominative</td>
</tr>
<tr>
<td>ACC</td>
<td>Accusative</td>
</tr>
<tr>
<td>GEN</td>
<td>Genitive</td>
</tr>
</tbody>
</table>
### Mood

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IND</td>
<td>Indicative</td>
</tr>
<tr>
<td>SBJ</td>
<td>Subjunctive</td>
</tr>
<tr>
<td>JSS</td>
<td>Jussive</td>
</tr>
<tr>
<td>NEG</td>
<td>Negative</td>
</tr>
<tr>
<td>CND</td>
<td>conditional conjunction</td>
</tr>
</tbody>
</table>

### Morphemes

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>verb</td>
</tr>
<tr>
<td>PX</td>
<td>prefix</td>
</tr>
<tr>
<td>SX</td>
<td>suffix</td>
</tr>
<tr>
<td>FUT</td>
<td>future marker</td>
</tr>
<tr>
<td>ASP</td>
<td>aspectual prefix</td>
</tr>
<tr>
<td>NEG</td>
<td>negative marker</td>
</tr>
<tr>
<td>S</td>
<td>subject</td>
</tr>
<tr>
<td>C</td>
<td>complement</td>
</tr>
</tbody>
</table>

### Levels of analysis

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>lex</td>
<td>lexical</td>
</tr>
<tr>
<td>phn</td>
<td>phonological</td>
</tr>
<tr>
<td>syn</td>
<td>syntactic</td>
</tr>
</tbody>
</table>
Appendix C: CS comparison by Tense, Aspect, Modality

Appendix C-1: Collective results for verb level

Total verb counts = 3118

<table>
<thead>
<tr>
<th>PRF count = 1058</th>
<th>PX-IMP-SX count = 1364</th>
<th>FUT-PX count = 171</th>
<th>PCP count = 233</th>
<th>MD count = 292</th>
</tr>
</thead>
<tbody>
<tr>
<td>prf pss11 sx0</td>
<td>px0 imp00 sx0</td>
<td>fut1 px1 imp00 sx2</td>
<td>pcp00 sx2</td>
<td>md00 sx0</td>
</tr>
<tr>
<td>prf pss11 sx1</td>
<td>px0 imp00 sx1</td>
<td>fut1 px1 imp11 sx0</td>
<td>pcp01 sx1</td>
<td>md00 sx1</td>
</tr>
<tr>
<td>prf pss11 sx2</td>
<td>px0 imp00 sx2</td>
<td>fut1 px1 imp11 sx2</td>
<td>pcp01 sx2</td>
<td>md00 sx2</td>
</tr>
<tr>
<td>prf pss22 sx2</td>
<td>px0 imp01 sx0</td>
<td>fut2 px0 imp02 sx2</td>
<td>pcp02 sx2</td>
<td>md00 sx0</td>
</tr>
<tr>
<td>prf00 sx0</td>
<td>px0 imp01 sx1</td>
<td>fut2 px1 imp pss11 sx2</td>
<td>pcp11 sx1</td>
<td>md00 sx1</td>
</tr>
<tr>
<td>prf00 sx1</td>
<td>px0 imp01 sx1</td>
<td>fut2 px1 imp pss11 sx2</td>
<td>pcp11 sx2</td>
<td>md00 sx1</td>
</tr>
<tr>
<td>prf00 sx2</td>
<td>px0 imp02 sx0</td>
<td>fut2 px1 imp02 sx2</td>
<td>pcp11 sx2</td>
<td>md00 sx0</td>
</tr>
<tr>
<td>prf01 sx0</td>
<td>px0 imp02 sx2</td>
<td>fut2 px1 imp11 sx0</td>
<td>pcp00 sx2</td>
<td>md00 sx0</td>
</tr>
<tr>
<td>prf01 sx1</td>
<td>px0 imp11 sx0</td>
<td>fut2 px1 imp12 sx2</td>
<td>pcp20 sx2</td>
<td>md11 sx0</td>
</tr>
<tr>
<td>prf01 sx2</td>
<td>px0 imp11 sx2</td>
<td>fut2 px2 imp pss22 sx2</td>
<td>pcp22 sx2</td>
<td>md11 sx1</td>
</tr>
<tr>
<td>prf02 sx0</td>
<td>px0 imp22 sx2</td>
<td>fut2 px2 imp00 sx0</td>
<td>pcp00 sx2</td>
<td>md12 sx2</td>
</tr>
<tr>
<td>prf02 sx2</td>
<td>px1 imp pss00 sx2</td>
<td>fut2 px2 imp00 sx2</td>
<td>md20 sx2</td>
<td></td>
</tr>
<tr>
<td>prf10 sx2</td>
<td>px1 imp pss11 sx0</td>
<td>fut2 px2 imp02 sx0</td>
<td>md22 sx0</td>
<td></td>
</tr>
<tr>
<td>prf11 sx0</td>
<td>px1 imp pss11 sx1</td>
<td>fut2 px2 imp02 sx2</td>
<td>md22 sx2</td>
<td></td>
</tr>
<tr>
<td>prf11 sx1</td>
<td>px1 imp pss11 sx2</td>
<td>fut2 px2 imp11 sx2</td>
<td>md22 sx2</td>
<td></td>
</tr>
</tbody>
</table>

198
<table>
<thead>
<tr>
<th>prf11 sx2</th>
<th>68</th>
<th>px1 imp00 sx0</th>
<th>13</th>
<th>fut2 px2 imp12 sx2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>prf12 sx2</td>
<td>6</td>
<td>px1 imp00 <strong>sx1</strong></td>
<td>16</td>
<td>fut2 px2 imp22 sx0</td>
<td>26</td>
</tr>
<tr>
<td>prf20 sx0</td>
<td>2</td>
<td>px1 imp00 sx2</td>
<td>39</td>
<td>fut2 px2 imp22 sx2</td>
<td>18</td>
</tr>
<tr>
<td>prf20 sx2</td>
<td>8</td>
<td>px1 imp01 sx0</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf21 sx2</td>
<td>1</td>
<td>px1 imp01 <strong>sx1</strong></td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx0</td>
<td>24</td>
<td>px1 imp01 sx2</td>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td>109</td>
<td>px1 imp02 sx0</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px1 imp11 sx0</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px1 imp11 <strong>sx1</strong></td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px1 imp11 sx2</td>
<td>105</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px2 imp pss22 sx2</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px2 imp00 sx0</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px2 imp00 <strong>sx1</strong></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px2 imp00 sx2</td>
<td>253</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px2 imp01 sx2</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px2 imp02 sx0</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px2 imp02 sx2</td>
<td>396</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px2 imp11 sx0</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px2 imp11 <strong>sx1</strong></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px2 imp11 sx2</td>
<td>28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px2 imp12 sx0</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prf22 sx2</td>
<td></td>
<td>px2 imp12 sx2</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>px2 imp20 sx0</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>px2 imp20 sx2</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>px2 imp21 sx2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>px2 imp22 sx0</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>px2 imp22 sx2</td>
<td>138</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix C-2: Collective results for subject-verb with no intervening markers

Total subject-verb counts = 1524

<table>
<thead>
<tr>
<th>SUB-PRF = 562</th>
<th>SUB-PX = 655</th>
<th>SUB-FUT = 83</th>
<th>SUB-PCP = 160</th>
<th>SUB-MD = 64</th>
</tr>
</thead>
<tbody>
<tr>
<td>s000 prf pss11 1</td>
<td>s000 px0 5</td>
<td>s000 fut1 1</td>
<td>s000 pcp00 2</td>
<td>s000 md00 1</td>
</tr>
<tr>
<td>s000 prf00 9</td>
<td>s000 px1 18</td>
<td>s001 fut1 1</td>
<td>s000 pcp01 1</td>
<td>s000 md22 2</td>
</tr>
<tr>
<td>s000 prf01 1</td>
<td>s000 px2 38</td>
<td>s010 fut1 1</td>
<td>s000 pcp02 1</td>
<td>s002 md00 1</td>
</tr>
<tr>
<td>s000 prf02 9</td>
<td>s000 px1 1</td>
<td>s110 fut1 1</td>
<td>s000 pcp00 2</td>
<td>s002 md02 15</td>
</tr>
<tr>
<td>s000 prf11 4</td>
<td>s002 px0 1</td>
<td>s112 fut1 1</td>
<td>s000 pcp20 1</td>
<td>s002 md11 3</td>
</tr>
<tr>
<td>s000 prf12 2</td>
<td>s002 px1 42</td>
<td>s000 fut2 5</td>
<td>s000 pcp22 9</td>
<td>s002 md20 3</td>
</tr>
<tr>
<td>s000 prf22 3</td>
<td>s002 px2 51</td>
<td>s002 fut2 11</td>
<td>s002 pcp00 1</td>
<td>s002 md22 6</td>
</tr>
<tr>
<td>s001 prf pss11 1</td>
<td>s010 px0 1</td>
<td>s020 fut2 28</td>
<td>s002 pcp01 2</td>
<td>s012 md11 1</td>
</tr>
<tr>
<td>s001 prf01 6</td>
<td>s010 px1 6</td>
<td>s022 fut2 13</td>
<td>s002 pcp02 2</td>
<td>s020 md00 2</td>
</tr>
<tr>
<td>s001 prf11 6</td>
<td>s011 px1 2</td>
<td>s102 fut2 1</td>
<td>s002 pcp11 6</td>
<td>s020 md02 9</td>
</tr>
<tr>
<td>s002 prf pss11 1</td>
<td>s012 px1 11</td>
<td>s112 fut2 3</td>
<td>s002 pcp20 1</td>
<td>s020 md12 1</td>
</tr>
<tr>
<td>s002 prf pss22 1</td>
<td>s012 px2 1</td>
<td>s220 fut2 10</td>
<td>s002 pcp22 5</td>
<td>s020 md22 7</td>
</tr>
<tr>
<td>s002 prf00 45</td>
<td>s020 px0 6</td>
<td>s222 fut2 7</td>
<td>s010 pcp00 2</td>
<td>s022 md22 1</td>
</tr>
<tr>
<td>s002 prf01 5</td>
<td>s020 px1 29</td>
<td></td>
<td>s010 pcp01 3</td>
<td>s110 md11 2</td>
</tr>
<tr>
<td>s002 prf02 24</td>
<td>s020 px2 160</td>
<td></td>
<td>s010 pcp11 1</td>
<td>s112 md20 1</td>
</tr>
<tr>
<td>s002 prf11 22</td>
<td>s022 px0 1</td>
<td></td>
<td>s010 pcp22 1</td>
<td>s220 md00 1</td>
</tr>
<tr>
<td>s002 prf12 1</td>
<td>s022 px1 21</td>
<td></td>
<td>s012 pcp11 1</td>
<td>s220 md02 3</td>
</tr>
<tr>
<td></td>
<td>prf</td>
<td>px1</td>
<td>px2</td>
<td>pc</td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>s002 prf22</td>
<td>12</td>
<td>49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s010 prf pss11</td>
<td>1</td>
<td></td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>s010 prf01</td>
<td>3</td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>s011 prf01</td>
<td>2</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>s011 prf11</td>
<td>6</td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>s012 prf00</td>
<td>1</td>
<td></td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>s012 prf01</td>
<td>2</td>
<td></td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>s012 prf02</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>s012 prf11</td>
<td>2</td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>s012 prf22</td>
<td>1</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>s020 prf pss11</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>s020 prf pss22</td>
<td>4</td>
<td></td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>s020 prf00</td>
<td>29</td>
<td></td>
<td></td>
<td>70</td>
</tr>
<tr>
<td>s020 prf01</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>s020 prf02</td>
<td>39</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>s020 prf11</td>
<td>18</td>
<td></td>
<td></td>
<td>37</td>
</tr>
<tr>
<td>s020 prf12</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s020 prf22</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s022 prf pss22</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s022 prf00</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s022 prf01</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s022 prf02</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>Value</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s022 prf11</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s022 prf20</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s022 prf22</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s110 prf pss11</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s110 prf pss22</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s110 prf00</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s110 prf01</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s110 prf11</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s110 prf20</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s111 prf pss11</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s111 prf01</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s111 prf11</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s112 prf pss11</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s112 prf00</td>
<td>34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s112 prf01</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s112 prf02</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s112 prf11</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s112 prf12</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s112 prf22</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s120 prf01</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s120 prf11</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s220 prf pss22</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s222 pcp01</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s222 pcp02</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s222 pcp11</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s222 pcp22</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s000 pcp00</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s222 pcp22</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s222 pcp11</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Key</td>
<td>Value</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s220 prf00</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s220 prf01</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s220 prf02</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s220 prf10</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s220 prf11</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s220 prf20</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s220 prf22</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s222 prf pss11</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s222 prf00</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s222 prf02</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s222 prf11</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s222 prf22</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix C-3: Collective results for complement-verb with no intervening markers

Total complement-verb counts = 2182

<table>
<thead>
<tr>
<th>COMP-PRF =</th>
<th>COMP-PX =</th>
<th>COMP-FUT =</th>
<th>COMP-PCP =</th>
<th>COMP-MD =</th>
</tr>
</thead>
<tbody>
<tr>
<td>c000 prf</td>
<td>c000 px1</td>
<td>c002 px1</td>
<td>c002 px0</td>
<td>c002 px2</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>8</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>c000 prf</td>
<td>c002 px1</td>
<td>c020 px1</td>
<td>c020 px2</td>
<td>c020 px1</td>
</tr>
<tr>
<td>pss11</td>
<td>19</td>
<td>32</td>
<td>31</td>
<td>27</td>
</tr>
<tr>
<td>c000 prf</td>
<td>c000 px1</td>
<td>c000 px1</td>
<td>c002 px2</td>
<td>c002 px0</td>
</tr>
<tr>
<td>pss22</td>
<td>113</td>
<td>119</td>
<td>94</td>
<td>93</td>
</tr>
<tr>
<td>c000 prf</td>
<td>c000 px1</td>
<td>c000 px1</td>
<td>c002 px2</td>
<td>c002 px1</td>
</tr>
<tr>
<td>pr00</td>
<td>1</td>
<td>6</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>c000 prf</td>
<td>c000 px1</td>
<td>c000 px1</td>
<td>c002 px2</td>
<td>c002 px1</td>
</tr>
<tr>
<td>pr01</td>
<td>5</td>
<td>6</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>c000 prf</td>
<td>c000 px1</td>
<td>c000 px1</td>
<td>c002 px2</td>
<td>c002 px1</td>
</tr>
<tr>
<td>pr02</td>
<td>5</td>
<td>6</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>c000 prf</td>
<td>c000 px1</td>
<td>c000 px1</td>
<td>c002 px2</td>
<td>c002 px1</td>
</tr>
<tr>
<td>pr10</td>
<td>5</td>
<td>6</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>c000 prf</td>
<td>c000 px1</td>
<td>c000 px1</td>
<td>c002 px2</td>
<td>c002 px1</td>
</tr>
<tr>
<td>pr11</td>
<td>5</td>
<td>6</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>c000 prf</td>
<td>c000 px1</td>
<td>c000 px1</td>
<td>c002 px2</td>
<td>c002 px1</td>
</tr>
<tr>
<td>pr12</td>
<td>5</td>
<td>6</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>c000 prf</td>
<td>c000 px1</td>
<td>c000 px1</td>
<td>c002 px2</td>
<td>c002 px1</td>
</tr>
<tr>
<td>pr20</td>
<td>5</td>
<td>6</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>c000 prf</td>
<td>c000 px1</td>
<td>c000 px1</td>
<td>c002 px2</td>
<td>c002 px1</td>
</tr>
<tr>
<td>pr22</td>
<td>5</td>
<td>6</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>c000 prf</td>
<td>c000 px1</td>
<td>c000 px1</td>
<td>c002 px2</td>
<td>c002 px1</td>
</tr>
<tr>
<td>pr00</td>
<td>5</td>
<td>6</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>c000 prf</td>
<td>c000 px1</td>
<td>c000 px1</td>
<td>c002 px2</td>
<td>c002 px1</td>
</tr>
<tr>
<td>pr01</td>
<td>5</td>
<td>6</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>c000 prf</td>
<td>c000 px1</td>
<td>c000 px1</td>
<td>c002 px2</td>
<td>c002 px1</td>
</tr>
<tr>
<td>pr02</td>
<td>5</td>
<td>6</td>
<td>32</td>
<td>31</td>
</tr>
<tr>
<td>Code</td>
<td>Type</td>
<td>Value</td>
<td>Code</td>
<td>Type</td>
</tr>
<tr>
<td>------------</td>
<td>----------</td>
<td>-------</td>
<td>------------</td>
<td>----------</td>
</tr>
<tr>
<td>c002 prf11</td>
<td></td>
<td>14</td>
<td>c022 px1</td>
<td></td>
</tr>
<tr>
<td>c002 prf20</td>
<td></td>
<td>1</td>
<td>c022 px2</td>
<td></td>
</tr>
<tr>
<td>c002 prf22</td>
<td></td>
<td>5</td>
<td>c101 px1</td>
<td></td>
</tr>
<tr>
<td>c010 prf pss11</td>
<td></td>
<td>5</td>
<td>c110 px0</td>
<td></td>
</tr>
<tr>
<td>c010 prf00</td>
<td></td>
<td>3</td>
<td>c110 px1</td>
<td></td>
</tr>
<tr>
<td>c010 prf01</td>
<td></td>
<td>4</td>
<td>c110 px2</td>
<td></td>
</tr>
<tr>
<td>c010 prf11</td>
<td></td>
<td>5</td>
<td>c111 px1</td>
<td></td>
</tr>
<tr>
<td>c011 prf00</td>
<td></td>
<td>1</td>
<td>c111 px2</td>
<td></td>
</tr>
<tr>
<td>c011 prf01</td>
<td></td>
<td>1</td>
<td>c112 px1</td>
<td></td>
</tr>
<tr>
<td>c011 prf11</td>
<td></td>
<td>5</td>
<td>c112 px2</td>
<td></td>
</tr>
<tr>
<td>c012 prf00</td>
<td></td>
<td>4</td>
<td>c120 px2</td>
<td></td>
</tr>
<tr>
<td>c020 prf pss22</td>
<td></td>
<td>1</td>
<td>c122 px1</td>
<td></td>
</tr>
<tr>
<td>c020 prf00</td>
<td></td>
<td>49</td>
<td>c122 px2</td>
<td></td>
</tr>
<tr>
<td>c020 prf01</td>
<td></td>
<td>4</td>
<td>c200 px1</td>
<td></td>
</tr>
<tr>
<td>c020 prf02</td>
<td></td>
<td>72</td>
<td>c220 px0</td>
<td></td>
</tr>
<tr>
<td>c020 prf10</td>
<td></td>
<td>1</td>
<td>c220 px1</td>
<td></td>
</tr>
<tr>
<td>c020 prf11</td>
<td></td>
<td>12</td>
<td>c220 px2</td>
<td></td>
</tr>
<tr>
<td>c020 prf12</td>
<td></td>
<td>1</td>
<td>c222 px0</td>
<td></td>
</tr>
<tr>
<td>c020 prf20</td>
<td></td>
<td>2</td>
<td>c222 px1</td>
<td></td>
</tr>
<tr>
<td>c020 prf22</td>
<td></td>
<td>25</td>
<td>c222 px2</td>
<td></td>
</tr>
<tr>
<td>c022 prf pss22</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c022 prf00</td>
<td></td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>Count</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c022 prf02</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c022 prf11</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c022 prf12</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c022 prf22</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c102 prf11</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c110 prf pss11</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c110 prf00</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c110 prf01</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c110 prf02</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c110 prf11</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c110 prf22</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>c111</strong> prf00</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>c111</strong> prf01</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>c111</strong> prf11</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c112 prf00</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c112 prf02</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c112 prf10</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c112 prf11</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c112 prf12</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c112 prf22</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c120 prf00</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c120 prf02</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Code</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>c222 md02</td>
<td>2</td>
</tr>
<tr>
<td>c222 md20</td>
<td>1</td>
</tr>
<tr>
<td>c222 md22</td>
<td>5</td>
</tr>
<tr>
<td>Code</td>
<td>Value</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
</tr>
<tr>
<td>c120 prf11</td>
<td>1</td>
</tr>
<tr>
<td>c122 prf00</td>
<td>3</td>
</tr>
<tr>
<td>c122 prf02</td>
<td>2</td>
</tr>
<tr>
<td>c122 prf11</td>
<td>3</td>
</tr>
<tr>
<td>c202 prf02</td>
<td>3</td>
</tr>
<tr>
<td>c220 prf00</td>
<td>19</td>
</tr>
<tr>
<td>c220 prf01</td>
<td>1</td>
</tr>
<tr>
<td>c220 prf02</td>
<td>7</td>
</tr>
<tr>
<td>c220 prf11</td>
<td>2</td>
</tr>
<tr>
<td>c220 prf22</td>
<td>5</td>
</tr>
<tr>
<td>c222 prf00</td>
<td>14</td>
</tr>
<tr>
<td>c222 prf02</td>
<td>8</td>
</tr>
<tr>
<td>c222 prf22</td>
<td>1</td>
</tr>
</tbody>
</table>
### Appendix C-4: Collective results for case-marked complement-subject combinations

<table>
<thead>
<tr>
<th>DR48.</th>
<th>c011</th>
<th>s011</th>
<th>prf11 sx1</th>
<th>wa štaːval-a r-raʔsu jaybā</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR306.</td>
<td>c001</td>
<td>s001</td>
<td>prf11 sx1</td>
<td>fa ʔinna l-lāha qad ʔabd-al-a sayyiʔāt-i-ka kulliha ʔila ◊hasanāt</td>
</tr>
<tr>
<td>HR132.</td>
<td>c001</td>
<td>s111</td>
<td>a0 px0 imp00 sx1</td>
<td>lā ya?man-u makr-a lā ʔilla l-qawm-u l-xāsir-ʔn</td>
</tr>
<tr>
<td>HR161.</td>
<td>c111</td>
<td>s001</td>
<td>a1 px1 imp11 sx0</td>
<td>ʔallahu ya-tawaffa-ʔ l-ʔanfus-a hīna mawt-i-ha</td>
</tr>
<tr>
<td>MR424.</td>
<td>c001</td>
<td>s001</td>
<td>a1 px1 imp00 sx1</td>
<td>fa ʔulāʔika yu-baddil-u lā sayyiʔāt-i-him ◊hasanāt</td>
</tr>
</tbody>
</table>

وشتغل الرأس شيبا
And my head is shining with grey hair (Qur’an 19:4)

فإن الله قد أبدل سيئاتك كلها إلى حسنات
Allah has changed you evil deeds to good deeds (equivalent to Qu’an 25:7)

لا يأمن مكر الله إلا القوم الخاسرون
None feels secure from the Plan of Allah except the people who are the losers (Qur’an 7:99)

الله يتوفي الأنفس حين موتها
It is Allah Who takes away the souls at the time of their death (Qur’an 39:42)

فأولئك يبدل الله سيناتهم حسنات
Allah has changed you evil deeds to good deeds (Qur’an 25:7)
And Allah sets forth parables (Qur’an 14:25)
Appendix D: CS comparison by discourse

Appendix D-1: MSA verb stems with ECA aspectual prefix

**Religious discourse**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Verb Stem</th>
<th>ECA Aspectual Prefix</th>
<th>English Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DR 33. a2 px1 imp11 sx2</td>
<td>بِيَاقِبِلَهَا</td>
<td>b-yu-qābilha</td>
<td>it contrasts with it</td>
</tr>
<tr>
<td>DR 32. a2 px2 imp11 sx2</td>
<td>بَيْتِلْفَنَا</td>
<td>b-ti-lfit-na</td>
<td>it draws out attention</td>
</tr>
<tr>
<td>HR 113. a2 px2 imp11 sx2</td>
<td>بَيْاَسِمِرَ</td>
<td>b-a-stamir</td>
<td>I continue</td>
</tr>
<tr>
<td>HR 114. a2 px2 imp11 sx2</td>
<td>بَيْانَظِرَ</td>
<td>b-a-nāfur</td>
<td>I look at</td>
</tr>
<tr>
<td>HR 154. a2 px2 imp11 sx2</td>
<td>بَيْاَسْتِخِفَ</td>
<td>b-a-stajif</td>
<td>I infer</td>
</tr>
<tr>
<td>MR 440. a2 px2 imp11 sx2</td>
<td>مَا بِيْاَطاَلِعَ</td>
<td>ma-b-ya-t'ali'f</td>
<td>he does not know</td>
</tr>
<tr>
<td>MR 448. a2 px1 imp11 sx2</td>
<td>بَيْتِيَنَسُ</td>
<td>bi-tu-ya??is</td>
<td>it drives to despair</td>
</tr>
<tr>
<td>SR 233. a2 px1 imp11 sx2</td>
<td>بَيْاَنَكِمْلَ</td>
<td>bi-nu-kmil</td>
<td>we complete</td>
</tr>
<tr>
<td>SR 335. a2 px1 imp11 sx2</td>
<td>بَيْىَنَتِقَلَ</td>
<td>bi-ya-ntaqil</td>
<td>we move to</td>
</tr>
<tr>
<td>SR 338. a2 px1 imp11 sx2</td>
<td>بَيْىَنَتِقَلَ</td>
<td>b-ta-ntaqil</td>
<td>it is transferred to</td>
</tr>
<tr>
<td>SR 340. a2 px1 imp11 sx2</td>
<td>بَيْىَنَتِقَلَ</td>
<td>b-ta-ntaqil</td>
<td>they are transferred to</td>
</tr>
</tbody>
</table>

**Political discourse**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Verb Stem</th>
<th>English Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP 65a. a2 px1 imp11 sx0</td>
<td>بِيَسْتَخْدِمُونَهَا</td>
<td>b-ya-staxdim-u-ha</td>
</tr>
<tr>
<td>DP 160. a2 px1 imp11 sx2</td>
<td>بَيْيَرْتَزِقُونَوَ</td>
<td>b-ya-rtaziq-û</td>
</tr>
</tbody>
</table>
DP 44. a2 px2 impl2 sx2 باتوجه b-a-twaggih I head to
DP 122. a2 px2 impl2 sx2 بيتعدى b-yi-tʃadda he transgresses
HP 133. a2 px2 impl1 sx0 بناديها bi-n-ʃaddi-ʃ-ha we perform it
HP 135. a2 px2 impl1 sx2 بنعقد bi-nu-ʃqid we hold
HP 163. a2 px2 impl1 sx2 بيمثله bi-y-məθill-u he represents him
HP 169. a2 px2 impl1 sx2 بتمثلها bi-t-məθill-ha it represents it
HP 179. a2 px1 impl1 sx0 ببن بغتهما bi-na-btayi-ʃ-ha we desire it
HP 182. a2 px1 impl1 sx2 بتمثلهم bi-tu-məθill-hum it represents them
MP 52. a2 px1 impl1 sx2 بيزداد b-ya-zdād it increases
MP 53. a2 px1 impl1 sx2 بيننظر b-ya-nəˈur he views

Sports discourse

DS 106. a2 px1 impl1 sx2 بتصلح bi-ta-sliuh it fits
DS 222. a2 px2 impl1 sx2 باتتابع b-a-təbī` I follow
MS 2. a2 px1 impl1 sx2 بنحاور b-na-tahāwar we converse
MS 21. a2 px2 impl1 sx2 بيتوقع b-ti-twaqqaf you expect
MS 91. a2 px1 impl1 sx2 بتجتهد b-ta-gtahid I try hard
MS 68. a2 px2 impl1 sx2 بيراود b-ya-rāwid it entices
MS 71. a2 px2 impl1 sx0 وباسي b-a-sʃa-ʃ I seek
MS 199. a2 px2 impl1 sx2 ما بينتوافق ma-b-ti-twāfiq-f it does not accord with
bi-yu-dir  he manages
b-ya-ryab  he wants
b-a-gtahid  I try hard
b-ya-s'dur  it is issued
b-ti-t'allaq  it relates to
b-ti-ngiz  it accomplishes
Appendix D-2: MSA Future markers in corpus

Religious

DR 8.  c000 s001 fut1 px1 imp00 sx2 sa yakūn ismu-hu ʾala namat⁵-in yēr masbūq
       his name will be in an unprecedented fashion

Political

DP 40.  c000 s001 fut1 px1 imp11 sx2 ʾanna-hu sa yu-haqqiq ribh
       he will realize a profit

DP 41.  c000 s110 fut1 px1 imp11 sx0 min ʾayna sa-ta-ʾtī hādī-hi ij-farika bi hādīhi l-ʾarbāh
       this will company come up with these profits

DP 52.  c112 s010 fut1 px1 imp11 sx0 ʾanna-hu sa-ya-talaqqā mlāl hādīhi l-fawāyid
       that he will get like these profits

HP 9.   c000 s112 fut1 px1 imp11 sx0 sa-ta-bqā niqābat il-muhāmiyyīn bi l-qāhira tahta l-hirāsa
       the lawyers syndicate will remain under guard

MP 161. c110 fut1 px1 imp11 sx2 sa-tanʿfat⁶ hādīhi l-ʾamaliyya
        this process will be active

ستأتي هذه الشركة بهذه الأرباح
this will company come up with these profits
Appendix D-3: MSA Negative markers in corpus

**a-Imperfective**

**Religious**

| SR | n1 a1 px1 imp11 sx1 | lā-ya-statīr-∅ | لا يستتر | he does not hide |
| DR 51. | n1 a1 px1 imp00 sx2 | lā ya-mūt-∅ | لا يموت | he does not die |
| DR 61. | n1 a1 px1 imp01 sx2 | lā ya-qraʾ-∅ | لا يقرأ | he does not read |
| DR 62. | n1 a1 px1 imp01 sx2 | lā ya-ktub-∅ | لا يكتب | he does not write |
| HR 120. | n1 a0 px1 imp11 sx0 | lā ta-nūr | لا تنظر | do not look |
| HR 127. | n1 a0 px0 imp01 sx1 | lā ḥā-man-u | لا آمن | I do not feel safe |
| HR 146. | n1 a0 px0 imp00 sx0 | lā ḥa-rḍā | لا أرضي | I am not satisfied |
| HR 214. | n1 a0 px1 imp11 sx1 | lā ta-ḍul-u | لا تعدل | it does not equate |
| HR 215. | n1 a0 px1 imp00 sx0 | lā tu-sāwī | لا تساوي | it is not equivalent to |
| HR 221. | n1 a0 px1 imp pss11 sx1 | lā yu-nāl-u | لا ينال | it is not obtained |
| HR 222. | n1 a0 px1 imp11 sx2 | lā ḥa-sīd-∅ | لا أسعد | I do not feel happy |
| SR 234. | n1 a0 px1 imp pss11 sx2 | lā-ḥū-gad-∅ | لا يوجد | it does not exist |
| SR 288. | n1 a1 px1 imp11 sx1 | lā ya-statīr-∅ | لا يستتر | he does not hide |
| MR 390. | n1 a0 px1 imp11 sx2 | lā ta-ṣīḥ-∅ | لا صحيح | it is not correct |
| MR 467. | n1 a0 px1 imp11 sx1 | lā ḥū-bāli | لا أبالي | I do not care |
| MR 469. | n1 a0 px1 imp00 sx1 | lā tu-ṣrīk-∅ | لا تشترك | do not become polytheistic |
MR 489. n1 a1 px1 imp00 sx1 lā ya-ɣfir-θ  لا يغفر he does not forgive
HR 158. n1 a0 px1 imp11 sx2 lam ya-tanāwal-θ-ha  لم يتناولها He did not address it
DR 83. n1 a0 px1 imp01 sx0 lam ya-qul  لم يقل he did not say
DR 88. n1 a0 px0 imp01 sx0 lam ?a-kum  لم أكن I was not
DR 103. n1 a0 px0 imp11 sx0 lam ta-tazawwag  لم تتزوج she did not get married
DR 111. n1 a0 px1 imp00 sx1 lam ya-kun  لم يكن he was not
HR 157. n1 a0 px1 imp00 sx1 lam ta-kun  لم تكن it was not
MR 533. n1 a0 px1 imp pss11 sx0 lam yu-qbal-θ  لم يقبل he did not accept
SR 287. n1 a1 px1 imp pss11 sx1 mā yū-ʔaðab-ān  ما يعذبان they do not torture

Political

DP 158. n1 a0 px1 imp11 sx1 lā ʔa-drī  لا أدرى I’m not aware of
DP 162. n1 a0 px1 imp11 sx2 lā ya-liq-θ  لا يلبق it is not appropriate
MP 58. n1 a1 px1 imp11 sx2 lā ta-gid-θ  لا تجد you do not find
HP 10. n1 a0 px1 imp11 sx2 lā ya-ʔlam-θ-ha  لا يعلمها he does not know it
MP 97. n1 a1 px1 imp00 sx2 lā ʔu-dāfiẓ-θ  لا أدافع I do not defend
MP 101. n1 a0 px1 imp01 sx2 lā ta-tˤul-θ  لا تطلول it does not take long
MP 117. n1 a1 px1 imp11 sx0 lā ʔarā-ha  لا أراها I do not see that
MP 125. n1 a0 px1 imp00 sx0 lā urdˤi-ha  لا يرضيها It does not please her
I do not agree
we do not agree
it does not interfere
it is accused
that it is not enforced
that it is not
they did not heed
it was not
it did not vote
it was not
it did not finish
it did not complete
I did not deal
it did not approve
it did not slack
it does not accept
we did not hear of
it did not forbid
<table>
<thead>
<tr>
<th>b- Participle</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Religious</strong></td>
</tr>
<tr>
<td>DR 1. n1 pcpl1 sx2 ēr wārid غير وارد</td>
</tr>
<tr>
<td><strong>Political</strong></td>
</tr>
<tr>
<td>DP 172. n1 pcpl1 sx2 ēr munsīfa غير منصفة</td>
</tr>
<tr>
<td>DP 178. n1 pcpl1 sx2 ēr mund'abīt'a غير منضبطة</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>c-Modal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Religious</strong></td>
</tr>
<tr>
<td>HR 70. n1 md11 sx2 lā tagib لا تجب</td>
</tr>
<tr>
<td><strong>Political</strong></td>
</tr>
<tr>
<td>DP 82. n1 md11 sx0 lā ya-nbāyī لا ينبغي</td>
</tr>
<tr>
<td>DP 89. n1 md11 sx1 lā ya-stat′ī'īn-ūn لا يستطيعون</td>
</tr>
</tbody>
</table>