



2015-03-01

Conditional Sentences in Egyptian Colloquial and Modern Standard Arabic: A Corpus Study

Randell S. Bentley

Brigham Young University - Provo

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>



Part of the [Linguistics Commons](#)

BYU ScholarsArchive Citation

Bentley, Randell S., "Conditional Sentences in Egyptian Colloquial and Modern Standard Arabic: A Corpus Study" (2015). *All Theses and Dissertations*. 4440.

<https://scholarsarchive.byu.edu/etd/4440>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu.

Conditional Sentences in Egyptian Colloquial and Modern Standard Arabic:
A Corpus Study

Randell S. Bentley

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of
Master of Arts

Dan P. Dewey, Chair
Deryle Lonsdale
Mark Davies

Department of Linguistics and English Language
Brigham Young University
March 2015

Copyright © 2015 Randell S. Bentley

All Rights Reserved

ABSTRACT

Conditional Sentences in Egyptian Colloquial and Modern Standard Arabic: A Corpus Study

Randell S. Bentley

Department of Linguistics and English Language, BYU

Master of Arts

This thesis examines the difference between conditional phrases in Egyptian Colloquial (EC) and Modern Standard Arabic (MSA). It focuses on two different conditional particles: *'ida* and *law*. Verb tenses featured after the conditional particle determine the difference between EC and MSA usage. Grammars for EC and MSA provide a prescriptive approach for a comparison with empirical data from Arabic corpora. The study uses data from the ArabiCorpus along with a corpus of Egyptian Colloquial that were compiled specifically for this study. The results of this study demonstrate that each particle (*'ida* and *law*) and register (EC and MSA) favors a certain tense. Also, the data contrast with rules prescribed by grammars for MSA. Present tense verbs appear in the proposed condition for particle *law* a total of 22 out of 400 tokens (5.5%). Verb tense also plays an important role in determining the connecting particle for MSA sentences. The results demonstrate that the selection of connecting particles for *law* does not occur by chance but is instead systematic in nature. An apodosis containing a past tense verb strongly favors the connector *la*, while one with a non-past tense verb strongly favors the connector *fa*.

Keywords: [Egyptian Colloquial, conditionals, Modern Standard Arabic, corpus]

ACKNOWLEDGEMENTS

There are many individuals without whom I could never have completed this study. I wish to thank the members of my committee: Dr. Dan P. Dewey for this help with the statistical analysis and guidance throughout the study, Dr. Mark Davies for assisting in the methodology, and Dr. Deryle Lonsdale for his patience and thorough reading of this study.

Finally, I wish to acknowledge my family who continued to support and encourage me throughout this time. I dedicate this thesis to them, with all my love and appreciation.

TABLE OF CONTENTS

1. Introduction.....	1
2. Review of Literature	4
2.1 Modern Standard Arabic and Colloquial Dialects.....	4
2.2 Conditional Sentences.....	5
2.3 Conditional Sentences in Arabic	6
2.4 Conditional Sentences in CA.....	7
2.5 Conditional Sentences in Modern Standard Arabic.....	11
2.6 Egyptian Arabic.....	16
2.7 Summary.....	24
3. Methodology	26
3.1 The corpora	26
3.2 The procedure.....	28
3.3 Quantity and type of tokens collected.....	28
3.4 Classifying the samples.....	29
3.5 Statistical Analysis.....	29
4. Results and Statistical Analysis.....	31
4.1 Comparing EC and MSA.....	31
4.1.1 <i>'Ida</i> Protasis.....	32
4.1.2 <i>'Ida</i> Apodosis.....	32
4.1.3 <i>Law</i> Protasis	35
4.1.4 <i>Law</i> Apodosis: EC compared with MSA	36
4.2 Comparing between <i>'ida</i> and <i>law</i>	38
4.2.1 <i>'Ida</i> and <i>law</i> MSA: Protasis.....	38
4.2.2 <i>'Ida</i> and <i>law</i> MSA: Apodosis.....	38
4.2.3 <i>'Iza</i> and <i>law</i> EC: Protasis	38
4.2.3 <i>'Iza</i> and <i>law</i> EC: Apodosis	42
4.3 The connecting particles.....	42
5. Discussion of Results	45
5.1 EC and MSA.....	45
5.1.1 <i>'Ida</i> Protasis	45
5.1.2 <i>'Ida</i> Apodosis.....	46
5.1.3 <i>Law</i> Protasis	47
5.1.4 <i>Law</i> Apodosis.....	48
5.2 <i>'Ida</i> and <i>law</i>	48
5.2.1 EC protasis	48
5.2.2 EC apodosis.....	49
5.2.3 MSA Protasis.....	50
5.2.4 MSA Apodosis	50
5.3 Connectors.....	50
5.4 The Corpora.....	52
6. Conclusion	55
6.1 Answering the research questions	55
6.2 Limitations.....	56
6.3 Future studies.....	57
References.....	59

LIST OF TABLES

Table 2-1: Features of CA conditionals.....	11
Table 2-2: Features of conditionals in MSA.....	16
Table 2-4: Levels of hypotheticality with <i>law</i> protasis	23
Table 2-5: Uses of <i>'in</i>	23
Table 2-6: Uses of <i>'ida</i>	24
Table 2-7: Uses of <i>law</i>	24
Table 4-1: Tense comparison between EC and MSA, <i>'ida</i> protasis.....	33
Table 4-2: Tense comparison between EC and MSA, <i>'ida</i> apodosis	34
Table 4-3: Tense comparison between EC and MSA, <i>law</i> protasis.....	36
Table 4-4: Tense comparison between EC and MSA, <i>law</i> apodosis.....	37
Table 4-5: Comparison between <i>'ida</i> and <i>law</i> , MAS protasis.....	39
Table 4-6: Comparison between <i>'ida</i> and <i>law</i> , MSA apodosis	40
Table 4-7: Comparison between <i>'ida</i> and <i>law</i> , EC protasis.....	41
Table 4-8: Comparison between <i>'iza</i> and <i>law</i> , EC apodosis	43
Table 4-9: Comparison of connecting particles in MSA	44

LIST OF FIGURES

Figure 3-1: Search in ArabiCorpus 27

1. Introduction

Arabic contains a wide range of dialects varying greatly from country to country despite its status as the unifying pan-Arab language. Many debate the definition of what Arabic consists of precisely. Modern Standard Arabic (MSA) commonly describes the formal language in many outlets such as newspapers, government documents, literature, and news broadcasts. However, this does not include the language of everyday life used by Arabs at home or in the street. Many nations throughout the Arab World constitute diglossic speech communities. In this situation, two varieties of a single language exist side by side, MSA enjoying official status while at least one vernacular is also spoken in the sphere of daily activities (Bassiouny, 2009). Ferguson's definition of diglossia accurately depicts the status of Arabic in Arab nations:

Diglossia is a relatively stable language situation in which, in addition to the primary dialects of the language (which may include a standard or regional standards), there is a very divergent, highly codified (often grammatically more complex) superposed variety, the vehicle of a large and respected body of written literature, either of an earlier period or in another speech community, which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any sector of the community for ordinary conversation (Ferguson, 1972, p. 345).

The highly codified language for this speech community is MSA; however, the dialect will vary from country to country and often within the same country. This study is a two-point comparison that contrasts MSA and EC and common prescriptive grammars

for both MSA and EC with empirical data. EC differs from MSA in many features such as vocabulary, pronunciation, and syntax. Many social and educational institutions (schools, colleges, mosques, news media, and Arabic Language academies) frequently subject MSA to prescribed standards. Usually these institutions will have professional correctors to ensure that the language meets minimal standards (Parkinson, 1996).

This thesis discusses the construction of conditional clauses in both MSA and EC with a particular focus on the most commonly used verb tenses and their possible effect on the choice of a connecting particle. I used a corpus-based analysis relying on empirical data rather than native speaker intuition to answer the following questions:

1. What are the most frequent verb tenses in conditional clauses in both MSA and EC?
2. Are these verb tenses specific to the conditional particle and register? Can we expect the same frequency in MSA and EC for the same conditional particle?
3. Do structures that are traditionally viewed as EC structures appear in MSA writing?
4. Is there a relationship between the verb tense of the proposed result and the connecting particle?

In order to answer these questions, samples of MSA and EC conditional sentences were collected from two different corpora: one containing mostly MSA samples, the other consisting only of EC. I compared the samples of MSA and EC in order to determine which verb tense forms were more frequent and whether these tenses may have some role in deciding the connecting particle. The chapters that follow elaborate upon these topics. In Chapter 2, I review many approaches to forming conditionals in MSA and EC and

even Classical Arabic (CA). This multi-genre comparison allows for a more in-depth explanation that seeks to align Arabic speakers' tendencies along with constructions from different grammar resources. In Chapter 3, I describe the methods I used in collecting and analyzing the corpus data. In Chapter 4, I present the numerical results of the statistical analyses to determine whether the findings were significant. In Chapter 5, I discuss the results of Chapter 4 in depth, offering possible explanations. In Chapter 6, I conclude this study, discussing implications, limitations, and possible directions for future work.

2. Review of Literature

This review of literature motivates a corpus analysis to answer the research questions in the introduction. First, this chapter treats the relation between Classical Arabic (CA), Modern Standard Arabic (hereafter MSA) and Egyptian Colloquial (EC). Then, this chapter discusses common guidelines for forming conditional clauses in both MSA and EC and to a certain extent Classical Arabic (CA). This cross-linguistic analysis provides another perspective that prepares the way for a corpus-based approach.

2.1 Modern Standard Arabic and Colloquial Dialects

The written and spoken forms of Arabic often differ in multiple ways. The written language tends to be subject to prescriptive grammatical analyses that may attempt to explain and codify spoken phenomena or try to standardize the language in a way that is different from its spoken form. The status of Arabic as a pan-Arab cultural link motivates the perpetuation of a standard language that can help bridge linguistic gaps that have formed between regions. Native Arabic speakers make use of the formal written language as a means of communication in various media such as newspapers, books, and journals. Scholars commonly refer to this form as MSA. It is important to recognize, however, that MSA (and its Arabic equivalent) is not a culturally salient concept in the Arab World, and most Arabs would not know what exactly defines the term MSA. They use terms like “newspaper Arabic”, “media Arabic” and, most frequently the term *fusha*, which includes any kind of formal Arabic that follows the rules of Classical Arabic grammar. Thus, the term is somewhat broader than what MSA refers to. Furthermore, there is not one agreed-

upon definition of what Arabs consider to be ungrammatical MSA. In a study by Parkinson (1991) testing Egyptian native Arabic speakers' ability with *fusha*, he concluded that native Arabic speakers have neither a specific point in the communicative continuum nor a list of specific features that they agree upon as being MSA. This lack of consensus influences some issues to be addressed later in this text and others that will not, since they are beyond the scope of this study.

Colloquial Arabic differs from MSA in many ways ranging from pronunciation and vocabulary to syntax and morphology. The relationship between the written and spoken language allows for a comparison revealing common as well as conflicting tendencies. Attempts to document the range of differences and set rules often face challenges. While carrying out a comparative study of multiple spoken Arabic dialects, Brustad notes, "the speaker's grammar is never complete, but always evolving. Rules of a grammar can never be exhaustively documented, because they vary over time and in different sociolinguistic contexts" (2000, p. 8). In order to acquire a more complete picture of the tendencies of Arabic in the written and spoken form, I consulted multiple sources from varying times. Also, in order to study the differences of conditional sentences throughout different forms of Arabic, an explanation of conditional sentences in general follows.

2.2 Conditional Sentences

Conditional sentences consist of two structurally independent clauses that contain propositions, the validity of one depending on the validity of the other. Only if the proposition in the conditional clause (the protasis) is realized can the proposition of the

answering clause (the apodosis) be deemed valid. This relationship between the two portions exists whether the conditional sentence describes events in a possible future world or in the past or present, whether impossible or uncertain, or when an unobtainable condition is imagined to have been realized, e.g. "If he had agreed, he would have regretted it" (Holes, 1995, p.238). Sometimes, as in sentences such as "If he agreed, he may/will have regretted it", the temporal/modal meaning of the verb in the conditional clause is only interpretable with the help of the apodosis. So in this example, the condition "if he agreed" refers to a past world about which the speaker is uncertain, rather than to some future unlikely world (1995, p. 238).

2.3 Conditional Sentences in Arabic

Conditional sentences in Arabic share certain features even across the Classical, Modern Standard, and Colloquial forms. Arabic uses an introducing particle in the protasis, each traditionally conveying a different degree of probability. All are roughly equivalent to English "if": *'in*, *'ida*, *law*. After each particle a verb-initial clause normally follows. When the protasis does not feature a verb, then the verb *kān* 'to be' is used (Holes, 1995).

The first particle, *'in*, introduces a proposition of uncertain but possible realization (1). The second particle, *'ida*, presents a contrasting meaning that implies that the proposition in its clause will definitely occur and that it is only the time which remains uncertain; it can thus be translated with either 'if' or 'when' (2). Finally the particle *law* conveys a meaning that the proposed condition or event is impossible or has already been proven not to have occurred (3) (Holes, 1995, p.238).

- (1) *'in māṭ zayd'*
 if die-3SG.PST Zayd
 'If Zayd dies/should die (e.g. of his present illness).'
- (2) *'ida māṭa zayd'*
 when die-3SG.PST Zayd
 'When(ever) Zayd dies (i.e. like all men, he eventually will).'
- (3) *law māṭa zayd'*
 if die-3SG.PST zayd
 'If Zayd had died.'

2.4 Conditional Sentences in CA

Holes (1995) notes that the choice of conditional particles plays a role in Arabic that is largely absent in English. Along with the choice of particles that introduce the condition, the sequence of verb forms indicates the speaker's understanding of the degree of probability of the event in question. Although these features (the particles and the sequence of tenses) work together to form conditional sentences in Arabic, Holes notes that often Arabic conditional sentences are ambiguous out of context and require more than just the complete conditional sentence to determine whether the conditional is real, possible, or unreal.

The most important features that Holes (1995) describes are the levels of probability each particle conveys. In CA, *'in* is for events that are highly likely to occur. *Law* is for events that are contrary to fact or impossible. Finally, *'ida* is a temporal marker rather than a true conditional like *'in* and *law* and is often translated as 'whenever' as opposed to 'if'.

The second point of interest here is the verb tense sequences. Regarding the verb in the previous examples (1-3), CA rules stipulate that the form of the verb remains the

same—simple perfect—regardless of the degree of probability portrayed. The simple perfect serves the function of presenting Zayd’s death, in the past or future, as a reference point from which the following events could take place. In other words, it has no past tense meaning. Thus, different shades of probability in CA conditionals are signaled by the choice of the particle, and not, as in English, by the form of the verb.

One question addressed in this study pertains to verb tense sequence: what types of tenses appear in the protasis and apodosis. Do scholars assign specific verb tenses to the protasis and the apodosis based on the conditional particles? CA rules give several options for verb forms in the protasis and the apodosis depending on these conditional particles. For example, Socin (1945) says the simple perfect in the protasis and the apodosis can convey both the present and the future sense (p. 133).

- (4) *‘in* *fa’alta* *dālika* *halakta*
 if do-2SG.PST that perish-2SG.FUT
 ‘If you do that, you will perish.’

Further possibilities for *‘in* include the jussive in both the protasis and apodosis (Socin, 1945). The jussive mood is related to the imperfect and marks non-occurring events (Bedawi, 2004). Also a jussive in the protasis may also be followed by an imperfect in the apodosis (Socin, 1945).

When addressing the same topic of verb tense combinations Thatcher (1958) provides a group of possible combinations for *‘in* and *‘ida*. Although the combinations all feature *‘in*, he notes that these combinations hold for all particles with few exceptions.

The exceptions that Thatcher (1958) mentions are that after *'ida* in a conditional sense, the jussive is rarely used. With *law* the perfect appears in both parts and rarely the imperfect (p. 232).

- (5) *law šā'* *rabbuka laja'ala l-nās*
'umma-taan waḥida-taan
if desire-3SG.PST lord-your then-make-3SG.PST the-
people nation one
'If thy Lord had chosen, He would surely have made (all) mankind one people.'

There is also a less frequent usage of *kān* repeated before the apodosis conveying a meaning similar to that of the pluperfect (p. 323):

- (6) *law kunt-u 'alimt-u dālika la-ḡarab-t-uka*
if be-1SG.PST know-1SG.PST this then-beat-1SG.PST-
you
'If I had known this, I would have beaten you.'

Finally, different connecting particles are possible after the protasis and before the apodosis. Holes (1995) notes that whenever there is not a parallel structure after *'in* or *'ida*, e.g. simple perfect in both the protasis and the apodosis, then the particle *fa* is inserted before the verb in the apodosis (p. 239).

- (7) *ida raġib-ta fī ḡajz-i tadkarat-in, fa-'alayka*
'*an ta-dfa'-a musabbaq-an*
if want-2SG.PST to reserve ticket, then-upon.you
to pay advance
'If you want to reserve a ticket, (then) you must pay in advance.'

Hence the verb after *fa* is always imperfect, whereas without a *fa* it would usually be perfect. In the example (4), the protasis contains a perfect, while a *fa* appears in the apodosis with a non-verbal sentence. Because non-verbal sentences—equational

sentences—are imperfect the particle *fa* precedes apodosis. Thatcher (1945) lists three circumstances that require the usage a *fa* for the particles: A nominal sentence, a verbal sentence expressing a wish, command or prohibition (with the verb in the imperative or jussive), or a verbal sentence that begins with one of these particles: *sawf*, *sa*, *qad*, *ma*, or *lan*.

For conditional sentences introduced with *law*, the connecting particle is *la*. However, CA rules do not provide any specific guidelines for the use of this particle (Thatcher, 1945, p.323).

- (8) *law arāda ‘an yaxūn-a la-kāna ‘axada*
l-kīs-a bi-rumma-tihi
 if want-3SG.PST to betray-3SG.PRS then-be-3SG.PST take-
 3SG.PRS the-bag in-whole-his
 ‘If he had wished to be treacherous, he would have taken the purse in its entirety.’

As previously mentioned, the two conditional particles, *‘ida* and *‘in*, make use of *fa* before the apodosis, depending on the nature of the clause and its verb or lack thereof; this does not appear to be the case when *la* is used with the particle *law* since it is purely optional.

Table 2-1 summarizes the mains points of comparison for CA.

Table 2-1: Features of CA conditionals

Conditional Particle	Probability	Verb Tense Sequence Protasis, Apodosis	Connecting particle
<i>'in</i>	high	Jussive, Jussive Past, Past Past/Jussive, Imperative	Use of <i>fa</i> before: 1) Nominal sentence 2) The particles: <i>sawf, sa, lan, ma, and qad</i>
<i>'ida</i>	None, only use for temporal	Jussive, Jussive Past, Past Past/Jussive, Imperative	Use of <i>fa</i> before: 1) Nominal sentence 2) The particles: <i>sawf, sa, lan, ma, and qad</i>
<i>law</i>	Contrary to fact, impossible	(kān +) Past, (kān +) Past	Optional use of <i>la</i>

2.5 Conditional Sentences in Modern Standard Arabic

We first compare the level of probability for each particle for CA and MSA. The clear distinction in the use of the three particles, *'in*, *'ida*, and *law*, when referring to probability, has not completely carried over into MSA. For example, the distinction in meaning between *'in* and *'ida* in CA has been mostly lost in MSA. *Law*, on the other hand, retains its 'contrary to fact' meaning in MSA (Bedawi, 2004). The particle *'in* exists in MSA, although mainly in frozen and religious expressions (Bedawi, 2004); in most contexts, *'ida* replaces it in modern usage. Bedawi's (2004) circumstances for the use of the particle *law* match that of other the grammarians reviewed thus far: it heads clauses whose validity is either impossible, highly unlikely, hypothetical or contrary to fact. These degrees of probability did not vary between the sources used for this study:

they all refer to *law* as the contrary-to-fact particle and associate ‘*ida* with events of high probability.

MSA sources differ in their treatment of verb tense usage; some, but not all, prefer to discuss "preferences" over rules, thus taking a more descriptive approach. Holes (1995) presents preferences (rather than rules) for verb sequences in Arabic conditional sentences.

MSA favors a simple perfect form for the protasis of ‘*ida* while allowing all forms for the apodosis. In contrast, CA grammars often involve either a jussive or simple perfect for the apodosis and a matching verb form for the apodosis. For *law*, the rules do not indicate strict limitations to possible forms but generally prescribe a simple perfect or *kān* followed by a simple verb for the protasis and a simple perfect for the apodosis (Holes, 1995).

A more in-depth look at conditional sentences by Bedawi (2004) offers more information on ‘*ida* and verb form combinations. Bedawi (2004) generally associates ‘*ida* with perfect in the protasis and the apodosis, or jussive in both. He writes that other combinations are possible in theory and does not list each possible combination. He later explains that the regular structures of CA are not very common in MSA and finds very few examples in the course of his study. Most of these symmetrical structures of CA are perfect+perfect or jussive+jussive and occur in proverbs and other formulaic structures.

Bedawi (2004) states that the protasis is always perfect and the verb of the apodosis is frequently imperfect. For the particle *law*, Bedawi (2004) says that these conditional sentences tend to exhibit the simple perfect tense in the protasis and apodosis.

This typically corresponds to the English pluperfect. The apodosis may also be elided in certain contexts pertaining to wishes and hopes (2004, p. 649):

- (9) *law* *rahimt-a* *ʔabd-a-k* *ya* *mawlaya*
if pardon-2SG.PST servant-your oh lord
'If you pardon your obedient servant, my Lord!'

Some grammars considerably shorten and simplify rules and possibilities of conditional sentences in favor of a more open explanation that does not necessarily cover all possible structures. This may be an effort to simplify and focus on those forms that are most commonly present in MSA for the target audience of students in their first or second year of studying Arabic. Ryding (2010), in her MSA grammar book, describes *'ida* as containing a verb in the past tense for the protasis and does not specify any preferred tense for the apodosis but rather lists examples containing imperative and imperfect forms.

Ryding's (2010) contrary-to-fact conditionals are introduced by the particle *law* and followed by a past tense verb. The verb tense of the apodosis is not addressed except through examples, many of which feature a past tense verb.

Other sources such as textbooks also display a simplified approach targeting students rather than grammarians. In the textbook *Al-Kitaab* (Al-Batal, Al-Tonsi, & Brustad: 2004), the level of student proficiency and the scope of the text do not favor lengthy, in-depth explanations like those in other grammars. *Al-Kitaab* targets college-level MSA and is widely used since few other textbooks exist for study at this level. This text's descriptions are concise and clear while not covering different verb form combinations seen in other grammars. *Al-Kitaab* defines the particle *'ida* as the most

common conditional particle. It must introduce a past tense verb despite the intended meaning of the sentence, whether past, present or future. The apodosis does not have a specifically suggested tense and the text presents examples featuring imperfect forms (p. 274).

(10) *wāliḍa-tu* *mahā* *ta-ğḍab-u* *'ida* *ta'axar-at*
fī *l-'awḍa-ti* *'ilā* *l-bayt-i*
 mother Maha become.angry-3SG.PRS if be.late-3SG.PRS
 in the-returning to the-home
 'Maha's mother gets angry if she is late returning home.'

(11) *law* *kunt-u* *l-ra'īs-a* *la-sā'adt-u* *kull-a* *l-nās*
 if be-1SG.PST president then-help-1SG.PST all the-people
 'If I were president I would help everyone.'

Neither Ryding's grammar nor *Al-Kitaab*, provides any details about the usage of connecting particles. Instead they mention the connecting particles as the usual connectors that are elided when the order of the protasis and apodosis are reversed and the apodosis precedes the protasis (Ryding, 2005). *Al-Kitaab* describes the connecting particle *fa* as usually introducing the apodosis for conditionals headed by *'ida*, whereas *la* corresponds with *law*. Their frequency, behavior, or usage are not mentioned (Al-Batal, Al-Tonsi, & Brustad, 2004).

Bedawi (2004) does not list specific rules that dictate the use of *fa* to connect the protasis and apodosis of a conditional sentence featuring *'ida*. Instead, he opts for a semantic explanation for its usage. When the apodosis is not the logical result of the conditional class, the particle *fa* is used. He also explains that it emphasizes the temporal and/or sequential aspect of the apodosis. He also notes that a conditional sentence introduced by *'ida* may feature the connecting particle *la* instead of *fa* which is normally

used (2004). Bedawi (2004) mentions an additional meaning to *law*, which loses its counterfactual quality where the connecting particle helps mark this change. This occurs in two types of sentences: the first have *fa* in the apodosis, expected for normal CA non-verbal apodosis (p. 647):

- (12) *'id law kān-a li-l-šaxš-I l-wāhid-I 'aktar-u min*
ra'sisin mubasirin fa-la natija-ta li-dalika ġayr-u
 for if be.3SG.PST to-person the-one more then
 boss direct so-no result for-that only
 'For if one person has more than one direct boss [then] the only result of that is.'

In the second type, *fa* is optional in the apodosis, with a future negation realized as *lan* or any other unmarked verb. (p. 647):

- (13) *wa-law sa'alt-a 'ahad-a-hum min 'ayna 'atayt-a bi-*
hāda l-haqqi... fa-lan tajid-a radd-an
 and-if ask.2SG.PST one-them from where get-2SG.PST with-
 this the-truth then-NEG find-2SG.PRS answer
 'and if you ask one of them where did he get this truth from..., [then] you will not find an answer.'

- (14) *law daxal-nā sibāq-a tasalluḥ-in nawawiyy-in fi l-*
miṭṭaqa-ti lan ya-ntahiy-a
 if enter-1PL.PSTTrace arm nuclear in the-
 region NEG end-3SG.PRS
 'If we enter a nuclear arms race in the region it will never [lit. will not] end.'

Table 2-2 lists a comparison of particle attributes for MSA.

Table 2-2: Features of conditionals in MSA

Conditional Particles	Probability	Verb Tense Sequence	Connecting Particle
<i>'in</i>	Used mostly in religious and set phrases	Past, Imperfect	<i>fa</i> listed as either optional or semantically motivated
<i>'ida</i>	High probability	Past, Imperfect	<i>fa</i> listed as either optional or semantically motivated
<i>law</i>	Contrary to fact impossible	Past, Past	Use of <i>la</i> is optional and use of <i>fa</i> can portray a highly probably meaning rather than contrary to fact

2.6 Egyptian Arabic

Although an ample amount of studies and grammars address EC when compared to other dialects, this does not match the work on MSA. In his EC grammar, Bedawi (2009) offers a slight distinction between the types of conditional sentences: the simple, and the hypothetical. His simple conditional is a sentence that lays down a condition and then states a result, the condition being a requisite for the completion of the result. The most important factor in these sentences is that both the condition and the result can occur (p. 51).

- (15) *'iza rāh ḥa-rūḥ ma²-āh*
 if go-3SG.PST go-1SG.FUT with-him
 'If he goes I will go with him.'

Hypothetical sentences also involve the process of laying down a condition; however, in contrast to the simple sentence the condition does not take place, hence the result is deemed not possible (Bedawi, 2009, p. 54).

- (16) *law kān ḥa-yrūḥ kunt a-rūḥ ma²-aah*

if be-3SG.PST g-3SG.FUT be-1SG.PST go-1SG.PRS with-him
 ‘If he were going to go, I would go with him.’

Bedawi’s rules include a caution that these are the most common usages and that widespread variations to these rules exist. Simple sentences are headed by the particle *‘ida* (commonly pronounced as *‘iza* in EC) or *‘in*. The constituent perfect, or *kān*, is obligatorily added before the verb. If the sentence is equational, then the constituent perfect is added after the participle *‘iza* or *‘in*. The possibility of a future verb in the apodosis preceded by a perfect form of *kān* yields a contrast between the two forms (2004, p. 51):

(17) *‘iza rāḥ*
 if go-3SG.PST
 ‘If he goes.’

(18) *‘iza kān ḥa-yrūḥ*
 if be-3SG.PST go-3SG.FUT
 ‘If he is going to go.’

(19) *‘iza kān-it bi-ti-dris ingiliizi fi-g-gam’a di*
ha-‘a-bilha f-yūm mil‘ayyām
 if be-3SG.PST study-3SG.PRS.HAB English in-the-university this
 meet.1SG.FUT-her in-day of-day-PL
 ‘If she is studying English at this university, I will meet her someday.’

(20) *‘iza kunt-i b-ti-‘ri ktāb kull yūm, ha-t-xallaṣi*
l-kutub di f‘usbū’
 if be.2SG.PST read-2SG.PRS.HAB book every day complete-
 2SG.FUT the-book-PL these in-week
 ‘If you are reading a book a day, you will complete (reading) these books in a week.’

Hypothetical conjunctions feature the particle *law* to introduce the condition. The use of *kān*, or the constituent perfect, is optional in this structure. If the verb in the protasis is also perfect then it is left unaltered. However, if the verb bears an aspect prefix,

i.e., *bi*, then the verb appears in its corresponding subjunctive form. The presence of the subjunctive form here means 'not possible, not attainable' (2004, p. 54).

- (21) *law* (*kānū*) *ha-y-rūḥu* *maṣr* *kānū* *'ālūlnā*
 if (be.3PL.PST) go-3PL.FUT Egypt be-3PL.PST tell-3PL.PST-to-us
 'If they were going to go to Egypt they would have told us.'

Some of the variations that are covered in this grammar include the interchangeable use of all three particles, *'iza*, *in*, and *law*, even in a simple or hypothetical conditional sentence (2004, p. 56):

- (22) *law/'iza/in* *rāḥ* *ḥa-rūḥ* *ma'-aa*
 if go-3SG.PST go-1SG.FUT with-him
 'If he goes I will go with him.'

Bedawi (2004) also notes the use of the invariable form *yib'a* as a means to introduce the result of a simple conditional sentence. He characterizes the distribution of *yib'a* as non-arbitrary but also difficult to describe. Sometimes its presence evokes the meaning 'then it is true, then it is a fact that, then it must be a fact that'. (p. 57)

Bedawi's (2004) explanations are similar to those which we have seen for MSA: a perfect must be used after *'iza* and *law* when introducing a conditional sentence. It differs from some MSA sources somewhat by not prescribing a tense for the apodosis as we saw with the textbook *Al-Kitaab*, which prescribed a perfect for both the protasis and apodosis for conditionals headed by *law*. It also differs by not making any mention of the connecting particles *fa* or *la*, which is expected since they are mainly confined to MSA. Instead, he mentions the optional use of *yib'a* for connecting the protasis and apodosis, though the rules for its usage are not clear. Finally he notes that great variations exist in the use of the conditional particles; some speakers use them interchangeably.

In other intermediate-level textbooks for Egyptian Colloquial Arabic, the guidelines for forming a conditional sentence are more flexible than those listed in previously discussed grammars. The textbook *Kullu Tamam* targets first-year university Arabic courses in the Netherlands and prepares students to further study MSA if they so desire. In the textbook *Kullu Tamam*, conditional sentences that refer to a possibility that can be realized are termed realis. The second type of conditional sentences is called contrafactive or irrealis. All the introducing particles, ‘*iza*, *law* and occasionally ‘*in*, appear in either type (Hein-Nasr, 2005).

The suggested manner of forming realis conditional sentences is rather broad. ‘*iza* can be followed either by the perfect or by *kān* with an imperfect, a participle, a nominal phrase, or a prepositional phrase. The particle *law* can also be followed by a simple imperfect. These possibilities are present in the following examples that are all translated with the same sentence (Hein-Nasr, 2005, p. 221):

- (23) ‘*iza* *kunte* *thibbe* *timši* *itfaḍḍal*
if be-2SG.PST want-2SG.PRS go-2SG.PRS go-2SG.IMP
‘*iza* *ḥabb-et* *tī-mši* *itfaḍḍal*
if want-2SG.PST go-2SG.PRS go-2SG.IMP
law *ḥabb-et* *tī-mši* *itfaḍḍal*
if want-2SG.PST go-2SG.PRS go-2SG.IMP
law *tī-ḥibbe* *tī-mši* *itfaḍḍal*
if want-2SG.PRS go-2SG.PRS go-2SG.IMP
‘If you want to go, then go.’

An irrealis conditional sentence is only introduced by ‘*iza* or *law*. In these sentences the particles are mostly followed by *kān*, and the main clause itself is always introduced by *kān*. The conditional in an irrealis clause can refer to the present or the

future. When it refers to the past, the main clause always has *kān*+perfect (Hein-Nasr, 2005).

The authors mention that in practice these rules are not always strictly applied, and determining whether the condition is realizable or not usually requires the full context.

When dealing with the variety of conditional clauses in EC, Eisele (2000) notes a great deal of variability among the different forms and also in their description in grammar books. In order to treat this topic more thoroughly, Eisele refers to a cross-linguistic analysis carried out by Comrie. Comrie (1986) does not categorize conditionals as either factual or counterfactual; instead, he proposes a continuum that allows for a wider range of what he refers to as "degrees of hypotheticality". In this scale, a proposition marked having "greater hypotheticality" indicates "lower probability" while one marked as having "lower hypotheticality" indicates "higher probability" (Comrie 1986, p. 88). Another aspect in conveying a degree of hypotheticality is the use of backshifting, or the "use of morphologically past tense with present (or future) time reference and a pluperfect with past time reference" (p. 94). This is similar to what Holes mentioned about perfect verbs in the protasis not carrying any past tense meaning. Comrie (1986) also suggests that conditionals with past time reference convey a higher degree of counter-factuality than conditionals with a future time reference, and conditionals with present time reference occupy a middle ground.

Upon these findings Eisele (2000) highlights some key features of conditional sentences in Egyptian colloquial: EC uses back shifting in conditional sentences and

distinguishes between low and probability in conditional sentences. EC favors future tenses for low probability and past tense for high probability.

He also makes note of distinguishing features of CA when compared to English (2000): EC distinguishes between high and low probability with other devices besides *'iza* for high probability and *law* for low probability. The particle *'iza* always requires backshifting in the protasis, whereas *law* only requires past tense to express contrary-to-fact conditionals. Eisele concludes that the *law* appears to convey high probability when it does not feature a past tense. He compares this high level of probability of *law* to *'iza+past*. This is somewhat different from the scale of other grammar which reserve *'iza+past* for high probability and *law+past* for low probability.

With this information, Eisele's (2000) explanation of possibilities in EC conditionals seems to account for some of the variation. He states that *'iza* is used for low hypothetical conditionals, and that backshifting is generally obligatory in the protasis whereas it is not used at all in the apodosis. He also adds that the use of the simple perfect in the apodosis is very rare and only appeared once in the examples that he collected.

He also adds a few examples where *'iza* may be used without the use of a past tense. These usually occur in instances with a certain number of active participles or modal verbs (2000).

Eisele (2000) notes that the usage of past tense with *law*, is very similar to English conditionals, as its use with a backshifted verb signals a counterfactual conditional. *Law* can also be used without backshifting in the protasis to express a simple, low hypothetical conditional, like *'iza*. *Law* may also be used with simple past tense verbs in the protasis

for either low or high hypotheticality. In these instances the tense of the apodosis helps determine whether it is to be read as high or low hypotheticality. A past tense in the apodosis (usually marked by *kān* + verb), results in the whole conditional as being read as high hypothetical. A non-past in the apodosis results in a low hypothetical. Apodoses with a simple perfect (whether of *law*- or *'iza*-conditionals) are not frequent in Eisele's material.

In his data, high hypothetical conditionals with *law* occur with the following examples (2000, p. 75):

- (24) *law ḥiseṇ 'ayyan kān abūk*
ittāṣal bi-d-duktar
 if Hussein become.sick-3SG.PST be-3SG.PST father-your
 call-3SG.PST the-doctor
 'If Hussein (is) sick, your father would have contacted the doctor.'
- (25) *law rāḥ il-madrasa, kunti ruḥti ma'-ah*
 if go-3SG.PST the-school be-1SG.PST go-1SG.PST with-him
 'If he went to school, I would have gone with him.'
- (26) *law kunti 'ult-il-i kunti sa'idt-ak*
 if be-2SG.PST tell-2SG.PST-to-me be-1SG.PST help-1SG.PST
 'If you had told me, I would have helped you.'

Eisele (2000) summarizes the different possibilities with *kān* (was) and their corresponding level of hypotheticality. If *kān* appears in the protasis, then the conditional conveys high hypotheticality, and the apodosis must also feature *kān*. If *kān* does not appear in the protasis, then the conditional may convey high or low hypotheticality: If *kān* appears in the apodosis, then the conditional always indicates high hypotheticality, whether the protasis contains past or non-past verb. If *kān* does not appear in the apodosis, then the conditional indicates low hypotheticality. Table 2-3 summarizes the possibilities for expressing different levels of hypotheticality: L = low, H = high, * = non-occurring, ? = rare.

Table 2-3: Levels of hypotheticality with *law* protasis

PROTASIS	APODOSIS	HYPOTHETICALITY
<i>law</i> + non-past	non-past	L
<i>law</i> + simple past	non-past	L
<i>law</i> + <i>kān</i>	non-past	**
<i>law</i> + non-past	simple past	?L
<i>law</i> + simple past	simple past	?L
<i>law</i> + <i>kān</i>	simple past	**
<i>law</i> + non-past	<i>kān</i> + verb	H
<i>law</i> + simple past	<i>kān</i> + verb	H
<i>law</i> + <i>kān</i>	<i>kān</i> + verb	H

This perspective of high and low probability adds a new range of combinations and also insight in to some possible explanations for trends that may seem to go against prescribed tense (i.e. the use of imperfect in the protasis for *law* as opposed to the traditionally prescribed perfect). A three-way comparison of the attributes of each particle in CA, MSA, and EC is below: prescribed tense i.e. the use of imperfect in the protasis for *law* as opposed to the traditionally prescribed perfect. A three-way comparison of the attributes of each particle in CA, MSA, and EC is below:

Table 2-4: Uses of ‘*in*’

‘ <i>in</i> ’	Probability	Verb Tense Sequence Protasis, Apodosis	Connecting Particle
CA	High	Jussive, Jussive Perfect, Perfect Perfect/Jussive, Imperative	Use of <i>fa</i> before: 1)Nominal sentence 2)The particles: <i>sawf, sa, lan, ma, and qad</i>
MSA	High	Perfect, Imperfect	<i>fa</i> listed as either optional or semantically motivated
EC	High	Perfect, Imperfect	<i>yib ’a</i> optional

Table 2-5: Uses of 'ida

<i>'ida</i>	Probability	Verb Tense Sequence Protasis, Apodosis	Connecting Particle
CA	Temporal	Jussive, Jussive Perfect, Perfect Perfect/Jussive, Imperative	Use of <i>fa</i> before: 1)Nominal sentence 2)The particles: <i>sawf, sa, lan, ma, and qad</i>
MSA	High	Perfect, Imperfect	<i>fa</i> listed as either optional or semantically motivated
EC	High	Perfect, Imperfect	<i>yib 'a</i> optional

Table 2-6: Uses of law

<i>law</i>	Probability	Verb Tense Sequence Protasis, Apodosis	Connecting Particle
CA	Contrary to fact impossible	(<i>kān +</i>)Perfect, (<i>kān +</i>) Perfect	Optional use of <i>la</i>
MSA	Contrary to fact impossible	Perfect, Perfect Perfect, Imperfect	Use of <i>la</i> is optional and use of <i>fa</i> is usually followed by a high probability result
EC	High or low	Perfect/Imperfect, Perfect/Imperfect	<i>yib 'a</i> optional

2.7 Summary

Our comparison of different grammars has shown the focus of some aspects of conditional sentences sometimes involves prescribing certain verb sequences and conditions for adding or eliding connecting particles. However, other sources prefer to take a more descriptive approach that does not adhere to a single form or even two or three, but rather leaves a very open approach that attempts to give meaning to almost every possible combination. In comparing CA sources to MSA sources, we found that

CA sources favor an explanation that covers possibilities for both the protasis and apodosis. On the other hand, MSA sources tend to focus solely on the protasis. EC sources take this trend even further and list very open possibilities. Although each source prescribes certain tenses to some extent and makes note of less commonly used tenses, none presents an expected frequency for each possibility. This study helps elucidate which forms and sequences may be more common than others and to what extent we can expect a conditional particle to take certain tenses in comparison to other particles. A cross-comparison between EC and MSA by means of a corpus-based study will also allow insights on how similar or different MSA and EC trends impact conditional sentences and whether or not some of the tendencies are due to varying degrees of hypotheticality. This comparison between EC and MSA also allows for the extent of EC's presence in some MSA writings. It also demonstrates whether structures previously deemed as EC appear in MSA writings.

3. Methodology

This chapter explains my methods for gathering data on conditional sentences in both MSA and EC. Sample data were gathered from one corpus for MSA and another for EC.

3.1 The corpora

MSA data came from the ArabiCorpus compiled by Dilworth Parkinson (<http://arabicorpus.byu.edu/>). This corpus features a wide variety of written Arabic. The sections of this corpus used for this study include: Egyptian newspapers (Al-Ahram and Ash-Shuruq) and Modern Literature. Egyptian newspapers were used as a base to compare the practices of MSA from Egypt to the Egyptian's colloquial dialect. Similarly, only Egyptian samples were taken from the Modern Literature section to facilitate comparison with EC. The section contained 17,959,738 words. The total number of tokens for *'ida* and *law* was 35,699 and 15,468 respectively. For the literature section, the total word count was 1,026,171 words. The total number of tokens for *'ida* and *law* was 3,705 and 1,902 respectively. Figure 3-1 is a sample of a typical search in the ArabiCorpus.

Figure 3-1: Search in ArabiCorpus

rank	Arabic text	English translation	Category
1	أثر العولمة على المشاركة الاقتصادية للمرأة وإدارة الموارد المائية، المرأة وانتخابات	النوع الإلكتروني على شبكة الانترنت تجد دراسات بعنوان المرأة والعولمة، أثر	SPECIALFILE
2	أمم هضاروا أمريكيين أكثر من المحافظين الجدد، ولدينا مسوقون أمام الرهائية	المسوق يتماهى مع من سحقه ويزين عليه، ولدينا مسوقون أمام	OPINION
3	أن لها أثرها الإيجابي، ورغم ذلك مات الملايين في العالم من	وأهداف جمال مبارك أن	ECONOMY
4	وراءه وأنه لا غر منها.	بنفسه ووضع دستورها، دولة افترضية في خيلته، لكنه يكتشف أن	ARTS
5	أن والتقدم العلمي والتكنولوجي وسهولة الاتصال بين الأفراد والجماعات وسهولة الانتقال	مشيرا إلى أن	ACCIDENTS/ISSUES
6	أهداف الرأسمالية) كما تتلشى مظاهر الفساد والش	المساحية الليبرالية الجديدة تتميز بالفردية الضيقة والشراء الاستهلاكي، (أحد أهداف	OPINION
7	إحصار والمصالح الدولية. لذلك كنت أود أن ألق على وجهة نظر	مواجهة المتغيرات السياسية والاجتماعية والاقتصادية الشرسة القادمة على جناسي إحصار	OPINION
8	إلى السياسية والاقتصادية وأدوات «القوة الناعمة» التي خدمت مصالح والشطن.	أمريكا وحاول بوش تدميرها، إنها بالتأكيد دعوة لإعادة الاعتبار إلى	WORLDNEWS
9	إلى التي نخلنا إليها بقوة، وأصبحت جزءا من تغيراتها.	والثقافة الرياضية تسيطر أكثر... وجزء من ذلك أنه لا يعود إلى	SPECIAL
10	ين أدت إلى وجود هوة كبيرة بين الشمال والجنوب، موضحا أنه	من المؤتمر، أمن الأول، تحت عنوان «مصلح الأدب الجديد» إن	HOTTOPICS
11	إنها التي تمت ترجمتها في كرة القدم عبر نظم «الاحتراف» وأنها	إنها	OPINION
12	إيجابيات وتقليل سلبياتها قدر الإمكان.	2010، فيما تدعو هذه الحقيقة عالم الجنوب للاستفادة من إيجابيات	SPORTS
13	البيات الجيزة والقاهرة.	السعراء للحرمان من الترحمة الأعلى في أسماها الإنساني باستخدام البيات	SPORTS
14	استخدام للتعبير عن التماسك بين الثقافات، وأن التكنولوجيا كان لها دور	هوة كبيرة بين الشمال والجنوب، موضحا أنه لا يتم استخدام	HOTTOPICS
15	الاقتصادي الطاغية. عندما استمعت للدكتور سعدون، لم أعرف هل أصف له	بخطن قد يراه البعض غير مبرز في إطار التقدم الاقتصادي	OPINION
16	الاتفاق على لغة الإعلام، والمعكس لذلك على الخطاب العلوي الجماهيري، والدور	وأثر الاتفاق	LAST
17	الاتفاق، والعولمة،	علاج مشكلة اجتماعية تتركب الشعب الصيني خاصة بعد مرحلة الاتفاق	CINEMA
18	والظروف السياسية المتناقضة والطبيعة الاجتماعية وعدم التقارب والتوازن بين النظام	وغياب الوازع الديني والعروج عن المبادئ والتقاليد في ظل التعريب	OPINION
19	الجنس وقال: «المسرحية الغنائية ليست خاطئة، ولا أسى هذا النجاح الذي	الاتفاق عليها منذ كتابة النص المسرحي للثلاثة على أرمع الجنس	ARTS
20	الحر وبخاصة في قطاع الغزل والنسيج الذي تكثر برامج الإصلاح	مصر تكرر ما مثل غيرها في دول العالم بديناميات الاقتصاد الحر	HOTTOPICS
21	الحرمة وأوبئة العصر.	معموصة، ضحايا الاستبداد والبطالة والبطورسة والخصخصة والفساد والاستمرار في السوق الحر	OPINION
22	الخصوصية وإزالة العوائق نتيجة تقدم وسائل الاتصال ساعدت كثيرا على تغيير	التي كانت عليها في السابق، مناطق متناخلة تدعم فيها الخصوصية،	SPECIAL
23	الدين والمجتمع والثقافة، وكلهم من متلقى نهاية الألفية ولهم مجلدات ومراجع	أو هانز أو بيير في علم الاجتماع الديني، وعلاوة الدين	OPINION
24	الدين فبقنا لم نترجم كتابات مرجعية عالية القيمة لأسماء مثل: حيفا	وإلى يومنا هذا بالرغم من أن الكال يتحدث عن الدين	OPINION
25	الدين والدين والعولمة والثقافة، وسلف بروين (الديناميات والدين)، وريسا ألبرت (أصوات	والدين/العولمة والثقافة السياسية في العالم الثالث). وبيتر بيير (صاحب الدين	OPINION
26	السلام وحقوق الإنسان، وكان البطل الأول لها هو الحوار الذي وصل	على مقدرات الحياة بالقوة العسكرية والقوة السياسي والاقتصادي ويدعوى السلام	HOTTOPICS
27	الكتاب ب«الديمير الركاثر الكافية»»، وشبهها بعثت برمودة، وعاجوا النظام بسبب ما	والم الكتاب	HOTTOPICS
28	المتوحشة ويتقاطع مع الكثير من الهومو الإنسانية الحالية وإشكاليات التعامل مع	وقد عرض على خشبة المسرح القومي ويناقش فكرة الرأسمالية المتوحشة	LIKETODAY
29	المرأة أثر العولمة على المشاركة الاقتصادية للمرأة، وإدارة الموارد المائية،	على مؤلفه الإلكتروني على شبكة الانترنت تجد دراسات بعنوان المرأة	SPECIALFILE
30	التمسي والذي أكل الأخضر والباقس في طريقه ولم يترك دولة إلا	مرة أن المسوخ النجلى ظهر بالفعل متملا في الوحش المسمى	SPECIALFILE
31	المعطرات ووصول براك أوباما إلى السلطة هي كلها مؤشرات على دخولنا	لا نظير له في كل تاريخ البشرية، وإن ثورة المعلومات	WORLDNEWS

I sought to compile a corpus with a variety of EC forms by including the following eight sources: a novel (*Taxi*), a play (*Awladna fi London*), transcriptions of seven episodes of an Egyptian television series (*Hikaayat Hiya*) and of eighty hours of telephone conversations (from the CALLHOME Egyptian Arabic project), blog posts from an Egyptian author (<http://wanna-b-a-bride.blogspot.com>), articles from interviews carried out by Egyptians (<http://langmedia.fivecolleges.edu/culturetalk/egypt/index.html>), articles from an Egyptian magazine *Ihna*, and comments from Egyptian chat websites that are part of the colloquial section of Parkinson’s ArabiCorpus. The novel *Taxi* was chosen because it is written in colloquial style and contains large amounts of dialogue. The play *Awlanda fi London* was included because it contains many dialogues between Egyptians. The television series was chosen because it is a modern series freely available for viewing on the Internet and is set in Cairo. I included the magazine *Ihna* because it

covers a variety of topics and is very unique being one of the only magazines in EC. The blog posts were chosen because they are written by an Egyptian and are almost exclusively in EC. I selected the interviews because only Egyptians conducted and participated in these interviews. Finally, the chat conversations were added because they were written in EC form and are demonstrative of typical spoken conversations. The total word count for this combined corpus was 382,050 words. The total tokens for *'iza* and *law* was 824 and 1,456 respectively.

3.2 The procedure

I retrieved data for the MSA corpus using the newspaper and literature sections equally for each particle. I performed a simple search in the ArabiCorpus for each particle. Results appeared in over 200 groups of 100 sentences. I randomly selected 10 of these groups and then selected every tenth token from these groups to include in my analysis. For the EC corpus, I performed a simple word search for the particle *law* in a word document of the collected texts. Because there were significantly fewer entries, I selected every fourth entry to span across all the types of texts in the corpus. For the particle *'iza*, there were not enough tokens to allow for every fourth particle. I therefore analyzed every other entry.

3.3 Quantity and type of tokens collected

I collected 800 examples of each particle—400 from the EC corpus and 400 from ArabiCorpus (half EC and half MSA). The samples from the ArabiCorpus were half from modern literature and half from newspapers (200 from each section). The total number of examined sentences in this study was 1,600: 800 for *'ida* and 800 for *law*; half MSA and

half EC each.

3.4 Classifying the samples

The verb tense of both was classified into one of the following groups:

perfect
imperfect
future
equational/no verb
bi
kān+perfect
kān+imperfect
kān+future
kān+equational
kān+*bi*
kān+*kān*+perfect
kān+*kān*+imperfect
kān+*kān*+future
kān+*kān*+equational
kān+*kān*+*bi*

Also, the use of the connecting particles *fa* or *la* was recorded. When frequencies were too low for statistical comparisons, group frequencies were occasionally combined with other similar groups to facilitate chi-square analyses. These combinations are explained in greater detail in the results chapter.

3.5 Statistical Analysis

In order to determine whether or not the variations between MSA and EC and their frequencies were significant, chi-squared statistical analyses were performed as described by Weisstein (1999). The chi-squared statistic is useful for comparing the frequencies of different variations to determine whether these variations are a result of more common trends or simply produced by chance. In this case, the goal was to

determine whether certain combinations of verb forms in MSA and EC have different or similar frequencies. Statistically significant results indicate that the variations demonstrate trends related to the particular form of Arabic (i.e., MSA or EC). Chi-square results that were not significant indicated that the variations could have happened by chance due to random variation (Welkowitz, Cohen, & Ewen, 2006).

4. Results and Statistical Analysis

This chapter presents conditional sentence data from the two corpora to answer the primary research questions of the present study. The first section of this chapter addresses frequent verb tenses for the conditional particles *fa* and *law*, comparing verb tense frequencies between EC and MSA. The second section will compare frequency differences between the conditional particles within EC and MSA. The third section will examine the relationship between connecting particles and verb tenses, particularly concerning conditional clauses headed by *law*. Chi-square analyses determine whether tense differences between conditional particles and register are due to chance or features that lead to systematic options.

4.1 Comparing EC and MSA

The first question of this study concerns the most frequently used verb forms in conditional sentences in both EC and MSA. The second question relating to this point is the relation between the two and how they compare. Here each particle will be compared by observing the frequency of each verb form used after the conditional particle in the protasis. Thus, a two-part comparison will be made for each particle for its use in EC and MSA, one comparison for the protasis and the other for the apodosis.

4.1.1 *'Ida Protasis*

The chi-square results for this comparison (see table 4-1) suggest that the differences between EC and MSA with this particle is statistically significant and not occurring by chance: ($\chi^2 = 91.63$, $df = 6$, $N = 800$, $p < .001$).

This comparison demonstrates some key differences between EC and MSA. The use of *bi* as a present habitual marker in EC does not occur in MSA. The use of *kān* as a means of backshifting is also very prevalent in much of the data. There are a few instances of imperfect equational in EC. Most of these included active particles such as *'āyiz*.

4.1.2 *'Ida Apodosis*

The chi-square results for this comparison (see table 4-2) suggest that the difference between EC and MSA with this particle is statistically significant and not occurring by chance: ($\chi^2 = 87.52$, $df = 8$, $N = 800$, $p < .001$).

The same difference between EC and MSA is evident with the presence of *bi*-prefix for a verb. The absence of an apodosis, or a semi-conditional sentence was much more common in EC than in MSA.

Table 4-1: Tense comparison between EC and MSA, ‘*ida* protasis

		Register		Total
		EC	MSA	
Tense <i>kān</i> + equational	Count	145	85	230
	Expected Count	115.0	115.0	230.0
	Std. Residual	2.8	-2.8	
<i>kān</i> + <i>bi</i>	Count	22	0	22
	Expected Count	11.0	11.0	22.0
	Std. Residual	3.3	-3.3	
<i>kān</i> + future	Count	17	3	20
	Expected Count	10.0	10.0	20.0
	Std. Residual	2.2	-2.2	
<i>kān</i> + present	Count	18	47	65
	Expected Count	32.5	32.5	65.0
	Std. Residual	-2.5	2.5	
<i>kān</i> + past	Count	17	17	34
	Expected Count	17.0	17.0	34.0
	Std. Residual	.0	.0	
Past	Count	166	248	414
	Expected Count	207.0	207.0	414.0
	Std. Residual	-2.8	2.8	
Equational	Count	15	0	15
	Expected Count	7.5	7.5	15.0
	Std. Residual	2.7	-2.7	
Total	Count	400	400	800
	Expected Count	400.0	400.0	800.0

Table 4-2: Tense comparison between EC and MSA, *‘ida* apodosis

			Register		Total
			EC	MSA	
Tense Imperative	Count	44	20	64	
	Expected Count	32.0	32.0	64.0	
	Std. Residual	2.1	-2.1		
<i>yikūn</i> + verb	Count	3	13	16	
	Expected Count	8.0	8.0	16.0	
	Std. Residual	-1.8	1.8		
<i>kān</i> + verb	Count	2	9	11	
	Expected Count	5.5	5.5	11.0	
	Std. Residual	-1.5	1.5		
<i>bī</i> + verb	Count	33	0	33	
	Expected Count	16.5	16.5	33.0	
	Std. Residual	4.1	-4.1		
Future	Count	58	74	132	
	Expected Count	66.0	66.0	132.0	
	Std. Residual	-1.0	1.0		
Present	Count	117	147	264	
	Expected Count	132.0	132.0	264.0	
	Std. Residual	-1.3	1.3		
Past	Count	38	39	77	
	Expected Count	38.5	38.5	77.0	
	Std. Residual	-.1	.1		
Equational	Count	57	87	144	
	Expected Count	72.0	72.0	144.0	
	Std. Residual	-1.8	1.8		
None	Count	48	11	59	
	Expected Count	29.5	29.5	59.0	
	Std. Residual	3.4	-3.4		
Total	Count	400	400	800	
	Expected Count	400.0	400.0	800.0	

4.1.3 *Law Protasis*

The chi-square results for this comparison (see table 4-3) suggest that the difference between EC and MSA with this particle is statistically significant and not occurring by chance: ($\chi^2 = 153.25$, $df = 6$, $N = 800$, $p < .001$).

Notable differences between EC and MSA include the higher tendency of EC to favor equational sentences whereas MSA features only a few. There is also a stronger tendency for the past tense to occur in MSA.

Table 4-3: Tense comparison between EC and MSA, *law* protasis

		Register		Total
		EC	MSA	
Tenses <i>kān</i> + verb	Count	10	28	38
	Expected Count	19.0	19.0	38.0
	Std. Residual	-2.1	2.1	
<i>kān</i> + Equational	Count	31	87	118
	Expected Count	59.0	59.0	118.0
	Std. Residual	-3.6	3.6	
<i>bi</i> + verb	Count	17	0	17
	Expected Count	8.5	8.5	17.0
	Std. Residual	2.9	-2.9	
Future	Count	7	0	7
	Expected Count	3.5	3.5	7.0
	Std. Residual	1.9	-1.9	
Present	Count	24	6	30
	Expected Count	15.0	15.0	30.0
	Std. Residual	2.3	-2.3	
Past	Count	197	263	460
	Expected Count	230.0	230.0	460.0
	Std. Residual	-2.2	2.2	
Equational	Count	114	16	130
	Expected Count	65.0	65.0	130.0
	Std. Residual	6.1	-6.1	
Total	Count	400	400	800
	Expected Count	400.0	400.0	800.0

4.1.4 *Law* Apodosis: EC compared with MSA

The chi-square results for this comparison (see table 4-4) suggest that the difference between EC and MSA with this particle is statistically significant and not occurring by chance: ($\chi^2 = 282.45$, $df = 7$, $N = 800$, $p < .001$).

The most prominent differences include in the use of the *bi*-prefix and a much stronger tendency to use past for MSA, possibly opposed to future or present equational.

Table 4-4: Tense comparison between EC and MSA, *law* apodosis

		Register		Total
		EC	MSA	
Tenses Imperative	Count	27	5	32
	Expected Count	16.0	16.0	32.0
	Std. Residual	2.8	-2.8	
<i>kān</i> + verb or equational	Count	25	50	75
	Expected Count	37.5	37.5	75.0
	Std. Residual	-2.0	2.0	
<i>bi</i> + verb	Count	62	0	62
	Expected Count	31.0	31.0	62.0
	Std. Residual	5.6	-5.6	
Future	Count	121	91	212
	Expected Count	106.0	106.0	212.0
	Std. Residual	1.5	-1.5	
Present	Count	89	51	140
	Expected Count	70.0	70.0	140.0
	Std. Residual	2.3	-2.3	
Past	Count	11	185	196
	Expected Count	98.0	98.0	196.0
	Std. Residual	-8.8	8.8	
Equational	Count	63	16	79
	Expected Count	39.5	39.5	79.0
	Std. Residual	3.7	-3.7	
None	Count	2	2	4
	Expected Count	2.0	2.0	4.0
	Std. Residual	.0	.0	
Total	Count	400	400	800
	Expected Count	400.0	400.0	800.0

4.2 Comparing between '*ida* and *law*

I now compare the two conditional particles within the same register.

4.2.1 '*Ida* and *law* MSA: Protasis

The chi-square results for this comparison (see table 4-5) suggest that the difference between EC and MSA with this particle is statistically significant and not occurring by chance: ($\chi^2 = 38.47$, $df = 4$, $N = 800$, $p < .001$).

The difference between the two is not large. However, *law* appears to allow more tenses not expected for a protasis in MSA, such as present and present equational.

4.2.2 '*Ida* and *law* MSA: Apodosis

The chi-square results for this comparison (see table 4-6) suggest that the difference between EC and MSA with this particle is statistically significant and not occurring by chance: ($\chi^2 = 218.52$, $df = 6$, $N = 800$, $p < .001$).

Notable differences include imperative occurring more in '*ida* and simple past being much more frequent for the apodosis for *law*.

4.2.3 '*Iza* and *law* EC: Protasis

Overall '*iza* has a stronger tendency to favor combination tenses with *kān* whereas *law* more frequently features the simple past and present equational.

The chi-square test for this comparison (see table 4-7) suggests that the differences in tenses between EC and MSA for this particle in the protasis are not by chance: ($\chi^2 = 249.23$, $df = 6$, $N = 800$, $p < .001$).

Table 4-5: Comparison between *'ida* and *law*, MAS protasis

		Register		Total
		<i>'ida</i>	<i>law</i>	
Tense <i>kān</i> + verb	Count	67	28	95
	Expected Count	47.5	47.5	95.0
	Std. Residual	2.8	-2.8	
<i>kān</i> + Equational	Count	85	87	172
	Expected Count	86.0	86.0	172.0
	Std. Residual	-.1	.1	
Present	Count	0	6	6
	Expected Count	3.0	3.0	6.0
	Std. Residual	-1.7	1.7	
Past	Count	248	263	511
	Expected Count	255.5	255.5	511.0
	Std. Residual	-.5	.5	
Equational	Count	0	16	16
	Expected Count	8.0	8.0	16.0
	Std. Residual	-2.8	2.8	
Total	Count	400	400	800
	Expected Count	400.0	400.0	800.0

Table 4-6: Comparison between *'ida* and *law*, MSA apodosis

			Register		Total
			EC	MSA	
Tenses Imperative	Count		27	5	32
	Expected Count		16.0	16.0	32.0
	Std. Residual		2.8	-2.8	
<i>kān</i> + verb or equational	Count		25	50	75
	Expected Count		37.5	37.5	75.0
	Std. Residual		-2.0	2.0	
<i>bi</i> + verb	Count		62	0	62
	Expected Count		31.0	31.0	62.0
	Std. Residual		5.6	-5.6	
Future	Count		121	91	212
	Expected Count		106.0	106.0	212.0
	Std. Residual		1.5	-1.5	
Present	Count		89	51	140
	Expected Count		70.0	70.0	140.0
	Std. Residual		2.3	-2.3	
Equational	Count		11	185	196
	Expected Count		98.0	98.0	196.0
	Std. Residual		-8.8	8.8	
Past	Count		63	16	79
	Expected Count		39.5	39.5	79.0
	Std. Residual		3.7	-3.7	
None	Count		2	2	4
	Expected Count		2.0	2.0	4.0
	Std. Residual		.0	.0	
Total	Count		400	400	800
	Expected Count		400.0	400.0	800.0

Table 4-7: Comparison between *'ida* and *law*, EC protasis

		Register		Total
		<i>'iza</i>	<i>law</i>	
Tense <i>kān</i> + verb	Count	74	10	84
	Expected Count	42.0	42.0	84.0
	Std. Residual	4.9	-4.9	
<i>kān</i> + equational	Count	145	31	176
	Expected Count	88.0	88.0	176.0
	Std. Residual	6.1	-6.1	
Future	Count	0	7	7
	Expected Count	3.5	3.5	7.0
	Std. Residual	-1.9	1.9	
Present	Count	0	24	24
	Expected Count	12.0	12.0	24.0
	Std. Residual	-3.5	3.5	
Past	Count	166	197	363
	Expected Count	181.5	181.5	363.0
	Std. Residual	-1.2	1.2	
Equational	Count	15	114	129
	Expected Count	64.5	64.5	129.0
	Std. Residual	-6.2	6.2	
<i>bi</i> + verb	Count	0	17	17
	Expected Count	8.5	8.5	17.0
	Std. Residual	-2.9	2.9	
Total	Count	400	400	800
	Expected Count	400.0	400.0	800.0

4.2.3 '*Iza* and *law* EC: Apodosis

The chi-square test for this comparison (see table 4-8) suggests that the difference in tenses between EC and MSA for this particle in the protasis is not by chance: ($\chi^2 = 109.73$, $df = 7$, $N = 800$, $p < .001$).

The difference is not as dramatic. However, there appears to be a stronger tendency for *law* to favor a future tense verb whereas present and simple past are more common for '*iza*.

4.3 The connecting particles

Here we review the different connecting particles and the environments in which they appeared in our data. The connecting particles data pertained exclusively to MSA. As mentioned earlier, the particle *law* can feature *la* or *fa* as a connecting particle. Our interest here is the comparison between the use of present and past tense with respect to the connecting particle.

The results for '*ida* did not include much variation. Out of the 400 instances, 225 featured *fa* while only one featured *la*. The remaining 174 did not contain a connecting particle.

The results of the chi-square test suggest (see table 4-9) that the selection of connecting particles for *law* does not occur by chance but rather is systematic in nature ($\chi^2 = 212.39$, $df = 1$, $N = 263$, $p < .001$).

The notable trend here is *fa*'s frequent presence in present tense verses *la*'s frequent presence in past tense.

Table 4-8: Comparison between *'iza* and *law*, EC apodosis

		Register		Total
		<i>'iza</i>	<i>law</i>	
Tense Imperative	Count	44	27	71
	Expected Count	35.5	35.5	71.0
	Std. Residual	1.4	-1.4	
<i>kān</i> + verb or equational	Count	5	25	30
	Expected Count	15.0	15.0	30.0
	Std. Residual	-2.6	2.6	
<i>bi</i> + verb	Count	33	62	95
	Expected Count	47.5	47.5	95.0
	Std. Residual	-2.1	2.1	
Future	Count	58	121	179
	Expected Count	89.5	89.5	179.0
	Std. Residual	-3.3	3.3	
Present	Count	117	89	206
	Expected Count	103.0	103.0	206.0
	Std. Residual	1.4	-1.4	
Past	Count	38	11	49
	Expected Count	24.5	24.5	49.0
	Std. Residual	2.7	-2.7	
Equational	Count	57	63	120
	Expected Count	60.0	60.0	120.0
	Std. Residual	-4	.4	
None	Count	48	2	50
	Expected Count	25.0	25.0	50.0
	Std. Residual	4.6	-4.6	
Total	Count	400	400	800
	Expected Count	400.0	400.0	800.0

Table 4-9: Comparison of connecting particles in MSA

		Tenses		Total
		Past	Present	
Connector <i>la</i>	Count	181	5	186
	Expected Count	132.3	53.7	186.0
	Std. Residual	4.2	-6.6	
<i>fa</i>	Count	6	71	77
	Expected Count	54.7	22.3	77.0
	Std. Residual	-6.6	10.3	
Total	Count	187	76	263
	Expected Count	187.0	76.0	263.0

5. Discussion of Results

In this chapter I discuss the results presented in chapter 4: first by register, then by conditional particle, and finally by connecting particles. The first section compares EC and MSA by means of each conditional particle starting with the protasis then moving on to the apodosis. The second section compares the tendencies of the two particles in both EC and MSA as opposed to comparing the trends of the registers. The third section treats the usage of the connecting particles particularly with *law* in MSA. The fourth section discusses the use of the two corpora and whether or not it was an effective tool for collecting and analyzing this type of data.

5.1 EC and MSA

The comparison between the two registers was revealing in many aspects. Trends for both the protasis and apodosis suggest that each register favors certain tenses over others. Furthermore, certain tenses that were unexpected in both EC and MSA appeared multiple times in MSA.

5.1.1 *ʿIda* Protasis

Many MSA grammars prescribe a past tense of some sort as mandatory after the particle in the protasis. Not surprisingly our MSA samples follow the strict rule whereas the EC samples show a few examples of present tense equational sentences. These mostly occurred as active particles such as *ʔāyiz* (wanting) or *ʿādir* (capable). The example features an active particle in the protasis for an EC conditional headed by *ʿiza*.

(27) *ʿiza l-ʿāb miš muwāfiʿ wa-l-bint muwāfiqa al-bint*

biḥāwil *tiqnaʿ* *ʿabūhā*
 if the-father NEG agree and-the-girl agree the-girl
 try-3SG.HAB convince-3SG.PRS father-her
 ‘If the father doesn’t agree and the daughter agrees, the girl tries to convince her father.’

The fewer total examples of these forms (15 instances) suggests that *ʿida* in EC mostly patterns the trends of MSA in the protasis. In MSA there was not a single instance of a present tense verb or equational sentence. The most common tense in MSA or EC for the protasis here was the past tense.

5.1.2 *ʿIda* Apodosis

The differences between the EC and MSA in the apodosis are the clearest by the presence of the *bi* prefix. The following example features the *bi* prefix on the verb in the apodosis in an EC conditional sentence.

(28) *ʿizā* *štaḡalit* *kitīr* *bitāxud* *ugar* *ʿakbar*
 if work-2SG.PST a lot take-2SG.PRS wage larger
 ‘If you work a lot you get a bigger wage.’

The second notable difference is the tendency of *ʿida* to be used in sentences that do not actually present both the conditions of the protasis and apodosis. Woidich (2005) lists these structures as indirect questions. The samples include a fair amount of these types of questions but also statements that include the same feature. The example contains a statement usage of *ʿiza* in EC without the apodosis (p. 22).

(29) *yigī* *masalān* *yišūf* *ʿizā* *kānit* *ʿāwza*
tīšrib *mayya*
 come-3SG.PRS for example see-3SG.PRS if be-3SG.PST wanter
 drink-2SG.PRS water
 ‘He comes for example to see if he wants to drink water.’

Present tense verbs appeared as the most commonly used tense in the apodosis for MSA conditionals with ‘*ida*’ (147 instances).

5.1.3 *Law* Protasis

Comparing EC and MSA structures of *law* produced multiple points of interest for this study. As mentioned by Eisles (2000), *law* permits high probability as well as very low probability whereas ‘*ida*’ is restricted to high probability only. Our observations confirmed this in the number of instances for *law* that feature a present tense verb or present equational sentence in the protasis (138 for EC and 22 for MSA). Seven instances of the protasis even contained a future tense verb in EC. In light of backshifting, these tenses suggest a higher probability of straying from many of the MSA and even some EC grammars that assign *law* as a low probability marker and ‘*ida*’ for high probability. Some grammars claim that speakers sometimes use these particles interchangeably, hence *law* merely replaces ‘*ida*’ as a marker for a conditional structure but the level of probability remains at the low level which ‘*ida*’ displays. However, the data for ‘*ida*’ EC usage in the protasis featured only 15 instances of present tense. *Law*, on the other hand, featured 138. In other words, our data indicates a trend of *law* permitting present tense in the protasis for high probability.

The few instances (22) of some form of a present tense in MSA for the protasis may indicate that the habits of EC have some effect on the grammar of MSA. Grammar sources claimed that a perfect was required in the protasis and some grammars such as *Al-Kitaab* required both the protasis and the apodosis to have a past tense form. While past tense forms clearly appeared most often (378 instances), the few instances of present

tense indicate that this is something that occurs and is not necessarily an idiosyncratic feature of a single author. The example contains an equational sentence in the protasis of a conditional sentence in an Egyptian newspaper.

- (30) *wa-lākinn-a l-‘amr-u law ṣaḥīḥ fahuwa yastaḥiqq-u l-dirāsa-ta wa-ta’ammul-a*
 however the-matter if true then-it deserve-3SG.PRS the-study and-pondering
 ‘If the matter is true then it deserves study and pondering.’

5.1.4 *Law* Apodosis

The main difference between EC and MSA again lies in EC’s usage of the verb prefix *bi* (62 instances). The contrast between the number of past tense forms (MSA 236 and EC 36) is also of interest. These numbers confirm that there exists a strong tendency for MSA conditionals headed by *law* to feature a past form in the apodosis. However, it is not the only form; future and present also form a substantial portion of the samples (MSA 142). The higher frequency of future tense instead of another past tense may also suggest more usage of higher probability rather than solely lower probability.

5.2 *‘Ida and law*

This next group of comparison deals with the particles themselves in each of their four possible environments: EC or MSA and the protasis or apodosis.

5.2.1 EC protasis

Our earlier analyses show that *‘iza* retains a stricter syntactic favoring of past tense and very few instances of present (15 instances of present equationals). Given *law*’s 114 instances, *law* has much more flexibility in EC than *‘iza*. Even though both particles

featured a simple past verb as the most frequent form (166 for *'iza* and 197 for *law*), *law*'s more flexible grammar sets it apart from *'iza*.

'Iza allegedly marks higher probability conditions, whereas *law* is for lower probability. However, given *law*'s larger range of application for probability the data suggest that the use of present in the protasis helps indicate a condition that is high in probability. Conversely, the use of a past aligns more closely to its commonly prescribed role as intended for conditions of lower probability. Since we do not see this same range from *'iza*, with the exception of certain active participles, the strong use of past tense may be exclusively for backshifting and not for marking probability as it appears overwhelmingly in the past tense.

5.2.2 EC apodosis

The comparisons for the apodosis are not quite as revealing as the protasis but the findings are still informative. They suggest that *'iza* is used more for conditional sentences that do not contain two conditions or what some may deem as “incomplete”. In other words, the condition lacks a result and only contains a proposed circumstance. The following example was taken from a phone conversation between two Egyptians:

- (31) *'ala l-'amūm bass miš 'ārīf 'iza kān dā*
ḥaī'i yi'ni
 upon the-generality NEG knower if be-3SG.PST this
 true
 ‘Anyway, but I don’t know if that’s true.’

5.2.3 MSA Protasis

The same difference between *law* and '*ida*' in EC is still present to some extent in MSA. While the instances of present tense forms in the protasis are few for *law* (22), there are none for '*ida*'.

- (32) *law ya-s'alūna-ni man huwa henry miller fa-sa'jibuhum*
'*anna-hu tabīb*
if ask-3PL.PRS who he Henry Miller then-answer.1SG.FUT
that-he doctor
'If they ask me who Henry Miller is, then I will answer then that he is a doctor.'

5.2.4 MSA Apodosis

The data for the apodosis suggests that *law* tends to trigger past more than '*ida*' (236 *law* and 62 '*ida*'). This is opposite of the trend that we saw in EC where the apodosis had a much lower frequency of past forms for *law* (36) whereas '*ida*' had around a similar number (82). This comparison indicates that *law*'s status as a marker of low probability is more common in MSA rather than its status in EC, which had a much wider range of probability.

5.3 Connectors

We next discuss the different connectors and their possible relations. The connectors used for '*ida*' did not provide much variation. Out of the 400 samples that were examined in MSA, 174 samples did not contain any connecting particle, 225 samples featured *fa* as a connector and only one instance featured *la*.

The particle *law* in MSA revealed a strong trend. Out of 400 samples, 77 featured the connecting particle *fa*. Out of these findings, a past tense verb followed six while a form of an imperfect such as an imperative, future or present tense followed the other 71.

In addition, 186 instances of the use of *la* occurred in our data. Five of these featured an imperfect form while 181 featured a past tense form. These trends support the wider range of *law* when compared to ‘*ida*. ‘*Ida* appears to have a much narrower range as it focuses mainly on high probable conditionals that tend to take an imperfect in the apodosis. The connector for these conditionals, *fa*, seems to follow that same trend when used with *law*. This allows for a much more open usage when compared to that of ‘*ida*, which only had 5 instances of the use of *la* as a connector. The following examples contain conditional particles headed by ‘*ida* featuring the connecting particle *fa* (67) and *la* (68):

- (33) *wa-‘ida najad-u sayāda-ti* *naḍarna ‘anna ‘iqlīm-a l-‘utmāniyya-ti* *‘ila kūsūfā* *l-tārīx-i* *l-ḥadīt-i kāna* *fa-‘anna-na taḥta l-*
and-if look-1PL.PST to the-history the-modern then-we
find-1PL.PRS that territory Kosovo be-3SG.PST under the-
rule the-Ottoman
‘And if we look at modern history we find that the Kosovo territory was under Ottoman rule.’
- (34) *wa-‘ida l-qarārāt-i ḥukūma-ti* *naḍarna la-staša’rna* *fī* *falsafa-i* *hādihi*
and-if look-1PL.PST in philosophy this
the-decision.PL then-aware.1PL.PST upon the-instance idea the-
government
‘If we look at the philosophy of this decision we will be aware of the government’s idea.’

The following examples of conditional headed by *law* in MSA contain the connecting particle *fa* (69) and *la* (70):

- (35) *law ‘innā l-addīna* *‘ann-a hunāka sayafūz-u* *ntixābāt-in* *nazīha-tan* *bi-l-kāmil-i* *fa-*
if that there election-PL free in-complete then-
those who win-3SG.FUT they the-independent-PL
‘If there are free completely free elections then those who will win are the independents.’

- (36) *wa-law naḍarna fi qaṣīda l-kūlira li-nāzik la-wajada-na-*
ha mu'allifa-tan min 'arba'a-ti maqāfi'-in
 and-if look-1PL.PST in poem Colera to-Nazek then-find-
 1PL.PST-it composed of four part-PL
 'And if we look at the Colera poem of Nazek then we find that it is composed of
 four parts.'

5.4 The Corpora

The empirical evidence gathered for this study came from two corpora. I discuss here the process of using the corpora and how it is beneficial and also some drawbacks to the current approach.

Using the ArabiCorpus greatly facilitated the process of gathering data from multiple sources. The variety included newspapers and literature, allowing for a focus on a certain aspect of MSA as opposed to religious texts that may favor CA. Also, all the data here are naturally occurring language that appear in authentic real-life contexts, as opposed to sentences contrived to display a grammatical principle or capability. The corpus approach is also a departure from the use of native speaker intuition, which could involve large amounts of metacognition rather than natural production in authentic discourse.

While the benefits of corpora research outweigh the drawbacks, a few limitations still exist that should be mentioned. The EC corpus consists of speech from a smaller group of speakers. With a little over 40 articles, most having a unique author, along with a novel and a play, the number of authors for the EC corpus is still lower when compared to the amount of authors for MSA, particularly the newspaper section. The literature section of the MSA corpus does have a similar issue since a large portion contains writings from one particular author, Naguib Mahfouz. This smaller group of speakers

may in turn give more influence to idiosyncratic features of their grammar. The other portions of the EC may also suffer from the same problem as the speech from the CallHome portion features long conversations between two speakers and the episodes of the series may also have a larger influence by a smaller group of writers. The only section that most likely does not have this issue would be the internet chat section as various users comment.

A final difficulty faced during the use of the corpora was the use of a different register than that which was expected. While many Arabs may feel that EC and MSA are not the same means of communication, an agreed-upon unit of features that marks something specifically as MSA does not exist (Parkinson 1991, p. 33). However there are some features that seem uncharacteristic of a certain register. These may include lesser frequent vocabulary items such as *ḥāsūb* for computer in MSA but commonly a borrowed form in other EC, *ṣāḥib* in EC as friend but the meaning differs in MSA meaning owner instead. Also *'urīd* in MSA meaning 'I want' as opposed to *ʔāyiz* in EC. Some grammatical features include *miš* as a negator in EC instead of *laysa* in MSA, using *mā+š* to negate instead of *la* or *lam* as normally used in MSA, and also the use of *'illī* to head a definite relative clause instead of *l-ladī*. These features are would seem out of place if they occurred in another register. It was along these lines that a distinction between EC and MSA was made if deemed necessary. This generally occurred in two areas: some quotes in newspapers that contained EC instead of the expected MSA, and some comments in the chat section consisted of an overwhelming amount of features not common of the EC register. If the sample featured only one of these markers of register,

then it was marked as the same register as expected: newspaper sources remained MSA and internet chat remained EC.

The following example was taken from the Egyptian chat section of the corpus and was not included in the study's statistics as it clearly displays MSA features such as vocabulary and negators.

- (37) *'ida 'aradt-a 'an tastabaqi sa'āda-taka l-*
zawjiyya-ta fa-la tantaqid
 if want-2SG.PST to preserve-2SG.PRS happiness-your the-
 marital then-NEG criticize-2SG.IMP
 'If you want to preserve your marital happiness then don't criticize.'

This following conditional sentence, which occurred in an Egyptian newspaper, was not included among the EC samples as it contained words clearly EC specific words.

- (38) *'iza kānit sākta 'ala ziyāda 'as'ār 'aklna*
'āyzāha titkallim 'an 'as'ār il-lā'bīn
 if be-3SG.PST quiet upon increase price.PL food-our
 want-her talk-3SG.PST about price-PL the-player.PL
 'If she's quiet about the increase of the prices of our food, I want her to talk about the prices of the players.'

6. Conclusion

The aim of the present study was to use data from two Arabic corpora to relationships between verb tense and connecting particles in both EC and MSA. The first section of this chapter presents the conclusions of the present study with regards to the primary research questions. The second section addresses limitations of this study and the third section covers possible directions for further studies.

6.1 Answering the research questions

This research sought to answer four questions concerning conditional sentences in EC and MSA:

1. What are the most frequent verb tenses in conditional clauses in both MSA and EC?
2. Are these verb tenses specific to the conditional particle and register? Can we expect the same frequency in MSA and EC for the same conditional particle?
3. Do structures that are traditionally viewed as EC structures appear in MSA writing?
4. Is there a relationship between the verb tense of the proposed result and the connecting particle?

Each conditional particle demonstrated certain tendencies for verb tenses we analyzed in the chi-square test. The two particles '*ida* and *law* differed in preferred tenses. These differences were also present when comparing EC to MSA. In other words, each

particle and register displays a frequency indicative of its register and conditional particle.

The comparison between the two revealed many different contrasts and one that particularly stands out. The tenses used in EC allow for a wider range, particularly in the protasis. The protasis is almost always followed by perfect verbs. However, EC strays from this and features many examples contained the imperfect tense, whether a present tense verb or an equational sentence.

It appears that EC may be having an influence on MSA, particularly in the use of present tense verbs in the protasis for conditions headed by *law*. *Law* was described in the prescriptive grammars as having a wider range of probability and this can even be seen in MSA. When compared to ‘*ida*, *law* allows for some present tense verbs much more than ‘*ida* in both EC and MSA. MSA grammars described *law* as being followed by a past tense. However, the data demonstrates that uses of present tense in the protasis also occur in MSA and not just EC.

The conditional particle *law* displaced a strong tendency to take the particle *la* when followed by a past tense verb and *fa* when followed by an imperfect. The particle ‘*ida* almost exclusively featured *fa* as the connecting particle while *la* was only used once. Verb tense clearly has a relation to the connecting particle.

6.2 Limitations

Although the present study was successful in answering trends for certain particles in EC and MSA, it wasn't without its limitations. The first problem was that the scope of the study only included MSA and EC. While the corpora did contain sources of CA such as the Quran, medieval grammars, and Hadiths, a complete study of CA, MSA,

and EC would require the study of instances of *'in* as well. This presents many ambiguities and false hits as the spelling of Arabic does not allow for quick searches for specific meanings of *'in*. A use a more

6.3 Future studies

The present study was particularly unique because of its corpora based approach. This method provided statistical evidence to examine common trends in authentic language use. This differs from many studies that utilize native speaker intuition and a qualitative approach to discover trends. Such a statistical approach is not as common for Arabic due to the lack of corpora when compared to other languages. Because of this unique approach this study revealed much information about EC and MSA conditionals and may also serve as a starting point for many other studies. The current study focused on the relation between EC and MSA as written in Egypt. Further studies could treat the relationship between the colloquial and written language in other Arab nations such as Morocco, Syria, and Iraq. A cross-dialectal comparison could also reveal similarities and contrasts across the Arab world.

Also, the relationship between local dialects and MSA practices of that region can be further studied. Corpora such as ArabiCorpus that contain texts spanning decades would serve as a tool for measuring influences of dialects in MSA over time. Future corpora studies of Arabic could also consider CA and the particle *'in*, providing statistics on the semantic changes that occurred as *'ida* replaced *'in* for high probability conditions. Statistical studies of overall frequency of the particles in CA, MSA, and any dialect would also help further document a clearer picture of how conditionals differ between

dialects and MSA. The use of corpora in linguistics is a powerful tool with much potential to provide information that is not readily noticed on the surface and thus requires large-scale analyses.

Not only does this study reveal overlapping trends in MSA and EC, but it also can serve as evidence of expected trends in other languages. The leak of colloquial practices into standardized means of communication may occur despite efforts to curb these practices with prescribed grammars.

References

- Al-Batal, M., Al-Tonsi, A., & Brustad, K. (2004). *Al-Kitaab: A textbook for beginning Arabic*. Washington, DC: Georgetown University Press.
- Abdel-Massih, E., Abdel-malek, Z., & Badawi, S. (2009). *A reference grammar of Egyptian Arabic*. Washington, DC: Georgetown University Press.
- Badawi, E., Carter, M., & Gully, A. (2004). *Modern written Arabic: A comprehensive grammar*. New York: Routledge.
- Brustad, K.. (2000). *The syntax of spoken Arabic: A comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington, DC: Georgetown University Press.
- Comrie, B. (1986). *Conditionals: a typology*. In Traugott et al. (Eds.), 1986:77-99.
- Eisele, J. (2000). *Arabic verbs in time: Tense and aspect in Cairene Arabic*. *Anthropological Linguistics*, 42, 71-75.
- Hein-Nasr, R., & Woidich, M. (2005). *Kullu tamam! An introduction to Egyptian colloquial Arabic*. Cairo: The American University in Cairo Press.
- Holes, C. (1995). *Modern Arabic: Structures, functions, and variations*. London: Longman.
- Parkinson, Dilworth. (1981). *Searching for modern fusHa: Real-life formal Arabic*. *Al-Arabiyya*, 24, 31-64.
- Parkinson, Dilworth. (1996). *Variability in Standard Arabic Grammar Skills*. *Understanding Arabic: Essays in Contemporary Arabic Linguistics in Honor of El-Said Badawi*. Cairo: The American University of Cairo Press, 91-102.
- Parkinson, Dilworth. (2014, June). *arabicorpus.byu.edu*. Provo: Brigham Young University.
- Ryding, Karin. (2010). *A reference grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.
- Socin, A. (1942). *Arabic grammar: Paradigms, literature, exercises and glossary*. New York: G. E. Stechert & Co.
- Thatcher, A. D. (1958). *Arabic grammar of the written language*. London: Percy Lund, Humphries.
- Weisstein, E. W. (2014). *Chi-squared test*. Retrieved from A Wolfram Web Resource: <http://mathworld.wolfram.com/Chi-SquaredTest.html>.

Welkowitz, J., Cohen, B. H., & Ewen, R. I. (2006). *Introductory statistics for the behavioral scientist*. New York: Wiley.

Wright, W. (1975). *A grammar of the Arabic language*. London: Cambridge University Press.