Spring 2019

Acquisition of the English Copula by Arabic Speaking ESL Learners: Evidence for Feature Reassembly

Jenna Steiner

ACQUISITION OF THE ENGLISH COPULA BY ARABIC SPEAKING ESL LEARNERS: EVIDENCE FOR FEATURE REASSEMBLY

by

Jenna Steiner

Bachelor of Arts
Georgetown University, 2009

Master of Arts
American University in Cairo, 2011

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Linguistics

College of Arts and Sciences

University of South Carolina

2019

Accepted by:

Mila Tasseva-Kurktchieva, Major Professor

Paul Malovrh, Committee Member

Sherry Warren, Committee Member

Atta Gebril, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

ABSTRACT

This study aims to identify whether the acquisition of the English copula by Arabic-speaking learners of English provides evidence for a performance or representational-based account of errors. The representational theory tested in this study is the Interpretability Hypothesis (Tsimpli & Mastropavlou, 2007) which proposes that language learners have only partial access to Universal Grammar (UG), making some language structures unacquirable for second language learners. The performance theory tested in this study is Feature Reassembly (Lardiere, 2008), which proposes that the source of errors lies with the mapping of features onto morphology rather inside the core computational component of grammar.

In Arabic, there is no overt copula in present tense sentences; however, in past tense, the copula is overt. This stands in contrast to English in which an overt copula is required in both present and past tenses. Therefore, copula omissions in English committed by Arabic speakers learning English can be analyzed in order to determine whether a representational or performance-based theory best accounts for the omissions. Previous research looked exclusively at copula omission in written production only under the condition of tense and framed results only in light of L1 transfer without considering the role of access to UG. This study builds upon previous research by providing a more thorough analysis of copula omission and errors than previous studies and by framing

these results in a more modern theoretical context of performance versus representational deficit.

The participants of this study consisted of 45 Arabic-speaking students studying in English language programs at large public U.S. universities. The participants were divided into two proficiency groups, with 24 participants in the "High" proficiency group and 21 participants in the "Low" proficiency group. Data were collected via a time-pressure grammaticality judgement task and an elicited imitation task. With these two tasks, the effects of tense, verb type, and number were investigated across several syntactic environments. The conditions of tense, verb type, and number are analyzed statistically, and a descriptive analysis is presented for syntactic environments and non-omission copula errors.

Feature Reassembly is best able to account for the data of this study. The Low Group performed better on past tense items than present tense items, better on main verb items than auxiliary items, and better on the singular items than the plural items. All of these results are predicted by Feature Reassembly as constrained by Slabakova (2009). Moreover, these deficits were overcome by the High Group, which is also predicted by Feature Reassembly. Additionally, the descriptive results of this study such as the omission of *-ing*, the doubling of the copula, and the tense and agreement errors on the copula are also predicted by FR. This all suggests that copula errors by highly proficient Arabic speakers learning English are performance-based, not representational, in nature.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

AG ..... Accepting Grammatical

AGR ..... Agreement

EA ..... Egyptian Arabic

EI ..... Elicited Imitation

FDH ..... Fundamental Difference Hypothesis

FFFH ..... Failed Functional Feature Hypothesis

FR ..... Feature Reassembly

FTFA ..... Full Transfer/Full Access

GJT ..... Grammaticality Judgment Task

IH ..... Interpretability Hypothesis

L1 ..... First Language

L2 ..... Second Language

MA ..... Moroccan Arabic

MSIH ..... Missing Surface Inflection Hypothesis

NP ..... Noun Phrase

NPI ..... Negative Polarity Item

RU ..... Rejecting Ungrammatical

SA ..... Standard Arabic

SLA ..... Second Language Acquisition

TP ..... Tense Phrase

UG ..... Universal Grammar

VP ..... Verb Phrase

CHAPTER 1

INTRODUCTION

The present study investigates the acquisition of the English copula by Arabic-speaking learners of English. There is an abundance of research which documents the fact that Arabic speakers learning English tend to omit the copula in production (Scott & Tucker 1974; Asfoor 1978; Assubaiai 1979; Al-kasimi, Topan & Khan 1979; Beck 1979; Kambal 1980; Sharma 1981; Al-Muarik 1982; El-Badarin 1983; Thompson-Panos & Thomas-Ruzic 1983; Abu-Ghararah 1989; Al-Zahrani, 1993; Kassem, 2000; Al-Shayban, 2007; Al-Buainain 2009; Chalikandy 2015). There have been three studies, Al-Zahrani (1993), Kasem (2000), and Al-Shayban (2007), which have specifically investigated issues of learnability of the English copula by Arabic speakers. These studies focused solely on the extent to which first language (L1) transfer plays a role in copula omission, with Al-Zahrani (1993) and Al-Shayban (2007) finding evidence which they claim supports L1 transfer, and Kasem (2000) finding evidence which he claims does not support L1 transfer. However, these studies suffer from important methodological and theoretical limitations, and it is these limitations which this study aims to address.

One major methodological limitation is that the previous research on copula omission by Arabic-speaking learners of English looked exclusively at written production. Looking exclusively at written production provides only a limited insight into the interlanguage of the ESL learner. In writing, learners have time to think about what it is they want to say, how they want to say it, the rules they know for how to

express certain grammatical structures, and they can even revise and edit as they write (James, 2013). In many cases, where errors have been erased or crossed out, this leaves only a modified form of the learner's interlanguage on the paper. For instance, in the case of the English copula, a writer could initially omit the copula in a sentence in their writing, but later when rereading their writing realize that their sentence is missing a verb and add it in. In this case, it would look like there was no error in copula omission in the writing, even though the learner's initial or more spontaneous language omitted the copula. It is these types of errors in particular, errors of a more spontaneous nature, that this paper is concerned with.

An additional methodological limitation of previous research is that only two of the studies investigated the role of tense in copula omission, while none of the studies investigated the role of copular function (i.e. copula as a main verb and copula as an auxiliary verb), or number. The copula is a particularly interesting grammatical structure to investigate because the Arabic copula in the past tense behaves the same way as the English copula, but the Arabic copula in the present tense behaves differently from the English copula. Therefore, Arabic speakers learning English already have the copular structure that they need for producing the English copula available to them in their first language, but they need to learn to produce it in environments that are different from their first language. These differences in environment are based on tense, which also interacts with verb type and number agreement. This study uses the progressive aspect as a way of investigating the acquisition of the copula as an auxiliary verb. In formal Arabic and in many dialects, the construction of the progressive aspect is subject to the same past/present distinction as mentioned above. In the past tense, the progressive aspect in

Arabic is essentially the same as in English in the sense that it consists of a copula plus a verb, but unlike English there is no progressive marker on the verb (e.g. -*ing*). In the present tense progressive form, the copula is still omitted, and since there is no progressive marker on the verb, the difference between present simple and present progressive is understood contextually rather than grammatically. Looking at the acquisition of the copula as an auxiliary verb allows for the investigation of the acquisition of the copula in a new environment, the present tense, plus the additional grammatical feature of aspect. Additionally, by comparing the acquisition of the copula in singular and plural forms, this study can investigate the acquisition of the copula in the new environment of present tense, plus the additional grammatical feature of number. By considering both verb type and number agreement in the acquisition of the copula, this study aims to determine whether the acquisition of the copula alone is problematic for Arabic-speaking learners of English or whether the features being realized on the copula also play a role. In turn, the answer to the above question will inform a broader theoretical debate as to the learnability of grammatical features.

The previous research on copula omission by Arabic speakers learning English takes a generative, rather than functional, approach. The present study also adopts a generative approach to second language acquisition, so discussion of previous research is limited to generative-based research. In terms of a generative approach, the major theoretical limitation of previous research on copula omission is that the researchers looked at their results only in light of L1 transfer without considering the role of access to Universal Grammar (UG). There are Second Language Acquisition (SLA) theories which state that second language learners have no access to UG, meaning that any language structure that

is not available to the learner through their L1 will never be fully acquired (Bley-Vroman, 1990; Clahsen & Felser, 2006). Other SLA theories state that language learners have partial access to UG, making some language structures which are absent in the L1 acquirable while others are not (Hawkins & Chan, 1997; Tsimpli & Mastropavlou, 2007). Lastly, some SLA theories state that language learners have full access to UG, meaning that target-like syntactic representations of all language structures are fully acquirable (Schwartz & Sprouse, 1996; Vainikka & Young-Scholten, 1996). In the end, the no-UG-access theories needed to be able to account for the fact that it appears that second language learners are frequently able to learn certain language structures absent in their L1. The partial-UG-access theories needed to come up with a systematic account of why some language structures are acquirable while others are not. Lastly, the full-UG-access theories needed to account for the fact that errors seem to persist even at advanced levels of proficiency, which should not be the case if all learners have full access to target-like language structures.

The no-UG-access and partial-UG-access theories can be grouped as representational-based language errors. Specifically, an error that is representational in nature means that the interlanguage syntax is not target-like (i.e. there is a deficit), and the error can be attributed to this deficit. To put it simply, the error is in the syntax itself. Representational research in recent years has focused on identifying which language structures are unacquirable (i.e. errors in syntax persist even at the highest levels of proficiency) versus those that are fully acquirable. A representational account stands in contrast to a theory of performance-based errors, in which the interlanguage syntax is fully target-like, and the source of the error lies outside of the core computational

component of grammar. Essentially, this means that the non-native like performance of highly proficient second language learners is due to the fact that although these language learners have acquired the abstract syntactic properties of the language, a failure to consistently associate them with the correct morphological forms persists even at the highest levels of proficiency. This approach explains how a learner may have full access to UG but continue to make errors regardless of proficiency. Therefore, in further building upon previous research, this study ultimately aims to identify whether the acquisition of the English copula by Arabic-speaking learners of English provides evidence for a performance or representational-based account of errors. In order to answer this question, this study examines whether a partial-UG-access theory or a full-UG-access theory best accounts for the data of this study.

The partial-UG-access theory being tested in this study is the Interpretability Hypothesis (IH) (Tsimpli & Mastropavlou, 2007). According to the IH, interpretable features are acquirable by second language learners, but uninterpretable features are not. Drawing on White, et al. (2003), inherent features on a head, such as gender on a head noun, are interpretable. Alternatively, those features that are not inherent but rather reflect the features found on another head, such as gender on an adjective, are uninterpretable. It is these uninterpretable features that IH predicts to be permanent deficits in second language learners. In terms of the copula, tense is an inherent feature of a verb, so it is interpretable. Therefore, IH would predict tense to be acquirable on the copula. Alternatively, number is an inherent feature of a noun, not of a verb. Number on a verb reflects the number feature on the noun, so it is uninterpretable. Therefore, IH

would predict number to not be acquirable on the copula. Thus, under the IH, we would expect participants to make relatively fewer errors in tense and more errors in number.

Alternatively, the full-UG-access theory being tested in this study is Feature Reassembly (FR) (Lardiere, 2008). Under FR, second language learners can have target-like syntax while still making variable errors in inflectional morphology. Thus, under FR, we would expect to find evidence that participants have acquired the syntax of the English copula (as evidenced by the fact that they perform equally well regardless of tense or verb type), but that they make occasional errors of omission or over rely on default forms when it comes to tense or number.

This study builds upon previous research by providing a more thorough analysis of copula omission and errors than previous studies and by framing these results in a more modern theoretical context of performance versus representational deficit. The driving question is: Do Arabic EFL learners at intermediate and advanced levels of proficiency have target-like syntactic representations of the copula but simply fail to produce them accurately due to a "performance deficit", or alternatively, is it the case that participants do not have the target-like underlying representation of the copula and consequently have a "representational deficit"? Data to answer this question were collected via a Grammaticality Judgment Task (GJT), in which participants were asked to mark written items as either grammatical or ungrammatical, and an Elicited Imitation (EI) task, in which participants listened to and repeated grammatical items. In aiming for a more thoroughly articulated study of these phenomena, the present study investigates participants' performance on these tasks along four parameters: tense (i.e. present or past), verb type (i.e. main verb or auxiliary verb), number (i.e. plural or singular) and

syntactic environment (negation, subject-auxiliary inversion, embedded clauses and *wh*-questions). Tense is included as a parameter because Arabic uses the copula in the past tense, but not in the present, whereas English uses the copula in both tenses. Therefore, if participants perform significantly better on the past tense than on the present, this may be indicative of a representational deficit. Verb type is included as another parameter because progressive aspect is encoded differently in Arabic in present and past tense sentences. In present tense progressive sentences, the progressive aspect is encoded contextually in Modern Standard Arabic because there is no overt copula. To express the past tense progressive, the past tense copula precedes the lexical verb. Therefore, if participants were to perform quite poorly on present progressive items as opposed to past progressive items, this may be indicative of a representational deficit. Furthermore, it would be interesting to compare participants' accuracy on main verb items and auxiliary verb items to investigate whether or not the added challenge of learning how to encode progressive aspect in English would cause participants to perform worse on auxiliary items than on main verb items. Number is included as a parameter because it allows us to investigate a feature which is argued by many to be uninterpretable on the verb, and we can compare the acquisition of number to the acquisition of a different feature, tense, which is generally considered to be interpretable on the verb. Interpretable and uninterpretable features are discussed in more detail in Section 2.1. Lastly, four syntactic environments are included as an attempt to compensate for the lack of free production data in this study. The goal was to include items containing as many different contexts for copula omission as possible while also keeping the number of items to a reasonable amount. Therefore, while syntactic environment is not a primary focus of this study, the

results are analyzed descriptively in the hopes that future research will build upon the very preliminary syntactic environment results of this study. Ultimately, the results of this study suggest that the participants of this study do have a target-like syntactic representation of the copula, and that the errors committed by the participants can be best accounted for by Feature Reassembly (Lardiere, 2008), thus suggesting that the errors are performance-based in nature.

CHAPTER 2

A REVIEW OF THE LITERATURE

The theoretical framework for this study is one that assumes that the source of second language learner errors is either performance or representational in nature. A performance deficit account argues that second language learners are fully capable of achieving native-like underlying representations of functional categories and their features via access to UG. They further argue that errors in inflectional morphology are not necessarily representative of the underlying grammar of a learner, but rather can be accounted for by looking at processing limitations, mapping problems, etc. A representational deficit account, on the other hand, would argue that there is an impairment or deficit in the functional categories, features, or feature strength in the interlanguage of a second language learner. Many representational deficit theories argue that there are deficits or impairments in the interlanguage syntax which are permanent, and that the learner can never achieve native-like underlying representations of the target language. This is generally attributed to the fact that adult language learners do not have access, or only have partial access, to UG. By focusing on the acquisition of the English copula by Arabic speakers, we are able to investigate the interplay between syntax and the relevant syntactic features, which are realized morphologically. By looking at the rate of copula omission across tense, verb type, and number, we can investigate the variable errors of copula omission, and how those errors interact with the tense, aspect,

and number features. This will ultimately inform whether the errors are fundamentally representational or performance-based in nature.

The representational deficit theories are presented in Section 2.1. The section first presents the no-UG-access theory, the Fundamental Difference Hypothesis (Bley-Vroman, 1990) before moving into the partial-UG-access theories of the Failed Functional Features Hypothesis (Hawkins & Chan, 1997) and the Interpretability Hypothesis (Tsimpli & Mastropavlou, 2007), which is the representational theory adopted in this study. The performance deficit theories are presented in Section 2.2. The section starts with a discussion of Full Transfer/Full Access (Schwartz & Sprouse, 1996) and Minimal Trees (Vainikka & Young-Scholten, 1996), as these were fundamental theories which set the stage for future research investigating the disassociation between syntax and morphology, such as the Missing Surface Inflection Hypothesis (Prevost & White, 2000), and the Bottleneck Hypothesis (Slabakova, 2008). Section 2.2 ends with a detailed description of Feature Reassembly (Lardiere, 2008), the performance-deficit theory tested in this study. Section 2.3 discusses the representation of the copula in Arabic syntax. The null copula analysis and the small clause analysis are rejected in favor of the functional projection analysis (Benmamoun, 2000), which is the syntactic representation adopted in this study. Section 2.4 evaluates the limited research that deals specifically with the acquisition of the English copula by Arabic speakers and identifies the theoretical and methodological limitations that this study aims to address. Lastly, Section 2.5 presents the research question of this study.

2.1 Representational Deficit Theories

The theoretical research on representational deficits rejects the idea of full access to UG, meaning that there are certain structures and features that will never become native-like in terms of representation in an L2. This inability of L2 learners to acquire target-like representations of certain functional categories and/or features is generally taken as evidence against full access to UG. As mentioned in the Introduction, there are two categories of representational deficit theories: Those which state that there is no UG access and those which state that there is partial UG access. Although this study does not test a no-UG-access theory, the Fundamental Difference Hypothesis (Bley-Vroman, 1990) is still discussed as it marked an important precursor of much of the modern representational research. For the partial-UG-access theories, the Failed Functional Features Hypothesis (Hawkins and Chan, 1997) is presented briefly for the purpose of demonstrating the groundwork upon which the Interpretability Hypothesis (Tsimpli & Mastropavlou, 2007) was built. Then there is a thorough discussion of the Interpretability Hypothesis, which is the partial-UG-access theory tested in this study.

2.1.1 No access to Universal Grammar theories. The Fundamental Difference Hypothesis (FDH) is a theory proposed by Bley-Vroman (1990), and reconceived in Bley-Vroman (2009), which argues that the same fundamental processes cannot control both child L1 acquisition and adult L2 acquisition; in other words, child L1 acquisition and adult L2 acquisition are fundamentally different. Bley-Vroman (1990) proposes that adult L2 acquisition is much more similar to general adult problem-solving processing than it is to child language development. In support of this theory, Bley-Vroman points out ten characteristic ways in which adult L2 acquisition differs from child L1 acquisition

including the overall lack of success, variation in success, fossilization, and importance of instruction among others. He uses these differences to argue that the logical conclusion is that adults do not have access to the domain-specific language acquisition system (i.e. UG) that children do. He goes on to offer a theory which he claims can account for both the high level of success achievable by an adult L2 learner and also the wide variation in success which is seen in L2 acquisition. Essentially, he proposes that L2 learners are able to "reconstruct" much of the UG structure via access to knowledge about language universals which is present in the L1. It is this (partial) knowledge of universal constraints, combined with general problem-solving and analytical skills, which makes L2 acquisition possible without (direct) access to UG.

In 2009, Bley-Vroman reconceived the FDH under new developments in linguistic theory, specifically the move from "rich" UG to the Minimalist Program (Chomsky, 1995). The original FDH states that children rely on "rich" UG (i.e. access to all language universals, not just those present in L1) for L1 acquisition, and that adult learners are able to acquire certain universal constraints through their L1, without having access to "rich" UG. However, Bley-Vroman (2009) states that under the Minimalist Program, the concept of this "rich" UG no longer exists. He goes on to describe the "essence" of UG under Minimalism as nothing more than "the language property itself" (p. 9). Moreover, Bley-Vroman (2009) is basing his reconceived theory on the argument that under Minimalism, all language principles are instantiated in all languages. Taking this point to its logical conclusion, Bley-Vroman quotes Hale (1996) in saying that under Minimalism, "[t]he distinction [between L1 and UG] becomes vitiated (Bley-Vroman, 2009, p. 317).

In the reconceived FDH, Bley-Vroman concludes that there is no observable distinction between L1 and UG, since all language principles are instantiated in all languages. On this premise, Bley-Vroman goes on state that research which aims to distinguish between access to UG and access to UG through an L1 which instantiates all UG principles is "without content" (Bley-Vroman, 2009, p. 12). Therefore, the Minimalist Program marked a shift in generative SLA research because it no longer makes sense to try to test the FDH by testing for the acquisition of a universal constraint which is "inactive" in the L1, but "active" in the L2, since it is no longer possible to tease apart the effects of L1 (which instantiates all UG principles) and the effects of UG itself. This shifted the focus of the research from those purely testing poverty-of-stimulus problems to those looking more closely at whether or not there is syntactic or semantic evidence of a representational deficit regardless of any variability in surface inflectional morphology.

2.1.2. Partial access to Universal Grammar theories. While FDH is a representational deficit model which states that there is no access to UG, there are representational deficit models which state that there is "partial" access to UG. Under a partial-UG-access model, L2 learners are never able to acquire target-like representations of certain functional categories and/or features. However, unlike FDH, a partial-UG-access theory states that L2 learners can acquire target-like representation of other functional categories and/or features. In essence, this "partial" access allows for some functional categories and features to be acquired while others remain unacquirable. Hawkins and Chan (1997) and Tsimpli and Mastropavlou (2007) subscribe to a partial-UG-access model. Under Hawkins and Chan (1997)'s Failed Functional Features

Hypothesis (FFFH), features of functional categories are inaccessible to adults, meaning second language learners do not acquire target-like representations of these features. However, under FFFH, a non-native speaker may produce target-like language in the L2 without actually having target-like representations of that language. This target-like production is attributed to the second language learner using features available in their L1 in order to imitate target-like production. For instance, when investigating *wh*-movement by Chinese L2 learners of English, Hawkins and Chan argued that Chinese L2 learners do not actually acquire *wh*- operator movement, as it is a functional feature unavailable to Chinese learners of English. Rather, the learners' mental representations involve pronominal binding, a feature available through their L1, which has the result of making their surface language more target-like. Therefore, under the FFFH, a Chinese-speaking learner of English could produce target-like *wh*- movement without actually acquiring *wh*- operator movement, and consequently, the underlying representations of those second language learners diverge from those of native speakers.

Based on this premise, Hawkins and Chan argue that since the underlying representations of the second language learners are not target-like, there must not be full access to UG. Moreover, Hawkins and Chan (1997) say that their theory addresses a problem that FDH cannot. Since FFFH proposes a second language learner can produce target-like language without an underlying target-like representation, the theory can account for the fact that some studies suggest that parameter resetting is possible while other studies suggest that it is not. Essentially, Hawkins and Chan would say that when it appears that parameters have been reset, it is really only that the learners have adopted solutions different from native speakers to account for the perceived differences between

their production and target-like production. Naturally, this theory has been much criticized in that it is simply too abstract to be testable. If we propose that L2 learners can perform in a native-like manner despite a representational deficit, then there is no way to test whether or not their underlying representation is target-like or not. This criticism set the stage for a partial-UG-access theory which could make testable hypotheses about what was acquirable in the L2 and what was not.

Tsimpli and Mastropavlou (2007) propose a theory of partial access that is more relevant here as it is more directly testable, which is why it is adopted and tested in this study as the partial-UG-access theory. The Interpretability Hypothesis (IH) states that all UG principles and operations are available in second language acquisition, but that there are certain features of a grammar which are "uninterpretable" and resist resetting (i.e. cannot be acquired by a second language learner) and other features which are "interpretable" which can be reset (i.e. can be acquired in by a second language learner). Tsimpli and Mastropavlou (2007) differentiate between "interpretable" and "uninterpretable" features by stating that those features that are visible at LF due to their semantic weight are interpretable, and those that are only used in syntactic derivations are uninterpretable. However, Tsimpli and Mastropavlou, while being clear that they are arguing that uninterpretable features resist resetting, did not develop the distinction between interpretable and uninterpretable features to any great extent. Therefore, this study looks to White, et al. (2003) which follows Chomsky (1995) for a more in-depth description of what makes a feature interpretable or uninterpretable.

In an agreement relationship, the inherent properties of one element determine the morphological form of another element. For instance, in Romance languages,

grammatical gender is an inherent lexical feature on head nouns. These inherent features found on the head are the interpretable features (White, et al., 2003). Adjectives and determiners, on the other hand, reflect the gender agreement found on the head noun. Gender is not an inherent feature on adjectives and determiners, but rather it is relational to the head noun. Therefore, gender on determiners and adjectives is uninterpretable, and feature checking needs to occur (White, et al., 2003). Consequently, under IH, an English speaker trying to learn Spanish would be able to acquire gender on nouns, but they would not be able to acquire gender agreement (White, et al., 2003). Under this understanding of interpretable and uninterpretable features, in English, number is an interpretable feature on the head noun, as it is an inherent feature. Number on the verb, on the other hand, is an uninterpretable feature, as it is a relational feature on the verb controlled by the head noun in subject position. Alternatively, both aspect and tense are inherent features of the verb, rather than relational features controlled by another head, so they are interpretable features.

Tsimpli and Dimitrakopoulou (2007) cite their study of *wh*- interrogatives as evidence for the IH. The study included 48 native Greek-speaking participants who were intermediate and advanced-level ESL learners. In Greek, a resumptive pronoun is used in subject and object *wh*- interrogatives as opposed to the "gap strategy" found in English. The examples of Greek below, paired with the English counterpart, show the resumptive pronoun *ton* (him) in instances where a resumptive pronoun would be considered ungrammatical in English.

(1)

a. Pjon ipes oti (**ton**) prosevalan xoris logho?

Whom said.$_{2SG}$ that **him**-insulting.$_{2PL}$ without reason

'Who did you say that they insulted **(*him)** without a reason?'

b. Who did you say that they insulted without a reason?

(2)

a. Pjon fititi ipes oti (**ton**) aperipsan sti sinedefski?

Which student said.$_{2SG}$ that **him**-rejected.$_{3PL}$ at-the interview

'Which student did you say that they rejected **(*him)** at the interview?'

b. Which student did you say that they rejected at the interview?

The participants completed a timed grammaticality judgment task in which they were asked to rate the acceptability of English subject and object *wh*- interrogative items containing either the English "gap strategy" (i.e. no resumptive pronoun), which is grammatical in English, or the Greek resumptive pronoun strategy, which is ungrammatical in English. A grammatical and corresponding ungrammatical object extraction item are shown in (3a) and (3b) below, and a grammatical and corresponding ungrammatical subject extraction item are shown in (3c) and (3d) below for illustrative purposes.

(3)

a. Who do you think Jane likes?

b. *Who do you think Jane likes him?

c. Who have you suggested should not resign?

d. *Who have you suggested he should not resign?

Tsimpli and Dimitrakopoulou (2007) make the case that resumptive pronouns are clusters of uninterpretable case and agreement features in Greek. They argue that the resumptive pronouns "double" the features of the extracted subject or object, or for our purposes, the features on the resumptive pronoun are relational to the features of the subject or object. For this reason, they are uninterpretable and would be predicted to resist parameter resetting. This explanation fits well with the definition of interpretable and uninterpretable features adopted in this study, which is that head features are interpretable and relational features are uninterpretable. In fact, this resistance to parameter resetting of uninterpretable features is precisely what Tsimpli and Dimitrakopoulou (2007) found. At even the most advanced levels, participants continued to accept a resumptive pronoun at the site of the subject or object extraction in *wh*-interrogatives. Tsimpli and Dimitrakopoulou (2007) conclude that the results provide evidence for the idea that uninterpretable features are unacquirable in the second language, and consequently there is a deficit in the interlanguage syntax of a second language learner.

For the purposes of this study, the IH is the most relevant and testable representational-based theory. The IH distinguishes between interpretable and uninterpretable features, and under IH, uninterpretable features resist parameter resetting, making them unacquirable by second language learners who do not have these features in their L1. Arabic speakers learning English have the copular structure available to them in their L1, but they need to learn to realize the structure in a new environment, the present tense. It is clear that this hypothesis would suggest that the participants of this study can acquire the interpretable feature of tense on the verb, and therefore, barring any

performance deficit, should be able to produce it. However, this hypothesis would also predict that the participants would struggle with the uninterpretable feature of number agreement on the verb.

2.2 Performance Deficit Theories

In contrast to the representational-deficit literature above, the performance-deficit models maintain that second language learners can fully acquire a second language grammar, as evidenced by their complete comprehension of the language. However, even highly proficient second language learns commit various errors in speaking. Under a performance-deficit account, these errors are not caused by a representational deficit because syntax has been fully acquired. Rather, the errors are caused by the incorrect mapping between the syntax and the appropriate morphological realization. Section 2.2.1 looks at the historical development of these performance-deficit theories starting with Full Transfer/Full Access (Schwartz & Sprouse, 1996) and Minimal Trees (Vainikka & Young-Scholten, 1996), as they were fundamental theories which set the stage for future research investigating the disassociation between syntax and morphology, such as the Missing Surface Inflection Hypothesis (Prevost & White, 2000), and the Bottleneck Hypothesis (Slabakova, 2008), presented in Section 2.2.2. Lastly, Section 2.2.2. concludes by presenting a detailed description of Feature Reassembly (Lardiere, 2008), the performance-deficit theory tested in this study.

2.2.1 Full Transfer/Full Access and Minimal Trees. Schwartz and Sprouse (1996) proposed the Full Transfer/Full Access (FTFA) hypothesis. Schwartz and Sprouse propose that the full L1 grammar constitutes the initial state of L2 acquisition (Full Transfer), but since L2 learners have full access to UG, parameter resetting (i.e.

target-like acquisition) is possible (Full Access). In this analysis, both functional (e.g. TP, NegP, AgrP, etc.) and lexical projections (e.g. VP, NP, PP, AP, and AdvP) are argued to be transferred from L1. On the other hand, Vainikka and Young-Scholten (1996) proposed the "Minimal Trees" (MT) analysis which differs from FTFA in that lexical projections, but not functional projections, are transferred from the L1. From this initial state, as proposed in FTFA, full UG access allows for parameter resetting. Thus, despite the fact that FTFA and MT differ in their view of L1 transfer (full vs. partial, respectively), they both agree that L2 learners have full access to UG and therefore, are capable of achieving native-like representation of L2 grammar. Taken to the logical conclusion, second language learners should be able to achieve native competence in the target language. However, this is almost never the case. Therefore, what both FTFA and MT were unable to do is provide a viable explanation as to why even the most proficient second language speakers continued to make errors despite full access to UG.

The inability to answer this question is what set the stage for new theories trying to reconcile full access to UG with persistent errors in the second language. Essentially, this question set the stage for theories which posit that participants are committing performance errors rather than representational errors (cf. Bley-Vroman, 1990; Hawkins & Chan, 1997; Clahsen & Felser, 2006; and Tsimpli & Mastropavlou, 2007). It is in this light that the work on performance deficit has been investigated more thoroughly in the Missing Surface Inflection Hypothesis (Haznedar & Schwartz, 1997; Prévost & White, 2000), the Bottleneck Hypothesis (Slabakova, 2008), and Feature Reassembly (Ladiere, 2008). All three of these theories argue that target-like representations of syntactic

structures can indeed be acquired by L2 learners, and that errors in overt morphology are not necessarily indicative of a representational deficit.

2.2.2 Missing Surface Inflection Hypothesis. Prévost and White (2000) extended the research of Haznedar and Schwartz (1997), which posited that errors in overt morphology were not necessarily indicative of a syntactic deficit in child L2 learners, to adult L2 learners. Prévost and White (2000) analyzed spontaneous production of two adult learners of French and two adult learners of German. The researchers proposed, as part of the Missing Surface Inflection Hypothesis (MSIH), that main verbs lacking verbal morphology are still finite. This means the MSIH would predict that finite forms only occur in finite position and seemingly non-finite forms (or forms lacking verbal morphology) would occur in both non-finite and finite positions. This is due to the fact that the non-finite form is sometimes used as a "default" form morphologically, even though it is grammatically finite. Furthermore, they proposed that agreement, when present, would be accurate as the relevant features and feature-checking are assumed to have been acquired. Results showed that, in line with the MSIH, finite forms did not occur in non-finite contexts (indicating variability is constrained) and that agreement, when present, was highly accurate. This led the researchers to conclude that L2 learners can represent finiteness and agreement at an abstract level and that optionality does not imply a major impairment in the interlanguage grammar.

White et al. (2003) and White (2003) provide additional support for the MSIH. The results of White et al. (2003) provide evidence against the representational-deficit idea that L2 learners can only access features instantiated in the L1. The study looked at

the acquisition of Spanish as an L2 by native speakers of French and native speakers of English. Spanish is a language which has gender on the noun and gender agreement on the determiner and adjective. French also has gender on the noun and gender agreement on determiners and adjectives, whereas English does not have gender agreement. Therefore, if one were to assume that adult language learners do not have access to features that are not already present in their L1, then we would expect the French participants of the study to have an advantage over the English participants of the study in terms of gender acquisition. However, this was not the case. The results of a sentence interpretation task, in which participants were asked to choose the correct picture that corresponded with the given sentence, indicated no correlation between accuracy on the task and L1. This means that the native French speakers did not perform any differently than the native English speakers on a task specifically testing gender acquisition. In order for this to be the case, White et al. (2003) argue that the English speakers must have access to the feature of gender even though it is not instantiated in their L1. This provides support for their performance-based account that second language learners are not restricted to uninterpretable formal features already instantiated in the L1.

The focus of White (2003) was slightly different from that of White et al. (2003) in that this study focused on the divergence between inflectional morphology and relevant syntactic knowledge rather than the ability or inability to acquire a feature absent in the L1. Production data were collected from a Turkish participant whose grammar was considered to be in the end-state. The production data showed that the participant's suppliance of both verbal and nominal morphology was variable with suppliance of verbal morphology averaging 80 percent and suppliance of nominal morphology ranging

from 60-87 percent. White (2003) analyzed the participant's data for syntactic properties related to both verbal and nominal morphology. For verbal morphology, White (2003) argues that even when main verbs are left uninflected, there is still evidence that they are in fact finite. This was evidenced by the fact that there was not a single error in pronoun case. The subject of a sentence, regardless of whether the verb was inflected for tense and agreement or not, was always produced in the nominative case. White (2003) argues that this is evidence for the finiteness of the verb regardless of a lack of overt verbal morphology.

In terms of nominal inflection, White (2003) demonstrates the acquisition of a definiteness/indefiniteness distinction regardless of production of relevant articles. The data in the study showed that the participant never used a definite article in an indefinite position or vice versa, but rather the problem was article omission. White looked at the participant's data for violations of the definiteness effect in order to demonstrate that the participant had indeed acquired the definiteness feature. Since definite NPs cannot be used in existential "there" constructions, it was hypothesized that if the acquisition of definiteness had indeed taken place, the participant would not violate this definiteness effect. Results showed that the participant violated this effect (i.e. used a definite NP in an existential "there" construction) only 1.7 percent of the time. This rate was low enough for White (2003) to conclude that the feature of definiteness was acquired regardless of variability in article production. White (2003) takes this as further evidence that variation in verbal and nominal inflection is a result of a breakdown in computation rather than in representation. She described the computation problem as an "interface problem" which reflects trouble in accessing underlying representations, rather than a

problem in the underlying representations themselves. Essentially, the syntax has been acquired but is not being accurately reflected in the morphology.

While the MSIH (White 2003) demonstrated that there can be a divergence between syntactic competence and morphological performance, Slabakova (2008)'s Bottleneck Hypothesis extended this argument to semantics as well. The Bottleneck Hypothesis assumes that morphology is the "tight spot" of language and that syntactic reflexes can be very accurate even when the related functional morphology is produced inconsistently, as argued by White (2003). What makes the Bottleneck Hypothesis distinct is that Slabakova adds evidence from semantics to further support the idea that functional categories can indeed be fully acquired despite variability in overt inflectional morphology. She argues that similar to syntax, knowledge of phrasal semantics also requires the acquisition of functional categories. Therefore, acquisition of phrasal semantics can also be used to demonstrate that relevant functional categories have been acquired, regardless of variability in production of morphology. More specifically, Slabakova (2008) concludes that "performance hurdles" are responsible for "flawed morphosyntax production" but that learners can still achieve native-like syntactic and semantic representations. While the work of White and Prevost (2000) and Slabakova (2008) provided evidence from both syntax and semantics for the divergence between syntactic competence and morphological performance, Lardiere's Feature Reassembly is the theory that really goes one step further to provide a full account as to *why* this is the case. In the discussion of Feature Reassembly that follows, we will see that when FR is constrained in the way proposed by Slabakova (2009), it is capable of making testable predictions about where and when this divergence will occur. In this way, Feature

Reassembly is the most developed of the performance-deficit theories and is most useful for the purposes of this study.

Lardiere (1998), in a precursor to what would eventually be known as Feature Reassembly, also provides support that an L2 learner's performance of inflectional morphology is not necessarily reflective of their syntactic competence. The participant in this study, Patty, who had reached her L2 end-state grammar, omitted verbal agreement 96 percent of the time. Under a representational-deficit account, such as the Interpretability Hypothesis, this omission of verbal agreement would be taken as evidence that the participant has not acquired the agreement feature. However, Lardiere argues that the relevant features for agreement have indeed been selected, but it is the encoding of these features in the morphology which has not been acquired. The evidence for this claim comes from the fact that Patty demonstrates acquisition of the [-strong] feature in Agr. In English, there is a [-strong] feature in Agr which accounts for the fact that verbs do not raise to I. As a result, in English, adverbs precede the verb, as the verb remains in a lower position. In contrast, if we were to compare English to a language like French, which has a [+strong] feature, this feature would trigger the verb to raise to I. As a result, adverbs follow the verb in French. In example (4a), we see that in a grammatical French sentence, the adverb 'often' follows the verb 'watch', in contrast to English, as seen in (4b), where the adverb 'often' must precede the verb 'watch'.

(4)

a. Jean **regarde souvent** la television
John **watch.$_{3SG}$ often** the television
*'John watches often television.'

b. John **often watches** television

Patty, as a native Chinese speaker, is a native speaker of a [+strong] language. This means in her native language, the negation and adverbs should follow the verb, as we saw in the French examples above. Therefore, if Patty had not acquired the [-strong] feature of English, we would expect to find optionality in her verb raising. In other words, if she had not acquired the English [-strong] feature, we would expect her to sometimes place the verb before the negation or adverbs, as a result of non-target-like V-to-I movement). However, there was not a single instance where Patty produced a sentence in which the verb preceded the negation or the adverb. This seems to strongly indicate that Patty had fully acquired the [-strong] feature of English.

Lardiere uses this evidence to conclude that learners can determine both feature strength and the status of verb-raising in the L2, even if verbal morphology is never acquired. Due to this "robust dissociation" between inflectional morphology and syntactic knowledge of formal features, Lardiere concludes that the fundamental differences between L1 and L2 acquisition lie outside of UG or rather outside the core computational component of grammar. It is subsequently in Lardiere (2008) that a more thorough explanation of the theory of Feature Reassembly is provided.

Lardiere (2008) continues to argue that variability in inflectional morphology is not necessarily a result of a failure or inability to reset parameters, as a representational deficit theory would suggest, but rather it is an inability to map the relevant features onto "new or different formal configurations". In this case, variability is defined as "the variable omission, underspecification, overreliance on default forms, and/or apparent optionality vs. obligatoriness of the morphophonological expression of grammar

properties" (p. 107). The heart of Feature Reassembly is that it is not the selection of appropriate features which is problematic, but rather the assembly of these selected features. In order to illustrate this point, Lardiere (2008) provided examples of definiteness and *wh*- movement in the speech of the same Chinese participant (Patty) as Lardiere (1998). The examples of definiteness and *wh*-movement below demonstrate the distinction between selecting a feature and accurately assembling (or reassembling) that feature in order to produce it in a target-like manner.

The discussion of the acquisition of the definiteness feature by a Chinese-speaking learner of English includes a much-simplified explanation of definiteness and number in Chinese. It has been argued that while Chinese lacks a definite article, it does indeed have a definiteness feature. By suffixing the Chinese plural/collective marker *men* to a noun, the noun can then only be interpreted as definite. This can be seen in the example below where the plural noun *xuesheng-men* can only be interpreted as "the students" and not "students" or "some students".

(5)

Ta hui dai xuesheng-men hui jia

He will bring student.$_{PL}$ back home

'He will bring the students back home'

'*He will bring (some) students back home'

Moreover, similarly to English, definite nouns cannot be used in existential constructions. This can be seen in the examples below where (6a) is ungrammatical due to the fact that a noun affixed with *–men* is being used in an existential construction whereas (6b) is grammatical since there is no *–men*. However, it is worth noting that with the removal of

–men, the noun is no longer necessarily plural. Thus, number and definiteness are intricately related in Chinese.

(6)

a. *you ren-men

Have person-$_{PL}$

'There are some persons'

b. You ren

Have person

'There is/are some person(s)'

Since it has been argued that Chinese does indeed have a definiteness feature, then a Chinese speaker learning English would not encounter a problem with *selecting* the definiteness feature, as it is already selected in their L1. However, according to Feature Reassembly, we would expect the Chinese speaker to still face trouble in the acquisition of the definite/indefinite articles since they are realized in a much different way in English than in Chinese. In Chinese, definiteness is inextricably intertwined with plurality and is realized as a suffix on a noun. In English, alternatively, definiteness is relatively independent of number and is realized as a definite or indefinite article. Thus, a Chinese speaker would first need to disassemble or "tease apart" the features of definiteness and number, as realized with *–men*, and then correctly reassemble them so that definiteness is realized on articles. Therefore, we would expect a Chinese speaker learning English to face a pretty big challenge in acquiring the English articles despite the fact that the speaker has already acquired the definiteness feature. In looking at data from Patty, this is precisely what was found. Patty produced articles variably and in non-

native like ways; however, her acquisition of the definiteness feature was demonstrated via the definiteness effect (White, 2003). As mentioned above, definite NPs cannot occur in an existential "there" construction, and Patty never produced a definite article in this context. This shows that a speaker who has fully acquired a feature such as definiteness can still face challenges in realizing that feature, and Feature Reassembly is able to account for this fact via the disassembling and reassembling of the realization of features.

Lardiere (2008) additionally looks at the acquisition of the [+Q] feature by Patty. In English, the verb in yes/no questions undergoes movement from I to C and in *wh*-interrogatives, there is fronting of the *wh*- element to Spec C with the possible insertion of dummy "do". Alternatively, in Mandarin Chinese, there is no I-to-C movement and *wh*- expressions remain *in situ.* An example of a Chinese *wh*- interrogative is provided below.

(7)

Women jintian wanshang chi shemen

We today evening eat what

'What are we having for support tonight?'

The data from Patty ultimately shows that Patty has fully acquired I-to-C movement as evidenced by the fact that she is consistently able to produce the correct movement even with the copula, modals, and auxiliaries. Lardiere makes the claim that this implicates the presence of a [+Q] feature in C, which means Patty was able to switch from the [-Q] feature in Chinese to the [+Q] feature in English. Under a representational deficit account, the resetting of the value of this feature would not be considered to be possible.

To conclude, Lardiere argues that there is a distinct morphological competence that must be acquired by the learner, which she boils down not to feature selection but rather to learning which forms "go with" which features. White (2009) points out that Feature Reassembly is very much "in line" with the Missing Surface Inflection Hypothesis; it seems to be distinct only in that it aims to explain *why* there is a divergence between acquired syntactic properties and their morphological realizations. Additionally, Slabakova (2009) applauds the idea that feature selection alone is insufficient as an explanation of language variation. The performance-deficit accounts have moved beyond the idea that variation among language learners is explainable solely by looking at whether an L1 has selected a certain feature and whether or not that feature can be transferred to an L2 or a new feature can be selected in an L2. The fact that learners must figure out how the feature is encoded in the new language (i.e. the morphology) is at the root of the problem, which is very much in line with the Bottleneck Hypothesis. However, Slabakova (2009) cautions that Feature Reassembly needs to be constrained, and that parameters should not be ignored altogether, as they offer strong explanatory power. If Feature Reassembly zeroes in on every distinction and complexity between features, and simply predicts "difficulty" where there is a difference, while simultaneously ignoring parameters altogether, then predictive power will be lost in terms of the degree of difficulty and order of acquisition. Consequently, Slabakova (2009) brings attention to Ramchand and Svenonius (2008)'s proposed constraints on Feature Reassembly. Ramchand and Svenonius (2008) essentially argue that all languages are capable of expressing universal meaning, and that the syntactic structure for all grammatical meaning is present in all languages, regardless of whether the grammatical

meaning is encoded morphologically in the language. Therefore, as an example, each language is able to express the universal meaning of definiteness and indefiniteness, regardless of whether a language has articles or not. It is how the universal meaning of definiteness and indefiniteness is expressed (morphologically or contextually) in different languages which accounts for language variation. Ramchand and Svenonius (2008) point out that by looking at how languages express the meaning of definiteness, you can group the languages into "types", with expression by morpheme being one type and expression by discourse context being another type.

In the end, Slabakova (2009) combines the categorization power of Ramchand and Svenonius (2008)'s analysis with Lardiere (2008)'s prediction about "mismatched" features in the L1 and L2 and develops a scale of difficulty in the acquisition of L2 features. L2 learners whose L1 encodes gender contextually would face the greatest challenge in acquiring a language in which gender is encoded grammatically. Alternatively, moving from an L1 where gender is encoded grammatically in a different way than how it is encoded grammatically in the L2 would prove slightly less challenging. Lastly, it is predicted that if both languages grammatically encode gender in the same way (i.e. no reassembly is required), then this should be the least arduous acquisition. The predictions made by this scale of difficulty were supported by Slabakova and Cho (2015), which found that English and Korean L2 learners of Russian acquired Russian morphemes that had a corresponding morpheme in the L1 with the same featural representation before they acquired Russian morphemes that did not have overt realizations in the L1.

Feature Reassembly is the most relevant performance deficit theory for the purposes of this study, as this study is looking at the acquisition of the features relevant to the English copula. In light of the constraints proposed by Slabakova (2009), Feature Reassembly is able to make predictions as to when a divergence between syntactic competence and morphological performance will occur and as to how difficult this divergence will be to overcome. Thus, the results of this study will be evaluated to discover whether Feature Reassembly, as a performance-based theory, can account for the data of this study better than the representational-based theory, the Interpretability Hypothesis, discussed above.

2.3 Representation of Copula in Arabic Syntax

In order to understand how a representational deficit could be responsible for copula omission by native Arabic speakers, it is necessary to understand the syntactic composition of the copula in Arabic and how it contrasts with the English copula. In English, every sentence must contain a fully spelled out verb. However, in Arabic, the copula is absent in the present tense. Therefore, the direct translation of (8a) is "The man sick". (Note- the difference in case marking on the adjective *mari:d* is discussed in Section 2.3.1)

This distinction between tenses holds true for expressing progressive aspect as well. For instance, example (9a) can have the interpretation of either "The students are studying" or "The students study". In contrast, (9b), which is in past tense, requires the use of the copula, as evidenced by the obligatory insertion of the past tense, third person, plural form of the copula *kunna.*

(8)

a. r-radʒul-u mari:d-un **SA (Standard Arabic)**

the-man-NOM sick-NOM

'The man is sick'

b. r-radʒul-u ka:na mari:d-an **SA**

the-man-NOM was sick-ACC

'The man was sick'

(9)

a. t-ta:liba:t-u ja-drus-na **SA**

the-students.$_{PL}$-NOM $_{3}$-study-$_{PL}$

'The students study/are studying"

b. t-ta:liba:t-u ***(kunna)** ja-drus-na **SA**

the-students.$_{PL}$-NOM were.$_{3PL}$ $_{3}$-study-$_{PL}$

'The students were studying"

The problem for Arabic syntax is figuring out how to account for these present-tense nonverbal predicates, which can be either NPs (10), PPs (11), or APs (12). In the past tense, the predicates are verbal, as seen in (10b), (11b), and (12b).

(10)

a. Omar muʕellim **MA (Moroccan Arabic)**

omar teacher

'Omar is a teacher'

b. Omar **ka**:**na** muʕellim

omar **was.$_{3SG}$** teacher

‘Omar was a teacher’

(11)

a. al-ktab fuq l-mekteb **MA**

the book on the desk

‘The book is on the desk’

b. al-ktab **ka**:**na** fuq l-mekteb

the-desk book **was.**$_{3SG}$ the-desk

‘The book was on the desk’

(12)

a. d-dar kbira **MA**

the-house big

‘The house is big’

b. d-dar **ka**:**na** kbira

the-house **was.**$_{3SG}$ big

‘The house was big’

There have been three major analyses proposed to account for these nonverbal predicates: the null copula analysis, the small clauses analysis, and the functional projection analysis. Each of these analyses will be presented and then evidence from Arabic data will demonstrate that Benmamoun (2000)’s functional projection analysis most strongly and consistently accounts for these data on two accounts: (1) the fact that there cannot be a VP projection in nonverbal predicates; and (2) the fact that there must be a functional projection, such as TP, in nonverbal predicates.

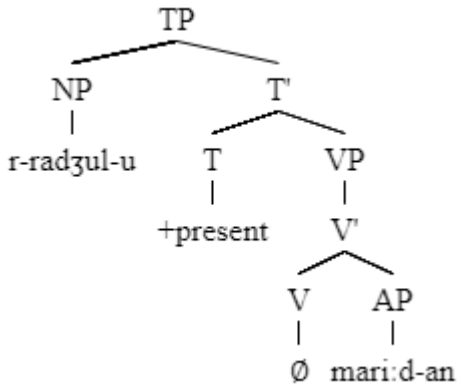
2.3.1 Null copula analysis in Arabic. There are two main null copula analyses for Arabic. The first analysis proposes that the copula is projected and subsequently deleted by a deletion rule (Bakir, 1980). The second analysis proposes that the copula is not deleted, but rather is only spelled out in specific environments (Fehri, 1993). While the details of the two analyses differ in their explanation as to why the copula does not surface except in certain environments, the relevant point for the purposes of this study is that both analyses propose that a copula is projected as the head of a VP. For instance, under both null copula analyses, the argument is that the nonverbal predicate in (13) would still have a full VP projection as illustrated in (14). Since nonverbal predicates still have a full VP projection, nonverbal and verbal predicates are identical in underlying structure, as seen by comparing (14) and (16). However, Benmamoun (2000) presents data that suggests it would be impossible to have a VP projection in nonverbal predicates, which provides strong evidence against the null copula hypothesis.

(13) r-radʒul-u mariːd-un **SA**

the-man-NOM sick-NOM

'The man is sick'

(14)



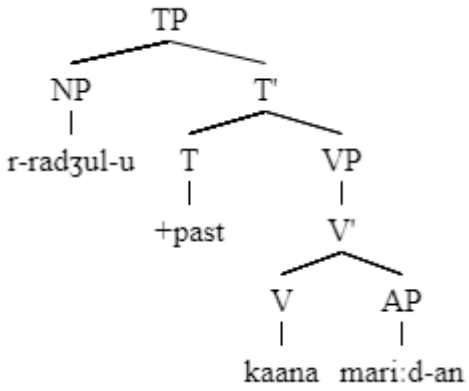
(15)

r-radʒul-u mariːd-an **SA**

the-man-NOM was.$_{3SG}$ sick-ACC

'The man was sick'

(16)



Benmamoun (2000) provides evidence demonstrating that there is no verbal projection in nonverbal predicates. His first piece of evidence comes from case assignment. It can be seen in (13) that the subject and the predicate both receive nominative case (i.e. *r-radʒul-u* 'the man' receives nominative case from T in spec of TP and the predicate *mariːd-un* 'sick' also receives nominative case). The structure in (14) illustrates what example (13) would look like under the null copula hypothesis. In (14), we have a VP projection, and since V assigns accusative, we would expect *mariːd* to be accusative. However, the predicate *mariːd* "sick" receives nominative. The absence of accusative case on *mariːd* suggests that there is no VP projection in the nonverbal predicate. Alternatively, in (15) we see that the subject, "the man" receives nominative case and the predicate "sick" receives accusative case in the presence of the verb *kaːna*. Because both of the accounts of the null copula analysis propose that the underlying

structures of (13) and (15) are identical, the fact that the predicates in (13) and (15) receive different case challenges the validity of the null copula analysis as it is unable to account for these case differences (i.e. the fact that the AP appears with accusative case in one instance and not in the other). Meanwhile, the proposal by Benmamoun, which is that nonverbal predicates in Arabic do not contain a VP, would be able to account for this difference in case based on the presence or absence of a VP projection.

The second point that Benmamoun (2000) uses to demonstrate the absence of a VP is by appeal to principles of "minimality". In many varieties of Arabic, bipartite negation is used to negate verbs. This means that the verb is circumfixed by two negation particles. In the Egyptian example below, the verb *štarret* "bought" can be seen circumfixed by the two negation particles *ma-* and *-š*.

(17)

ma- štarret- š men r-ragel dah **Egyptian Arabic (EA)**

NEG-bought.1SG- NEG from the-man this

'I did not buy from this man'

What is of particular note with bipartite negation is the fact that in Moroccan Arabic, not only can verbs raise to merge with the bipartite negation as a circumfix, but adjectives can also raise and merge with negation if the predicate is nonverbal as seen in (18). This indicates that the adjective is able to raise to merge with negation as seen in (19).

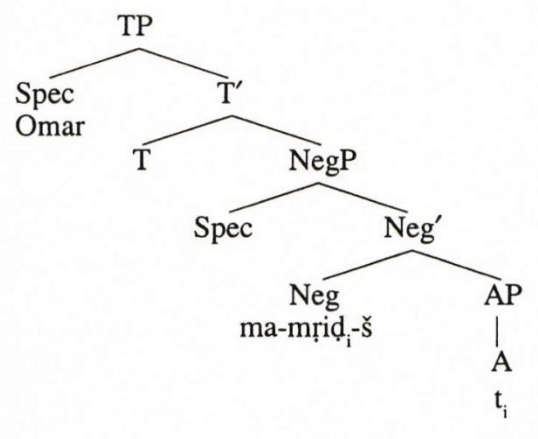
(18)

Omar ma-mrid-š **MA**

Omar NEG-ill- NEG

'Omar is not ill'

(19)



However, the same raise/merge cannot happen if the predicate is verbal. In (20), we have the same sentence as above in (18), but it is now in the past tense rather than in the present tense. This means that it now contains the copula and has become a verbal predicate. With the addition of the copula in (20), the adjective can no longer raise and merge with negation, as evidenced by the ungrammaticality of (20). Benmamoun argues that the VP, when projected and filled by a V (i.e. the copula in (20)) intervenes between the adjective and negation, and thus raising cannot occur without violating the minimal link condition. Therefore, in order to negate the sentence in (18), it is the verb *kana* that raises and merges with negation, as seen in (22). The structure in (22) shows that *kana* can move from V, to Neg, and finally to T, leaving traces in V and Neg.

(20)

*Omar ma-kan-š **MA**

*Omar $_{NEG}$-ill-$_{NEG}$ was.$_{3.SG}$

'Omar was not ill'

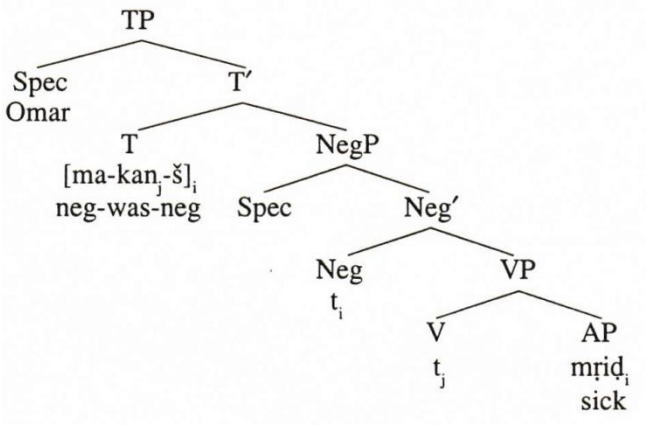
(21)

Omar ma-kan-š mriḍ **MA**

Omar NEG-was.3.SG-NEG ill

'Omar was not ill'

(22)

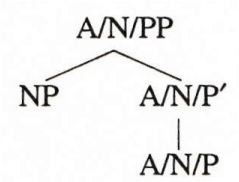


Benmamoun uses the fact that adjectives can raise and merge with negation in sentences with nonverbal predicates, but not with verbal predicates, to support his argument that there is a VP projection in verbal predicates, but not in nonverbal predicates. In turn, Benmamoun argues that the evidence from both case assignment, specifically that predicates in present tense receive nominative case and predicates in past tense receive accusative case, and principles of minimality, specifically that adjectives can raise and merge with negation in present tense but not in past tense, demonstrates that there is no VP projection in present tense nonverbal predicates.

2.3.2 Small clause analysis in Arabic. As an alternative to the null copula analysis, some linguists proposed that the nonverbal predicates found in both Arabic and Hebrew are actually small clauses, meaning that the nonverbal predicates do not contain either a VP or TP projection. According to Hazout (2010), Mouchaweh (1986) was the first to propose nonverbal predicates as root small clauses. Root small clauses have a semantic subject and predicate, which is why they are called clauses at all, but they lack a

specified TP projection, which is necessary to make it a full clause. In English, there are particular verbs that license small clauses such as *consider* and *want.* For instance, in the sentence "I want you well-rested for the exam", the verb *want* licenses the small clause "you well-rested…". Not only does the small clause function as the predicate of the sentence, but within the small clause there is a semantic subject and predicate (i.e. "you" and "well-rested…"). (23) is presented to illustrate that a small clause consists of an NP plus either an AP, NP, or PP, with no intervening TP projection. It is important to note that even within English, there is dispute as to what qualifies as a small clause.

(23)



Hebrew linguists such as Rapaport and Rothstein adopted and extended this analysis for Hebrew. Adopting a small clause analysis for Arabic nonverbal predicates would not meet the same criticisms as the null copula analysis explained above. Since the criticisms of the null copula analysis rely on evidence suggesting there is no VP projection in nonverbal predicates, a small clause analysis would not be subject to this criticism since there are no VP projections in small clauses. However, Benmamoun (2000) strongly argued against a small clause analysis for Arabic, and importantly, many of his criticisms hold true for Hebrew as well. The first two criticisms of a small clause analysis presented below hold true for Arabic as well as Hebrew and two additional criticisms from Arabic-specific data follow.

First, embedded nonverbal clauses have an independent temporal reference. In (24a), despite the fact that "say" is in the past tense, the interpretation is that Omar is

now, presently, in the house. In order to say "He said Omar was in the house", you would need an overt embedded past tense copula (25). The corresponding Hebrew examples are presented in (24b) and (25a) and demonstrate that this is also the case in Hebrew. Therefore, Benmamoun concludes that if embedded nonverbal sentences can have an independent temporal interpretation, they must have their own TP projection.

(24)

a. qal belli Omar f-d-dar **MA**

said.$_{3.SG}$ that Omar in-the-house

'He said that Omar is in the house'

b. amarti še Izzy babajit **Hebrew**

said.$_{1.SG}$ that Izzy in.the.house

'I said that Izzy is in the house'

(25)

a. qal belli Omar kan f-d-dar **MA**

said.$_{3.SG}$ that Omar was.$_{3.SG}$ in-the-house

'He said that Omar was in the house'

b. amarti še Izzy hajah babajit **Hebrew**

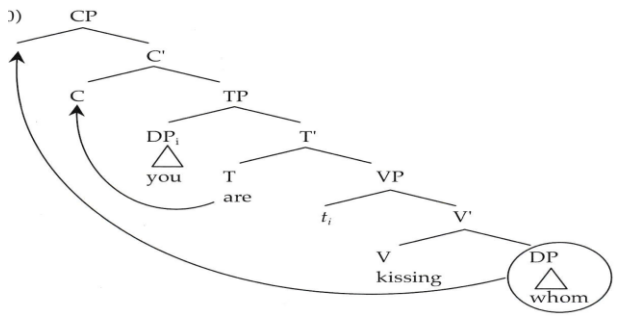
said.$_{1.SG}$ that Izzy was.$_{3.SG}$ in.the.house

'I said that Izzy was in the house'

Second, it is additionally true for Arabic and Hebrew that both the subject and the object of a nonverbal sentence can undergo *wh*- movement. In *wh*- movement, the *wh*-word is generated in either the subject or object position and then moves to Spec CP. In (26) below, the structure of an English example, "Whom are you kissing", is presented.

"Whom" is generated as the object of V, and then it undergoes movement to Spec CP, resulting in the standard English word order of *wh*- questions. In order to have *wh*-movement, there must be a CP in the same clause for the subject or object to move to. In a small clause, there is a semantic subject and predicate, but there is no VP, TP, or CP. Therefore, it is informative that in Arabic, *wh*- questions can be formed from nonverbal sentences. In order for this to be the case, there must be a CP layer in these nonverbal sentences, as the subject/predicate needs to be able to raise to Spec CP to form the *wh*-question. For instance, the Arabic example in (27a) contains a nonverbal predicate and an object *wh*- question. The *wh*- question word *fin* or "where" has moved from its position as the object of V to the Spec CP in order to get the *wh*-question word order of "Where Omar?". Additionally, we see this with an Arabic subject *wh*-question in (27b). Benmamoun uses the fact that sentences with nonverbal predicates can undergo *wh*-movement in order to support the fact that nonverbal sentences must be a full CP clause, rather than a small clause. If there were no VP, TP, or CP, as suggested by the small clause analysis, then we would not expect to be able to get the word order "Where Omar?" or "Who in the house" because there would be no position for the subject or object *wh*-word to raise to. The corresponding Hebrew examples in (27c) and (27d) demonstrate the same phenomena.

(26)



(27)

a. fin Omar **MA**

Where Omar

'Where is Omar?'

b. škun f-d-dar **MA**

who in-the-house

'Who is in the house?'

c. ajfo Izzy? **Hebrew**

Where Izzy

'Where is Izzy?'

d. mi babajit? **Hebrew**

Who in.the.house

'Who is in the house?'

The third criticism, which can only be supported with Arabic data since Hebrew no longer realizes case, relates to the fact that the subject of a nonverbal predicate is assigned nominative case. In an English small clause, the subject cannot be nominative. For instance, in English, we say "I wish **him** well" or "I consider **him** rude", but never "I wish **he** well" or "I consider **he** rude". The fact that the subject of a small clause cannot take nominative case is viewed as evidence that there is no T projection in the small clause (Basilico, 2003). In Arabic, the subject of a nonverbal predicate must be nominative. In order to have nominative case assigned, the presence of a T to assign or check nominative case is assumed to be required. Example (13) from above is presented

again below in order to show that in Standard Arabic, the only form of Arabic which still extensively utilizes case marking, the subject of the nonverbal predicate, and its predicate, are both marked for nominative case. Benmamoun argues that this suggests the presence of a T projection, consequently discrediting the theory that the nonverbal sentences are small clauses.

(13) r-radʒul-u mariːd-un **SA**

the-man-NOM sick-NOM

The man is sick

The fourth criticism is related to the distribution of negative polarity items (NPI). Broadly speaking, NPIs are items whose use is restricted to negative contexts. For instance, in English, the word *yet* is an NPI. There must be a negation above *yet* in order to license the use of *yet.* For instance, you cannot say "*I have written a book yet" because there is no negation in the sentence to license the use of *yet.* Naturally, the correct sentence would be "I have not written a book yet". Languages vary in terms of which contexts license NPIs, and there are some distinct differences between Arabic and English. For instance, in English, an NPI can be licensed by a negative in the matrix clause even if the NPI is contained within an embedded finite clause, such as in the sentence "He does not think I have *any* money". In contrast, in Moroccan Arabic, a negative contained within the matrix clause cannot license an NPI that is contained within an embedded finite clause. This is demonstrated in ungrammatical Moroccan Arabic example presented in (28a) where the NPI *hetta wahed* "anyone" is contained within the embedded finite clause "that Nadia met anyone" and the negation is in the matrix clause "He does not think…". Example (28a) is ungrammatical because the NPI

hetta wahed cannot be licensed by the negation in the matrix clause. According to Benmamoun, the presence of a finite T in the embedded clause blocks the licensing of the NPI by the negation in the matrix clause.

On the other hand, similarly to English, an NPI in Moroccan Arabic is licensed by a negation when it is contained within the same clause. In English, the sentence "I have not written a book" is an example of this type. Example (28b) illustrates the NPI *hetta wahed* "anyone" being licensed by negation when it is contained within the same clause as the negation (i.e. the matrix clause). The argument is that there is no additional finite T (i.e. there is only the matrix clause T) to block the licensing of the NPI. To further illustrate this point, Benmamoun gives the example in (28c) which shows that the NPI *hetta wahed* can be licensed by negation when it is contained within a nonfinite clause. Although the negation and the NPI are contained in different clauses, the embedded clause does not contain a finite T, but rather a nonfinite T, so the licensing of the NPI is not blocked. These examples together demonstrate that in Moroccan Arabic, an NPI in an embedded clause can be licensed by a negation in the matrix clause as long as there is no finite T in the embedded clause.

(28)

a. *ma-ta-j-den belli Nadia tlaq-at hetta wahed **MA**

* NEG- ASP-3.SG-think that Nadia met-3.SG any one

('He does not think that Nadia met anyone')

b. Nadia ma mʕa hetta wahed **MA**

Nadia NEG with any one

'Nadia is not with anyone'

c. ma-bɣa-h j-tlaqa hetta wahed **MA**

NEG-wanted.3.SG-him meet.INF any one

'He does not want him to meet anyone'

The interesting point to make is that an NPI contained within an embedded nonverbal clause cannot be licensed by negation in the matrix clause (29). According to the examples above, an NPI in an embedded clause can be licensed by negation in the matrix clause as long as there is no finite T in the embedded clause. Since an NPI contained in an embedded nonverbal clause is not licensed by negation in the matrix clause, this suggests not only the presence of a separate T in the embedded clause, but also that this T is finite. Under the small clause analysis, the example in (29) would have a structure where the embedded nonverbal predicate does not have its own TP, and thus, the small clause analysis would incorrectly predict (29) to be grammatical.

(29)

*ma-ta-j-den belli Nadia mʕa hetta wahed **MA**

* NEG-ASP-3.SG-think that Nadia with any one

('He does not think that Nadia is with any one')

To sum up, Benmamoun (2000) presents evidence supporting the fact that, contra the null copula analysis, there is no VP projection in nonverbal predicates in Arabic. Benmamoun first points out that adjectives in nonverbal predicates receive nominative case, whereas adjectives in verbal sentences receive accusative case. If there were a VP projection in the nonverbal predicate, as suggested by the null copula analysis, then we would expect the adjective in a nonverbal predicate to receive accusative case. Additionally, Benmamoun points out that in Moroccan Arabic, verbs can raise and merge

with negation in nonverbal predicates, but not in verbal predicates. Benmamoun attributes this to the absence of a VP projection in nonverbal predicates, as the presence of an intervening VP projection would block the verb from raising and merging with negation. Benmamoun (2000) also presents evidence supporting the fact that, contra the small clause analysis, there is a TP projection in nonverbal predicates in Arabic. Benmamoun first demonstrates that embedded nonverbal clauses have an independent temporal reference, which he argues could only be the case if there is a TP in the embedded nonverbal clause. In addition, Benmamoun argues that since nonverbal predicates can undergo *wh*-movement, then they must be a part of a clause containing a TP and CP layer, rather than just a small clause. Moreover, Benmamoun points to the fact that subjects in nonverbal sentences are assigned nominative case, which would only be expected to occur in the presence of a T projection. Lastly, Benmamoun points out that in Moroccan Arabic, an NPI cannot be licensed by a negative in the matrix clause if the NPI is contained within an embedded finite clause. Similarly, an NPI contained within an embedded nonverbal clause cannot be licensed by negation in the matrix clause. Benmamoun uses this similarity to suggest that the embedded nonverbal clause must have its own TP in order to disallow the licensing of an NPI in the matrix clause. In concluding that nonverbal predicates in Arabic have a TP, but do not have a VP, Benamoun proposed what this paper will refer to as the "functional projection analysis".

2.3.3 Functional projection analysis in Arabic. According to Benmamoun (2000), nonverbal predicates in Arabic contain a functional projection that dominates a nonverbal predicate. This analysis provides solutions for all of the criticisms of the small clause analysis discussed above. When the evidence that nonverbal predicates must have

a TP is combined with the criticisms of the null copula analysis (case and minimality), which disprove the presence of a VP, this evidence together provides support for Benmamoun's proposal that nonverbal predicates have a TP but no VP.

The final point that Benmamoun (2000) addresses, which is relevant for the purposes of this study, is *why* there is no VP in nonverbal sentences. He argues that it is not universal that all languages have a [+D] and [+V] feature in every tense/mood. An example of this in English would be the imperative, which does not require an overt subject because the TP in this structure does not have a [+D] feature that needs checking, in contrast with other tenses and moods which do. He goes on to argue that the Arabic TP, when marked for present tense, contains a [+D] feature only (for purposes of EPP), whereas any TP marked for past and future tense have both [+D, +V] features that need checking by an overt subject and overt verb, respectively. Therefore, tense does not need to be licensed by a VP because it does not contain a [+V] feature which needs checking. In contrast, all tenses in English contain a [+V], and consequently TP must always be licensed by a VP in English. However, before introducing the evidence that Benmamoun (2000) and Aoun, Benmamoun, and Choueiri (2010) provide for their analysis of the V feature in Arabic, it is worthwhile to introduce the [D] and [V] features a bit more formally, as well as look at some of their explanatory power outside of the scope of this paper.

Chomsky (1995) first proposed that T is specified for the categorical features of [D] and [V]. In proposing a categorization of languages into either D or V-prominent, Davies and Dubinsky (2001) were able to provide an account for the "unified behavior" of subjects in certain types of languages which inevitably were not shared in other types

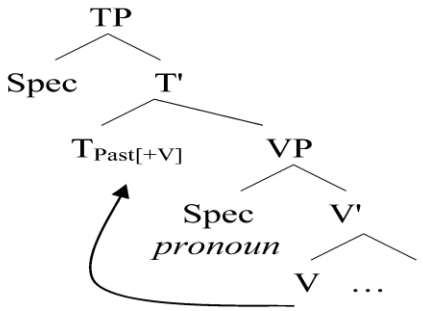
of languages. Rather than trying to argue for some unified semantic notion of subjects, a categorization of languages into D or V-prominent languages is able to account for the fact that there seems to be a cluster of subject properties present in one type of language, yet completely absent in another type. For instance, the [D] feature accounts for the dependency between tense and subject, triggering subject movement to TP in languages such as English and French, whereas the [V] feature accounts for the dependency between tense and the verb in languages such as Bulgarian and Irish, which triggers V to T movement.

In building upon this framework, Benmamoun (2000) and later Aoun, Benmamoun, and Choueriri (2010) argue that Arabic is not simply a D or V-prominent language, but rather that the past tense in Arabic is V-prominent and the present tense in Arabic is D-prominent. Benmamoun (2000) cites evidence from negation, word order, and agreement and morphological realization across the Moroccan and Egyptian dialects in addition to Modern Standard Arabic, in order to support his featural distinction between the present and past tenses. His evidence is described below and demonstrates how the featural distinction in tense is capable of accounting for many tense asymmetries in Arabic.

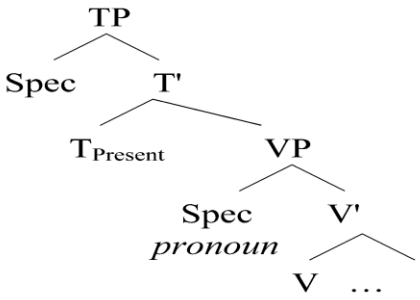
Benmamoun (2000) and Aoun, Benmamoun, and Choueiri (2010) argue that the [+V] feature in past tense triggers verb raising as illustrated in (30a) (Sultan, 2011) whereas verb raising is not obligatory in the present tense since there is no [+V] that needs checking as seen in (30b) (Sultan, 2011).

(30)

a.



b.



This obligatory verb raising in the past tense, but not in the present tense, is exemplified in negation in Egyptian Arabic. In the past tense, the verb must obligatorily raise and merge with negation as evidenced by the grammaticality of (31), where the verb *katab* has merged with the negative particles *mi-sh* and the ungrammaticality of (32) where the negative particles *mi-š* precede the verb *katab*. The grammatical structure of past tense negation in Egyptian Arabic can be seen in (33).

(31) Omar ma-katəb-š ig-gawɑːb **EA**

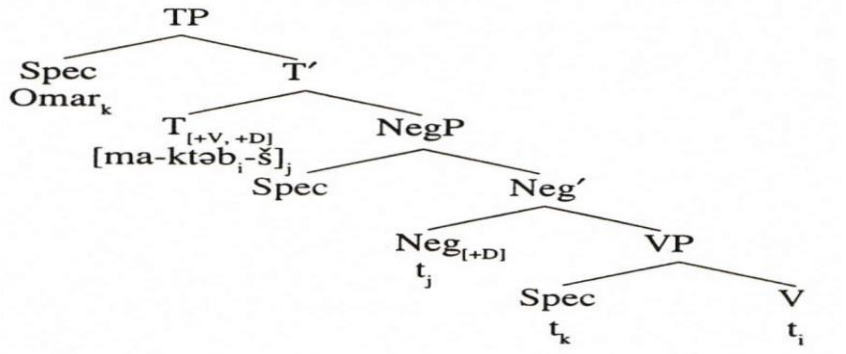
Omar NEG-wrote.3.SG-NEG the-letter

'Omar did not write the letter'

(32) *Omar mi-š katəb ig-gawɑːb **EA**

*Omar NEG-NEG wrote.3.SG the-letter

(33)



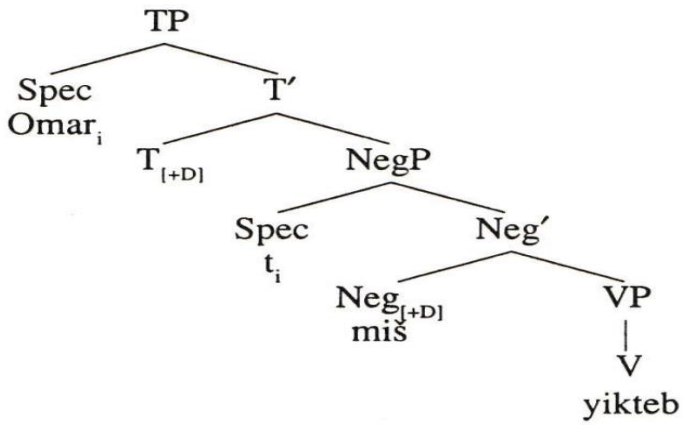
In contrast, the verb in present tense negation does not merge with negation as evidenced by (34), where the verb *jikteb* follows the negative particles *mi-sh*. The structure of the present tense negation in Egyptian Arabic, as shown in (35), would be ungrammatical in the past tense. Benmamoun argues that this is strong evidence for the presence of a [+V] feature in the past and future tense which is responsible for triggering obligatory verb movement to T, but this movement is not obligatory in the present tense.

(34) Omar mi-š jikteb ig-gawɑ:b **EA**

Omar NEG-NEG write.3.SG the-letter

Omar is not writing the letter

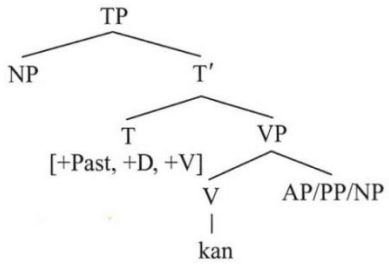
(35)



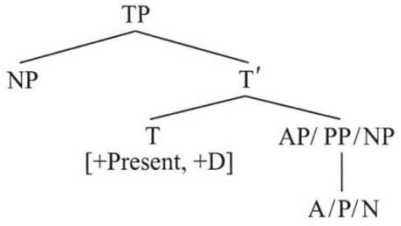
Assuming Benmamoun (2000)'s account of nonverbal predicates in Arabic, the English and present-tense Arabic structures of the copula are distinct. It is of particular importance to note that this is the case not just when the copula is being used as a main verb, but also when the copula is being used as an auxiliary verb in progressive aspect. Present tense progressive clauses in Arabic do not contain the copula, but past tense progressive clauses do. The English copular structure, regardless of tense, is identical to the Arabic past tense copular structure shown in (36), where the presence of a [+V]

requires that TP always be licensed by a VP, which results in the overt realization of the copula in all tenses. On the other hand, the present tense Arabic nonverbal predicate can be seen in (37) where the lack of a [+V] feature allows for the licensing of TP without a VP, resulting in the lack of a copula in the present tense. In the past and future tense, the Arabic structure would look like (36), which would also be the underlying structure for all English tenses.

(36)



(37)



Thus, when investigating the acquisition of the English copula by Arabic speakers, it would be more accurate to say that it is the acquisition of the English [+V] feature in a [+present] TP that is specifically of interest. Following this featural difference between L1 Arabic and L2 English, the results of this study may be indicative as to whether the source of copula omission by Arabic speakers learning English is performance or representational in nature. This, in turn, would have implications for the broader performance vs. representational deficit debate.

2.4 Research on English Copula Acquisition by Arabic Speakers

There have been a handful of studies which directly investigated the acquisition of the English copula by Arabic speakers. It is important to highlight that in all of these studies, acquisition was investigated only via written production. Al-Zahrani (1993) and Al-Shayban (2007) found a statistically significant difference between omissions of

copula in present tense versus the other tenses, but Kasem (2000) did not investigate copula omission across tense.

Al-Zahrani (1993) conducted a study with 36 native speakers of Arabic studying English in an intensive English program at a large US public university. The participants were all at the intermediate or advanced level. Written data were collected from 36 participants and results indicated that participants tended to omit the copula in the present tense (27%) more than in other tenses (<1% in past, 0% in future). He concludes that the omission of the copula in the present tense is evidence of negative L1 transfer, and that the retention of the copula in the past and the future tense is positive L1 transfer.

Al-Shayban (2007) conducted a study with 100 male Saudi EFL learners studying in an intensive English program in Saudi Arabia. Half of the participants were at the intermediate level and half were at the advanced level. One writing sample was collected from each participant. Results indicated, similarly to Al-Zahrani (1993), that participants were most likely to omit the copula in the present tense and that the number of copula omission errors reduces as proficiency level increases. She also concludes that her results were indicative of L1 transfer.

Kasem (2000) investigated the acquisition of the English copula by native speakers of Lebanese Arabic. A 6-month longitudinal study was conducted with 10 Arabic EFL learners, ages 13-17, who had arrived in Australia less than 6 months before the start of the study. He describes the participants as having a "very moderate knowledge of English" and as being "highly motivated" with "positive attitudes". Written data were collected every two months for six months in order to observe a continuum of development. The omission errors in Kasem (2000) were categorized into

three types of omission: existential "there", AP predicates, and NP predicates. In the first stage, the majority of errors (87) were found with the existential "there" structure, followed by AP predicates (62) and NP predicates (59). In stage 2, the participants stopped omitting the copula altogether, but made errors in agreement with the copula 41.6 percent of the time. This figure dropped to 20.1 percent in stage 3. With these results, Kasem (2000) argues strongly that these types of errors are developmental and are not an indication of L1 transfer. It is disappointing, however, that Kasem (2000) did not discuss the role of tense in any of his data. Therefore, we cannot determine if there was a higher tendency to omit the copula in the present tense than in the past and future tenses.

To sum up, research has established that Arabic-speaking English language learners tend to omit the copula, specifically in written production. Furthermore, research from Al-Shayban (2007) and Al-Zahrani (1993) indicates that participants are more likely to omit the copula in the present tense than in any other tense, which they suggest is a result of L1 transfer. However, previous research has not looked at the rate of copula omission where the copula is an auxiliary rather than a main verb. Furthermore, previous research has only looked at copula omission in writing. Therefore, this study will not only investigate the role of tense in copula omission, but also the role of copular function and number, and it will do so via a test of competence (Grammaticality Judgment Task) and a test of spoken quasi-production (Elicited Imitation).

2.5 Summary and Research Question

The aim of this study is to investigate whether Arabic-speaking English language learners at intermediate and advanced levels of proficiency have target-like syntactic representation of the copula (via UG access), but are simply failing to produce them accurately due to a "performance deficit", or alternatively, is it the case that participants do not have the target-like underlying representation of the copula and consequently have a "representational deficit"? Therefore, the research question of this study asks which theory, Feature Reassembly or Interpretability Hypothesis, best accounts for the results of this study as they relate to the acquisition of English copula by native speakers of Arabic.

Under the Interpretability Hypothesis, interpretable features, defined in this study as head features, are accessible and fully acquirable. Alternatively, uninterpretable features, defined in this study as features which are not inherent to the head, but rather are relational to the head features, are uninterpretable. Uninterpretable features resist parameter resetting even at the most advanced levels of proficiency, resulting in a deficit in the interlanguage syntax of a second language learner. Under Feature Reassembly, errors in a certain feature are attributable to a failure to learn how the feature is encoded in the new language (i.e. the morphology), not a deficit in the L2 learner's underlying representation of the feature. Slabakova (2009) used Lardiere (2008)'s prediction about "mismatched" features in the L1 and L2 to develop a scale of difficulty in the acquisition of L2 features. L2 learners whose L1 encodes a feature contextually would face the greatest challenge in acquiring a language in which a feature is encoded grammatically. Alternatively, moving from an L1 where a feature is encoded grammatically in a different way than how it is encoded grammatically in the L2 would prove slightly less

challenging. Lastly, it is predicted that if both languages grammatically encode a feature in the same way (i.e. no reassembly is required), then this should be the least arduous acquisition. Therefore, in light of the constraints proposed by Slabakova (2009), Feature Reassembly is able to make predictions as to when a divergence between syntactic competence and morphological performance will occur and as to how difficult this divergence will be to overcome.

For the purposes of this study, the Interpretability Hypothesis would predict more accurate production of tense than of number agreement on V due to the fact that tense is an interpretable feature on V and number agreement is not. In addition, since the deficit is predicted to be permanent, we would not expect to see the more proficient group in this study, the High Group, overcome the deficit, meaning they would be expected to continue making the same errors. In contrast, Feature Reassembly, as constrained by Slabakova (2009), would generally predict that participants would perform better on past tense than on present tense, on main verb items than on auxiliary verb items, and on singular items than on plural items. In terms of tense, the present tense copula in Arabic is not realized at all and is therefore encoded contextually, meaning context, rather than overt syntax or morphology, is used to distinguish between simple present tense and identical surface forms such as present progressive. Alternatively, in English, the present tense copula is realized grammatically. Therefore, present tense copula would be expected to be more difficult to acquire than past tense copula, which requires only the learning of a new grammatical encoding. Relatedly, the present progressive in Arabic is also encoded contextually, as in Standard Arabic, the present simple and present progressive are identical and distinguished by context. Alternatively, in English, the

present progressive requires the overt realization of the copula (unlike present progressive in Arabic), and it also requires the addition of the suffix *-ing* on the main verb. This would be a change from contextual to grammatical encoding, plus the added step of acquiring the *-ing* suffix, which would again be predicted to be very difficult. The past progressive, on the other hand, is formed in Arabic with the copula plus the imperfective form of the verb, so the only difference between the Arabic and English past progressive is the English addition of the *-ing* suffix. This would predict that the acquisition of the past progressive should only be moderately difficult when compared to simple past, and specifically it would be expected that the participants would omit the *-ing.* Lastly, in terms of number, plural in Arabic is realized on the suffix in the SVO word order, but it is not realized at all in VSO word order. Therefore, it would be expected that plural would be only moderately difficult to acquire in the SVO word order, but it would be more difficult to acquire in the VSO order (i.e. subject-auxiliary inversion).

Research question: Does the Interpretability Hypothesis (Tsimpli & Mastropavlou, 2007) or Feature Reassembly (Lardiere, 2008) best account for the data in this study?

CHAPTER 3

PILOT STUDY

3.1 Research Design

In order to answer the research questions above, the participants of this study completed both an imitation task and a comprehension task. The use of both a quasi-production task and a comprehension task allows for the comparison of the participants' performance and competence, which has not been done in previous studies. Furthermore, previous research looked exclusively at written production, and consequently, were limited to analyzing only the errors that occurred in the free production. Since this study wanted to look at copula omission across a variety of parameters, free production was not a feasible option, as it would be extremely unlikely to get enough occurrences of copula omission for statistical analysis. For this reason, an elicited imitation task (EI) was chosen in order to get an idea of the types of errors that may occur in free production, although it is expected that these errors would be made at a lower rate due to the nature of the task. According to Lee (2003), errors in repetition signal "underdeveloped psycholinguistic mechanisms", so if a learner makes an error in repetition, it is possible that they have not fully acquired that particular feature. The time-pressure GJT was designed to measure the participants' competence of the English copula. According to Ellis, et al. (2015), "timed and untimed GJTs measure different constructs" and this study proposes that a time-pressure GJT is distinct from an untimed GJT in that it gives insight into the true competence of a learner's language which may or may not be dissimilar

from a learner's performance on an untimed GJT in which they can resort to the use of rule application and more cognitive problem-solving skills. Gutierrez (2012) analyzed the grammatical and ungrammatical portions of the GJT separately, concluding that the grammatical section of a time-pressure GJT constitutes a measure of implicit knowledge whereas the ungrammatical section of a time-pressure GJT is a measure of explicit knowledge. Therefore, this study adopts a time-pressure GJT and also analyzes the grammatical and ungrammatical sections of the GJT separately since research suggests they may measure different types of knowledge. With these two tasks, the effects of tense, verb type, number, and syntactic environment can be investigated both within tasks and across. For instance, we can compare whether or not a participant performed better on present tense items than past tense items on the GJT and whether or not this held true for the EI as well. Environments where participants' performance deviates greatly either within or across tasks is informative of the type of knowledge they have of a given structure.

3.2 Pilot study

The pilot study was conducted in the summer of 2015 in an intensive English program at a private college with six Arabic-speaking participants enrolled in a low-intermediate-level course. The pilot study consisted of an untimed GJT with 30 target items and 30 fillers. Of the 30 target items, 15 were grammatical (i.e. they included the copula) and 15 were ungrammatical (i.e. the copula was omitted). There were five syntactic environments: Simple sentences, complex sentences, negation of simple sentences, subject-auxiliary inversion questions and *wh*- questions. For each syntactic environment, there was one item in the past, present, and future tense. The pilot study

also included an elicited imitation task of the same 30 target items with 10 fillers, which was completed before the GJT task.

3.3 Pilot study results

Since the number of participants in the pilot study was quite small, results from any sort of statistical analysis would not be significant. Therefore, in order to get the best picture of any emerging patterns in the data, the participants' results were combined and are organized and presented by tense. In the past tense, it appears that participants did negligibly better on the GJT than on the EI. It is interesting to note that for both simple sentences and negation, there were participants who performed completely accurately on the GJT items, but then omitted the copula in the EI. This may be suggestive of discrepancy between the participants' performance and competence. However, the results of the past tense *wh*- question item seem to suggest that the participants of this study have simply not acquired copula in *wh*-questions yet. This can be concluded since the same four participants all rejected the grammatical item, accepted the ungrammatical item, and then failed to produce the copula. This could alternatively suggest that at this level, some of the participants have yet to acquire the English CP, and that this consequently is interfering with their ability to accurately perceive or produce the English copula in *wh*- questions.

Table 3.1

Past tense scores by task (Pilot)

Past	**EI**	**Accept Ungrammatical**	**Reject Grammatical**
Simple	4/6	3/6	6/6
Embedded	6/6	6/6	6/6
Negation	1/6	6/6	3/6
***Wh*- questions**	2/6	0/6	2/6

S-Aux Inversion	6/6	6/6	6/6
Total out of 30:	19/30	21/30	23/30

**Number of correct responses out of total responses*

In the present tense, there were actually significantly more errors on the GJT than on the EI. The data in Table 3.2 suggests that 3 students who accepted the ungrammatical present-tense *wh*- question item proceeded to accurately produce the copula on the EI. Moreover, four students who committed errors on the GJT on present-tense negation items also went on to accurately produce the copula on the EI. These results could simply suggest that the participants' have acquired the ability to accurately produce the copula in the present tense. However, it is worth pointing out that all six participants contracted the present-tense negation item in their production to "They're". Participants were also prone to using the same contraction incorrectly in the past tense. This was also the case for the simple sentences. In some cases, both "She is very tired today" and "She was very tired yesterday" were produced as "She's very tired today/yesterday". Therefore, it is possible that the participants have internalized common contractions such as "she's", "I'm", "you're" as a chunk. For this reason, it seemed worthwhile to change the pronouns in the target items to nouns and proper nouns as a way to mitigate against this possible variable.

Table 3.2

Present tense scores by task (Pilot)

Present	**EI**	**Accepted Ungrammatical**	**Rejected Grammatical**
Simple	6/6	6/6	6/6
Embedded	6/6	6/6	6/6
Negation	6/6	4/6	4/6
***Wh*- questions**	3/6	0/6	3/6
S-Aux Inversion	6/6	6/6	6/6
Total out of 30:	27/30	22/30	25/30

**Number of correct responses out of total responses*

Overall, the future tense was problematic as seen in Table 3.3. Despite 3 students making errors in the simple future on the GJT, they all produced the copula on the EI. Despite 4 students making errors in the embedded future on the GJT, again they all produced the copula on the EI. While it makes sense that the two students who accepted future negation without the copula also produced the future negation without the copula, it is interesting to point out that there were still 2 other students who rejected the grammatical future negation (i.e. the presence of the copula) but produced the negation in the EI. In addition, while it also makes sense that the student who rejected the grammatical future *wh*-question would not produce the copula on the EI, it is once again interesting to note that there were four students who made errors on the future *wh*-question GJT but went on to produce the copula in the EI. Lastly, 5 students made errors on the future subject-auxiliary inversion GJT items, but they all produced the copula in the EI. Essentially, the participants performed much more poorly on the GJT in future tense than the past or present tense and significantly better on the EI in the future tense than on the past or present tense.

Table 3.3

Future tense scores by task (Pilot)

Future	**EI**	**Accepted Ungrammatical**	**Rejected Grammatical**
Simple	6/6	3/6	4/6
Embedded	6/6	3/6	5/6
Negation	4/6	4/6	4/6
***Wh*- questions**	5/6	4/6	3/6
S-Aux Inversion	6/6	3/6	4/6
Total out of 30:	27/30	17/30	19/20

**Number of correct responses out of total responses*

To conclude, it appears that the participants performed inconsistently at best on the future tense. It very well may be the case that the participants have fully acquired the copula in English in the future tense, but that the word order or presence of a modal is what students have yet to acquire, and this is what is causing them to reject grammatical future tense items and/or accept ungrammatical future tense items. There is some evidence for this interpretation; despite producing the copula in the future tense the overwhelming majority of the time, the participants produced the incorrect word order 7 times, deleted the modal 3 times, and failed to produce a future-tense item at all 3 times. Therefore, since the presence of the modal "will" and the use of the infinitive form of the copula in the future tense introduced (at least) two additional variables potentially influencing participants' performance on the tasks, the future tense was eliminated from the main experiments.

Overall, the participants omitted the copula in their production on the EI task a tenth of the time for the present and future tense and a third of the time in the past tense. The question then becomes: Why is there such a large body of research suggesting that Arabic speakers learning English regularly omit the copula? One possibility is that despite the justification above for using EI as a measurement of production, perhaps EI's are just not an astute enough measurement of production. For this reason, the current study does not make any claim that the rate of copula omission would be similar in an EI and in free production. However, the EI is still effective for giving insight into the environments and conditions in which a participant may omit the copula. Secondly, it is possible that Arabic speakers tend to omit the copula not when it is the main verb in the sentence, but rather when it is used as an auxiliary verb, due to the fact that Arabic does

not use the copula as an auxiliary verb as discussed above. For this reason, progressive items were added to the items of this study.

3.3.1 Changes to instruments. The pilot study consisted of an untimed GJT and an EI. Due to a growing body of research on the inability of untimed GJTs to tap into implicit knowledge (Ellis, et al., 2015), the present study utilized a time-pressure GJT instead of an untimed GJT. This study argues that this is more representative of the participants' competence.

3.3.2 Changes to items. The following changes were made based on the data collected in the pilot study. The items for the pilot study contained only items where the copula was being used as the main verb. The present study adds an equal number of target structures where the copula is being used as an auxiliary verb in order to have a better representation of the environments in which participants are most likely to omit the copula. Furthermore, the items for the pilot study contained items in past, present and future tense. However, participant performance in future tense items seemed to indicate that the presence of the modal "will" and/or the infinitive form of the copula could confound results. Therefore, the items in the present study consist of only past and present tense, which will be sufficient to investigate whether or not tense plays a role in the omission of the copula by Arabic speakers. The items also include an equal number of singular and plural items to allow for more insight into number agreement errors, an important point of divergence in predictions between IH and FR. Recall that IH would predict more errors with number agreement than with tense whereas FR would predict more errors with plural items than singular. The simple sentence items were eliminated in favor of more items in the more complex syntactic environments of embedded clauses,

negation, subject-auxiliary inversion, and *wh*-questions. Lastly, results of the pilot EI demonstrated that the participants were very likely to add a contracted copula (e.g. she's; you're) regardless of tense and regardless of whether the target item actually included the copula or not. For instance, in the sentence "He does not like to eat pizza", a number of the participants responded with "He's not like to eat pizza". In order to avoid an overgeneralization of the contraction due to habitual use such as "she's" and "I'm", the pronouns in the target items were changed to nouns and proper nouns. To sum up, based on the results of the pilot study, future tense was eliminated, simple sentences were eliminated, and all subjects were made nouns or proper nouns, rather than pronouns, to avoid use of contractions. See Section 4.2 for a detailed explanation of the items used in the main study and see Appendix A for a list of the actual items used.

CHAPTER 4

MAIN STUDY

4.1 Participants

The participants of this study were 48 native Arabic-speaking English language learners, but 3 participants were eliminated from the study due to the incomprehensibility of their EI audio recording, leaving a total of 45 participants. All 45 of the participants were studying at an intensive English language program at a large public university in the United States, and they participated in the study during their class time. Participants were divided into a "high" (n=24) and "low" (n=21) proficiency group, based on their performance on Slabakova (2000)'s cloze passage. The cloze passage had a gap every seventh word with a total of 40 gaps. The "exact match" method of scoring was used, meaning participants received one point if they provided the exact word from the passage and no points if they provided any other word. There appeared to be a natural "gap" in the scores because 23 participants scored 10 points or higher and 21 participants scored 8 points or lower. Only one participant received a score of 9, and since this participant was in the most advanced class used for this study, this participant was placed in the "high" group. The placement of all participants into either the "high" or "low" group matched their placement in the intensive English program into either "advanced" or "intermediate" classes, lending support to the use of this cloze passage assessment for proficiency assessment. The mean score for the "high" group was 14.4/40 and the mean score for the

"low" group was 5.6/40, as compared to a mean score of 22/40 for native speakers as cited in Slabakova (2000).

While research has documented the close correlation of cloze passage performance and reading comprehension (Gellery & Elbrow, 2013; Tabatabaei & Mirzaei, 2014), not much research has been conducted evaluating the correlation between performance on a cloze passage and spoken proficiency. While it is true that the present study utilizes both a comprehension and quasi-production task, it was not feasible to conduct both a comprehension and production proficiency test with the limited amount of class time offered for the experiment. Thus, Slabakova (2000)'s cloze passage was used as a quick and efficient proficiency measurement, and results were then compared to the class level of the participants in the program. It is worth noting that only one participant in a lower-level class scored into the "high" group for the study and no participant in a higher-level class scored into the "low" group. Therefore, taken together, the class placement and cloze passage provide a sufficient measurement of proficiency for the purposes of this study.

4.2 Materials

4.2.1 Items. After the changes to the target items discussed in Section 3.3.2, the main study ended up having 64 target items for the GJT and 32 target items for the EI.[1] The first condition was tense. Half of the target items were in present tense and half of the items were in past tense. For the GJT, this meant 32 present tense items and 32 past tense items. For the EI, this meant 16 present tense items and 16 past tense items. The

[1] See Section 4.2.2 for a discussion on why only grammatical items were used in the EI, whereas the GJT used both a grammatical and ungrammatical version of each item, resulting in double the amount of items as the EI.

second condition was copula type. In half of the items, the copula was used as a main verb, and in half of the items, the copula was used as an auxiliary verb in a progressive construction. The third condition was number. In half of the items, the subject was singular, and in half of the items, the subject was plural. This resulted in a 2 (tense) by 2 (verb type) by 2 (number) design. This means that in the EI, there were 4 items for each condition. In the GJT, there were 8 items for each condition; the 8 items included the grammatical items which were identical to the EI items plus the corresponding ungrammatical items in which the copula was omitted. This design allowed for a statistical analysis of the results, discussed in more detail in Section 4.5.

The main study used syntactic environment as a type of control for the items. If you use only one type of syntactic environment, there is the risk that there is something about that specific syntactic environment which is confounding results, and results would not be generalizable to other syntactic environments. For instance, if all items were embedded declaratives, the results would only stand true for this specific syntactic environment, and there would be no way to know if something particular to this syntactic environment was confounding results. The goal of this study was to see what patterns of copula omission shined through even across multiple syntactic environments with the hope that this would make the results regarding the copula more generalizable. Therefore, items were equally divided among four syntactic environments: negation of simple sentences, embedded declaratives, *wh*- questions, and subject-auxiliary inversion. This means in the EI, there were 8 items for each syntactic environment, and in the GJT there were 16 items for each syntactic environment, due to the addition of the corresponding ungrammatical items. However, the decision was made to not use

syntactic environment as a condition. If syntactic environment were used as a condition, the number of target items would have needed to be quadrupled in order to be able to conduct a statistical analysis with the same confidence as what is presented in this study. This would be too many items given the procedural constraints presented in Section 4.3, namely the limited class time in which to conduct the experiment and the risk of participant fatigue. Although syntactic environment was not used as a condition in this study, the results of each syntactic environment are presented descriptively in Section 5.1. The hope is that any interesting findings, while not statistically significant, may help inform future research in this area.

4.2.2 Elicited imitation. The EI was designed to identify the environments in which copula omission by an Arabic-speaking English language learner take place. According to Ambridge and Rowland (2013), elicited imitation is a "valuable index of production ability" as speakers extract the meaning of the sentence but reproduce it from scratch using their standard production. However, in order for this to occur, the phonological storage must be disrupted, which they suggest can be done simply by having participants wait in silence before repeating. Therefore, participants in this study were instructed to wait in silence for five seconds until they heard a "beep", and then repeat the sentence. However, it is worth noting that they do caution that the overall error rate may be lower in EI than in free production as there is arguably some retention of the presented sentence. While error rates may be lower in elicited imitation, it is still believed that specific inaccuracies are reflective of differences between the subject's grammar and the target grammar (Bley-Vroman & Chaudron, 1994) and this is the focus of this study.

It has also been repeatedly attested that general content EI's strongly correlate with oral proficiency, and it is consequently argued that EI's tap into the same linguistic system that is used for spontaneous communication (Burdis, 2014). Research has shown strong correlations between OPI ratings and EI scores in English (Graham et al., 2008), in Japanese (Matsuhita, Lonsdale & Dewey, 2010), and French (Millard & Londsdale, 2011) and between ACTFL scores and EI scores in Russian (Burdis, 2014). The correlation was so high in Burdis (2014) that it accounted for 86% of score variance. Ortega and Wu (2013) found a "significant relationship" between EI performance and narrative performance for 80 L2 Chinese learners across proficiency levels including both heritage and non-heritage learners. While it is undeniable that naturalistic data is preferable in looking at production errors, the benefits of the EI in allowing the targeted investigation of such a wide variety of environments outweighs the disadvantage of not being able to accurately measure rate of copula omission in production.

The EI contained 32 target items and 16 fillers. All items were controlled for both syllable length and word length ranging between 7-10 syllables and 6-9 words. It is common for there to be fewer fillers on an oral task where the participants do not see the items. In addition, the items varied in tense, number, verb type, and syntactic environment. Therefore, one-third of all items were chosen to be fillers in order to balance the desire to have a sufficient number of fillers with procedural limitations. Moreover, traditionally only grammatical items are used in an EI, rather than using a combination of grammatical and ungrammatical items, as used in the GJT.

Fillers consisted of 4 sentences with embedded object pronouns in past and present tense. For instance, "We think he was happy last week" is a past tense example

of an embedded object pronoun. The fillers also included 12 simple sentences with subject-verb agreement which varied based on person and number (e.g. first person singular and plural, second person singular, third person singular and plural). This resulted in the 16 fillers for the EI as mentioned above. See Appendix B for the full list of fillers.

4.2.3 Grammaticality judgment task. The time-pressure GJT was designed to measure the participants' competence (Ellis et al. 2015) of the English copula in a variety of structures. The GJT was designed using Qualtrics software, in which the participants were presented with an item and asked if the item was grammatical. Participants only needed to click "Yes" or "No". The GJT contained the exact same grammatical target items as the EI plus 32 ungrammatical target items (i.e. the copula was absent). Therefore, there were 32 items in each tense, verb type, and number, and 16 items in each syntactic environment. Moreover, the GJT contained 32 fillers. While it is more advisable to have at least half of all items be fillers, it was decided to keep the fillers to one-third of the total items for a number of reasons. First, both participant fatigue and limited access to the participant were of concern. The researcher had access to the participants for one hour, and it was important to be mindful of not overworking the participants to the point where they just started guessing on the GJT. However, it was important to include enough target items so that statistical analyses could be run for the conditions of tense, verb type, and number. Therefore, as mentioned in the previous section, since there was such a wide variety of structures and environments for the items including *wh*-interrogatives, subject-auxiliary inversion, negation, and embedded clauses, only 32 fillers were used for the GJT. The fillers consisted of the 16 grammatical fillers

described in Section 4.2.2 for the EI plus the corresponding ungrammatical items, for a total of 32 fillers. See Appendix B for a complete list of fillers.

For the time-pressure GJT, participants were given 6 seconds to judge the grammaticality of each target item. In a trial run of the time-pressure GJT, advanced participants averaged 7 seconds in their response time to the GJT items. This average response time was decreased by 1 second in order to have participants pressured into answering more quickly than they would without a time-pressure component.

4.3 Procedure

All tasks were completed during class time in the intensive English language program. First, the participants were informed that the study was voluntary and would not affect their class grade, and those that agreed to participate signed a consent form. Next, the participants were given ten minutes to complete a cloze task borrowed from Slabakova (2000) in order to assess their level and group them accordingly into two levels corresponding roughly to intermediate (Low Group) and advanced (High Group).

For the first task, the participants read written instructions for the EI followed by oral instructions. Then, the participants sat at a computer with a headset containing a microphone. Each computer had Audacity to record the sound and VLC media player to play the sentences to be imitated. When the student was ready, the researcher went to each computer to start recording with Audacity and to play the sound with VLC. All of the headsets, microphones, and software were checked before the participants arrived to the room. Once the sound started, participants listened to a sentence, sat in silence for 5 seconds, and then heard a "beep," which was their cue to repeat the sentence they heard. The participants were instructed that the first sentence would be a "practice" sentence for

them to be able to get familiar with the procedure. The participants repeated this process for all 32 randomized grammatical target items and 16 fillers.

When participants finished the EI, they were instructed to raise their hand. The researcher went to each computer, stopped the recording, and saved the recording to a USB. At this point, the researcher gave the participant four short-answer questions to answer in writing as a distractor between the EI and the GJT. The goal was that the ten minutes of writing would be enough time and enough of a distractor to limit any priming effect from hearing the grammatical items in the EI. The four short-answer questions asked students about the best gift they ever received, their favorite city, what they did the previous summer, and what they were studying at university. The participants were instructed to write a paragraph for each question. While a few participants wrote only one sentence, most participants wrote at least several sentences for each question. Therefore, it is not believed that the chances of any priming effect were limited but not eliminated. See Section 6.2 for a discussion of possible priming effects.

When the participants completed the four short-answer questions, they raised their hands, and the researcher collected the paper from them. The researcher pulled up the Qualtrics survey, and the participants started by reading written directions for the GJT. The participants were informed in those written directions that they would have three practice questions in order to familiarize them with the Qualtrics webpage and the timing of the items. For each item, participants were asked the same question: "Is the following sentence grammatical", and the participant clicked "Yes" or "No". As mentioned above, participants had 6 seconds to answer each item or the survey automatically moved to the next question. The participants completed the time-pressure grammaticality judgment

task (GJT) with the same randomized 32 grammatical target items as the EI plus 32 ungrammatical target items and 32 fillers. All participants conducted the tasks in this same order, as conducting a balanced distribution of the EI and GJT where half the participants took the EI first and half took the GJT first was not feasible since data were collected across more than 12 different classes with varying numbers of participants in each class.

4.4 Analysis of data

4.4.1 Elicited imitation. For the elicited imitation task, two main analyses were conducted. The first analysis dealt strictly with copula omission. In this analysis, if participants omitted the copula in their repetition of a target sentence, that was "incorrect" and they received 0 points. If they produced any form of the copula, they received 1 point towards their total score of 32 (for 32 potentially correct items). For instance, if the item "I know that the boys are sick today" was produced as "I know that the boys is sick today", they received 1 point for including the copula. This allows us to look strictly at copula omission without "punishing" participants for an error in tense or number agreement. For the second analysis, the non-target-like productions of the copula are presented descriptively. This means any errors related to the copula in terms of tense, number agreement, omission of the *–ing* in auxiliary items, or copula doubling are discussed. For instance, several students produced the item "Is David your best student this year" as "Is David is your best student this year". This section allows for further investigation into the types of errors that participants made when actually producing the copula. By conducting both of these analyses, it allows us to look at both the

environments in which participants produced the copula *and* the ways in which they produced the copula in a non-target-like manner.

4.4.2 Grammaticality judgment task. For the GJT, participants received one point for correctly rejecting an ungrammatical item for a score out of 32. The participants also received one point for correctly accepting a grammatical item for an additional score out of 32. The "accepting grammatical items" portion of the GJT (AG) and the "rejecting ungrammatical items" portion of the GJT (RU) are statistically analyzed as separate tasks. When a participant accepts a grammatical item, this does not definitively demonstrate that the participant has acquired the particular feature being tested, only that the participant can recognize that an item "looks right". (Ellis, et al., 2015). For the purposes of this study, when a participant rejects an ungrammatical item after having accepted the grammatical version of that same item, this shows the researcher that the participant knows the item is grammatical when the copula is present and additionally knows the item is ungrammatical when the copula is absent. This is important because Arabic does have a copula, so the presence of a copula may "look right" to Arabic speakers, but Arabic does not have the copula in present tense. Therefore, we need to know that the participants not only recognize that having the copula is grammatical, but also that copula omission is ungrammatical. If the participant both accepts the grammatical items and accepts the ungrammatical items, this could be a case of positive transfer for the grammatical items and negative transfer for the ungrammatical items. Unfortunately, the design of this study is unable to distinguish this possibility of transfer from the possibility that the participants simply have a bias towards "accepting" items, which is a common effect of a task of this type. Alternatively, if

participants were to reject grammatical present tense items, this could suggest that the presence of a copula in the present tense does not look right to them, but we also have no way of knowing that it was not another part of the sentence that caused them to reject it. However, if that same participant also accepts the ungrammatical version of the item, then this more strongly supports the argument that the problem is the copula.

4.4.3 Statistical analysis. A statistics laboratory was consulted for the statistical analysis in this study. A series of repeated measures ANOVA tests were run for each condition using SPSS. This means that for each task, a repeated measures ANOVA was run for present versus past, main verb versus auxiliary verb, and singular versus plural. When a repeated measures ANOVA is used with just two repeated conditions, the results match that of a paired-samples t-test. In this case, the sphericity assumption is typically violated, so the results presented in this study are from the Wilks' Lambda test. It is acknowledged that by running a series of tests, the chances of making a Type I Error increase. Therefore, to guard against a Type I Error, this study uses a more stringent significance level of $p < .03$ (Perry, 2005). Lastly, the number of items for each syntactic environment was not sufficient for statistical analysis; thus, the results for syntactic environment are presented descriptively.

CHAPTER 5

RESULTS

The results of this study will be organized into two main portions: in Section 5.1, the scores on the GJT and the scores for copula production on the EI are presented for the participants as a whole before being discussed separately for the lower proficiency group (Low Group) and the higher proficiency group (High Group). The GJT analysis is broken down into the 32 items in which participants were asked to accept grammatical items (AG) and the 32 items in which the participants were asked to reject ungrammatical items (RU). As mentioned in the methodology section, the EI consisted of only the 32 grammatical items. After the results for the GJT and EI copula production are presented, Section 5.2 consists of a discussion of non-target like production of the copula in the EI including double copulas, *-ing* omission, and tense and number agreement errors.

5.1 Overview of GJT and EI Copula Production Results

The goal of this section is to look at whether the participants correctly accepted grammatical items (AG), correctly rejected ungrammatical items (RU), and produced the copula, in any form, on the Elicited Imitation (EI). For this reason, as described in Section 4.4, the participants were given 1 point if they produced any form of the copula in the elicited imitation and 0 points only if they omitted the copula entirely. For instance, if the target item were "We were reading last night," and the participant produced "We was reading last night," the participant still received 1 point for the answer

because it did include the copula. The analysis in this section focuses exclusively on whether the copula was produced in any form on the EI, and Section 5.1 will analyze errors in copula production other than copula omission. This resulted in very high scores for the elicited imitation task in this Section, but this does not mean the copula production was fully target-like. All non-target-like productions of the copula are discussed in Section 5.2.

Overall, it is most noteworthy that there were very few variations observed based on the selected parameters. Generally, regardless of the parameters, the participants performed best on the Elicited Imitation (EI), followed by Accepting Grammatical (AG) items and worst on Rejecting Ungrammatical (RU) items. When looking at the parameters within a given task, the participants scored better on the past tense than the present tense on the RU, better on the main items than the auxiliary items on the EI, and better on the singular items than the plural items on the AG and the EI. However, upon closer analysis, these differences held true only for the Low Group and not for the High Group. While the Low Group's results on the RU in terms of tense and verb type may be suggestive of some L1 transfer, there are no observable differences across tense, verb type, or number for the High Group. This seems to suggest that any evidence of L1 transfer observed in the Low Group is no longer present in the High Group. This appears to be in contrast to previous research such as Al-Zahrani (1993) and Al-Shayban (2007) who both argue for persistent L1 transfer even at the advanced level.

5.1.1 Overview of total scores across task.

All participants total score across task. As a whole, the 45 participants performed best on the Elicited Imitation (EI), followed by the Accepting Grammatical

items (AG), and worst on the Rejecting Ungrammatical items (RU). The mean scores and standard deviations for all participants are presented in Table 5.1. The RM ANOVA indicates that there is a significant difference between the participants overall accuracy between the AG and RU (Wilks Lambda = .690, F(1,44) = 19.756, p = .000), the AG and EI (Wilks Lambda = .583, F(1,44) = 31.429, p = .000), and the EI and RU (Wilks Lambda = .418, F(1,44) = 61.321, p = .000).

Table 5.1

Descriptive Statistics Total Score across Tasks

	Mean	**Std. Deviation**	**N**
Total AG (All)	27.111	3.8977	45
Total RU (All)	23.578	7.0598	45
Total EI (All)	29.78	3.096	45

However, when the scores are broken down by group, there is no statistically significant difference in performance for the High Group on the AG and RU tasks, which is discussed in more detail below.

Low Group total score across tasks. Following the whole group pattern, the Low Group scored poorest on RU items with a mean score of 17.2, better on the AG with a mean of 24.3 and best on the EI with a mean of 28. See Table 5.2 for descriptive statistics broken down by group. The RM ANOVA indicates a statistically significant difference between the Low Group's performance on all tasks (AG/RU: Wilks Lambda = .352, F(1,20) = 36.783, p = .000); AG/EI: Wilks Lambda = .532, F(1,20) = 17.578, p = .000); EI/RU: Wilks Lambda = .099, F(1,20) = 183.00., p = .000).

High Group total scores across task. Similarly to the Low Group, the High Group scored poorest on RU items with a mean of 29.2, better on AG items with a mean of 30 and best on the EI items with a mean of 31.4. See Table 5.2 for the descriptive

statistics broken down by group. The RM ANOVA shows that the High Group's differences in performance between the AG and EI (Wilks Lambda = .552, F(1,23) = 18.678, p = .000) and the RU and EI (Wilks Lambda = .571, F(1,23) = 17.261, p = .000) are both statistically significant. However, there is no statistically significant difference between the AG and the RU (Wilks Lambda = .979, F(1,23) = .499, p = .487). This seems to suggest that the deficit present at the lower proficiency which causes a discrepancy in performance on the AG and the RU is no longer present at the higher proficiency level.

Table 5.2

Descriptive Statistics Total Score across Tasks by Group

	Mean	**Std. Deviation**	**N**
Total AG (Low)	24.333	3.7193	21
Total RU (Low)	17.190	4.6111	21
Total EI (Low)	27.95	3.584	21
Total AG (High)	29.542	1.9556	24
Total RU (High)	29.167	2.6485	24
Total EI (High)	31.38	1.209	24

5.1.2 Scores by tense. The results for each task are broken down and presented to investigate the role of tense. This allows us to see if the participants performed better on one tense than the other, specifically better on past tense than present tense, as would be predicted by FR but not by IH. First, the scores for present tense and past tense are presented and compared within each task as a whole before breaking results down for each group. Next, the scores for the present tense are compared across tasks and the scores for the past tense are compared across tasks. This allows us to see if there were any parameters which broke with the overall pattern in which participants scored best on the Elicited Imitation (EI), followed by the Accepting Grammatical items (AG), and

worst on the Rejecting Ungrammatical items (RU). For instance, if one of the groups performed significantly better on the AG present tense items than the EI present tense items, this may be demonstrative of a competence/performance distinction.

Tense within tasks. As a whole, the participants' performance was not statistically significantly different based on tense on the AG (Wilks Lambda = .997, F(1,44) = .121, p = .730) and the EI (Wilks Lambda = .983, F(1,44) = , p = .393), but the they performed significantly better on the past tense items than on the present tense items on the RU (Wilks Lambda = .824, F(1,44) = 9.426, p = .004). The descriptive statistics are presented in Table 5.3. This demonstrates that the participants were able to recognize grammatical items as being grammatical and were able to produce the copula at the same rate regardless of tense. It is only on the RU that the participants performed significantly better on the past tense items than on the present tense items. Discussion of tense by group is presented in the following section.

Table 5.3

Descriptive Statistics Tense within Task

	Mean	**Std. Deviation**	**N**
Present AG (All)	13.622	2.1668	45
Past AG (All)	13.489	2.4920	45
Present RU (All)	11.178	3.8922	45
Past RU (All)	12.400	3.6519	45
Present EI (All)	14.82	1.614	45
Past EI (All)	14.96	1.651	45

Low Group Tense. The Low Group's performance is presented by tense in Table 5.4, and overall it shows that the Low Group recognizes the presence of the copula as correct regardless of tense, but they are still accepting an omitted copula in the present tense. The mean scores in Table 5.4 show different results for present and past tense

items on the AG, with participants unexpectedly performing better on the present tense items. However, the results of the RM ANOVA demonstrate that this difference is not statistically significant (Wilks Lambda = .928, F(1,20) = 1.550, p = .228), meaning the Low Group's performance on the AG does not appear to have been affected by tense. The mean scores in Table 5.4 show that participants performed better on past tense items than on present tense items on the RU, which would be predicted by Feature Reassembly. The results of the RM ANOVA demonstrate that this difference is, indeed, statistically significant (Wilks Lambda = .759, F(1,20) = 6.335, p = .020), meaning the Low Group's performance on the RU does appear to have been affected by tense. Lastly, the mean scores in Table 5.4 show that participants performed slightly better on past tense items than on present tense items on the EI; however, the results of the RM ANOVA demonstrate that this difference is not statistically significant (Wilks Lambda = .968, F(1,20) = .665, p = .424), meaning the Low Group's performance on the EI does appear to have been affected by tense. These results suggest that the Low Group appears to be able to recognize items containing the copula as grammatical regardless of tense. However, at this stage, the Low Group is still accepting present tense items without a copula as grammatical, even though they are able to reject past tense items without a copula as ungrammatical. This result may be suggestive of L1 transfer, but the High Group's performance suggests that this L1 transfer is overcome at higher proficiencies.

High Group Tense. The High Group's performance by tense is presented in Table 5.4, and the results indicate that the High Groups performance was not affected by tense on any task. The mean score for all present tense items on the AG is 14.5 items correct and 15.04 items correct for the past tense. However, the results of the RMANOVA

Table 5.4

Descriptive Statistics Tense within Task by Group

	Mean	**Std. Deviation**	**N**
Present AG (Low)	12.619	2.5194	21
Past AG (Low)	11.714	2.4727	21
Present RU (Low)	7.714	2.5326	21
Past RU (Low)	9.476	3.0597	21
Present EI (Low)	13.86	1.852	21
Past EI (Low)	14.10	1.972	21
Present AG (High)	14.500	1.3188	24
Past AG (High)	15.042	1.0826	24
Present RU (High)	14.208	1.6934	24
Past RU (High)	14.958	1.6545	24
Present EI (High)	15.67	.637	24
Past EI (High)	15.71	.751	24

demonstrate that this difference is not statistically significant (Wilks Lambda = .867, F(1,23) = 3.524, p = .073). Thus, the High Group's performance on the AG does not appear to have been affected by tense. The mean score for all present tense items on the RU is 14.21 items correct and 14.96 items correct for the past tense. However, the results of the RM ANOVA demonstrate that this difference is not statistically significant (Wilks Lambda = .877, F(1,23) = 3.218, p = .086). Thus, High Group's performance on the RU does not appear to have been affected by tense. The mean score for all present tense items on the EI is 15.67 items correct and 15.71 items correct for the past tense. However, the results of the RM ANOVA demonstrate that this difference is not statistically significant (Wilks Lambda = .996, F(1,23) = .087, p = .770). Thus, the High Group's performance on the EI does not appear to have been affected by tense.

While the Low Group's performance on the RU was affected by tense, the High Group's performance was not affected by tense on any task. This demonstrates that the

deficit present in the Low Group which caused them to perform better on past tense items than present tense items has been overcome at the higher proficiency level as evidenced by the High Group's performance.

Tense across task. Looking at tense across task allows us to see whether any of the parameters deviate from the overall pattern of the participants' performance on the tasks. When looking at the total scores for the Low Group in Table 5.1, the Low Group performed highest on the EI, next highest on the AG, and lowest on the RU. It would be informative to know if the participants' performance under a specific condition was so drastically different that it deviates from the pattern of total scores. However, this was not the case for either group. When looking separately at the Low Group's performance on present tense items across tasks and past tense items across task, there was no deviation from the overall pattern, meaning the Low Group still performed highest on EI, next highest on AG, and lowest on RU. The results of the RM ANOVA indicate that these differences are significant for the present tense (AG/RU: Wilks Lambda = .690, $F(1,44) = 19.478$, $p = .000$); AG/EI: Wilks Lambda = .700, $F(1,44) = 18.857$, $p = .000$); EI/RU: Wilks Lambda = .412, $F(1,44) = 62.868$, $p = .000$) and for the past tense (AG/RU: Wilks Lambda = .883, $F(1,44) = 5.816$, $p = .020$); AG/EI: Wilks Lambda = .662, $F(1,44) = 22.512$, $p = .000$); EI/RU: Wilks Lambda = .555, $F(1,44) = 35.224$, $p = .000$). These results suggest that the Low Group has the ability to recognize the presence of the copula as grammatical, but they are also still accepting copula omission as grammatical. The fact that the highest score was on the EI is likely due to the fact that elicited imitations are better measures of error types than of error rates, as discussed in Section 4.2.

When looking at the total scores for the High Group in Table 5.1, the High Group performed best on the EI, and there was no statistically significant difference between the performance on the AG and RU. When looking separately at their performance on present tense items across tasks and past tense items across tasks, there was no deviation from the overall pattern; the High Group performed best on the EI and there was no statistically significant difference between their performance on the AG and RU. The results of the RM ANOVA indicate that the differences between AG and EI and RU and EI are significant for the present tense (AG/EI: Wilks Lambda = .546, $F(1,23) = 19.102$, $p = .000$); EI/RU: Wilks Lambda = .514, $F(1,23) = 21.757$, $p = .000$) and for the past tense (AG/EI: Wilks Lambda = .802, $F(1,23) = 5.662$, $p = .026$); EI/RU: Wilks Lambda = .812, $F(1,23) = 2.308$, $p = .028$). However, the difference between the AG and RU is not statistically significant for present tense (Wilks Lambda = .975, $F(1,23) = .580$, $p = .454$) and for past tense (Wilks Lambda = .998, $F(1,23) = .047$, $p = .831$). This suggests that the High Group has overcome the deficit that caused the Low Group to perform significantly better on the AG than the RU. Overall, when looking at tense across tasks, neither group's performance across tasks deviated from the pattern of their overall scores. This means that the Low Group presents with a defect in tense that causes them to perform better on AG items than RU, but this defect has been overcome by the High Group.

To summarize the results for tense, there was only one instance where performance seems to have been affected by tense: The Low Group performed significantly better rejecting ungrammatical past tense items than present tense items. In

all other tasks, there was no statistically significant difference between tense nor did the scores for tense deviate from the overall score pattern (i.e. total scores).

5.1.3 Scores by verb type.

The results for each task are broken down and presented to investigate the role of verb type. This allows us to see if the participants' performance was affected by whether the copula was a main verb or an auxiliary verb. FR would predict a temporary difficulty in acquiring the copula in present progressive with the added difficulty of learning to map the *-ing* morphology to the verb in both tenses. First, the scores for main verb and auxiliary verb items are presented and compared within each task as a whole before breaking results down for each group. Next, the scores for the main verb items are compared across tasks and the scores for the auxiliary verb items are compared across tasks. This allows us to see if there were any parameters which broke with the overall pattern in which participants scored best on the Elicited Imitation (EI), followed by the Accepting Grammatical items (AG), and worst on the Rejecting Ungrammatical items (RU).

Verb type within tasks. When looking at the results for verb type as a whole, the only difference in performance was on the EI in which participants performed much better on the main verb items (Mean = 15.18) than on the auxiliary verb items (Mean = 14.6) as seen in Table 5.5. The results of the RM ANOVA show that this difference is statistically significant (Wilks Lambda = .821, $F(1,44) = 9.582$, $p = .003$). However, when results are broken down by group, only the Low Group performed significantly better on main verb items than on auxiliary verb items.

Table 5.5

Descriptive Statistics Verb Type within Task

	Mean	**Std. Deviation**	**N**
Main AG (All)	13.489	2.1065	45
Aux AG (All)	13.622	2.2792	45
Main RU (All)	11.622	3.6760	45
Aux RU (All)	11.956	3.6801	45
Main EI (All)	15.18	1.284	45
Aux EI (All)	14.60	1.982	45

Low Group Verb Type. The Low Group's performance on the AG is presented by verb type in Table 5.6. The mean scores presented in Table 5.6 show that the participants performed slightly better on auxiliary items than on main verb items on the AG; however, the results of the RM ANOVA demonstrate that this difference is not statistically significant (Wilks Lambda = .990, F(1,20) = .202, p = .658). Thus, Low Group's performance on the AG does not appear to have been affected by verb type. The mean scores presented in Table 5.6 show that participants performed slightly better on the auxiliary items than on the main verbs items on the RU; however, the results of the RM ANOVA demonstrate that there is no significant difference (Wilks Lambda = 1.000, F(1,20) = .007, p = .934) . Thus, Low Group's performance on the RU does not appear to have been affected by verb type. The mean scores presented in Table 5.6 for the EI show that the participants performed better on main verb items than on auxiliary items, and the results of the RM ANOVA demonstrate that this difference is, indeed, statistically significant (Wilks Lambda = .549, F(1,20) = 16.419, p = .001) . Thus, the Low Group's performance on the EI does appear to have been affected by verb type with scores for the main verb items being significantly higher.

These results demonstrate that in production, the Low Group is more likely to omit the copula for auxiliary items than for main verb items. This is interesting because

just like main verb items, auxiliary items in the past tense in Arabic contain a copula, but in the present tense there is no copula. Therefore, if participants have learned to produce the copula in present tense main verb items, as demonstrated in the results on tense presented above, then it would be expected that the participants would also be able to produce the copula in present tense auxiliary items. However, in progressive items in English, the participants have the added burden of grammatically encoding progressive aspect on the verb, which is contextually encoded in Arabic. It is interesting that the Low Group, rather than simply omitting *-ing* (i.e. the grammatical encoding of progressive aspect), appears to have omitted the entire copula. This may suggest that the copula omission in this instance is due to the failure to encode relevant features, rather than due to a representational deficit in which they have not yet acquired the English copula. However, it is possible that that the participants are simply producing the simpler morphological realization.

Table 5.6

Descriptive Statistics Verb Type within Task by Group

	Mean	**Std. Deviation**	**N**
Main AG (Low)	12.048	1.9615	21
Aux AG (Low)	12.286	2.4524	21
Main RU (Low)	8.571	2.8909	21
Aux RU (Low)	8.619	2.3765	21
Main EI (Low)	14.62	1.564	21
Aux EI (Low)	13.33	2.244	21
Main AG (High)	14.750	1.2597	24
Aux AG (High)	14.792	1.2847	24
Main RU (High)	14.292	1.6280	24
Aux RU (High)	14.875	1.3929	24
Main EI (High)	15.67	.702	24
Aux EI (High)	98%	.624	24

High Group Verb Type. The High Group's performance on the AG is presented by verb type in Table 5.6. The mean scores for the AG show that participants performed slightly better on auxiliary items than on main verb items, but with the different being so minute, naturally, the results of the RM ANOVA demonstrate that there is no statistically significant difference (Wilks Lambda = .999, F(1,23) = .016, p = .901). Thus, the High Group's performance on the AG does not appear to have been affected by whether the copula introduced a verbal or nonverbal predicate. The mean scores on the RU show that participants performed better on auxiliary items than on main verb items. However, the results of the RM ANOVA demonstrate that this difference is not statistically significant (Wilks Lambda = .859, F(1,23) = 3.769, p = .065) . Thus, the High Group's performance on the RU does not appear to have been affected by verb type. The mean scores on the EI show that the participants performed slightly better on auxiliary items than on main verb items, but the results of the RM ANOVA demonstrate that this difference is not statistically significant (Wilks Lambda = .994, F(1,23) = .138, p = .714). Thus, the High Group's performance on the EI does not appear to have been affected by verb type.

To conclude, the High Group did not perform statistically differently on any task based on whether the copula was a main verb or an auxiliary verb, but the Low Group performed significantly better on main verb items than on auxiliary items on the EI.

Verb type across tasks. Looking at verb type across tasks allows us to see if the participant's performance on either main verb or auxiliary items was significantly different from their overall performance on each task, as indicated by their total scores in Table 5.2. Regardless of whether the copula was a main verb or an auxiliary verb, the Low Group performed best on the EI, then the AG and worst on the RU. The results of

the RM ANOVA indicate that these differences are significant for the main verb items (AG/RU: Wilks Lambda = .440, F(1,20) = 25.473, p = .000); AG/EI: Wilks Lambda = .417, F(1,20) = 28.012, p = .000); EI/RU: Wilks Lambda = .126, F(1,20) = 138.446, p = .000) and for auxiliary items (AG/RU: Wilks Lambda = .454, F(1,20) = 24.062, p = .000); AG/EI: Wilks Lambda = .825, F(1,20) = 4.231, p = .053); EI/RU: Wilks Lambda = .185, F(1,20) = 87.823, p = .000). This means that the deficit that causes the Low Group to perform better on the AG than the RU persists regardless of verb type.

When looking at the total scores for the High Group in Table 5.2, the High Group performed best on the EI, and there was no statistically significant difference between the performance on the AG and EU. When looking separately at their performance on main verb items across tasks and auxiliary verb items across tasks, there was no deviation from the overall pattern; the High Group performed best on the EI, next best on the AG, and worst on the RU. The results of the RM ANOVA indicate that the differences between the AG and EI and the RU and EI are significant for the main verb items (AG/EI: Wilks Lambda = .597, F(1,23) = 15.547, p = .001); EI/RU: Wilks Lambda = .551, F(1,23) = 18.762, p = .000) and for the auxiliary verb items (AG/EI: Wilks Lambda = .727, F(1,23) = 8.616, p = .007); EI/RU: Wilks Lambda = .185, F(1,23) = 87.823, p = .000). However, the difference between the AG and RU is not statistically significant for the main verb items (Wilks Lambda = .920, F(1,23) = 2.001, p = .171) or for the auxiliary verb items (Wilks Lambda = .998, F(1,23) = .055, p = .817). This means there is no statistically significant difference between performance on the AG and the RU regardless of verb type which in turn suggests that the High Group has overcome the deficit that caused the Low Group to perform significantly better on the AG than the RU. Overall, when

looking at verb type across tasks, neither group's performance across tasks deviated from the pattern of their overall scores.

To summarize the results for verb type, there was only one task in which there was a statistically significant difference between verb type. The Low Group performed more accurately on the main verb items than the auxiliary items on the EI. To put it differently, the Low Group omitted the copula more often for auxiliary items on the EI than for main verb items. Otherwise, there were no differences within tasks and the results across tasks patterned with overall scores without any statistically significant deviation.

5.1.4 Scores by number.

By comparing performance based on number, we get to the heart of the difference between FR and IH. FR would predict an initial defect in which participants at a lower proficiency perform better on singular items than on plural items, but this defect is predicted to no longer be present at higher proficiencies. IH would predict that since number on a verb is a relational rather than a head feature, it remains unacquirable even at higher proficiencies.

Number within task. Number is the parameter which saw the greatest number of differences in performance within tasks. As a whole, the participants performed better on singular items than plural items on both the AG and the EI as seen in Table 5.7. The RM ANOVA indicates that these differences are significant (AG: Wilks Lambda = .562, $F(1,44) = 34.269$, $p = .000$; EI: Wilks Lambda = .879, $F(1,44) = 6.056$, $p = .018$). However, when broken down by group, this distinction holds true only for the Low Group.

Table 5.7

Descriptive Statistics Number within Task

	Mean	**Std. Deviation**	**N**
Singular AG (All)	14.467	1.5754	45
Plural AG (All)	12.667	2.6884	45
Singular RU (All)	11.978	3.7142	45
Plural RU (All)	11.622	3.5438	45
Singular EI (All)	15.16	1.364	45
Plural EI (All)	14.69	1.869	45

Low Group Number. The Low Group's performance by number (i.e. plural vs. singular) is presented in Table 5.8. The mean scores on the AG show that participants performed much better on singular items than on plural items. The results of the RM ANOVA demonstrate that this difference is statistically significant (Wilks Lambda = .223, F(1,20) = 69.605, p = .000). Thus, the Low Group's performance on the AG does appear to have been affected by number, with the group scoring significantly higher on singular items than plural items. The mean scores for the RU show that the participants performed slightly better on singular items than on plural items. The results of the RM ANOVA demonstrate that this difference is not statistically significant (Wilks Lambda = .954, F(1,20) = .975, p = .335). Thus, the Low Group's performance on the RU does not appear to have been affected by number. The mean scores on the EI show that participants performed better on the singular items than on the plural items, and the results of the RM ANOVA demonstrate that this difference is, indeed, statistically significant (Wilks Lambda = .792, F(1,20) = 5.268, p = .030). Thus, the Low Group's performance on the EI does appear to have been affected by number.

These results show that the Low Group was significantly more likely to accept singular grammatical items than plural grammatical items, meaning they tended to reject

grammatical plural items. Moreover, the Low Group was significantly more likely to produce the copula in singular contexts than in plural contexts.

Table 5.8

Descriptive Statistics Number within Task by Group

	Mean	**Std. Deviation**	**N**
Singular AG (Low)	13.71	1.736	21
Plural AG (Low)	10.62	2.312	21
Singular RU (Low)	8.810	2.9089	21
Plural RU (Low)	8.381	2.0366	21
Singular EI (Low)	14.48	1.662	21
Plural EI (Low)	13.62	2.202	21
Singular AG (High)	15.125	1.0759	24
Plural AG (High)	14.458	1.4136	24
Singular RU (High)	14.750	1.3593	24
Plural RU (High)	14.458	1.5598	24
Singular EI (High)	15.75	.608	24
Plural EI (High)	15.63	.711	24

High Group Number. The High Group's performance is presented by tense in Table 5.8. The mean scores for the AG show that participants performed better on the singular items than on the plural items. However, the results of the RM ANOVA demonstrate that this difference is not statistically significant (Wilks Lambda = .856, $F(1,23) = 3.874$, $p = .061$). Thus, the High Group's performance on the AG does not appear to have been affected by number. The mean scores for the RU show that participants performed slightly better on singular items than on plural items, but the results of the RM ANOVA demonstrate that this difference is not statistically significant (Wilks Lambda = .953, $F(1,23) = 1.146$, $p = .295$). Thus, the High Group's performance on the RU does not appear to have been affected by number. The mean scores for the EI

show that participants performed slightly better on the singular items than on the plural items. However, once again, the results of the RM ANOVA demonstrate that this difference is not statistically significant (Wilks Lambda = .946, F(1,23) = 1.302, p = .266). Thus, the High Group's performance on the EI does not appear to have been affected by number.

Similarly to both tense and verb type, the results under the condition of number show a defect in the Low Group, causing them to perform better on singular items than plural items on both the AG and the EI, but this defect appears to dissipate in High Group. The fact that the High Group performs statistically similar regardless of number suggests that they have acquired an uninterpretable feature contra IH.

Number across tasks. Looking at number across tasks allows us to see if the participant's performance on either singular or plural items was significantly different from their overall performance on each task, as indicated by their total scores in Table 5.2. Regardless of whether the copula was singular or plural, the Low Group performed best on the EI, then the AG and worst on the RU. The results of the RM ANOVA indicate that the differences between the AG and RU and the RU and EI are statistically significant for singular items (AG/RU: Wilks Lambda = .316, F(1,20) = 43.214, p = .000); EI/RU: Wilks Lambda = .175, F(1,20) = 94.533, p = .000) and for plural items (AG/RU: Wilks Lambda = .560, F(1,20) = 15.722, p = .001); EI/RU: Wilks Lambda = .143, F(1,20) = 120.278, p = .000). The difference between the AG and EI for singular items was not statistically significant (Wilks Lambda = .835, F(1,20) = 3.945, p = .061), but the difference between the AG and EI for plural items was statistically significant (Wilks Lambda = .446, F(1,20) = 24.868, p = .000).

When looking at the total scores for the High Group in Table 5.2, the High Group performed best on the EI, and there was no statistically significant difference between the performance on the AG and EU. When looking separately at their performance on singular items across tasks and plural items across tasks, there was no deviation from the overall pattern; the High Group performed best on the EI, next best on the AG, and worst on the RU. The results of the RM ANOVA indicate that the differences between the AG and EI and the RU and EI are significant for the singular items (AG/EI: Wilks Lambda = .801, $F(1,23) = 5.731$, $p = .025$); EI/RU: Wilks Lambda = .636, $F(1,23) = 13.143$, $p = .001$) and for the plural items (AG/EI: Wilks Lambda = .581, $F(1,23) = 16.574$, $p = .000$); EI/RU: Wilks Lambda = .637, $F(1,23) = 13.105$, $p = .001$). However, the difference between the AG and RU is not statistically significant for the singular items (Wilks Lambda = .950, $F(1,23) = 1.220$, $p = .281$) or for the plural items (Wilks Lambda = 1.000, $F(1,23) = .000$, $p = 1.0$). This means there is no statistically significant difference between performance on the AG and the RU regardless of number.

To summarize the results for number, the Low Group performed significantly better on singular items than on plural items on both the EI and the AG portion of the GJT. This means the Low Group correctly accepted more singular grammatical items than plural grammatical items, and that they also were more likely to omit the copula in plural items than in singular items. Moreover, as discussed above, there was a deviation from overall scores for the Low Group in their performance on singular items on EI and AG, as there was no statistically significant difference.

5.1.5 Scores by Syntactic Environment

The target items in this study consisted of embedded sentences, negative sentences, yes/no questions and *wh*- questions. The primary purpose of including a variety of syntactic environments was to eliminate a potentially confounding variable and to make the results more generalizable. For instance, if this study were to use only sentences with embedded clauses, then the argument could be made that there is something about this particular environment which is responsible for any observed results. By including a variety of items, any statistically significant results reported in this study hold true across a variety of syntactic environments. As a secondary purpose, the results from the various syntactic environments may be helpful for future research. Therefore, the results of the syntactic environments are presented descriptively below.

Overview for both groups. Table 5.9 below is an overview of the percent correct for each group for each task by syntactic environment (i.e. embedded clause, negation, subject-auxiliary inversion and *wh*-questions). The two groups performed quite similarly. For the Accepting Grammatical items (AG), the highest scores for both groups were on the negative items and the lowest scores were on the subject-auxiliary inversion items. For the Rejecting Ungrammatical items (RU), the groups, once again, performed most accurately on the negative items with very similar (relatively much lower) scores for the other syntactic environments. Lastly, on the Elicited Imitation (EI), both groups scored very highly on the subject-auxiliary inversion and negative items and seemingly much worse on the embedded and *wh*- items. While the rankings varied quite a bit across task (e.g. subject-auxiliary inversion being the lowest score for the AG and the highest for the EI), it appears that the two groups performed similarly to one another.

Table 5.9

Overview Syntactic Environments

	Low Group AG	**Low Group RU**	**Low Group EI**	**High Group AG**	**High Group RU**	**High Group EI**
Neg	83%	85%	95%	95%	95%	100%
WH	79%	**45%**	86%	89%	90%	95%
Embedded	75%	51%	**71%**	94%	**82%**	**93%**
S-AUX	**67%**	52%	97%	**86%**	90%	100%

**Lowest score is bolded*

The first observation for the Low Group is that subject-auxiliary inversion is ranked lowest for the AG and that only slightly more than half of the ungrammatical subject-auxiliary inversion items were correctly rejected. Despite this, subject-auxiliary inversion was produced in the EI with the highest rate of accuracy- 97%. With this being the case, an argument could be made that the EI is not a good measurement when the target structure is contained in the first word of the item. It is possible that the unusual level of markedness put on the copula when it is sentence-initial is responsible for the extremely accurate scores for both the Low Group and the High Group in the EI. Despite this discrepancy, the trends for the rest of the syntactic environments are very similar. Negative items were the most accurate environment across both groups and WH and embedded items were the least accurate.

Similarly to the Low Group, the first observation for the High Group from Table 5.9 above is that subject-auxiliary inversion is ranked lowest for the AG, and despite this, subject-auxiliary inversion was produced in the EI with a 100% accuracy rate. This seems to provide further evidence for the argument that the EI may not be a good measurement when the target structure is contained in the first word of the item due to the markedness of the sentence-initial copula. Also similarly to the Low Group, negative

items were the most accurate environment *and wh-* and embedded items were the least accurate. The fact that both groups performed better on negative items than any other items and across all tasks may be a good point to further investigate in future research.

Syntactic environments by tense, verb type, and number.

As discussed in the methodology, in order to be able to cover a wide range of potential syntactic environments for copula omission, the ability to run statistical analyses was sacrificed for the syntactic environments. However, it is still worth investigating each environment and its interactions with tense, verb type and number as a potential starting point for future research. Therefore, for each syntactic environment, the mean scores were compared for past/present tense, main/auxiliary verb type, and singular/plural number. For the majority of the results, the means were within ten percentage points of each other. For instance, on the AG, Low Group had a mean of 81% on *wh-* present tense items, and a mean of 76% on *wh-* past tense items. Considering the number of participants in this study and the number of items in each syntactic environment, a difference in mean of only a handful of percentage points is likely not suggestive of anything.

Therefore, in order to identify the most worthwhile results, the results below only present conditions in which there was a greater than ten percentage points difference in the means between corresponding items. For instance, if the participants had a mean score of 65% on embedded present tense items, and a mean score of 40% on embedded past tense items, then embedded past tense is listed as a "problematic" environment below, as the discrepancy in means was more than ten percentage points. For the purposes of this study, ten percentage points is essentially an arbitrary cut-off point, but it

serves its purpose in allowing us to identify a handful of the most problematic environments for potential future research. These results are just meant as an indication that within one of the syntactic environments, there was a difference in performance based on either tense, verb type or number, despite the fact that this difference cannot be statistically supported.

Table 5.10 below shows that on the AG, the Low Group scored much more highly on singular items than plural items across three of the syntactic environments. Moreover, in the subject-auxiliary inversion environment, participants scored more than 20 percentage points higher on the auxiliary items than the main verb items. On the RU, the Low Group performed worse on present tense subject-auxiliary inversion and embedded items than past tense items. Moreover, plural *wh*- items once again proved problematic. On the EI, plural items were once again problematic in the *wh*- and embedded environments. In contrast to their performance on the RU, the Low Group scored higher on the main verb embedded items than the auxiliary items on the EI. The important take-aways from this are that for the Low Group, the present tense and plural items were particularly problematic across several syntactic environments, but the results were inconsistent for verb type items.

Table 5.10

Problematic Syntactic Environments by Tense, Verb, and Number

	AG	**RU**	**EI**
Group 1	*Wh*- plural SAI plural Embedded plural SAI main	*Wh*- plural SAI present Embedded present *Wh*- aux Embedded main	*Wh*- embedded *Wh*- plural Embedded aux
Group 2	SAI plural SAI main	*Wh*- plural Embedded present	(*Wh*- plural)

As for the High Group, there are far fewer instances where there is a ten percent difference or more between corresponding items. However, the few instances where there are large differences are worth noting and comparing to the Low Group. For the AG, the subject-auxiliary inversion plural items received many more errors than the subject-auxiliary inversion singular items. This was also true for Low Group on the AG. On the RU, the embedded present tense items received many more errors than the past tense items. This also held true for the Low Group's performance on the RU. On the EI, there were no distinctions of 10 percentage points, possibly due to the fact that the scores on the EI were overall so high. This could be indicative of the fact that the EI is simply not an astute enough measurement of advanced students' omissions to find distinctions with such few items for each environment. The important take-aways here are that the only two problematic areas for the High Group were also problematic areas for the Low Group- plural items in subject-auxiliary inversion environment and present tense items in the embedded environment. To sum up the results for the syntactic environments, both groups generally performed much better in the negative syntactic environments than any other syntactic environment regardless of the task. When looking at the interaction of syntactic environment and tense, verb type, and number, both groups shared a couple of a similarities. Both groups struggled with subject-auxiliary inversion plural items and embedded present tense items. Since these problems persisted across both groups, they are certainly an area deserving future research. Additionally, for the Low Group, the present tense was particularly problematic in two syntactic environments and plural items were very problematic across several syntactic environments and tasks. This highlights the present tense and plural items as being problematic more generally. On the other

hand, the results for the verb type items were inconsistent, with the Low Group sometimes performing better on the main verb items and other times performing better on the auxiliary items.

5.2 EI Additional Errors

5.2.1 Overview of EI additional errors. The copula errors in the EI besides copula omission are plentiful enough that they deserve discussion. The EI errors besides copula omission consisted of omitting *–ing* in progressive items, doubling of the copula in interrogatives, and tense and agreement errors. The table below shows that the Low Group made the error of omitting *–ing* at a rate of 5% whereas the High Group did not make this error at all. The Low Group doubled the copula in interrogatives 8% of the time and the High Group did so 5% of the time. In terms of tense and agreement errors, the Low Group made both tense and agreement errors at a rate of 14% whereas the High Group made tense errors 6% of the time and agreement errors 8% of the time. To put this in perspective, 35% of the items produced by the Low Group contained errors other than copula omission, as compared to the rate of copula omission at 13%. This means that even at lower proficiency levels, the rate of copula error is higher than the rate of copula omission. For the High Group, 17% of the total items produced contained errors besides copula omission, as compared to the rate of copula omission at 2%. This means that while the amount of copula errors and copula omissions decreased as proficiency increased, the rate of copula errors was still greater than the rate of copula omission. With the instances of these errors being so high, they deserve to be addressed. Each error type is discussed in more detail below to investigate its interaction with tense, verb type, and number, as was done with copula omission above.

Table 5.11

Overview of EI Production Errors

	Omit –ing	**Double Copula**	**Tense**	**Agr**	**Total Errors***	**Copula Omissions**
Low Group	5%	8%	14%	14%	35%	13%
High Group	0%	5%	6%	8%	17%	2%

**Excluding copula omission errors*

5.2.2 Omission of –ing. The Low Group was the only group to commit errors of *–ing* omission. Since only half of the items contained the progressive form of a verb, there are not enough items to conduct a statistical analysis. However, the fact that *-ing* omission occurs at all is an interesting finding, as FR would predict that even after acquiring features related to progressive aspect, second language learners may not morphologically encode those features, as we saw with *-ing* omission in the Low Group. This falls in line with the FR prediction that a feature such as aspect can be acquired but that the production of its morphological realization may lag.

5.2.3 Doubling of copula in interrogatives. Another EI error that was committed by both groups is the doubling of the copula in interrogatives. For instance, the item "Is Layla reading a book?" was repeated as "Is Layla is reading a book?". This also occurred to a lesser extent in *wh-* questions. This error occurred a total of 8% of the time with the Low Group and 5% of the time with the High Group. This error indicates that participants are filling both positions in which the morphology could surface, meaning the participants are not lacking the proper representations but rather are unsure as where to map them. This supports a performance rather than a representational-based theory. The implications of both of these possibilities are discussed in further detail in Section 6.1.

5.2.4 Errors in tense and agreement. The errors in tense and agreement are presented together as they are closely connected, and many errors included an error in both tense and agreement. Table 5.12 below presents a breakdown of the copula productions by tense and agreement for the Low Group in addition to the in the rate of copula omission for comparison. The left column is the target or prompt item and the top row is what the participant produced. For instance, when the prompt contained "is" (i.e. singular, present tense item), it resulted in omission 8% of the time, target-like "is" 82% of the time, "are" 1% of the time, "was" 8% of the time and "were" 0% of the time. Thus, for the singular present tense items, there was a change in number 1% of the time and a change in tense 8% of the time. Interestingly, there was never a change in both.

Table 5.12

Copula Production by Tense and Agreement

Prompt\result	**Ø**	**Is**	**Are**	**Was**	**Were**
Is (Low)	8%	82%	1%	8%	0%
Are (Low)	19%	18%	54%	9%	0%
Was (Low)	10%	21%	0%	69%	0%
Were (Low)	14%	10%	9%	19%	55%
Is (High)	2%	94%	0%	4%	0%
Are (High)	3%	8%	81%	3%	5%
Was (High)	2%	3%	0%	96%	0%
Were (High)	2%	1%	3%	18%	76%

For the Low Group, "is" was the most accurately produced (82%) followed by "was" (69%) followed by "are"/"were" at 54%/55%. This means the singular items were much more likely to be produced accurately than the plural items and that the present tense singular item was much more likely to be produced accurately than the past tense singular item. To look at the results another way, it is interesting to note how often each

form of the copula was produced in a non-target like structure. For instance, how often were the other copula forms incorrectly produced as "is"? There were 84 changes from the other forms to "is", 61 changes to "was", 17 changes to "are" and 0 changes to "were". This follows the same pattern as above with the singular present tense being the most commonly "misused" form, followed by singular past, plural present, and plural past. Thus, the preference for copula form for the Low Group appears to look like the following where singular present is preferred to singular past which is preferred to plural present which is preferred to plural past.

Singular present > Singular past > Plural present > Plural past

Another interesting point to mention is that there were instances where the participant changed the subject to match their non-target like copula form. This resulted in a grammatical, yet not target-like, production. For instance, the item "The girls were reading yesterday" was produced as "The girl was reading yesterday". This was counted as a change from "were"→"was" in the chart below, despite the fact it can be argued that the participant simply never perceived the plurality of the subject and consequently produced a singular subject and singular copula form. To be specific, 11 of the 31 changes from *are* to *is*, and 24 of the 32 changes from *were* to *was*, correctly matched the singular subject that the participant produced. However, even if we were to "subtract" these from the total errors due to the possible conflict with perception, the preference for copula form remains the same as specified above.

For the High Group, the most accurately produced copula forms were "was"/"is" at 96%/94%, followed by "are" at 81% and "were" at 76%. This is similar to the Low Group's performance in that the singular items were much more likely to be produced accurately than the plural items, but the tense distinction seems to be less for the High

Group with “was” and “is” being produced at nearly the same rate and “are” and “were” produced at only slightly different rates. In terms of number of non-target like productions, “was” was produced in a non-target like environment 49 times, followed by “is” at 28 times, followed by “were” at 10 times and “are” at 5 times. Once again, we see that singular items are much preferred to plural items, but that the preference for present tense has seemingly started to dissipate at this point. It is important to note that a whopping 30/35 times that “were” was changed to “was”, the participants also changed the subject to a singular subject. This subject change additionally occurred 9/15 times for the change from “are” to “is”. Furthermore, 6/10 times that “were” was produced in a non-target-like item, the adverb was changed to match the tense, again making the production grammatical but not target-like. To conclude, it seems like the High Group’s preference for copula form relies much more heavily on number than tense. In this case, the order of preference would be as follows where singular present and singular past are more strongly preferred than plural present and plural past.

Singular present, singular past >Plural present, plural past

5.3 Summary of All Results

A summary of all of the results of this study is presented here. As a whole, participants scored best on the EI, followed by the AG and worst on the RU. On the RU, the participants scored significantly better on past tense items than on present tense items. On the EI, they scored significantly better on main items than on auxiliary items. On the EI and the AG, they scored significantly better on singular items than on plural items. However, when their performance is broken down by group, these trends hold true only for the Low Group. The High Group, when investigated separately, performed best on

the EI, and with no statistically significant difference between the AG and RU, and there were no observable differences in their performance regardless of condition.

For the syntactic environments, it was found that both groups performed extremely well on the negative items and performed poorly on the subject-auxiliary inversion items on the GJT, but very accurately on the subject-auxiliary inversion items on the EI. This discrepancy may be attributed to the markedness of having a sentence-initial copula in the subject-auxiliary inversion construction. Both groups performed especially poorly with the subject-auxiliary inversion plural and with the embedded present tense items. For the Low Group, both present tense items and plural items were problematic across multiple syntactic environments.

For the additional EI errors, it was found that the Low Group omitted the *–ing* in progressive constructions 5% of the time. It was also found that the Low Group produced a double copula in interrogative constructions 8% of the time and the High Group did so 5% of the time. Lastly, the tense and agreement errors seem to suggest that the Low Group had a preference for singular items and present tense items. On the other hand, the preference for present tense items seems to have dissipated for the High Group, but the preference for singular items remains.

CHAPTER 6

DISCUSSION

6.1 Discussion of the Research Question

The driving question for this study is whether the Interpretability Hypothesis or Feature Reassembly best accounts for the data of this study, which would inform whether copula omission by Arabic speaking learners of English is a representational or performance deficit. Under the Interpretability Hypothesis, interpretable features, defined in this study as head features, are interpretable and fully acquirable. Alternatively, uninterpretable features, defined in this study as features which are not inherent to the head, but rather are relational to the head features, are uninterpretable. Uninterpretable features resist parameter resetting even at the most advanced levels of proficiency, resulting in a deficit in the interlanguage syntax of a second language learner. Under Feature Reassembly, errors in a certain feature are attributable to a failure to learn how the feature is encoded in the new language (i.e. the morphology), not a deficit in the L2 learner's underlying representation of the feature. Slabakova (2009) used Lardiere (2008)'s prediction about "mismatched" features in the L1 and L2 to develop a scale of difficulty in the acquisition of L2 features. L2 learners whose L1 encodes a feature contextually would face the greatest challenge in acquiring a language in which a feature is encoded grammatically. Alternatively, moving from an L1 where a feature is encoded grammatically in a different way than how it is encoded grammatically in the L2 would prove slightly less challenging. Lastly, it is predicted that if both

languages grammatically encode a feature in the same way (i.e. no reassembly is required), then this should be the least arduous acquisition. In determining which of these theories best accounts for the data of this study, we would gather insight into whether copula errors by Arabic-speaking English language learners are representation or performance-based in nature.

For the purposes of this study, the Interpretability Hypothesis would predict more accurate production of tense than of number agreement on V due to the fact that tense is an interpretable feature on V and number agreement is not. In addition, since the deficit is predicted to be permanent, we would not expect to see the more proficient group in this study, the High Group, overcome the deficit, meaning they would be expected to continue making errors in number. In contrast, Feature Reassembly, as constrained by Slabakova (2009), would predict in terms of tense that since the present tense Arabic copula is encoded contextually, unlike the English present tense copula which is encoded grammatically, the present tense copula would be expected to be more difficult to acquire than past tense copula, which requires only the learning of a new grammatical encoding. Relatedly, the present progressive in Arabic is also encoded contextually, as in Standard Arabic, the present simple and present progressive are identical and distinguished by context. Alternatively, in English, the present progressive requires the overt realization of the copula (unlike present progressive in Arabic), and it also requires the addition of the suffix *-ing* on the copula. This would be a change from contextual to grammatical encoding, plus the added step of acquiring the *-ing* suffix, which would again be predicted to be very difficult. The past progressive, on the other hand, is formed in Arabic with the copula plus the infinitive form of the verb, so the only difference between

the Arabic and English past progressive is the English addition of the *-ing* suffix. This would predict that the acquisition of the past progressive should only be moderately difficult when compared to simple past, and specifically it would be expected that the participants would omit the *-ing*. Lastly, in terms of number, plural in Arabic is realized on the suffix in the SVO word order, but it is not realized at all in VSO word order. Therefore, it would be expected that plural would be only moderately difficult to acquire in the SVO word order, but it would be more difficult to acquire in the VSO order (i.e. subject-auxiliary inversion and *wh*- questions). Most importantly, though, unlike IH, FR would predict that this obstacle would be overcome by higher proficiency learners.

All in all, the predictions made by Feature Reassembly as constrained by Slabakova (2009) do a remarkable job of accounting for the data of this study. First, the Low Group performed statistically significantly better on present tense items on the RU than past tense items, but there was no statistically significant difference in performance between present and past tense items for the High Group. This may indicate that in the Low Group, there were remnants of L1 transfer as the Low Group accepted the presence of the copula at similar rates regardless of tense, but they also accepted the omission of the copula at similar rates regardless of tense. However, the deficit that caused the Low Group to perform better on present tense items than on past tense items was overcome by the High Group.

Second, the Low Group performed statistically significantly better on main verb items than auxiliary verb items on the EI. This means that the Low Group was more likely to produce the copula when it was the only verb in the sentence than when it was acting as an auxiliary verb in a progressive construction. Additionally, in the auxiliary

verb items, we saw the omission of *-ing* in the EI. This omission occurred in both the past and present tense even when the participant had correctly produced the copula. This suggests that these participants had acquired the relevant features for progressive aspect on the English copula as evidenced by the production of the copula, yet the morphological encoding for progressive aspect took additional time to be acquired, as suggested by FR. Then this deficit is overcome, as evidenced by the fact that High Group did not commit this error.

Another piece of evidence that may lend support to the idea that the mapping of morphology is at the heart of the copula errors in this study is that participants produced "double copulas." For instance, for the item "Is Layla reading a book?" the participants produced "Is Layla is reading a book?". It is a bit ironic that in a study on copula omission, one of the errors was the production of the copula twice. However, this error could be explained in the same way as the other errors- it is simply a result of the fact that the participants do not know where to map the morphology, and it ends up mapped twice. While this is an interesting idea it is also necessary to note that this error could also be explained as a syntactic error in which the participants are failing to delete the lower copy. This study was not designed to differentiate between these two types of errors.

The participants' performance on number provides some of the strongest support for FR. The Low Group performed statistically significantly better on singular items than on plural items on the AG and EI. In fact, the Low Group performed so significantly better on singular items than on plural items that their AG score was not statistically different from their EI score for the only time in the study. This is particularly interesting because unlike tense, there is no L1 transfer that would cause Arabic speakers to omit the

copula based on number. Somehow, just the presence of a plural subject made participants more likely to omit the copula in production. This could mean that at least in the Low Group, when the participants had not fully acquired number on the V, they simply omitted the copula. Both FR and IH would predict this initial difficulty in number. However, unlike IH, FR would predict that this obstacle would be overcome by more highly proficient learners, and this is exactly what the data of this study shows, as the High Group did not perform statistically differently regardless of number. It is also worth noting that FR would predict that acquiring number in VSO word order would be more difficult than in SVO word order, since Arabic does not realize number on the verb when it precedes the subject. While the results from the syntactic environments are not supported by statistical analysis, it is worth pointing out that plural *wh-* items were the single worst category for the Low Group. Alternatively, this category was not at all problematic for the High Group, suggesting that they have overcome this obstacle.

The results from the tense and agreement errors in copula production are also better accounted for by FR. As discussed in Section 5.2, the Low Group made errors in agreements 14% of the time and errors in tense 14% of the time. The High Group made errors in agreement 6% of the time and errors in tense 8% of the time. Someone may point to this fact and argue that despite their performance on the GJT, even the High Group is still making errors in agreement in their copula production, and so this provides evidence that number is unacquirable. However, it is important to note that the rates of errors in tense and agreement are very similar, despite the fact that IH would predict that tense would be fully acquirable and number would be unacquirable. Therefore, IH does not account well for this data. Alternatively, FR would account for this data because the

theory states that both tense and number should be acquirable but that errors can persist even at higher levels of proficiency. These errors are attributable to morphological mapping problems rather than to syntax. A strong piece of evidence to support this point is the fact that in both groups, when the copula was produced in a non-target-like manner, the participants often changed the subject to match the form of the copula that they produced. The example given above is that the item "The girls were reading yesterday" was produced as "The girl was reading yesterday." For the High Group, 30/35 of the changes from "were" to "was" also included a changed subject that matched the verb. This is strong evidence that number agreement has been acquired. Moreover, 6/10 times that the past tense was changed to the present tense, the adverb was also changed to match the tense that was produced. This again supports the idea that the participants have acquired tense but that there is a hinderance not rooted in syntax that is causing them to produce the copula in a non-target-like manner. Therefore, it appears that Feature Reassembly better accounts for the tense/agreement errors encountered in this study. It does appear that the participants have acquired the feature of agreement, in contrast to the predictions of IH, and that errors in agreement are due to a lack of proficiency in mapping agreement in a target-like way.

Lastly, it is worth emphasizing one last time that at the heart of difference between IH and FR is the fact that IH predicts that some features are unacquirable whereas FR predicts that all features are acquirable. In support of the theory that all features are acquirable, it is worth pointing out that five participants from the High Group received perfect scores on both portions of the GJT and the EI. It would be hard to argue that these participants have a representational deficit despite being able to consistently

produce the copula on the EI and recognize all items containing a copula as grammatical and all items missing a copula as ungrammatical regardless of the conditions or syntactic environments presented in this study. Therefore, FR does a better job of accounting for copula omission and errors in Arabic-speaking EFL learners.

6.2 Limitations

With a grammaticality judgment task, it is always a concern that participants have a bias towards accepting items. While this may have even been the case for the GJT in this study and may help to explain the drastic difference in accuracy between accepting grammatical items and rejecting ungrammatical items, the results are still informative in the sense that the same bias would have held across all structures. Therefore, comparing the results of each structure (verb type, tense, number, and syntactic environment) within the items that require participants to accept/reject a grammatical item will show that even though participants tended to accept ungrammatical items, they broke this bias for particular structures. A bias towards accepting simply means that comparing the rate of errors across the AG, RU and EI does not hold as much weight, but this also holds true for the EI which cannot be said to be necessarily reflective of the rate of omission of free production. An additional limitation is that despite the writing task between the EI and the GJT, there may have been some sort of priming effect from the participants hearing the grammatical items before completing the GJT. Although the participants were asked to write a paragraph for each of the four questions, some wrote only a sentence or two.

Furthermore, when wanting to investigate performance errors, it goes without saying that free production data, in both written and spoken form, is the most authentic form of data collection in terms of types of errors and rate of errors. Therefore, the fact

that the methodology of this study relied on an elicited imitation task for data collection rather than free production is a major limitation. For instance, it is possible that the subject-auxiliary inversion items on the elicited imitation task had such high accuracy rates because of the saliency of the copula as the first word of the item. It may be the case that in free production, some participants may have omitted the copula relying only on tone to signal a question. Secondly, it is possible that the role of perception was a confounding variable in the elicited imitation. While it is generally the case that perception precedes production, it is possible that in the cases of copula omission on the elicited imitation, the participants failed to perceive the copula, which is arguably distinct from accurately perceiving the copula but failing to produce it. However, this study maintains that the elicited imitation was the most efficient method of collecting quasi-production data across a wide variety of structures. While the methodology of this study may not have allowed for the discovery of all possible environments in which copula omission takes place, nor allowed for an accurate prediction of the rate of copula omission by Arabic speakers learning English, it did allow for a comparison of rate of omission between verb type, tense, number, and syntactic environment within each task as well as an insight into the types of copula errors that Arabic speakers learning English make beyond just those of copula omission. Based on the results of this study, future research (to be discussed in more detail in Section 6.3) has a starting point for identifying the sort of structures that would be most beneficial to hone in on and then collect naturalistic data for these particular structures in a much more economical way.

Another limitation of this study is the number of items for each structure. While there were enough items in tense, verb type, and number for statistical analysis, there

were not enough items in each syntactic environment to achieve any sort of statistical significance. This study aimed to investigate a large variety of structures in order to compensate for the fact that no spoken naturalistic data was collected. However, this had to be balanced with both participant fatigue and limited participant access. The researcher had access to the participants for one hour and this allowed for completion of an elicited imitation of 48 items and a time-pressure grammaticality judgment task of 96 items after directions were given. Ideally, future research will select a smaller number of structures to focus on based on the results of this study. This limitation also applies to the analysis of the tense, agreement, and double copula errors in the elicited imitation. A larger sample of production data whether it be EI or naturalistic would allow for a better picture of the role of tense and agreement in the acquisition of the English copula which in turn may give more insight to the acquisition of tense and agreement in general.

6.3 Directions for Future Research

First and foremost, future research should be conducted to address the limitations of this study. Most importantly, a study utilizing free production that verifies the results of this study would be ideal. First, the free production data would be able to support to refute the argument of this study that both groups performed so accurately on subject-auxiliary inversion items on the EI due to the markedness of the sentence-initial copula. It would be interesting to compare how often Arabic learners of English omit the copula in the subject-auxiliary inversion structure in free production. Furthermore, naturalistic data would allow for an accurate picture of at what rates the copula is omitted under different conditions such as tense, number, verb type, syntactic environment, etc. This is something the EI simply did not allow for.

Moreover, as mentioned above, future research would ideally use the results of this study to narrow the focus of their items. Specifically, it was found that *wh-* plural items were problematic on all tasks for both groups, so they should be of utmost importance. Furthermore, embedded items were particular problematic in both the EI and RU. Therefore, it is possible a lot could be learned from further investigation of copula omission in embedded structures.

Lastly, it would be ideal if future research could replicate the intent of this study with participants who had a wider range of proficiency levels. Using two adjacent proficiency levels does not give as much insight into the progression of development as would using more than two proficiency levels or two non-adjacent proficiency levels. In addition, it would be interesting to compare Arabic-speaking English learners to other learners whose L1 does contain the copula. This would allow us to investigate whether or not performance-based copula omission occurs in all EFL learners or only those whose L1 does not contain the copula.

6.4 Conclusions

The participants of this study completed a grammaticality judgement task and an elicited imitation task to investigate whether copula omission by Arabic speaking English as a second language learners is a representational or performance deficit, as evidenced by whether the Interpretability Hypothesis or Feature Reassembly best accounts for the data of study. Feature Reassembly is best able to account for the data of this study. The Low Group performed better on past tense items than present tense items on the RU, better on main verb items than auxiliary items on the EI, and better on the singular items than the plural items on the AG and EI. All of these results are predicted by Feature

Reassembly as constrained by Slabakova (2009). Moreover, the statistically significant deficits in the Low Group were overcome by the High Group, again predicted by FR and not IH. Additionally, the descriptive results of this study such as the omission of *-ing*, the doubling of the copula, and the tense and agreement errors on the copula are also predicted by FR, suggesting errors of copula omission by highly proficient Arabic speakers learning English are performance-based, not representational, in nature.

WORKS CITED

Abu Ghararah, A. H. (1989) Syntactic errors committed by Arabic EFL learners. *Linguistic Communication, 1*(2), 112-123.

Al-Buainain, H. (2009). Students' writing errors in EFL: A case study. *Journal of Faculty of Education, 19*(1), 311-351.

Al-Kasimi, A., Topan, F. and Khan, Y. (1979). Error analysis of Saudi freshmen's English. Research sponsored by Educational Research Centre, Riyadh University, Saudi Arabia, Riyadh.

Al-Muarik, S. A. (1982) Error analysis and English learning strategies among intermediate and secondary school students in Saudi Arabia. Unpublished Ph.D. Indiana University.

Al-Shayban, A.S. (2007). Copula omission by EFL Arab Learners. Master's thesis, Colorado State University.

Al-Zahrani, M.S. (1993). Copula omission and retention by Arab learners of English. M.A. Thesis, University of South Carolina at Columbia.

Ambridge, B. and Rowland, C. (2013). Experimental methods in studying child language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*(2), 149-168.

Aoun, J.E., Benmamoun, E., & Choueiri, L. (2010). *The Syntax of Arabic (Cambridge Syntax Guide series)*. The United Kingdom: Cambridge University Press.

Asfoor, A. A. (1978). An analysis of selected errors of Arabic speakers learning English. Unpublished Ph.D. University of Colorado.

Assubaiai, S. H. (1979) Empirical bases for interlanguage of Arab students learning English. Unpublished Ph.D. Georgetown University. Washington D.C.

Bakir, M. (1980). *Aspects of Clause Structure in Arabic.* Ph.D. Dissertation, Indiana University, Bloomington.

Basilico, D. (2003). The topic of small clauses. *Linguistic Inquiry, 34(1), 1-35.*

Beck, R. D. (1979). An error analysis of free composition. *Teachers of English, 30,* 20-26.

Beck, M.L. (1997). Viruses, parasites and optionality in L2 performance. Paper presented at the Second Language Research Forum, Michigan State University, October 1997.

Belikova, A. and White, L. (2009). Evidence for the fundamental difference hypothesis or not? *SSLA, 31,* 199-223.

Benmamoun, E (2000). *The Feature Structure of Functional Categories: A comparative study of Arabic dialects.* New York, NY: Oxford University Press.

Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistics Analysis, 20,* 1-49.

Bley-Vroman, R. and Chaudron C. (1994). Elicited imitation as a measure of second-language competence. In *Research Methodology in Second Language Acquisition*. Elaine Tarone, Susan Gass and Andrew Cohen (Eds.). Lawrence Erlbaum Associates: Hillsdale, NJ.

Bley-Vroman, R. (2009). The evolving context of the fundamental difference hypothesis. *SSLA, 31,* 175-198.

Bown, J., Burdis, J. and Cox, T. (2015). Exploring proficiency-based vs. performance-

based items with elicited imitation assessment. *Foreign Language Annals, 48*(3), 35-371.

Burdis, J. (2014). Elicited imitation as a predictor of language proficiency for learners of Russian (Unpublished doctoral dissertation). Brigham Young University, Provo, UT.

Chalikandy, M.A. (2015) Omani undergraduate students' writing errors. In *Methodologies for Effective Writing Instruction in EFL and ESL Classrooms.* (Eds.) Rahma Al-Mahrooqi, Vijay Singh Thakur and Adrian Roscoe. IGI Global: Hershey, PA

Cho, J.H. (2012). Remapping nominal features in the second language. (Unpublished doctoral dissertation). University of Iowa, Iowa City, IA.

Chomsky, N. (1995). *The Minimalist Program.* Cambridge, MA: MIT Press.

Clahsen, H. and Felser, C. (2006). Continuity and shallow structures in language processing. *Applied Psycholinguistics, 27,* 107-126.

Dubinsky, S. and Tasseva-Kurktchieva, M. (2009). The distribution of subjects and predicates in Bulgarian: An (EEP) V-feature account. *Formal Approaches to Slavic Linguistics, 17,* 237-252.

Eisenbeiss, S. Production methods in language acquisition research. In E. Blom and S. Unsworth (Eds.), *Experimental Methods in Language Acquisition Research* (pp. 11-34). Philadelphia, PA: John Benjamins Publishing.

El-Badarin, M. N. (1983). Transfer, strategies, and structural complexity in the acquisition of English syntax by Arabic speakers. Unpublished Ph.D. The University of Texas at Austin

Ellis, R., Gass, S., Godfroid, A., Jung, S., Loewen, S., and Park, J-H. (2015). Timed and untimed grammaticality judgments measure distinct types of knowledge. *Studies in Second Language Acquisition, 37,* 269-97.

Eubank, L. (1993/94). On the transfer of parametric values in L2 development. *Language Acquisition, 3,* 183-208.

Fassi Fehri, A. (1993). *Issues in the Structure of Arabic Clauses and Words*. The Netherlands: Kluwer Academic Publishers.

Gellert, A. S., & Elbro, C. (2013). Cloze tests may be quick, but are they dirty? development and preliminary validation of a cloze test of reading comprehension. *Journal of Psychoeducational Assessment, 31*(1), 16-28.

Graham, R., Lonsdale, D., Kennington, C., Johnson, A., & McGhee, J. (2008). In Proceedings of the 6th International Language Resources and Evaluation Conference. Marrakech, Morocco: European Language Resources Association.

Guttierez, X. (2012). Implicit Knowledge, Explicit Knowledge, and Achievement in Second Language (L2) Spanish. *Canadian Journal of Applied Linguistics, 15*(1), 20-41.

Hale, K. (1996). Can UG and L1 be distinguished in L2 acquisition? *Behavioral and Brain Sciences, 19,*726-728.

Hawkins, R. and Chan, Y. (1997). The partial availability of Universal Grammar in second language acquisition: A missing uninterpretable feature account. *Second Language Research, 22,* 269-301.

Hawkins, R. & Liszka, S. (2003). Locating the source of defective past tense marking in advantage L2 English speakers. In van Hout, R. et al. (Eds.) *The Lexicon-Syntax Interface in Second Language Acquisition,* 21-44. Amsterdam: John Benjamins.

Haznedar, B. and Schwartz, B. (1997). Are there optional infinitives in child L2 acquisition? In E. Hughes, M. Hughes & A. Greenhill (Eds.), *Proceedings of the 21st Annual Boston University Conference on Language Developmnt*, 21, 257-268. Somerville, MA: Cascadilla Press.

James, C. (2013). Errors in Language Learning and Use: Exploring Error Analysis. New York, NY: Routledge.

Kambal, A. (1980). An analysis of Khartoum University students' composition errors with implications for remedial English in the context of Arabicization. Unpublished Ph.D. The University of Texas at Austin

Kasem, A. (2000). The acquisition of the English copula by native speakers of Lebanese Arabic. In Z. Ibrahim, S. Aydelott & N. Kassabgy (Eds.) Diversity in Language: Contrastive studies in Arabic and English theoretical and applied linguistics. Cairo: The American University in Cairo Press.

Lardiere, D. (1998). Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second Language Research, 14*(4), 359-375.

Lardiere, D. (2004). Representation and processing really are stuck with each other. *Studies in Second Language Acquisition, 26*(3), 461-467.

Lardiere, D. (2008). Feature assembly in second language acquisition. In J. Liceras, H. Zobl & H. Goodluck (Eds.), *The Role of Formal Features in Second Language Acquisition* (pp. 106-140). New York: Lawrence Erlbaum Associates.

Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research, 25*(2), 173-227.

Lee, J. (2003). Cognitive and linguistic perspectives on the acquisition of object pronouns

in Spanish as a second language. *Spanish Second Language Acquisition: State of the Science,* 98-129.

Matsushita, H., Lonsdale, D., & Dewey, D. (2010). Japanese elicited imitation: ASR-based oral proficiency test and optimal item creation. *Corpus, ICT, and language education,* 161-172. Glasgow: University of Strathclyde Publishing.

Millard, B., & Lonsdale, D. (2011, March). Developing French sentences for use in French oral proficiency testing. Paper presented at the Linguistic Symposium on Romance Linguistics, Ottawa, Canada.

Mouchaweh, L. (1986). De la syntaxe de petites propositions. Doctoral dissertation, Universite de Paris.

Ortega, L. and Wu, Shu-Ling (2013). Measuring global oral proficiency in SLA research: A new elicited imitation test of L2 Chinese. *Foreign Language Annals, 46*(4), 680-704.

Perry, F.L. (2005). *Research in Applied Linguistics.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Prévost, P. and White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research, 16,* 103-133.

Ramchand, G. and Svenonius, P. (2008). Mapping a parochial lexicon onto a universal semantics. In Biberauer, M.T., editor, *Limits of Syntactic Variation.* John Benjamins: Amsterdam, 219-245.

Rapaport, T. (1987) Copular, nominal and small clauses: A study of Israeli Hebrew. Doctoral dissertation, MIT, Cambridge, MA.

Rothstein, S. (1995). Small clauses and copula constructions. *Syntax and Semantics, 28,* 27-47.

Schulz, B. (2011). Syntactic creativity in second language English: *wh*-scope marking in Japanese-English interlanguage. *Second Language Research, 27*(3), 313-348.

Schwartz, B.D. (1990). Unmotivating the motivation for the fundamental difference hypothesis. In Proceedings of the 10th Meeting of the Second Language Research Forum. Eugene, Orgeon.

Schwartz, B.D. and Sprouse, R.A. (1996) L2 cognitive states and the full transfer/full access model. *Second Language Research 12*, 40–72.

Schwartz, B.D. and Song, H.S. (2009). Testing the fundamental difference hypothesis: L2 adult, L2 child, and L1 child comparisons in the acquisition of Korean *wh*-constructions with negative polarity items. *SSLA, 31,* 323-361.

Scott, M.S. and Tucker, G.R. (1974). Error analysis and English-language strategies of Arab students. *Language Learning, 24*(1), 69-97.

Sharma, V. A. (1981) Syntactic errors as indices of developing language proficiency in Arabic speakers writing English at the intermediate and advanced levels of English as a second language. Unpublished Ph.D. Bloomington, Indiana University.

Slabakova, R. (2008). The bottleneck hypothesis. *Meaning in the Second Language.* New York: Mouton de Gruyter.

Slabakova. R. (2009). Features or parameters: Which one makes second language acquisition easier, or more interesting to study. *Second Language Research, 25*(2), 313-324.

Slabakova, R. and Cho, J. (2015). A feature-based contrastive approach to the L2 acquisition of specificity. *Applied Linguistics,* 1-23.

Soltan, U. (2011). On issues of Arabic syntax: An essay in syntactic argumentation. *Brill's Annual of Afroasiatic Languages and Linguistics, 3,* 236-280.

Tabatabaei, O., & Mirzaei, E. (2014). Correlational validation of cloze test and C- test against IELTS. *Journal of Educational and Social Research, 4*(1), 345.

Thompson-Panos, K. and Thomas-Ruzic, M. (1983). The least you should know about Arabic: Implications for the ESL writing instructor. *TESOL Quarterly, 17*(4), 1983.

Tsimpli, I.M. and Roussou, A. (1991). Parameter resetting in L2? *UCL Working Papers in Linguistics, 3,* 149-169.

Tsimpli, I.M. and Dimitrakopoulou, M. (2007). The interpretability hypothesis: evidence from *wh*-interrogatives in second language acquisition. *Second Language Research, 23*(2), 215-242.

Tsimpli, I.M. and Mastropavlou (2007). Feature interpretability in L2 acquisition and SLI: Greek clitic and determiners. In Liceras, J., Zobl, H. and Goodluck, H. (Eds.) *The Role of Formal Features in Second Language Acquistiion.* Mahwah, NJ: Lawrence Erlbaum.

Vainikka, A. & Young-Scholten, M. (1996), Partial transfer, not partial access**.** *Behavioral and Brain Sciences,* 19, 744-745.

White, L., Valenzuela, E., Kozlowska-Macgregor, M., and Ingrid, Y.L. (2003). Gender

agreement in L2 Spanish: Evidence against failed features. *McGill Working Papers in Linguistics, 17*(2), 153-182.

White, L. (2003). *Second Language Acquisition and Universal Grammar.* Cambridge Textbooks in Linguistics.

White, L. (2009). Some questions about Feature Reassembly. *Second Language Research, 25*(2), 343-348.

APPENDIX A – TARGET ITEMS

Grammaticality	**Syntactic environment**	**Main vs. auxiliary verb**	**Past tense (Singular)**	**Present tense (Singular)**	**Past tense (Plural)**	**Present tense (Plural)**
Grammatical	*WH-question*	*Main*	What was the name of the movie last night?	What is the name of your favorite movie?	What were the dates of the exams last week?	How many pages is the essay?
Ungrammatical	*WH-question*	*Main*	What the name of the movie last night?	What the name of your favorite movie?	What the dates of the exams last week?	How many pages is the essay?
Grammatical	*S-AUX*	*Main*	Was Layla your math teacher last year?	Is David your best student this year?	Were these women your students last year?	Are these men your teachers today?
Ungrammatical	*S-AUX*	*Main*	Layla your math teacher last year.	David your best student this year?	These women your students last year?	These men your teachers this year?
Grammatical	*Embedded*	*Main*	I know that Ahmed was mad last week.	We know that Layla is absent today.	They know that the teachers were mad last week.	I know that the boys are sick today.
Ungrammatical	*Embedded*	*Main*	I know that Ahmed mad last week.	We know that Layla absent today.	They know that the teachers mad last week.	I know that the boys sick today.
Grammatical	*Negation*	*Main*	The test was not very easy last week.	The boy is not very sick today.	The students were not sick yesterday.	The teachers are not happy today.

Ungrammatical	*Negation*	*Main*	The test not very easy last week.	The boy not very sick today.	The students not very sick yesterday.	The teachers not happy today.
Grammatical	*WH-question*	*Aux*	What was Ahmed doing in class yesterday?	What is Layla reading in class today?	What were the girls doing in class last week?	What are the students reading in class today?
Ungrammatical	*WH-question*	*Aux*	What Ahmed doing in class yesterday?	What Layla reading in class today?	What the girls doing in class last week?	What the boys reading in class today?
Grammatical	*S-AUX*	*Aux*	Was John writing an essay last week?	Is David reading in class right now?	Were the boys walking to school last week?	Are the students writing an essay right now?
Ungrammatical	*S-AUX*	*Aux*	John writing an essay last week?	David reading in class right now?	The boys walking to school last week?	Students writing an essay right now?
Grammatical	*Embedded*	*Aux*	We know that the girl was talking last week.	They know that John is listening now.	We know that the boys were reading last night.	I know that the girls are writing now.
Ungrammatical	*Embedded*	*Aux*	We know that the girl talking last week.	They know that John listening now.	We know that the boys reading last night.	I know that the girls writing now.
Grammatical	*Negation*	*Aux*	The teacher was not teaching yesterday.	The girl is not reading a book right now.	The boys were not writing an essay last night.	The students are not listening today.
Ungrammatical	*Negation*	*Aux*	The teacher not teaching yesterday.	The girl is not reading a book right now.	The boys not writing an essay last night.	The students not listening today.

APPENDIX B – FILLERS

Pronoun/Case Fillers (Syllable count/word count)

	Past	**Present**
Him (Ungrammatical)	We think him was happy last week (8/7)	We think him is happy right now (8/7)
He (Grammatical)	We think he was happy last week (8/7)	We think he is happy right now (8/7)
Her (Ungrammatical)	They think her was sad last week (7/7)	They think her is sad right now (7/7)
She (Grammatical)	They think she was sad last week (7/7)	They think she is sad right now (7/7)

SVA Fillers (Syllable count/word count)

	Grammatical	**Ungrammatical**
I	I read one book every week (7/6)	I reads one book every week (7/6)
He	He does not like to eat pizza (9/7)	He does not likes to eat pizza (9/7)
She	He likes to play football at night (8/7)	He like to play football at night (8/7)
They	They listen to music every week (9/6)	They listens to music every week (9/6)
We	We drive to class every morning (8/6)	We drives to class every morning (8/6)
You	You have writing class every Tuesday (9/6)	You has writing class every Tuesday (9/6)
I	I listen to music every week (9/6)	I listens to music every week (9/6)
He	He drives to class every morning (8/6)	He drive to class every morning (8/6)
She	She does not like to eat hot dogs (9/7)	She does not likes to eat hot dogs (9/7)
They	They read one book everyday (7/6)	They read one book everyday (7/6)
We	We walk to the library each day (9/7)	We walks to the library each day (9/7)
You	You listen to music all day (8/6)	You listens to music all day (8/6)