2-1-2016

Effectiveness of noun phrase accessibility hierarchy in predicting order of difficulty of relative clauses for AFL learners

Mohamed Amin

The American University in Cairo

School of Humanities and Social Sciences

Effectiveness of Noun Phrase Accessibility Hierarchy in Predicting Order of Difficulty of Relative Clauses for AFL Learners

A Thesis

Submitted to

The Department of Applied Linguistics

in partial fulfilment of the requirements

for the degree of Master of Arts in

Teaching Arabic as a Foreign Language

by

Mohamed Amin

(under the supervision of Dr. Raghda El Essawi)

December 2016

The American University in Cairo

School of Humanities and Social Sciences

Effectiveness of Noun Phrase Accessibility Hierarchy in Predicting Order of Difficulty of Relative Clauses for AFL Learners

A Thesis Submitted by

Mohamed Amin

Submitted to the Department of Applied Linguistics

December 2016

In partial fulfillment of the requirements for
The degree of Master of Arts
in Teaching Arabic as a Foreign Language

has been approved by

Dr. Raghda El Essawi _____
Thesis Supervisor
Affiliation: The American University in Cairo
Date Dec 27, 2016

Dr. Zeinab Taha _____
Thesis first Reader
Affiliation: The American University in Cairo
Date Dec. 27. 2016

Dr. Dala Abou El Seoud _____
Thesis Second Reader
Affiliation: The American University in Cairo
Date Dec 27, 2016

Dr. Amira Agameya _____
Chair, Department of Applied Linguistics
Date Jan. 15, 2017

Dr. Nathaniel Bowditch _____ 17 JAN 2017
Dean, School of Humanities and Social Sciences
Date _____

To my dear family:

Mother, Shereef, Emad, Eman,

and the soul of my father

Amin Abd El-Hamid Mostafa

11 November 1941 – 28 January 1998

Acknowledgment

I would first like to express my sincere gratitude to my advisor professor Raghda El Essawi, the director of Teaching Arabic as a Foreign Language M.A. Program at the American University in Cairo. She consistently allowed this paper to be my own work, but steered me in the right direction whenever she thought I needed it. Her guidance helped me through the stages of conducting the research and writing of this thesis. I was really lucky to have a great supervisor like Dr. El Essawi.

I'm most grateful to my thesis committee: Professor Zeinab Taha, associated professor at the Department of Applied Linguistics, American University in Cairo, and professor Dalal Aboul Seoud, chair of the department of Arabic language instruction, American University in Cairo for their precious comments and advices, which helped widen my research from various perspectives.

I would also like to acknowledge the Leiden Institute for Area Studies, Leiden University that provided me the opportunity to join their Arabic program as an intern and gave access to the laboratory and research facilities.

My sincere thanks also go to Dr. Hossam Ahmed, lecturer in Leiden Institute for Area Studies, Leiden University for all the help and support he provided me.

I would also like to thank Dr. Soren Wichmann, lecturer in Leiden University, Centre of Linguistics, and Dr. Maria Parafita Couto, lecturer in Leiden University, Centre of Linguistics for their generous help and consultations.

I would also like to express my deepest appreciation to all professors in the Department of Applied Linguistics, who gave me the most wonderful learning experience in my life: Dr. Atta Gebril, the great scholar and lecturer, who consistently provided us with all his knowledge and expertise; Dr. Ashraf Abdou, one of the best instructors and persons I've ever met; and Dr. Reem Bassiouny, a great lecturer and instructor.

A very special gratitude goes to Mrs. Mariam Salah Eddin, not only for being my great English teacher, but also for being a great and kind person.

Finally, sincere thanks go to Dr. Ayman Eddakrouri, Ms. Sara Tarek, and Ms. Sara Aboul Goukh for their co-operation and assistance.

Abstract

This study investigates the difficulty order of relative clause structures in Arabic second language acquisition by checking the applicability of NPAH to predict such order using data taken from the *Arabic Learner Corpus* (ALC). In order to examine the hypothesis, quantitative methodology is used to determine the hierarchies of frequency and accuracy of RC structures used in non-native learners' written and spoken production, and to determine whether L1 transfer affects acquisition order or not. Results reveal that the hierarchy of frequency is different from the hierarchy of accuracy, and both are different from the order suggested in NPAH. Findings support the significant role of language-specific characteristics in deciding the order of acquisition. Findings also support the common reported results, indicating that subject relative clauses structure is more accessible than direct object relative clauses. Regarding research, these findings suggest focusing on distinguishing features of individual languages and the effect of these features in the acquisition order of syntactic structures. Regarding pedagogy, results suggest that the different order of acquisition of subject relative clauses and direct object relative clauses can be used in evaluating learners' proficiency levels.

TABLE OF CONTENTS

List of tables

List of figures

Transliteration Symbols

Letter/Vowel/Diacritic	Symbol	Letter/Vowel/Diacritic	Symbol
ـَ	a	س	s
ـِ	i	ش	ʃ
ـُ	u	ص	sˤ
ـاً	an	ض	dˤ
ـٍ	in	ط	tˤ
ـٌ	un	ظ	ðˤ
ـّ	Doubling the consonant	ع	ʕ
ء	ʔ	غ	ɣ
ا	a:	ف	f
ب	b	ق	q
ت	t	ك	k
ث	θ	ل	l
ج	dʒ	م	m
ح	ħ	ن	n
خ	x	ه	h
د	d	و (Consonant)	w
ذ	ð	و (Long vowel)	u:
ر	r	ي (Consonant)	j
ز	z	ي (Long vowel)	i:

Chapter 1

Introduction

1.1 Rationale of the study and statement of the problem:

This is a quantitative corpus-based study that aims at improving Arabic as a foreign language teaching programs, contributing to research in the Arabic second language acquisition, and to enriching data available across human languages regarding relative clause acquisition and typological universals. This is done by investigating the acquisition order of relative clause structures in Arabic as a second language learning as they appear in learners' production, considering the universal difficulty hierarchy suggested by keenan and Comrie (1977). The study addresses questions and issues that have been dealt with by (Eckman, Bell, & Nelson, 1988; Gass, 1979; Hyltenstam, 1984a; Ozeki & Shirai, 2007) and other researchers, but through different perspectives. Thus, the main contributions of this study are: (1) it provides data about a language that have not been examined, which is the Arabic language; (2) it uses a sample that is much larger than all works reported in literature; (3) it exclusively deals with learners' free production, which better reflects learners' interlanguage systems.

Using a corpus, which is an electronic computerized database of a language, provides an opportunity to examine a quite large sample in short time. Software tools have been developed to undertake statistical and numerical analyses through a corpus. These tools present data, such as the number of occurrence of a specific word and total number of words (word list tool), and they present data about the context in which a specific word or phrase

appears among the corpus (concordance line tool).

The corpus employed in this study is a learner corpus - namely *Arabic Learner Corpus* (ALC). Hunston mentioned learner corpus as one of types of corpora. She defined it as "a collection of texts – essays, for example - produced by learners of a language" (Hunston, 2002, p. 15). This type of corpora aims to spot the differences between groups and individuals of the learners. It also allows researchers to identify the differences between the language of learners and the language of native speakers.

1.1.1 Main constructs of the study

This section briefly introduces the main principles and constructs related to the study. These principles and constructs are explained in details in chapter two.

Acquisition order

Order of acquisition or difficulty order is of great importance in second language acquisition research because of the useful information it provides regarding methodology of teaching and teaching materials (Nielsen, 1997). Hence, highlighting the difficulty order of complex syntactic structures, such as relative clause, will contribute to achieving better understanding of the process of acquisition of targeted structures in specific and to second language acquisition in general.

Typological universals

Typological universals research developed as a result of the emergence of universal grammar theory, which constituted with language transfer the two main theoretical views in language

acquisition.

Language transfer, which was suggested by behaviourism psychological theory, was considered as the main factor affecting second language learning (Noor, 1994). Behaviourists, who supported the exclusive role of L1 transfer, suggested that similar grammatical structures would be easy and would be acquired first, while the different ones would be difficult and, in turn, would be acquired later.

Alternatively, Chomsky (1965) suggested that there is an innate faculty in human's brains that facilitates acquiring languages, and that this innate faculty includes a set of grammatical principles existing in all languages and delimit the number of possible alternatives for constructing syntactic structures. This set of principles was called universal grammar (UG) (White, 2003).

Typological universals are conceived as the manifestation of these deep linguistic principles suggested by universal grammar. It has provided a lot of data proving the existence of underlying linguistic principles shared universally across human languages (Culbertson, 2012). Using data of wide range of human languages, research on typological universals induced features of different syntactic structures that are universally applied through human languages. Hence, the innate faculty theory suggested that UG principles are the main factors affecting the order of acquisition, and typological universals provided information about the surface structures underlying these principles and how they work.

Markedness

Research on typological universals revealed that structures differ in their distributions through languages. The notion of markedness basically means categorizing structures to marked or unmarked. Unmarked structures are those ones that widely exist in human languages, while marked structures are the ones that are less common universally. Typological universals suggest that the unmarked structures are easier to be acquired while the marked ones are more difficult in acquisition. For example, plural form is more common, universally, than dual form. Thus, dual is more marked than plural, which means that dual form is more difficult than plural form (Braidi, 1999). This hypothesis has been borne out empirically (Gass, 1979; Hyltenstam, 1984).

Relative clause

Relative clause (RC) structures have been heavily relied upon in evaluating theoretical hypotheses about syntax acquisition, such as markedness. This is due to the fact that variation in their word orders, among human languages and within the same language, provides an opportunity to investigate how the differences in structures affect the difficulty of processing syntactic structures throughout human languages (Gibson & Wu, 2013). This significant role can be attributed to the universality of RC structure across languages, its distinguishing syntactic characteristics, and its high frequency in actual language use (Izumi, 2003).

Keenan & Comrie (1977) adopted a semantically-based definition of RC. According to this definition, RC should refer to a specific set of individuals, which may contain only one member, through two steps: Firstly, specifying a larger set of individuals; secondly,

restricting this set to a sub-set of the larger one. For example, in the RC (the girl that John likes), the noun phrase (NP) (the girl) specifies a larger set of individuals, which is the set of girls. The clause (That John likes) restricting this larger set to a sub-set, one-member set, which is this individual girl who is the beloved of John. In this sense, RC includes the head noun phrase (NP), which is relativized in the restricting clause. To clarify, the restricting clause in this definition is the one considered as the relative clause in traditional grammar, while the RC notion is extended here to include the NP that is relativized (also called the head noun). Furthermore, a relative pronoun is not essential to have an RC. A clause like (the book I bought) is considered as RC, since its head NP (the book) specifies a large set (books) and the clause (I bought) restricts this set to a sub-set (the individual book that I bought).

Applying this definition to the Arabic language means that there are two possibilities of RCs.

First possibility is that the RC includes head noun, relative pronoun, and restricting clause, as explained in this example:

(Ex 1.1) قابلت الأستاذ الذي يدرس لابني

qa:baltu ʔalʔusta:ð ʔallaði: judarris liʔibni:

* I met the professor who teaches to my son

I met the professor who teaches my son.

This example includes head NP (the professor = الأستاذ), relative pronoun (who = الذي), and restricting clause (who teaches my son = الذي يدرس لابني).

Second possibility is that the RC includes only head NP (which is also the relative pronoun)

and restricting clause, as explained in this example:

(Ex 1.2) قابلت الذي يدرس لابني

qa:baltu ʔallaði: judarris liʔibni:

* I met who teaches to my son

I met the one who teaches my son.

This example includes head NP (who الذي), and restricting clause (teaches my son يدرس لابني). This means that relative pronouns in Arabic may function as NPs.

To elaborate, with application to Arabic, there are two types of relative pronouns in Arabic language: Relative particles and relative nouns. It is clear from the name that relative nouns are listed under the category of nouns in the Arabic parts of speech. This means that they may take the syntactic role of NP in a sentence as (Ex 1.2) showed. In the mentioned example, the relative noun (who الذي) functioned independently as NP, without referring to another preceding NP. Given that the Arabic language allows relative nouns to function as independent NPs and given that relative pronouns are not essential in the RC according to the definition followed in this study, sentences like the model presented in (Ex 1.2) are considered as RCs.

Moreover, the section of relative nouns conforms to the semantic conditions included in the RC definition presented by Keenan and Comrie (1977). To clarify, any noun in this section, when used as independent NP, refers to a specific group of individuals as the following examples show:

(Ex 1.3) الذي علمني هو أبي

ʔallaði: ʕallamani: huwa ʔabi:

*who taught me he my father.

The one who taught me was my father.

In this example, the relative noun (who الذي) specifies a large set of individuals, which is human masculine (men). Then, the clause (علمني) restricted this set to a one member-sub-set, which is that individual who taught the speaker.

(Ex 1.4) شاهدت اللاتي جئن بالأمس

ʃa:hadtu ʔalla:ti: dʒiʔna bilʔams

*I saw who came yesterday.

I saw (the women) who came yesterday.

The head NP here is the relative noun (اللاتي), and it specifies the large set of women groups. Members of this set are not individual women; members are the groups consisting of women. Any group of women is an individual member of this large set. Then, the clause (جئن بالأمس) restricted this large group to a sub-set, which is that specific group of women who came yesterday.

Similarly, the relative noun (ʔallaði:na الذين) refers to the large set of human groups, (man من) refers to the large set of individual humans, (ma: ما) refers to individual non-humans, etc.

Relative pronouns (ما) and (من) were not included in the study since Keenan and Comrie

(1977) built their model relying on only definite RCs, while these two relative nouns are indefinite.

Noun phrase accessibility hierarchy (NPAH)

One of the universals that are heavily examined in relation with second language acquisition research is the acquisition order of different strategies used in forming relative clause structures. The Noun Phrase Accessibility Hierarchy (NPAH), proposed by Keenan and Comrie (1977), is a hypothesis that investigated the different strategies used in relative clause structures in terms of the variety in syntactic positions of the relativized noun. Also, it suggested the order of acquisition of these structures based on their markedness level.

Following are the six patterns, ordered from the least marked structure to the most marked one, as explained in (Braidi, 1999, pp. 83–84):

1. SU = Subject, as in (The dog that bit the man ….)
2. DO = Direct object, as in (The man that the dog bit …)
3. IO = Indirect object, as in (The girl that I wrote a letter to …)
4. OBL = Oblique, meaning object of preposition, as in (The house that I talked to you about …)
5. GEN = Genitive as in (The family whose house I like …)
6. OCOMP = Object of comparative, as in (The woman that I'm taller than …)

The collected data showed that these different positions are not equally distributed among languages, but they form a hierarchical model of markedness according to the distribution of each structure through human languages. Using the third criterion of markedness hypothesis, the more common structure (most widely distributed) is considered as less marked structure.

Being less marked means being more accessible, and in turn, easier to be acquired. Therefore, Keenan and Comrei (1977) gave this model the name “Noun Phrase Accessibility Hierarchy” (NPAH). Data collected from about 50 human languages resulted in the following hierarchical model:

SU > DO > LO > OBL > GEN > OCOMP.

The sign ">" means “is more accessible than.”

Pronoun retention

Pronoun retention refers to the fact that some structures of relative clause contain a personal pronoun whose referent is the relativized noun. For example, the English sentence (I read the book which I bought) does not include a pronoun after the verb “bought” referring to the head noun “the book.” When translating this sentence to Arabic it will be (قرأت الكتاب الذي اشتريته = qaraʔtu ʔalkitaːb ʔallaði: ʔiʃtarajtuhu = *I read the book which I bought it), which includes a pronoun “ـه = it” whose referent is the head noun “الكتاب = the book.” This distinction is because the Arabic language applies the strategy of pronoun retention, while the English language does not.

Data investigated by Keenan and Comrei (1977, 1979) showed that retaining the pronoun is universally more common than deleting such pronouns, which means that applying pronoun retention is the unmarked strategy, and omitting the pronoun is the marked strategy.

Other hypotheses

Other theoretical proposals were suggested to determine and understand the order of

difficulty/acquisition of RCs structures through human languages. Perceptual difficulty hypothesis (PDH) proposed by Kuno (1974) and dependency locality theory (DLT) proposed by Gibson (1998, 2000) were two salient hypotheses used in examining the acquisition order of RC structures in second language. Both proposals relied on the complexity of cognitive processes running while comprehending/producing RC structures. DLT suggested the centre-embedding and incompatible word order as the source of complexity, and in turn difficulty in acquisition. Alternatively, DLT suggested more detailed criteria for calculating the complexity of structures, based on the distance between the basic NP and the basic VP in the main sentence.

It is noticeable that the perspective followed in these proposals has drawn the attention to the fact that language-specifics of the target language play a significant role in acquiring syntactic structures.

Both hypotheses were based on structures of fixed word order systems. Therefore, applying these hypotheses to Arabic language is not practical since the Arabic language is characterized by flexibility in word order.

1.1.2 Existing research

Many studies supported the validity of the typological universal principles regarding the relative clause acquisition to natural second languages (Doughty, 1991; Eckman et al., 1988; Gass, 1979; Hyltenstam, 1984; Pavesi, 1986). More recent studies reflected results denying the applicability of NPAH in predicting the difficulty order of RC acquisition (Gibson & Wu, 2013; Ju, 2014; Ozeki & Shirai, 2007). Other research indicated a significant role for language-specifics existing in the target language (Comrie, 2007; Lin, 2015; Marefat &

Rahmany, 2009). Finally, some studies reported mixed results supporting partial validity for NPAH (Izumi, 2003; Lin, 2015).

Proficiency level has been considered in many studies, for example (Marefat & Rahmany, 2009; Ozeki & Shirai, 2007), since differences (if exist) between levels of proficiency indicate differences in difficulty. For example, (Ozeki & Shirai, 2007) concluded that there is no difference in difficulty between SU, DO, and OBL because all these positions appeared in low proficiency levels as they appeared in higher proficiency levels.

This study investigates the effect of proficiency level through comparing the hierarchies of frequency and accuracy across two levels of the high school: The first year and the fourth year. The aim of this question is to provide a comprehensive image about the order of difficulty by considering as many variables as possible. Additionally, this question aims to explore changes happening as results of developing proficiency. These changes (if exist) may reveal the procedure in which the acquisition of RC structure is developed, and how this procedure relate to the hypothesis of NPAH and markedness. Finally, absence of differences between the two levels of proficiency confirms the generalizability of results - as they cannot be attributed to other factors.

Moreover, researchers in recent studies were careful to distinct between comprehension and production (Comrie, 2007). Additionally, it was noticeable that researchers used to collect samples of one communication mode for each study: Written (Gass, 1979; Marefat & Rahmany, 2009) or spoken (Jeon & Kim, 2007; Ozeki & Shirai, 2007).

Aiming to achieve a complete picture concerning the acquisition order of RC, the study also considers the variable of the mode of communication (written & spoken), especially that

previous studies used to investigate each mode apart from the other. Results of the study may indicate whether the mode of text reflects different difficulty order or not. If the answer is no, this means that results revealed in the study really reflect the ASL acquisition process and cannot be attributed to specific features related to the mode of communication.

Hence, this study considers all these factors: proficiency, mode, and production focus.

As for research on Arabic language, literature - to the best of my knowledge - lacks for research concerning acquisition order of RC. It has research regarding types of structures of RC in the Arabic language (Alotaibi & Borsley, 2013), while the research regarding acquisition order is about agreement (Nielsen, 1997).

1.1.3 Importance of the study and statement of problems

The field, however, is not without gaps. To begin with, to the best of my knowledge, the literature lacks any study that considers the applicability of NPAH predictions of difficulty order of the acquisition of relative clause patterns to the Arabic language.

Moreover, research in typological universals and syntax acquisition resulted in contrasting findings and interpretations. Some findings confirmed the applicability and universality of the NPAH hypothesis (Gass, 1979; Hyltenstam, 1984) while other findings firmly denied this applicability (Ozeki & Shirai, 2007).

From a pedagogic perspective,

Acquisition order has been a key issue in the research on second language acquisition for years, and with good reason. If it can be shown that learners proceed in a specific order when

learning a second language, it will not only provide teachers with guidelines as to the sequencing of teaching material, but furthermore, it will have far-reaching consequences as to how we could make second language learning more efficient (Nielsen, 1997, p. 49).

Therefore, a study that highlights the order of acquisition of relative clause structures in Arabic as a second language is expected to help improving ASL teaching programs by providing needed knowledge regarding the learners' interlanguage system. It is also expected to highlight a nearly blind area of study (acquisition order of AFL RCs) and to add more knowledge to more global areas of research (acquisition order in SLA, and typological universals).

1.2 Hypothesis and research questions:

As Gass (1979) proposed in her study, NPAH hypothesis can be proven if empirical research shows that learners produce the more accessible patterns (higher patterns in the hierarchy) more frequently and more accurately than the less accessible patterns (lower patterns in the hierarchy). As mentioned before, the study considers proficiency level and mode of communication, so the effect of these variables will be investigated. However, the study relies on the high school year as indicator of proficiency level because there is no data available about proficiency level of learners whose productions are investigated. Furthermore, Braidi (1999) pointed to the fact that confirming the NPAH hypothesis requires evidence that learners would apply pronoun retention (applying retention is unmarked while not applying is marked) strategy regardless of the existence of this strategy in their L1 backgrounds. This is because pronoun retention is unmarked. Hence, The first and second research questions in this study address NPAH hypothesis, while the third question examines

the role of L1 transfer against the concept of markedness regarding pronoun retention. Research questions are as follows:

1. What is the hierarchy in frequency and accuracy of AFL students' production of the six Arabic RC structures forming NPAH?
2. Is there any difference in hierarchies of frequency and accuracy between the first year and the fourth year of high school learners?
3. Is there any difference in hierarchies of frequency and accuracy between written production and spoken production?
4. Is there any significant difference in the rate of errors regarding pronoun retention between different groups based on L1 backgrounds?

1.3 Important definitions:

Universal grammar is

> the theory of the predetermined linguistic mechanisms. As it appears to be the case that what is predetermined is a uniform characteristic of the species, UG is in a sense a theory of linguistic universals, but only of the universals that are biologically necessary, not of the accidental universals that can occasionally arise as a historical contingency(Rizzi, 1989, p. 70).
>
> UG thus is an abstract characterisation of the notion possible human language, specifying what can vary and what remains constant across languages (Rizzi, 1989, p. 70).

Typological universals are "statements of structural dependencies that hold to varying degrees across a wide range of languages." (Braidi, 1999, p. 81).

L1 transfer refers to that

> individuals tend to transfer the forms and meanings, and the distribution of forms and meanings of their native language and culture to the foreign language and culture—both productively when attempting to speak the language and to act in the culture, and receptively when attempting to grasp and understand the language and the culture as practiced by natives (Lado (1957) as cited in Gass & Selinker, 2008, p. 89).

Interlanguage:

> At any stage of learning process, the oral and writing language competence of second language learners is a coherent system that is governed by internalized rules. This system and these rules are different from those of the native speakers. This system is referred to in the literature using different terms. The most common term used to describe this system is interlanguage and sometimes the language-learner language (Omaggio, 2001, p. 232).

Relative clause:

> We consider any syntactic object to be an RC if it specifies a set of objects (perhaps a one-member set) in two steps: a larger set is specified, called the domain of relativization, and then restricted to some subset of which a certain sentence, the restricting sentence is true.' The domain of relativization is expressed in surface structure by the head NP, and the restricting sentence by the restricting clause, which

> may look more or less like a surface sentence depending on the language (Keenan & Comrie, 1977, p. 63–64).

Pronoun retention means Retaining or copying the pronoun that the relative marker represents (Braidi, 1999).

Corpus is a systematic collection of authentically occurring texts (of written and/or spoken language). "Systematic" means that the structure and contents of the corpus applies certain extra linguistic principles, e.g. representation, size, sampling, and balance. In spite of a "corpus" can indicate to any non-randomly text collection, it is usually used in a narrower sense today, and is often only used to indicate to the computerized format of the systematic text collections (Hunston, 2002).

Concordance is a software tool that is used to "find every occurrence of a particular word or phrase" (O'keeffe, McCarthy, & Carter, 2007, p. 8). "The search word or phrase is often referred to as the 'node' and concordance lines are usually presented with the node word/phrase in the center of the line with seven or eight words presented at either side" (O'keeffe et al., 2007, p. 8).

Types and tokens, as explained in (Hunston, 2002), refer to two aspects of word counts in a corpus. Types are the word count of words contained in a corpus regardless how many times each word occurs in the corpus. Tokens are the word counts of all words occurring in a corpus, that is, any sequence of letters separated by a space or punctuation is counted as a token. Thus, if a word occurs four times in a corpus, it is counted as one type, and it is also counted as four tokens.

1.4 Operational definitions:

Relative pronouns in this study refer to the set of nouns constituting the category of definite relative nouns in Arabic language (ʔallaði: الذي , ʔallati: التي , ʔallaða:ni: اللذان, ʔallata:ni اللتان, ʔallaði:na الذين, ʔalla:ti: اللاتي, ʔalla:ʔi: اللائي).

Subject relativization (SU) includes the subject of active voice verbs, passive voice verbs, and nominal sentences, as shown in examples (Ex 1.6), (Ex 1.7), and (Ex 1.8) respectively:

(Ex 1.6) قابلت المؤلف الذي كتب الرواية

qa:baltu ʔalmuʔallif ʔallaði: kataba ʔalruwa:yah

I met the author who wrote the novel.

Relativized noun acts as the subject of the active voice verb (wrote).

(Ex 1.7) زرت المدينة التي أنشئت حديثا

zurtu ʔalmadi:nah ʔallati: ʔunʃiʔat ħadi:θan

I visited the town which was built recently

Relativized noun functions as the subject of the passive voice verb (was built).

(Ex 1.8) أفضل الفيلم الذي هو جاد

ʔufadˁil ʔalfilm ʔallaði: huwa dʒa:d

I prefer the movie which is serious.

Restricting clause is a nominal sentence (he is serious هو جاد), and the relativized noun functions as the subject of the sentence.

In case of relative clause consisting of a prepositional phrase, this study treats it as a nominal sentence whose subject is omitted as in the following example,

(Ex 1.9) الكتاب الذي عندي

ʔalkita:b ʔallaði: ʕindi:

The book which is my own

It actually means (ʔalkita:b ʔallaði: huwa ʕindi: الكتاب الذي هو عندي), which means that it will be counted as subject position.

Indirect object relativization (IO): An object is considered as IO if the verb of the restricting sentence is ditransitive, meaning that the verb may have two objects (with or without prepositions), so the first one is counted as IO.

(Ex 1.10) هذا هو الرجل الذي أعطيته الكتاب

haða huwa ʔarradʒul ʔallaði: ʔaʕtˤajtuhu ʔalkita:b

This is the man who I gave the book to

The restricting clause includes the ditransitive verb (gave). There are two objects, second one is the direct object (the book) as it is affected directly by the verb. First object, which the pronoun referring to the NP (the man), is the indirect object. The relativized noun here is (the man), so such RC is counted as IO position.

Genitives (GEN): Genitives in this study are defined according to the Arabic grammar, so some structures considered as OBL in English (such as with = مع, above = فوق, under = تحت) will be considered as GEN in the study, as these words are categorized under a specific set of

words called locative adverbs (ظروف الزمان وظروف المكان), which is a sub-category of nouns. More information about locative adverbs is available in Ryding (2005).

Proficiency level: In this study, the educational year of high school (year 1 & year 4) are used to indicate proficiency level since the corpus does not provide any information about learners' levels according to The American Council on the Teaching of Foreign Languages (ACTFL) or The Common European Framework of Reference for Languages(CEFR).

Adverbial noun refers to a special type of Arabic nouns that function as markers for position or time, such as (such as with = maʕa = مع, above = fawqa = فوق, under = taħta = تحت). Some of these nouns are considered as prepositions in English (i.e. with = مع)

Adverbial noun phrase refers to "Idafa" structure that includes an adverbial noun.

Full corpus means the corpus including the entire production of all **non-native** learners, including all mother tongues, all educational levels, and both modes of communications (written and spoken). However, this corpus (full corpus) is also considered as a sub-corpus of the *Arabic Learner Corpus* (ALC).

Written corpus is the corpus including the written production of non-native learners. However, this corpus (written corpus) is also considered as a sub-corpus of the full corpus.

Spoken corpus is the corpus including the spoken production of non-native learners. However, this corpus (spoken corpus) is also considered as a sub-corpus of the full corpus.

Year 1 corpus is the corpus including the entire production (written & spoken) for non-native learners enrolled in the first year of high school. However, this corpus (year 1 corpus) is also considered as a sub-corpus of the full corpus.

Year 4 corpus is the corpus including the entire production (written & spoken) for non-native learners enrolled in the fourth year of high school. However, this corpus (year 4 corpus) is also considered as a sub-corpus of the full corpus.

1.5 Abbreviations:

AFL = Arabic as a foreign language

ASL = Arabic as a second language

SLA = Second language acquisition

L1 = First language = Native language

CAH = Contrastive analysis hypothesis

RC = Relative clause

NPAH = Noun phrase accessibility hierarchy

NP = Noun phrase

AH = Accessibility hierarchy

SU = Subject relativiation, as in (The dog that bit the man ….)

DO = Direct object relativiation, as in (The man that the dog bit …)

IO = Indirect object relativiation, as in (The girl that I wrote a letter to …)

OBL = Oblique, meaning object of preposition relativiation, as in (The house that I talked to you about …)

GEN = Genitive relativiation as in (The family whose house I like …)

OCOMP = Object of comparative relativiation, as in (The woman that I'm taller than …)

HCs = Hierarchy constraints

ALC = Arabic Learner Corpus

+ retention = Apply pronoun retention

- retention = Do not apply pronoun retention

Chapter 2

Review of Literature

This chapter provides a detailed report of the information synthesizes research found in the literature relevant to the study. The chapter is divided to three main sections: Section one explains the theoretical framework of the study. Thus, it gives a preview of the main theories, constructs, and hypotheses relevant to the study. Section two offers a summary of results reported from empirical studies regarding the applicability of NPAH in predicting difficulty order of RC acquisition to different languages, giving detailed information about some salient studies. Third section draws the attention to some gaps in the field, which led to conducting this study.

2.1 Theoretical background

2.1.1 Acquisition order

Acquisition order is an important area of research that was heavily used in examining theoretical hypotheses, and in providing useful data for developing programs and teaching materials.

To elaborate, Brown (1973) and Dulay and Burt (1973, 1974a, 1974b, 1975) conducted their salient works on morpheme order studies to examine theories on the nature of language acquisition for children, namely the habitual proposal of behaviourists and innate faculty theory of mentalists. Later, Bailey, Madden, and Krashen (1974) investigated the same two

proposals for adult second language learners (as cited in Gass & Selinker, 2008). Proving the role of innate mental faculty in second language acquisition led to proposals determining specific order of acquisition of syntactic structures, such as question formation and relative clause. Therefore, checking the validity of these suggested proposals became a part of second language research, aiming to improve second language teaching.

For example, Eckman (1988) used the order of RC acquisition suggested by markedness and NPAH in determining the best generalization for instructions of relative clause in English second language teaching. The study aimed to examine the effect of structure-focused instructions on acquisition of RC structures. 36 students of ESL were divided to four equal groups. A pre-test of sentence-combining was administered to all learners. Results showed no significant differences between groups, which means that any difference appears in post-test is likely to be resulted from instructions. One of the four groups functioned as a control group, having no instructions. Each of the other three groups had intensive instructions on only one pattern: SU, DO, and OBL. Post-test was administered for the four groups. Results showed that number of errors of SU was less than DO, and DO was less than OBL for all groups, which support markedness hypothesis and NPAH. Moreover, results showed that the group that achieved the best scores in all structures was the one that had instructions on OBL. This means that instructions focusing on more marked positions lead to acquiring the less marked position. Thus, best generalization of instructions for RC is to concentrate on more marked structures.

This study is an obvious example of the importance of acquisition order for second language acquisition research and for second language teaching. It examined the hypotheses of NPAH

and markedness, highlighting the pedagogical implications based on these hypotheses.

2.1.2 Language transfer

Theories about second/foreign language acquisition have started off by highlighting the role of native language (and so L1 transfer) as the main factor affecting the acquisition of second language (Noor, 1994). For example, Lado (1957) maintained that the acquisition of second language is a task of overcoming native language (as cited in Noor, 1994). Behaviourists maintained the role of L1 transfer as the dominant factor affecting second language acquisition, so they suggested the contrastive analysis hypothesis (CAH) as an instrument predicting the difficulties of the acquisition of a specific second language for speaker of a specific native language. It is a comparative study determining the structural (syntactic, phonetic, etc.) differences between two languages in order to define the expected errors and difficulties, which leads to determining what structures that need to be taught and what structures that do not. According to the contrastive analysis hypothesis, similar structures are easy to be acquired and may not need to be taught, while the different structures are the ones making difficulties and need to be focused on in teaching (Gass & Selinker, 2008). Therefore, difficulty order and acquisition order are defined as follows: Similar is easy and in turn is acquired first, while different is difficult and in turn is acquired later.

Experimental research put the validity of CAH in question. It showed that different is not always difficult and similar is not always easy (Gass & Selinker, 2008).

2.1.3 Universal grammar

Alternatively, the concept of "universal grammar" (UG), presented by Chomsky (1965),

suggests that language acquisition is biased by set of constraints existing in human brain as innate faculty that facilitates language acquisition. Since the early years of research on UG, it focused on the relationship between UG and second language acquisition. In other words, it investigated if UG is applicable to only primary languages (L1) or also to second languages (White, 2003). A very important notion to be included in this area of research was "interlanguage." White (2003) explained the construct of interlanguage and its relationship with UG:

> In the late 1960s and early 1970s, several researchers pointed out that the language of second language (L2) learners is systematic and that learner errors are not random mistakes but evidence of rule-governed behavior (Adjémian, 1976; Corder, 1967; Nemser, 1971; Selinker, 1972). From this developed the conception of "interlanguage," the proposal that L2 learners have internalized a mental grammar, a natural language system that can be described in terms of linguistic rules and principles. The current generative linguistic focus on interlanguage representation can be seen as a direct descendent of the original interlanguage hypothesis. Explicit claims are made about the nature of interlanguage competence, the issues being the extent to which interlanguage grammars are like other grammars, as well as the role of Universal Grammar (UG).
>
> (White, 2003, p. 19)

Following the hypothesis that the interlanguage is governed by the innate faculty, principles of UG are the determining factor in deciding the acquisition order of language structures, not the L1 transfer as claimed by Behaviorists.

2.1.4 Typological universals

Another construct contributing to the theory of innate faculty is "Typological universals, which is conceived as the apparent phenomena that demonstrate the deep principles suggested by universal grammar (Culbertson, 2012).

Research in typological universals is based on inductive approach using a huge amount of data collected from a wide range of human languages. The term "typological universals" is taken from its work in classifying languages to different types, such as VSO (verb – subject - object word order), SOV (subject – object – verb word order). Furthermore, it investigates the structural dependencies associated with each type (Braidi, 1999). Comrie (1981) differentiates between two types of universals: Absolute universals (exist in all languages without exceptions) and tendency universals (have some exceptions) (as cited in Braidi, 1999). Similar to UG, research on typological universals did not stop on primary language, researchers also investigated the role of typological universals in second language acquisition.

2.1.5 Markedness

One important notion included in the study of typological universals is the notion of markedness. Different criteria are used to define the degree of markedness: Simplicity/complexity, frequency and distribution. First and second criteria are subjects to analysis through an individual language. For example, singular structure in English is simpler than plural structure. Therefore, in English language, singular is unmarked while plural is marked regarding to simplicity/complexity criterion. Similarly, singular in the English

language is used more frequent than plural, which means that singular is less marked than plural, in the English language, regarding the frequency criterion. The criterion that is relevant to typological universals research is the third one, namely distribution across languages. For example, plural form is more common, universally, than dual form. Hence, dual is more marked than plural (Braidi, 1999). Typological universals suggest that the acquisition order of different structures and different varieties of a specific structure is determined by its degree of markedness. In other words, less marked structures are early acquired than more marked ones.

This notion was the corner stone in determining whether typological universals have a role in second language acquisition or not, and the extent to which this role dominate the acquisition in comparison to L1 transfer. On one hand, research that relied on markedness revealed that typological universals don't have a dominant role in second language acquisition. On the other hand, reported results showed influence of L1 transfer, regarding the strategy of pronoun retention in particular (C. Doughty, 1991; Gass & Selinker, 2008; Hyltenstam, 1984; Pavesi, 1986). For example, Gass (1979) concluded that unmarked strategies were used more frequent and accurate than the marked ones, indicating that typological universals are the main factor deciding acquisition order. However, she pointed to L1 transfer as an intervening factor in using the strategy of pronoun retention in the more marked structures of relative clause in English second language learning. Pavesi (1986) reported that learners of English in informal learning context showed the same order of acquisition of RC patterns as learners of formal context. This supports the hypothesis of markedness since informal learners followed the model suggested by markedness though they did not have any instruction.

2.1.6 Relative clause

Relative clause is a syntactic structure that is characterized by variation in word order possibilities among human languages and within the same language (Gibson & Wu, 2013). Relative clause is also characterized by universally wide distribution and high frequent use in natural languages (Izumi, 2003). Due to these distinguishing syntactic features, relative clause structures were heavily relied upon in examining theoretical views, especially those about acquisition order. For example, it was used in examining noun phrase accessibility hierarchy (NPAH) as in (Gass, 1979b; Marefat & Rahmany, 2009; Ozeki & Shirai, 2007), markedness as in (Gass, 1979; Hyltenstam, 1984), dependency locality theory (DLT) as in (Lin, 2015), and perceptual difficulty hypothesis (PDH) as in (Lin, 2015).

2.1.7 Noun phrase accessibility hierarchy (NPAH)

Based on data collected from about 50 languages, Keenan and Comrie (1977) stated that, universally, there are six possible patterns of relative clauses, based on six different grammatical functions of the relativized noun in the RC. Following are the six patterns as explained in (Braidi, 1999, pp. 83–84) with the equivalent examples in Arabic:

1. SU = Subject, as in (The dog that bit the man ….)

   (الكلب الذي عض الرجل = ʔalkalb ʔallaði: ʕadˤdˤa ʔarradʒul)

2. DO = Direct object, as in (The man that the dog bit …)

   (الرجل الذي عضه الكلب = ʔarradʒul ʔallaði: ʕadˤdˤahu ʔalkalb)

3. IO = Indirect object, as in (The girl that I wrote a letter to …)

   (الفتاة التي كتبت رسالة إليها = ʔalfata:tu ʔallati: katabtu risa:lah ʔilajha:)

4. OBL = Oblique, meaning object of preposition, as in (The house that I talked to you about …) = (البيت الذي حدثتك عنه = ʔalbajtu ʔallaði: ħaddaθtuka ʕanhu)
5. GEN = Genitive as in (The family whose house I like …)

   (الأسرة التي أحب بيتها = ʔalʔusrah ʔallati: ʔuħibbu bajtaha:)
6. OCOMP = Object of comparative, as in (The woman that I'm taller than …)

   (المرأة التي أنا أطول منها = ʔalmarʔah ʔallati: ʔana: ʔatˁwal minha:)

The collected data showed that these different positions are different in their distribution through human languages. According to markedness, the more common pattern the less marked it is. Keenan and Comrei (1979) ordered the six structures in a hierarchy from the most unmarked to the most marked, considering that the less marked pattern is more accessible in acquisition. Thus, the model was given the name "Noun Phrase Accessibility Hierarchy" (NPAH), and is usually presented as follows:

SU > DO > LO > OBL > GEN > OCOMP.

The sign ">" means "is more accessible than."

Distribution of these suggested positions of noun phrases varies across languages. Meaning, not every single human language must allow all the six positions.

The model included what they called: Hierarchy constraints (HCs). These constraints were as follows:

1) all languages must apply subject relativization.

2) a language that applies a particular strategy should be able to apply all strategies preceding it in the hierarchy. For example, if a language applies the position OBL, it must

apply the higher position (less marked positions) in the hierarchy, namely IO, DO, and SU.

3) a language may cut-off applying the relative clause strategies at any point of lower positions of the hierarchy. This means that a language may apply only the first three positions (SU, DO, IO), and stop applying the lower three position. Another language may apply position 1-5 (SU, DO, IO, OBL, GEN) and don't apply the sixth position. For example, Arabic applies all the six positions of NPAH, while Fulani applies positions from 1-5 and do not apply the sixth position (OCOMP). Moreover, Maori language applies only the first four positions.

Another patterning issue revealed in the NPAH research is the pronoun retention, which refers to the fact that some structures of relative clause contain a personal pronoun whose referent is the relativized noun. The difference between Arabic and English languages provides good example to explain this issue.

(Ex 2.1) قرأت الكتاب الذي اشتريته

qaraʔtu ʔalkitaːb ʔallaði: ʔiʃtarajtuhu

*I read the book which I bought it

I read the book which I bought.

Arabic language applies the strategy of pronoun retention, so the Arabic RC includes the accusative third person pronoun which (ـه), which refers to the relativized NP (head NP), which is (الكتاب). In other words, the head NP was mentioned again in the restricting clause. This can be seen in the ungrammatical English translation, as it includes the pronoun (it).

On the other hand, the grammatical English translation does not include this pronoun, since English is a language that does not apply pronoun retention.

Data investigated by Keenan and Comrei (1977, 1979) proved that retaining the pronoun is universally more common than deleting such pronouns. Moreover, it was noted that the distribution of the phenomenon of pronoun retention within the hierarchy positions shows a tendency to be applied more in the most difficult – less accessible - positions.

Given that the NPAH is based on the notion of markedness and that simplicity/complexity is one criterion of markedness, NPAH suggests that each position in the hierarchy is more complex than its higher position and less complex than its lower position. For instance, the NPAH suggests that OBL position is more complex than IO and less complex than GEN. This can also be understood through the title of the hypothesis: Accessibility hierarchy. Accessibility refers to being more accessible in acquisition, which definitely means less complex and in turn less difficult.

Comrie (2007) pointed to a new version of NPAH, presenting it as follows:

Subject > direct object > other objects > genitive/possessor.

The version of 2007 gathered all objects, indirect object, object of pronoun, and object of comparison in one category (other objects). This can be interpreted as that these positions are three varieties of one syntactic category having the same difficulty level. It is worth mentioning here that Comrie presented this new version as his “own more recent thinking”, and he described it using the expression “essential intuition” (Comrie, 2007, p. 303).

On the other hand, Keenan (2014) kept using the basic version of NPAH.

This study follows the basic version since it still the one reported as result of analysing universal human languages data. Furthermore, results based on the old version can be compared to the new version, while the opposite cannot be done.

2.2 Applicability of NPAH predictions

NPAH used to be the common denominator in research on acquisition order of RC structure, including those studies that examined other hypotheses. Findings of studies concerning RC acquisition used to be compared to NPAH model (Gibson, 2000; Gibson & Wu, 2013; Ju, 2014; Lin, 2015; Marefat & Rahmany, 2009; O'Grady, Lee, & Choo, 2003). Moreover, many studies have been dedicated to examine the validity of markedness and NPAH hypotheses. Some studies focused on accuracy (number of errors or test score), such as (Eckman et al., 1988; Hyltenstam, 1984; O'Grady et al., 2003; Pavesi, 1986). Other studies relied on frequency (Lin, 2015; Ozeki & Shirai, 2007). There were also researches that considered both aspects (Gass, 1979). There were other important variables considered in acquisition order research: Proficiency level (Marefat & Rahmany, 2009; Ozeki & Shirai, 2007), and mode of communication (Jeon & Kim, 2007; Ozeki & Shirai, 2007).

Contrasting findings were reported in the literature regarding the applicability of NPAH and markedness in predicting difficulty order of RC acquisition in second language. (Doughty, 1991; Eckman et al., 1988; Gass, 1979; Hyltenstam, 1984; Pavesi, 1986) gave evidence of the applicability of markedness and NPAH. More recent studies reflected results denying the applicability of NPAH in predicting the difficulty order of RC acquisition (Gibson & Wu, 2013; Izumi, 2003; Ju, 2014; Ozeki & Shirai, 2007). Other findings reported mixed results supporting partial validity for NPAH. (Jeon & Kim, 2007; Lin, 2015; O'Grady et al., 2003).

Finally, some studies indicated a significant role for specifics existing in the target language (Comrie, 2007; Marefat & Rahmany, 2009; O'Grady, Lee, & Choo, 2003).

2.2.1 Studies supported the NPAH

One of the most cited works concerning the role of native language transfer in comparison with the role of typological universals was the study conducted by Gass (1979). Gass examined the acquisition of relative clauses of English language by speakers of nine different languages. Subjects of study were 17 English second language learners, who were native speakers of nine languages. These languages were Arabic, Chinese, French, Italian, Korean, Persian, Portuguese, Japanese, and Thai. She administered two tasks: grammaticality judgment and sentence combining. In the grammaticality judgment task, 29 sentences, including RCs, were given to participants. Thirteen sentences were formed correctly, while the other 16 included four types of errors: Relative clause marker omission, pronoun retention, wrong selection of the relative clause marker, and adjacency (separating the modified noun from the relative clause modifying it). In sentence combining task, participants were asked to combine two given sentences using relative clauses. A task of free writing was added to these two tasks in a larger study. In her discussion, Gass proposed two different areas of testing that can prove the AH hypothesis: Frequency and accuracy. The higher positions of the hierarchy should be more frequent and accurate in learners' production. Gass stated that the first hypothesis, regarding frequency, was proved (Gass, 1979, p. 339). Results of the sentence combining task gave the evidence of the second area - that learners produce the most accessible positions more accurately than the less accessible ones. Results of the grammaticality judgment task showed that L1 had no role in three types

of errors: Omission, wrong selection, and adjacency. Generally, the results supported the role of typological universals apart from any influence from L1, with only one exception, which was pronoun retention. This exception, pronoun retention, was not applicable in all instances. The transfer effect regarding pronoun retention was obvious only in the three more-marked position: OP, GEN, and OCOMP. Gass concluded that typological universals have the prominent role in the acquisition of relative clause structures of English as a second language, while the influence of L1 transfer is delimited by universal constraints.

Some limitations can be noticed in this study. Only 17 learners participated in the study, which is a small sample. Moreover, conclusions of the study were based on raw numerical data (total score of participants), while the statistical tests resulted in no significant differences.

Another important study was conducted by Hyltenstam (1984), who examined the validity of markedness in predicting the acquisition order regarding the pronoun retention strategy in forming relative clause, given that applying pronoun retention is the unmarked strategy while not applying it is marked. Subjects of the study were Swedish learners who are native speakers of four languages, two of which allow pronoun retention (Persian and Greek) and the other two do not allow pronoun retention (Spanish and Finnish). Results showed that all subjects did produce pronominal reflexes but to different extents according to the native language. Learners who were native speakers of Persian and Greek produced more pronoun retentions than Spanish-speaking and Finnish-speaking learners. These results proved that typological universals have basic influence in second language acquisition while the

influence of L1 only affects the extent to which this role (the role of typological universals) is applied.

Pavesi (1986) investigated the effect of learning context on RC acquisition order. Since typological universals suggested that acquisition order is governed by markedness, she presumed that both formal and informal learners should have had the same order of acquisition, which should be yielded to NPAH. She examined two groups of Italian-speaking subjects living in Britain. Group one consisted of 48 students of high school, representing the formal learning context. Group 2 consisted of 38 workers who had minimal amount of instructions and were exposed to English in their work places and homes. RCs were elicited from participants by asking them to describe characters appearing in pictures. Participants were given scores according to their accuracy in RC production. Results of both groups were generally compatible with NPAH; however, they had two inversions between IO and OBL and between GEN and OCOMP. Pavesi attributed these inversions to specific features in English syntax that are mutual between IO, OBL, and OCOMP, which made them almost one category.

When speaking about limitations, the study relied completely on elicited RCs, lacking to natural production.

Reporting exceptions is a noticeable phenomenon in studies that supported NPAH. (Gass, 1979a) and Hyltenstam (1984) reported an intervening influence of L1 transfer regarding pronoun retention. Moreover, (Hyltenstam, 1984) found that Spanish learners of Swedish language reverse the order of IO and OBL, which was considered unproblematic. The GEN position was the most salient and repeatedly reported in exceptions of NPAH predictions.

Gass (1979) found that learners responses to GEN RCs in sentence-combining task were more accurate than responses of DO and IO positions. In addition, Doughty (1991) reported that learners demonstrate earlier acquisition of GEN than IO. On the other hand, Pavesi (1986) found that GEN was acquired later than the most marked position OCOMP since findings showed more pronoun retention errors in GEN than OCOMP. One of the explanations suggested for this distinction of GEN is that salient language-specific peculiarities of the second language are acquired precisely. Both English and Swedish require a different relative pronoun (whose in English) that is exclusively used for GEN relativizations. Therefore, Gass (1980) concluded that universals determine the general outline of acquisition while other factors like L1 transfer and L2 specifics have influence on the aspects undetermined by the universals (as cited in Braidi, 1999). This means that typological universals have the main influence on acquisition, but they interact with other secondary intervening factors like language specifics and L1 transfer.

2.2.2 Studies rejected NPAH

Izumi (2003) examined three hypotheses of RC acquisition order - NPAH (Keenan & Comrie, 1977), PDH (Kuno, 1974), and SOHH (Hamilton 1994)- on English second language learning. 61 learners who were native speakers of 12 languages participated in the study. Data was taken from another study (Izumi 2000, 2002) that investigated the effectiveness of different instructional technique on RC acquisition. Three different tests were used in the study: Sentence-combining, interpretation (relating a given RC to one of several pictures) and grammatical judgement. As for NPAH, the study focused only on three

patterns: SU, DO and OBL. Results of three tests were not consistent, which led the researcher to conclude that NPAH was not accepted.

Ozeki and Shirai (2007) conducted two studies examining the applicability of difficulty predictions based on NPAH on Japanese second language learners' production. In the first study, they analysed, in terms of frequency - data of oral production of 90 learners of different proficiency levels, 30 of Mandarin Chinese, 30 of English, and 30 of Korean. Data are taken from learner corpus consisting of transcribed ACTFL oral proficiency interviews. This data was compared to data taken from similar interviews with 15 Japanese native speakers. Results showed that even participants of low proficiency levels produced relative clauses of positions DO and OBL, which indicates that these positions are not more difficult than SU.

The second study examined the accuracy through a sentence-combining task, given to 50 native speakers of Cantonese learning Japanese as a second language in intermediate and advanced levels. The materials used in the study included the three positions SU, DO, and OBL. Scores of the task regarding each structure showed that SU and DO had the same level of difficulty, while OBL was more difficult. Thus, they conclude that NPAH does not predict difficulty order of RC in Japanese.

Gibson and Wu (2013) conducted their study on Chinese RC processing from different perspectives. First, it focused only on native speakers, so it examined first language acquisition not second language. Second, it focused on comprehending RCs not producing. Reading pace was examined on 40 participants using software tool displaying a moving window while reading. Time spent, in reading each RC, was calculated by the software.

Results showed that participants read subject RCs slower than object RCs, which means that subject RCs were more difficult than object RCs. These findings oppose the priority of SU RCs hypothesized by NPAH for both L1 and L2.

Ju (2014) examined the applicability of NPAH and markedness to Korean native speakers and Korean second language learners. The study focused on checking differences in difficulty order of the different patterns forming NPAH between four groups, three of them are Korean second language learners (Chinese, Japanese, and English native speakers) and the fourth group consisted of Korean native speakers. Two experiments were conducted: One examined processing time in comprehending RCs during listening, and the other examined the reaction time needed before producing RCs in speaking activity based on visual stimulator, in which participants should have described a picture using RC. The study also aimed to examine whether NPAH and markedness are typological universals or the difficulty order of RC structures are language-specific matter. Results showed no significant differences between SU and DO RCs for all groups. Researcher concluded that NPAH and markedness do not apply for Korean first and second language acquisition, and so, NPAH and markedness are not universal.

2.2.3 Studies with partial support and mixed findings

Aiming to present an interpretation of the difference in accessibility between SU and DO, O'Grady et al. (2003) suggested a distance-based hypothesis (structural distance hypothesis) and examined RC acquisition for English-speaking learners of Korean language. They concluded that English-speaking learners of Korean prefer subjective relative clauses than objective relative clauses. They suggested that their hypothesis may interpret differences

between all structures forming NPAH, but Korean language was not the sufficient language for such research because it lacks IO and OCOMP positions. Regardless the hypothesis tested in the study, it considered the structures of NPAH, and its results agreed partially with it.

Jeon and Kim (2007) investigated the interaction between NPAH and typological characteristics of Korean language in Korean second language learning. As Korean language has two types of RCs, head internal and head external, the study investigated whether there were differences between the two types in applying NPAH. The study focused only on SU and DO positions. Data was taken from pre-test and post-test of a previous experiment (Jeon, 2004), which examined the effect of task-based instructions on different linguistic forms. Subjects were 40 English-speaking learners of Korean, ranged from high beginner to intermediate levels of Korean university language program in the United States. To elicit RC production, participants were given pictures with circles and were asked to describe the location of each circle in the picture. Scores of participants' accuracy in producing RCs were calculated for each structure (SU & DO). Frequency of each pattern was also calculated. Researchers concluded that SU had more advantage over DO in head-external RCs, whereas similar advantage was not proven in head-internal. They also pointed to that the advantage of SU in head-external may not be exclusively because of the syntactic form, but animacy of the head noun may have contributed to that result.

Lin (2015) tested three theories addressing acquisition order of RC: NPAH, PDH, and DLT. He examined natural production of Mandarin-speaking learners of English. Corpora of learners' production were built, including compositions collected from university students and high school students. The study relied on frequency of each RC structure in learners'

production. Results showed a full applicability of PDH, and partial success for NPAH and DLT. NPAH was proven to correctly predict the positions of GEN and OCOMP, but results showed a fail in predicting IO and OBL, that results reported that OBL>IO.

2.2.4 Language specific features

Some studies literature ended up with conclusion indicating the significant role of language-specifics in acquisition order of relative clause. NPAH was one of seven hypotheses tested by Marefat and Rahmany (2009) on native Persian learners of English. 39 participants had a comprehension test in which they were asked to identify the subject and object of each matrix sentence and the subject and object of the RC. Only SU and DO positions were examined, and they were confirmed to be compatible with NPAH. However, it was criticized because it did not explain the complete account of RC difficulty order since it did not consider the relation between the syntactic position of relativization and matrix sentence. As findings also supported SOHH which is based on word order system, researchers concluded that word order system affects RC acquisition in English second language, especially for learners whose L1 applies different word order.

Similarly, Jeon & Kim (2007) explained that Korean language has two types of RCs that differ in word order: Head-external RCs and head-internal RCs. They examined the compatibility of each type with predictions of NPAH (only regarding SU and DO) on English-speaking learners of Korean. They concluded that there are differences between the two types regarding NPAH predictions. The two types were as follows: Head-external revealed to be compatible with NPAH, while the head-internal RCs did not show a similar evidence.

On the other hand, some theoretical views imply a significant role for the specific features of each language in deciding the acquisition order of RC structures. For example, DLT, proposed by Gibson (1998, 2000) suggested that the extent of difficulty of a structure is determined by the intervening discourse items between the basic NP and VP of the sentence. This means that the word order system adhered by a language determines the difficulty level of each structure according to the number of intervening items required to produce this structure. Furthermore, Comrie (2007) clarified that the article of Keenan and Comrie (1977) did not deny the possibility of interaction between NPAH and other linguistic principles, and it pointed to the possibility that NPAH might be reflection of more fundamental psycholinguistic principles.

In conclusion, empirical research in the literature resulted in contradictory results: Totally confirming, partially accepting, and totally rejecting the NPAH hypothesis.

As for research on Arabic language, literature, according to my knowledge, lacks research concerning acquisition order of RC. It has research regarding types of structures of RC in the Arabic language (Alotaibi & Borsley, 2013), while the research regarding acquisition order is about agreement (Nielsen, 1997)

Chapter 3

Methodology

This chapter provides detailed information concerning research design, sample, instruments and tools, and procedures of data collection and analysis. It explains, step by step, how the study was conducted, giving the rationale of these frameworks and procedures, and how these processes have answered research questions of the study:

1. What is the hierarchy in frequency and accuracy of AFL students' production of the six Arabic RC structures forming NPAH?
2. Is there any difference in hierarchies of frequency and accuracy between the first year and the fourth year of high school learners?
3. Is there any difference in hierarchies of frequency and accuracy between written production and spoken production?
4. Is there any significant difference in the rate of errors regarding pronoun retention between different groups based on L1 backgrounds?

3.1 Research design

This study is a quantitative corpus-based study. As the study aims to generalize results to the population of Arabic as a second language learners, quantitative method based on large amount of data provided by corpus is needed.

Data was collected from the Arabic learner corpus (ALC) developed by Abdullah Alfaifi and Eric Atwell (2014).

The corpus used in this study can be considered as a convenient sample with a high level of representativeness. The corpus consists of written and spoken production of Arabic learners (native and non-native learners) in Saudi Arabia taken in the years 2012 and 2013. It includes 282732 words (tokens), 1585 texts (written and spoken) of 942 students from 67 nationalities and 66 different L1 backgrounds. 795 of these texts are produced by learners who are non-native Arabic speakers. Of these texts, 753 are written production and 42 are spoken production. Texts are collected from different educational levels: Pre-university (years 1-4 of high school) and university.

ALC online search allows running a search with different determinants, such as age, nationality, mother tongue, nativeness, mode of communication (written or spoken), level of education (year of school), etc. This tool helped in categorizing files, and, in turn, building a sub-corpus of non-native speakers and constituting groups of different L1 backgrounds, different modes of communication, and different education levels.

As for convenience, the corpus texts are available for download, which allows the researcher to use Wordsmith software tools in the analysis. Moreover, data base of the corpus provides all possible categorizations, which allowed the researcher to easily build a sub-corpus of non-native learners, which was needed to answer the first and second research questions. It also allowed building sub-corpora of written production, spoken production, first high school year written production, and fourth high school year written production, all of which are required to answer third and fourth research questions. Finally, it allowed the researcher to build two

sub-corpora of two groups of L1 backgrounds: Languages that apply pronoun retention, and languages that don't apply pronoun retention. These sub-corpora were needed to answer the fifth research question.

As for representativeness, the corpus includes a quite high number of texts and a wide range of L1 backgrounds. These high numbers of learners, nationalities, and L1 backgrounds provided a sample that is large in size and diverse in typology, which offered a quite representative sample of AFL.

The study examined all the RC sentences (about 1100) in the sub-corpus of non-natives to answer the first question. The number of texts was 795, produced by 325 students, and the number of tokens was 133227.

As for second and third questions, all RCs - appeared in written production sub-corpus, spoken production sub-corpus, 1st year written production sub-corpus, and 4th year written production sub-corpus - were analysed.

To answer the fourth question, two groups were selected: The first group included students whose L1 apply pronoun retention, while the second included students whose L1 do not apply pronoun retention. The sample size used to answer this question was relatively small due to the limited information of typological categories for many of learners' L1. Only two languages, among native languages of learners in ALC, were mentioned in Keenan and Comrie (1977) in the category of applying pronoun retention, namely Chinses and Persian. The number of texts of learners whose L1 was Chinese or Persian was 90 (of 17 students). A sub-corpus for the group of (+ retention) was built using these texts, containing 14150 tokens. Although the number of languages, in ALC, that do not apply pronoun retention (-

retention) exceeds 10 languages, only four languages were chosen to constitute the second group, in order to make the two groups balanced in size. These languages were English, Korean, Malay, and Turkish. The number of texts in this group was 97 (of 39 students), containing 14540 tokens.

3.2 Instruments

3.2.1 Wordsmith corpus tools

The instrument used for data collection in this study was Wordsmith corpus tool, developed by Mike Scott at the University of Oxford in the year 1996. The version used in the study is version six, developed in the year 2012. It has several software tools providing information about how a word or phrase used within its real context. It also provides general numerical and statistical information about a corpus, such as the total number of words (tokens) and types (i.e. words of the corpus without counting repetition) constituting the corpus. The main software tool used in the study was concordance. Concordance is a software tool that is used to "find every occurrence of a particular word or phrase" (O'keeffe et al., 2007, p. 8). "The search word or phrase is often referred to as the 'node' and concordance lines are usually presented with the node word/phrase in the centre of the line with seven or eight words presented at either side" (O'keeffe et al., 2007, p. 8). Wordsmith allows users to determine the number of letters displayed at either side. This offers a good opportunity to extend the excerpt displayed in the concordance line, which allowed the researcher to see the full context. Figure 1 shows how concordance lines appear in Wordsmith.

Figure 3.1:

Screen shot of a concordance search using Wordsmith

Wordsmith also allows saving the search results in Microsoft Excel format, which facilitated the procedure of analysis, especially quantitative analysis. More information about Wordsmith tool version six is available in Scott (2012).

3.2.2 Microsoft Excel and its add-ins

Microsoft Excel is a famous software program widely used for different administrative and arithmetic functions. In this study, it was used in categorizing, scoring, and calculating simple counting results. It was also used in reporting descriptive statistics, such as means, variance, and standard deviation. Additionally, Excel has the add-ins feature, which allows it to include specific downloadable external tools. XLSTAT is a software tool for data analysis

solutions that can be added to Microsoft Excel. It is a registered trade mark of Adinsoft SARL. More information about XLSTAT is available on ("XLSTAT | Trademarks and copyrights |," n.d.).

In this study, XLSTAT was used to run statistical tests needed to check the significance of differences between scores of RC structures, which appeared in raw numbers of occurrences and descriptive statistic results. Three types of tests were done in the current study: t-test, ANOVA one way test, and Tukey test.

T-test is used to check significance of results between only two groups/variables, by calculating the probability of obtaining these results randomly. In other words, it examines whether these results indicate a relation between variables or likely to happen by chance. Probability value of results is compared to alpha value, which is commonly determined in applied linguistics research by 5%. The test is based on "null hypothesis", meaning that there are no differences between samples. If the probability value (p) is greater than alpha (0.05), null hypothesis is accepted, which means that results do not indicate significance differences between samples. Conversely, if (p) value is less than alpha, null hypothesis is rejected, which means that differences are significant and can be generalized.

ANOVA test is used to check the significance of differences between more than two groups/samples. It is used in cases of comparing multiple independent variables concerning one dependent variable. Like t-test, it depends on "null hypothesis" using alpha value of 5%. If the results of ANOVA indicate significance in differences, these significant results may be between all variables or only between two of them. Therefore, a post-hoc test is needed in order to determine the source of significance.

Tukey test for homogeneous subsets difference (HSD) is the post-hoc test used in this study. It analyses significance of differences between each two variables/samples examined in ANOVA, and it reports the significance value of differences between each two variables (Dörnyei, 2007).

3.3 Data collection:

3.3.1 Framework of data collection

Targeted data in this study were all incidents in which RCs of definite relative nouns were used in AFL learners' production. Hence, this set of relative pronouns were searched for through the non-native learners' corpus, which resulted in listing all incidents of RCs. Having these data, all RCs in learners' free production, allowed the researcher to determine the different patterns used. Hence, the researcher can calculate the frequency score of each pattern. This also allowed the researcher to calculate target-like and non-target like RCs to detect the degree of accuracy of each pattern.

Answering the second question required more specific data. It required building two sub-corpora: One for written production (116724 tokens; 753 texts; 301 students) and another one for spoken production (16503 tokens; 42 texts; 24 students).

Like the second question, the third question also required building a sub-corpus of learners in first year in high school and another sub-corpus of learners in the fourth year in high school.

Finally, answering the fourth question required two groups of learners with different L1 backgrounds. Using the online search tool mentioned above, a search that only included

specific native languages was used to download files. Then, the researcher built a sub-corpus of texts whose learner's L1 apply pronoun retention (Chinese & Persian = 90 texts by 17 Ss & 14150 tokens) and another sub-corpus of texts whose learner's L1 do not apply pronoun retention (English, Korean, Malay, and Turkish = 97 texts by 39 Ss & 14540 tokens).

3.3.2 Procedure of collecting data for 1st questions

Step 1: A search for text files was run using the multi-determinants tool in ALC-search web page, namely the determinant of nativeness. The researcher marked the non-native choice. The web page indicated that 795 are available for download, so they were downloaded and saved in one folder constituting the corpus of non-native learners' production.

Step 2: Using the concordance tool, a multi-word search was run through the corpus to list all incidents in which targeted relative pronouns were used. In order to guarantee having all RCs, all the varieties in script were considered, regardless the accuracy in spelling. Hence, the search included all the following words: الذي ، الذى ، ألذي ، ألذى ، التي ، التى ، ألتي ، ألتى ، اللذان ، أللذان ، اللذين ، أللذين ، اللتان ، أللتان ، اللتين ، أللتين ، الذين ، ألذين ، اللاتي ، اللاتى ، ألاتي ، ألاتى ، اللائي ، اللائى ، ألاتي ، ألاتى.

> These varieties take in consideration the different case markers for dual form, the possibility of using "hamza" on the definite article, and the possibility of using "alif maqsˤuːra" instead of the letter "Yaː" at the end of the relative pronoun.

Step 3: Wordsmith tool generated a concordance file, including all the RCs in the corpus with the possibility to be saved in Wordsmith extension and in Excel sheet. The file was saved in both forms. The Wordsmith form (.conc) allowed the researcher to refer to in case the concordance line includes more than one relative pronoun to specify which one of them is

the node word. On the other hand, Excel sheet facilitated the analysis procedures: categorizing, calculating, and running statistical analysis.

3.3.3 Procedure of collecting data for 2nd question

Step 4: Another search for text files was run using the multi-determinants tool in ALC-search web page, but this time using determinants of nativeness, and mode of text. The researcher marked the non-native, and written choices. The web page indicated that 753 are available for download, so they were downloaded and saved in one folder constituting the corpus of written production. Similar procedure was done using the determinant of spoken instead of written to build the corpus of spoken production, which included 42 texts.

Step 5: repeating steps 2 and 3 for each of the two new corpora (written & spoken)

3.3.4 Procedure of collecting data for 3rd question

Step 6: Similar procedures of step 4 using suitable determinants for each corpus needed in this question (first high school year and fourth high school year).

Step 7: Repeating steps 2 and 3 for each of the new corpora (year 1 & year 4).

3.3.5 Procedure of collecting data for 4th question

Step 8: The researcher built a sub-corpus of (+ retention) language group. A search for text files was run using the multi-determinants tool in ALC-search web page. This time, the researcher used more than one determinant: Nativeness, marking the non-native choice; and

mother tongue, marking the choices of Chinese and Persian. The web page indicated that 90 are available for download, so they were downloaded and saved in one folder.

Step 9: The researcher built a sub-corpus of (- retention) language group. Same sequence of steps mentioned in step 4 were used, but this time different mother tongues were marked: English, Korean, Malay, and Turkish.

Step 10: The researcher run multi-word concordance search, typically like the one used in step 2, through each of the two new sub-corpora. Both concordance files generated by Wordsmith were saved in (.conc) format and Excel sheets.

After completing these steps, all needed data was collected and saved, so data became ready for analysis.

3.4 Data analysis

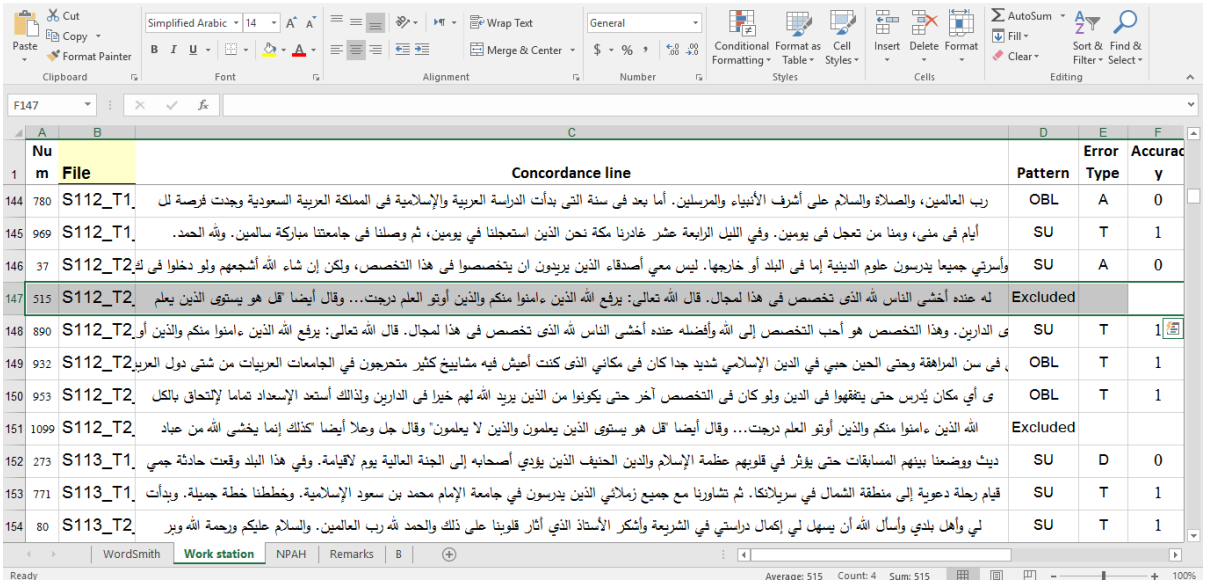
3.4.1 Framework of data analysis for 1st question

3.4.1.1 Frequency

To answer the first research question, full corpus of non-native learners was examined using the Excel sheet prepared in steps 1-3 in data collection. All concordance lines were analysed quantitatively. RCs that are quoted from Quran were excluded, since they do not represent the learners' production. Figure 3.2 explains this process.

Figure 3.2:

Determining RC structures using Excel

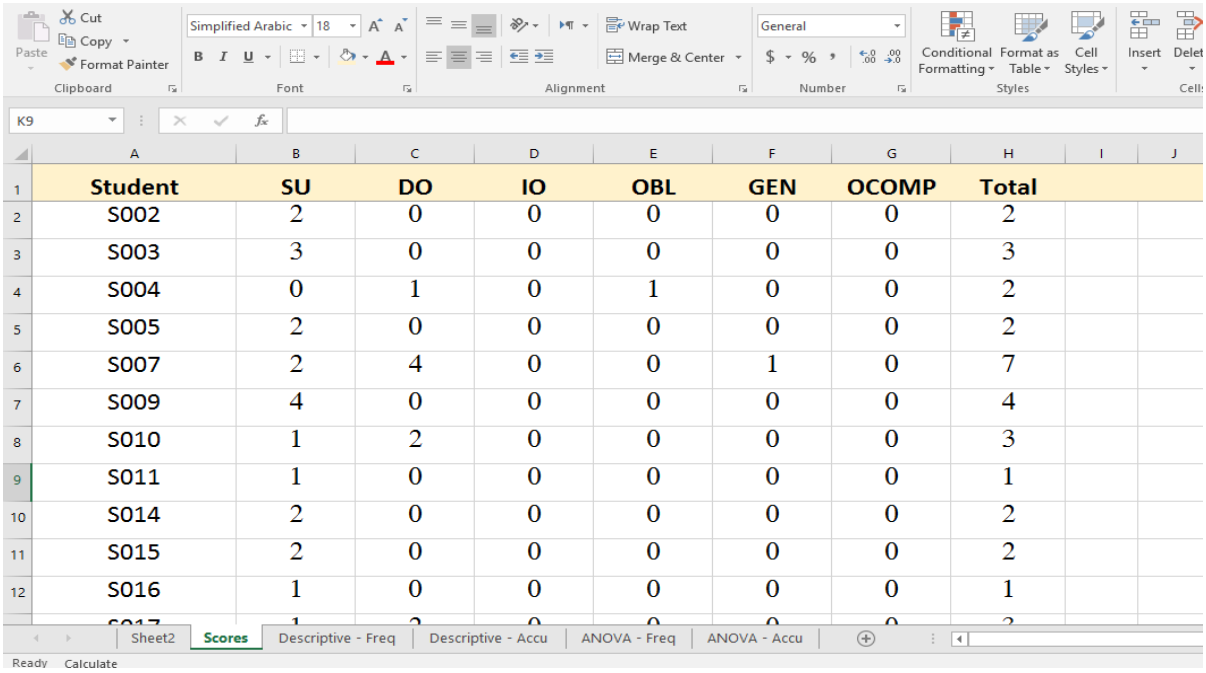


	Num	File	Concordance line	Pattern	Error Type	Accuracy
144	780	S112_T1_	رب العالمين، والصلاة والسلام على أشرف الأنبياء والمرسلين. أما بعد فى سنة التى بدأت الدراسة العربية والإسلامية فى المملكة العربية السعودية وجدت فرصة لل	OBL	A	0
145	969	S112_T1_	أيام فى منى، ومنا من تعجل فى يومين. وفي الليل الرابعة عشر غادرنا مكة نحن الذين استعجلنا في يومين، ثم وصلنا فى جامعتنا مباركة سالمين. ولله الحمد.	SU	T	1
146	37	S112_T2_	وأسرتي جميعا يدرسون علوم الدينية إما فى البلد أو خارجها. ليس معي أصدقاء الذين يريدون ان يتخصصوا فى هذا التخصص، ولكن إن شاء الله أشجعهم ولو دخلوا فى ك	SU	A	0
147	515	S112_T2_	له عنده أخشى الناس لله الذى تخصص فى هذا لمجال. قال الله تعالى: يرفع الله الذين ءامنوا منكم والذين أوتو العلم درجت... وقال أيضا "قل هو يستوى الذين يعلم	Excluded		
148	890	S112_T2_	ى الدارين. وهذا التخصص هو أحب التخصص إلى الله وأفضله عنده أخشى الناس لله الذى تخصص فى هذا لمجال. قال الله تعالى: يرفع الله الذين ءامنوا منكم والذين أو	SU	T	1
149	932	S112_T2_	فى سن المراهقة وحتى الحين حبي في الدين الإسلامي شديد جدا كان فى مكاني الذى كنت أعيش فيه مشايخ كثير متحرجون في الجامعات العربيات من شتى دول العربي	OBL	T	1
150	953	S112_T2_	ى أي مكان يُدرس حتى يتفقهوا فى الدين ولو كان فى التخصص آخر حتى يكونوا من الذين يريد الله لهم خيرا فى الدارين ولذالك أستعد الإسعداد تماما لإلتحاق بالكل	OBL	T	1
151	1099	S112_T2_	الله الذين ءامنوا منكم والذين أوتو العلم درجت... وقال أيضا "قل هو يستوى الذين يعلمون والذين لا يعلمون" وقال جل وعلا أيضا "كذلك إنما يخشى الله من عباد	Excluded		
152	273	S113_T1_	ديث ووضعنا بينهم المسابقات حتى يؤثر في قلوبهم عظمة الإسلام والدين الحنيف الذين يؤدي أصحابه إلى الجنة العالية يوم لاقيامة. وفي هذا البلد وقعت حادثة جمي	SU	D	0
153	771	S113_T1_	قيام رحلة دعوية إلى منطقة الشمال في سريلانكا. ثم تشاورنا مع جميع زملائي الذين يدرسون في جامعة الإمام محمد بن سعود الإسلامية. وخططنا خطة جميلة. وبدأت	SU	T	1
154	80	S113_T2_	لي وأهل بلدي وأسأل الله أن يسهل لي إكمال دراستي في الشريعة وأشكر الأستاذ الذي أثار قلوبنا على ذلك والحمد لله رب العالمين. والسلام عليكم ورحمة الله وبر	SU	T	1

The score of frequency was calculated in terms of the number of occurrences of each structure in learners' production, and it was recorded as shown in figure 3.3.

Figure 3.3:

Recording frequency scores for each RC structure



Student	SU	DO	IO	OBL	GEN	OCOMP	Total
S002	2	0	0	0	0	0	2
S003	3	0	0	0	0	0	3
S004	0	1	0	1	0	0	2
S005	2	0	0	0	0	0	2
S007	2	4	0	0	1	0	7
S009	4	0	0	0	0	0	4
S010	1	2	0	0	0	0	3
S011	1	0	0	0	0	0	1
S014	2	0	0	0	0	0	2
S015	2	0	0	0	0	0	2
S016	1	0	0	0	0	0	1

In order to test the significance of differences in scores, ANOVA and Tukey tests were done. Pattern of RCs were set as the independent variables, while frequency was set as the dependant variable.

Due to some features that are language specific in the Arabic language, such as parts of speech and possibility to omitting the subject, distinguishing some patterns of the NPAH required more precise definition as follows:

Subject relativization (SU) includes the subject of active voice verbs, passive voice verbs, and nominal sentences. In case of relative clause consisting of a prepositional phrase, this study treats it as a nominal sentence whose subject is omitted. For example, a sentence like (ʔalkita:b ʔallaði: ʕindi: الكتاب الذي عندي) will be considered as (ʔalkita:b ʔallaði: huwa ʕindi: الكتاب الذي هو عندي), which means that it will be counted as subject position.

Genitives (GEN): Genitives in this study are defined according to the Arabic grammar, so some structures considered as OBL in English (such as with = maʕa = مع, above = fawqa = فوق, under = taħta = تحت) will be considered as GEN in the study, as these lexical items are considered in the Arabic language as nouns.

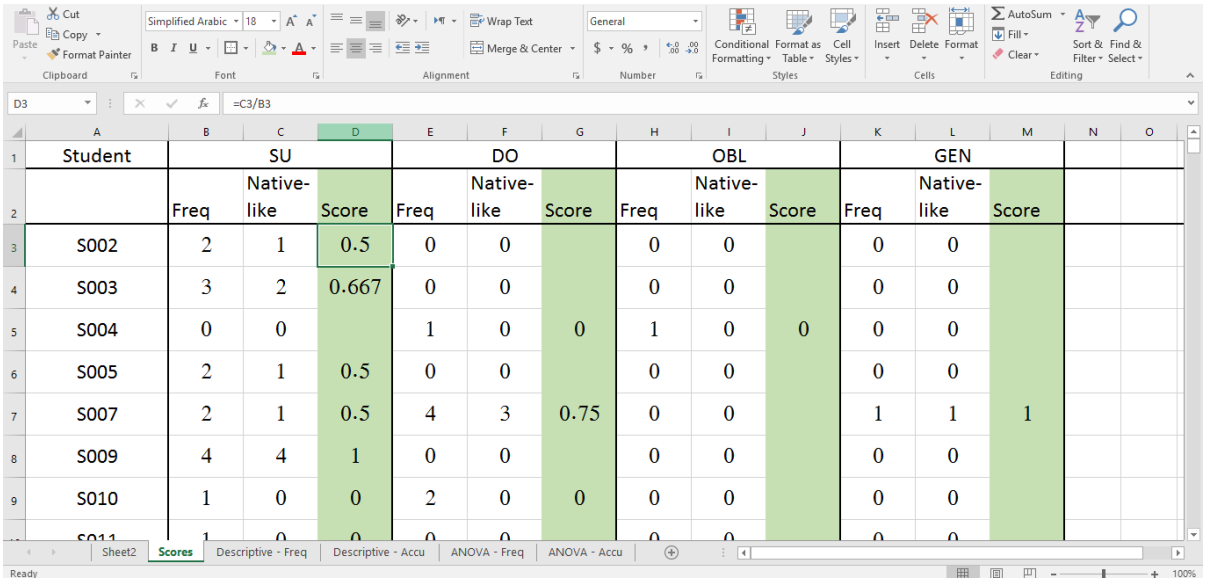
3.4.1.2 Accuracy

As for accuracy, the same data in the Excel sheet was analysed quantitatively in terms of accuracy. The construct of accuracy was quantified by calculating the score of accuracy in producing each structure. This score was calculated through dividing the number of “target like” incidents by the total number of production for the same structure. For example, if a learner produces 3 RCs of SU structure, and two of them were target-like, the accuracy score

will be 2/3 = 0.667. Figure 3.4 clarifies how this process appeared in Excel sheets.

Figure 3.4:

Calculating accuracy scores.



Student	SU			DO			OBL			GEN		
	Freq	Native-like	Score	Freq	Native-like	Score	Freq	Native-like	Score	Freq	Native-like	Score
S002	2	1	0.5	0	0		0	0		0	0	
S003	3	2	0.667	0	0		0	0		0	0	
S004	0	0		1	0	0	1	0	0	0	0	
S005	2	1	0.5	0	0		0	0		0	0	
S007	2	1	0.5	4	3	0.75	0	0		1	1	1
S009	4	4	1	0	0		0	0		0	0	
S010	1	0	0	2	0	0	0	0		0	0	

For testing the significance of differences in scores, ANOVA and Tukey tests were done. Pattern of RCs were set as the independent variables, while frequency was set as the dependant variable.

Four types of errors were counted as non-target-like production: Null, fragment, pronoun retention, and agreement. Vague and ambiguous RCs were excluded from the sample, as the researcher could not determine their patterns. As the purpose of the study is the structure of RC, errors that are not related to the RC and the relative pronoun were not counted as inaccurate production. The next lines explain in details how this study assign the accurate and inaccurate production.

3.4.1.2.1 What is excluded from the sample:

RCs quoted from Quran, for example,

(Ex 3.1) هذا الدين يرفع الله به كما قال" يرفع الله الذين آمنوا منكم والذين أوتوا العلم درجات"

haða: ʔaddi:n jarfaʕ ʔallahu bihi kama: qa:l "jarfaʕu ʔallahu ʔallaði:na ʔa:manu: minkum wallaði:na ʔu:tu: ʔalʕilma daradʒa:t"

*This religion raises up God with it as he said "raises up Allah who believed of you and who were granted knowledge degrees"

God raises up people by this religion as he said "Allah will raise up to high rank and degrees those of you who believe and who have been granted knowledge.

This clause is quoted from Quran, Al-Muja:dala (Chapter 28, Sura 58, verse 11).

Non-sense clauses, for example,

(Ex 3.2) الذي لا يعرف أهم الاسلام ودين شرعة ، أهم أصول الدين، الذي يرغب دعوة الاسلام طريقى وطريقة الإستخدام كيف إلى الإسلام

ʔallaði: la: jaʕrif ʔahamm ʔalʔisla:m wa di:n ʃarʕih, ʔahamm ʔusˁu:l ʔaddi:n, ʔallaði: jarɣab daʕwat ʔalʔisla:m tˁari:qi wa tˁari:qat ʔalʔistixda:m kajfa ʔila: ʔalʔisla:m

*Who do not know the most important of Al-Islam and religion religious law, the most important basics religion, who wants the call of Islam my way and method of use to Islam.

The one who knows the most important of Islam and religion of Islamic law, the most important of religion basics, who wants the call to Islam my way and the method of use how to Islam.

Some words of this proposition can be interpreted (the most important may be means the importance), but even after adapting these words, the whole proposition is vague and has no sense.

Ambiguous RCs, for example,

(Ex 3.3) التخصص الذي أرغب في دراسة بعد المعهد لغة العربية هي عقيدة في أصول الدين

ʔattaxasˁsˁusˁ ʔallaði: ʔaryabu fi: dira:sat baʕda ʔalmaʕhad luya ʔalʕarabijja hiya ʕaqi:dah fi ʔusˁu:l ʔaddi:n

*The specialization which I want in study after the institute language Arabic she a belief in basics the religion.

This clause is ambiguous since the absence of pronoun retention made it able to be understood in two possible meanings:

التخصص الذي أرغبه في دراسة بعد المعهد لغة العربية هي عقيدة في أصول الدين

ʔattaxasˁsˁusˁ ʔallaði: ʔaryabu fi dira:sat baʕda ʔalmaʕhad luya ʔalʕarabijja hiya ʕaqi:dah fi: ʔusˁu:l ʔaddi:n

The specialization which I want (it) in study after the Arabic language institute is the belief in religion basics

Or, التخصص الذي أرغب في دراسته بعد المعهد لغة العربية هي عقيدة في أصول الدين

ʔattaxasˁsˁusˁ ʔallaði: ʔaryabu fi dira:satihi baʕda ʔalmaʕhad luya ʔalʕarabijja hiya ʕaqi:dah fi: ʔusˁu:l ʔaddi:n

the specialization which I want to study (it) after the Arabic language institute is the belief in religion basics.

Thus, it was not possible to determine if the clause was under the DO category or the GEN category.

After excluding those types of clauses, remaining 1050 RCs were examined to fulfil the first and second research questions.

3.4.1.2.2 What is counted as inaccurate:

As mentioned above, there are four types of errors resulting in assigning an incident as inaccurate: Null, fragment, pronoun retention, and agreement.

A. **Null error** means that the student uses the relative pronoun when he/she must not use it. For example, a relative pronoun cannot be used to modify an indefinite noun. Moreover, using a relative pronoun as the head of a predicate clause makes the matrix sentence incomplete.

(Ex 3.4) ليس معي أصدقاء الذين يريدون ان يتخصصوا فى هذا التخصص

laysa maʕi: ʔasˁdiqa:ʔ ʔallaði:na juri:du:na ʔan jataxasˁsˁasˁu: fi haða: ʔattaxasˁsˁusˁ

Not with me friends who want to specialize in this specialization

‘There are no friends with me wanting to specialize in this major’

Another case of this category is when the learner adds the relative pronoun before a clause that should be the predicate of a subject, as in this example:

(Ex 3.5) لا أحد من زملائي وأصدقائي الذي يرغب في التخصص في هذا المجال

la: ʔaħada min zumala:ʔi: wa ʔasˁdiqa:ʔi: ʔallaði: jarɣab fi ʔattaxasˁsˁusˁ fi haða: ʔalmadʒa:l

'No one of my colleagues and friends who wants to specialize in this field'

The relative pronoun here makes the sentence fragment. The correct form should be as follows:

لا أحد من زملائي وأصدقائي يرغب في التخصص في هذا المجال

la: ʔaħada min zumala:ʔi wa ʔasˁdiqa:ʔi jarɣab fi ʔattaxasˁsˁusˁ fi haða: ʔalmadʒa:l

'No one of my colleagues and friends wants to be specialized in this field'

B. **Fragment** means that the relative clause is not complete grammatically. In the Arabic language, relative clause can only be a complete verbal sentence, a complete nominal sentence or a prepositional phrase/adverbial noun phrase that functions as a predicate of an omitted relativized subject. Other kinds of noun phrases are not acceptable as RCs. Despite the fact that the absence of pronoun retention makes the sentence incomplete grammatically, it is not counted as fragment because there is a separate category for pronoun retention, since it is one of the universal phenomena. This means that fragment error is concerned with grammatical features that are Arabic language specific.

(Ex 3.6) أبي هو الذي القائد في هذه الرحلة

ʔabi: huwa ʔallaði: ʔalqa:ʔid fi: haðihi ʔarriħlah

My father he who the leader in this trip

'My father is the one who was the leader in this trip'/who led this trip

In the example, the learner used a noun phrase, القائد في هذه الرحلة. In case of using a relative pronoun, the correct form is:

أبي هو الذي يقود هذه الرحلة

ʔabi: huwa ʔallaði: jaqu:d haðihi ʔarriħlah

My father he who leads this trip

'My father is the one who leads this trip'

C. **Pronoun retention** means that the student ignores adding the relator whose antecedent is the head noun.

(Ex 3.7) أدعو الله أن يتحقق هذه الأشياء التي تمنيت

ʔadʕu ʔallah ʔan jataħaqqaq haðihi ʔalʔaʃja:ʔ ʔallati: tamannajt

I ask God to make true these things which I hoped

'I ask God to make these things which I hoped come true'

In the example, the learner should have added a pronoun (relator) referring to the modified noun (things), but he/she ignored the pronoun. The correct form is:

أدعو الله أن يتحقق (يحقق) هذه الأشياء التي تمنيتها

ʔadʕu ʔallah ʔan jataħaqqaq (juħaqqiq) haðihi ʔalʔaʃja:ʔ ʔallati: tamannajtuha:

I ask God to make true these things which I hoped them

‘I ask God to make these things which I hoped come true’

This category also includes incidents in which the whole phrase of the pronoun is missing. For example,

(Ex 3.8) كذالك المكان الذي أبيت في المنى

kaða:lika ʔalmaka:n ʔallaði: ʔabi:t fi: ʔalmina

*As the place which I sleep in Mena

As the place on which I sleep in Mena’

In this example, an entire prepositional phrase is missing, namely فيه. The correct structure is as follows:

كذالك المكان الذي أبيت فيه في المنى

kaða:lika ʔalmaka:n ʔallaði: ʔabi:t fi:hi fi: ʔalmina

And so the place which I sleep in it in Mena

‘And so the place in which I sleep in Mena’

D. **Agreement** means that there is an error in the rules of verb-subject agreement or noun-noun agreement, considering that I only count errors in which the relative pronoun is a part of the combination in which the error occurs.

(Ex 3.9) أجهز الأسياء الذين يستحق رجوع معي

ʔudʒahhiz ʔalʔasja:ʔ ʔallaði:na jastaħiqq rudʒu:ʕ maʕi:

I prepare things who (he) deserve returning with me

'I prepare things who deserve to be taken back with me'

In the example, the learner used the relative pronoun of human masculine plural الذين with a non-human noun الأسياء (الأشياء). He/she should have used the relative pronoun of non-human plural التي. So the correct form is as follows:

أجهز الأشياء التي تستحق الرجوع معي

ʔudʒahhiz ʔalʔaʃja:ʔ ʔallati: tastaħiqq ʔalrudʒu:ʕ maʕi:

I prepare things which (she) deserve returning with me

'I prepare things which deserve to be taken back with me'

3.4.1.2.3 What is not counted as inaccurate:

As the purpose of the study is to investigate the acquisition of the syntactic structures of RC, non-syntactic errors, such as vocabulary and spelling, have been excluded. Furthermore, syntactic errors occurred out of the RC basic components, namely head noun, relative pronoun, and the word/phrase carrying the retention, were not counted as inaccurate production of RC. Following are examples of such incidents:

(Ex 3.10) هي الرحلة التي فعلناها إلى المدينة المنورة

hiya ʔarriħlah ʔallati: faʕalna:ha: ʔila: ʔalmadi:nah ʔalmunawwarah

It is the trip which we did it to Al-Madina Al-Monawwara

It is the trip that we took to Al-Madina Al-Monawwara

This error is related to vocabulary. The learner used the verb (فعل) while he/she should have used the verb (قام بـ). The incident was counted as target-like since the learner used the correct structure using a wrong word, which is related to the lexicon of his/her interlanguage and cannot be attributed to the syntactic system.

(Ex 3.11) قابلت الشص (الشخص) الذي لا أعرفه من قبل

qa:baltu ʔaʃʃaxsˁ ʔallaði: la: ʔaʕrifuh min qabl

I met the person who I don’t know him before

I met the person who I don’t know before

This incident included two irrelevant errors: spelling, writing (الشص) instead of (الشخص) and tense, using present tense instead of past tense. Both errors are irrelevant to the structure of the RC.

(Ex 3.12) الثلج الذي لعبته ليس حقيقيا

ʔaθθaldʒ ʔallaði: laʕibtuhu lajsa ħaqi:qijjan

The snow which I played it was not real

The snow which I played was not real

In this clause, the learner used a DO position while he/she should have used OBL. This error seemed to be related to lexicon, in terms of transitivity and intransitivity. The learner did not have the knowledge that this verb in this context should be followed by a preposition. The learner conceived this verb as a transitive and formed the structure correctly according to his/her conception. Thus, it was counted as target-like.

3.4.2 Framework of data analysis for 2nd, 3rd and 4th questions

Same considerations mentioned in the first question framework were applied in calculating frequency in third and fourth questions. Similarly, same considerations mentioned in the second question framework were applied in calculating accuracy scores for third and fourth questions.

As for the fifth question, more considerations were added. RCs of SU were excluded, since they are not relevant to pronoun retention error. Moreover, only pronoun retention error was investigated and other types were ignored. Finally, significance was tested using t-test, since we had only two groups to be examined.

3.4.3 Procedure of data analysis

Step 1: Concordance lines of the corpus of the full non-native learners were re-organized to be ordered by students.

Step 2: The researcher studied all concordance lines to determine which pattern of the six patterns forming the NPAH was used in each RC incident. Sentences and RCs that are non-sense, ambiguous, and quoted from Quran or common heritage texts were excluded from the sample. A column was added to the Excel sheet to assign the pattern of each RC.

Step 3: Using the column added in step 1, the researcher counted the number of occurrences of each pattern per learner, then scores of frequency for each pattern were recorded as explained in figure 3.3.

Step 4: Using XLSTAT, descriptive statistic (mean, standard deviation, etc.) were reported.

Step 5: ANOVA test and Tukey test were done to check significance of differences appeared in descriptive statistic report.

Step 6: The researcher studied each concordance line in the corpus of full non-native learners to determine the accurate and non-accurate incidents. Another column was added to the Excel sheet to assign whether a sentence was target-like or non-target-like.

Step 7: Number of target-like incidents were calculated for each pattern (per student), then the score of accuracy was calculated through dividing the number of accurate incidents by the total number of producing the structure, as explained in figure 3.4.

Step 8: Descriptive statistics were reported.

Step 9: ANOVA and Tukey tests were done. (End of 1st question).

Step 10: Steps 1-9 were repeated with each of the two sub-corpora of written and spoken production. (Answering 2nd question).

Step 11: Steps 1-9 were repeated with each of the two sub-corpora of first high school year and fourth high school year. (Answering 3rd question).

Step 12: Steps 1-2 were repeated for each of the two sub-corpora of (+ retention) and (- retention).

Step 13: RCs of SU position were excluded from both groups.

Step 14: The researcher studied each of the remaining concordance lines (without categorizing structures) in each group to determine the accurate and non-accurate incidents.

The construct of accuracy was quantified differently that time: Only the error type of pronoun retention was concerned.

Step 15: Step 7 (calculating accuracy scores) was repeated for each group.

Step 16: Descriptive statistics were reported.

Step 17: T-test was done (End of 4th question).

Chapter 4

Results

This chapter introduces the findings of each research question of the study, then it provides a summary of all results. For each research question, it starts by presenting simple count results, then it moves to previewing descriptive statics and reporting results of tests of significance. Finally, it states the answer of the question. Each of these data items is reported by numerical tables, diagrams, and verbal explanation.

Research questions:

1. What is the hierarchy in frequency and accuracy of AFL students' production of the six Arabic RC structures forming NPAH?
2. Is there any difference in hierarchies of frequency and accuracy between the first year and the fourth year of high school learners?
3. Is there any difference in hierarchies of frequency and accuracy between written production and spoken production?
4. Is there any significant difference in the rate of errors regarding pronoun retention between different groups based on L1 backgrounds?

4.1 Research question 1: What is the hierarchy in frequency and accuracy of AFL students' production of the six Arabic RC structures forming NPAH?

This question aims to determine the order of the six RC structures in terms of frequency and accuracy of their appearance in the full learners' production, which reflects the difficulty

order of the structures of NPAH in Arabic second language acquisition. The goal is to detect whether accuracy and frequency with which a certain structure is produced agree with the one suggested by NPAH hierarchy (the assumption in this hierarchy is that the less marked are produced more frequently and more accurately). A hierarchy that is similar to the NPAH indicates that "typological universals do play the main role in the frequency and ease of acquisition of relative clauses as predicted by markedness relationships on the NPAH" (Braidi, 1999, p.92)

4.1.1 Frequency

Addressing frequency requires analysing all RCs of the non-native corpus in ALC. Total number of RCs is 1102. 54 clauses are excluded according to the criteria explained in chapter three, so the number of RCs examined are 1048, produced by 325 learners.

Results show that SU structure records the highest number of occurrences and mean score (Sum = 567, M = 1.745, SD = 1.877), then DO which records (Sum = 250, M = 0.769, SD = 0.922), then OBL records (Sum = 181, M = 0.557, SD = 0.828), and GEN records (Sum = 50, M = 0.154, SD = 0.439). Table 4.1 provides a summary of descriptive statistics of frequency scores of the full non-native corpus.

IO position has only one incident of occurrence and OCOMP has no occurrences at all, so both of which are excluded from statistical tests.

Table 4.1:

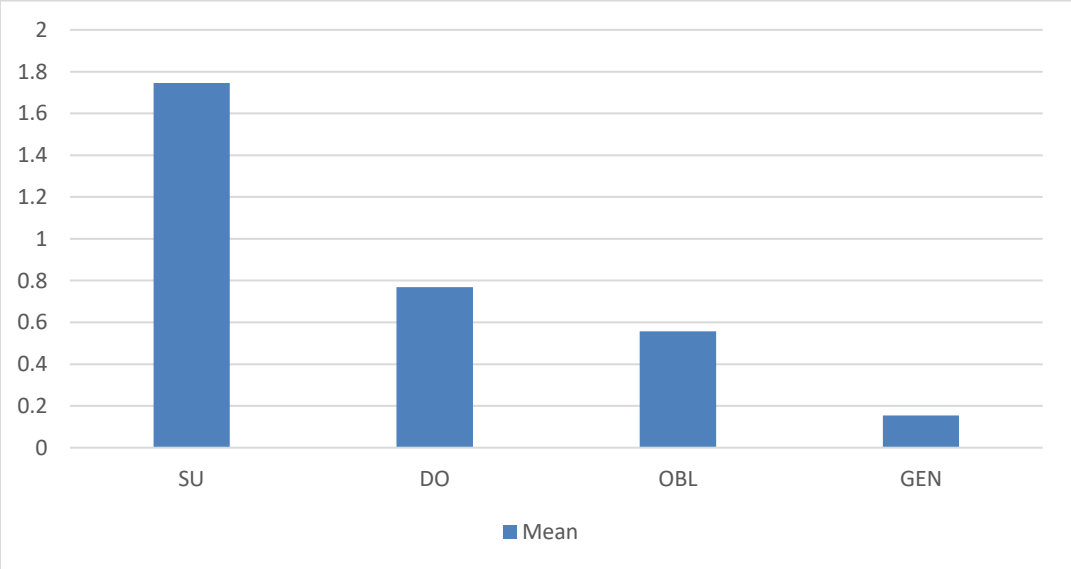
Descriptive statistics: Frequency scores of full non-native corpus.

Statistic	Nbr. of observations	Sum	Mean	Variance (n-1)	Standard deviation (n-1)
SU	325	567	1.745	3.524	1.877
DO	325	250	0.769	0.851	0.922
OBL	325	181	0.557	0.686	0.828
GEN	325	50	0.154	0.192	0.439

Results suggest differences between mean scores of patterns that are explained in figure 4.1.

Figure 4.1:

Differences in means of frequency scores between patterns.



As the diagram in figure 4.1 shows, means of frequency scores suggest a frequency order in which SU structure occupies the first place, DO takes the second place, OBL takes the third place, and GEN occupies the fourth place. IO can be considered as the fifth, second last, frequent pattern, as it records one occurrence versus zero for OCOMP, which comes in the

sixth and least place.

Differences in means are not generalizable unless they are proven to be true. Thus, one-way ANOVA test was run to check the significance of these differences. Summary of results of ANOVA test are listed in table 4.2.

Table 4.2:

ANOVA test: comparing means of frequency scores of full non-native corpus.

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	3	445.151	148.384	112.988	< **0.0001**
Error	1296	1702.000	1.313		
Corrected Total	1299	2147.151			

As shown in table 4.2, one-way ANOVA test examines differences in frequency (dependent variable) between the four RC patterns (independent variables) in the free production of non-native learners in *Arabic Learner Corpus* (ALC). As the p value is ($p < 0.0001$), which is so much less than alpha (0.05), null hypothesis is rejected, which means that there is significant difference(s) between all or some of the independent variables (RC patterns). To determine the significantly different patterns, Tukey (HSD) test was run. Results of Tukey test are presented in table 4.3.

Table 4.3:

Tukey (HSD): Checking significance of frequency means for full non-native corpus.

Contrast	Difference	Standardized difference	Critical value	Pr > Diff	Significant
SU vs GEN	1.591	17.695	2.572	< **0.0001**	Yes
SU vs OBL	1.188	13.212	2.572	< **0.0001**	Yes
SU vs DO	0.975	10.850	2.572	< **0.0001**	Yes
DO vs GEN	0.615	6.845	2.572	< **0.0001**	Yes
DO vs OBL	0.212	2.362	2.572	0.085	No
OBL vs GEN	0.403	4.484	2.572	< **0.0001**	Yes

As table 4.3 explains, probability values are calculated for differences between each two structures. A probability value that is less than 0.05 indicates that difference between these two structures is significant, which means that the difference is true for the sample and generalizable for the population. Differences between patterns are proven to be true except for DO and OBL, their means are not significantly different from each other.

Thus, the above results of significance tests, reveal that the real frequency hierarchy of RCs structures in non-native full corpus is as follows:

SU>DO=OBL>GEN> IO=OCOMP.

SU position is the most frequent structure; DO and OBL share the second position, and GEN is the third frequent structure. IO and OCOMP share the last position as the first recorded one occurrence, while the latter recorded zero occurrences.

4.1.2 Accuracy

Second research question aims to determine the level of accuracy for each of the six RC structures on the corpus of full non-native production, which is another factor indicating the difficulty order of the structures of NPAH in Arabic second language acquisition.

Descriptive statistics result in the following scores: SU (M = 0.699, SD = 0.379), DO which records (M = 0.413, SD = 0.431), OBL records (M = 0.684, SD = 0. 0.428), and GEN records (M = 0.837, SD = 0.358). Table 4.4 provides a summary of descriptive statistics of accuracy scores of the full non-native corpus.

Table 4.4:

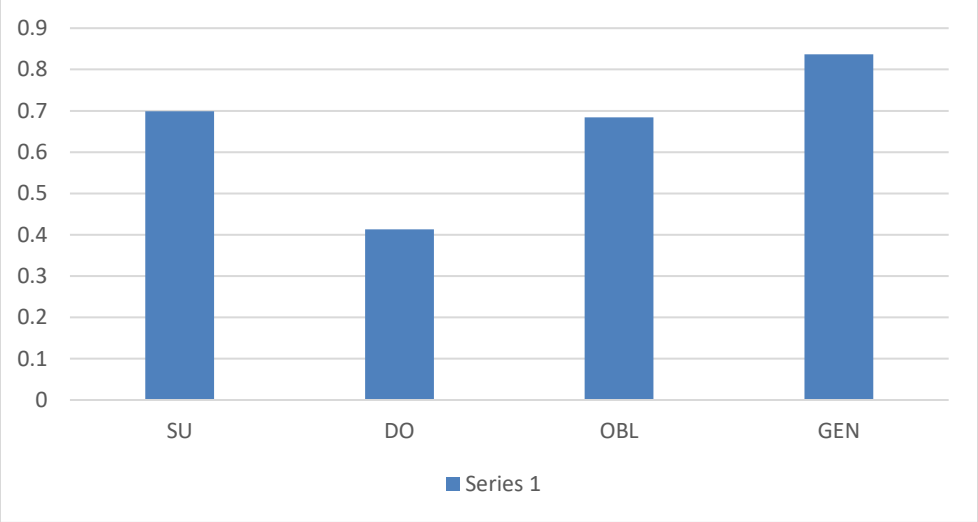
Descriptive statistics: Accuracy scores of full non-native corpus.

Statistic	Nbr. of observations	Sum	Mean	Variance (n-1)	Standard deviation (n-1)
SU	325	180.418	0.699	0.143	0.379
DO	325	68.917	0.413	0.185	0.431
OBL	325	86.150	0.684	0.183	0.428
GEN	325	35.167	0.837	0.128	0.358

Results suggest differences between mean scores of patterns that are explained in figure 4.2.

Figure 4.2:

Differences in means of accuracy scores between patterns.



As the diagram in figure 4.2 shows, means of accuracy scores suggest that the accuracy order to be as follows: GEN, then SU, then OBL, and finally DO.

As done in the first question, one-way ANOVA test was run to check the significance of these differences. Summary of results of ANOVA test are listed in table 4.5.

Table 4.5:

ANOVA test: comparing means of accuracy scores of full non-native corpus.

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	3	11.280	3.760	23.111	**< 0.0001**
Error	589	95.829	0.163		
Corrected Total	592	107.110			

As shown in table 4.5, the p value is (< 0.0001), which is so much less than alpha (0.05), so null hypothesis is rejected, which means that there is/are significant difference(s) between all or some of the independent variables (RC patterns). To determine the significantly different patterns, Tukey (HSD) test was run. Results of Tukey test are presented in table 4.6.

Table 4.6:

Tukey (HSD): Checking significance of accuracy means for full non-native corpus.

Contrast	Difference	Standardized difference	Critical value	Pr > Diff	Significant
GEN vs DO	0.425	6.099	2.576	**< 0.0001**	Yes
GEN vs OBL	0.154	2.137	2.576	0.143	No
GEN vs SU	0.138	2.056	2.576	0.169	No
SU vs DO	0.287	7.155	2.576	**< 0.0001**	Yes
SU vs OBL	0.016	0.355	2.576	0.985	No
OBL vs DO	0.271	5.695	2.576	**< 0.0001**	Yes

Results of Tukey test show that the accuracy score of DO position is significantly different from all other RC structures. On the other hand, there are no differences between SU, OBL, and GEN.

Hence, results of significance tests indicated a hierarchy of accuracy that is much different from the hierarchy suggested by mean scores. Therefore, the accuracy hierarchy of RCs structures in non-native full corpus is as follows:

SU=OBL=GEN>DO.

SU, OBL, and GEN share the same level of accuracy, while DO is proven to be produced less accurate than them.

4.2 Research question 2: Is there any difference in hierarchies of frequency and accuracy between the first year and the fourth year of high school learners?

The goal of this question is to check whether the mode of communication has influence on the acquisition order of RC structure or not. This is done through the same criteria decided in the study, which was suggested firstly by (Gass, 1979b) namely: Less difficult structures are produced more frequently and accurately. Hence, the frequency and accuracy hierarchies are searched for through the written sub-corpus and the spoken sub-corpus to be compared to each other.

Written corpus includes 1000 RCs. 51 clauses are excluded according to the criteria explained in chapter three, so the number of examined RCs are 949, produced by 301 learners.

Spoken corpus includes 102 RCs. 3 clauses are excluded according to the criteria explained in chapter three, so the number of examined RCs is 99, produced by 24 learners.

4.2.1 Frequency hierarchy of written and spoken corpora

As for written corpus, results show the following scores of frequency: SU (M = 1.698, SD = 1.886), DO (M = 0.748, SD = 0.888), OBL (M = 0.558, SD = 0.837), GEN (M = 0.150, SD = 0.433).

As for spoken hierarchy, following are values reported by the descriptive statistics software tool: SU (M = 2.333, SD = 1.685), DO (M = 1.042, SD = 1.268), OBL (M = 0.520, SD = 0.721), and GEN (M = 0.208, SD = 0.509).

Table 4.7, and diagram in figure 4.3 provide a summary of these numbers and suggest the orders inferred from them.

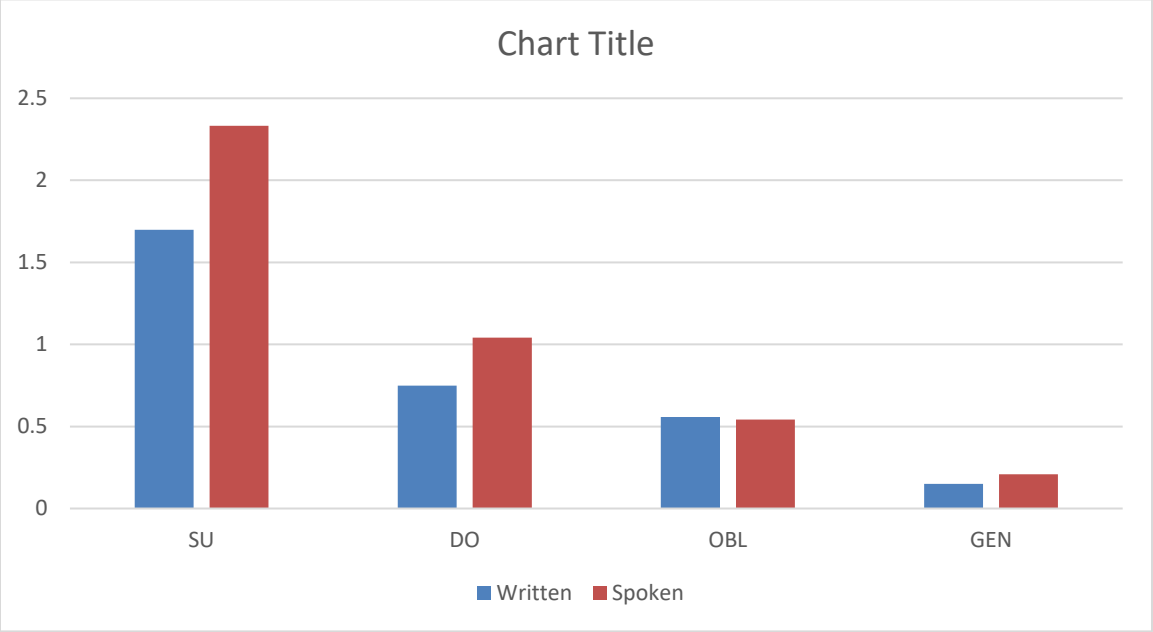
Table 4.7:
Descriptive statistics: Frequency scores of Written versus spoken corpora.

Statistic	Sum		Mean		Standard deviation	
	Written	Spoken	Written	Spoken	Written	Spoken
SU	511.000	56.000	1.698	2.333	1.886	1.685
DO	225.000	25.000	0.748	1.042	0.888	1.268
OBL	168.000	13.000	0.558	0.542	0.837	0.721
GEN	45.000	5.000	0.150	0.208	0.433	0.509

Results suggest frequency orders for RC patterns in both groups, explained in figure 4.3.

Figure 4.3:

Differences in means of frequency scores for written versus spoken corpora.



Similar hierarchies for both corpora are suggested by mean scores: SU>DO>OBL>GEN.

One-way ANOVA tests were run to check the significance of these differences. Summary of ANOVA test for written corpus are listed in table 4.8.

Table 4.8:

ANOVA test: Comparing means of frequency scores of full non-native corpus.

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	3	388.189	129.396	98.851	**< 0.0001**
Error	1200	1570.804	1.309		
Corrected Total	1203	1958.993			

As the results of ANOVA test indicate a significant value, Tukey test is required to detect the source(s) of this significant difference.

Table 4.9:

Tukey (HSD): Checking significance of frequency means for written corpus.

Contrast	Difference	Standardized difference	Critical value	Pr > Diff	Significant
SU vs GEN	1.548	16.600	2.573	< **0.0001**	Yes
SU vs OBL	1.140	12.219	2.573	< **0.0001**	Yes
SU vs DO	0.950	10.188	2.573	< **0.0001**	Yes
DO vs GEN	0.598	6.412	2.573	< **0.0001**	Yes
DO vs OBL	0.189	2.031	2.573	0.177	No
OBL vs GEN	0.409	4.382	2.573	< **0.0001**	Yes

Tukey test indicate significant differences between all pairs of RC structures except for DO and OBL.

Thus, the frequency hierarchy of written production of ALC non-native learners is as follows:

SU>DO=OBL>GEN> IO=OCOMP.

Results of ANOVA test for spoken corpus are reported in table 4.10.

Table 4.10:

ANOVA test: Comparing means of frequency scores of spoken corpus.

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	3	62.698	20.899	15.995	< **0.0001**
Error	92	120.208	1.307		
Corrected Total	95	182.906			

ANOVA test results in a probability value indicating significance in differences among RC patterns. Tukey test, reported in table 4.11, detect that the source of this significant value is the SU position.

Table 4.11:

Tukey (HSD): Checking significance of frequency for spoken.

Contrast	Difference	Standardized difference	Critical value	Pr > Diff	Significant
SU vs GEN	2.125	6.440	2.617	< **0.0001**	Yes
SU vs OBL	1.792	5.430	2.617	< **0.0001**	Yes
SU vs DO	1.292	3.914	2.617	**0.001**	Yes
DO vs GEN	0.833	2.525	2.617	0.063	No
DO vs OBL	0.500	1.515	2.617	0.433	No
OBL vs GEN	0.333	1.010	2.617	0.744	No

Results of sub-sets examination reveal that the mean score of SU is significantly different from means of all other patterns. Alternatively, differences between DO, OBL, and GEN are not significant, which means that they are equally produced by learners. Therefore, the frequency hierarchy of RC structures in spoken production is as follows:

SU>DO=OBL=GEN

4.2.2 Accuracy hierarchy of written and spoken corpora

Results regarding accuracy for written sub-corpus are as follows: SU (M = 0.698, SD = 0.384), DO (M = 0.408, SD = 0.433), OBL (M = 0.682, SD = 0.426), and GEN (M = 0.846, SD = 0.347).

Means and standard deviations of accuracy scores for spoken corpus have the following values: SU (M = 1.745, SD = 1.877), DO (M = 0.769, SD = 0.922), OBL (M = 0.557, SD = 0.828), and GEN (M = 0.154, SD = 0.439).

Table 4.12 provides a summary of all descriptive analysis.

Table 4.12:

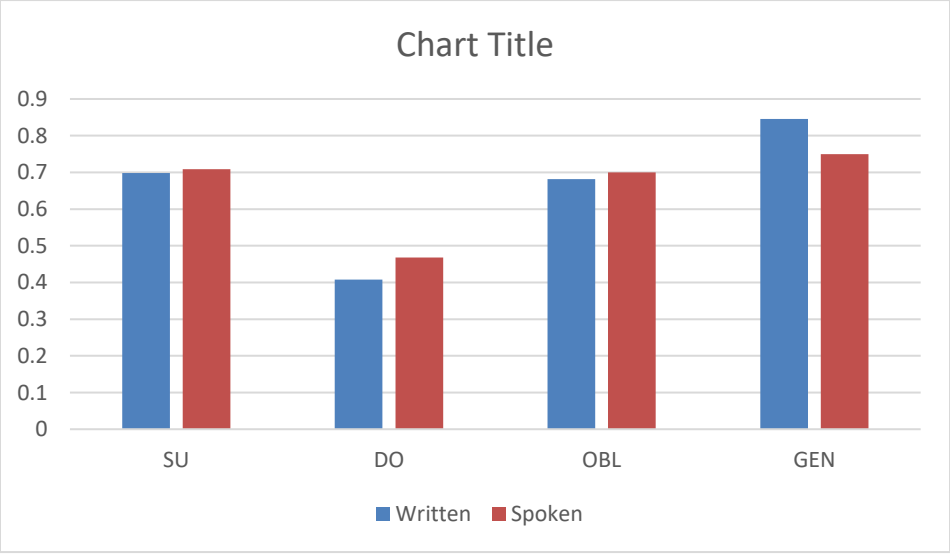
Descriptive statistics: Accuracy scores of Written versus spoken corpora.

Statistic	Sum		Mean		Standard deviation	
	Written	Spoken	Written	Spoken	Written	Spoken
SU	166.235	14.183	0.698	0.709	0.384	0.319
DO	62.833	6.083	0.408	0.468	0.433	0.416
OBL	79.150	7.000	0.682	0.700	0.426	0.483
GEN	32.167	3.000	0.846	0.750	0.347	0.500

Differences between structures can be better explained by figure 4.4.

Figure 4.4:

Differences in means of accuracy scores for written versus spoken corpora.



The same order of patterns is suggested by the diagram for both groups: GEN>SU>OBL>DO

One-way ANOVA tests were run to check the significance of these results. Summary of ANOVA test for written corpus are listed in table 4.13.

Table 4.13:
ANOVA test: Comparing means of accuracy scores of written corpus.

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	3	10.780	3.593	21.910	< **0.0001**
Error	542	88.887	0.164		
Corrected Total	545	99.667			

As the p value indicated significance, null hypothesis is rejected. This led to deciding to run Tukey test, whose results are reported in table 4.14.

Table 4.14:
Tukey (HSD): Checking significance of accuracy of written corpus.

Contrast	Difference	Standardized difference	Critical value	Pr > Diff	Significant
GEN vs DO	0.438	5.978	2.577	< **0.0001**	Yes
GEN vs OBL	0.164	2.169	2.577	0.133	No
GEN vs SU	0.148	2.092	2.577	0.157	No
SU vs DO	0.290	6.935	2.577	< **0.0001**	Yes
SU vs OBL	0.016	0.352	2.577	0.985	No
OBL vs DO	0.274	5.510	2.577	< **0.0001**	Yes

According to p values reported in Tukey results, the hierarchy of accuracy for learners' written production is as follows: SU=OBL=GEN>DO

As for spoken corpus, differences between means are proven to be not significant, that the p

value, as it appears in table 4.15, indicates that null hypothesis should be accepted.

Table 4.15:
ANOVA test: comparing means of accuracy scores of spoken corpus.

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	3	0.564	0.188	1.179	0.329
Error	43	6.861	0.160		
Corrected Total	46	7.425			

Based on ANOVA, all RC structures have the same level of accuracy in spoken production of AFL learners participating in ALC.

4.2.3 Comparing written & spoken

Findings reported above indicate that the frequency hierarchy of written corpus is (SU>DO=OBL>GEN), while the frequency hierarchy of spoken corpus is (SU>DO=OBL=GEN). Additionally, the hierarchy of accuracy for written production is (SU=DO=GEN>DO), while the ANOVA test reveals that there are no differences between patterns in spoken corpus. This means that there is an "apparent" difference in frequency and accuracy hierarchies between the two different modes of communication.

4.3 Question 3: Is there any difference in hierarchies of frequency and accuracy between written production and spoken production?

This question addresses another important variable, namely proficiency level, which is defined, in this study, by the general level of education. Therefore, a comparison is done

between the RC production of the first year in high school and the RC production of the fourth year in high school.

4.3.1 Frequency hierarchy of year 1 and year 4 corpora:

The total number of examined RCs in year one hierarchy is 203, produced by 46 students, while the examined RCs in the corpus on year four are 486, produced by 158 students.

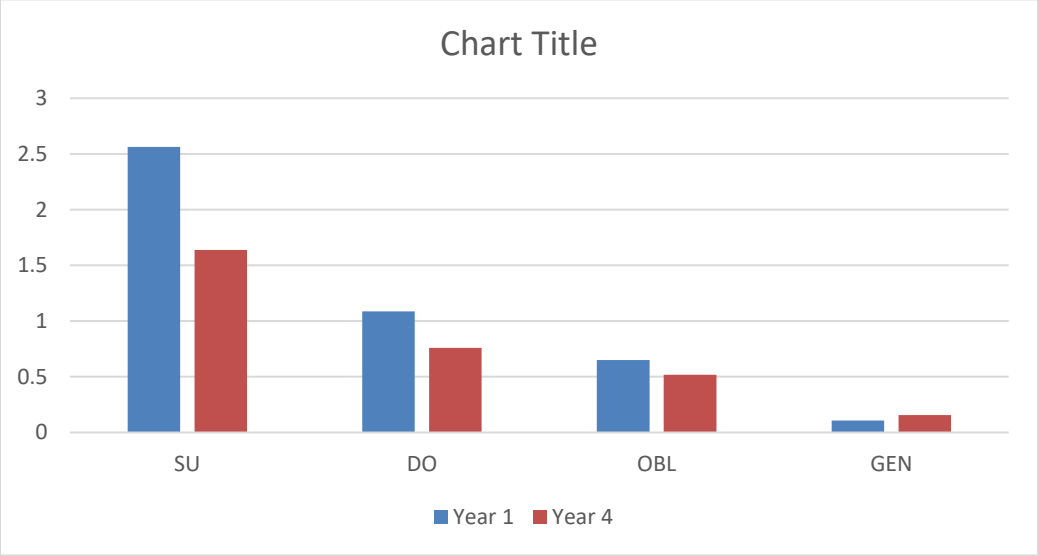
Table 4.16 includes values of frequency means for both corpora. As for year one, values recorded are SU (M = 2.565, SD = 3.606), DO (M = 1.087, SD = 1.050), OBL (M = 0.652, SD = 0.822), and GEN (M = 0.109, SD = 0.315). On the other hand, values of means on year four corpus are SU (M = 1.639, SD = 1.346), DO (M = 0.759, SD = 0.817), OBL (M = 0.519, SD = 0.865), and GEN (M = 0.158, SD = 0.499). Figure 4.7 highlights the order suggested by these values.

Table 4.16:
Descriptive statistics: Frequency scores of year one versus year four.

Statistic	Sum		Mean		Standard deviation	
	Year 1	Year 4	Year 1	Year 4	Year 1	Year 4
SU	118.000	259.000	2.565	1.639	3.606	1.346
DO	50.000	120.000	1.087	0.759	1.050	0.817
OBL	30.000	82.000	0.652	0.519	0.822	0.865
GEN	5.000	25.000	0.109	0.158	0.315	0.499

Figure 4.5:

Differences in means of frequency scores of year 1 versus year 4 corpora



Values explained in the diagram suggest that RC structures in both groups recorded the same order: (SU>DO>OBL>GEN). However, ANOVA and Tukey tests, indicate different hierarchies, as explained in tables 4.17, 4.18, 4.19, and 4.20.

Table 4.17:

ANOVA test: comparing means of frequency scores of year one corpus.

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	3	153.190	51.063	13.722	**< 0.0001**
Error	180	669.848	3.721		
Corrected Total	183	823.038			

Table 4.18:

ANOVA test: comparing means of frequency scores of year 4 corpus.

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	3	188.487	62.829	72.293	**< 0.0001**
Error	628	545.785	0.869		
Corrected Total	631	734.272			

Table 4.19:

Tukey (HSD): Checking significance of Frequency of year 1 corpus.

Contrast	Difference	Standardized difference	Critical value	Pr > Diff	Significant
SU vs GEN	2.457	6.107	2.593	**< 0.0001**	Yes
SU vs OBL	1.913	4.756	2.593	**< 0.0001**	Yes
SU vs DO	1.478	3.675	2.593	**0.002**	Yes
DO vs GEN	0.978	2.432	2.593	0.075	No
DO vs OBL	0.435	1.081	2.593	0.702	No
OBL vs GEN	0.543	1.351	2.593	0.532	No

Table 4.20:

Tukey (HSD): Checking significance of year 4 corpus.

Contrast	Difference	Standardized difference	Critical value	Pr > Diff	Significant
SU vs GEN	1.481	14.120	2.576	**< 0.0001**	Yes
SU vs OBL	1.120	10.681	2.576	**< 0.0001**	Yes
SU vs DO	0.880	8.388	2.576	**< 0.0001**	Yes
DO vs GEN	0.601	5.733	2.576	**< 0.0001**	Yes
DO vs OBL	0.241	2.293	2.576	0.101	No
OBL vs GEN	0.361	3.440	2.576	**0.003**	Yes

The p values calculated by ANOVA tests for both groups are less than 0.05, so null

hypothesis is rejected in both groups. As for year one, Tukey tests reveal that SU is significantly different from all other structures, while DO, OBL, and GEN are not significantly different from each other. As for year four, differences between structures are significant except for the difference between DO and OBL. Therefore, frequency order of RC patterns for year one is SU>DO=OBL=GEN, and order in year four is SU>DO=OBL>GEN. These results indicate a slight difference, regarding frequency, between the two levels of proficiency.

4.3.2 Accuracy hierarchy of year 1 and year 4 corpora

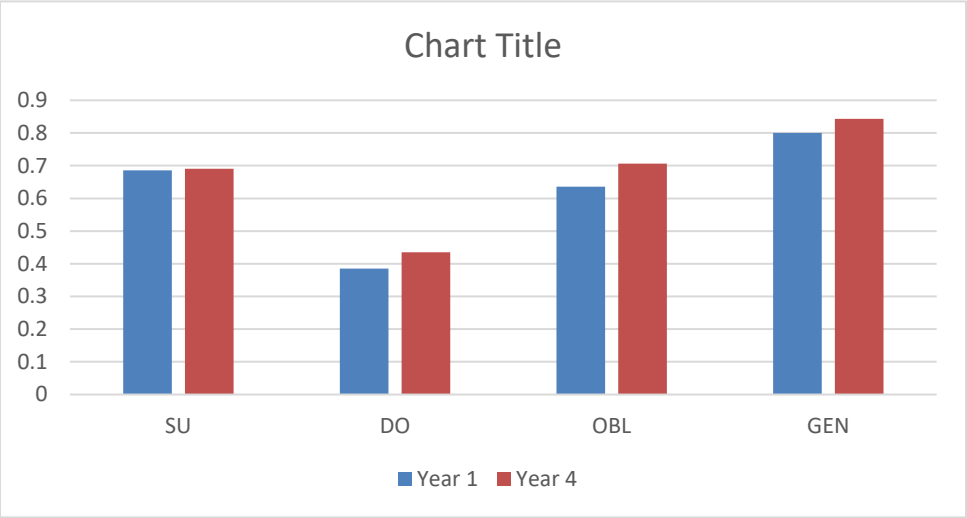
Table 4.21 includes values of accuracy means for both corpora. As for year one, values recorded are SU (M = 0.686, SD = 0.401), DO (M = 0.153, SD = 0.391), OBL (M = 0.219, SD = 0.468), and GEN (= 0.200, SD = 0.447). On the other hand, values of means on year four corpus are SU (M = 0.691, SD = 0.391), DO (M = 0.435, SD = 0.438), OBL (M = 0.706, SD = 0.404), and GEN (M = 0. 0.843, SD = 0.336). Figure 4.8 highlights the order suggested by these values.

Table 4.21:
Descriptive statistics: Accuracy scores of year 1 versus year 4 corpora.

Statistic	Sum		Mean		Standard deviation	
	Year 1	Year 4	Year 1	Year 4	Year 1	Year 4
SU	26.051	87.750	0.686	0.691	0.401	0.391
DO	11.167	37.833	0.385	0.435	0.391	0.438
OBL	14.000	38.817	0.636	0.706	0.468	0.404
GEN	4.000	15.167	0.800	0.843	0.447	0.336

Figure 4.6:

Differences in means of accuracy scores of year 1 versus year 4 corpora



Mean scores suggest (GEN>SU>OBL>DO) to be the model of accuracy hierarchy of year 1 corpus, and (GEN>OBL>SU>DO) to be the model of year 4. However, results of ANOVA and Tukey indicate different results, as shown in tables 4.22 and 4.23.

Table 4.22:

ANOVA test: Comparing means of accuracy scores of year 1 corpus.

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	3	1.832	0.611	3.518	**0.018**
Error	90	15.623	0.174		
Corrected Total	93	17.455			

Table 4.23:

Tukey (HSD): Checking significance in accuracy for year 1 corpus.

Contrast	Difference	Standardized difference	Critical value	Pr > Diff	Significant
GEN vs DO	0.415	2.545	2.618	0.175	No
GEN vs OBL	0.164	1.004	2.618	0.858	No
GEN vs SU	0.114	0.702	2.618	0.939	No
SU vs DO	0.301	1.843	2.618	**0.022**	No
SU vs OBL	0.049	0.302	2.618	0.971	No
OBL vs DO	0.251	1.541	2.618	0.150	No

Null hypothesis is accepted for comparisons between patterns, except for SU and DO. This means that SU, OBL, and GEN had the same level of accuracy. On the other hand, DO, OBL, and GEN were proven to also have the same level. However, there was a significant difference between SU and DO though both were equal to OBL and GEN. These complicated and conflicting results do not allow deciding a specific hierarchy, so the four patterns are considered as equal in the accuracy level.

Table 4.24:

ANOVA test: Comparing means of accuracy scores of year 4 corpus.

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	3	4.908	1.636	9.958	< **0.0001**
Error	283	46.498	0.164		
Corrected Total	286	51.406			

Table 4.25:

Tukey (HSD): Checking significance in accuracy for year 4 corpus.

Contrast	Difference	Standardized difference	Critical value	Pr > Diff	Significant
GEN vs DO	0.408	3.885	2.584	**0.001**	Yes
GEN vs SU	0.152	1.485	2.584	0.448	No
GEN vs OBL	0.137	1.243	2.584	0.600	No
OBL vs DO	0.271	3.879	2.584	**0.001**	Yes
OBL vs SU	0.015	0.226	2.584	0.996	No
SU vs DO	0.256	4.539	2.584	< **0.0001**	Yes

As for year 4, probability values indicate significance difference in comparing DO with the other three pattern, while differences between SU, OBL, and GEN are not significant. This means that the order of accuracy for RC patterns in year 4 corpus is SU=OBL=GEN>DO. Thus, there is a slight difference in accuracy orders between the two levels of proficiency.

4.3.3 Comparing results of year 1 and year 4 corpora:

Raw numeric results and descriptive statistical analyses suggest the similarity between the two levels of proficiency for both frequency and accuracy. However, significance tests reveal slight differences. Frequency hierarchy for year one is SU>DO=OBL=GEN, meaning that GEN is in the same rank with DO and OBL. On the other hand, frequency hierarchy for year 4 is SU>DO=OBL>GEN, meaning that GEN is less frequent than DO and OBL.

As for accuracy, year one records indicate that all patterns the same level of accuracy, while year four results indicate that DO was less accurate than the other three structures, which forms this hierarchy: SU=OBL=GEN>DO.

Thus, there are some differences in orders of both variables (frequency and accuracy) between the two levels of proficiency.

4.4 Research question 4: Is there any significant difference in the rate of errors regarding pronoun retention between different groups based on L1 backgrounds?

Two sub-corpora are used to answer this question: (+ retention) group and (- retention) group. The corpus of (+ retention) includes 90 texts written by native speakers of Chinese and Persian, and contains 14150 words. The corpus of (- retention) includes 97 texts written by native speakers of English, Korean, Malay, and Turkish, and it contains 14540 words. RCs of the pattern SU are considered irrelevant because all languages (including Arabic) are similar in not applying pronoun retention in this position.

The (+ retention) corpus lists 70 RCs, and the (-retention) corpus lists 176 RCs. After excluding RCs that include quotations from Quran, none-sense sentences, ambiguous clauses, and SU position, the investigated RCs are 31 in (+ retention) group and 73 in (- retention) group.

As reported in table 4.26, the mean of accuracy score of the (+ retention) group is 0.685 (SD = 0.394), while the mean of accuracy score of the (- retention) is 0.529 (SD = 0.434).

Table 4.26:
Descriptive statistics of accuracy scores of (+retention) and (- retention) groups.

Statistic	Nbr. of observations	Sum	Mean	Variance (n-1)	Standard deviation (n-1)
(+ retention)	39	11.650	0.685	0.156	0.394
(- retention)	39	20.650	0.529	0.188	0.434

The significance of the differences, appeared in descriptive statistics, is examined by t-test. As reported in table 4.27, the probability value is 0.2099, which is greater than 0.05. This means that there is no significant difference between the two groups regarding the accuracy of using pronoun retention strategy. This means that L1 transfer does not make a significant difference in the accuracy of applying pronoun retention strategy.

Table 4.27:

Results of t-test between the (+ retention and (- retention) groups

	F-test for equality of variances		t-test for equality of means		
	F	Sig.	t	df	Sig. (p value)
Equal variances assumed	1.2095	0.3507	1.2689	54	0.2099

4.5 Summary of results

Tables 4.28 and 4.29 provide a summary of results for questions 1 to 4, which helps understanding, interpreting, and discussing these results.

Table 4.28:

Hierarchies suggested by means

Frequency					**Accuracy**				
Full	**Written**	**Spoken**	**Year 1**	**Year 4**	**Full**	**Written**	**Spoken**	**Year 1**	**Year 4**
SU	SU	SU	SU	SU	GEN	GEN	GEN	GEN	GEN
DO	DO	DO	DO	DO	SU	SU	SU	SU	OBL
OBL	OBL	OBL	OBL	OBL	OBL	OBL	OBL	OBL	SU
GEN	GEN	GEN	GEN	GEN	DO	DO	DO	DO	DO
IO	IO								
OCOMP	OCOMP								

Table 4.29:

Hierarchies indicated by significance tests

Frequency					**Accuracy**				
Full	**Written**	**Spoken**	**Year 1**	**Year 4**	**Full**	**Written**	**Spoken**	**Year 1**	**Year 4**
S	S	S	S	S	S=O=G	S=O=G	No	No	S=O=G
D=O	D=O	D=O=G	D=O=G	D=O	D	D	hierarc	hierarc	D
G	G			G			hy	hy	
I=OC	I=OC								

(S = SU, D = DO, O = OBL, G = GEN, I = IO, OC = OCOMP)

As for the fourth question, t-test indicates that there is no significant difference regarding pronoun retention error between learners whose L1 apply pronoun retention and learners

whose L1 do not apply pronoun retention. Thus, it can be inferred that L1 background does not have a significant influence in RC acquisition in Arabic as a second language.

Chapter 5

Discussion

This chapter includes three sections. It starts with understanding what the results of the study indicate about the acquisition order of relative clause structures in Arabic as a second language and highlighting how results reported in the study relate to the hypothesis examined and other studies conducted regarding the same topic. Then, it moves to previewing pedagogical implications suggested by the findings of the study. Finally, it reports the limitations of the study.

5.1 Applicability of NPAH to Arabic language

5.1.1 Frequency

As mentioned in chapter two, frequency used to be examined in research concerning acquisition order of RC structure as an indicator of the level of difficulty in acquisition of structures. It is supposed that the more frequent a structure is the more ease in acquisition it is. In this study, the hierarchy of RC patterns regarding frequency is examined for the full production of learners participating in ALC, written production, spoken production, first year in high school, and fourth year of high school.

First and the most salient phenomenon found in results is the lack of OCOMP (0 incidents) and IO (1 incident) positions. The structure of IO records only one occurrence while OCOMP records no occurrences. Having no OCOMP RCs in the AFL learners' production may be understood, as this position is universally considered the most marked position as suggested in NPAH. However, IO pattern is counted as one of the high three positions in the

hierarchy, which means that it is expected to be more frequent than OBL and GEN. Two possible inferences can be considered for such results. First possible reason is that the IO structure in Arabic language is more complicated than OBL and GEN. This can be due to the fact that IO position in Arabic sentence is not indicated by a preposition unlike in English, for example, which indicates the IO by the preposition "to." For instance, the English sentence (Mohamed gave a pen to Ahmed) indicates the indirect object by the preposition "to." Alternatively, the equivalent sentence in Arabic (muħamed ʔaʕtˤa: aħmed qalam = محمد أعطى أحمد قلماً) does not include a preposition to indicate the indirect object (Ahmed). Moreover, the grammatical case of indirect object in Arabic language is similar to the direct object: Both are accusative, while the grammatical cases for DO are different from the case of IO in English. DO is accusative and IO is dative though that English language does not have inflectional morphemes indicating grammatical cases. Thus, having two accusative objects for one verb may be conceived as more complicated structure than object of preposition and genitive, as the first (OBL) is indicated by the preposition and the latter (GEN) is indicated by the combination between a noun and a suffix pronoun. Second possible inference is that the use of IO is generally less common than OBL and GEN in Arabic language. Because of the lack of incidents of these two positions, IO and OCOMP, in the corpus they are excluded from accuracy examinations, sub-corpora examinations, and statistical analyses and they are put together in the last rank in frequency hierarchy.

Second issue to be mentioned in this discussion is the order of RC patterns in terms of frequency of production. Scores of each corpus have been tested for significance, which results in suggesting the hierarchies previewed in table 5.1.

Table 5.1:

Frequency hierarchy of RC patterns (considering significant differences only)

Full	Written	Spoken	Year 1	Year 4
SU	SU	SU	SU	SU
DO=OBL	DO=OBL	DO=OBL=GEN	DO=OBL=GEN	DO=OBL
GEN	GEN			GEN
IO=OCOMP	IO=OCOMP			

Results based on significance tests show that the frequency hierarchy for the full, written, and fourth year (higher proficiency level) corpora have the same hierarchy, which is SU>DO=OBL>GEN. Patterns IO and OCOM are not included in analysis as they occupy the last position with almost no occurrences.

Alternatively, spoken corpus and first year (lower proficiency level) corpus also are similar in another hierarchy: SU>DO=OBL=GEN. Thus, the results of spoken and first high school year are different from the other corpora only in one position, namely GEN. Pattern GEN comes the last in full, written, and year 4 corpora, but it comes in the same level with DO and OBL in spoken and year 1 corpora. Moreover, the hierarchy showed in spoken and year 1 corpora differs from the hierarchy of the other corpora, but it does not contradict it. To explain, the hierarchy appearing in spoken and year 1 does not reflect a different order for GEN, it only reflects the absence of a significant difference between GEN and the higher positions (DO and OBL). It is also noticeable that the two corpora reflecting the different hierarchy (in which GEN have the same level in hierarchy with DO and OBL) are both characterized by being relatively small in size (Spoken = 24 students; year 1 = 46 students). As the absence of significance is a phenomenon that is associated with small-sized samples,

this small size may be the factor causing non-significant results. It is also worth mentioning here that the mean scores appear in descriptive statistical reports suggested the same hierarchy for all corpora (SU>DO>OBL>GEN), as stated in table 5.2.

Table 5.2:

Frequency hierarchy of RC patterns (considering mean scores)

Full	Written	Spoken	Year 1	Year 4
SU	SU	SU	SU	SU
DO	DO	DO	DO	DO
OBL	OBL	OBL	OBL	OBL
GEN	GEN	GEN	GEN	GEN
IO=OCOMP	IO=OCOMP			

Hence, the study reflects that SU is consistently the most frequent position and that DO and OBL are consistently equal in the level of frequency. This consistent order for these three structures is strongly proven through statistical analysis. Alternatively, the structure of GEN shows a slightly inconsistent ranking. Statistical tests reflect that GEN is lower in frequency than DO and OBL in full, written, and year 4 corpora, while it shares with them (DO & OBL) the same level of frequency in spoken and year 1 corpora. As the relatively small size of spoken and year 1 corpora may reduce the effectiveness of the statistical tests, it is worthy to consider the consistent results appearing in mean score values. Results of mean scores indicate that GEN is consistently the least frequent structure. As mean scores are not as strong as statistical tests, the inference taken from these results is that there is a "tendency" for the GEN position to be the least frequent structure in AFL learners' production.

To sum up, the study shows a strong evidence that AFL learners produce RC structures with the following order of frequency: SU>DO=OBL. The study also indicates a tendency to use

GEN as the least frequent structure. To generalize these findings, it can be inferred that there is a general tendency for RC structures in Arabic language to have the following order of frequency: SU>DO=OBL>GEN.

In order to compare findings of this study with the model suggested by the NPAH, it is better to include the patterns IO and OCOMP, so the hierarchy of frequency for RC structures in AFL learners' production will be as follows: SU>DO=OBL>GEN>IO=OCOMP.

As explained in table 5.3, these findings agree, to some extent, with the old NPAH with two major exceptions: Regression of IO and sharing the second place between DO and OBL. Alternatively, results are more compatible with Comrie (2007)'s version, as the IO and OCOMP positions are included with OBL in one category (other objects), except for the equal levels for DO and other objects.

Table 5.3:

Comparing results of frequency with old NPAH and new NPAH

ALC results (Frequency)	NPAH (old)	NPAH (new)
SU	SU	SU
DO = OBL (other objects)	DO	DO
GEN	IO	Other objects
IO = OCOMP	OBL	GEN
	GEN	
	OCOMP	

* In results of this study, OBL = Other objects

Results, regarding frequency, agree to some extent with those studies supported NPAH and markedness (Gass, 1979; Hyltenstam, 1984), as each of those studies also reported few exceptions. While Hyltenstam (1984) reported that Swedish second language learners whose

L1 is Spanish acquired IO position after OBL, results of this study indicate a more regression of IO to be placed after both OBL and GEN. Moreover, current study shows that OBL jumps not only to replace IO, but to share the same level with DO, which is a more dramatic movement. This agrees with the study conducted by Ozeki & Shirai (2007), which relied on frequency, and it also reported that DO and OBL were equally produced by high and low proficiency levels, which led to rejecting NPAH model as a predictor to difficulty order in Japanese second language.

Results also agree with many studies on east Asian languages (Izumi, 2003; Jeon & Kim, 2007; Lin, 2015; O'Grady et al., 2003), which only examined SU and DO. Those studies reported that SU was more accessible than DO, and the current study also shows a much higher frequency of SU compared to DO.

Back to the regression of IO and the lack of OCOMP, first, the position of IO records a dramatic regression from the three higher positions to the end of the hierarchy, almost equal to OCOMP. Not only does the IO position occupy a late rank in the hierarchy, but it only records one occurrence of more than one thousand RCs, so its proportion is zero percent. Trying to understand this phenomenon, first interpretation for such results is that the NPAH was not accurate in ordering positions or even in categorizing them. This interpretation agrees with the new version of NPAH, proposed by Comrie (2007), in which he re-categorized positions. In this version, he gathered the three positions IO, OBL, and OCOMP in one position (other objects). Another interpretation, which is explained in previous section, is that the specific features of Arabic syntactic systems increase the complexity of IO structure compared to OBL and GEN, which leads to this change in its place in the hierarchy.

As for the lack of OCOMP position (zero incidents), such a score may provide a strong evidence for NPAH (Keenan & Comrie, 1977, 1979). It indicates a maximum of difficulty, to the extent to be completely avoided through Arabic learners of 67 different native languages.

Therefore, results regarding frequency in this study agree with the NPAH model (Keenan & Comrie, 1977, 1979) in reporting the different advantages for SU, DO, GEN, and OCOMP (SU>DO>GEN>OCOMP), and it is strongly different from NPAH regarding IO and OBL positions, since they reveals a regression for IO to the lowest level and a noticeable progression for OBL to share the second level with DO.

On the other hand, these findings are more compatible with the new NPAH (Comrie, 2007), which gathers OBL, IO, and OCOMP in one category. This means that the only difference between results and the model of NPAH 2007 is that other objects are equal in frequency to DO structure. To elaborate, NPAH suggests the hierarchy SU>DO>Other objects>GEN, while the current study shows the hierarchy SU>all objects>GEN.

To sum up, results reported in frequency hierarchy of RC structures in AFL learners' production suggest that NPAH is accepted as a general model of markedness of RC structures, but language-specifics have a significant influence allowing each individual language to have its own distinguishing order of frequency, and in turn, difficulty in acquisition.

5.1.2 Accuracy of RC production

Results regarding accuracy, as stated in table 5.3, carry another two salient phenomena. The first one is that SU, OBL and GEN share the same level in the accuracy hierarchy in all

corpora, as shown in table 5.3 (SU=OBL=GEN). Consistency of this phenomenon regardless of the mode of communication and proficiency level indicates similarity in difficulty level for these three structures in Arabic second language acquisition.

The second phenomenon reflected in findings is the movement of GEN structure from fifth place in NPAH model and fourth place in frequency hierarchy of this study to the first place with SU and OBL structures. This may be due to the nature of GEN structure in the Arabic language. The GEN pattern of RC in Arabic consists of one word divided to two morphemes: Noun and inflectional morpheme (possessive pronoun). These are examples of GEN RCs:

(Ex 5.1) التخصص الذي أرغب في دراسته هو تدريب المعلمين

ʔattaxasˁsˁusˁ ʔallaði: ʔarɣab fi: dira:satihi huwa tadri:b ʔalmuʕallimi:n

*The major that I desire in studying it is training teachers

The major that I want to study is teacher training

(Ex 5.2) أنا طالب من البلد الذي أكثرها من غير مسلم

ʔana: tˁa:lib min ʔalbalad ʔallati: ʔakθaruha: min ɣajr muslim

*I student from the country that it’s majority from not muslim

I’m a student from the country whose majority are not muslims

(Ex 5.3) جمع كل الطلاب الذين جاءت أسماءهم في القائمة

dʒamaʕa kull ʔatˁtˁulla:b ʔallaði:na dʒa:ʔat ʔasma:ʔuhum fi: ʔalqa:ʔimah

*he gathered all students that came their names in the list

He gathered all students whose names were listed.

As shown in examples, relativized (head) nouns are repeated in the RCs as pronouns (ـه ، ـها ، ـهم) attached to nouns (دراسة ، أكثر ، أسماء). Thus, the GEN structure in RC is consistently synthesized from two simple items: A noun, which is a clear lexical item and a possessive pronoun. The GEN structure does not carry more than one grammatical item like DO, which is attached to a verb having another grammatical item (conjugation indicating subject). Thus, the nature of the pattern, being one word synthesized by two simple lexical and grammatical units, allows the structure to be conceived as one lexico-grammatical unit, which takes less memory space while processing. This leads to an advantage for GEN structure in memorizing and, in turn, in acquisition. This means that the resulted acquisition order regarding RC structure, which is different from NPAH and many other human languages, is attributed to specific features in the syntactic system of Arabic language. This conclusion supports other results indicating the influence of language-specifics, as explained later in this section.

Table 5.4:
Accuracy hierarchy of RC patterns (considering significant differences)

Full	**Written**	**Spoken**	**Year 1**	**Year 4**
SU=OBL=GEN DO	SU=OBL=GEN DO	SU=OBL=GEN=DO (No hierarchy)	SU=OBL=GEN=DO (No hierarchy)	SU=OBL=GEN DO

The second phenomenon noticed in results of accuracy is the regression of DO pattern to be the least accurate structure in full, written, and year 4 corpora. Its results in the other two corpora, spoken and year 4, are not much different, since it does not precede any of the mentioned patterns. It just came in the same level with other structures to result in “no

hierarchy." In other words, DO structure does not record any preference, concerning accuracy, over any other structure in all corpora. Furthermore, mean scores and raw numbers reported that DO is the least accurate structure in all corpora, including spoken and year 1 corpora, which can be seen in table 5.4. Thus, results of statistical tests show a consistent equality in accuracy level for positions SU=OBL=GEN and show inconsistent regression for DO: Less accurate in full, written, and year 4 corpora, while it shares the same level with the other three structures in spoken and year 1 corpora, which are small sized corpora. Given that DO consistently has the last place in the accuracy hierarchy according to mean score results, this indicates a tendency for DO structure to be less accurate than the SU, OBL, and GEN structures in AFL learners' production.

Thus, it can be inferred that this study indicates that there is a general tendency in AFL learners' production to show the following hierarchy of accuracy regarding RC structures: SU=OBL=GEN>DO.

Table 5.5:
Accuracy hierarchy of RC patterns (considering mean scores)

Full	**Written**	**Spoken**	**Year 1**	**Year 4**
GEN	GEN	GEN	GEN	GEN
SU	SU	SU	SU	OBL
OBL	OBL	OBL	OBL	SU
DO	DO	DO	DO	DO

Interpreting this phenomenon, regression of DO in the level of accuracy, requires looking at results of the fourth question, which addresses the strategy of pronoun retention. The structure of DO occupies the second place in frequency hierarchy in this study and is

reported as the second unmarked position, preceding OBL and GEN, in NPAH and most previous studies regarding NPAH. The most salient feature that may cause more complexity to DO position in Arabic is that the pronoun retention error tends to appear in DO position more than OBL and GEN. Gass (1979) reported that pronoun retention strategy tends to appear more in the lowest three positions (OBL, GEN, and OCOMP), while participants in her study tended to omit the pronoun in DO and IO. Omitting pronoun retention in Arabic DO RCs, unlike English, results in inaccurate RC. Not only do the results regarding pronoun retention indicate no significant difference between (+ retention) mother tongues and (- retention) mother tongues, they also show that both groups have low accuracy scores regarding this error type. In other words, groups are not different because both of them reflect low accuracy level in using pronoun retention strategy. The mean score of (+ retention) is 0.685 and the mean score of (- retention) is 0.529, which reflect accuracy score of 68% and 53% respectively. Given this relatively low accuracy in pronoun retention error, and the tendency of this error to occur with DO more than other structures, it can be inferred that pronoun retention is the most probable reason of retreating the DO structure to the fourth place in the accuracy hierarchy.

The hierarchy of accuracy, according to results, showed a difficulty order that is different from both versions of NPAH. Accuracy score of the existing four positions in ALC tend to be as follows: SU=OBL=GEN>DO. Table 5.6 includes a comparison between accuracy order reflected in the current study and hierarchies suggested by the two versions of NPAH.

Table 5.6:

Comparing results of accuracy hierarchy with old NPAH and new NPAH

ALC results (Accuracy)	NPAH (old)	NPAH (new)
SU=OBL=GEN	SU	SU
DO	DO	DO
	IO	Other objects
	OBL	GEN
	GEN	
	OCOMP	

* In results of this study, OBL = Other objects

Results agree with NPAH in only one aspect: SU is more accessible than DO. On the other hand, they dramatically differ from the examined model and previous research in structures of DO, OBL and GEN. Patterns of OBL and GEN, jump to be in the top of the hierarchy sharing the first place with SU. This means that NPAH cannot be used in predicting difficulty order for Arabic RC structures, at least regarding accuracy. Moreover, although the current results are not compatible with the exact model of NPAH, they are not consistent with other studies that rejected NPAH in the literature. To elaborate, Gibson and Wu (2013) reported that object RCs were less difficult that subject RCs, while results of this study showed that DO is more difficult than SU. Furthermore, the first experiment of Ozeki and Shirai (2007) revealed that DO and OBL are similar in difficulty to SU, while their second experiment showed similar level of difficulty for SU and DO. Those results differ from this current study, as results in this study showed a big difference between SU and DO in both frequency and accuracy.

The GEN structure was repeatedly reported in different places other than its place in NPAH. Doughty (1991) and Gass (1979) reported that GEN Pattern was acquired earlier than expected in the NPAH. Alternatively, Hyltenstam (1984) found that GEN was more difficult than OCOMP. The common interpretation for this phenomenon was that the GEN structure is a subject to be influenced by language-specifics (Braidi, 1999). This provides more support to language-specific characteristics to be more influencing in determining difficulty order of RC structures.

Summing up, this study reflects distinctive hierarchies of RC structures regarding frequency and accuracy. Both hierarchies record differences from the NPAH to different extents. The order revealed in accuracy is less compatible with NPAH, as it only agrees with the model of NPAH in one aspect, namely the preference of SU pattern over DO pattern. The hierarchy of frequency is more compatible with NPAH, as it agrees with NPAH in preferring SU over all structures, DO and OBL over GEN, and GEN over OCOMP (SU>DO=OBL>GEN>IO=OCOMP).

Disagreements between frequency and accuracy for AFL learners in one side and NPAH in the other side indicate an important influence for language-specifics in determining the acquisition order of RC structures in Arabic as a second/foreign language. One example of such specific features is diversity in sentence structures. The Arabic language allows non-verbal sentences. Such a feature influences the distribution of RC structures through the Arabic language and in turn the order of frequency of RC structures. To elaborate, English, for example, has one word order system (SVO). A sentence must include a subject (100%) and a verb. A verb can be transitive or intransitive. DO only appears with transitive verbs.

Thus, in each sentence, there are two possibilities: To have a transitive verb, and to have an intransitive verb. If the verb is transitive, the sentence will include a DO. Alternatively, if the verb is intransitive, the sentence will not include DO. This means that the advantage of SU over DO is 2:1.

As for Arabic language, there are three possible types of sentences: Sentence includes transitive verb, which requires a DO; sentence includes intransitive verb, which means that there is no DO, and non-verbal sentence, which also lack to DO. Thus, DO may only appear in one type out of the three types of sentences in Arabic language. This means that the specific feature of having a non-verbal sentence in Arabic language changes the advantage of SU over DO to be 3:1 instead of 2:1 in English, which, in turn, leads to a distribution of structures (being more or less common) in the Arabic language than the expected distribution in other languages.

Another example of the influence of language-specific features is the GEN structure, which is explained previously. A structure that is conceived as one lexico-grammatical unit has more chance to be acquired easier because it requires less memory space and, in turn, less energy in cognitive processes. Given that complexity and simplicity is one factor of markedness, such a structure will be less marked because of its simplicity.

Therefore, findings revealed in this study do not contradict NPAH, though they record different predictions. However, findings indicate a broader interpretation of the model suggested by NPAH. To illustrate, the hierarchy suggested by NPAH heavily relies on only one factor of markedness, namely distribution of structures across human languages. Braidi (1999) listed two more important factors: Complexity/simplicity, and distribution through the

target language. Both factors are related to language specifics. Thus, considering language-specific features agrees with markedness, which is a main principle in NPAH.

To conclude, results of this study indicate that NPAH and markedness constitute the main framework of acquisition order of RC structures allowing more than one variety in prediction models, while language-specifics play a significant role in determining which variety of these models is applied to each individual language.

5.1.3 Mode of communication and proficiency level

As discussed above, mean scores give typically similar results for both modes of communication. On the other hand, the differences given by significance tests reveal only two slight differences regarding GEN in frequency and DO in accuracy, and it is concluded that there are general dominant models of hierarchies for frequency and accuracy that are followed through different categories of learners. Same can be reported regarding corpora of year 1 and year 4. Thus, it can be inferred that findings of this study signalize that difficulty order in Arabic second language is demonstrated closely equal through different modes of communication and different proficiency levels. This indicates that an individual language tends to have its own acquisition order that is consistently applied to different situations and levels of proficiency of that language.

5.1.4 Pronoun retention

As the NPAH suggested that acquisition of RCs is governed by the universal constraints (markedness) not L1 transfer, testing the L1 transfer in learners' production is useful in examining the applicability of NPAH. This study examines L1 effect through the strategy of

pronoun retention. Two groups of learners' production were created: One includes texts of learners whose native languages apply pronoun retention strategy, and the other group includes texts of learners whose native languages do not apply pronoun retention strategy. Priority for NPAH over L1 transfer would be proven if both groups demonstrate the same level of accuracy regarding pronoun retention, since this indicates that L1 transfer does not affect the acquisition. On the other hand, L1 transfer would be considered as a dominant factor in acquisition of RCs if learners whose L1 do not apply pronoun retention record a significantly less score of accuracy , since this indicates that L1 characteristics affect their acquisition. Results of t-test approve the null hypothesis, meaning that L1 transfer does not have a significant role in acquisition order of RC structures. This result agrees with general findings reported in Gass (1979) and Hyltenstam (1984). Because of exceptions found in their findings, Gass and Hyltenstam concluded that L1 transfer has secondary influence in RC acquisition, but the universal constraints are the main and dominant factor affecting acquisition. While conclusions made by Gass (1979) and Hyltenstam (1984) depended on their observations, this study supports the same conclusion with a statistical evidence that is tested for significance, which makes it stronger and generalizable.

On the other hand, despite the fact that results indicate a great advantage for universal constraints over L1 transfer, the low level of accuracy for both groups (+ retention = 68%) and (-retention = 53%) is a phenomenon that puts the "full predominance" of universals constraints in question. To elaborate, these low scores make it difficult to conclude that the absence of significant differences between the two groups is exclusively related to the typological universals constraints. For more elaboration, (+ retention) group are expected to achieve a high level of accuracy, as both factors examined (L1 transfer and typological

universals) positively affect their production f pronoun retention strategy. This is because that their L1 is similar to the target language (Arabic) and the pronoun retention strategy is the unmarked strategy according to typological universals. Thus, if typological constraints, signified in this case by NPAH, completely dominate the order of RC acquisition, scores would reflect higher levels of accuracy, because the typological constraints would push (- retention) learners to produce the unmarked strategy as accurate as the (+ retention) group. However, results reflect the opposite: (+ retention) group do produce pronoun retention nearly as less-accurate as (- retention) group. This indicates that there may be more factors affecting the accuracy of pronoun retention strategy. One idea that may interprets this phenomenon is proficiency. Proficiency may be the third factor affecting the accuracy of pronoun retention strategy, meaning that learners do the error of pronoun retention because they do not have adequate competence to determine whether the relativization is SU or DO.

Therefore, results of this question may be interpreted as supportive to the typological universals over L1 transfer, or they may indicate the existence of more variables affecting the accuracy of RC regarding pronoun retention strategy.

To conclude, findings of this study reveal incompatible results. Some findings conform to the model suggested in NPAH for acquisition order of RC structures, while other findings differ from the model, indicating more influencing factors (i.e proficiency) and a major role for language-specifics in deciding the order of difficulty of acquisition of RC patterns.

5.2 Pedagogical implications

One of the pedagogical implications inferred from NPAH hypothesis was that teaching higher structures in the hierarchy (unmarked structures) is not needed and that teaching only

lower structures (more marked structures) is sufficient and can lead learners to acquire all higher structures. Applying this inference to the specific order of difficulty of ASL, as indicated in this study, suggests giving more focus on teaching DO as it is the less accurate RC structure in learners' production.

Regardless the differences in hierarchies, results indicate that SU RCs are strongly preferred over DO RCs in both frequency and accuracy. This means that the acquisition of DO is expected to be much later than SU. Thus, mastering DO position may be used as indicator of higher achievement and higher level of proficiency. This can be investigated by comparing frequency and accuracy scores between learner's productions of different proficiency levels.

As language specifics are reported in this study, and many other studies, as a strongly influencing factor in acquisition, highlighting the distinguishing features related to RC structures in Arabic language (Double function of relative pronoun, requiring a definite head noun, agreement, flexible word order, non-verbal sentences) is expected to increase students' language competence and in turn to enhance their acquisition. To elaborate, double function of relative pronouns (acting as a head noun and a relativizer) is related to the Arabic way in categorizing different lexical items in parts of speech. Setting relative pronouns under the category of nouns as a part of speech is the reason that relative pronouns can function as head nouns. Flexibility in word order (SVO & VSO) and non-verbal sentences (Subject-predicate = NP + NP) are other specific features in syntax of Arabic language that affect the distribution (frequency) and complexity of RC structure. Issues like parts of speech and this wide range of word order patterns are not unified through human languages, but they differ from an individual language to another. This means that the specific features of Arabic

regarding these issues are not expected to be acquired unconsciously. Thus, Explicit teaching of these syntactic aspects and their influence on the complexity and simplicity of RC structures is expected to improve learners' competence of these structures.

5.3 Suggestions

The current study sheds light on some issues that worth investigation in further research. To begin with, differences between the difficulty order appears in frequency hierarchy, and the difficulty order appears in accuracy hierarchy of AFL learners raise a question about the extent to which each variable (frequency & accuracy) reflects the status of acquisition. In other words, when evaluating the status of acquisition, should we rely more on frequency, rely more on accuracy, or rely equally on both?

When examining theories about acquisition, some findings indicated that there may be more intervening factors affecting the order of acquisition in addition to the universals and L1 transfer. For example, the low scores of accuracy of (+ retention) group suggest competence in grammar as an influencing factor. Moreover, differences between frequency hierarchy and accuracy hierarchy indicate differences in the factors affecting each variable. These intervening factors may be caused by conscious decisions made by learners. For instance, a learner may decide to produce more DO structures because of his/her self-evaluation. The learner, in this case, conceive him/herself as competent to produce this structure while he/she is not.

Extending the last point leads to a broader issue, which is investigating the relationship between typological universals and other factors affecting acquisition. Gass (1979) and Hyltenstam (1984) reported that their results reflected interaction between typological

universals and L1 transfer. Explaining the notion of typological universals, Culbertson (2012) stated that the principles or constraints of typological universals delimit the number of assumption done by a learner while acquiring a language. This limited number of hypotheses facilitates acquisition. When conceiving typological universals as a framework limiting the number of assumption that can be made by human brain, considering the interaction between universals and L1 reported in previous research, and considering results of this study that indicate more intervening factors, the interaction between typological universals and other factors affecting acquisition may be understood from another perspective: Two-level interaction. In this interaction, typological universals are in the higher level constituting a comprehensive umbrella providing the human brain with the possible assumption for a human language and limiting the brain from only these assumptions. The second level contains all influencing factors, such as L1 transfer, conscious learning, and language-specifics. These factors interact with each other resulting in the specific acquisition order of each individual language and maybe for each individual learner. For example, typological universals, signified by NPAH and markedness, determine the framework of RC structure as a model includes six possible patterns with a rudimentary order (SU> DO>IO>OBL>GEN> OCOMP). In the second level of interaction, L1 transfer, conscious decisions, and language-specifics play their roles in determining the specific order of acquisition for each individual language. This specific order does not contradict the main principles of typological universals. Results of this study give an example. The rudimentary order suggests that DO is less marked than GEN according to one criterion of markedness, which is distribution through human languages. Results of this study suggest that GEN is less difficult in acquisition, and it is attributed to another criterion of markedness, which is

simplicity/complexity. Thus, a feature that is language-specific (simplicity of GEN compared to DO) influences the order of acquisition under the comprehensive umbrella of typological universals (markedness criteria and the six structures of NPAH).

There is one more point that needs to be mentioned in this section, as an observation noticed by the researcher, which is considering applying pronoun retention strategy as unmarked universally. A more detailed analysis is needed - that (+ retention) may be unmarked in specific positions (i.e GEN & OCOMP) and marked in other positions (i.e DO).

5.4 Limitations

Education context, in which learners participated in the corpus were involved, has some features that are different from other contexts of Arabic teaching. Education system in Saudi Arabia pay much more attention to religion than any other education system. It was noticeable that religious topics were the dominant themes in most texts. This may delimit representativeness of the sample.

Level of proficiency is another limitation in this study since it is defined by the year of high school and is not accurately tested.

One more issue is about the nature of corpus search. Finding RCs in a corpus search can only be done by searching for incidents of using relative pronouns. There is no possibility to run a corpus search that reveals incidents in which students try to use RCs, but they mistakenly miss to right the relative pronoun. In other words, in my search I can find the null error (using relative pronoun when it should not be used), but I cannot find the opposing error (not using relative pronoun when it should be used). Detecting this type of error requires

analysing full content of all texts manually. Thus, the analysis of this study does not cover the full image of learners' RC production.

5.5 Conclusion

Data analyses done in this study reveal that the order in which RC structure are frequently used in Arabic as a second language production is obviously different from the order of accuracy. Moreover, hierarchy suggested in NPAH cannot predict the specific acquisition order of RC structures in Arabic second language acquisition. Findings agree with NPAH in the relationship between SU and DO structures. Language-specifics are highlighted as a major factor determining difficulty order of RC structures in Arabic as a second language learners', while NPAH and markedness are considered as the broad framework. Therefore, the study supports explicit teaching for unique aspects of Arabic syntax. Finally, findings suggest the acquisition of the DO pattern may reflect achievement of higher level of proficiency.

References

Alfaifi, A., & Atwell, E. (2014). *Arabic learner corpus (ALC) v2*. Kobe, Japan. Retrieved from http://www.arabiclearnercorpus.com/

Alotaibi, M., & Borsley, R. D. (2013). Gaps and resumptive pronouns in modern standard Arabic. In *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin* (pp. 6–26). Berlin.

Braidi, S. (1999). *The acquisition of second language syntax*. New York: Arnold.

Comrie, B. (2007). The acquisition of relative clauses in relation to language typology. *Studies in Second Language Acquisition*, *29*(02), 301–309.

Culbertson, J. (2012). Typological universals as reflections of biased learning: Evidence from artificial language learning. *Language and Linguistics Compass*, *6*(5), 310–329.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.

Doughty, C. (1991). Second language instruction does make a difference. *Studies in Second Language Acquisition*, *13*(04), 431–469.

Eckman, F. R., Bell, L., & Nelson, D. (1988). On the generalization of relative clause instruction in the acquisition of English as a second language1. *Applied Linguistics*, *9*(1), 1–20.

Gass, S. (1979). Language transfer and universal grammatical relations. *Language Learning*, *29*(2), 327–344.

Gass, S., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). New York: Routledge.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1–76.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). London: MIT Press.

Gibson, E., & Wu, H.-H. I. (2013). Processing Chinese relative clauses in context. *Language and Cognitive Processes*, *28*(1–2), 125–155.

Hunston, S. (2002). *Corpora in applied linguistics*. United Kingdom: Cambridge university press.

Hyltenstam, K. (1984). The use of typological markedness conditions as predictors in second language acquisition: The case of pronominal copies in relative clauses. In R. W. Andersen (Ed.), *Second languages: A cross-linguistic perspective* (pp. 39–58). Rowley, MA: Newbury House.

Izumi, S. (2003). Processing difficulty in comprehension and production of relative clauses by learners of English as a second language. *Language Learning*, *53*(2), 285–323.

Jeon, K. S., & Kim, H.-Y. (2007). Development of relativization in Korean as a foreign language: The noun phrase accessibility hierarchy in head-internal and head-external relative clauses. *Studies in Second Language Acquisition*, *29*(02), 253–276.

Ju, Y. (2014, February). *Typological universals of relative clauses*. Presented at the East Asian Linguistics Seminar (EALS), University of Oxford.

Keenan, E. L. (2014). *Universal Grammar (RLE Linguistics A: General Linguistics)*. New York: Routledge.

Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, *8*(1), 63–99.

Keenan, E. L., & Comrie, B. (1979). Data on the noun phrase accessibility hierarchy. *Language*, *55*(2), 333–351.

Kuno, S. (1974). The position of relative clauses and conjunctions. *Linguistic Inquiry*, *5*(1), 117–136.

Lin, Y. (2015). Developmental Trajectory of Relative Clauses among Mandarin Learners of English. *Concentric: Studies in Linguistics*, *41*, 63–94.

Marefat, H., & Rahmany, R. (2009). Acquisition of English relative clauses by Persian EFL learners. *Journal of Language and Linguistic Studies*, *5*(2), 21–48.

Nielsen, H. L. (1997). On acquisition order of agreement procedures in Arabic learner language. *Al-'Arabiyya*, *30*(1), 49–93.

Noor, H. H. (1994). Some implications of the role of the mother tongue in second language acquisition. *Linguistica Communicatio*, *6*(1&2), 97–106.

O'Grady, W., Lee, M., & Choo, M. (2003). A subject-object asymmetry in the acquisition of relative clauses in Korean as a second language. *Studies in Second Language Acquisition*, *25*(03), 433–448.

O'keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.

Omaggio, A. (2001). *Teaching language in context* (3rd ed.). Boston: Thomson.

Ozeki, H., & Shirai, Y. (2007). Does the noun phrase accessibility hierarchy predict the difficulty order in the acquisition of Japanese relative clauses? *Studies in Second Language Acquisition*, *29*(02), 169–196.

Pavesi, M. (1986). Markedness, discoursal modes, and relative clause formation in a formal and an informal context. *Studies in Second Language Acquisition*, *8*(01), 38–55.

Rizzi, L. (1989). The new comparative syntax: principles and parameters of universal grammar. *Cahiers Ferdinand de Saussure*, (43), 65–78.

Ryding, K. C. (2005). *A reference grammar of modern standard Arabic*. New York: Cambridge university press.

Scott, M. (2012). *WordSmith 6.0*. Stroud: Lexical analysis software. Retrieved from http://www.lexically.net/wordsmith/

White, L. (2003). On the nature of interlanguage representation: Universal grammar in the second language. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 19–42). Blackwell.

XLSTAT | Trademarks and copyrights |. (n.d.). Retrieved December 3, 2016, from https://www.xlstat.com/en/company/legal-informatio

Appendix: Sample of concordance lines of the corpus of non-native learners.

Num	Concordance line
1	التي أريد أنا اتحصص في دراستي في المستقبل الشريعة ، أرغب هذه التخصص لان في أصل
2	مسرو مسروران مسروران أيضاً؛ لأن أ أق أ– صعب أ– أ في بلدي كان كان صعب أ– الذي يستطيعون أن يحج؛ لأن فيه أ– فيه مقدار محدد محدد، وأنا الحمد لله اختار ال
3	و– الح الآن عندنا فرصة أن أتعلم اللغة جيداً أ– جيدةً و– أدرِّس الناس أ– الذي في بلدي، و– أريد أن إن شاء الله أن +أفتح +حتى المعهد صغير إن شاء الله في
4	ستعدنا اس– أ– عددنا كلما أشياء ال أ– أشياء التي نحتاج الذي نحتج نحتاج أ– التي يح نحتاج إلى إلى إلى السباحة، ثم ثم ذهبنا إلى مكان للسباحة، وبعد ساعتين
5	متي فاطمة خيره المشهورة بفاطمة جيبوتي فرحبت بنا أيما ترحيب كوننا أبناءها الذين لم ترهم إلا منذ زمن بعيد في جمهورية جيبوتي ، فوضعنا حقائبنا و غيرنا ملا
6	التخصص: اتتخصص الذي أحبه وأغب فيه وهى التخصص الشرعة ,لأنه من أهم التخصصاتقف الجامعةإذا تخرج ال
7	شاهد مكان الذ تقع المعركة بين المسلمين والمشركين وزرنا مقبرة الشهداء أحد التي دفنوا فيها حمزة بن عبد المطلب ثم بعد ذلك زرنا أصدقاءنا الذين يدرسون في ا
8	ما ذهبت إليه أبدا بسبب ذلك قررت أن أذهب مع أصدقائي أول مرة. تكلمت مع أحد الذي يعرف طريق جيدا و اتفقنا أن نذهب إن شاء الله. ولكن ما كان أمر سهلا. أولا،
9	د الجبال كبيرا طويلا، ولما انتهيت من هذا، وذهبت إلى بعد أماكن آخر ..... الذي ليس لي وقت أن أذكرهُ. وجدت في هذه الرحلة كثير من النّاس قد تركوا نفسهم ل
10	ما – أياشؤويا ، وفيها أيضاً كتوب كثيرة بعددهم وقديم. وقد زرنا مكاناً آخر الذي بنوا في أثناء المعركة. وقد تعلمت في هذه الرحلة، آثار التارخية وأثار القد
11	شمس كل الناس طلعوه للمزدلفة وهناك ازدحام شديداً وايضاٌ هناك المواقف أخرى التي حدثت قلة نظافة وقلة المياه الصافية والى اخره وفي البداية السفر هناك خربت

ت لكم تصاريح، وكذا وكذا. رجعنا خائبين من عنده واخضرت آلام السفر مرة أخرى التي كنا نتمنى نسيانها, وما كنت وحيدا في هذا الأمر بل كان معي زملاءي الآخرون, 12

ى تأشيرة الدخول في المم مملكة العربية السعودية، وكان معي خمسة طلاب آخرين الذين قد قُبلوا مثلي في جامعة الملك سعود. وفي شهر سبتمبر الماضي، ليلة السادس 13

قرآن والسنة، إذا فهموا المسلمين العلم الشريعة بطريق صحيح فينتهى عن أخطاء الذى فعلوا فى الصلاة وفى الزكاة وفى الحج وفى الحد شريعة. أو وفي غيرها الذي تك 14

لت معا في السيارة إلى بيتي وفي البيت التقيت بباقة الحب والحنان من اخواتي الذين كانو ينتظرونني بفارغ الصبر وتعانقت مهم تصافحنا بكل حرارة وتجاذبنا أطراف 15

فوصلنا بسلامة أيضاً والحمد لله وجمعنا أغراضنا على العربية فطلبنا إخواننا الذين جاءوا لأجلنا، وهناك وادعت أخي وزميلي الذي جئت معه من الرياض أن يصبر لما 16

اد المنح الدراسية قمنا هناك أسبوع وزرت جامعة الإسلامية وقابلت بعض إخواني الذين يدرسون في جامعة الإسلامية وبعد ذلك ذهبنا إلى جبل ورأيت شهداء أهدى وبعد 17

وهي إلى عاصمة بلدنا (أبوجا)؛ لبعض الإجراءات الأزمة عن أخواني الذين تخرجوا السنة .2013 /7 /4 الماضية من المرحلة الثانوية العربية لنحاول لهم القبول في إ 18

وكنا نرى بعض ملامح العاصمة في كلي حافة الطريق. وقبل أن نصل إلى بيت أخينا الذي قد كان ينتظرنا بكل شغف، حدث بنا ما حدث. وكان الموقف طريفا جداً. بدأت الس 19

وسيراليون النسبة المسلمين في بلدي خمس وسبعون في المائة ويوجد لدينا إذاعة التي تسمى صوت الإسلام لما أنشأت هذه الإذاعة أدّت على إسلام كثير من الناس معظم 20

جرت الأمور علي رحلتي السياحية, ورجعنا سالمين إلي بلدنا بعد ثلاثة اسابيع التي قضينا في ذلك المكان والحمد لله ربّ العالمين والصّلاة والسلام علي الرسول 21

ا حصل أثناء السفر وقبل السفر في السعودية وأبيدجان. ثم قمت بزيارة أساتذتي الذين تعلمت لديهم وانا صغير ثم قمت ببعض المواعظ والإرشاد في بعض المساجد. مكثت 22

م و الحديث الشريف فيصدق ني هذه اشياء مهمة جدا وفي العامة استطيع ان اساعد الذين لا يعلمون واستخدم الى دين الأسلام بقوة واتقان على اسلوب الدعوة والسلام 23

و و عندما أتيت هنا، وأستطيع أن أ أهو أدرس كان أهو أ الآن أدرس مع أستاذة الذي يتكلم بالعربية، وكل
24 +كتبي عربية، و أ– أتكلم كل يوم مع صديقتي باللغة العر

ة والدين في مدينتي . لأن عامة الناس يريدون أن يتعلموا ولكن لا يوجد أسخاص الذين يدرِّسنهم. وأرجو أن
25 أستفيد من الدراسة في هذه الجامعة المشمرة التي تقع ف

ازة بين فصلين عام دراسي 1433/1434هـ وقضيت هناك عشرة أيام. سافرت مع أسرتي التي هي موجودة
26 معي في الرياض. الاستعداد للرحلة كان صعبا قليلا لأننا انتقلنا إ

فدت من هذه الفرصة، ودخلت إلى اسطنبول. في المطار قابلتني صديقي من اسطنبول الذي يدرس معي في
27 المعهد التعليم اللغة العربية بجامعة الإمام محمد بن سعود هو ا

وروبا، واشتريت بعض الكتب المفيدة. الحمد لله . قضيت وقتا ممتعا في اسطنبول الذي اخبرتكم عنه والذي ما
28 استطعت أن أخبركم بسبب الوقت المحدود في أثناء الدرس.

ذهبت إلي النادي و سجلت اسمي لأشارك في الرحلة. بعد أسبوع حينما ظهر أسماء الذين يشارك في الرحلة
29 نظرت اسمي في تلك الأسماء، في يوم الخميس التالي ذهبنا إل

ذهبنا قرية التي يسمى يوني وجدنا في هذه القرية من المحاضرين. ثمانية أشخاص الذين دخلوا في الاسلام في
30 أثناء المحاضرات، وكان من بينهم مشركين وأخبرنا أصدقا

ن أكن خبير الاقتصاد و التاجر كبير مثل عثمان بن عفان رضي الله عنه. و اشيء الذي يحمني في هذا
31 التخصص يعني كيف سيرة التفكير الاقتصاد من أول الناس التقالدي

بي أستاذ اللغة العربية في جامعة في كوريا وحتى هو لا يعرف و لا يفهم أشياء التي نتعلمها هنا، و هو الذي
32 شجعني أن أدرس في هذه الجامعة.. لذلك أحاول و أتمنى

33 راسة اللغة العربية. نعم أسرتي يساعدوني لإختيار التخصص و أشكرهم لكل أشياء الذين هم يفعلون لى. ى.

فصلينا صلاة الفجر، وبعد ذلك استعدنا اس– أ– عددنا كلما أشياء ال أ– أشياء التي نحتاج الذي نحتج نحتاج
34 أ– التي يح نحتاج إلى إلى إلى السباحة، ثم ثم ذهبنا

م يوم الثانى ذهبنا إلى ميناء ملك عبدالعزيز الذى كان الدمام وشاهدنا أشياء التى تاتى من خارج المملكة
35 السعودية العربية وصعدنا إلى البرج وشاهد البحر، ثم ر

ير ذلك كمسجد قبلتين وغير ذلك لكن جبل عود هو حبل شائق جد، وسلمت على أصحاب الذى دفنا فى
36 العود، وبعد عائد إلى رياض باسلام ورحمة، هذا رحلة ممتعة جد فيه عا

وأسرتي جميعا يدرسون علوم الدينية إما فى البلد أو خارجها. ليس معي أصدقاء الذين يريدون ان يتخصصوا
37 فى هذا التخصص، ولكن إن شاء الله أشجعهم ولو دخلوا فى ك

اف الماء فكنت في الشاطئ جانب البحر و أخذنا بعد صور من هناك مع بعض أصدقاء الذين جاعوا من غانا
38 وسنغال وبنين مكثنا حتى إلى غروب الشمس ورجعنا إلى البيت وب

ل أحد ومقبرة شهداء وبعض المزرعة. ثم نصلي في المسجد القباء ونزور اصدقاءنا الذين يدرسون في جامعة
39 الإسلامية بمدينة المنورة قضينا يومين ثم نرجع إلى مكة ون

د انتهاء يصاحبنا أصدقاءنا إلى أسواق التمور لشراء التمر ثم ودعنا أصدقاءنا الذين يدرسون في جامعة
40 الإسلامية بالمدينة المنورة فعدنا إلى الرياض.

لوا في الاسلام في أثناء المحاضرات، وكان من بينهم مشركين وأخبرنا أصدقاءنا الذين كانوا من قريتهم أنّ من
41 الذين أسلموا وثنيون. وبعد إسلامهم كثروا أصنامهم

الشهداء أحد التي دفنوا فيها حمزة بن عبد المطلب ثم بعد ذلك زرنا أصدقاءنا الذين يدرسون في الجامعة
42 الإسلامية بالمدينة المنورة وهذه الجامعة كثير من الطلا

تهينا الصلاة جمع ذهبنا إلى الجامعة الأم القرى في مكة لزيارة إلى اصدقائنا الذين يدرسون في الجامعة الأم
43 القرى. بعد ذلك أكلنا الطعام و شربنا ماء زمزم في

من ثيابٍ وعطور وساعات وطاقيات. وبعد إسلام هؤلاء سمعنا أخبار من أصدقائنا الذين كانوا في تلك القرية
44 أنّ من بينهم من كان عبّاد الأصنام ورؤسائهم وبعد إسل

حات, كانت رحلتي في شهر جولي في يوم التاسع 2011/7/9م.ذهبت مع إحدى أصدقائي الذي هو عبد
45 السلام بللو, إلي الورن, طالما نشوق المشاركة مع الأصدقاء المتعلمين

ى حلاقة و حلقنا رأسنا ثم رجعنا إلى فندوق , بعد ذلك ذهبت إلى بيوت أصديقاء الذين يدرسون في جامعاة
46 مختلفة كمثل جامعة أم قرى و جامعة الدار الحديث لأجلي زي

ع أن أتحدث مع سعودية إن شاْ الله أذا أخرج الجامعة سأذهب إلي بلدي أعلمهم الذي هم يريد أن يدرسو اللغة
47 العربية" ي هم يريد أن يدرسو اللغة العربية"

ا بعد : كما تعرفون من العنوان رحلة إلى كابول أن كابول هي عاصمة أفغانستان الذي وقع في قارة أسيا وهي
48 مكان جميل زينة الأشجار والأزهار والجبال والبحار وفي

نك تدرس العقيدة وتزيل هذه العقيدة الباطلة. ما عندي أيّ اسرة أو أي أقرباء الذين حصلوا التخصص من علوم
49 الشرعي، بل أنصح لي معلمي الذي درّسنا في بلادنا. أن

لعام في ذلك أم في تلك السفارة، وقد بين لي بعض م بعض الاستفسارات منها، ال التي طلبت منه م منها
50 الأوراق المطلوبة للس الأوراق المطلوبة للتصديق منها س– أ–

واحد عند أ– لكل واحد لديهم أ– لديهم عمل أعمال كثيرة أعمال كثيرة التي ال الذي ترك في بلدنا الذي ترك
51 في بلده. ثم استع اس استع استعدنا للسفر، وبع ثم ركب

ن، لغة العربية، ولغة الفرنسية، في من يدرس اللغة الفرنسية فقط، ولكن ال ال الذين يدرسون اللغة العربية لا بد
52 من أن يدرسوا لغة الفرنسية أيضاً، يجمعون بينه

بعد فترة شعرت أنني لا أستطيع أن أتعلم هذا الدين تعلما حقيقيا ونافعاً إلا الذي يفتح لي علوم الدين (بالمفتاح)
53 أبوابه ألا وإن هذا المفتاح هو دراسة اللغة

قريب أخيراً قال تعالى فى محكم التنزيل: والعصر 1 إن الإنسان لفى خسر 2 إلا الذين آمنوا وعملوا الصالحات
54 وتواصوا بالحق وتوصوا بالصبر 3 هذه السورة تدلنى عل

من أهم الأشياء التي قمت بها. فكان الحصول على التأشيرة صعب لسبب الإجرءات التي تتعلق بالحصول
55 عليها. يوم السفر كان يوم جميل وكنت مسرورا جد من جهة ، لكن

درجني الله سبحانه وتعالى في تلك الزمرة المباركة، هذا الشعور وهذا الإحساس الذي قد جعلني أسجد لله سبحانه
56 وتعالى، وجعلتي وجعل عيوني تسح الدموع سحا، وبعد

مرحلة الإعداد وأنا على باب الدخول في مرحلة جديدة حقيقية في عالم الأحلام التي كنت أعيشه. ففي ليلة
57 السادس من سبتمر الماضي خرجت من الدار وذهبت إلى مطار

يد ولا يقل مما نحتاجة لهذ الرحلة، مما أعجبني لهذه الرحلة الصحبة الإخائية التي قوة فيما بيننا، وعندما كنا في
58 الطريق إلى المقصد التفت الأ،ظار إلى جبال ر

طقة اسم "بارااليا" ممتزاً. كنت احب البحر والشاطئ الرملي وكل الشياء الاخرى التي توفّرت إمكانية الاستراحة
59 الجيدة. ورحلتى لم تقتصر إلى زيارة المنطقة الوحد

ى كل طاب أن يحصل على علوم الشريعة جيدا أولا ثم يتعلم يتعلم العلوم الأخرى التي يستفاد بها في الدين
60 والدنيا، كما يدعو الناس لله سبحانه و تعالى، فيقول في

لكن أنا عرفت منهم أن في أصول الدين لا يوجد أصول الفقه وبعض المواد الأخرى التي أنا لم أتعلمها من قبل
61 والتي أحْتاج إليها شديدا. فقررت بالشريعة التي معرو

هـ بعد ثلاثة أيام في المراكش. على الرغم من هذا والأحداث المختلفة الأُخرى التي واجهنا قضيتُ هناك إحدة
62 من أجمل فترات في حياتي. تي.

بلحظات شرد البال قليلا وأنا أتأمل فى المنظر الخارجي وأتأمل اللون الأخضر الذى غطى البلد ومع هذا
63 الشرود. بعد ما وصلنا الذي كان فى تمام الساعة السابعة و

حماية البيئة في الشريعة الإسلامية لا يقتصر على حماية الإنسان من الأخطار التي يتعرض لها في وجوده
64 وصحته فحسب، لكنه يتجاوز هذة النظرة؛ ليشمل حماية جميع

ين.وأرجوالله أن يعينني على هذه التخصص ,نعم بعد الإستخا رة استشرت الإخوان الذين سبقوني على هذه
65 التخصص.وحكي لي كل ما يحتا هذه التخصص إبتدءا بقوله:إحذرأو

كعبة. ما كان لي إلا أن أشرع في العمرة وأتناول المساعدة من قبل بعض الإخوة الذين كان قد سبقت لهم
66 العمرة. فكان هذا أمراً وموقفاً لن أنساه. ولله الحمد وال

لنا بالسلامة، لما وصلنا إلى المطار قمنا ب– بعض الإجراءات، بعد ذلك الإخوة الذي إ جاؤوا وقابلونا في
67 المطار––، و– أول مرة في الرياض يعني ما كنت أتوقع أنن

مئنين مرتاحين لبركتها حيث لما وصلنا إلى مطار القاهرة استقبلنا بعض الإخوة الذين يدرسون في الأزهر ثم
68 أخذونا إلي سكن الطلابالمسمى ب(مدينة البعوث الإسلامي

هـ تعالى : " يرفع الله .....والذين أوتوا العلم درجات" فهذه الأدلة التي تدل على أهمية العلوم الشرعية
69 ومكانتها عند الله سبحانه و تعالى فيا إخواني

لفلبين في بداية العطلة الصيفية قائلا هل سنسافر بعد غد؟ أي اليوم الأربعاء الذي كان في الحجز :– ألا نأجل
70 السفر؟ قلت بإعتبار أن الدراسة ستبدأ بعد اسبوعين

حج ثم وصلنا إلى مكة الأرض المقدسة التي هي أحب المكان لله على الوجه الأرض التي قد مشي عليها
71 الأنبياء والرسل والصحابة والتابعين وأنعم اله علي أن أكون مم

ن يعطينا من الأجانب الفرصة لتعلم اللغة العربية بطريقة صحيحة لأن الأساتذة الذين يدرسوننا كانوا من
72 المخصصين في تعليم اللغة العربية لغير الناطقين بها. وخ

الثاني(قمة الدراسة في معهد التعليم اللغة العربية).أشكركم جميعا الأساتذة الذين سعدوا وتعبوا ورشدوا وتبرعوا علياّ بعلمهم.أسأل الله أن يرضينا ولكم. حينم 73

ن هناك لم يفهموا ب بشكل جيد عن دينه عن دينهم ف– ف– من بعض نصانح الأساتذة التي الذي كنت أدرس معهم ف– إذا أنت تلتحق بدراسة ال– في السعودية مثلاً في جامع 74

لناس. هم يشجعونني على أن اتعلم القرآن وأكون متخصصا به. وأيضا من الأساتذة الذين علموني ودرسوني يحفّظوني . ويقولون لي أنهم موجودون خلفي لأي المساعدة الت 75

ها أشعار تنشطني كلما سمعت منها شيئا.ومن الأشياء تجرني إليها قلة الأساتذة الذين تخصصوا فيها ، والحكم والأمثال مثل قول الشاعر : ومن ملك البلاد بغير حرب 76

في اختيار هذا المال لأنني أحبه وأرى في هذا المجال كثير من المواد الأساسي التي سأدرس فيها. مثل العقيدة، التفسير، الحديث وغيره وأيضا إذا تخرجت من هذا ال 77

ظ على هذه الكليات والمقاصد بل جعلتها مقصودها الأعظم. ومن الأمور الأساسية التي حرص عليها رسول الله –صلى الله عليه وسلم– الإلتزام بالبيئة، وقد أوصى بها 78

لديهم من الخبرة والمعارف في هذا المجال، ويقدمون إليّ بعض الدروس الأساسية التي يروها أنّ لها أهمية في هذا المجال، لو كما أنهم يرشدونني إلى بعض الكتب ال 79

لي وأهل بلدي وأسأل الله أن يسهل لي إكمال دراستي في الشريعة وأشكر الأستاذ الذي أثار قلوبنا على ذلك والحمد لله رب العالمين. والسلام عليكم ورحمة الله وبر 80

اً في هذا التخصص، ولأنه تخصص يحتاجه بلدي ومنطقتي، فأنا الوحيدة في الأسرة التي وصلت إلى هذا المستوى من العلم، ولذا يجب عليّ الإجتهاد في هذا التخصص. فزو 81

الله عنه ونسلم علهم ثم انطلقنا إلى مسجد القباء هو أول المسجد في الإسلام الذي وصل النبي صلى الله عليه وسلم وسلم إلى المدينة المنورة وصليت فيه العصر ثم 82

آله وصحبه أجمعين. وبعد! فهذه الجامعة جامعة الإمام محمد بن سعود الإسلامية التي أتاحت لنا الفرصة الثمينة والقيمة التي نستفيد منها جميعا نعمة من نعم الكب 83

أصيلة في المعارف الإسلامية فإننا اكتفينا باعتبار دائرة المعارف الإسلامية التي أنشأها المستشرقون وأصدروها بعدة لغات مرجعاً هاماً نعود اليه ويعتمده الكث 84

لصين اللشمالية , قد زرت المسلمين الأهليين , قدمتهم هدية من كتب الإسلامية التى اخذتما من السعود,قدمنى
85 اطعمة لذيذة وفواكه طازجة ندرس بعضنا بعض , يساعد ب

لوب منا أن الستفيدين من القرآن الكريم والأحاديث الشريف والمواد الإسلامية التي نستفيد منها في المستقبل
86 كنتُ في الطفوليت اتمنى أن أكون متخصصاً في العلوم

لذي يعالج الأمراض الناس الروحية وبدأت بالدراسة بالتعليم الدواء الإسلامية التي فيه شفاء للناس لأحقق ذالك
87 لغابة بل هي وظيفة الأنبياء الدعوة إلى الله بل

ب الجدد في هذه الجامعة الإمام محمد ابن سعود وموجود اسمي أبضاً في الأسماء التي قبل فيه والحمد لله
88 بعدما سمع خبر واستعدت عن السفر إلى بلدي لكي اكتملت ال

لخبر سعيد عن القبول في جامعة الإمام محمد بن سعود في السعودى وفيه الأسماء التي قبل فيه موجود اسمي
89 والحمد لله أشكر الله تعالى ثم سافرت إلى بلادنا لكي إج

ى سنة وزيادة حتى وفقني الله بالقبول في مكة المكرمة فأن بدأت أجهز الأسياء الذي يستحق رجوع معي
90 فالذالك رجعت ذاك الوقت.

ة الدكتورات فى هذه جامعة الأميرة بنت عبد الرحمن، قد لاحظت أن بعض الأشخاص الذين يأخذون
91 التخصص الدينية أرى إليهم نور الإمكان على وجههم أسأل الله أن أكون

جمعاً، وتناولنا العشاء في المطعم. في الطريق المكة شاهدنا كثير من الأشيا التى لم نرها من قبل: مثل
92 الجمال، والصحراء والجبال. ولما وصلنا إلى الميقات يسم

إخترت أن أتخصص في علوم شرعية في كلية شريعة بإذن الله استفسرت عن الأشياء التي تتميز عن كليات
93 أخرى فنصحني معظم الطلاب والأساتذة بدخولي في كلية الشريعة

ووقتا سائدا بها. وما قابلنا بالصعوبات في هذه الرحلة ولله الحمد. الأشياء التي تعلمت في هذه الرحلة منها
94 السفر. الذي نعرف بها الأماكن الغريبة تتعرف بالن

من استعداد الرحلة كان الحصول على التأشيرة وشراء التذكرة من أهم الأشياء التي قمت بها. فكان الحصول .
95 على التأشيرة صعب لسبب الإجرءات التي تتعلق بالحصول ع

جهز بعض الأشياء مثل الشراب ونحو ذلك. أنا مع أخي الكبير نشتري بعض الأشياء التي نطلبها لإقامة الرحلة
96 هناك مثل الخيمة والإشعال النار والحطب، والكبريت وكل

في حاجة ماسة إلى من يتعلم العلوم الإسلامية ثم يقوم بتعليمها. من الأشياء التي تساعدني إن شاء الله في
97 اختياري هذا أن بعض الزملاء اخترو المجال نفسه ولهذ

مدينة التى هاجر إليها نبيّنا محمد صلى الله عليه وسلم أتذكر عن كل الأشياء التي فعلناها خلال هذه الرحلة
98 كأنها أمس، لقد ذهبنا إلى الفندق واسترحنا حتى الس

السعودية في ساحل العاج وسألت و– سألتهم، وقام منهم واحد أخبرني عن الأشياء التي ينبغي لي أ– أن أقوم
99 بها، أخذت منهم استبانةً وذهبوا وذهبت بها إ– إلى الدا

ه أن يكون من الدعاة السعودية (المبعوثين). أدعو الله أن يتحقق هذه الأشياء التي تمنيت وأن يوفقني على هذا
100 المجال. مجال.

افلة' وفي الطريق شاهدنا الأشجار والجبال المرتفعة, وتعجبنا في هذه الأشياء الذي شاهدناها, وعند ما جئنا إلى
101 الشاطئ اغتسلنا في البحر والبعض يستبقون بالسبا

كانو يدرسون معي . في الصباح جاهزنا أنفسنا وأخذنا الحقائب وداخلهم الأشياء الذي نحتاجها في هذه الرحلة
102 ثم ركبنا الحافلات وإنطلقت الحافلات الى الطريق . ال

فى الساعة الثانية وكنت أتحدث مع صديقي شيخ نور عالم فى السيارة عن الأشياء التى ما رأيتها من قبل حتى
103 وصلت فى احدى المحطة ونزلنا من السيارة، وهناك، أكلنا

والأذكار المسنون. وإذا وصلنا إلى مكة نزلنا من السيارة ووضعنا كل الأشياء التي معنا في الفندق. وذهبنا
104 بالقدم إلي بيت الله (الكعبة) وبدأنا الطواف من الو

لخبر هل هو صحيح، فأخبروه في في أن هذا الخبر صحيح، ثم سأ سألهم ما الأشياء التي +ينبغي له أن يقوم
105 بها حتى يعني يكونُ الذهاب إلى السعودية سريعاً جداً؛ لأ

العصر والطهر معا قصرا وجمعاً. وايضا زرنا جبل الأحد وشاهدنا قبور الأصحاب الذين قتلوا هناك. وهذا
106 حدث في يوم غزوة أحد. فدعونا لهم المغفرة والرحمة نسأل ا

ليشرع لهم شرعاً شافيا حتي يستفيدوا استفادة قيمة . وكان معي بعض الأصدقاء الذين اختروا هذا التخصص
107 ,هم يشجعونني عليه وكذلك أشجعهم حتي نتدرب بين انفسنا ل

ئك كالأنعام بل هم أضل" فما بال الناس! يعبدون سوى الله، ونسي هدفهم الأصلى الذي لها خلقهم الله تبارك
108 وتعالى، فقال " وما خلقت الجن والإنس إلا ليعبدون" وب

ك– القراء لكن كنت أستطيع أن أقرأ القرآن. ذهبت إلى قريتنا، قريتنا الأصلية التي ولد فيه جدي، أقص أقصد
109 أبُ الأبِ، في قرية التي كان جدي يعبد الأوثان أ غفر

التعليمية المادة على التالي: 1– تعديل اتجاهات الطالب السلوكية في الإطار الذي رسمه الإسلام وحدّدته.
110 2– تنيمة الروح الدينية لدى الطالب وترفيه ثقافته ال

يرة نورة بنت عبد الرحمن ولما أخذنا إشعارنا استعددنا ملابسنا وبعض الأغراض الذين نحتاج إليه ثم ذهبنا إلى
111 عاصمتنا، ومكثنا هناك 15 يوما بانتظار إلى تأشير

التعلم فقال: " أقرأ بسم ربك الذي خلق خلق الإنسان من علق اقرأ وربك الأكرم الذي علّم بالقلم علّم الإنسان ما
112 لم يعلم" فكل علم الذي يفيد الإنسان في حياته

أنواع تقريبا. هذا الطعام يقدمه على ورقة شجر الموز ونأكل منها كان الأكلات التي أكلنا مفيدة ونافعة جدّاً
113 لأجسامنا وصحتنا. بعد المكوث في القريب السكن، ذه

ة الجيدة للأبجادية اليونانية. كانت هذه الرحلة مغامرة كبيرة، رأيت الأماكن التي شهدت في الزمان القديم ميلاد
114 الديمقراطية والتي كانت الوطن للمفكرين المتعد

وكذلك تجوولنا في بعض المدينة للدعوة والزيارة واستفدنا وزرنا بعض الأماكن الذي يعبد بغير الله ودعوناهم
115 إلى الإسلام الصحيح والحمد لله بعضهم قبل وبعضهم ر

صبح ذهبت إلى المسجد الحرام وصليت الصبح هناك وذهبت مع أخي لكي أرى الأماكن التى كتبت أتمنى أن
116 أراها من قبل وفي الثامن ذهبت إلى منى وبت هناك وفي الصباح ي

وصلنا إلى مكة استعدنا للعمرة، وبعد إنتهينا من العمرة، وذهبنا إلى الأماكن الذين في مكة لنعرف لماذا يسمى
117 هذا المكان، وزرنا بعض المشايخ في هناك، وأن الشي

بح ذهبت إلى المسجد الحرام وصلت الصبح هنالك وذهبت مع خالي لكي أرى الأماكن التى كنت أتمنى أن
118 أراها من قبل وفى اليوم الثامن ذهبت إلى منى وبت هنالك وفى ال

مندوا عاصمة نيبال بالحافلة. أولاً حجزنا مقعدين للحافلة، ثم جهزنا الأمتعة التي تساعدنا في أثناء الرحلة من
119 الطعام والشراب. لما صادفنا الموعد خرجنا معاً

د عال عال العالم الذي ماهر الذي هو ماهر في الشريعة، وعندما ي– يأتي الأمر التي الذي يتعلق بالشريعة
120 فهو لا ير لا يريد يعني لا يجيد أ– جيداً. و– أنا أحب

لمباركة لآتعلم هذه اللغة وأتخصص فيها. وبالتالي يأيت التفصيل في هذا الأمر الذي يعيش في بلاد الغرب
يعرف أن المسلمين في حاجة ماسة إلى من يدعوهم إلى الإسل 121

وسأقوم بدورات الإسلامية في بلادي إنشاء الله. وكذلك سأقوم بتدريس الأمهات اللاتي لا تعرفون الدين صحيحاً
لأن تتمنى أن أكون عالما في الدنيا والآخر أجد أجر 122

الصبح ركبنا الحافلة ورجعنا إلى الرياض بسرور مع زملائي ومتشاورا عن الأمور التي قضينا في الحج من
الصعوبات ومن المشكلات. بعد صلاة العصر وصلنا الرياض ونزل 123

قبل، ومالا يمكث أن ترى إلا إلا في هذه المنطقة. كما تعجبنا من كثرة الأناس الذين يرتادون هذه المنطقة إذا
إنّ طبيعتها مناسبة للسياحة، كما قمنا بأخذ الصور 124

ن هو أعظم كلام لأنه كلام الله سبحانه وتعالى الي خلقته، وما أجمل الإنسان الذي يفهم كلام إلهه. وبسبب
الثاني أريد أن أفهم التفاسير القرآن بدقة لأجل الله 125

اس في المدينة المنورة أخلاقهم خير من الناس في مكة لأنهم من أصلاب الأنصار الذين ناصروا الإسلام
والمسلمين وهذه الرحلة تكون سعدة لنا جميعا والحمد لله. 126

ا، هذه المهارة تحتاج الطلاب العلم أن يدخلون فيها ويديرون ويصلحون الانظمة التي وُجد الآن، أن تَخَرجتُ
من كلية التجارية من أحد جامعة من جامعات الكندي ول 127

لى: زرنا بعض المتاحف في مكة، وفي المتحف شاهدنا بعض الآثار. ومنها الأوانى التى الأصحاب النى
صلى الله عليه وسلم يستخدمون، والجلود التى كانوا يجعلون فيها 128

نها شيئا أو إضافةً إلى ذلك ما جَهزت للسفر ولا استحمَمْتُ من كثرة الأوراق التي أصحّحَ جاءتني فكرة أن
أعطي الأوراق لواحد من الإخوان ثم جهّزت للسفر وتوجه 129

الله لي حفظ حديث رسوله صلى الله عليه وسلم. اكتفى بهذا القدر لقلة الأوراق التي قدمّ لي. وإلا لتحدثت أكثر
من هذا. . 130

ام أدرس الطلاب الصغار الذين ماعندهم خبر في اللغة العربية خلال هذه الأيام التي قضيتها في بماكو. المواد
التي كنت أدرسهم عليه :القرآن كل الصباح ثم أكون م 131

لعربية إلى غيرها. مذ طفولتي أحب هذا التخصص لأنني شاهدت في بلدي أن الأئمة الذين درسوا اللغة العربية
وغيرها من اللغات يدخلون الناس في الدين دائما لأنهم 132

ابه أجمعين أما بعد : في اجازة الحج الماضية، سافرت إلى مدينة "خفر الباطن" التي تقع في محافظة المنطقة
133 الشرقية وتبعد حوالي 100 كم عن حدود دولة الكويت. زر

لى الفندق تناولنا العشاء ثم نمنا وبعد اليوم الثالث. ذهبنا إلى شاطئ البحر الذي يسمونه نصف القمر، فقضينا
134 هناك يوماً كاملاً، ثم رجعنا إلى الجامعة بسلامة.

رحلة إلى شطئى البحر في اسبوع الماضي ذهبت مع اصدقاء على شطئ البحر التي يقع في الدمام احد
135 من المناطق في البلاد سعودية انطلاقنا من الرياض في الساع

ة العشيش التي امتلأ ببيوت الخشب بلا كهرباء في البر وأيضا سمعت قصص البدون الذين لا يحتملون أي
136 جنسية ولا إثبات ولا هوية فصدمت كثيرا. كيف يوجد هؤلاء النا

أثرت هذا الموقف البطولي الذي قام هذا الشيخ، ولذا أقول إن الأعمال البديعة التي انتشرت في الربوع المسلمين
137 ،و هي الطاهرة قد فشت في التجمع الإسلامي في الآ

ل صلى الله عليه وسلم صلينا فيه وفي ليلته ذهبنا إلى إستراحة أقمنا البرامج التي جهزناها في الحافلة وفي
138 الصباح عدنا إلى السكن.

ذهبنا إلى الفندق قريباً في المسجد، ودخلنا بعد قليل أخبر المشرف البرنامج التي نقوم بها في يومين، نازرو
139 الأماكن مختلف، نبدأ من مجسد قبلتين والمسجد القب

دخلنا المدينة وذهبنا إلى الأسواق. كان البضائع رخيصة بشدّة. رأينا البضائع التي نرى في أسواق في بلدان
140 الأخر بسئر رخيص وفهمنا أن لهذه الأماكن هم الأماكن

و ركبنا سيارة, في الصباح خرجنا من الرياض, في وقت العصر وصلنا إلى البطحى الذي خدود بين المملكة
141 العربية السعودية و الإمارات. ثم دخلنا إلي التفتيش و بعد

عضهم ، هذا الشيخ الجليل، على هذا العمل العظيم، وتأثرت هذا الموقف البطولي الذي قام هذا الشيخ، ولذا
142 أقول إن الأعمال البديعة التي انتشرت في الربوع المسلم

ذلك سأدرس العقيدة بالجهد. فكّرت فى هذه المادة تفكيراً كثيراً لأنّ البلاد الذي أسكن فيها وخاصة في القرية
143 التي أسكن فيها، في أطرافها الأربعة يعيش مذهب ب

ت في البيت وفكرت عن الإكتمال الدراسي في المرحلة ميجستر ثم فكرت عن البلاد الذي أريد أن أدرس فيه
144 السعودية ومصر لكن السعودية لابد بالشروط ثم قال أخ الكبي

ورِ الد أ– الدُفُن هو أراد أن يرجع عند شيخه في نفس البلاد أ– فينفس البلد الذي كان فيه ليتعلم، فلما أراد
145 رفض أمه، رفضت أمه، فقالت إذا ذهبت من الذي يكون

ت بِلد آخرِ. ولكنْ ذلك سَعر وذهبتُ مُباشَرة بعد أَخدتُ تَقشِيْرَة. البلد الذي ذَهَبتُ إِليه يعْني سعود عَرَبية. ذهَبتُ
146 إلى هُنك لأدرس اللغَةِ العَربية

فالمسلم لابد أن يعرف ما هو الإسلام خاصة لطلاب المسلم، فأنا طالب من البلد الذي أكثرها من غير مسلم.
147 كثير منهم لا يعرفون الإسلام. فلذلك أنا أحب أن أعرف ا

ى جمعة الشمال وكان معنا تقريبا أربعة عشر طلابا ثم بعد ذلك جئنا إلى البلد التي كانت في طريق الشمال
148 وقمنا به للإستراحة والفطور وقضينا الاستراحة تقريبا س

انت تهطل بغزارة ، لأن الوقت كا في موسم الأمطار علما أن أثيوبا من البلدان التي تكثر فيها الأمطار ،
149 استراح دهني من عناء الدراسة . أما أهل المدينة كانوا

الإسلام، كيف يصلون وكيف يفضلون الصلات، سمعت أصول دين طيبة للبعد البلدان، الذي لا يعرف أهم
150 الاسلام ودين شرعة ، أهم أصول الدين، الذي يرغب دعوة الاسلام ط

لبناء الشامخ الكبير فذهبت إليه وسألت رجلا عن البناء فقال : هذا هو البناء الذي يجلس في الملوك ويصدرون
151 الأمور الملكي، وسرير الملك منخفض من محراب المسجد

ونظرنا مناظر بلدي اسنتشعرنا بالراحة والمريحة، لأننا قد انقطعنا من البيئة التى نشأنا فبها مدة طويل. وكنا
152 نتمني الرجوع إليها في الجامعة. وقد قضينا إجازة

المقابلة ‘ثم رجعنا إلى مكة المكرمة وأردنا الذهاب إلى الرياض لكن التاشيرة التي معنا لا تسمح لنا بالدخول إلى
153 الرياض’ فقدمنا أوراقنا لأخينا الحبيب محمد ت

ثم ذهبتُ إلى عمادة الشؤون الطلاب لأخذ جواز السفر. ووجدتُ خطأ في التأشيرة الذي التاريخ العودة قبل شهر
154 رمضان. فأسال مدير العمادة وطلبت منه تغير التأشيرة

و". وكان الاختيار موفق فقد استمتعت كثيراً بالسفرة على الرغم من التحذيرات التي سمعت قبل السفر. أنا اصلا
155 من ولاية "داغستان" الذي تقع في غرب الروسية، ولذ

156 التحصص الذي أريد هي كلية الشريعة. أخترت كلية الشريعة لأكون عام في الفقه لأن بلادي غان

الحمد لله جيداً وكيف دراستك مريم الحمد لله ممتاز دراسة كومل ما هو التخص الذى ترغبين في الدراسة او
157 العمل فيه في المستقبل أيامك؟ مريم ارغب فى تخصص الشر

التخصص الذي أرغب في دراستي: – التخصص الذي أرغب في دراستي، وهي إن شاء الله إن كان عمري باق
158 أريد أن أخصص في الدّين و

بعض المواد الدراسية التي أعتقد فيه ينفعني لديني واللغة العربية من التخصص الذي أحبه كلية الشريعة
159 الإسلامية. خطة التي أتخذت للمستقبل القراءة ولإستماع لل

160 ما هو التخصص الذي أريده عندما كنت صبيا كان التمني أن أكون طبيبا حتى أعالج الناس من الأمراض

خيرته اللغوية وارتقى مستوى اللغوي عن المستوى السابق الدراسية وأنا التخصص الذي أحبه كلية حديث فقد
161 إمتدى هذه المادة لمعالجة أبواب وفضائل الأعمال والأدعي

له والصلاة والسلام على رسول الله وعلى آله وأصحابه أجمعين أما بعد: التخصص الذي أرغب دراسته في
162 المستقبل اللغة العربية. وأريد أن أكون مدرسا في بلدي للغة

ن عبدالله وعلى آله وصحابته الكرام ومن تبعهم بإحسان إلى يوم الدين. التخصص الذي أريد أن أتخصص فيه
163 هو كلية الشريعة، أريد أن أتخصص في كلية الشريعة؛ لأن عم

164 التخصص الذي أرغب في دراستي: – التخصص الذي أرغب في دراستي، وهي إن شاء الله إن كان عمر

لّ علم فيه تخصص يوضع هذه التخصصات يكون تيسيراً من تعلّم وقد اخترت التخصص الذي احبه وهو
165 الحديث، ووصول إلى هذا التخصص يحتاج جهداً كبيراً ولكن همتي تساعد

166 التخصص الذي أختار اليه "أدرس في جامعة الإمام , الحمد لله, الان تقدمت ألى مستوى الرابع

التخصص فى الشريعة "بسم الله الرحمان الرحيم أما بعد التخصص الذى أرغب فى دراستى، هو التخصص فى
167 الشريعة، لأننا نسكن بين الكفار ، المسلمون فى

ام عليكم ورحمة الله وبركاته، أما بعد أريد أن أوجه مقولاتي بالنسبة التخصص الذي أخترته، ها أنا أردد وأقول
168 كما جاء في الحديث أن " طلب العلم فريضة على كل

ما كنت أدع الله أن يوفقني مجيئ إلى مكة المكرمة في جامعة أم القرى. التخصص التي أريد أن أدرس في
169 الجامعة هي الشريعة الإسلامي حتىّ أفهم الدّين الإسلام بدق

التخصص في مستقبل التخصص الذي احب عن اوصل دراستي في مستقبل الى كلية اصولدين. احب لهذه
170 كلية لانني في راي

171 التخصص الذي أرغب هي الترجمة منذ صبايا كنت أريد أن أسمع أو ان أتحدث عدة اللغات، لذلك أ

هي من دراستي في هذا المعهد بعد سنة. ولذلك لا بد علي أن أفكر وأختر التخصص الذي سوف أتعلم في
172 الكلية أو بعد انتهاء المعهد. لحظة من لحظة، أصبح يكون عندي م

وكنا نتحدث عن المستقبل فسألتني صديقتي: عندما ستخرجين في المعهد أي التخصص الذي ترغبين في
173 دراسته أو العمل في المستقبل أيامك؟ فقلت: سؤالك هم جدّاً بالنسب

م درست خمس سنوات هنك في كلية الءتصالات ولهذا ما كانت أي علاقة بين التخصص الذي تعلمت
174 والجامعة وبين اللغة العربية وثقافة الإسلامية. وبعد التخرج عملت في

175 التخصص الذي أرغب في دراسته هو تدريب المعلمين أولا: بسم الله الرحمان الرحيم قصدي في اخ

176 التخصص الذي أحببته. اسمه #معلومة شخصية محذوفة# و طالب في مرحلة الماجستير في قسم البلا

وتبرعوا علياّ بعلمهم.أسأل الله أن يرضينا ولكم. حينما تتكلم على أي التخصص الذي ترغب فيه دائما أفكر
177 سواء أريد كلية الشريعة أو أصول الدين.أقول لك حقا أنه

178 التخصص الذى أود فيه بعد ما انتهيت من معهد تعليم اللغة العربة للناطقات بغيرها على وهي

179 التخصص التخصص الذي أرغب دراستها في المستقبل إن شاء الله هو "اللغة العربية" لكوني مسلمة أحتاج

ت التي أواجه في الدراسة هي عدم وجود من يشجعني في دراسة ويرغبك على التخصص التي تخصصت،
180 وارى أن التشجيع والترغيب قد يساعدان للطالب فى دراسته ويصل إلى آخر

181 التخصص الذي ارغب في دراستي وهو الشريعة وعملي فيه مستقبل الداعية إلى السنة صلى الله عل

كلية الشريعة لابد منها التخصص الذي أريد أن أدرس فيه هو الشريعة، فاخترتُ هذاه التخصص قبل ستة
182 السنوات،لأنّ معظ

التخصص الاقتصاد بسم الله الرحمن الرحيم التخصص الذي سأختار بعد انتهيت من معهد اللغة العربية الى
183 وهو تخصص الاقتصاد. أفضل هذا ا

وبحمد الله وتوفيقه لقيت ما كنت أصبو إليه من أ– أ– مواد مختلفة في التخصص الذي أحببته، ومدرسين
184 متمكنين، ودكاترة بارعين في مجالهم، وأنشطة وندوات ومحاضرا

ن قسم البلاغة والنقد ومنهج الأدب الإسلامي. أ– طفولتي لها علاقة ب– التخصص الذي اخترته، أ– وذلك أ–
185 أنني وأنا طفل صغير كنت كان يغلب علي الجانب الشعوري، أ

186 التخصص الذي أرغب في دراسته هو الدراسات الإسلامية ولقد اخترت هذا التخصص لأنشر الإسلام

التخصص الذي أرغب في دراستي هو "التخصص في المصرفية الاسلامية"
187 الفقه، والتخصص الذي أرغب للعمل في المستقبل هي مصرفية ال

اتخرج من معهد التعلم اللغة العربية سالتحق بكلية اصول الدين لأكمل التخصص التي ارغب اليه واشتاق لأشد
188 الناس إلى صراط الله المستقيم وإلى الهداية نسأل الل

كتبت كتبتها يعني عن رغبتي في– في الدراسة في قسم الثقافة إسلامية، التخصص الذي أ– أريد أن أرغب فيه
189 يعني الث الثقافة إسلامية، لماذا لأن ي ينبغي لكل مسلم

رغتبي في الدراسة بقسم الثقافة الإسلامية التخصص الذي أرغب في دراسته أو العمل فيه في المستقبل هو
190 التخصص العلوم الشرعية وفي جامع

التخصص في الشريعة التخصص الذي أرغب هو التخصص في الشريعة / قسم كلية الحديث وكلية التفسير إن
191 وفق الله إنش

التخصص الذي أرغب لدراسته في مستقبل أيّامي بسم الله الرحمن الرحيم التخصص الذي أرغب لدراسته في
192 مستقبل أيّامي هو القرآءآت، وقد اخترت هذا التخصص لوجود معي

193 التخصص الذي أرغب في دراسته لقد رغبت ان أختر على نفسي ان أدخل في كلية الشريعة . وقبل ذ

194 : التخصص الذي أرغب فية هو التخصص في العلوم الشرعية، لأنني بحاجة في هذا التخصص لأكون داع

التخصص بسم الله الرحمن الرحيم التخصص الذي كنت أريد وَأَخترت أن أعمله في مستقبل على وهي أصول
195 الدين، وهذه مهمة جداً ف

ول تخصصي) لأعبّر وأبيّن ما في نفسي لهذا التخصّص تبعاً، اخترتُ هذا التخصص التي هي الشريعة (كلية
196 الشريعة) لأنني أحبها أريد أن أكون عالماً متقناً عارفاً

الدراسات الإسلامية التخصص الذى أرعبوه في مستقبلى الدراسات الإسلامية لأني أحب هذا التخصص لكي
197 أدافع عن الإ

198 التخصص الذي أرغب في دراستي أو العمل في مستقبل أيّامي هي التخصص عن كليّة القرآن اخترته

199 التخصص الذي أرغب فى دراستي هي: كلية القرآن. 2- أخترته لأنني أحبه القرآن كثير ولأن الر

التخصص الدراسات الإسلامية. التخصص الذى أرغب فيه: هو الدراسات الإسلامية، لأننى جئت إلى
200 هذه البلد لأتعلم الإسلام،